## Contents

### Preface ........................................................................ 13
- Summary of New Features .................................................. 14
- Conventions ........................................................................ 22
- Related Publications ................................................................. 23
- Information You Should Have .......................................................... 23
- User Feedback ...................................................................... 24
- iWay Software Training and Professional Services ........................................ 24

### 1. Introduction to DataMigrator .................................................. 27
- Data Management and Migration Overview .............................................. 28
- What Is DataMigrator? ............................................................... 28
- DataMigrator Architecture ............................................................. 29
- DataMigrator Server ................................................................ 29
- DataMigrator Data Management Console ............................................... 31
- Server Web Console .................................................................. 34
- Designing Data Flows and Process Flows ............................................... 35
- Revised Functionality ................................................................... 36
  - Testing the Server. .......................................................... 36
  - Quick Copy. .............................................................. 37
  - Keep Server Connection Alive. .................................................. 37
  - File Listener. ................................................................ 38
  - COMT and TABT Target Files. .................................................. 38
  - Custom Reports for the Log and Statistics Tables. ................................... 38
  - Run Deferred (Legacy). ........................................................... 39
  - Test Run and Run. ........................................................... 39
  - Database Default Values. .......................................................... 39
  - Unassigned Fields. .............................................................. 39
- Getting Help . ........................................................................ 40
- Information Builders and iWay Software on the Web ...................................... 40

### 2. Setting Up DataMigrator ...................................................... 41
- DataMigrator Setup At a Glance ....................................................... 42
- Remote Server Setup At a Glance .......................................................... 44
DataMigrator Setup .......................................................... 45
Initiating Server Administration Tasks ................................ 53
Authorizing DataMigrator Server Usage and Administration ... 53
Managing Application Directories and Configuring the Application Path ..................... 57
Configuring Data Adapters ..................................................... 65
Creating Metadata .................................................................. 71
Testing Sample Data and a Process Flow ........................................... 83
Performing Server Administration Tasks ................................................ 84
  Configuring DataMigrator Server Operations........................................ 84
Using Secure Shell FTP (SFTP) .................................................. 99
  Using SFTP from Windows .................................................... 101
Upgrading DataMigrator From Release 4.3 or 5.1 ...................................... 102
Upgrading DataMigrator from Release 5.2 and Higher ..................................... 102
  Port Number .................................................................. 102
  Configuration Files.......................................................... 103
  Profiles and Applications.................................................... 103
  Data Flows ................................................................ 103
  Scheduler ................................................................ 103
  Data Management Console.................................................. 103
Upgrading DataMigrator from Release 7.6 .............................................. 105
Upgrading DataMigrator to Release 7.7.07 or 8.2 ........................................... 105

3. Planning a DataMigrator Application ................................................. 107
  Designing Your Application ................................................. 107
  Using Synonyms for Data Sources and Data Targets ........................................ 108
  Reviewing Your Data Sources .............................................. 112
    Accessing Data Sources .................................................. 113
  Reviewing Your Data Targets .............................................. 115
    Relational Data Targets ............................................... 116
    Non-Relational Data Targets ........................................ 117
    Remote Target Destinations ........................................... 117
  Establishing Security for DataMigrator ............................................. 119
    Running Flows ................................................................ 119
4. Working in the Data Management Console ........................................125

Data Management Console Interface .................................................125

Application Menu ........................................................................127

Quick Access Toolbar ...................................................................128

Ribbon .........................................................................................129

Ribbon Organization ...................................................................130

Ribbon Tabs ...............................................................................130

Style Menu ................................................................................130

About Button ...............................................................................131

Shortcut Keys for Navigating the Workspace ................................131

Navigation Pane .........................................................................132

Copying, Cutting, and Deleting Procedures, Synonyms, and Other Files. 137

DataMigrator Workspace .............................................................138

Data Flow Tab ............................................................................139

Working With Columns, Variables, and Functions..........................141

Process Flow Tab .........................................................................144

Text View Tab .............................................................................147

Text Editor ................................................................................147

Aerial View ...............................................................................150

Ribbon Tabs ...............................................................................150

Home Tab .................................................................................151

Flow Tab ....................................................................................153

User Functions Tab ....................................................................157

Reports Tab ...............................................................................158

Metadata Tab .............................................................................159

Text Editor Tab ...........................................................................159

Console Log ...............................................................................159

Server Session Log .....................................................................161

Setting Flow Properties ...............................................................161

Viewing Properties .....................................................................168
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application Directory Properties</td>
<td>171</td>
</tr>
<tr>
<td>Flow or Procedure Properties</td>
<td>172</td>
</tr>
<tr>
<td>Setting User Preferences</td>
<td>172</td>
</tr>
<tr>
<td>User Privileges</td>
<td>188</td>
</tr>
<tr>
<td>License Code</td>
<td>190</td>
</tr>
<tr>
<td>Configuring National Language Support</td>
<td>191</td>
</tr>
<tr>
<td>Ending a Data Management Console Session</td>
<td>191</td>
</tr>
<tr>
<td>5. Designing a Data Flow</td>
<td>193</td>
</tr>
<tr>
<td>The Data Flow Interface</td>
<td>194</td>
</tr>
<tr>
<td>Simple Data Flow at a Glance</td>
<td>205</td>
</tr>
<tr>
<td>Complex Data Flow at a Glance</td>
<td>206</td>
</tr>
<tr>
<td>Creating a Data Flow</td>
<td>207</td>
</tr>
<tr>
<td>Adding and Deleting Objects in Your Data Flow</td>
<td>208</td>
</tr>
<tr>
<td>Connecting and Disconnecting Objects in Your Data Flow</td>
<td>211</td>
</tr>
<tr>
<td>Saving a Data Flow</td>
<td>212</td>
</tr>
<tr>
<td>Specifying Data Sources</td>
<td>213</td>
</tr>
<tr>
<td>Updating the List of Data Sources</td>
<td>218</td>
</tr>
<tr>
<td>Creating a Source Transformation in a Data Source Object</td>
<td>219</td>
</tr>
<tr>
<td>Order of Evaluation</td>
<td>233</td>
</tr>
<tr>
<td>Changing the Order of Evaluation</td>
<td>234</td>
</tr>
<tr>
<td>Joining Data Sources</td>
<td>234</td>
</tr>
<tr>
<td>Adding a SELECT to a Data Flow</td>
<td>245</td>
</tr>
<tr>
<td>Adding a UNION to a Data Flow</td>
<td>246</td>
</tr>
<tr>
<td>Selecting Columns</td>
<td>247</td>
</tr>
<tr>
<td>Aggregating Columns</td>
<td>253</td>
</tr>
<tr>
<td>Creating an SQL Calculation</td>
<td>254</td>
</tr>
<tr>
<td>Using an SQL Function</td>
<td>258</td>
</tr>
<tr>
<td>Testing the SQL Statement</td>
<td>265</td>
</tr>
<tr>
<td>Specifying Filters in a Data Flow</td>
<td>267</td>
</tr>
<tr>
<td>Specifying Sorts</td>
<td>273</td>
</tr>
<tr>
<td>Selecting a Data Target</td>
<td>276</td>
</tr>
<tr>
<td>Specifying Data Target Options</td>
<td>283</td>
</tr>
</tbody>
</table>
7. Tutorial: Creating Data and Process Flows With DataMigrator

What You Are Going To Do

Startup and Setup Tasks

Start Your DataMigrator Server and Data Management Console
Configure a Server Node From the Data Management Console (Optional)
Create Sample Data and Metadata and Create an Application Directory

Exercise 1: Creating the Data Flow

Selecting and Joining Data Sources

Add Your First Data Source
Add and Join a Second Data Source to the First
Create A Virtual Column in the Second Data Source
Assign Join Properties

Selecting and Managing Columns

Select and Create Columns
Filter the Data

Adding Data Targets

Map Columns in an Existing Data Target
Create a Data Target
Save and Name the Data Flow

Submitting the Data Flow

Data Flow Review

Exercise 2: Creating the Process Flow

Starting the Process Flow

Adding a Data Flow Object to the Process Flow
Save and Name the Process Flow

Branching in the Process Flow
Create the Procedures
Connect the Procedures to the Data Flow
Specify Your Connection Logic

Submitting the Process Flow
8. Generating Reports, Viewing the Log File, and Printing Flows .......................... 467
   DataMigrator Log, Statistics, and Summary Flow Reports ................................ 467
   Scheduler Reports .............................................................................................. 472
   Impact Analysis Reports on Synonyms, Procedures and Columns ....................... 475
   Working With a Report ....................................................................................... 481
   Viewing a Detailed Flow Report and Log File ................................................... 488
   Printing a Graphical Representation of a Data or Process Flow ............................ 492

9. Optimizing Performance .................................................................................... 493
   Improving Data Extraction Performance ............................................................ 493
   Using Automatic Passthru ................................................................................... 494
     Avoiding Non-Automatic Passthru. ................................................................... 495
   Improving Data Loading Performance ............................................................... 497
   Improving Lookup Performance ......................................................................... 498

10. Working With Synonyms .................................................................................. 501
    Synonym Editor Layout ..................................................................................... 501
    Viewing and Editing Synonym Attributes ......................................................... 515
    Using Variables in a Synonym ........................................................................... 533
    Variables Attribute Summary ............................................................................ 535
    Use a Variable in a Synonym ............................................................................. 536
    Creating and Enhancing Synonyms Using Modeling View ............................... 536
    Adding Indexes to a Synonym ............................................................................ 538
    Adding Virtual Columns (DEFINE) in a Synonym ............................................ 540
    Viewing Data Profiling Characteristics ............................................................. 543
      Data Profiling a Synonym or Segment. ............................................................. 544
      Data Profiling a Single Column. ....................................................................... 546
    Viewing Data Lineage for a Column .................................................................. 552

11. Advanced DataMigrator Tips and Techniques ............................................... 555
    Overview of Examples ....................................................................................... 556
    Creating Sample Data and Metadata .................................................................. 557
    Horizontal Partitioning with Multi-Table Load .................................................. 558
    Using the Decode Function ............................................................................... 562
    Generating a Surrogate Key .............................................................................. 566
Generating a Surrogate Key Using Slowly Changing Dimensions ........................................... 571
Mapping to a Sequential File of a Specific Format .............................................................. 575
Creating and Using Reusable Transformations ............................................................... 578
Using Multiple Targets to Create a Single File With Multiple Records Types ...................... 583
Creating a Target File With Multiple RECTYPEs ............................................................ 586
Combining Multiple Sources Rows in a Single Target Row ............................................... 590
Using Lookups to Create Multi-Table Flows Without Joins .............................................. 594
Record Logging .................................................................................................................. 599
Using Variables in Data Flows ........................................................................................... 602
Creating a Data Flow Using a Target Based on a Predefined XML Schema ......................... 607
Loading a Target Table Using the File Listener ............................................................... 610
Sending a Target File to an FTP or SFTP Server ............................................................. 619
Loading a Star Schema ...................................................................................................... 623
Iterating Through Multiple Tables .................................................................................... 631
Using Subqueries .............................................................................................................. 640
Loading Images Into a Database Table ............................................................................. 642
Updating Values in a Table ............................................................................................... 646

12. Change Data Capture ................................................................................................. 655

Enabling the Database for Logging in Relational Databases ............................................ 655
   Configuring CDC in Microsoft SQL Server .................................................................... 655
   Configuring CDC in MySQL .......................................................................................... 657
       Preparing the Environment for CDC with MySQL ...................................................... 657
       Java. ............................................................................................................................. 657
       Connector/J. ................................................................................................................. 658
       Utilities for Processing Binary Log Files .................................................................. 658
       Configuring MySQL. ................................................................................................. 659
   Configuring CDC in Oracle. .......................................................................................... 659
   Configuring CDC in IBM Db2 Universal Database on Windows and UNIX ................. 661
   Configuring CDC in IBM Db2 Universal Database on System Z .................... 661
   Configuring CDC in IBM Db2 Universal Database on iSeries ................................... 662
       Creating a Journal Library. ......................................................................................... 663
       Turning a Library Into a Journaling Library. ............................................................ 663
Server Run-Time Error Codes ................................................................. 713
DataMigrator Management Problem Resolution ............................................. 714
Including Run-Time Messages in the ETL Log .................................................. 715
Performing a Trace to Diagnose Server Problems ............................................. 715

D. Customizing Your Environment .......................................................... 717
DataMigrator Parameters .................................................................................. 717

E. Source Management ............................................................................. 721
Overview of Source Management ................................................................. 721
Source Control Security Server and Client Setup .............................................. 723
Configuring Source Management and Setting Up Source Control ...................... 727
   Setup, Configure, and Associate DataMigrator to Work With ChangeMan DS. .... 727
   Setup, Configure, and Associate DataMigrator to Work With ClearCase. .......... 729
   Setup, Configure, and Associate DataMigrator to Work With Concurrent Versions
      System (CVS). ....................................................................................... 731
   Setup, Configure, and Associate DataMigrator to Work With Perforce. .......... 735
   Setup, Configure, and Associate DataMigrator With Subversion .................... 739
   Setup, Configure, and Associate DataMigrator to Work With Team Foundation Server
      (TFS). .................................................................................................... 746
   Setup, Configure, and Associate DataMigrator to Work With Visual SourceSafe. .. 753
Using a Source Control System to Manage Procedures and Synonyms ............... 756
Preface

This content describes how to install and use iWay DataMigrator, formerly ETL Manager. It is intended for the user responsible for transforming data by designing and executing data and process flows.

In some instances, DataMigrator is abbreviated as DM.

For up-to-the-minute information, please refer to the release notes.

How This Manual Is Organized

This manual includes the following chapters:

<table>
<thead>
<tr>
<th>Chapter/Appendix</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  Introduction to DataMigrator</td>
<td>Provides background information about DataMigrator architecture, components, concepts, and uses.</td>
</tr>
<tr>
<td>2  Setting Up DataMigrator</td>
<td>Provides setup information for DataMigrator developers and administrators, including overviews for local and remote DataMigrator Server configurations.</td>
</tr>
<tr>
<td>3  Planning a DataMigrator Application</td>
<td>Offers conceptual and practical information to be considered before designing a DataMigrator application.</td>
</tr>
<tr>
<td>4  Working in the Data Management Console</td>
<td>Describes the Data Management Console (DMC) interface.</td>
</tr>
<tr>
<td>5  Designing a Data Flow</td>
<td>Provides detailed information about all aspects of Data Flow design.</td>
</tr>
<tr>
<td>6  Designing a Process Flow</td>
<td>Provides detailed information about all aspects of Process Flow design.</td>
</tr>
<tr>
<td>8  Generating Reports, Viewing the Log File, and Printing Flows</td>
<td>Enables you to view reports on how flows run, what the DataMigrator Server environment looks like, and what error messages the DataMigrator Server generates.</td>
</tr>
</tbody>
</table>
## Summary of New Features

This table describes the new features and enhancements for Version 7.7.07.

<table>
<thead>
<tr>
<th>New Feature</th>
<th>For more information, see...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Management Console</td>
<td></td>
</tr>
</tbody>
</table>

---

## Chapter/Appendix

<table>
<thead>
<tr>
<th>Chapter/Appendix</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 Optimizing Performance</td>
<td>Provides tips for ensuring successful deployment of your data mart and data warehouse implementations.</td>
</tr>
<tr>
<td>10 Working With Synonyms</td>
<td>Provides information about creating synonyms in the Data Management Console and using the Synonym Editor.</td>
</tr>
<tr>
<td>11 Advanced DataMigrator Tips and Techniques</td>
<td>Provides a series of examples to demonstrate the capabilities of DataMigrator.</td>
</tr>
<tr>
<td>12 Change Data Capture</td>
<td>Describes how to use Change Data Capture to enable reading from database logs to determine what rows in a table have changed.</td>
</tr>
<tr>
<td>A DataMigrator Log Statistics and Tables</td>
<td>Provides information about DataMigrator internal information.</td>
</tr>
<tr>
<td>B Event-Based Scheduling</td>
<td>Describes event-based scheduling options that enable execution a flow outside of DataMigrator.</td>
</tr>
<tr>
<td>C Problems, Errors, and Troubleshooting</td>
<td>Assists you in resolving problems that may occur during the development of your DataMigrator application.</td>
</tr>
<tr>
<td>D Customizing Your Environment</td>
<td>Provides information about customizing the DataMigrator environment.</td>
</tr>
<tr>
<td>E Source Management</td>
<td>Source control programs enable you to manage flows, stored procedures, and synonyms, typically providing file backup, change tracking, and version control</td>
</tr>
<tr>
<td>New Feature</td>
<td>For more information, see...</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>--------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Column Name Display Strategy</strong></td>
<td><em>How to Change the Information Displayed for Columns</em> on page 143</td>
</tr>
<tr>
<td>When the strategy is something other than Name, the header label now shows, for example, Display Name (Title). This helps to clearly show what is being displayed. In addition, the field Name can now be displayed even when the display strategy is not Name.</td>
<td></td>
</tr>
<tr>
<td><strong>Text Editor</strong></td>
<td><em>Text Editor Shortcut Menu</em> on page 148</td>
</tr>
<tr>
<td>When using the Text Editor with a stored procedure (or focexec) with one or more lines selected, a new shortcut menu allows you to comment or uncomment multiple lines.</td>
<td></td>
</tr>
<tr>
<td><strong>Download</strong></td>
<td><em>File or Synonym Shortcut Menu Options</em> on page 137</td>
</tr>
<tr>
<td>When selecting a file in the navigation tree, a new Download option downloads the selected file to your Downloads folder.</td>
<td></td>
</tr>
<tr>
<td><strong>Using Find in Properties Panes</strong></td>
<td><em>Setting Flow Properties</em> on page 161</td>
</tr>
<tr>
<td>When viewing a Properties pane in a synonym or flow, a new Find option in the shortcut menu allows you to search for text in the Attributes column.</td>
<td></td>
</tr>
<tr>
<td><strong>Security Type</strong></td>
<td><em>Server Node Configuration Dialog Box</em> on page 169</td>
</tr>
<tr>
<td>New Single Sign On capability is available when connecting from the DMC to a DataMigrator or WebFOCUS Reporting Server on Linux with Kerberos enabled. The IWA security type has been replaced with NTLM for Windows and Kerberos for use with Linux.</td>
<td></td>
</tr>
<tr>
<td><strong>Synonym Editor</strong></td>
<td></td>
</tr>
</tbody>
</table>
## Summary of New Features

<table>
<thead>
<tr>
<th>New Feature</th>
<th>For more information, see...</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pivot</strong></td>
<td><em>Synonym Editor - Column/Field Shortcut Menu</em> on page 507</td>
</tr>
<tr>
<td>New option allows you to transpose XML, JSON, Hive Arrays, and Structures into columns, or multiple columns into rows.</td>
<td></td>
</tr>
<tr>
<td><strong>Parameterize Value</strong></td>
<td><em>Creating a Synonym for a Log Table</em> on page 673</td>
</tr>
<tr>
<td>The Parameterize Value dialog box now allows you to overwrite existing variables, as well as adding new ones.</td>
<td></td>
</tr>
<tr>
<td><strong>Assign SCD Columns</strong></td>
<td><em>How to Assign SCD Columns on a Synonym</em> on page 524</td>
</tr>
<tr>
<td>New option allows you to assign Slowly Changing Dimension attributes to columns. The Surrogate Key Field can now also be explicitly specified.</td>
<td></td>
</tr>
<tr>
<td><strong>Internal and Needs Value Fields</strong></td>
<td><em>Synonym Editor - Field View Tab</em> on page 501</td>
</tr>
<tr>
<td>Internal fields can be distinguished in the Synonym Editor by a grayed out icon.</td>
<td></td>
</tr>
<tr>
<td>In addition, the new Access Property column shows these values for fields with these properties.</td>
<td></td>
</tr>
<tr>
<td>New icons are also available to identify field types that do not allow Sample Data.</td>
<td></td>
</tr>
<tr>
<td><strong>Data and Process Flows</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Expand/Collapse on Flow Objects</strong></td>
<td><em>Data Source Object Options</em> on page 217</td>
</tr>
<tr>
<td>The shortcut menu <em>Toggle</em> for objects in data or process flows has been replaced by <em>Expand/Collapse</em>, which reflects the current state of the object. In addition, the shortcut menu for a flow has a new option of <em>Expand All/Collapse All</em>.</td>
<td></td>
</tr>
<tr>
<td>When multiple expanded objects are selected, you can collapse them using a new shortcut menu option of <em>Collapse All</em>.</td>
<td></td>
</tr>
<tr>
<td>New Feature</td>
<td>For more information, see...</td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td>------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Target Display Name</strong></td>
<td>#unique_13</td>
</tr>
<tr>
<td>The name displayed on a Flow can be changed</td>
<td></td>
</tr>
<tr>
<td>from the default application and synonym name.</td>
<td></td>
</tr>
<tr>
<td><strong>Long Display Name</strong></td>
<td></td>
</tr>
<tr>
<td>Long synonym and procedure names are now</td>
<td></td>
</tr>
<tr>
<td>shown truncated with ellipses. The full name is</td>
<td></td>
</tr>
<tr>
<td>shown if you hover the mouse over the name.</td>
<td></td>
</tr>
<tr>
<td><strong>Set default target database</strong></td>
<td>Navigation Pane: Adapters Folder on page 69</td>
</tr>
<tr>
<td>A default target database can be set from the</td>
<td></td>
</tr>
<tr>
<td>Change Common Adapter Settings shortcut menu</td>
<td></td>
</tr>
<tr>
<td>option. The value for this Default ETL Target</td>
<td></td>
</tr>
<tr>
<td>DBMS, ETL-TRG-DBMS, is now respected for SQL</td>
<td></td>
</tr>
<tr>
<td>(adapter) targets only. If a flow uses a new</td>
<td></td>
</tr>
<tr>
<td>target other than SQL adapter type, then the</td>
<td></td>
</tr>
<tr>
<td>ETL-TRG-DBMS setting is ignored.</td>
<td></td>
</tr>
<tr>
<td><strong>Email Attachments</strong></td>
<td>How to Add an Email Object to the Process Flow on page 410</td>
</tr>
<tr>
<td>Files of any type (or extension) can now be</td>
<td></td>
</tr>
<tr>
<td>attached with the E-mail object of a process flow.</td>
<td></td>
</tr>
<tr>
<td><strong>Email Option for Send HTML as inline message</strong></td>
<td>How to Set the Flow Email Notification Properties on page 165</td>
</tr>
<tr>
<td>The HTML output from a report can now be sent in</td>
<td></td>
</tr>
<tr>
<td>an email message on failure, as well as on</td>
<td></td>
</tr>
<tr>
<td>successful completion.</td>
<td></td>
</tr>
<tr>
<td><strong>Email Recipients</strong></td>
<td>How to Set the Flow Email Notification Properties on page 165</td>
</tr>
<tr>
<td>Email messages can now be sent to recipients as</td>
<td></td>
</tr>
<tr>
<td>a Cc or Bcc.</td>
<td></td>
</tr>
</tbody>
</table>
### New Feature

<table>
<thead>
<tr>
<th>New Feature</th>
<th>For more information, see...</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Optimize Load</strong></td>
<td><em>How to Set the Flow Execution Properties</em> on page 162</td>
</tr>
<tr>
<td>When source and target tables are in the same relational database and an Insert/Update load type is used, a new flow property <em>Optimize Load</em> for databases that support it, a SQL MERGE command is generated. For Insert only operations, an INSERT INTO ... SELECT command is generated.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Calculators</strong></th>
<th><em>SQL Functions</em> on page 259</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Format Autocorrect</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>DTRUNC Function</strong></th>
<th><em>Using Functions</em> on page 347</th>
</tr>
</thead>
<tbody>
<tr>
<td>The simplified DTRUNC function, which returns the first day of a time period (YEAR, MONTH, QUARTER), has been extended to return the first day of a WEEK and the last day of any time period.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>New Available Functions</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>The following functions are now available:</td>
<td></td>
</tr>
<tr>
<td>- CEIL</td>
<td></td>
</tr>
<tr>
<td>- FLOOR</td>
<td></td>
</tr>
<tr>
<td>- GREATEST</td>
<td></td>
</tr>
<tr>
<td>- LEAST</td>
<td></td>
</tr>
<tr>
<td>- MOD</td>
<td></td>
</tr>
<tr>
<td>New Feature</td>
<td>For more information, see...</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>REPLACE Function</td>
<td>Using Functions on page 347</td>
</tr>
<tr>
<td>The REPLACE function will replace or remove a substring within a string.</td>
<td></td>
</tr>
<tr>
<td>Regular Expressions</td>
<td></td>
</tr>
<tr>
<td>Regular expressions can now be used in a WHERE condition for screening.</td>
<td></td>
</tr>
<tr>
<td>Data Profiling</td>
<td></td>
</tr>
<tr>
<td>Hex Views</td>
<td>Data Profiling a Single Column on page 546</td>
</tr>
<tr>
<td>A new data profiling report shows the hexadecimal value of any field.</td>
<td></td>
</tr>
<tr>
<td>Charts</td>
<td>User Preferences - Run Options on page 187</td>
</tr>
<tr>
<td>A new option in the Run Options dialog box allows you to omit NULL values from charts.</td>
<td></td>
</tr>
<tr>
<td>Reporting</td>
<td></td>
</tr>
<tr>
<td>Dependencies Analysis</td>
<td></td>
</tr>
<tr>
<td>A new report shows all of the flows or procedures that the selected flow or procedure invokes.</td>
<td></td>
</tr>
<tr>
<td>Data Lineage for Procedures</td>
<td></td>
</tr>
<tr>
<td>For a selected flow or procedure, a new report shows the name of each field used in the procedure, as well as how it is used, its synonym name, and the name of the table and column in the underlying database.</td>
<td></td>
</tr>
<tr>
<td>Application Path</td>
<td>Upgrading DataMigrator to Release 7.7.07 or 8.2</td>
</tr>
<tr>
<td>The maximum length of an application path has been increased from 64 characters to 256 characters.</td>
<td></td>
</tr>
</tbody>
</table>
### Summary of New Features

<table>
<thead>
<tr>
<th>New Feature</th>
<th>For more information, see...</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analysis Reports for Flow</strong></td>
<td><em>Navigation Pane: Procedures</em> on page 195</td>
</tr>
<tr>
<td>The Impact and Dependency analysis reports have been grouped under a new shortcut menu, Analysis reports. The Data Lineage report is also included.</td>
<td></td>
</tr>
<tr>
<td><strong>Running</strong></td>
<td></td>
</tr>
<tr>
<td><strong>External Password with CMRUN</strong></td>
<td><em>How to Start Flows With CMRUN</em> on page 696</td>
</tr>
<tr>
<td>When initiating a flow with CMRUN, you can specify the location of a password file on any platform. Previously, this option was only available for zOS.</td>
<td></td>
</tr>
<tr>
<td><strong>File Listener Option to Archive File on Local Server</strong></td>
<td><em>File Listener Attributes</em> on page 526</td>
</tr>
<tr>
<td>When the File Listener is processing remotely located files, they can be archived on either the remote or the local server.</td>
<td></td>
</tr>
<tr>
<td><strong>Do Now Allow Multiple Instances</strong></td>
<td><em>Properties Pane for Flows and Procedures</em> on page 407</td>
</tr>
<tr>
<td>When working with a Process Flow that initiates another data flow, you can now prevent the called flow from starting if another instance of the flow is running, by selecting the <em>Do not allow multiple instances</em> check box.</td>
<td></td>
</tr>
<tr>
<td><strong>Adapters</strong></td>
<td></td>
</tr>
<tr>
<td><strong>MS SQL Server ODBC</strong></td>
<td></td>
</tr>
<tr>
<td>The ODBC adapter for MS SQL Server can be used on Linux, in addition to Windows.</td>
<td></td>
</tr>
<tr>
<td>New Feature</td>
<td>For more information, see...</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td><strong>Extended Bulk Load Options</strong></td>
<td><em>Target Load Options</em> on page 285</td>
</tr>
<tr>
<td>Extended Bulk Load is now available MS SQL Server, ORACLE, PostgreSQL, Informix, Netezza, and Exasol</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>New Adapters</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ <strong>Apache Phoenix for HBase.</strong> New adapter for data stored in HBase.</td>
<td></td>
</tr>
<tr>
<td>✓ <strong>Apache Drill.</strong> New adapter for data in HDFS, HBase, and other sources.</td>
<td></td>
</tr>
<tr>
<td>✓ <strong>Cloudera Impala.</strong> New adapter for Cloudera/ Simba JDBC driver</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Connection for [Delimited] Flat File Targets</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>For Delimited and Flat File targets, you can now select the connection, FTP or SFTP server, where the data should be written.</td>
<td>#unique_28</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Column Names in Tutorials</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>By default, column names are created in UPPER CASE. They can now be created in lower case, which is preferable for some databases.</td>
<td><em>How to Create Column Names in Lower Case</em> on page 80</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>File Attributes From File Listener</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>File attributes, such as the file size and last modified timestamp, are now also available when reading the files with a pickup strategy of IMMEDIATE.</td>
<td><em>Prototype File Monitoring Synonym</em> on page 532</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Change Data Capture for MySQL and MariaDB</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DataMigrator support for Change Data Capture now includes the MySQL and MariaDB databases.</td>
<td><em>Configuring CDC in MySQL</em> on page 657 and <em>Configuring CDC in MariaDB</em> on page 667</td>
</tr>
</tbody>
</table>
## New Feature

<table>
<thead>
<tr>
<th>SFTP from Windows</th>
<th>For more information, see...</th>
</tr>
</thead>
<tbody>
<tr>
<td>With an SFTP client program installed, an sftp server can be used as a source or target from a server running on Windows.</td>
<td>Using SFTP from Windows on page 101</td>
</tr>
</tbody>
</table>

## Quick Copy

### Delimited Flat File Targets
Now supported.

### Specify Keys
The key columns in the target table can be derived from the source table or specified by the user.

## Conventions

The following conventions apply throughout this manual:

<table>
<thead>
<tr>
<th>Convention</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>THIS TYPEFACE</strong> or <strong>this typeface</strong></td>
<td>Denotes syntax that you must enter exactly as shown.</td>
</tr>
<tr>
<td><strong>this typeface</strong></td>
<td>Represents a placeholder (or variable) in syntax for a value that you or the system must supply.</td>
</tr>
<tr>
<td><strong>underscore</strong></td>
<td>Indicates a default setting.</td>
</tr>
<tr>
<td><strong>this typeface</strong></td>
<td>Represents a placeholder (or variable), a cross-reference, or an important term. It may also indicate a button, menu item, or dialog box option that you can click or select.</td>
</tr>
<tr>
<td><strong>Key + Key</strong></td>
<td>Indicates keys that you must press simultaneously.</td>
</tr>
<tr>
<td>Convention</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>{ }</td>
<td>Indicates two or three choices. Type one of them, not the braces.</td>
</tr>
<tr>
<td>[ ]</td>
<td>Indicates a group of optional parameters. None are required, but you may select one of them. Type only the parameter in the brackets, not the brackets.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>Indicates that you can enter a parameter multiple times. Type only the parameter, not the ellipsis (...).</td>
</tr>
<tr>
<td>. . .</td>
<td>Indicates that there are (or could be) intervening or additional commands.</td>
</tr>
</tbody>
</table>

### Related Publications

Visit our Technical Documentation Library at [http://documentation.informationbuilders.com](http://documentation.informationbuilders.com). You can also contact the Publications Order Department at (800) 969-4636.

### Information You Should Have

To help our consultants answer your questions most effectively, be ready to provide the following information when you call:

- Your six-digit site code (xxxx.xx).
- Your iWay Software configuration:
  - The iWay Software version and release. You can find your server version and release using the Version option in the Web Console. **Note:** the MVS and VM servers do not use the Web Console.
  - The communications protocol (for example, TCP/IP or LU6.2), including vendor and release.
  - The stored procedure (preferably with line numbers) or SQL statements being used in server access.
The database server release level.
The database name and release level.
The Master File and Access File.
The exact nature of the problem:
    Are the results or the format incorrect? Are the text or calculations missing or misplaced?
    Provide the error message and return code, if applicable.
    Is this related to any other problem?
Has the procedure or query ever worked in its present form? Has it been changed recently? How often does the problem occur?
What release of the operating system are you using? Has it, your security system, communications protocol, or front-end software changed?
Is this problem reproducible? If so, how?
Have you tried to reproduce your problem in the simplest form possible? For example, if you are having problems joining two data sources, have you tried executing a query containing just the code to access the data source?
Do you have a trace file?
How is the problem affecting your business? Is it halting development or production? Do you just have questions about functionality or documentation?

User Feedback

In an effort to produce effective documentation, the Technical Content Management staff welcomes your opinions regarding this document. You can contact us through our website http://documentation.informationbuilders.com/connections.asp.

Thank you, in advance, for your comments.

iWay Software Training and Professional Services

Interested in training? Our Education Department offers a wide variety of training courses for iWay Software and other Information Builders products.
For information on course descriptions, locations, and dates, or to register for classes, visit our website ([http://education.informationbuilders.com](http://education.informationbuilders.com)) or call (800) 969-INFO to speak to an Education Representative.

Interested in technical assistance for your implementation? Our Professional Services department provides expert design, systems architecture, implementation, and project management services for all your business integration projects. For information, visit our website ([http://www.informationbuilders.com/support](http://www.informationbuilders.com/support)).
DataMigrator, formerly ETL Manager, represents a broad category of tools designed to facilitate and automate the extraction and integration of data from various sources. At various points during the process, from source extraction through target load, data is transformed through the application of business rules. Once the transformation is complete, the data is loaded into table structures that have been optimized for a particular application.

DataMigrator automates this process.

For up-to-the-minute information, please refer to the release notes in the Help menu of the Web Console.

In this chapter:

- Data Management and Migration Overview
- What Is DataMigrator?
- DataMigrator Architecture
- DataMigrator Server
- DataMigrator Data Management Console
- Server Web Console
- Designing Data Flows and Process Flows
- Revised Functionality
- Getting Help
- Information Builders and iWay Software on the Web
Data Management and Migration Overview

DataMigrator management and migration is the process of extracting data from one or more data sources and moving that data into one or more data targets.

- A data source often contains operational data, such as sales, orders, customers, invoices, and inventories. The data source is designed for online transaction processing and high-speed updates. Systems that maintain these data sources are often highly volatile and mission critical. Within a company, there may be many such systems containing related information with no means for cross-structure comparison.

- A data target, referred to as the data warehouse, data mart, or another type of target, is often relational and is generally located on a machine with optimal performance and availability. This data target typically stores information historically, and is geared for high-performance online analysis.

Often, the process of ETL management involves data cleansing, such as numeric aggregation, improvements to cryptic names, or the calculation of new columns. This process yields many benefits, including:

- **Ease of use.** The data warehouse may be optimized for analysis, and the data cleansing process ensures that data is stored in a more readable format.

- **Reduced load on legacy systems.** Since your analysts are no longer accessing your operational data sources, production systems are not impacted by complex queries.

- **Better response time.** The data warehouse can be situated on a local department machine. You no longer have to search the network to find your data.

- **Ownership.** Many departmental data warehouses can be created, each one containing summary information for that department only.

What Is DataMigrator?

DataMigrator is a suite of software components that automates the process of building and maintaining a data warehouse or data mart. Using DataMigrator, you can:

- Access source data in numerous formats and operating systems.

- Integrate multiple data sources into a single target or multiple data targets.

- Apply powerful data cleansing rules and transformation logic.

- Aggregate data and create roll-ups to aid decision support.
- Use specialized high-volume data loaders.
- Schedule data updates at user-defined intervals, triggered by events, or based on conditional dependencies.
- Load a Star Schema with Slowly Changing Dimensions.
- Monitor and manage key server functions.
- View comprehensive logging and transaction statistics.

Since DataMigrator is a true client/server middleware solution, underlying communications protocols and data source subsystems are transparent.

**DataMigrator Architecture**

DataMigrator is comprised of two basic components and one supplementary component:

- **DataMigrator Server.** A server must be installed and configured to support DataMigrator operations. The server houses and executes data and process flows, which are the key DM design elements. For details about server operations, see *DataMigrator Server* on page 29.

- **DataMigrator Data Management Console (DMC).** The main component of the DMC is the DataMigrator graphical user interface. It runs on a Windows PC, where you design, test, and run data and process flows. For details about elements of this console, see *DataMigrator Data Management Console* on page 31.

- **Server Web Console.** The Server Web Console is installed with the server. It enables you to perform browser-based server administration tasks. For a list of DataMigrator-related server administration tasks, see *Server Web Console* on page 34.

You must install and/or configure these components before you can begin using DataMigrator to build the data flows and process flows that will build your data targets.

**DataMigrator Server**

A DataMigrator Server must be installed and configured for DataMigrator. The server houses and executes the data and process flows from which a data warehouse, data mart, or other data targets are constructed. The server consists of the following components:

- Engines for data sorting, transformations, email notification, scheduling, and security rules.
- Read and write adapters, which are used to extract and load the data.
- Metadata or synonyms (.mas and .acx files) that describe column information and access rules to data sources and/or data targets.
Data flows, which are procedures that contain the necessary instructions for extracting, transforming, and loading the data targets. These procedures are built by the DMC and stored on the server in two files:

- A procedure with the extension .fex, which is ready to run.
- A file with the extension .etg, which contains detailed information needed for the graphical display.

Process flows, which contain the necessary instructions for interleaving data flows with other objects. Like data flows, these instructions are built by the DMC and stored on the server as .fex and .etg files.

DataMigrator internal tables, containing the DM log and run-time statistics.

Stored procedures, which contain instructions for specific tasks. These procedures are edited using the text editor of the DMC.

iWay remote server (optional). Access to a remote server is required when source data is on a different platform from the DataMigrator Server. Adapters to extract the source data reside on the remote server (subserver). Though optional, one or more subservers are used at many DataMigrator sites.

The following figure illustrates a DataMigrator Server environment in which both the data source and the data target reside on the DataMigrator Server.
The following figure illustrates a DataMigrator Server environment in which the data target (whether a data warehouse, data mart, or another type of target) resides on the DataMigrator Server and the data source resides on a remote server (subserver).

DataMigrator Data Management Console

The DataMigrator Data Management Console (DMC) is a graphical design interface that runs on Windows 7, 8.2, 10, and Windows Server 2008 and 2012. You can use the DMC to:

- Design and test data flows and process flows. For more information, see Designing Data Flows and Process Flows on page 35.
- Create and run stored procedures.
- Add and configure data adapters.
- Test synonyms and data adapter configurations.
- Manage metadata by creating and editing synonyms.
- View log and report statistics.

**Note:** Ancillary functions, such as scheduling and email notification for processing, can also be performed from the Web Console, which you can launch from the DMC. For more information, see Server Web Console on page 34.

This manual focuses on the tasks you perform from the DMC. For more information on the DataMigrator Server and the Server Web Console, see the Server Administration manual.
The following image shows the main window of the DMC in Data Flow mode.

It includes the following components:

- **The Navigation pane.** On the left, always displays a list of servers that can be accessed by DataMigrator and on which you can store flows.

  Each application directory displays the files for data flows, process flows, any supporting stored procedures created for an application, and metadata.

  **Note:** When you are logged on to an iWay Server, you can view the Application Directories, Adapters, and Workspace folders in this pane. Until you log onto a server, only the servers appear.
Each application directory also displays the synonyms that are available for source and target data in the data flows of an application. The synonyms in the pane shown in the following image describe data for a variety of database systems.

![Browser screenshot](image)

**Note:** If you do not see the synonyms, you must select the synonyms button 🔍, located in the Filter group of the Home tab.

- **The Data Flow workspace.** At the right, where the flow is designed.
- **The Ribbon.** At the top, provides buttons for the actions most commonly performed from the DMC. Tools available on the ribbon correspond to the task currently being performed. The tabs that can appear on the ribbon are:
  - A Home tab for familiar tasks, such as language selection.
  - A Flow tab for key aspects of the design tasks.
  - A Report tab for tasks associated with report output.
  - A Metadata tab for editing synonyms.
  - A Text Editor tab for editing procedures and synonyms in text view.
- **The Console Log.** Appears below the workspace and displays application and server messages.
For detailed information about these features of the DMC, see *Working in the Data Management Console* on page 125.

**Server Web Console**

Installed with the DataMigrator Server, the Web Console, accessed through a browser, can be used to perform DataMigrator Server administration and monitoring tasks, including the following:

- Controlling server security.
- Monitoring and administering running agents.
- Configuring email notification.
- Monitoring the execution of flows.
- Debugging data and process flows.
- Configuring the DataMigrator Server and remote servers.
- Accessing DM and Scheduler reports.
- Configuring data adapters.
- Creating metadata required for data extraction and loading.
- Accessing help documents and release notes.

The following image shows the Web Console, which can be accessed directly or from the DMC. From here you can open web pages on which all of these functions can be performed.
For information about key server functions required for DataMigrator, see *Initiating Server Administration Tasks* on page 53.

In addition, server administration is fully documented in the *Server Administration* manual, which is accessible as online help from the Web Console.

**Designing Data Flows and Process Flows**

At the core of DataMigrator is the ability to design data flows and process flows, which are types of procedures. The DataMigrator Data Management Console facilitates the design task using two tabs, Data Flow and Process Flow, and provides a Text View tab from which you can see the code generated by the graphical tools.

- **Data Flow.** Normally, you design a data flow from a set of objects that identify your data source(s), your data target(s), the columns of data to include in the data target(s), and the business rules that govern how the data is extracted from the source(s), transformed, and loaded into the target(s).

  You can join multiple data sources, filter and sort the extracted source data, and specify column mappings from the source to the target.

- **Process Flow.** You structure how the data flow is processed by positioning objects in your workspace and defining their interactions. You can combine:
  
  - A Start object on a process flow defines the destination of the log and output of the submitted process, and defines the schedule of the process.
  
  - Data flow objects that indicate what data to extract and move from data sources to data targets.
  
  - Email objects that notify users about the status of the process at specified points.
  
  - Stored procedure objects that perform a variety of supplementary tasks before or after the extraction and load steps defined in a data flow.
  
  - Variable objects that set values for the variables used in a flow.
  
  - Connector objects, represented as arrows, that specify execution logic for the objects in the process flow (for example, what to do next when a step succeeds or fails).
  
  - Group objects that specify parallel, rather than sequential, execution of a subset of objects within the total flow. Objects that follow are not executed until processing for the grouped objects is complete.
Text View. This tab enables you to toggle between the graphical interface and the underlying code at any point in your design process.

Example: Designing a Data Flow

This data flow consists of a data source object, an SQL object in which the columns to be extracted are defined, and a target object that contains the data target you are loading. Double-click these objects to view the columns of each object. In the SQL column selection object, you can add filtering and sorting criteria, arrange columns, insert columns, and create expressions to generate columns of new data. In the target object, you can perform transformations that apply calculations to columns (for example, you can multiply two columns or use functions to perform specialized calculations), edit synonyms, view sample data, and specify mapping logic.

The navigation pane at the left lists available servers and available synonyms for source and target objects.

A data flow typically becomes an object in a process flow. To open a data flow from a process flow, right-click the data flow and select Open. For details about the creation and relationships between data and process flows, see Designing a Data Flow on page 193 and Designing a Process Flow on page 385.

Revised Functionality

Starting with Release 7.7 DataMigrator has some functionality changes from previous releases. The following list describes these changes.

Testing the Server

Prior to release 7.7.07, right-clicking a server name from the DMC opened a shortcut menu that contained a Test Server option. Selecting this option opened a dialog box that contained options for SQL, Focus, and Stored Procedure. The user could select one of these options and then enter a select statement, FOCUS request, or stored procedure to test the server. This option has been removed.

These operations are now done by performing the following steps:

1. Right-click an application directory, such as foccache, select New and then click Procedure.
2. In the procedure area, enter a TABLE request. Alternately, you can type SQL, followed by a SELECT statement or EXEC followed by a procedure name.
3. On the Text Editor tab, in the Run group, click Run.
When you are done with your testing, you can discard the procedure by performing the following steps:

1. Go to the tab that shows the procedure name, for example, foccache/flow1.fex, and click X to close the window.
   
The Save Changes dialog box opens.

2. Click No.

**Quick Copy**

In previous server releases, Quick Copy provided the capability to partition the data for faster loading and to select the columns to copy. While these features are no longer available from the DMC, they are available from the Custom Copy option on the Web Console.

**Keep Server Connection Alive**

The Data Management Console normally connects to the server for each request, for example, when running a flow or stored procedure. When the task completes, the connection to the server is severed. This mode of operation is similar to a web browser and helps to keep server resources available to all users.

In previous releases, the Keep Server Connection Alive option kept a constant connection to the server for the duration of the DMC session. This tied up the server resources regardless of whether they were being used or not.

However, this functionality allowed for tasks to be completed such as setting a value for a variable, issuing a SET or a FILEDEF command, or establishing a DEFINE FUNCTION.

These tasks can now be accomplished with the Edit my profile option. Add any needed commands to your profile. Since the profile is run for every connection to the server, they will be in effect for each session.

**Note:**

- Any changes to variables, or other settings, by running stored procedures will be overridden by the profile when connecting to the server for the next task.

- Global variables can be retained with a setting. For more information, see *How to Make a Global Variable Persistent in a Session* on page 98.
File Listener

In the previous release, an agent based File Listener configured on the Web Console facilitated real time processing. While this allowed processing files to arrive in a specified directory, only Direct Load flows could be used and they could not be scheduled. The flow also did not complete until the agent was terminated, so the log could not be viewed, nor could an email message be sent after processing.

This facility has been replaced in the current release with a new adapter based file listener capability. All configuration is done in a synonym that describes the source files.

The new File Listener has added capabilities. It supports both data flows (SQL based) as well as Direct Load flows and these flows can also be scheduled. The flow completes when a file or time limit is reached, and an email message can be sent on completion.

COMT and TABT Target Files

In previous releases a target with an Adapter of Formatted File could have Format of COMT. DataMigrator would create a comma-delimited file with double quotes around all character columns and the column names as the first line of the file. The TABT format used a tab character as a delimiter instead of a comma.

In Release 7.7 support for COMT and TABT files has be superseded with the support for Delimited Flat Files that allow any delimiter to be specified, not just comma and tab. The enclosure character for strings can also be specified as any character. Inclusion of the column names as the first line is now optional.

Any data flows created in earlier releases with a target type of COMT and TABT will still run. However, when they are opened in the Data Management Console they will be converted to have a target of Delimited Flat File, with the Header set to yes, Enclosure set to a double quote, and Delimiter set to a comma or tab.

Since previous releases did not permit mapping to a predefined synonym describing a delimited file, they allowed adding columns to the target file in the Target Transformations panel. In the current release, a synonym that describes a delimited file can be used as a target. Any columns needed in the target file can be included in the target synonym and assigned in the Target Transformations panel.

Custom Reports for the Log and Statistics Tables

The capability to customize the standard log and statistics reports has been discontinued.
Run Deferred (Legacy)

The capability to submit flows directly to the deferred execution scheduler has been discontinued.

Instead, use Submit or Submit with options and select *Do not wait for completion* to run at a specific date and time.

Test Run and Run

The Run option has been replaced by Submit, which submits a flow to run on the server.

Test Run has been replaced by Run on the browser shortcut menu for flows and on the toolbar for flows.

As with Test Run, when a flow is Run data is loaded and the output appears in the DMC console log (and not in the ETL Log or an email message).

However, Run is not subject to the Maximum number of rows for test reports limit found in the Run Options preferences. To limit the number of rows processed, set the limit from the Run with limits dialog box.

Database Default Values

If a database column has a default value assigned in the database, and no value is assigned in Target Transformations, the database will now assign the default value.

In previous releases, fields with no value assigned in Target Transformations had a NULL value assigned for NULLABLE columns and a space or zero for NOT NULL columns.

**Note:** If a column is described as NOT NULL and has no default value, DataMigrator now requires a value to be assigned, the same as when using a database insert or update statement.

Unassigned Fields

In releases prior to 7.7.04, if no value was assigned to a field in Target Transformations it was assigned an arbitrary value as shown. This value replaced the default value in the database if applicable.

<table>
<thead>
<tr>
<th>Column Described As</th>
<th>Value Assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td>NULLABLE (MISSING=ON)</td>
<td>NULL</td>
</tr>
<tr>
<td>NOT NULL (MISSING=OFF)</td>
<td>Space or zero</td>
</tr>
</tbody>
</table>
In release 7.7.04 the assignment has been changed to correspond with expected relational database behavior:

<table>
<thead>
<tr>
<th>Column Described As</th>
<th>Default Value in Database</th>
<th>Value Assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td>NULLABLE (MISSING=ON)</td>
<td>NO</td>
<td>NULL</td>
</tr>
<tr>
<td>NULLABLE (MISSING=ON)</td>
<td>Yes</td>
<td>Database default value</td>
</tr>
<tr>
<td>NOT NULL (MISSING=OFF)</td>
<td>No</td>
<td>Database error</td>
</tr>
<tr>
<td>NOT NULL (MISSING=OFF)</td>
<td>Yes</td>
<td>Database default value</td>
</tr>
</tbody>
</table>

To avoid an error from the relational database, assign a value for all NOT NULL fields.

**Getting Help**

You can access the DataMigrator online help from within the DMC.

**Procedure:** How to Access Online Help

To access the online help, click the About button arrow and click Help Topics.

You can also go directly to the documentation pertaining to this release by clicking the About button arrow and clicking PDF User's Guide.

**Information Builders and iWay Software on the Web**

You can visit the Information Builders and iWay Software websites where you can ask questions, get help, and look at new products and information for iWay Software. The websites are located at [http://www.informationbuilders.com](http://www.informationbuilders.com) and [http://www.iwaysoftware.com](http://www.iwaysoftware.com).

You can also access these sites by clicking the About button arrow and selecting Information Builders on the web or iWay Software on the web.

For the latest information on DataMigrator, visit the iWay data warehouse technical support site located at [https://techsupport.informationbuilders.com/tech/contents/iway_toc.html](https://techsupport.informationbuilders.com/tech/contents/iway_toc.html), or select iWay Software on the Web from the DataMigrator Help menu.
Chapter 2

Setting Up DataMigrator

Before you create data flows and process flows in DataMigrator, setup steps must be completed.

The Data Management Console (DMC) is used to perform the major setup and administration tasks that are required before you can build flows. As a developer, you will probably perform most of these tasks yourself.

Some DataMigrator setup tasks may be performed by a server administrator, if that is a separate function in your organization. A server administrator may use the Web Console to perform these tasks, which may vary somewhat based on how and where data is stored at your site. As an alternative, the Web Console can also be used to perform the setup and administration tasks that are available in the DMC.

You can use this chapter as a checklist to ensure that, regardless of the distribution of setup tasks at your site, all of the preparation steps are completed before your design work begins.

In this chapter:

- DataMigrator Setup At a Glance
- Remote Server Setup At a Glance
- DataMigrator Setup
- Initiating Server Administration Tasks
- Authorizing DataMigrator Server Usage and Administration
- Managing Application Directories and Configuring the Application Path
- Configuring Data Adapters
- Creating Metadata
- Testing Sample Data and a Process Flow
- Performing Server Administration Tasks
- Using Secure Shell FTP (SFTP)
- Upgrading DataMigrator From Release 4.3 or 5.1
- Upgrading DataMigrator from Release 5.2 and Higher
- Upgrading DataMigrator to Release 7.7.07 or 8.2
DataMigrator Setup At a Glance

The two principal components of DataMigrator are the iWay DataMigrator Server, which includes the Web Console, and the Data Management Console.

1. Install the DataMigrator Server on the platform that will contain your target data (this is the platform you will be moving data to).

   DataMigrator Server is typically installed on the platform where your data target (a data warehouse, data mart, or another type of target) will reside. Your first task is determining that location. The following list of tasks assumes that your source data resides on that same platform.

   **Note:** At many sites, source data is on a different platform from the one on which the DataMigrator Server is installed. If this is the case at your site, you need to install, configure, and test one or more remote servers, using the process described in Remote Server Setup At a Glance on page 44.

2. For Windows platforms, the DMC is installed with the DataMigrator server. For non-Windows platforms, you need to install the DMC separately. For more information, see How to Install the Data Management Console on page 46.

   The DMC can be installed on each Windows computer that will be used to build flows or on a shared network drive.

3. Start the DataMigrator Server.

4. Start the DMC by opening *iWay Software* and then *iWay 77 DM Client* from the Windows Start menu. For Windows servers, the DMC can be found in *iWay Software\iWay 77 DataMigrator Server*.

5. Configure one or more nodes in the DMC configuration file.

   This step identifies the name and port of the DataMigrator Server you will be connecting to and adds the server name(s) to the DMC server list. For more information, see How to Add a Node to Identify the Server to the Data Management Console on page 50.

   **Tip:** If, during the installation of the DMC, you configured it to access the DataMigrator Server, this step is optional. When installing a Windows server, a default LOOPBACK node is configured. This node should not be deleted.

6. Connect to and test the DataMigrator Server from the DMC. You can test that the server is started and able to perform basic data retrieval (against sample data). For more information, see How to Test the Server on page 52. You can also test the server from the Home page of the Web Console.

7. Configure and test data adapters for the source data you will be using from the DMC Adapters folder. The test verifies that the configuration was successful and you can access the data adapter.
Note: If all of your source data is stored on remote servers, omit this step and Step 8 and proceed to Step 9.

8. Create and test the metadata for the source data from the DMC. The test verifies data retrieval functionality for the specific data sources you are using. You can also create and test the metadata from the Applications page of the Web Console.

   Note: If you are using more than one data source for the same adapter, repeat Step 8 as many times as necessary before going to Step 9. If, however, you have data sources in a variety of adapters, repeat Step 7 and Step 8, then proceed to Step 9.

9. Configure and test the data adapter for the data target from the DMC. You can also configure and test adapters from the Adapters folder of the Web Console.

   This step is necessary if the data target is of a different database type than the data source(s) or if all of your data is on remote servers. After completing this step, proceed to Step 10.

   Note: If you have more than one data adapter for your data targets, repeat Step 9 as many times as necessary.

   Step 9 is not necessary if the data target(s) and data source(s) are of the same database type. Proceed from Step 8 to Step 10.

10. Create and test the metadata for one or more data targets from the DMC. You can also create and test the metadata from the Applications page of the Web Console.

   Note: This step assumes that you are moving data to existing target tables. If you are building new target tables in the DMC, this step is not required.

11. Open and test the sample flow dmdemo, which is provided for this purpose. For each one, run and view the last log to confirm that the DataMigrator Server is configured and properly processing flows. For more information, see Testing Sample Data and a Process Flow on page 83.

12. Configure server operations, including security, run-time IDs, and log and statistics tables. See Performing Server Administration Tasks on page 84.

   Note: Authorize any users who will be using DataMigrator from the Data Management Console or Web Console. For details, see Authorizing DataMigrator Server Usage and Administration on page 53.

13. Configure the application path and set up additional application directories. For more information, see Managing Application Directories and Configuring the Application Path on page 57. These procedures can also be performed from the Web Console. For more information, see the Server Administration manual or the Web Console online help.

14. Configure email notification and the scheduler from the Data Management Console. For more information, see Configuring DataMigrator Server Operations on page 84.
For detailed information about the Web Console pages used to set up DataMigrator for local data access, see **Performing Server Administration Tasks** on page 84 and the **Server Administration** manual.

**Remote Server Setup At a Glance**

Although the DataMigrator Server is typically installed on the same platform as the target data for DM applications, there is no such restriction for source data. In fact, the DataMigrator Server frequently resides on one platform while the source data resides on one or more platforms that are remote from the DataMigrator Server itself. In these situations, you must configure the remote servers (also called subservers) to make the data accessible to the DataMigrator Server for processing. For an illustration of this setup, see the second diagram under **DataMigrator Server** on page 29.

First, complete the basic DataMigrator Server configuration steps described in **DataMigrator Setup At a Glance** on page 42. After completing these steps, you are ready to start the setup for one or more remote servers.

To set up the remote server, complete these steps:

- Configure remote servers for data sources and create metadata.
- Return to the DataMigrator Server and configure it to access the remote servers so that you can use the data residing there.

The following steps describe this process:

1. Install (if necessary) the remote server on the platform required by your site. (In many situations, the remote server may have already been installed.)

2. Start the DMC by selecting *iWay Software* and then *iWay 77 DM Client* from the Windows Start menu. For Windows servers, the DMC can be found in *iWay Software\iWay 77 DataMigrator Server*.

3. Configure one or more nodes for the remote server in the DMC configuration file.

   This step identifies the name and port of the remote server you will be connecting to and adds the server name(s) to the DMC server list. For more information, see **How to Add a Node to Identify the Server to the Data Management Console** on page 50.

4. Connect to and test the remote server from the DMC. You can test that the server is started and able to perform basic data retrieval (against sample data). See **How to Test the Server** on page 52. You can also test the server from the Home page of the Web Console.

5. Configure and test the data adapter(s) for the data source(s) you will be using from the DMC. The test verifies that the configuration was successful and you can access the data adapter.
6. Create and test the metadata for the data source(s) you will be using from the DMC. The test verifies data retrieval functionality for the specific data sources you are using. You can also create and test the metadata from the Managing Metadata page of the remote server Web Console.

**Note:** If you are using more than one data source for the same adapter, repeat Step 6 as many times as necessary before going to Step 7. If, however, you have data sources in a variety of adapters, repeat Step 5 and Step 6, then proceed to Step 7.

7. Connect the DataMigrator Server to the remote servers. Once you have installed and configured the remote servers, with their data sources, you must establish a connection from the DataMigrator Server to the remote servers so that DataMigrator will be able to access the data that resides on those servers. To do this, you return to the DataMigrator Servers Web Console Home page and select Remote Servers from the navigation pane. To add the connection to the remote server(s), you click Add in the navigation pane.

8. Create metadata to mirror the metadata on the remote server(s) from the DataMigrator Server Web Console Remote Servers page. (Note that for remote servers, you create metadata from the Remote Servers page rather than from the Metadata page that you used for the DataMigrator Server.)

This step makes the data on the remote server accessible to the DataMigrator Server.

For detailed information about the Web Console pages used to set up DataMigrator for remote data access, see the *Server Administration* manual.

**DataMigrator Setup**

In addition to providing a graphical interface for designing flows, the DMC can also be used to perform the major setup and administrative tasks.

**Note:** When you install a Windows server on a computer where you are building flows, the DMC is installed automatically. For other server types or for access to remote Windows servers, the DMC must be installed separately.

**Reference:** WebFOCUS Environments

The DMC allows you to configure the mid-tier environment for a WebFOCUS Server.

**Reference:** Requirements for Installing the Data Management Console

To install the DMC, you need the following:

- Pentium 550 mHz processor or higher.
- 400 MB of hard disk space.
- 512 MB of RAM.
SVGA or higher resolution graphics card set to at least 800 x 600 pixels and high color (16 bit). (The preferred setting is 1024 x 768, with a color depth of 16 bits.)

- A mouse or compatible pointing device supported by Windows.
- A network (TCP/IP) connecting to the DataMigrator Server.
- Windows Vista, 7, 8.1, 10, or Server 2012.

**Procedure:** How to Install the Data Management Console

1. Close any open programs.
2. Unzip the iserver.zip file to a directory on your hard drive. In Windows Explorer, navigate to the directory and run setup.exe.
   The setup program begins and the InstallShield Wizard opens.
3. In the Choose Setup Language window, select a language and click Next.
   The License Registration window opens.
4. Enter your license key in the corresponding field. Your license key is found on the packing slip you received with your iWay 77 Data Management Tools CD. Click Next.
   The InstallShield Wizard displays the License Agreement. Click Yes.
5. The InstallShield Wizard displays the Select the Program Folder and Standard Location Prompt dialog box, as shown in the following image.

6. The default Program Folder name, as it will appear in the Windows Start menu, is IWAY 77 Client. If you wish to use a different folder name, type it here. (You can also choose a folder from the Existing Folders list.)

7. Accept the default root and product configuration directories, or browse to directories in which you prefer to install your software. Click Next.

The Configure Basic Server Information dialog box opens.
8. When you install the software, you can set up communication to a server. When you start the software, the server you identify here appears in the Server list in the navigation pane. The following image shows the Configure Basic Server Information dialog box.

- If you enter a user ID and password in the Server Administrator ID and Password boxes, the iWay client saves the entries, setting security to Explicit by default. You will not be prompted for them again from the DMC. This step is optional.
In the HTTP Listener box, enter the port number of the DataMigrator Server. If you used the defaults when you installed your DataMigrator Server, the HTTP Listener port number is 8117.

**Note:** The server will use a group of six ports, starting one less than the port entered here.

In the Host Name box, type the computer name or IP address where the DataMigrator Server is running.

9. Click **Next**.
   Review the selected installation parameters listed.

10. Click **Continue**.

   Files are copied to the designated locations. Percent of progress is noted during this process.

   Informational messages indicate that setup is configuring the product based on your specifications and that the program folder and icons are being created.

11. Optionally, you can start the DataMigrator Server and Web Console, or start the Data Management Console as soon as the installation is complete.

12. Click **Finish**.

**Procedure:**  **How to Launch the Data Management Console and Connect to a Server**

**Note:** For non-Windows server installations, the DMC can be found in iWay Software\iWay 77 DM Client\Data Management Console. For Windows server installations, the DMC can be found in iWay Software\iWay 77 DataMigrator Server\Data Management Console.

1. On a Windows system, click the **Start** menu, and click **Programs**. If using Windows 10, click **All Apps**.

2. Click **iWay Software** (or **Information Builders**, if you have a WebFOCUS server installation), then iWay 77 DM Client\Data Management Console or iWay 77 DataMigrator Server\Data Management Console.

   The DMC opens with the available servers listed in the navigation pane on the left.

3. Right-click the server you wish to use and click **Connect**.

   If you have not entered connection information, the DMC login dialog box opens. Enter the credentials in the fields provided.

4. If your server has security enabled, type your user ID and password in the User Name and Password fields. If you are connecting with a domain user ID, select the domain from the drop-down menu.
Windows requires that your user name be no longer than 32 characters.

**Note:** If you select Save credentials, you do not need to enter these credentials each time you connect, and the dialog box does not appear. However, you must reenter this information when you change your password for the server. To do so, right-click the server name in the navigation tree and click Properties to reopen the Remote Server Configuration dialog box.

5. Click **Connect**.

   For data retrieval and processing to occur, you must perform this connection step each time you start the DMC.

**Procedure:** **How to Add a Node to Identify the Server to the Data Management Console**

You can set up the DMC to connect to any number of servers in your enterprise.

When the DMC opens for Windows servers, the default LOOPBACK node is listed under Servers in the navigation pane. It does not appear if you open the DMC from a non-Windows installation.

**Warning:** Do not delete the LOOPBACK node, since it is required for internal processing.

To add another server node in this location:

1. Right-click **Servers** and click **Add Server Node**.
The Server Node Configuration dialog box opens, as shown in the following image.

**Note:** The DMC connects to server nodes using an HTTP listener port and the properties specified in the Host section. Applications that connect using the EDA API, like Resource Analyzer/Governor, use a TCP listener port and the properties specified in the Advanced Parameters section.

2. In the Host box, type the computer name or IP address where the DataMigrator Server is running. For a server running on the same PC as the client, you can type localhost.

3. In the HTTP Listener Port box, enter the port number of the DataMigrator Server, which is 8117 if you used the installation defaults. This port listens to all incoming requests from the DMC and the Web Console.
4. In the Node Name box, type a name for the Server node you wish to add.

5. Optionally, type a description of the node in the Node Description box.

6. Optionally, use a Secured Sockets Layer (SSL) by selecting that check box in the Advanced Parameters section.

7. Security is set to *Explicit* by default. When this setting is in effect, the User ID and password you used to log on to Windows are not passed to the server. If security is required, you must enter it either at connection time or in the Security credentials fields to the right.

   Alternatively, for servers running on Windows, you can choose *IWA* (Integrated Windows Authentication), which passes the Windows User ID and password to the server when a connection is made.

   For information about configuration parameters, see *Viewing Properties* on page 168.

8. Click *OK*.

   The node is added to the Servers tree, as shown in the following image.

   ![Servers tree](image)

   You can test the connection between the DMC and the DataMigrator Server you have added.

**Procedure: How to Test the Server**

Before you proceed, it is recommended to test the connection between the DMC and the DataMigrator Server.

To ensure that the server is running and that you are able to retrieve sample data:

1. In the DMC, double-click the server name. If you can connect to the server, multiple folders, including Application Directories, Adapters, and others, open in the navigation pane.

2. Expand the *Application Directories* folder.

   If you can see the list of application directories, you are connected to the server.
Initiating Server Administration Tasks

Although all required server administration and monitoring tasks can be performed by a designated administrator directly from the DataMigrator Server Web Console, an authorized developer can conveniently initiate the more common administration tasks from the DMC using a server administration ID.

To perform any of these tasks, your server must be running so that you can connect to it from the DMC.

Procedure: How to Launch the Web Console

You can access the Web Console directly from the DMC to perform additional server administration tasks. This procedure assumes that the server has already been started.

From the navigation pane, right-click a server name, and click Access Web Console.

The Web Console opens in your browser.

Authorizing DataMigrator Server Usage and Administration

DataMigrator Server administration functions require a connection with a server administration user ID, such as the user ID that was used to install and configure the server. This user can add additional server administration user IDs.

Note: To use DataMigrator from the Data Management Console or Web Console, a user must be at least an Application Administrator.

Procedure: How to Add Administration IDs

1. In the DMC, expand the Access Control folder.
2. Right-click the Roles folder and click Register PTH <internal> User.

Note: The Access Control folder options will vary depending on the Security Mode setting. If the Security Mode is set to OFF, no other options will be available.
3. From the drop-down menu, select either Manual or Import from provider, as shown in the following image.

If you select Import From Provider, the security provider will provide the list of users available for registration.

4. If you select Manual, complete the following steps:
   a. Fill in the User, Description, and E-Mail boxes.
   b. Select a group from the Inherit Privileges from drop-down menu:
      
      **Server Administrator**
      Permits all functions.
      
      **Server Operator**
      Permits monitoring of server processes.
Application Administrator

Permits management of flows. This administration level displays flows in the Web Console. It is the minimum level required for DM developers who want to use the Web Console to view their flows and/or to use the DMC to create or edit flows.

Basic User

Provides no administrative or monitoring capabilities.

None

Does not provide privileges.

c. Type values in the Password and Confirm Password boxes.
d. Click Add and Register.
e. Click OK in the Confirm window to save your changes and register the new user.

5. If you select *Import from provider*, complete the following steps:

a. Clear the *Filter for multiple users* check box.
b. Click Next.
The User Registration dialog box opens, as shown in the following image.

```
<table>
<thead>
<tr>
<th>Security Provider</th>
<th>PTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inherit Privileges from</td>
<td>Server Administrator</td>
</tr>
<tr>
<td>override the existing registration</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>User</th>
<th>Description</th>
<th>E-Mail</th>
<th>Registered Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>srvadmin</td>
<td></td>
<td></td>
<td>Server Administrator</td>
</tr>
</tbody>
</table>
```

[Diagram of User Registration dialog box]
c. Select the checkbox next to a user and double-click the Description and E-Mail boxes to type in values.

d. Click Register.

6. In the Activate Providers dialog box, select the status for each security level and click Save Provider's Status.

Managing Application Directories and Configuring the Application Path

Procedures and the synonyms they use reside in application directories. The Application Path is a list of viewable application directories. It determines their order of appearance in the navigation pane.

The Application Path is the default search criteria for the server. The server searches each of the directories listed in the Application Path sequentially to find the requested procedure or synonym.

By default, the server profile (EDASPROF.PRF) is run for all users when they connect to the DataMigrator Server to provide access to all application directories in the server search path.

However, you can control a user’s access to application directories by creating individual user profiles. Each user can then:

- Access only the application directories specified in the Application Path specified for that profile.
- Use synonyms and procedures in the specified Application Path.

Grouping procedures and synonyms into application directories provides better control. For your convenience, three directories are included in the server search path during installation:

- `baseapp` is included in the search path for every user profile the DM administrator creates. This directory cannot be removed.

- `ibisamp` is the default location for sample files you can use to create procedures, as well as sample procedures.

- `foccache` is a special directory that allows you to create and reuse files for the duration of your Data Management Console session, or until the maximum retention time is reached (set by the server administrator).

You can, of course, set up additional application directories in the server search path.
The following operations are available from the Application Directories folder by right-clicking in the DMC or Web Console.

- Creating sample procedures and data.
- Setting up application directory mappings for different user profiles.
- Configuring the application path by adding, removing, or reordering application directories.
- Filtering the procedures or synonyms that appear in the application directories.
- Creating new procedures and synonyms.

You can also use the Configure Application Path window to set the viewable directories for DataMigrator, as well as add new directories and mappings. To use DM procedures that reside in a directory other than \ibi\apps\baseapp, the directory must be added to the list of viewable directories.

You can also manage application directories and configure the Application Path from the Applications page of the Web Console. For more information, see the Server Administration manual or the Web Console online help.

**Note:** The instructions in the next several sections are for changing the server profile (edasprof.prf), which applies to all users. However, each user can also have their own profile which is used if security is enabled on the server. An Application Administrator can, by default, change their own application path. A Server Administrator can also change the application path of any user.

To change the application path of a user, expand the Access Control and then Roles folder. Expand the user’s role folder and right-click on user’s ID and click Application Path. On the Application Path Configuration panel, click Change Profile Preference and select a value other than Inherit from previously executed profiles.

**Procedure:** How to Create an Application Directory

1. In the navigation pane, right-click Application Directories, point to New, and then click Application Directory.
The Create New Application dialog box opens, as shown in the following image.

2. Enter the name of the new application directory, and, optionally, a description. Leave the Add directory to APPPATH check box selected.

   **Caution:** The directory name may not contain spaces, special characters, or capital letters. Capital letters are converted to lowercase. If an application directory is created outside of the DMC or Web Console, the same rules apply.

3. Click OK.

**Procedure:** How to Remove an Application Directory From the Application Path

1. Right-click Application Directories in the navigation pane and click Application Path.
The Application Path Configuration dialog box opens, as shown in the following image.

2. Find the application directory you wish to delete. Right-click the path and click Remove from path.

3. Click Save.

**Procedure:** How to Delete an Application Directory

**Note:** You must be connected to a server from a Server Administration ID.

1. If the application directory is still located in the application path, remove it. For more information, see How to Configure the Application Path on page 62.

2. Right-click Application Directories and click Application Preferences.
   - Select the Show Applications not in PATH check box.
   - Skip this step if Hide Application Directories not in PATH is the available option.
   - If neither option appears, make sure you are connected to the server from a Server Administration ID.
3. Click Update.

4. Expand the Inactive Directories folder in the navigation pane.

5. Right-click the directory you want to delete and click Delete.

6. In the Confirm dialog box, click OK.

Procedure: How to Add Application Directory Mappings

You can create a new mapping to a directory anywhere in the file system and include it in the server path. Application directory mappings can be added and deleted on a per profile basis.

1. Right-click Application Directories in the navigation pane, point to New, and then click Application Directory.

   The Create New Application dialog box opens.

2. From the Application Type drop-down menu, select Application Mapping to Disk.

   Additional options appear in the Create New Application dialog box, as shown in the following image.

3. In the Application Name box, type an application name of up to 64 characters.

4. In the Physical location box, enter the full path of the physical directory to be mapped, in the format required by your platform. If needed, select Create application if it does not exist from the Map to drop-down menu.
5. Optionally, enter a description in the Description field.

6. Optionally, select the Add directory to APPPATH check box.

7. If you select Add directory to APPPATH, you can also select from the Position in APPPATH and Profile drop-down menus.

8. Click OK.

   The new mapping appears in the Application Configuration dialog box.

**Procedure:** How to Delete Application Directory Mappings

1. In the navigation pane, right-click Application Directories and click Application Path.
   
   The Application Path Configuration dialog box opens.

2. Find the mapping you wish to delete. Right-click it and click Remove from Path.

3. Click Save.

4. Click OK to confirm the deletion.

**Procedure:** How to Configure the Application Path

You must add an application directory or mapping to the Application Path for it to be accessible.

1. In the navigation pane, right-click Application Directories and click Application Path.
   
   The Application Path Configuration dialog box opens.

2. Click Switch Profile to select the profile of the user whose Application Path you wish to set.

3. To add a directory or mapping to the application path, right-click it in the Available Applications section and click Add to Path.

4. Optionally, right-click any directory or mapping you want to remove from the application path and click Remove from Path. This allows you to remove the directory or mapping from the path without deleting it completely.

5. Set the directory search order by dragging a directory or mapping to its desired position.

   You may want to consider moving the ibisamp directory to the bottom of the list to avoid confusion. If you do not plan to use the sample data, you can remove ibisamp from your Application Path completely.

6. When you have added and arranged all necessary directories and mappings, click Save.

7. Click OK in the confirmation dialog box.

8. Right-click Application Directories in the navigation pane and select Refresh to see the directory or mapping in the tree.
**Note:** If there is a long list of application directories, you can click the *Find* button on the top right or use Ctrl + F to find a text string in the name, physical location, or description.

**Reference:** **Application Directories Shortcut Menu Options**

If you right-click *Application Directories* in the DMC navigation pane, the following shortcut menu options are available:

**Refresh**

Refreshes the application directories tree.

**Explorer**

Opens a Windows Explorer style window on the right side of the Data Management Console that shows detailed information about the currently selected server object.

**New**

Allows you to add tutorials, flows, procedures, synonyms, and files on the server.

**Search Files**

Allows you to search on a number of parameters, including file name, extension, description, and content, as well as set filtering criteria.

**Sort by**

Allows you to sort by name, size, type, and modify date.

**Application Preferences**

Allows you to set preferences for how objects in an application directory appear. You can select whether to show subdirectories last, show applications not in the path, and show object statistics.

**Application Settings**

Opens an Application Settings dialog box that allows you to set settings for the foccache and home application root directories.

**Manage External Repository**

This feature is not available in DataMigrator.

**Application Path**

Allows you to add and delete application directories and mappings from the application path.

**Create My Home**

Allows you to create a home application folder.
Schedule and E-mail

Has the following options:

- **Scheduler Agents.** This report runs for the user whose ID value is stored in the sched_scan_id of a Scheduler. The Scheduler must be active to run the Scheduler Agents report. For more information on this report, see *Scheduler Agents Report Columns* on page 474.

- **Scheduled Events.** Opens a dialog box that shows all jobs scheduled to run during a specified time range.

  The Scheduler must be active to run the Scheduler Events report.

  This report shows all flows and procedures scheduled to run during the time period you selected. The Run Date and Run Time columns show their scheduled run times.

  The User ID column shows the user ID used when running the flow or procedure:

  - **sched_run_id**
    
    Shows what user ID the flow is run under.
  
  - **user**
    
    The user ID that saved the flow or procedure.
  
  - **server_admin_id**
    
    The first server administrator ID.

- **Force Scheduler Scan.** Forces the scheduler to rescan the directories immediately, in order to find and run any new or changed flows or procedures according to their schedules.

Log and Statistics

Opens the Log and Statistics dialog box that enables you to view a log or statistics report.

The following options are available:

- **Report Type**
  
  Allows you to select either a log or statistics report.

- **Report Range**
  
  Allows you to select to view the last report, all reports, or a report specified within a select time range.

- **Completion Status**
  
  Allows you to view all reports, all completed reports, all successfully completed reports, or all completed reports that contained errors.
Run by

Allows you to view reports for all flows, including those initiated by the scheduler or submitted manually. This includes flows initiated from the DMC, WC, and CMRUN.

Application or Flow Name

Allows you to select a specific application or flow name for the report.

Impact Analysis

Opens a sub-menu that allows you to view impact analysis reports. For more information, see Impact Analysis Reports on Synonyms, Procedures and Columns on page 475.

Tutorials

Opens the Create Tutorial Framework dialog box that allows you to create sample procedures and data for DataMigrator. For more information, see How to Create Sample Procedures and Data on page 78.

Configuring Data Adapters

Data adapters must be configured to allow the DataMigrator Server to access data on different types of databases. Once a data adapter is configured, you can create metadata for the tables or views in the form of synonyms. Synonyms are used to both access the data as sources, and to write the data to targets.

To access data in the DM process, the data adapters you need must be defined to the server. You need to add and configure a data adapter for every database type you want to access.

You can also configure and test data adapters from the Configuring Data Adapters page of the Web Console. For more information, see the Server Administration manual or the Web Console online help.

For complete documentation on configuring data adapters and for a detailed explanation of the associated connection parameters, refer to the Adapter Administration manual.
**Procedure:** How to Add and Configure a Data Adapter

1. Expand the **Adapters** folder in the navigation pane and then expand the **Available** folder, as shown in the following image.

![Diagram of available adapters](image)

2. Expand the adapter group and folder of available adapters, for example, **SQL**.

3. Right-click the adapter and click **Configure**.
The Add to Configuration dialog box opens, as shown in the following image.

4. Configuring the adapter consists of specifying connection and authentication information for each of the connections you want to establish. Enter the appropriate information and click **Configure**.

   A confirmation message appears.

5. Click **Close**.

   For complete documentation on configuring data adapters and for a detailed explanation of the associated connection parameters, refer to the *Adapter Administration* manual.

**Procedure:** **How to Test a Data Adapter**

1. Expand the Adapters folder in the navigation pane if it is not already expanded, and then expand the *Configured* folder.

2. Expand the folder for the data adapter and right-click the configured adapter connection.
3. Click Test, as shown in the following image.

The Test Results open.

**Note:** The output will vary depending on the adapter.
Reference:  Navigation Pane: Adapters Folder

The Adapters folder shown in the following image allows you to add and configure adapters from the DMC.

The Adapters folder has the following subfolders:

**Configured folder**

Contains the adapters that have been configured for this server, as well as the Remote Servers folder.

**Available folder**

Contains all of the available adapters that can be configured for this server.

If you right-click the Adapters folder, the following options are available:

**Complete Adapter List**

Shows a list of data adapters that are available on all platforms.

**SQL APT Status**

Displays a report of SQL functions and keywords that show how they are handled for each supported database. The keywords are:

- **pass.** Passes the function to the database as is, without doing a conversion.

- **convert.** Converts the function into database specific implementation and passes it to the database.
focus. Forces SQL to be processed by FOCUS.

SQL Optimization Report
Opens the SQL Optimization Report dialog box, which allows you to specify settings for, and run, the report. The SQL Optimization report shows which functions are optimized to their SQL counterparts by adapter.

Data Types
Opens the Filter Data Types Report dialog box, which allows you to specify settings for, and run, the report. The Data Type report contains mapping options for different data types.

Troubleshooting Tips
Opens a submenu for each configured database adapter that opens a help page for the selected adapter.

Change Common Adapter Settings
Opens the Change Settings for Common Adapter dialog box, which allows you to change the default adapter settings.

Sort Connections
Opens a submenu with the choices of Alphabetically and In Profile Order.

For any adapters for which there are multiple connections configured, this controls the order in which they are displayed.

If you right-click an adapter folder, the following options are available:

Add connection
Allows you to configure another connection for that adapter.

Change Settings
Allows you to change optimization settings, controlling session parameters, and miscellaneous settings, depending on the adapter.

Remove
Removes the configured adapter from the server.

Help
Launches the Web Console Help for the selected adapter.
If you right-click a connection, the options shown in the following image are available:

**Properties**

Allows you to view and change the connection properties.

**Create or Update Synonym**

Allows you to create or update a synonym for that connection.

**Test**

Allows you to test the connection.

**Impact Analysis**

Shows synonyms and flows that use the selected connection.

**Delete**

Allows you to delete the connection.

**Creating Metadata**

Synonyms constitute the metadata that DataMigrator requires to define unique names (or aliases) for each table or view that is accessible from the server. A synonym must be created for every table that DataMigrator accesses.

Once you have created the metadata (synonym), it is a good idea to test it to make sure that the data is available for DM processing.
You can also create and test metadata from the Managing Metadata page of the Web Console. For more information, see the *Server Administration* manual or the Web Console online help.

**Procedure:** How to Create Synonyms for Database Tables

1. In the Data Management Console, right-click an application directory, point to *New*, and then click *Synonym (Create or Update)*.

   The Select adapter to configure dialog box opens, showing the adapters that you have configured.

2. Expand a configured adapter folder, select a connection, and click *OK*.

   The Select Synonym Candidates dialog box opens, as shown in the following image.

   ![Select Synonym Candidates Dialog Box](image)

   **Note:** The contents of this dialog box will depend on the adapter selected.
3. To use the database specified when the connection was added, leave Default Database as is, or select a different database from the drop-down menu. This option is only available for Informix, Microsoft SQL Server, and Sybase ASE adapters.

4. Optionally, restrict candidates based on object type by selecting or clearing the corresponding check boxes. The object types will vary depending on the adapter.

5. Optionally, filter the possible candidates by name and owner. Selecting this option adds Owner and Table Name parameters to the dialog box.

   - For the Owner parameter, type a string for filtering the owner IDs and insert the wildcard character (%) as needed at the beginning and/or end of the string. For example, enter ABC% to select tables or views whose owner IDs begin with the letters ABC, %ABC to select tables or views whose owner IDs end with the letters ABC, or %ABC% to select tables or views whose owner IDs contain the letters ABC at the beginning, middle, or end.

   - For the Table Name parameter, type a string for filtering the table or view names and insert the wildcard character (%) as needed at the beginning and/or end of the string. For example, enter ABC% to select all tables or views whose names begin with the letters ABC, %ABC to select tables or views whose names end with the letters ABC, or %ABC% to select tables or views whose names contain the letters ABC at the beginning, middle, or end.

6. Click Next.
The Create Synonym dialog box opens, as shown in the following image, showing all objects that meet the specified criteria.

7. If you have tables with identical table names, you can assign a prefix or a suffix to distinguish them. For example, if you have identically named human resources and payroll tables (most likely with different owners), assign the prefix HR to distinguish the synonyms for the human resources tables. Note that the resulting synonym name cannot exceed 64 characters. Alternately, you can specify different application directories for different schemas or owners.

When synonyms already exist, you can select the Update or Select option. If you select Update from the Update or Create Metadata drop-down menu, you have the option of selecting which attributes to update. For more information, see How to Update a Synonym on page 76.

8. To specify that this synonym should overwrite any earlier synonym with the same fully qualified name, select the Overwrite existing synonyms check box.
9. To reflect the current cardinality (number of rows or tuples) in the table during metadata creation, select the **Cardinality** check box. The use of cardinality is for equi-joins. The order of retrieval is based on the size (cardinality) of the table. Smaller tables are read first.

   **Note:** If the cardinality of the tables to be used in the application is dynamic, it may not be beneficial to choose this setting.

10. To specify that the synonym should not contain column information, select the **Dynamic columns** check box.

    If this option is selected, column data is retrieved dynamically from the data source at the time of the request.

11. Expand the **Customize data type mappings** section to access additional settings, as shown in the following image.

![Customize data type mappings](image_url)

12. Optionally, change the scale and precision settings.

13. From the Decompose Date fields into components drop-down menu, select **OFF - Turn off**. This option must be off when using DataMigrator.

14. Complete your table or view selection:

    To select all tables or views in the list, select the check box to the left of the **Default Synonym Name** column heading.

    To select specific tables or views, select the corresponding check boxes.

15. The **Default Synonym Name** column displays the name that will be assigned to the synonym. To assign a different name, replace the displayed value.

16. Click **Next**.

   Synonyms are created and added under the specified application directory.
The status field of the Create Synonym window shows that the synonyms were created successfully.

17. Click Close.

**Procedure: How to Create a Synonym for a Relational Stored Procedure**

If you create a synonym for a relational stored procedure, you can use that stored procedure as a data source in a flow. When you create the synonym, you will be prompted for values for the parameters of the procedure. For information on creating the flow, see *How to Use a Relational Stored Procedure as a Data Source* on page 215.

We are using Oracle as an example for this procedure.

1. In the navigation pane, right-click a configured relational Oracle connection and click *Create or Update Synonym*.

   The Select Synonym Candidates (Step 1) dialog box opens.

2. Select *Stored Procedure* from the drop-down menu and the *Filter by owner/schema and object name* check box. Enter an Owner/schema or Object name or both and click Next.

   The Select Synonym Candidates (Step 2) dialog box opens.

3. Select the radio button for a stored procedure and click Next.

   The Select Synonym Candidates (Step 3) dialog box opens.

4. Select check boxes for the parameters and supply a value for testing. Click *Create Synonym*.

   **Note:** Values entered are used only when creating the synonym and are not saved.

   A confirmation message appears.

5. Click Close.

   The synonym now appears in the navigation pane listed under the application directory it was created in.

**Procedure: How to Update a Synonym**

To update a synonym from the Select Synonym Candidates dialog box:

1. Select *Update* from the Update or Create Metadata drop-down menu.

2. Select a synonym from the list and click Next.
The Update Synonyms dialog box opens, as shown in the following image.

3. Select the attributes that you want to override those from the existing synonym and click Next.

The selected synonym will now be updated

**Procedure:** How to Test Synonyms for Tables or Files

1. In the Data Management Console, right-click a server name and click Connect.
2. Expand the Application Directories folder, and expand the application directory where you created your synonyms.

**Note:** If the synonyms are not listed, on the Home tab, from the Filter group, click Synonyms.
3. Right-click the synonym you wish to test and click **Sample Data**. Sample data appears in the workspace, and looks similar to the following image.

![Sample Data Table](image)

**Procedure:** How to Create Sample Procedures and Data

Sample procedures and data are needed for *Tutorial: Creating Data and Process Flows With DataMigrator* on page 433 and for the exercises in *Advanced DataMigrator Tips and Techniques* on page 555.

1. From the DMC, right-click **Application Directories** in the navigation pane, point to **New**, and click **Tutorials**.
The Create Tutorial Framework dialog box opens, as shown in the following image.

2. From the Tutorial drop-down menu, select one of the following options:
   - WebFOCUS - Retail Demo
   - WebFOCUS - State Population Demo
   - WebFOCUS - Custom SQL Security Provider
   - DataMigrator - General
   - DataMigrator - Iterator
   - DataMigrator - File Listener
Note: The WebFOCUS and Legacy sample tables and files are not used with DataMigrator. The Retail Demo option loads tables with a pre-determined amount of data designed for WebFOCUS reports.

3. From the Select DBMS drop-down menu, choose a DBMS type, and click Create. DATREC, a binary file format, will be the only choice if you do not have other adapters configured.

For the Prefix for SQL Tables field, use the default value unless you have a reason to avoid tables starting with the indicated default. If so, type a different value.

For the Application directory field, use the default, ibisamp directory to make it easier to follow the tutorial described in Tutorial: Creating Data and Process Flows With DataMigrator on page 433, Advanced DataMigrator Tips and Techniques on page 555, and other examples in this manual.

When the creation process is complete, a confirmation message appears. The message will depend on the type of samples you create, for example, DataMigrator - General, DataMigrator - Iterator, DataMigrator - File Listener, or DataMigrator - Star Schema.

The message also displays in the Console Log.

The procedures and synonyms appear in the ibisamp application directory, which is created and added to the server search path during installation. For FOCUS, the sample data is also stored in ibisamp. For other adapters, it is stored in the database of the configured adapter. The source data for the file listener flow is stored in ibisamp in four files with names starting with dmord and an extension of data.

4. Ensure that the ibisamp application directory is in the server path. It will be in the path unless you delete it.

 Procedure: How to Create Column Names in Lower Case
**Note:** By default, column names are created in the database with UPPER CASE column names. This is problematic for databases, such as PostgreSQL and those based on it such as Greenplum, where column names are expected to be in lower case.

1. In the DMC, expand the Adapters folder, followed by the Configured Adapters folder.
2. Right-click the folder name of the adapter you wish to configure and click Change Settings.
   
The Change Settings dialog box opens.
3. In the Metadata section, from the ALIAS_CASE drop-down menu, select `LOWER - Enforce lower case`.
4. Click Save.
   
This setting is also used by DataMigrator when creating new tables.

**Procedure:** **How to Create Sample Procedures and Data for Star Schema**

Star Schema sample procedures are needed for *Loading a Star Schema* on page 623.

1. From the DMC, right-click Application Directories in the navigation pane, point to New, and then click Tutorials.
   
The Create Tutorial Framework dialog box opens. Your display will depend on the specific adapters you have configured.
2. Select DataMigrator - Star Schema from the Tutorial drop-down menu.
3. From the Select DBMS drop-down menu, select a relational database that you have configured.

   For the Prefix for SQL Tables field, use the default value unless you have a reason to avoid tables starting with the indicated default. If so, type a different value.

   For the Application directory field, use the default, ibisamp directory to make it easier to follow the tutorial described in *Tutorial: Creating Data and Process Flows With DataMigrator* on page 433, *Advanced DataMigrator Tips and Techniques* on page 555, and other examples in this manual.
4. Click Create. The following status message should appear:

   Star Schema sample procedures created for *database*

   where:

   *database*

   indicates the database type you have selected.

   The procedures and synonyms appear in the ibisamp application directory, which is created and added to the server search path during installation.
Procedure: **How to Create Sample Procedures and Data for File Listener**

From the DMC:

1. Right-click *Application Directories* in the navigation pane, select *New*, and then click *Tutorials*.

   The Create Tutorial Framework dialog box opens. Your display will depend on the specific adapters you have configured.

2. Select *DataMigrator - File Listener* from the Tutorial drop-down menu.

3. From the Select DBMS drop-down menu, select a relational database that you have configured.

   For the Prefix for SQL Tables field, use the default value unless you have a reason to avoid tables starting with the indicated default. If so, type a different value.

   For the Application directory field, use the default, ibisamp directory to make it easier to follow the tutorial described in *Tutorial: Creating Data and Process Flows With DataMigrator* on page 433, *Advanced DataMigrator Tips and Techniques* on page 555, and other examples in this manual.

4. Click *Create* to create the sample procedures and data for file listener.

Procedure: **How to Create Sample Procedures and Data for Iteration**

Additional sample procedures and data are needed for *Iterating Through Multiple Tables* on page 631. These instructions should be followed after you have loaded the sample data and procedures for DataMigrator.

1. From the DMC, right-click *Application Directories* in the navigation pane, point to *New*, and then click *Tutorials*.

   The Create Tutorial Framework dialog box opens. Your display will depend on the specific adapters you have configured.

2. Select *DataMigrator - Iterator* from the Tutorial drop-down menu.

3. From the Select DBMS drop-down menu, select a relational database that you have configured.

   For the Prefix for SQL Tables field, use the default value unless you have a reason to avoid tables starting with the indicated default. If so, type a different value.

   For the Application directory field, use the default, ibisamp directory to make it easier to follow the tutorial described in *Tutorial: Creating Data and Process Flows With DataMigrator* on page 433, *Advanced DataMigrator Tips and Techniques* on page 555, and other examples in this manual.
4. Click Create. The following status message should appear:

Sample procedures created for database

where

database

Indicates the database type you have selected.

The procedures and synonyms appear in the ibisamp application directory, which is created and added to the server search path during installation.

5. Load the created tables and expand the ibisamp directory.

Right-click the dmplntpf flow and click Submit.

Testing Sample Data and a Process Flow

Once you have verified that all required DM setup tasks have been successfully completed, you can perform three simple tests to ensure that your flows will run properly. We created some sample data and process flow procedures for you to use in testing your setup.

We recommend that you run these tests before you create and attempt to run your own flows.

Example: Testing the Sample Data Flow DMDEMO

1. Expand the ibisamp directory in the navigation pane.

2. Double-click the dmdemo data flow. The data flow opens in the workspace at the right.

   This data flow extracts data from the data sources named dmord and dminv and loads an existing table named dmrpts.

3. From the Flow tab, in the Run group, click Run.

4. Click Submit from the drop-down menu.

5. When the data flow run has been completed, the following message appears in the Server Console Log:

   (ICM18763) Request ibisamp/dmdemo complete

6. Right-click the dmdemo data flow from the ibisamp directory, point to Logs, and then click Log and Statistics.

   The Log and Statistics dialog box opens. From the Report Type drop-down menu, select Log and then click the View Report button. The log, which is displayed in the workspace, indicates that the table named dmrpts has been loaded with 178 rows.

7. From the Home tab, in the Windows group, click Windows and then click Close All to close the log and the data flow window.
Next Steps: If you have received these results, you can feel confident that your client/server setup is running correctly and you can proceed to the design phase.

If you are not familiar with DM processing, we urge you to read *Planning a DataMigrator Application* on page 107, before you begin creating flows to ensure that you make optimal decisions about such issues as data structures and security to meet the needs of your site.

Performing Server Administration Tasks

You can control many key aspects of DataMigrator Server operations directly from the DMC. The DMC allows you to:

- Configure the application path.
- View and run flows by expanding the approot folder and its application directories.
- Create sample procedures and data.
- Set up log and statistics options.
- Run reports.
- View run-time information.
- Schedule flows.
- Set Scheduler settings.
- Set email notification.

Configuring DataMigrator Server Operations

You can set up your DataMigrator Server to send an email message that notifies recipients of the successful completion or failure of every flow execution.

In addition, from the DMC, you can enable the distribution of email messages at any point in a process flow.

However, to enable either variation on the email notifications process, you must first configure a SMTP service. (You need to set up your server for email before it will work anywhere in the flow.)

Note: You must be logged on as a Server Administrator with the appropriate privileges to make these configurations.
Procedure: How to Edit the Scheduler Configuration

1. In the DMC, expand the Workspace folder, followed by the Special Services and Listeners folder.

2. Right-click SCHEDULER and click Properties.
   The Scheduler Configuration dialog box opens.

3. Make your changes, and click Save and Restart Scheduler. For more information on the available options, see Scheduler Configuration Dialog Box on page 86.
Reference:  **Scheduler Configuration Dialog Box**

The DataMigrator Scheduler runs as a task under the iWay DataMigrator Server. The Scheduler must be running to submit flows and stored procedures and to execute scheduled data flows and flows that run in parallel groups in a process flow. The following image shows the Scheduler Configuration dialog box.
The Scheduler Configuration dialog box contains the following fields/options:

**Scanning**

**sched_autostart**

Specifies whether the Scheduler should check for flows automatically. The options are:

- **y**
  
  Automatically starts the Scheduler when the server starts. This is the default value.

- **n**
  
  Does not start the Scheduler when the server starts.

  **Note:** The Scheduler must be running to execute not only scheduled data flows, but also restarted flows and flows run in parallel in a process flow.

**dfm_autostart**

If this option is enabled, the DFM listener will start with the server.

**dfm_app**

Specifies the application name of the location of deferred requests and reports.

**dfm_dir**

Specifies the physical location of the deferred requests and reports, if the dfm_app value is not set.

**Execution**

**sched_service**

This is the agent service that the DataMigrator Scheduler will use when running DM flows. If this is not set, the DEFAULT service will be used.

**sched_restart_interval**

Specifies the time interval in seconds between restart attempts for all restartable flows. The default value is 60.

  **Note:** This value cannot be less than 0.

**sched_restart_failed**

Specifies the number of times that failed agents will be restarted. If the value is 0, they will not be restarted. If it is -1, they will be restarted indefinitely. The default value is -1.
**CMASAP_clear_interval**

Specifies how often CMASAP agents are removed from the agents list, in seconds. When set to -1, this is never. The default value is 300.

**Logging and Output**

**sched_log_lines**

Maximum number of lines that the DataMigrator Scheduler is to write to the log for each flow. When the maximum number of lines is reached, logging stops. When set to -1, the number of lines is unlimited. When set to 0, there is no logging.

The default value is set to 20.

**Note:** Some essential lines will still be imposed into the ETL LOG, even when the number of lines (sched_log_lines) is limited to a positive integer. For more information, see *List of Essential ETL LOG Message Codes* on page 692.

**sched_standalone_log**

Select this option to have a separate log file for every job.

**sched_dmlog_app**

The name of the application in which the individual log and statistics files are located. This option is only applicable if you select sched_standalone_log.

**sched_log_commit_interval**

Specifies the maximum time interval in seconds that can elapse before log data is committed to the DataMigrator log. The default value is 10.

**sched_log_commit_maxlines**

Specifies the maximum number of log lines to collect before committing them to the DataMigrator log. The default value is 1000.

**Merge formatted output into log**

Select this option to set the output from a report to go to the log. Only text output is displayed to a maximum width of 80 characters. This is the prior behavior.

**Note:** To view report output, in the browser, right-click on the flow name and click *Logs*. Then, click *Last Output* to view the output from the last run, or select *Log and Statistics*. For Report Range select *All*, and then right-click the job id in the report and click *Show Output*.

**dfm_maxage**

Specifies the maximum number of days that deferred reports are kept.
dfm_maxoutput

Specifies the maximum size of a deferred report.

E-mail Notification

On Start

Sends an email when the Scheduler starts.

On Completion

Sends an email when the schedule completes.

On Failure

Sends an email if the schedule fails.

sched_email_address

Enter an email address, user ID, or procedure name.

dfm_email

If this option is enabled, the schedule specified for server wide email will be followed.

dfm_email_from_userid

If this option is enabled, the email address associated with the user ID will be retrieved from the security provider.

Note: Before you can use email notification, you must configure an email server. For more information, see How to Configure the Server for Email Notification on page 91.

Procedure: How to Start or Stop the Scheduler

1. From the DMC, connect to the server with an administrator user ID.
2. Expand the Workspace folder, and then expand the Special Services and Listeners folder.
3. Right-click SCHEDULER.
   - If the Scheduler is running, select Stop.
   - If the Scheduler is not running, select Start.

Procedure: How to Update Passwords

APPLICATION-level administrators can change their configured passwords from the DMC.

1. In the DMC, right-click the server name, point to My Console, and then click Change Password.
   - The User Information dialog box opens.
2. Enter your current operating system password in the Current Password field.

3. Enter the password in the New Password and Confirm Password fields.

4. Select the Update operating system password or Update password in admin.cfg, server will be restarted check box.

5. Click Update.

**Note:** Changes to the password will take effect the next time the server or scheduler is restarted.
**Procedure:** How to Configure the Server for Email Notification

To enable email notifications, you must first configure an SMTP service. (You need to set up your server for email before it will work anywhere in the flow.)

1. In the DMC, right-click the Workspace folder, point to Workspace SET, and click SMTP Server Settings.

   The E-mail SMTP Server dialog box opens, as shown in the following image.

   ![E-mail SMTP Server dialog box](image)

   **Note:** If an antivirus program is running on this computer, please make sure that the tscom3.exe, edapgwy.exe, and edaplog.exe processes are excluded.

2. Enter the following information:

   - **smtp_host.** Defines the IPv4 address of the SMTP email server. The address can be in standard dot notation (for example, smtpmail.yourcompany.com).
   
   - **port.** Specifies that the mail server is bound to the TCP port that is shown.
   
   - **smtp_user.** The user ID to the SMTP server. This value is optional.
   
   - **smtp_password.** The password to the SMTP server. This value is optional.
   
   - **sender_email.** Defines the email address of the sender, which appears in the 'From' field in emails (for example, yourname@yourcompany.com).
server_admin_email. Defines the email address of the recipient (for example, yourname@yourcompany.com).

3. Click Save and Restart Server.

4. After configuring SMTP services, go to the Web Console Scheduler Configuration page and select a sched_email option:

    - **off**: Never send email.
    - **completion**: Send email when a flow finishes.
    - **failure**: Send email when a flow fails.

   **Note**: This setting applies to every flow.

**Procedure**: How to Edit the Server Configuration

1. In the DMC, right-click the Workspace folder, point to Workspace SET, and then click Miscellaneous Settings.
The Miscellaneous Settings dialog box opens, as shown in the following image.

![Miscellaneous Settings Dialog Box](image)

The following Miscellaneous Settings are saved in the edaserv.cfg file:

**automatic_recovery**

Determines whether the server will automatically recover crashed listeners and/or special services. The default is *n*. 
**transaction_coordination_mode**

Selecting `y`, activates the iWay Transaction Coordinator. A Two-Phase Commit will be used when XA-Compliant databases participate in a transaction. The default is `n`.

**tscom3_kill_delay1**

Defines the time (in seconds) allowed for an agent to stop before soft kill is sent. The default is 16.

**tscom3_kill_delay2**

Defines the time (in seconds) allowed for a soft kill to work before a hard kill is sent. The default is 32.

**crashed_cleanup**

By default, crashed agents are kept in the agents table for diagnostics purposes, and are only removed when they are explicitly killed. When this set to `y`, crashed agents are automatically removed from agents list when the Workspace Manager detects a crash. The default value is `n`.

**crashed_max_dump**

If crashedCleanup is set to `y`, this determines the number of memory dump files retained for later diagnosis.

**stats_runningavg**

Defines the time window in seconds for the calculation of running response time averages. The default is 600.

**htmlencode**

Specifies whether to encode HTML tags in data as plain text in order to prevent attacks by inserting executable code into data. Valid values are `y` and `n`. The default value is `n`.

**proxy_server**

Defines a proxy server, which acts as an intermediary on requests for resources from other servers.

**proxy_port**

 Defines the port number of the proxy server.

**proxy_user**

Defines a user ID for the proxy server.

**proxy_password**

Defines a password for the user ID connecting to the proxy server.
**outbound_ssl_certificate_file**
Required for SSL when using self-signed certificates on i5 systems.

**outbound_ssl_certificate_passphrase**
Required for SSL when using self-signed certificates on i5 systems.

**outbound_ssl_certificate_label**
Required for SSL when using self-signed certificates on i5 systems.

**outbound_ssl_certificate_keyfile**
Required for SSL when using self-signed certificates on i5 systems.

**outbound_ssl_auth_server**
Required for SSL when using self-signed certificates on i5 systems.

**seslog_max_lines**
Defines the upper limit for the number of rows in a session log. When the limit is exceeded, only the latest records will be kept in the session log and the number of records will equal the limit. The default is 1000.

**popup_timer**
Defines the time in seconds to wait before the popup timer appears. The default is 5.

2. Make your changes, and click **Save**.

For details on the server configuration, see the **Server Administration** manual.

**Procedure:** How to Set a Value for Global Variables

1. In the DMC, right-click the **Workspace** folder, point to **FOCUS Sets and Info**, and then click **Core Engine Global Variables**.

   The Core Engine Global Variables dialog box opens.

2. Click **Add New**.

3. From the Command drop-down menu, select **DEFAULT** or **SET**.

4. In the Name box, enter the name for the global variable. Do not use ampersands.

5. In the Expression box, enter the value of the variable, enclosed in single quotes for character values.

6. Click **Save**.

7. Click **Cancel**.
**Procedure:** How to Create a Backup of Server Configuration Files

1. In the DMC, expand the server and then expand the Workspace folder.
2. Right-click the Configuration Files folder and click New Backup.
3. Click OK to confirm the backup.

**Procedure:** How to Restore Server Configuration Files

1. In the DMC, expand the server and then expand the Workspace folder.
2. Right-click the Configuration Files folder and click Manage Backups.
   
   The Configuration Files Backup dialog box opens.
3. Right-click the file you want to restore, or expand it and right-click on a specific file. Click Restore.

**Procedure:** How to Update Application Settings

Application settings control some aspects of use of application directories. You must be a server administrator to change any of these settings, and they require a restart of the server to take effect.

When you connect to the server, a temporary directory called foccache is added as the first directory in the search path. When you want to be able to reuse data within the same session, you can store the data in the form of a HOLD, SAVE, or SAVEB file in the foccache directory. As long as the session remains active, the files stored in the foccache directory can be referenced in requests. Foccache is also used to store some system created temporary files.

To change server application settings:

1. From the DMC, log on with a server administrator user ID.
2. Right-click the Workspace folder, point to Workspace SET, and click Application Settings.
The Application Settings dialog box opens, as shown in the following image.

The following options are available:

**foccache_dir**

Defines the location of the foccache directory. The default location is edatemp/foccache/....

**foccache_maxage**

Defines the maximum number of minutes that a foccache directory will last. When the maximum foccache retention time is reached, the foccache directory is removed by the server and can no longer be referenced. The default is 180 minutes.

**nested_app**

When set to y, nested applications are enabled. This is the default value. To disable, set to n.

**upload_allowed**

Lists the file extensions that are allowed for upload files.
A home application directory provides each user with a private application directory with full access privileges to store private focexecs and data for the test runs and ad hoc reports.

**homeapps**

The operating system specific pathname. Defines the location of home application root directory. If defined, each connected user will see the *My Home* applications folder on the Application tree. User can create new applications (home applications) in this folder. These applications will be visible and available only to the connected user and the server administrators. The user will have full access to their home applications. For this setting to work, nested_app should be set to y.

3. Make any necessary changes and click *Save and Restart Server*.

**Procedure:** How to Make a Global Variable Persistent in a Session

Global variables can be used by various users, regardless of the fact that each user has their own profile, by changing a listener configuration property.

**Note:** This is available for HTTP Listener only.

1. In the DMC, expand the *Workspace* folder.
2. Expand the *Special Services and Listeners* folder and right-click *TCP/HTTP*.
3. Click *Properties of HTTP*.

   The Listener Configuration dialog box opens.

4. Expand the *Sessions Control* section.
5. Select 1 in the PERSISTENT_GLOBAL drop-down menu, as shown in the following image.

6. Click Save and Restart Server.

The values of any global variables will now be kept active as long as you stay connected to the server.

**Using Secure Shell FTP (SFTP)**

Before using SFTP with DataMigrator, you need to set up SSH connections without passwords.
For example, if your DataMigrator Server is on a Linux system called local and you want to retrieve or write files on a server called remote, you should log on to your local system with your user ID, for example, user. Then, you need to generate a key file:

/home/user/$ ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/home/user/.ssh/id_rsa):
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/user/.ssh/id_rsa.
Your public key has been saved in /home/user/.ssh/id_rsa.pub.
The key fingerprint is:
/home/user/$

This process creates two files in your .ssh directory: your private key id_rsa and your public key id_rsa.pub. Then, you need to add your public key to the authorized users on the remote server. You can do this using FTP. Enter the name of your remote server for remote.

**Note:** When you issue the command to create the remote directory (.ssh), it will fail if it already exists. This is not an issue.

/home/user/$ ftp remote
Connected to remote (172.19.23.42).
220 (vsFTPd 2.2.2)
Name (remote:user):
331 Please specify the password.
Password: password
230 Login successful.
Remote system type is UNIX.
Using binary mode to transfer files.
ftp> lcd .ssh
Local directory now /home/user/.ssh

ftp> mkdir .ssh
257 "/home/user/.ssh" created
ftp> cd .ssh
250 Directory successfully changed.
ftp> append id_rsa.pub authorized_keys
local: id_rsa.pub remote: authorized_keys
150 Ok to send data.
226 Transfer complete.
223 bytes sent in 4.2e-05 secs (5.2e+03 Kbytes/sec)
ftp> quit
221 Goodbye.

You can verify if this worked by using SFTP. You should not get prompted for a password.
Your connection is now set up. In the Data Management Console, expand the Adapters folder. Then expand the Configured folder. Right-click Delimited Flat File (or Flat File) and click Add Connection. You can now select SFTP as the transfer protocol.

Using SFTP from Windows

While Windows does not include an SFTP client program, there are programs available from both open source and commercial sources that implement the protocol.

One free source is PuTTY, which supplies a collection of utilities including a client program (psftp.exe) with similar functionality. They also provide puttygen.exe for generating private or public key pairs and converting an existing private key to a file in their .ppk format.

To use PuTTY’s PSFTP:

1. Go to the PuTTY Download Page, download the MSI installer, and run it to install and configure their utilities. Alternatively, download psftp.exe and puttygen.exe to a directory in your PATH.

2. Download a private key to your Windows system or generate a new key using puttygen. Save it to a directory of your choice, for example, C:\putty\id_rsa.ppk.

3. If necessary, add your public key to ~/.ssh/authorized_keys on the remote server.

4. Test the connection from the command line using the following lines:

   drive:\> psftp -v -i "drive:\putty\id_rsa.ppk" user@remote
   Using username "user"
   The server's host key is not cached in the registry
   ...
   Store key in cache? (y/n) y
   Remote working directory is /home/user
   psftp> ls
   psftp> quit

   where:

   drive
   Is the drive on which the private key is saved.

   user
   Is your user ID.

   remote
   Is the name of your remote server.
Upgrading DataMigrator From Release 4.3 or 5.1

You should not be prompted for a password or passphrase. However, if you have not previously connected to the server, you will be prompted to store the key in the cache. Reply y and you will not be prompted again. You must do this so that the DataMigrator Server does not receive this prompt.

5. In your PATH, create a file called sftp.bat with the following lines:

```bash
@echo off
psftp -i "drive:\putty\id_rsa.ppk" %1 %4 %5 %6 -batch
```

where:

`drive`

Is the drive on which the private key is saved.

If your DataMigrator Server is running, shut it down and then start it again. You can now configure a connection for a (delimited) flat file for a remote connection using SFTP.

Upgrading DataMigrator From Release 4.3 or 5.1

To upgrade from predecessor products Copy Manager or ETL Manager, you must first upgrade to Release 5.2, 5.3, 7.1, or 7.6 following the instructions in the appropriate DataMigrator User’s Guide.

Upgrading DataMigrator from Release 5.2 and Higher

For most prior releases, upgrading to Release 7.7 always results in a new installation of the DataMigrator Server. This allows you to run your new and old server installations in parallel, if needed.

**Note:** This is not the case when replacing Release 7.6. For more information on upgrading from that release, see *Upgrading DataMigrator from Release 7.6* on page 105.

This section lists some things to consider when upgrading DataMigrator. Refer to *Revised Functionality* on page 36 to review functional changes found in Release 7.7 to see if they affect your flows.

Port Number

Each DataMigrator Server listens on a unique series of port numbers for connections from the Data Management Console, a web browser, and other IP services. By default, port 8116 is used for TCP services and 8117 through 8119 are used for HTTP services.

To have multiple servers running on the same system, such as the 7.6 and 7.7 servers, they must each have their own unique set of port numbers.
Configuration Files

A new installation means that you will need to recreate or migrate your configuration information. Consider whether you want the new and old installations to use the same database connections. For instructions on how to migrate the configuration files, see *How to Migrate Configuration Information* on page 103.

**Note:** If you choose to migrate configuration information from an older release, you should do so before entering any new configuration settings. This prevents overriding existing configuration information by the migration process.

Profiles and Applications

If you have installed your new server installation on a different system from the prior release, you may also need to migrate your APPS directories where flows, procedures, and metadata are stored. You may also need to migrate your PROFILES directory, which is where individual user profiles are stored. This can be done by using operating system copy commands or from Explorer on Windows.

Data Flows

Data flows and procedures that ran in prior release should run in this release without modification. They will be automatically upgraded to take advantage of new functionality when they are opened and resaved from the DMC.

Scheduler

If you have scheduled flows, you should consider whether you want the scheduler on both the new and old releases to initiate scheduled flows from the same application directories.

Data Management Console

The Data Management Console (DMC) and Server releases should both be the same. When upgrading the DataMigrator Server, the new version of the DMC must also be installed for each user that connects to the server.

**Note:** Two or more installations of the DMC can be on one PC in order to connect to different server releases.

**Procedure:** How to Migrate Configuration Information

To migrate configuration information from a previous release:

1. In the navigation pane of the DMC, expand the server node and then right-click the *Workspace* folder.
2. Click Migrate.
   The Migrate dialog box opens.

3. Enter the full path of the configuration instance directory (EDACONF). On Windows, for example, this path is C:\ibi\srv76\dm.

4. Click Migrate.

**Procedure:** How to Migrate DataMigrator From One System to Another

**Note:** If the new DataMigrator Server has been installed with security and data adapters have been configured, you do not need the configuration files from the previous environment.

1. Locate the following files:
   - edaserve.cfg
   - edaenv.cfg (If it exists)
   - edasprof.prf

2. Backup these files using the DataMigrator Server Web Console:
   a. Click the Workspace tab.
   b. Right-click the Configuration Files folder and click New Backup.
   c. Click OK.

3. Copy the following files into a directory path on the new environment so they can be used with the migration process. The paths used are examples for a Windows system:
   - From the C:\ibi\srv77\dm\bin directory into the C:\ibi\cnfg\bin directory:
     - edaserve.cfg
     - edaenv.cfg (If it exists)
   - From the C:\ibi\srv77\dm\etc directory into the C:\ibi\cnfg\etc directory:
     - edasprof.prf

4. After migrating the configuration files, start the DataMigrator Server.

5. Place the following files from the previous environment into the new environment. They should be manually copied and placed into the same directories in the new environment:
   - From the C:\ibi\profiles directory on the old machine, the admin.cfg file and any user profiles (.prf) files.
   - From the C:\ibi\apps directory on the old machine, all of the sub application directories.

Upgrading DataMigrator from Release 7.6

If you have a previous installation of DataMigrator Release 7.6 installed on your system, a message like the one below will appear that gives you the opportunity to replace your current installation with Release 7.7.

![Question dialog box](image)

If you click No, then a new installation is performed, as described above. This the preferred option for most users, as it allows for parallel operation if needed.

If you click Yes, then your 7.6 installation is replaced with a 7.7 installation. You will then have a Release 7.7 installation in a directory identified (by default) as SRV76, but you will not be able to use Release 7.6 unless you re-install it.

In this case, there is no need to migrate your configuration files. However, the DataMigrator Log and Statistics table formats have changed and they must be re-created, as explained in How to Recreate the DataMigrator Log and Statistics Tables on page 106.

Upgrading DataMigrator to Release 7.7.07 or 8.2

In this release, the length of the application path names have been increased from 64 to 256 characters. This requires that the DataMigrator Log and Statistics tables are recreated.
Procedure: How to Recreate the DataMigrator Log and Statistics Tables

You can recreate the DataMigrator Log and Statistics tables using a Server Administrator user ID from the Data Management Console or the Web Console, by following these steps:

1. Click the Workspace folder or tab.
2. Expand the Logs and Traces folder.
3. Right-click DataMigrator Log and Statistics and click Recreate.
4. Click OK in the Confirm dialog box.
Planning a DataMigrator Application

Although DataMigrator facilitates all aspects of your design task, experience has convinced us that the results of your efforts will be more rewarding if you do a little advanced planning.

This chapter provides important background information that will help you make optimal decisions to fulfill data management needs of your site.

In this chapter:

- Designing Your Application
- Using Synonyms for Data Sources and Data Targets
- Reviewing Your Data Sources
- Reviewing Your Data Targets
- Establishing Security for DataMigrator

Designing Your Application

Data warehouses and/or data marts are created for a variety of reasons. Knowing the factors leading to the decision for extracting or moving data will help you identify the issues for DM management that will have the greatest impact on your success. The ultimate goal is always to have an automated process that moves and converts all of your data in a very small period of time. However, it helps to know what issues are most important in your company’s criteria for a successful data warehouse. Are you consuming large amounts of time in the management of 3GL extraction programs and load routines? Are you uncomfortable with ad hoc user access to production legacy systems? Is your network unable to handle the performance requirements for decision support and analysis?

If you consider the reasons for warehousing, your attention will be properly focused on the critical variables. DataMigrator provides you with the flexibility to control the variables inherent at your site and in your data.
From an application perspective, you must evaluate what will direct your overall data movement strategy. Considering the following questions will help you learn more about the capabilities of DataMigrator and how to control them.

- **What is the typical volume of data being migrated?** One million bytes, 100 million, or several gigabytes of data?

- **How often will data be migrated?** Once every night or every 15 minutes? A single large load once per month?

- **What time of day or night is optimal for moving the data?** Some sites cannot get access to data until after the nightly production jobs are complete (which may be at 5:30 AM). When is the warehouse required? Are different time zones involved? This complex question impacts not only your performance requirements, but also your scheduling methodology.

- **Is there a key value in the source data that will identify changed rows?** DataMigrator can locate changed data using an SQL WHERE criteria, unless an extract file exists with delta changes. Is this file available? Do you know what the WHERE criteria should look like? Using Change Data Capture, an optional, add-on component available for selected relational databases, you can load a data target with only the records that have changed.

- **How many designers will be developing extractions with DataMigrator?** Some sites designate a single DataMigrator administrator who is responsible for all the extract and load routines that are used to develop their data marts. (There are rarely any conflicts in such an environment; one user has complete control over the requests, the source data, and the target locations that will be written.) Other sites have a team of developers interacting to create a collective data warehouse.

### Using Synonyms for Data Sources and Data Targets

iWay Servers, including the DataMigrator Server, access local and remote data sources and data targets using synonyms, which contain the metadata for a data source. This metadata consists of the attributes that describe the data source or target.

Since DM processing depends upon the appropriate metadata being available to the DataMigrator Server and any remote servers in your configuration, you must create synonyms for all data sources and any existing data targets before you can use them in a flow. For information about providing metadata to servers, see *DataMigrator Setup At a Glance* on page 42 and *Remote Server Setup At a Glance* on page 44.

The Data Management Console application directory tree lists the synonyms available to the DataMigrator user for data sources and data targets. This structure is also known as the users Application Path on the server.
Reference: General Guidelines for Synonyms

- The first character in a synonym name must be a letter. The remaining characters can be any combination of letters, numbers, and underscores (_). The synonym name cannot exceed 64 characters. When naming a synonym, you must avoid the following:
  - Names containing capital letters.
  - A name beginning with the letters ETL or SYS. These prefixes are reserved for DataMigrator internal use.
  - In a FOCUS/FDS or XFOCUS data source, a non-unique field name.
  - Names that are used as commands within DataMigrator. These are HOLD, SQLIN, SQLGET, and MODIFY.
  - Names containing spaces, dashes, and other non-standard characters in both the source and target data columns (*, &, #, $, %, @).

- Synonyms must not exceed any of the following:
  - Approximately 3,000 columns.
  - 30K of total column name character length.
  - 16K for single LRECL (total length of field formats, one instance of a row).
  - Column names cannot exceed 128 characters.

See the Adapter Administration manual for suggestions on how to access your structures without pressing these limits.

- For column names, avoid the following:
  - Special characters. Only alphanumeric characters and underscores are permitted in column names.
  - Names beginning with the capital letter E followed by a two or three digit number, for example, E01 - E999.
  - Reserved words, including:
Synonyms created using the Dynamic Columns option (NOCOLS parameter) should not be used with DataMigrator because their columns and formats will not display when building flows in the DMC.

**Reference:** Inclusion of Application Directory Names With Flow Components

DataMigrator Server supports inclusion of application directory names in flow component labels. This means that data and process flows can store information about the application directories where the synonyms or procedures that they use are located.

Prior versions of DataMigrator used application directories to group synonyms and requests (flows) into logical groups, and by default did not provide support for duplicate or multiple synonyms or requests with the same name, even if they were in different application directories.

For example, if you had two synonyms or requests with the same name and your application path was

```plaintext
ibisamp
demo
baseapp
```
it would not matter which synonym or request you selected. You would get the one from the application directory that appeared first in the application path, since the earlier version of the server did not support inclusion of the application directory. If you had synonyms or requests with the same name in more than one application directory, there would be no way to distinguish them, and the server would always use the one in the first application directory in the application path. Some users take advantage of this capability by constructing a run-time application path to control which synonym or flow is actually used.

For upward compatibility, when you build data flows, by default, they will not include application directory names. That means that the flow will not save the application directory name with the source or existing target. If you use a synonym that exists in more than one application directory, you will get a warning, which is shown in the following image.

![Duplicate File](image)

The flow will pick up the source from the first application directory in the application path.

To include application directory names in data flows, on the Home tab, in the Tools groups, click Options. Click Data Flow and select the Use application directory name with flow components check box from the Data Flow Designer page. This way your flow will use exactly the synonym you specify.

For process flows, the default is to include application directory names.
Reference: Guidelines for Variable Length Character and Text Columns in Synonyms

If you are using variable length character or text columns in your synonyms, you should be aware of the following special considerations.

Variable Length Character Columns

DataMigrator supports variable length character columns that are described in a relational database as VARCHAR. This attribute enables you to copy such columns without padding them with spaces to the full column width or removing trailing spaces. It enables you, for example, to copy a column containing a single space.

This support is enabled by default.

When new synonyms are created, they are created with variable length columns described as

\[ AnV \]

where:

\( n \)

Is the maximum column width.

Note: For flat file data targets, the length of the value precedes the data value itself.

TEXT (TX) Columns

DataMigrator also supports columns described with ACTUAL=TX for Text columns. This includes data types MS SQL Server VARCHAR(MAX) and Text, ORACLE LONG, and Db2 LONG VARCHAR. Sample Data displays only the first 50 characters of these columns.

Note: Columns described as TX cannot be used as a key column.

Reviewing Your Data Sources

Aspects of your data sources contribute to the performance of DataMigrator. Consider the following issues when reviewing your data sources:

- **Number of columns in an extract.** iWay supports about 400 columns in an SQL command. Previous servers, however, support fewer. If more columns are required, you may need to have multiple flows or concatenate columns.

- **Number of tables.** The DataMigrator Server supports joining of up to 63 data sources, but the complexity increases as the number of files increases. Thorough knowledge of the data is required to determine if a multi-table join makes sense in any given environment. As a general rule, try to keep the number of tables to a minimum.
Location of your tables. Is a cross-platform join required (for example, an z/OS table to a UNIX table)? Is all the data in one data source on one platform? Perhaps a cross-DBMS join (for example, C-ISAM to Oracle) is needed? If so, this may impact performance.

Data types. What data is being moved to the warehouse? Are all the data types supported by the current server environment? Also, be sure that the data adapter supports the ability to read and extract the data type required by your warehouse. Be sure that the joins you require can be supported by the data types of your columns. Consider the scale (number of decimal points) on floating-point columns, or site-specific columns that are often found in old data sources, such as company-specific date fields that require the use of a home-grown assembler routine.

Accessing Data Sources

You can use DataMigrator to access data sources across an enterprise and move data from those data sources into various data targets. The DataMigrator Server accesses any iWay-supported data source, which includes numerous formats on over 35 platforms.

The DataMigrator Server can access data that resides locally, that is, on the DataMigrator Server itself. This is the simplest configuration, but not necessarily the most typical.

If all or some of your source data is located on one or more remote servers, the DataMigrator Server establishes a connection with the remote server(s) to access this data.
The following image illustrates how data can be moved or extracted to the DataMigrator Server from a source that is local to the DataMigrator Server.
The next image illustrates how data can be moved or extracted to a data target on the DataMigrator Server from a data source that is local to the DataMigrator Server, as well as one that is located on a remote server. With this configuration, despite the different locations of the data sources, the DataMigrator Server can join them together and access the joined files.

**Reviewing Your Data Targets**

The data targets you use for your application affect the performance of DataMigrator. Consider the following issues that pertain to the tables or files being written by DataMigrator:

- **Row processing logic.** Is this a total refresh? Will the request be deleting all rows or dropping the table before starting? Is an update or delete necessary? The processing logic is specified in your data flows.

- **Location of the tables.** Are the tables local to the machine where the server will be running? What connection information will be needed to resolve the names for remote data sources? Can it be handled in the Access File for the synonym? The DMC allows you to specify these items, but in many cases it is best to allow the iWay Hub Server to manage the location of tables. Review the server profile (EDASPROF.PRF) or the profile for your user ID, as well as the Access File that is used with the DataMigrator Server to establish table names and their storage locations.

DataMigrator extracts the data from a data source and loads it into a data target, which may be:

- A relational data target. For details, see *Relational Data Targets* on page 116.
A non-relational data target. For details, see Non-Relational Data Targets on page 117.

Remote destination. For details, see Remote Target Destinations on page 117.

Relational Data Targets

Depending on the platform, supported relational data targets currently include the following: Access, Db2, Informix, MySQL, Nucleus, ODBC, Oracle, SQL Server, Sybase, Red Brick, and Teradata.

In the following diagram, DataMigrator creates a target table or writes to an existing target table in a local RDBMS which resides on the DataMigrator Server (the recommended configuration for a data target).

Note: EDA is only supported for existing targets. The name of the remote server must match the real server name of the subserver. This option should only be used for small data volumes. For large volumes, you should install the DataMigrator server on the platform where the data will be loaded.
Non-Relational Data Targets

You can create non-relational data targets of the following types on the DataMigrator Server (the recommended configuration for a data target): a FOCUS/FDS or XFOCUS data target, or a flat sequential file of different formats. A flat sequential file is created in the character set appropriate for the server platform (for example, ASCII on UNIX or EBCDIC on z/OS). In the following diagram, DataMigrator creates or writes to an existing non-relational data target.

Remote Target Destinations

Many relational data source management systems provide for remote table access. If properly configured, the DM server supports the syntax necessary to write to remote targets (for example, SQL*NET provides access to remote Oracle tables).
The following figure illustrates how a DBMS located on the DataMigrator Server uses networking capabilities to write to a remote data target.

**Reference:** General Guidelines for Data Targets

- DataMigrator supports up to 64 character for the synonym name used with a data target or the limit set by the operating system or DBMS. The 64-character limit is for the total name, including data source and owner details.
  
  For example, if you are loading a relational data target, the full name may exceed 64 characters since you can add data source and owner details to the synonym name.

- Avoid the following when naming data targets:
  
  - Names beginning with the capital letter E followed by a two or three digit number (for example, E01 - E999).
  
  - Names beginning with the capital letter T followed by a number (for example, T1 - T16).

- When you write or move data to an existing relational target, ensure that the correct access rights are assigned to the approved group of designers. New relational targets generally default to assigning an owner ID equal to the logon ID, although this may vary based on the type of data source selected.

- DataMigrator is often used to create data targets in the local relational subsystem. This requires that the user ID processing flows has appropriate access to the target RDBMS.

  - If the data target is a new table, the user ID must have permission to issue CREATE TABLE statements.
If the data target is an existing table, transaction update access is necessary.

- Relational data targets created in DataMigrator are owned by the user ID specified in the connection string in the server profile.

- Flat, FOCUS/FDS, and XFOCUS file data target attributes are controlled independently by the designer. Therefore, they may only create conflict if users select the same file name (operating system dependent) for their output. This can be avoided by establishing a convention for where users will write their extracted answer sets.

- Where flat file data targets are being created, the user ID of the processing request must have access to the subdirectory or data set that is currently allocated to receive the answer set. Flat files created in DataMigrator are owned by the server user ID.

**Note:** For the best throughput loading relational tables, the DataMigrator server should be on the same system as the RDBMS server. Throughput to a remote server is dependent on your network speed and traffic.

### Establishing Security for DataMigrator

If only a single DataMigrator user will be designing flows, no alterations to the DataMigrator Server configuration are necessary. However, if the DataMigrator Server supports multiple users, the server administrator may need to establish separate user IDs and profiles to control access to DM application directories.

**Note:** Establishing security for iWay Servers, including the DataMigrator Server on z/OS requires additional consideration. For details, refer to Step 7. Configure Server Security in Chapter 4, *Server Installation for z/OS* in the *Server Installation* manual.

### Running Flows

The iWay agent created when you connect to a DataMigrator Server has an associated logon user ID. Local file, directory, and resource security is controlled by that user ID.

For scheduled flows, the DM components that a user ID can see and run from its Application Path are controlled from `sched_run_id` option, accessible using the Type setting in the New Tenant dialog box, accessed from the Manage Scheduler Userids dialog box.

When `sched_run_id` is set to:

- **server_admin_id.** Scheduled DM jobs are run under the first ID that appears in the list of server administrators displayed on the Access Control page. If the ID does not have a password specified in the Access Control tab, a profile for that user ID must be created. `server_admin_id` is the default.
Scheduled DM jobs are run under the user ID that was used to save the flow. The Application Path specified in the user profile is utilized.

Scheduled DM jobs will run under the specified user ID. If you select this option, an additional drop-down menu appears where you can select a user ID from a list of configured user IDs.

If security is ON and you set sched_run_id to user to run a scheduled flow for a certain user ID:
1. The user ID must be a valid user on the system.
2. The user ID must be set to an access level of either SERVER or APPLICATION from the Access Control page.
3. The password for the user must be set. A SERVER-level administrator can set the password for a SERVER-level ID from the Access Control page when adding a user.

If a DataMigrator user with an APPLICATION-level ID wants to run scheduler requests:
1. A SERVER administrator must make them an APP administrator from the Access Control page.
2. The user must set their password on the User Information page.
3. The scheduler must be restarted. (Restarting the server will also restart the scheduler.)

For more information, see Scheduler Configuration Dialog Box on page 86.

Restricting the Application Paths Available to a User

By default, the server profile (EDASPROF.PRF) is run for all users when they connect to the DataMigrator Server to provide access to all application directories in the servers search path. However, an administrator can control a user’s access to application directories by creating individual user profiles. Each user can then:

- Access only the application directories specified in the Application Path specified for that profile.
- Use synonyms in the specified Application Path.

For details, see Authorizing DataMigrator Server Usage and Administration on page 53.

It follows that the user ID that a flow runs under determines the user profile that is run. The profile controls the application directories available to the flow, as well as access to relational databases or source servers.

- If there is a profile associated with the user ID, then it is used.
- If there is no profile, then EDASPROF is used instead.

The user can only access the application directories defined in the profile being used.
You can set the Application Path from the DMC or the Web Console. For information on setting the Application Path from the DMC, see *Managing Application Directories and Configuring the Application Path* on page 57. For information on setting the Application Path from the Web Console, see the *Server Administration* manual or the Web Console online help.

**Running Scheduled Flows Under a User ID**

By default, scheduled flows are run using the server admin ID.

To run all scheduled flows under the user ID that saved them, you need to:

1. Define a scheduler tenant and change the sched_scan_id and scan_run_id.
2. Create a new user (if the user ID does not already exist). This procedure will depend on your operating system.
3. Add users who can run flows as an Application Administrator.
4. Have the new users change their security settings.
5. Connect to the server as the new user in the DMC, schedule a flow, and save it.

**Procedure: How to Define a Scheduler Tenant and Change the sched_scan_id and scan_run_id**

1. Connect to the server with a Server Administration user ID.
2. In the navigation pane, expand the server, followed by the *Workspace* folder.
3. Expand the *Special Services and Listeners* folder.
   
   **Note:** If there is a *Start* option, the scheduler is not running. To run the scheduler, click *Start*.
4. Right-click *SCHEDULER* and click *Manage Tenants*.
   
   The Manage Scheduler Userids dialog box opens.
5. Click *New*.
   
   The New Tenant.
6. To set the sched_scan_id for the tenant, select a user ID from the ScanID drop-down menu, which lists the configured user IDs from admin.cfg.
7. To set the sched_run_id for the tenant, select one of the following options from the Type drop-down menu:
   
   - Select *server_admin_id* to run all jobs as the Server Admin ID, which is the first user ID in admin.cfg.
   - Select *user* to run each job as the user that saved the flow.
Select RunID to specify a user ID for all jobs. If you select this option, an additional drop-down menu appears where you can select a user ID from a list of configured user IDs.

8. Click OK to configure the tenant.
9. Click Save and Restart Scheduler.

**Note**: It is recommended that each ScanID has its own Application Path set in their user profile. The same application directory should not appear in the application path of more than one ScanID.

**Procedure**: How to Add the New User as an Application Administrator

**Note**: If you want to run all scheduled flows under a user ID that does not already exist, you must create one using an operating system-specific procedure.

1. In the DMC, expand a server and then expand the **Access Control** folder.
2. Expand the **Roles** folder, right-click **Application Administrator**, and click **Register PTH <internal> User**.
   
   The User Registration dialog box opens.
3. In the User ID box, type the new user name.
4. Optionally, type a description and the email address of the user.
5. Click **Next**.
   
   The User Registration dialog box opens.
6. From the Inherit Privileges from drop-down menu, select **Application Administrator**.
7. Click **Register**.
8. Click **OK** to save your changes and register as a new user.

**Procedure**: How to Change a Password for Running Scheduled Flows

1. Log in to the DMC with an Administrator user ID.
2. In the navigation pane, expand the server and then the **Access Control** folder.
3. Expand the **Roles** folder and then expand the folder of the desired Role.
4. Right-click the user ID you want to manage and click **Properties**.
5. In the Optional password for scheduled runs section, type the new password, and re-type it to confirm the password.
6. Click **Update**.
**Procedure:** How to Connect to the Server as a New User and Schedule a Flow

1. In the DMC, right-click the server and click *Properties*.
2. Change the *User ID* and *Password* boxes in the Security section to the new values and click *Save*.
3. Disconnect and reconnect the server.
4. Open a process flow in the DMC and add a Schedule to the Start object.
5. Save the flow.

The Scheduled Events report will now list scheduled flows by the user ID that saved them.
Working in the Data Management Console

This topic describes the Data Management Console interface and some key tasks you will perform as you design your flows.

In this chapter:

- Data Management Console Interface
- Navigation Pane
- DataMigrator Workspace
- Text Editor
- Aerial View
- Ribbon Tabs
- Metadata Tab
- Text Editor Tab
- Console Log
- Setting Flow Properties
- Viewing Properties
- Setting User Preferences
- User Privileges
- License Code
- Configuring National Language Support
- Ending a Data Management Console Session

Data Management Console Interface

The Data Management Console integrates all of the major tasks required to build flows. You can build flows, access the Web Console, and create, manipulate, and access the metadata associated with those flows. You can also create and run stored procedures.
The DMC interface has a navigation pane on the left, an area that will become the DataMigrator workspace on the right, and the Console Log on the bottom.

The interface contains:

- Servers that can be accessed by the DMC. The server application directories appear when you are logged on to a server. Until you connect to a server, only the servers are displayed.

- Application Directories listing the directories in the server path, which contain flows, procedures, synonyms, reports, adapters, and administrative functions.

- Application Directory Folders that contain flows, synonyms, procedures, and reports.

- Adapters and Workspace folders appear if you are logged on with a Server Administrator user ID. Monitor folder appears if you are logged on with an Application Administrator ID.

- The Console Log that displays the console log, including, but not limited to, messages when you connect and disconnect from the server, or execute flows and stored procedures.

- The ribbon that represents the most commonly performed activities in the DMC. The ribbon tab changes to reflect the currently open object.

- The Home tab in the ribbon, which lets you select which server objects are displayed, and which flows, procedures, or synonyms to use. The tab also lets you select explorer view when displaying information.

- The workspace where flows are built, synonyms are edited, reports are viewed, and the explorer view is shown.
Application Menu

In the upper-left corner of the Data Management Console, click the Application button to open the Application menu.

From this menu, you can:

- Create a new object.
- Save an active document.
- Save an active document under a new name.
- Preview or print an active document.
- Refresh the explorer tree.
- Delete a directory.
- Close an active document.
- Access the Options dialog box where you can set user preferences.
- Close an active document and exit the DMC.

**Note:** In the Application menu, the terms document and active document represent all files that you can create in the Data Management Console.

**Tip:** Double-click the Application button to exit the DMC.
The following image shows the Application menu.

Quick Access Toolbar

The Quick Access Toolbar provides access to the most commonly used functions. It is located to the right of the Application menu button and is always visible no matter which options are selected.

The Quick Access Toolbar contains the following buttons, as default. Other buttons can be added as required.

- **New.** Creates new flows, stored procedures, synonyms, specialized flows, or server connections, as well as configuring adapters.
- **Save.** Saves the current procedure under a specified file name.
- **Save As.** Saves the current procedure under a specified file name.
**Refresh.** Refreshes the Explorer tree in the navigation pane.

On the Quick Access Toolbar, click the arrow to open the Customize Quick Access Toolbar menu, as shown in the following image.

![Customize Quick Access Toolbar](image)

From this menu, you can click:

- The check marks next to the New, Save, Save As, and Refresh options to add or remove these buttons from the toolbar.

- **More Commands** to select and add new commands to the toolbar.

- **Show Below the Ribbon** or **Show Above the Ribbon** to move the toolbar above or below the ribbon.

- **Minimize the Ribbon** to hide the ribbon.

**Note:** You can right-click the ribbon, groups, or commands on the ribbon, to access a modified version of the Customize Quick Access Toolbar menu. This menu contains the Add to Quick Access Toolbar, Customize Quick Access Toolbar, Show Quick Access Toolbar Below the Ribbon, and Minimize the Ribbon options.

**Ribbon**

The DMC ribbon replaces traditional toolbars with a single, rectangular region that spans the width of the application window.

From the ribbon, you can:

- Select DMC commands to build and customize applications.
Customize the appearance of the DMC interface from the Style menu.

Open the Help menu to view the DMC technical documentation and online Help.

The following image shows the default ribbon in the Data Management Console.

![Default ribbon in DMC](image)

**Ribbon Organization**

The DMC ribbon is composed of tabs. Each tab represents a subset of controls and commands that are organized into related groups depending on function or relationship. All tabs are named, all groups are labeled, and all icons represent a control or command.

**Ribbon Tabs**

When you launch the DMC for the first time, only the Home tab is visible. The Home tab is the only static tab and does not close or reopen as you work within the console. All other tabs on the ribbon are contextual, and will only open or close as you develop using the associated features.

The DMC tabs include:

- **Flow.** Opens when you are working with a data or process flow.
- **Direct Load Flow.** Opens when you are working with a Direct Load flow.
- **Text Editor.** Opens when you are editing files as text.
- **Metadata.** Opens when you are working with a synonym in the Synonym Editor.
- **Report.** Opens when you are viewing a report.

**Style Menu**

In the upper-right corner of the ribbon, click the Style menu.

Select a classic theme to customize the Data Management Console. The themes available to you are determined by the themes you have installed for Microsoft® Office. The Style menu is always available.

**Note:** When you close the DMC, the theme you selected is saved, and appears the next time you launch the product.
About Button

In the upper-right corner of the ribbon, click the About button.

Select Help Topics to access the online Help system or click About to view the product licensing information. The Help button is always available.

The following image shows the About button.

Note: You can also press the F1 button on your keyboard to access the online Help system.

Shortcut Keys for Navigating the Workspace

Shortcut keys enable you to move around the DMC interface using key strokes instead of a mouse. If your keyboard shortcuts are not underlined in the DMC, you can change the settings on your Windows system to enable the underlining of such keys by performing the following steps:

1. From your Windows Control Panel, click Ease of Access Center.
2. Click Make the keyboard easier to use.
3. Select the Underline keyboard shortcuts and access keys check box.
4. Click Apply.
5. Click OK to close the window.

The following combinations are available:

<table>
<thead>
<tr>
<th>Key Combination</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alt + 0</td>
<td>Moves focus to the navigation pane.</td>
</tr>
<tr>
<td>Key Combination</td>
<td>Result</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------</td>
</tr>
<tr>
<td>Alt + 1</td>
<td>Opens the New dialog box.</td>
</tr>
<tr>
<td>Alt + 2</td>
<td>Opens the Save dialog box.</td>
</tr>
<tr>
<td>Alt + 3</td>
<td>Performs the action of the 3rd item on the Quick Access Toolbar.</td>
</tr>
<tr>
<td>Alt + 4</td>
<td>Performs the action of the 4th item on the Quick Access Toolbar.</td>
</tr>
<tr>
<td>Alt + 8</td>
<td>Shows or hides the browser.</td>
</tr>
<tr>
<td>Alt + 9</td>
<td>Shows or hides the Console Log.</td>
</tr>
<tr>
<td>Alt + U</td>
<td>Toggles full screen mode.</td>
</tr>
<tr>
<td>Shift + F10</td>
<td>Opens the right-click menu.</td>
</tr>
<tr>
<td>Menu key</td>
<td>Opens the right-click menu.</td>
</tr>
<tr>
<td>Ctrl + 2</td>
<td>Shows or hides the status bar.</td>
</tr>
<tr>
<td>Ctrl + E</td>
<td>Shows or hides explorer view.</td>
</tr>
</tbody>
</table>

### Navigation Pane

The navigation pane on the left side of the interface contains servers, which in turn contain application directories. The application directories contain flows, procedures, reports, synonyms, adapters, and administrative functions.

The navigation pane has three or more top-level folders, depending on your administration level:

- **Application Directories folder.** Provides access to flows, procedures, and reports that can be generated from a flow. It also provides access to synonyms that can be used as data sources and data targets. For details, see *Navigation Pane: Procedures* on page 195.

- **Adapters folder.** Allows you to add and configure adapters. This folder is only available if you are logged in as an Application Administrator.

  For details, see *Configuring Data Adapters* on page 65.
Resource Management. Allows you to configure Resource Management, if you have a license.

Workspace folder. Allows you to configure the server. This folder is only available if you are logged in as a Server Administrator.

Access Control. Allows you to control what users can access the server and their access level. This folder is only available if you are logged in as a Server Administrator.

Scalability. Used for performance testing.

Monitor folder. Allows you to monitor server operations. This folder is only available if you are logged in as an Application Administrator.

Reference: Server Shortcut Menu Options

Server shortcut menu options are available if you right-click a server or highlight one and click the Menu key. The following options are available:

Connect/Disconnect. Connects or disconnects the DMC to or from the server.

Start/Stop Server. Starts or stops the server from running. This option is only for a local LOOPBACK server.

Alternate Authentication. This option is only available if a third party cookie is used, such as Netweaver.


Other Options. The following options are available:

My Console:

My General Privileges. Shows the user privileges associated with the selected server.

Note: This option is only available when connecting to a server with security enabled.

Manage My Agents. Runs the Data Services Agents report that shows all currently active agents.

Save My Last Report. Allows you to save a focexec with contents USERINFO.

Edit My profile. Allows you to edit your server profile.
Note: This option is only available for a Server Administrator or when connecting to a server with security disabled.

- Login Info. Displays your current log in status including your user ID, profile, and security role.

- Change Password. Allows you to change your password on the server.
  Note: This option is only available when connecting to a server with security enabled.

- Stop My Requests. Stops all Web session requests for the server.


- Help. The following options are available:
  - Contents and Search. Displays the Help system.
  - Version. Displays information about the release of the server that you are connecting to, and the operating system.
  - New Features. Displays new feature documentation.
  - Licenses. Displays information about software licenses for Information Builders or 3rd party products.

- Server Session Log. Displays the server log.

- Refresh. Refreshes the server application directories.

- Delete. Deletes the server node from the DMC.

- Copy Server Node. Opens the Server Node Configuration dialog box, which displays information for the selected server. The Node Name is prefaced by COPY_ and can be edited to reflect the desired name.

- Properties. Provides access to the server connection configuration parameters. For more information, see Server Node Configuration Dialog Box on page 169.
Reference: Directory Shortcut Menu Options

Directory shortcut menu options are available if you right-click any of the application directories in the navigation pane. The following options are available:

- **Refresh.** Refreshes the list of objects in the directory.
- **New.** Opens a sub-menu that allows you to create a new object in the selected directory.
- **Paste.** Moves or copies a directory to the selected location. This option is only available after you have copied or cut a directory.
- **Quick Copy.** Creates a simple Data Flow that selects all rows and columns from a source table.
- **Schedule and E-mail.** Has the following options:
  - **Scheduler Agents.** Shows all flows and procedures that contain a schedule, as shown in the following image.

![Scheduler Agents screenshot]

**Note:** This report runs for the user whose ID value is stored in the sched_scan_id of a Scheduler. The Scheduler must be active to run the Scheduler Agents report.

For more information on the commands available from the Report tab while viewing a Scheduler Agents report, see *Reports Tab* on page 158.

- **Scheduled Events.** Opens a dialog box where you can view jobs scheduled to start during a specified time range. The Scheduler must be active to run the Scheduler Events report.

  When you click View Report, all jobs scheduled to run during the time range are shown.

  For more information about scheduler options, see *Schedule Attribute Pane* on page 427. For more information on email notifications, see *How to Set the Flow Email Notification Properties* on page 165.
Log and Statistics. Opens the Log and Statistics dialog box that enables you to view a log or statistics report. The following options are available:

- **Report Type.** Allows you to select either a log or statistics report.
- **Report Range.** Allows you to select to view the last report, all reports, or a report specified within a select time range.
- **Completion Status.** Allows you to view all reports, all completed reports, all completed successfully reports, or all completed reports that contained errors.
- **Run by.** Allows you to view reports for all flows, including those initiated by the scheduler or submitted manually. This includes flows initiated from the DMC, WC, and CMRUN.
- **Application or Flow Name.** Allows you to select a specific application or flow name for the report.

Impact Analysis. Opens a sub-menu that allows you to view impact analysis reports. For more information, see Impact Analysis Reports on Synonyms, Procedures and Columns on page 475.

- **Copy.** Copies the selected directory.
- **Delete.** Deletes the selected application directory.
- **Rename.** Allows you to rename the selected application directory.
- **Cut.** Cuts the selected directory.
- **Explorer.** Opens an explorer window in the workspace area.
- **Privileges.** Shows the user privileges associated with the selected directory.

  **Note:** This option is only available when connecting to a server with security enabled.

- **Properties.** Opens a properties window for the selected directory that displays the following information:

  - **Location.** Physical location of the application directory on the disk.
  - **Number of files.** Number of files in the application directory.
  - **Last modified.** Date and time the last changes were made to the directory.
  - **Description.** A user entered description for the application directory. This is the only field that can be edited.
Reference: File or Synonym Shortcut Menu Options

Synonym shortcut menu options are available if you right-click a synonym in the navigation pane. The following additional options are available:

- **Download.** Downloads the selected file to your Downloads folder. In Windows, the default location of this folder is C:\users\username\Downloads.

- **Properties.** Opens a dialog box that shows the physical location of the file and when it was last modified. For a synonym, this shows the Adapter, Connection, and real table name or file name. It also shows the Description (if any), which can be edited.

  The Properties dialog box has the following buttons:

  - **Save.** Saves the description.
  - **Check File.** Runs a report that shows the number of segments and fields, and the total length of all of the fields.
  - **Synchronize.** For synonyms for relational database tables, refreshes the synonym from the database, retaining descriptions, remarks, and defined fields.

Copying, Cutting, and Deleting Procedures, Synonyms, and Other Files

Procedures, synonyms, and other files on the server can be copied, cut, and deleted from the navigation pane.

Note that multiple objects can be selected when performing these procedures.

- To select multiple adjacent objects, click on the first object you wish to select, and then hold the Shift key while selecting the last object.

- To select multiple non-adjacent objects, click on the first object you wish to select, and then hold the Ctrl key while selecting any additional objects.

Procedure: How to Copy Procedures and Synonyms From the Application Directories Folder

From the Application Directories folder of the navigation pane:

1. Open an application directory.
2. Right-click a procedure or synonym and click *Copy*.
3. Right-click an application directory and click *Paste*.
   - If you select the same directory, you will be prompted for a different name.
4. Enter a name in the Name field and click *Save*.
**Procedure:** How to Cut Procedures, Synonyms, and Other Files From the Application Directories Folder

From the Application Directories folder of the navigation pane:

1. Open an application directory.
2. Right-click a procedure, synonym, or other file and click **Cut**.
3. Right-click an application directory and click **Paste**.
   
   If you select the same directory, you will be prompted for a different name.
4. Enter a name in the Name field and click **Save**.

**Procedure:** How to Delete Procedures, Synonyms, or Other Files From the Applications Directories Folder

From the Application Directories folder of the navigation pane:

1. Open an application directory.
2. Right-click a procedure, synonym, or other file and click **Delete**.
3. Click **OK** to confirm the deletion.

**DataMigrator Workspace**

Flows are built and modified in the DataMigrator workspace. Objects can be dragged into the workspace from the navigation pane and, when in process flow mode, from the process flow toolbar.

For data flows and process flows, the workspace consists of the following tabs:

- **Data Flow tab.** Build data flows to extract and move data from data source objects to data target objects. Arrows connect the objects and support the flow.
- **Process Flow tab.** Build process flows. Arrows connect the objects and support the flow.
- **Text View tab.** Allows you to view the underlying code for a flow.

Other tabs are available when creating different types of data flows:

- **Direct Load Flow tab.** Builds direct load flows. For information, see *Creating a Direct Load Flow* on page 377.
- **DBMS SQL Flow tab.** Builds DBMS SQL flows. For information, see *Creating a DBMS SQL Flow* on page 379.

**Note:** When building different types of data flows, a Process Flow and Text View tab are also available.
For synonyms, the workspace consists of the following tabs:

- **Field View.** Shows a combined view of the segment and field names used. A tree of the segments and columns in the synonym are shown on the left, with the attributes and values of the selected item on the right.

- **Segment View.** Shows the name used to identify the specified segment in a tree format.

- **List View.** Shows the field names used in a grid format.

- **Modeling View.** Shows a graphical representation of the synonym, which can be double clicked to access columns, sample data, and sample data with parent keys.

- **Text View.** Shows the code that describes the segments and fields of a synonym.

- **Access File Text View.** Shows the syntax behind an Access File of a synonym.

For details about synonyms, see *Working With Synonyms* on page 501.

**Data Flow Tab**

The Data Flow tab is where you build data flows to extract and move data from data source objects to data target objects. Arrows connect the objects and support the flow.

The following image is a sample of a simple data flow in the Data Flow tab.

![Data Flow Diagram](image)

**Procedure: How to Validate a Flow**

The flow is checked for errors.

2. If there are no errors, the Valid dialog box opens, confirming that no errors or warnings were found.

   If errors are found, a warning message appears.

3. To exit the window, click OK. For more information about the error, click Details in the Warning window.

Reference: Data Flow Objects in the Workspace

Data flow objects appear as follows in the workspace:

<table>
<thead>
<tr>
<th>Object</th>
<th>Function</th>
</tr>
</thead>
</table>
| ![Data source icon](image) | Data source - `sourcename` (Tn) or `appdir/sourcename` (Tn)  
Indicates a data source for the data flow, where `sourcename` is the name of the synonym and Tn is the table number.  
A key icon in this object indicates that the synonym has a key identified. |
| ![Inner Join icon](image) | Inner Join  
Indicates that two data sources are connected by an inner join and that only rows that appear in both tables used in the Join are extracted. |
| ![Left Outer Join icon](image) | Left Outer Join  
Indicates that two data sources are connected by a left outer join and that all rows are extracted from the left data source, as well as the columns from the right source that match. |
| ![Right Outer Join icon](image) | Right Outer Join  
Indicates that two data sources are connected by a right outer join and that all rows are extracted from the right data source, as well as the columns from the left source that match. |
| ![Full Outer Join icon](image) | Full Outer Join  
Indicates that two data sources are connected by a full outer join and that all rows are extracted from both data sources. |
<table>
<thead>
<tr>
<th>Object</th>
<th>Function</th>
</tr>
</thead>
</table>
| ![Cross Join](image1.png) | Cross Join  
Indicates that two data sources are connected by a cross join, or Cartesian product of two tables. It consists of all possible pairs of rows between the two tables. |
| ![SQL/Select Columns](image2.png) | SQL/Select Columns  
Contains the generated SQL used to extract the data, based on selected columns, filters, and sorts used in the data flow.  
A filter, or sort, icon in this object indicates when a filter or sort has been specified. |
| ![Union](image3.png) | Union  
Indicates that two SQL objects are connected by a union intersect or minus. A different icon appears when Intersect or Minus is selected.  
A pages icon appears when ALL is selected. |
| ![Target - targetname or appdir/targetname](image4.png) | Target - targetname or appdir/targetname  
Indicates a target for the data flow, where targetname is the name of the synonym.  
A key icon indicates if the target table has key columns identified. A red star on the icon indicates a new target table. |

**Reference:**  
**Moving Objects in the Workspace**  
Multiple objects can be selected in the workspace by selecting them while pressing the Shift or Ctrl keys. You can also hold down the Ctrl key while clicking the right mouse button. This will allow you to move or delete the selected items as a group.

**Working With Columns, Variables, and Functions**  
DataMigrator features a number of calculators for creating transformations and filters using columns, variables, and functions. These columns, variables, and functions can be used in the Synonym Editor to create virtual columns.  
These items can be displayed either in a tree or in a list. When displayed in a list, they can be sorted either ascending or descending. You can also search for items.
You can customize the information that is available for columns, as well as how they are displayed.

**Note:** Tree view is the default for the calculators and for Synonym Editor. List view is the default for Join Properties and Column Selection (SQL object).

**Procedure:** How to Change the Display of Columns and Variables in a Calculator

To open a calculator, click the *Insert* icon or the *Compute* icon, or right-click the SQL object and click *WHERE Filter* or *HAVING Filter*.

The Columns/Variables tab defaults to the tree view. To change it:

1. Right-click the *Display Name* heading bar, as shown in the following image.

2. Click *List View*.

The columns and variables are now shown in a list, as shown in the following image.
3. Optionally, you can right-click the Display Name header and select the following options:

- **Sort Ascending or Sort Descending.** This option is only available while in List View.
- **Best Fit.** Resizes the column to fit its data.
- **Customize.** Lets you display additional columns. The available columns are: Format, Description, and Category.

**Procedure:** How to Change the Display of Functions in a Calculator

To open a calculator, click the Insert icon or the Compute icon, or right-click the SQL object and click WHERE Filter.

Calculators open to the Columns/Variables tab using the default tree view.

To change the display of functions:

1. Click the Functions tab.
2. Right-click the Display Name heading bar.
3. Click List View.

   The functions are now shown in a list.

4. Optionally, you can choose Sort Ascending or Sort Descending.
5. Optionally, you can choose Best Fit, which will resize the column to fit its data.

**Procedure:** How to Change the Display of Columns in the Synonym Editor

Double-clicking a synonym opens the Synonym Editor.

1. Right-click the Display Name heading bar.
2. Select List View.

   The columns are now shown in a list.

3. Optionally, you can choose Best Fit, which will resize the column to fit its data.

**Procedure:** How to Change the Information Displayed for Columns

The information displayed for columns is controlled by the Column Management pane of the Options dialog box.

In a calculator or the Synonym Editor:

1. Right-click the Display Name heading bar.
2. Point to Customize and then click a column to add or remove from the shortcut menu.

3. Optionally, click More to add or remove multiple columns or reorder them.

   The Column Management pane of the Options dialog box opens.

   For more information on the options available in the Column Management pane, see User Preferences - Column Management on page 176.

4. Select the columns you want to appear by selecting their check boxes.

5. Optionally, you can change the display order of the columns by moving them up or down in the Selected Columns field.

6. Optionally, you can change the Column Name Display Strategy to control what appears in trees and grids. By default, the Title is displayed, but you can change it back to Name, which is more readable in calculations. You can have columns listed by Name, Alias, Title, or Description. If no Title or Description exists, the display will default to the Alias.

7. Click OK.

   The columns will now appear in the order you specify in calculators or the Synonym Editor.

**Process Flow Tab**

The Process Flow tab is where you build process flows. Arrows connect the objects and support the flow.

The following image is a sample of a simple process flow in the Process Flow tab.

![Process Flow Example](image)

**Reference:** Process Flow Objects in the Workspace

Process flow objects appear as follows in the workspace:
<table>
<thead>
<tr>
<th>Object</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start</td>
<td>Indicates the start of the process flow. Start is always the first object in the process flow.</td>
</tr>
<tr>
<td></td>
<td>The Start object contains the scheduler attributes for the process flow. You can schedule a flow to run once, on a recurring basis, or on</td>
</tr>
<tr>
<td></td>
<td>a specified days of the week or month, on an interval you define.</td>
</tr>
<tr>
<td>Data Flow</td>
<td>Indicates an embedded data flow that contains data source and data target objects, along with business rules that govern what is extracted</td>
</tr>
<tr>
<td></td>
<td>and moved from source to target objects. To open the data flow, right-click the flow and select Open.</td>
</tr>
<tr>
<td></td>
<td><strong>Note:</strong> This icon appears when the data flow is embedded in a process flow.</td>
</tr>
<tr>
<td>Flow</td>
<td>Indicates that a complete data flow or another process flow is being executed from the current process flow. You can nest one flow within</td>
</tr>
<tr>
<td></td>
<td>another. To open the flow, right-click the flow and select Open.</td>
</tr>
<tr>
<td></td>
<td><strong>Note:</strong> This icon appears when a previously saved data flow or process flow is executed from the current process flow.</td>
</tr>
<tr>
<td>Parallel Group</td>
<td>Indicates that a group of flows or procedures will be executed simultaneously. The group object is generally followed by a Wait object.</td>
</tr>
<tr>
<td><strong>Object</strong></td>
<td><strong>Function</strong></td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
</tr>
</tbody>
</table>
| ![Wait Icon](image) | **Wait**  
Indicates that processing of the next object in the flow will wait until execution of all processes in the parallel group have been completed.  
A Wait object generally follows a Parallel Group object, but is not required. |
| ![E-mail Icon](image) | **E-mail**  
Indicates an email message is sent at the designated points in the flow.  
Email can be sent to multiple email addresses at any point in a process flow. Remember that the server must be configured for sending email. |
| ![Stored Procedure Icon](image) | **Stored Procedure**  
Indicates a saved procedure (other than a data or process flow). The stored procedure can be any FOCUS program that is executed during a process flow. |
| ![Connector Icon](image) | **Connector**  
Represented as arrows, connector objects specify the logic that governs what is executed next in a process flow. In some cases, the logic is unconditional. In others, the logic is conditional (set either by defaults or by conditions you define). The color of the arrows indicates these relationships:  
- Black indicates unconditional  
- Green indicates a success condition  
- Red indicates a failure condition  
- Purple indicates a custom-defined condition |
### Object

<table>
<thead>
<tr>
<th><strong>Object</strong></th>
<th><strong>Function</strong></th>
</tr>
</thead>
</table>
| ![Set variables](image) | Set variables  
Set default values for variables used by other objects in the flow. |

### Text View Tab

The Text View tab allows you to view the underlying code for a flow.

**Note:** You can use Ctrl + F or select *Find* from the Edit menu to search the text. After you close the Find dialog box, you can press the F3 key to find the next occurrence and Shift + F3 to search backwards.

![Text View Code](image)

### Text Editor

The Data Management Console has a graphical user interface for editing flows, synonyms, and workspace configuration. It is recommended that this interface should be used since it ensures proper syntax. However, for advanced users, the Data Management Console also has a text editor that can be used to edit the underlying files.

**Note:** Changes made to a flow using the text editor may be lost when the flow is opened normally.

Use caution when adding comments to the first lines (header) of a stored procedure or a flow. Avoid comments starting with `-DM`, since such comments are reserved for DataMigrator internal use. A best practice is to put a space between `-` and the rest of your comment.
**Reference:**  
**Text Editor Shortcut Menu**

Depending on where you right-click in the Text Editor, the following shortcut menu options are available:

**Undo**
- Undoes the last edit operation.

**Redo**
- Undoes the last undo.

**Cut**
- Cuts the selected text to the clipboard.

**Copy**
- Copies the selected text to the clipboard.

**Delete**
- Deletes the selected text.

**Paste**
- Inserts text from the clipboard in the specified location.

**Select All**
- Selects the entire file.

**Find**
- Opens the Find dialog box, which allows you to enter a text string and search parameters.

**Find Next**
- Finds the next occurrence.

**Find Previous**
- Finds the previous occurrence.

**Replace**
- Opens the Replace dialog box, which allows you to enter a find and replace string.

**Go To**
- Opens the Go To dialog box, which allows you to enter a specific line number to go to.
Change Case

This option is available when text is selected. The following options are available:

- Invert Case. Changes all uppercase letters to lowercase and all lowercase to uppercase. For example, iWay will be changed to Iway.
- Upper Case. Changes selected text to all uppercase.
- Lower Case. Changes selected text to lowercase.

Comment Selection

Adds the comment symbol -* in front of the selected lines. This option is only available for procedures.

Uncomment Selection

Removes the comment symbol -* when it appears at the start of the selected lines. This option is only available for procedures.
Aerial View

Aerial View is available in the View main menu option by clicking Aerial View. For large and complex process flows, data flows, and Synonym Editor Modeling Views, the Aerial View allows you to select the portion of the flow or synonym that appears in the workspace. The green area in the Aerial View shows the area that is viewed, as shown in the following image.

Ribbon Tabs

The ribbon represents the activities performed in the DMC. Commands are active or inactive depending on the activity being performed. These include a general toolbar for operations like opening and saving files, a process and data flow toolbar for operations like running and deleting flows, a specialized process flow toolbar for adding objects to the workspace, and a reports toolbar for operations like displaying data in a spreadsheet.
Home Tab

The Home tab consists of the following groups and commands:

**View Group**

- **Status Bar**
  Displays the status bar at the bottom of the DMC.

- **Browser**
  Displays the browser in the DMC.

- **Console Log**
  Displays the Console Log in the DMC.

- **Full Screen**
  Displays the workspace in full screen mode.

- **Explorer**
  Opens a Windows Explorer style window on the right side of the DMC that shows detailed information about the currently selected server object.

- **Back**
  Goes back one directory level. This option is available when in Explorer view.

- **Forward**
  Goes forward one directory level. This option is available when in Explorer view.

- **Up**
  Goes up one directory level. This option is available when in Explorer view.

**Filter Group**

- **Clear**
  Clears any of the filters set for an application directory.

- **Procedures**
  Shows saved flows, procedures, and user functions in the application directory.

- **Synonyms**
  Shows synonyms that describe data.

- **Advanced**
  Allows you to specify what you want to see in the navigation pane. For more information, see *Advanced Filter* on page 203.
Find File Group

Find What

Allows you to specify a term to search for in an application directory. You can specify that the search needs to match a whole word or case, by selecting the appropriate check box below the find field.

Tools Group

Console Log File

Opens the console log file in a text editor, such as Notepad.
Selecting the arrow opens a shortcut menu with the following additional options:

- View Console Log File. Opens the console log file in a text editor, such as Notepad.
- Clear Console Log Window. Deletes all information from the message window.
- Clear Console Log File. Deletes all information from the message window log file.

License Code

Opens the Product License Code window where you can view or change your license code.

Language Selection

Opens the Language Selection window where you can configure a localized version (for selected languages) of the Data Management Console and the language used for text entry and display.

Options

Opens the Options dialog box where you can set user preferences. For more information, see Setting User Preferences on page 172.

Configure Source Control

Opens the Configure Source Control dialog box where you can select a SCC Provider if you have one installed on your PC.

Windows Group

Windows

Has the following options from the shortcut menu:

- Close. Closes the active window in the workspace.
- Close All. Closes all the open windows in the workspace.

- Close All But This. Closes all open windows in the workspace except for the one currently open.

**MDI Tabs**

Has the following options from the shortcut menu:

- **New Horizontal Tab Group.** Creates a new horizontal tab group and opens the last object there.

- **New Vertical Tab Group.** Creates a new vertical tab group and opens the last object there.

- **Move to Next/Previous Tab Group.** Changes the focus to the next or previous tab group. This option is available only if multiple tab groups are open.

**Flow Tab**

The Flow tab consists of the following groups that apply to both a data and process flow. Some of these commands are only available when working with a process flow.

**Undo/Redo Group**

- **Undo**
  
  Undoes the last operation.

- **Redo**
  
  Undoes the last undo.

**Clipboard Group**

- **Delete**
  
  Deletes the selected object.

- **Cut**
  
  Cuts the selected item.

- **Copy**
  
  Copies the selected item.

- **Paste**
  
  Pastes the cut or copied item.
**Editing Group**

**Find**
Opens the Find dialog box, which allows you to enter a text string and search parameters. This option is available when in Text or Edit as Text view.

**Find Next**
Finds the next occurrence. This option is available when in Text or Edit as Text view.

**Find Previous**
Finds the previous occurrence. This option is available when in Text or Edit as Text view.

**Select All**
Selects all of the text in the file. This option is available when in Text or Edit as Text view.

**Replace**
Opens the Replace dialog box that lets you enter a text string to find and replace it with. This option is only available when in Edit as Text view.

**Go To**
Opens the Go To Line dialog box, which allows you to enter a specific line number to view. This option is available when in Text or Edit as Text view.

**Invert Case**
Changes all uppercase letters to lowercase and all lowercase to uppercase. For example, iWay will be changed to Iway.

**Upper Case**
Changes selected text to all uppercase.

**Lower Case**
Changes selected text to all lowercase.

**Comment Selection**
Adds the comment symbol -* in front of the selected lines. This option is only available for procedures.

**Uncomment Selection**
Removes the comment symbol -* when it appears at the start of the selected lines. This option is only available for procedures.
View Group

Data Flow
Switches to the Data Flow view in the workspace.

Process Flow
Switches to the Process Flow view in the workspace.

Text View
Switches to the Text view in the workspace.

Aerial View
Opens the Aerial View window.

Find Flow Object Group

Find What
Allows you to specify a term to search for in the workspace. You can specify that the search needs to match a whole word or case, by selecting the appropriate check box below the Find What field.

Insert Group (Data Flow Only)
If you are working with a data flow, the Insert group consists of the following commands.

Source
Creates a source in the workspace.

Join
Joins two objects in the workspace.

Select
Creates a Select object in the workspace.

Union
Creates a Union object in the workspace.

Existing Target
Adds an existing target in the workspace.

New Target
Creates a new target in the workspace.

Insert Group (Process Flow Only)
If you are working with a process flow, the Insert group consists of the following commands.
Procedure
Creates a Stored Procedure object in the workspace.

Flow
Creates a Flow object in the workspace.

E-Mail
Creates an Email object in the workspace.

Parallel Group
Creates a Parallel Group object in the workspace.

Wait
Creates a Wait object in the workspace.

Set Variables
Sets variables for the process flow.

Run Group
Run
Launches a drop-down menu for Submit, Run, Run Advanced, Impact Analysis, Dependencies Analysis, and Data Lineage.

Reports Group
Impact Analysis
Runs the impact analysis report for the flow that is currently open.

View Last Log
Opens the last log for the flow that is currently open.

View Flow Report
Runs the flow report for the flow that is currently open.

Tools Group
Arrange Items
Rearranges the objects in the workspace.

Validate
Validates the flow.
Properties
Displays the property pane for the flow.

User Functions Tab
The User Functions tab becomes active when you edit functions. It contains the following additional groups:

View Group
- **Tree View**
  Shows the selected function in a tree view.
- **List View**
  Shows the selected function in a list view.
- **Text View**
  Shows the selected function in a text view.

Insert Group
- **Function**
  Creates a new function for the function.
- **Parameter**
  Creates a new parameter for the function.
- **Define**
  Creates a new define for the function.

Run
- **Test Function**
  Opens the Test Function dialog box, where you can enter values to test the function.

Tools Group
- **Properties**
  Opens the Properties pane for whatever is selected in the workspace.
Reports Tab

The Report tab becomes active when you run a report or view the Server Session Log. The Report tab has the following additional groups.

Print Group

Print

Opens the Print dialog box to print the report.

Print Preview

Displays the preview of the item to be printed.

Run Group

Refresh

Reruns the report.

Next Result

Displays the next report in the output.

Note: This option is only available for output with multiple pages.

Automatic Refresh

Automatically reruns the report at the refresh interval specified.

Refresh Interval (in seconds)

Allows you to specify how often a report reruns.

Tools Group

Export to Excel

Exports the report as an Excel spreadsheet.

Note: This option is available for Alpha and Original output only.

Copy as Text

Copies the selected report text to the clipboard.

Filter

Allows you to specify what information is shown in the report. The options available are FOCUS, SQL/Mdx, Transform, Errors, Messages, Timestamp, and No Delay. This option is available when viewing a Server Session Log.
Clear Messages

Deletes all messages from the log and controls what types of messages appear in the server session log. This option is available when viewing a Server Session Log.

Choose Columns

Opens the Scheduler Agent page where you can select the columns that appear on the scheduler agents page. This option is available when viewing the Scheduler Agent.

Choose States

Opens the Select the states that will appear on the Scheduler Agents dialog box where you can select what states will appear on the Scheduler Agents page. This option is available when viewing the Scheduler Agent.

Activate Selected

Allows you to select multiple flows or procedures and set the schedule status as active in a single operation. When using this command, it is recommended that you use Choose columns and select the Status column to display. This option is available when viewing the Scheduler Agent.

Deactivate Selected

Allows you to select multiple flows or procedures and set the schedule status as inactive in a single operation. When using this command, it is recommended that you use Choose columns and select the Status column to display. This option is available when viewing the Scheduler Agent.

Metadata Tab

The Metadata tab becomes active when you are working with a synonym. For more information, see Working With Synonyms on page 501.

Text Editor Tab

The Text Editor tab becomes active when you are using the Text Editor. It contains the following additional group:

Run Group

Run

Launches a drop-down menu for Run, Run with Options, and Impact Analysis.

Console Log

The DataMigrator Console Log shows results from the DMC communications with the server.
The Console Log displays:

- Log messages when you connect or disconnect from the server, run reports, or save flows.
- Server messages when you execute flows or stored procedures.

To toggle the Console Log, select or deselect Console Log from the View menu or press Alt+9.

If you right-click the console log, the following options are available:

**View Console Log File**

Displays the log file in Notepad.

**Clear Console Log Window**

Deletes all information from the message window.

**Clear Console Log File**

Deletes all information from the message window log file.

**Copy Selection**

Copies items highlighted in the Message window to the clipboard. Note that you can select multiple lines by pressing the Shift or Ctrl key while selecting each line.
Server Session Log

The Server Session Log shows the actual commands that are sent from the DMC to the server, which may be useful for problem determination. It is available by right-clicking a server in the navigation pane, pointing to Other Options, and clicking Session Log from the shortcut menu. The log will then open in the workspace, as shown in the following image.

![Server Session Log Image]

For more information on the commands available when viewing this report, see Reports Tab on page 158.

Setting Flow Properties

Flows properties allow you to view or set properties that apply to the entire data flow. These include:

- **General properties.** Includes who created the flow, the last time it was updated, a description, and comments.

- **Execution properties.** Includes stop processing options.

- **Restart properties.** Includes the number of restart attempts and their start points. A restart attempt occurs if the flow execution fails. Restart attempts can be made from the first record or from the last commit point.

- **Record Logging.** Shows actions to be taken when processing records with errors.

Specify flow properties in the Properties pane by clicking Properties in the Tools group of the Flow tab.
If you right-click anywhere in the Properties pane, a shortcut menu opens, with the following options available:

- **Categorized.** Displays all of the available properties for the selected flow by category.
- **Alphabetic.** Displays all of the available properties for the selected flow in alphabetic order.
- **Find.** Allows you to search for text in the Attributes column.

**Procedure:** How to Set the Flow General Properties

   The Properties pane opens, as shown in the following image.

   ![Properties Pane](image)

   The General attribute section of the Properties pane has the following fields and options:

   **Description**
   - Is a description of the flow. You can enter up to 150 characters. Descriptions appear in the navigation pane next to the flow name.

   **Comment**
   - Are comments regarding the flow.

   This section also provides information on the flow name, the user who created the flow, and the date it was last modified.

2. Make your desired changes and click *Save* in the Quick Access Toolbar.

**Procedure:** How to Set the Flow Execution Properties

   The Properties pane opens.
2. Expand the *Execution* attribute, as shown in the following image.

![Execution attribute in Properties pane](image)

The *Execution* attribute section of the Properties pane has the following fields and options:

**Optimize Load**
- Allows faster loading when the source and target are in the same relational database. This option can be used with databases that use the MERGE command, as well as Insert only operations.

**Continue processing when a partial answer set is returned**
- When selected, a flow will continue to run, even if an error is generated.

**Stop after DBMS errors**
- Causes the flow to fail with a non-zero return code if the number of errors specified is exceeded. Applies to relational targets when Insert/Update or Insert records from memory is used. Any rows written to the RDBMS are committed to the RDBMS. The default value is 10000000, unless this is overridden in the Run Options found in the Tools group of the Home tab.

A relational database may return an error for a constraint violation, such as a Unique Index constraint, a Not Null constraint, or a user-specified constraint.

If such errors are not expected, you may want to fix their cause without waiting for the entire flow to be processed. Using a low value, such as 10, can provide enough information to determine the cause of the problem, without running the entire job.

If such errors are expected, and you want to continue processing anyway, use a high value, one greater than the number of records in the source, or do not select the *Stop processing after* check box.

**Note:**
- When using the Load Type setting of Bulkload utility using a disk file, this setting is not used.
This setting is not available if the Optimize Load option is enabled.

Stop if 0 rows selected

Allows you to stop processing if no data is returned. When this option is selected, and no data is returned, code 18708 occurs.

3. Make your desired changes and click Save in the Quick Access Toolbar.

Procedure: How to Set the Flow Restart Properties


   The Properties pane opens.

2. Expand the Restart attribute, as shown in the following image.

   ![Properties pane with Restart attribute expanded]

   The Restart attribute section of the Properties pane has the following fields and options:

   Number of Attempts

   Is the number of times a flow will be restarted if the execution fails.

   **Note:** The amount of time between restart attempts can be set for the server. The default value is 60 seconds. For more information, see Setting Up DataMigrator on page 41.

   Restart From

   Is the point from which the next attempt will be made. The options are Beginning or Last Commit.

   **Note:** Last Commit cannot be used for flat file, delimited file, or XML targets. Last Commit can only be used for existing targets, and cannot be used with Prior to Load Options of Delete all rows from table, Truncate table, or Drop table.

3. Make your desired changes and click Save in the Quick Access Toolbar.
Procedure: **How to Set the Flow Email Notification Properties**

**Note:** Your DataMigrator Server must be configured for email notification to use this option.

1. Right-click the Start object in a process flow and click *Properties*.
   
   The Properties pane opens.

2. The E-mail attributes open, as shown in the following image.

![Properties Pane](image)

The Log and output destinations section has the *Send HTML output as inline message*. When selected, if any stored procedures called by the flow generate an output, the message is included as message text, instead of as an attachment.
Note: This option for a flow schedule works in conjunction with the Merge formatted output into log option that is available for a Scheduler Special Service. The following considerations should be taken into account:

- When the Send HTML output as inline message and Merge formatted output into log options are selected, the submitted flow emails the attached ETLLOG file and the HTML output is embedded into the body of the email.

- When the Send HTML output as inline message and Merge formatted output into log options are not selected, the submitted flow emails the attached ETLLOG file and the HTML output is embedded into the attached ETLLOG file

E-Mail On Start, E-Mail On Completion, and E-Mail On Failure all have the following fields and options:

Enable

Select this to enable when an email should be sent.

Mail To

The email addresses to which the messages will be sent. Place each email address (or user ID configured on the server with an associated email address) on a new line with no separators.

Advanced

Select this option to get the following options:

- Mail to addresses defined in procedure. You can send the messages to email addresses defined in the procedure by either typing in the location of the procedure, or clicking the ellipses and manually selecting it from the Select Procedure dialog box.

- Importance. Select the importance for the email from the drop-down menu. Choices are Low, Normal, or High.

- Subject. The subject line of the message.

- E-Mail Message. The body of the message.

Note: Your DataMigrator Server must be configured for email notification in order to use this option. See How to Configure the Server for Email Notification on page 91.

3. Make your desired changes and click Save in the Quick Access Toolbar.
Procedure: How to Set the Flow Record Logging Properties

Note: These settings are not available if the Optimize Load option is enabled.


   The Properties pane opens.

2. Expand the Record Logging attribute, as shown in the following image.

   ![Record Logging Attribute Section](image)

   The Record Logging attribute section of the Properties pane has the following fields and options:

   **DBMS Errors**
   
   Logs transactions rejected by an RDBMS due to a constraint violation or data type error. For flat file targets, format errors are also recorded.

   **Duplicate Rejections**
   
   Logs transactions rejected due to a duplicate key.

   **Invalid Data**
   
   Logs transactions rejected for not meeting a VALIDATE condition

   **No Match Rejections**
   
   Logs transactions rejected when a required key is not found.

   **Format Errors**
   
   Logs transactions rejected due to a format error. Format errors only apply to direct load flows.
3. Make your desired changes and click Save in the Quick Access Toolbar.

### Viewing Properties

During the course of your DM design work, you may need to view and edit the properties of your server connection. For example, you may wish to:

- Update expired user IDs and passwords.
- Change port numbers.

You can perform these server-connection related tasks from the Server Node Configuration dialog box within the DMC.

In addition, you can view, but not edit, the properties for the Application Directories folder, specific application directories within this folder, and procedures within an application directory.

**Procedure: How to View Server Information**

Server Information displays information about the release of the server to which you are connecting. It is used primarily for customer support, if you ever need to report a problem.

From the navigation pane, right-click the DataMigrator Server, point to Help, and then click Version.

**Procedure: How to View the Server Properties**

The Server Node Configuration dialog box shows the parameters you set during installation or when configuring an additional server node. See *How to Install the Data Management Console* on page 46 and *How to Add a Node to Identify the Server to the Data Management Console* on page 50.

From the navigation pane:

1. Right-click the server and click Properties.
   - The Server Node Configuration dialog box opens.
2. Make any desired changes.
3. Optionally, if you are using the EDA API to connect an application like iWay Service Manager to the server, expand the Advanced Parameters section to set connection and wait limits.

4. Click Save.

Reference: **Server Node Configuration Dialog Box**

To access the Server Node Configuration dialog box, right-click the server and click *Properties*.

![Server Node Configuration Dialog Box](image)

**Note:** The general and security settings in the Server Node Configuration dialog box are used by the DMC.
The Server Node Configuration dialog box contains the following fields and options:

**Host**
- **Host**
  
  Is the machine name or IP address where the DataMigrator Server is running.

**HTTP Listener Port**
- The DMC connects to the server using the HTTP Listener Port. When you install a DataMigrator Server, the HTTP Listener is assigned to port 8117 by default. When you add new server nodes using the DMC, the default port is 8101.

**Node Name**
- Is the DataMigrator Server name.

**Node Description**
- Provides additional information about the node that will be visible in the navigation pane.

**Security**

**Security Type**
- Is the security setting.

The following options are available:

- **Explicit.** With this setting, if security is required, you must enter it either at connection time or in the Security credentials fields to the right. This is the default value.

- **NTLM.** Use when connecting to a Windows server. The Windows user ID and password to the server are passed to the server when a connection is made.
  
  **Note:** Only when OPSYS security is used.

- **Kerberos.** Use when connecting to a Linux server. The user ID and password to the server are passed to the server when a connection is made.
  
  **Note:** Only when OPSYS security is used.
**Note:** Settings on the Server Node Configuration Advanced section are used by applications connecting to the server using the EDA API, including Resource Analyzer and Resource Governor.

**Advanced Parameters**

**TCP Listener Port**

Allows you to set your own TCP port, rather than the default, when connecting an application to the server using the EDA API. Applications like Resource Analyzer and Resource Governor use the EDA API to connect to the server through this port. When you install a DataMigrator Server, the TCP Listener is assigned to port 8116 by default. When you add new server nodes, the TCP Listener Port defaults to one less than the HTTP Listener Port.

**Use Secure Sockets Layer (SSL)**

Sets a Secure Sockets Layer (SSL), which is a commonly-used protocol for managing the security of a message transmission over HTTP. To use the HTTP Listener Secure Sockets Layer, you must first configure the HTTP Listener for SSL.

**Service Name**

Is the name of the service to which the user ID is connecting. If it is left blank, the user ID will connect to the default service.

**Connect Limit**

The maximum time that the client will wait for a TCP connection response from the server.

**Maximum Wait Limit**

The time that the client will wait for a response from the server in seconds.

**Compression**

Activates data compression in data transfer between the client and the server.

**Encryption**

Defines the encryption method used in data transfer between the client and the server. Encryption is not supported in PIPE protocol. Only DES is supported in HTTP protocol.

**Application Directory Properties**

View the properties of an application directory using any of the following methods:

- In the navigation pane, right-click the *Name* heading bar, point to *Customize*, and then click *More*.
The Column Management window opens. This is where you can customize what property columns are displayed in the navigation pane.

- In the navigation pane, right-click *Application Directories* and click *Explorer*.

The application directories that are configured are now displayed in the workspace. Right-click the heading bar in the workspace, point to *Customize*, and then click *More*.

The Column Management window opens. This is where you can customize what property columns are displayed in the navigation pane.

For more information on the Column Management window, see *User Preferences - Column Management* on page 176.

**Flow or Procedure Properties**

View the properties of a flow or procedure using any of the following methods:

- In the navigation pane, right-click the *Name* heading bar, point to *Customize*, and then click *More*.

The Column Management window opens. This is where you can customize what property columns are displayed in the navigation pane.

Expand the application directory to see the properties for the contents of that directory.

- In the navigation pane, right-click an application directory and click *Explorer*.

The contents of the directory are now displayed in the workspace. Right-click the heading bar in the workspace, point to *Customize*, and then click *More*.

The Column Management window opens. This is where you can customize what property columns are displayed in the navigation pane.

For more information on the Column Management window, see *User Preferences - Column Management* on page 176.

**Setting User Preferences**

User preferences allow you to set up your DMC environment. On the Home tab, in the Tools group, click *Options* to access these settings.

The Options dialog box opens with General settings in the right pane.

**Procedure:** How to Set User Preferences

1. On the Home tab, in the Tools group, click *Options*.

The Options window opens and the General settings page appears in the right pane.
You can set preferences for the following:

- **General.** For details, see *User Preferences - General* on page 173.
- **Format.** For details, see *User Preferences - Format* on page 175.
- **Column Management.** For details, see *User Preferences - Column Management* on page 176.
- **Traces.** For details, see *User Preferences - Traces* on page 179.
- **Data Flow Designer.** For details, see *User Preferences - Data Flow Designer* on page 180.
- **Target Defaults.** For details, see *User Preferences - Target Defaults* on page 182.
- **Process Flow Designer.** For details, see *User Preferences - Process Flow Designer* on page 183.
- **Synonym Editor.** For details, see *User Preferences - Synonym Editor* on page 185.
- **Run Options.** For details, see *User Preferences - Run Options* on page 187.

2. Select the check boxes for the settings that you want to apply.
3. Click OK.

**Reference:**  *User Preferences - General*

General user preferences include using background images, automatically switching and synchronizing trees, determining the location of messages, and controlling scrolling speed.
On the Home tab, in the Tools group, click Options to access these settings.

The General settings dialog box has two sections: General and Others.

The General section has the following fields and options:

**Use background images**
- Set a blue background image for the workspace.

**Automatically switch filters**
- Switches filters automatically depending on what flow view is open. Clicking the Data Flow tab will switch on the Synonyms filter, while clicking the Process Flow tab will switch on the Procedures filter.

**Start LOOPBACK server on startup**
- This option is enabled if you start the DMC from a Windows installation. The local LOOPBACK server will start automatically, if it is not already running, when you start the DMC. When you close the DMC, the LOOPBACK server is stopped. This option is not enabled if you start the DMC from a client installation.
Show tabbed view selector at top of each frame

Moves workspace tabs, like Data Flow, Process Flow, and Text View, to the top. If this option is not selected, by default, the workspace tabs appear at the bottom of the frame.

Auto Save for existing files (in minutes)

You can select the number of elapsed minutes before open files are automatically saved. By default this feature is not enabled.

The Others section has the following options:

Scrolling speed

Determines the speed at which a flow scrolls when you drag an object outside the viewable window. It can be set in relative units of 1 to 10. The default is 1.

Recently used file list

You can select the number of files to display in the recently used file list from 0 to 16. The default is 4.

Enable Developer Studio Integration

Enables the use of Synonym Editor in Developer Studio.

Display WebFOCUS Environments*

If not selected, the WebFOCUS Environments option is disabled and does not appear in the navigation pane. This option requires restarting the DMC.

Include images when exporting data to Excel

Enables data to be exported with images included.

Clear User Preferences*

Clears all user set preferences and returns them to default values. This option requires restarting the DMC.

Save GUI State

Used for diagnostics, this option saves the current state. This file is only needed if requested by customer support.

Ribbon*

Select this option to enable the Ribbon. Deselect it to use the classic toolbar. This option requires restarting the DMC.

Reference: User Preferences - Format

Format user preferences include settings, colors, and fonts for the Data Console interface.
On the Home tab, in the Tools group, click Options to access these settings.

The Fonts Category section has the following fields and options:

**Category list box**

Allows you to set the font for text in the process flow workspace, reports, text views, and log views.

**Font**

Launches a dialog box for changing font settings.

**Reset fonts to defaults**

Restores the default fonts.

**Display full component names**

Allows you to display the full name of a component.

**Reference:** User Preferences - Column Management

Column Management user preferences enable you to select which columns to display on the grids for transformations, column selection, joins, sorts, sources, and targets. You can also set the column display order. The settings apply to all column-related dialog boxes.
On the Home tab, in the Tools group, click *Options* to access these settings.

The Column Management settings include the following fields and options:

**Customize column display**

Lists columns that can be added to the grid using the check boxes.

**Column Name Display Strategy**

Controls the information that appears in trees and grids. The available options are Name, Title, and Description. By default, the Title is displayed, but you can change it back to Name, which is more readable for calculations. If no Title or Description exists, the display will default to the Alias. When the strategy is something other than the Name, this is indicated in the header row. For example, Display Name (Title).
Function display strategy
Controls how functions are identified on the functions tab in the calculators. By default, the function syntax is displayed. It can be changed to a short description, which displays a brief description of what the function does instead.

Use segment to qualify field reference
Enables you to fully qualify field names in the JOIN criteria of a cluster synonym. The available options are Always and For duplicate fields. The default value is For duplicate fields, which allows you to maintain compatibility with earlier releases.

The default settings are shown. Not all columns are on every grid. The following are examples of columns that can be added:

Expression
Is the expression for the column.

Type
Is the type of object in an application directory.

Length
Is the column length.

Scale
Is the maximum number of digits to the right of the decimal.

Alias
Assigns an alternative name for a column, or the real column name for a DBMS synonym.

Title
Supplies a title to replace the column name that is normally used in reports.

SCD Type
Is used for processing slowly changing dimensions.

Field Type
Indicates that a field is an index (I) and/or is read-only.

Table
Is the synonym that contains the column.

Format
Is the type and length of column data as stored.
**Description**

Is a description or comment about the column.

**Nulls**

Indicates whether the column can contain null data.

**Note:** You can change the display order of the columns by moving them up or down using the arrows. You can also change the *Column Name Display Strategy* to control what appears in trees and grids.

**Reference:** User Preferences - Traces

On the Home tab, in the Tools group, click *Options* to access these settings.

![User Preferences - Traces dialog box](image)

The Traces settings dialog box has the following options:

**Enable**

Select this option to perform tracing of the application to debug problems. You can choose the components activated in the trace files tscom3.trc and connection.trc. These trace files reflect the operations between the client and the server. You can view these trace files from the DMC.

**View Trace Files**

Select which file to view from the drop-down menu. The options are tscom3.trc and connection.trc.
View

Click this button to view the trace file.

Choose the trace components to activate

Allows you to specify which components will be traced.

Reference: User Preferences - Data Flow Designer

Data Flow Designer user preferences include using application directories, automatically selecting and mapping all columns, adding Join objects and conditions, and performing validation.

On the Home tab, in the Tools group, click Options to access these settings.

The Data Flow Designer settings dialog box has two sections: On Add Source and On Add Target.

Use application directory name with flow components

Includes the application directory name for objects in data flows.
Show Notes

Allows you to add comments to data flow objects, directly from the workspace.

Undo/Redo Limit

Specifies the maximum number of undo operations allowed in the Data Flow Designer.

The On Add Source section has the following fields and options:

Automatically select all columns

Selects all columns in a source.

Add Join Object if needed

When two sources are dragged onto the workspace, this option connects the sources to each other and to the SQL Object using Join objects. This is the default.

Automatically add join conditions

Creates an inner join if two data sources have a column with a primary/foreign key relationship or with the same name and format. The column must be a key column in one of the sources. This is the default. This option also creates source/lookup column joins when using the DB_LOOKUP function.

Show adapter type

This option appears for Source and Target. When selected, the adapter type (for example, Flat File or MS SQL Server) appears under the synonym name.

Note: On Add Source settings affect new sources that are added to a flow.

The On Add Target section has the following fields and options:

Automatically map all columns

Automatically maps all source columns to target columns of the same name.

Show adapter type

This option appears for Source and Target. When selected, the adapter type (for example, Flat File or MS SQL Server) appears under the synonym name.

Data Flow Line Colors

Is a drop-down menu of available colors for data flow lines.

Reset colors

Restores the default colors.


**Reference:**  **User Preferences - Target Defaults**

Target Defaults user preferences allow you to set default Prior to Load Values for each category of targets. You can also set commit levels for targets. All targets in the current and subsequent flows will use the commit level default set here.

On the Home tab, in the Tools group, click *Options* to access these settings.

![Target Defaults settings dialog box](image)

The Target Defaults settings dialog box has two sections: Prior to Load Values and Flow Level.

The Prior to Load Values section has the following fields and options:

**Relational**

- **No changes**
  - Does not delete the records already in a data target.

- **Delete all rows from table**
  - Deletes all rows and then creates a database log.

- **Truncate table**
  - Deletes all rows from the table but does not generate a database log. This is a faster option than Delete all rows from table because no database log is generated. If the target database does not support truncate, a delete is performed.

**Focus**

- **No changes**
  - Does not delete the data target.
Drop Table

Deletes and recreates the data target.

Flat File

No changes

Does not delete the data target.

Delete File

Deletes and recreates the data target.

The Flow Level section has the following fields and options:

Commit every row(s)

Specifies the number of rows to process before issuing a commit to a relational database target or a CHECK to a FOCUS target.

A value of zero (0) means that a COMMIT is only issued when the flow completes.

Note: This setting will be applied to all targets in a flow.

When several flows are loading multiple tables, it may be desirable to not issue a COMMIT until all the flows have completed successfully. In this case, a special value of NONE can be used to mean that DataMigrator should not issue any COMMIT commands. This means that it is the responsibility of the user to issue a COMMIT (or ROLLBACK) when appropriate. This can be done by creating a process flow that runs any required data flows and checks the return code for each data flow. When all flows have completed successfully, the process flow should then call a stored procedure that issues the COMMIT command, or if any data flow has failed, a ROLLBACK command.

This option should be used with caution. If no COMMIT or ROLLBACK is issued, the target tables could be locked.

Reference: User Preferences - Process Flow Designer

Process Flow Designer user preferences include using application directories.
On the Home tab, in the Tools group, click Options to access these settings.

The Process Flow Designer settings dialog box has the following options:

**Use application directory name with flow components**

Includes the application directory name for objects in process flows.

**Add flows to run as stored procedures**

When flows are added to a process flow, the Run as stored procedure check box is automatically selected.

**Show Notes**

Allows you to add comments to process flow objects, directly from the workspace.

**Process Flow Line Colors list box**

Allows you to set colors for connector lines in process flows.

- **OK Default Condition.** Changes the color of the OK connector in process flows.
- **Failed Default.** Changes the color of the Failed connector in process flows.
- **Unconditional.** Changes the color of connectors in data flows, and unconditional connectors in process flows.
- **Custom Line Color.** Changes the color of the custom connector in process flows.
**Reset colors**

Restores the default colors.

**Reference:** User Preferences - Synonym Editor

Synonym Editor user preferences include using the application name with the synonym, supporting the extended option, and showing parent segments in the Join Editor.

On the Home tab, in the Tools group, click Options to access these settings.

The Synonym Editor settings dialog box has the following fields and options:

**Use application directory name with synonym**

If selected, an application directory name will be used when selecting a synonym name for both referencing an existing synonym, and a transformation with db_lookup.

**Undo/Redo Limit**

Specifies the maximum number of undo operations allowed in the Synonym Editor.
Support extended options
If selected, the Sort Object and Style buttons appear in the Insert menu on the Metadata tab, in the Insert group.

Enable insert copy of existing synonym
If selected, this option will be available in the shortcut menu when a synonym or segment name is selected.

Insert child segment with snowflakes
If selected, when you insert a child segment (dimension), any of its child segments are inserted automatically.

Automatically detect new segment relations
When you create a new synonym in the Modeling View of Synonym Editor with the Automatically detect new segment relations option selected, and select or drag these three tables into the work area (modeling view), the relationships (joins) will automatically be created for you based on the foreign key information in the access file.

Automatically arrange segments/folders in Modeling View
If you select or drag multiple, or individual, tables while in Modeling View, they will be automatically arranged for you.

Show parent segments in Join Editor
If selected, this option controls whether or not columns in parent segments are displayed in the Join Editor for a cluster join, or for a synonym that references, or includes, other synonyms.

Automatically detect entry segment

Default Join Type
Allows you to select the default Join type. Options are Join All or Join Unique.

Auto Calculate Missing
This option controls the behavior when you insert a DEFINE field into a synonym where at least one operand in the expression has a MISSING-ON attribute. The available options are:

- **PROMPT.** Asks *Do you want to change MISSING to ON ALL* whenever a new DEFINE is added. This is the default value.

- **ON.** Expressions are automatically adjusted so that the result of the expression is MISSING if any operand is missing.
Expressions are not adjusted. The result of the expression is MISSING only if all of the operands are missing.

**Modeling View Line Colors list box**

- **Default.** Changes the color lines in the Modeling View.
- **Highlighted.** Changes the color of the highlighted lines in the Modeling View.
- **No keys.** Changes the color of lines with no keys in the Modeling View.

**Reset colors**

Restores the default colors.

**Reference:** User Preferences - Run Options

Run Options user preferences include determining the number of rows to retrieve when sampling data and keeping your server connection alive.

On the Home tab, in the Tools group, click Options to access these settings.

![Run Options settings dialog box]

The Run Options settings dialog box has the following fields and options:

**Maximum number of rows for test reports**

Sets the number of rows retrieved to produce sample data when testing transformations or SQL. The default is 50.
Maximum number of columns for test reports

Sets the number of columns retrieved to produce sample data when testing transformations or SQL. The default is to retrieve all columns (with a highest value setting of 999999).

Test reports default format

Sets the format of reports for retrieving sample data when testing transformations, or SQL. The default is Default. There are seven report formats available:

- **Default.** Formats numeric and date columns based on edit options in the synonym.
- **HTML.** Produces the report in HTML format.
- **HTML - plain text.** Produces the report in plain text format.
- **Active Report.** Produces an HTML active report designed for offline analysis.
- **PDF.** Produces the report in PDF format.
- **Unformatted.** Does not apply formatting to numeric and date columns.

Omit missing (null) values in Data Profiling charts

Enable this option to generate Data Profiling bar and pie charts that omit missing values.

Stop processing if 0 rows selected

Allows you to stop processing if no data is returned. This setting is the default for new flows. You can override this option for an individual flow by resetting the execution properties of the flow.

Stop after DBMS errors

Allows you to stop processing the number of DBMS errors received if the number of errors specified is exceeded. This setting is the default for new flows. You can override this option for an individual flow by resetting the flow execution properties for that flow.

User Privileges

User privileges associated with a directory or other server objects, such as procedures or synonyms, can be managed from the DMC.

Note: These privileges can only be controlled by a user that is logged on to a secure server with a user ID with Server Administrator privileges. If you are logged on with an ID that is not a Server Administrator, you will only have access to view your own privileges. You cannot change them or view the privileges of another user.
Procedure: **How to Access Privileges**

There are two ways to access the Manage Privileges dialog box:

- Right-click an application directory and click **Privileges**.
- or
- Right-click an object (procedure, synonym, flow, and so on) and click **Privileges**.

Reference: **Manage Privileges Dialog Box**

To access the Manage Privileges dialog box, right-click an application directory, or an object, such as a procedure or synonym, and click **Privileges**.

The Manage Privileges dialog box has the following columns:

**Subject**

This column shows roles, users, and groups.

**Roles**

There are five roles that can be used for managing privileges. Additional roles can be added by the server administrator. The default roles are:

- **Server Administrator**. Has full control of the workspace, adapters, synonyms, and procedures.
• **Application Administrator.** Can create and run synonyms and procedures.

• **Server Operator.** Can start and stop the server, as well as monitor and kill agents.

• **Basic User.** Can execute procedures.

• **None.** Has no privileges.

Each role can have users or groups of users associated with it. These are shown indented under each role name.

**Privileges**

The privileges that are associated with each directory are:

• **Read.** The user can read the directory or file.

• **Write.** The user can write, create, edit, or delete the directory or file.

• **Execute.** The user can run procedures or access metadata.

• **List.** Lists files.

**Inherit Subject**

Shows if the privileges are inherited from those of a role, by showing that role.

**Inherit Directory**

Shows if the privileges are inherited from those of a directory, by showing that directory.

**Description**

An optional description for the directory or file.

**License Code**

A valid license code is required to run DataMigrator. The license code is available from the Tools menu. The license code used to install the product is displayed.

**Procedure:** How to View or Change Your License Code

1. On the Home tab, in the Tools group, click *License Code*.

   The Product License Code window opens.

2. Enter changes in the License Code field.

3. Click *OK*.

**Note:** If you purchase DataMigrator after installing the client and/or server, you will be provided with a license code to activate the DataMigrator features.
Configuring National Language Support

You can configure National Language Support (NLS) directly from the DMC. National Language Support enables the connector or server to translate both double-byte and single-byte National Character sets.

Procedure: How to Configure NLS


   The Language Selection dialog box opens, as shown in the following image.

   ![Language Selection dialog box]

   The Language Selection dialog box has the following fields and options:

   **Select Language**

   Controls the default value of the code page and the language of server error messages if they are available for translation. The default value is AMENGLISH (American English) on all platforms. Localized versions are noted by an asterisk (*).

   2. Select a Language from the drop-down menu. (Note that localized versions are noted with an asterisk (*).)

   3. Click OK.

Ending a Data Management Console Session

When you are finished working in DataMigrator, you should disconnect from the server and close the DMC.

Procedure: How to End Your Data Management Console Session

1. Save all changes.

2. Right-click the server name and select Disconnect.

3. Close the DMC by clicking the control button in the upper-right corner.

   or
Selecting *Exit* from the Application menu.

or

Double-clicking the Application menu icon (DMC).

**Note:** When you close the DMC, you will be prompted to save changes if you did not already save them.
Designing a Data Flow

A data flow contains the parameters for extracting information from data sources and loading it into data targets. You create data flows in the Data Management Console using the data flow workspace.

Typically, you build a data flow from a set of objects that identify data sources, data targets, data columns to include, and business rules governing how data is extracted, transformed, and loaded to form a data warehouse, data mart, or any other type of target.

You can join multiple data sources, filter and sort the extracted source data, and specify column mappings from the source to the target.

You can also create specialized data flows, such as Direct Load and DBMS SQL flows.

This chapter describes how to design and create the different kinds of data flows.

In this chapter:

- The Data Flow Interface
- Simple Data Flow at a Glance
- Complex Data Flow at a Glance
- Creating a Data Flow
- Specifying Data Sources
- Creating a Source Transformation in a Data Source Object
- Joining Data Sources
- Adding a SELECT to a Data Flow
- Adding a UNION to a Data Flow
- Selecting Columns
- Testing the SQL Statement
- Specifying Filters in a Data Flow
- Specifying Sorts
- Selecting a Data Target
- Specifying Data Target Options
- Mapping Transformations
- Using Variables in a Flow
- Using Functions
- Column Formats and Mapping
- Setting Properties of a Flow
- Running a Data Flow
- Creating a Direct Load Flow
- Creating a DBMS SQL Flow
- Creating a Quick Copy
The Data Flow Interface

The Data Flow interface consists of:

- The workspace, where you build your data flows to extract and move data from data source objects to data target objects. Arrows connect the objects and support the flow.
- Three tabs under the workspace, Data Flow, Process Flow, and Text View, which provide different views of the flow.
- The Flow tab in the ribbon, which enables you to set data flow properties, run flows, and arrange flow objects. The same tab is active when you create process flows.
- The navigation pane, where you choose and store flow components. It has six folders:
  - The Application Directories folder provides access to flows, procedures, and reports that can be generated from a flow. It also provides access to synonyms that describe data. You can use them as data sources or data targets in your flow.
  - The Adapters folder allows you to add and configure adapters.
  - The Resource Management folder is used with Resource Analyzer and Resource Governor.
  - The Workspace folder allows you to configure the server.

As shown in the following image, when you create a new flow, the Data Flow tab opens by default, with an SQL object in the workspace to anchor the flow. By default, the Synonyms button in the navigation pane is active and ready for you to add data source objects. You can change this from the Options dialog box by clearing the Automatically switch filters check box on the General page. The Flow tab is also active.
You can add data source objects and data target objects, and use the SQL object to establish rules for extracting and moving data between them.

For more information on the Data Flow Interface, see *Working in the Data Management Console* on page 125.

**Reference:** **Navigation Pane: Procedures**

On the Home tab, in the Filter group, clicking *Procedures* shows saved flows, procedures, and user functions. You can add flows and procedures to a process flow by dragging them from the navigation pane into the workspace.

If you right-click a flow or a procedure, or highlight one and press the Menu key, a shortcut menu opens.

The following options are available:

**Open**

Adds the flow or procedure to the workspace. A flow opens in the flow editor, while a procedure opens in the text editor.

**Edit as Text**

Opens a text view of the flow in the workspace. This option is only available for flows.

**Note:** If you change the text behind the flow, those manual changes will be lost if you edit and re-save the flow in the DMC.

**Submit**

Submits the flow to the server. This option is only available for a flow. Log information is written to the DataMigrator Log.

**Run**

Runs the flow or stored procedure immediately. Log information appears in the console log window.

**Run Advanced**

Has the following options:

**Submit with E-Mail**

Submits the flow or stored procedure to run in the background. The Submit flow dialog box opens.
The following options are available:

- **Run at a specified date and time.** Allows you to schedule a procedure to run at a specified date and time.

- **Do not wait for completion.** Allows you to continue your work as you wait for the submitted flow or stored procedure to complete. When this option is selected, an additional check box appears which allows you to specify that the flow or procedure should be run at a specified date and time.

- **Submit changing defaults.** Runs a procedure with a prompt to override default values for variables in a procedure or flow, then run it.

- **Submit with Limits.** This option is only available for DataMigrator flows, not stored procedures. Allows you to specify the start and end of the loading records.

- **Report Name.** The optional report name.

- **Schedule Status.** Determines if there is a schedule and if so, whether the schedule is active or inactive.

- **ETLLOG.** Sends the log to ETLLOG and the output to disk.

- **EDAPRINT.** Sends the log to EDAPRINT.

- **E-Mail.** Sends the log and output as an email attachment, if the email notification settings *On Completion* or *On Failure* are requested.

- **Send HTML output as inline message.** When selected, if any stored procedures called by the flow generate an output, the message is included as message text, instead of as an attachment.

  **Note:** This option for a flow schedule works in conjunction with the *Merge formatted output into log* option that is available for a Scheduler Special Service. The following considerations should be taken into account:

  - When the *Send HTML output as inline message* and *Merge formatted output into log* options are selected, the submitted flow emails the attached ETLLOG file and the HTML output is embedded into the body of the email.

  - When the *Send HTML output as inline message* and *Merge formatted output into log* options are not selected, the submitted flow emails the attached ETLLOG file and the HTML output is embedded into the attached ETLLOG file.
E-mail Notification Condition. Allows you to specify conditions for sending email. The options are:

E-Mail On Start
Sends an email when the procedure begins.

E-Mail On Completion
Sends an email when the procedure completes.

E-Mail On Failure
Only used if the procedure fails.

Mail To
Is the field where you enter the email address that the message is being sent to.

On Completion or Failure, the email message will contain attachments of the log and output from the procedure, if there are any.

Run Stress
Generates a script for scalability testing. For more information, see the Server Administration manual for UNIX, Windows, OpenVMS, IBM i, and z/OS.

Schedule and E-mail
Has the following options:

Manage. Allows the schedule and email notification settings for a flow or stored procedure to be changed without opening the flow.

For more information about Scheduler options, see Schedule Attribute Pane on page 427. For more information on email notifications, see How to Set the Flow Email Notification Properties on page 165.

Scheduler Agents. This report runs for the user whose ID value is stored in the sched_scan_id of a Scheduler. The Scheduler must be active to run the Scheduler Agents report.

Scheduled Events. Opens a window that shows all jobs scheduled to run during a specified time range. The Scheduler must be active to run the Scheduler Events report.

Logs
Has the following options:

Last Log. Shows the last log available for the selected flow or stored procedure.
☐ **Last Output.** Shows the last output, if any, for the selected flow.

☐ **Log and Statistics.** Opens the Log and Statistics dialog box that enables you to view a log or statistics report. The following options are available.

**Report Type**

- Allows you to select either a log or statistics report.

**Report Range**

- Allows you to select to view the last report, all the reports, or a report specified within a select time range.

**Completion Status**

- Allows you to view the log for all flows, all completed flows, all flows completed successfully, or all completed flows that contained errors.

**Run by**

- Allows you to view reports for all flows, including those initiated by the scheduler or submitted by the user. This includes flows initiated from the DMC, WC, and CMRUN.

**Application or Flow Name**

- Allows you to specify a specific application, flow, or procedure to run the report on.

**Analysis Reports**

Has the following options:

☐ **Impact Analysis.** Opens a report that shows every flow or stored procedure that calls the selected flow.

☐ **Dependencies Analysis.** Opens a report that shows all synonyms used by the selected data flow and all flows and procedures used by the selected process flow.

☐ **Data Lineage.** Opens a report that shows the name of each field used in the procedure, how it is used, its synonym name, and the name of the table and column in the underlying database.

**Flow Report**

- Provides a summary of the attributes of the flow. This option is only available for flows.

**Download**

- Allows you to download the selected file to the Downloads directory in your system.

**Cut, Copy, Delete, Rename**

- Cuts, copies, deletes, or renames the flow or stored procedure.
Privileges

This option is only available if the server is started with security enabled. For more information, see Manage Privileges Dialog Box on page 189.

Source Control

This option only appears if you have a source management program configured on your machine. It has the following options:

Add/Check In

Adds or checks files into source management.

Check Out

Checks files out of source management.

Undo Check Out

Cancels a check out.

History

Provides the source management history of a file.

Properties

Opens a Properties dialog box of the selected flow or stored procedure.

Procedure: How to Manage Scheduled Flows

1. Right-click Application Directories or an individual application directory, point to Schedule and E-mail, and then click Scheduler Agents.

The Scheduler Agents for window opens, as shown in the following image.

2. Select or clear the check boxes for any flow you want to make active or inactive.

3. On the Report tab, from the Tools group, click Activate Selected or Deactivate Selected.
Reference: **Navigation Pane: Synonyms**

On the Home tab, in the Filter group, click Synonyms to show synonyms that describe the data. You can use them as data sources or data targets in your flow. The synonym type is listed after its name. The type of synonyms available will depend on the adapters that are configured on your server.

If you right-click a synonym, or highlight one and click the Menu key, the following options are made available:

**Open**

Opens the synonym in the workspace. The file, segments, and columns appear on the left, and their attributes and values appear on the right.

**Edit as Text**

Places the synonym code in the workspace where it can be edited.

**Edit Access File as Text**

Places a relational synonym Access File in the workspace where it can be edited. This option is not available for FOCUS files.

**Sample Data**

Shows a sample of the data associated with the synonym.

**Data Profiling**

Provides the data characteristics of the selected synonym columns. The following options are available:

**Statistics**

Provides a data profile of each column of the selected synonym. For more information, see *How to View a Data Profiling Report* on page 550.

**Count**

Shows a count of the number of rows in the selected source table. For a multi-segment synonym, it shows the count for each segment.

**Key Analysis**

Analyzes the data source to find out which columns or combination of columns are unique and could potentially be used as key columns.
**Metadata Management**

The following options are available:

**Create Cluster Synonym**

Opens a new synonym in the Synonym Editor with a reference to the selected synonym. References to additional synonyms can be added to create a cluster synonym comprising multiple tables.

**Refresh Synonym**

Refreshes the selected synonym. Updates the synonym from the database with changed field information while preserving its old title, description, usage, DEFINE field, and DBA information.

**Data Management**

The following options are available:

**Recreate DBMS Table**

Deletes the data and recreates the underlying DBMS table of the synonym.

Synonyms created in Release 7.7.04 or later include information about database indexes.

**Note:** Use caution when using this command if the database table was created using the data description language of a particular RDBMS. Some attributes of relational database tables, including the column order and database constraints, are not reflected in the synonym, and therefore, cannot be recreated.

**Delete ALL Data**

Deletes the data from a table or file described by a synonym. This is useful to clear a synonym that will be used as a data target.

**Drop Table**

Drops the underlying table from the database or deletes a file from the disk.

**Insert Sample Data**

Loads the synonym with sample records (for relational and flat file synonyms). You will be prompted for the number of sample records to load into the table.

**Note:** If you are inserting sample data into a flat file, the synonym must contain a DATASET statement indicating where to put the data.

**Show/Modify Data**

Allows you to view and edit data in a database table.
Quick Copy

Creates a simple data flow that selects all rows and columns from a source table. For more information on this option, see Creating a Quick Copy on page 381.

Impact Analysis

Allows you to assess the impact of a change to the synonym on this server and any subservers that use it. Lists all flows, stored procedures, or reports using the synonym. For data flows, it will indicate whether the synonym is used as a source or a target. Flows and stored procedures are referred to generically as procedures.

Dependencies Analysis

For a synonym that contains a cluster join or multiple tables, all synonyms directly referenced by the selected synonym will be shown.

Download

Allows you to download the selected file to the default Downloads directory in your system. When a synonym is selected, a shortcut menu allows you to select either Master File or Access File. These are the two files that comprise a synonym.

Source Control

This option only appears if you have a source management program configured on your PC. It has the following options:

Add/Check In

Add or checks files into source management.

Check Out

Checks files out of source management.

Undo Check Out

Cancels a check out.

History

Provides the source management history of a file.

For more information regarding source management programs, see Source Management on page 721.

Cut, Copy, Delete, Rename

Cuts, copies, deletes, or renames the synonym.
**Privileges**

This option is only available if the server is started with security enabled. For more information, see *Manage Privileges Dialog Box* on page 189.

**Properties**

Opens a Properties dialog box of the selected synonym.

**Reference:** Advanced Filter

The Advanced Filter gives you finer control over the file types that appear in the navigation pane. For example, you could specify that you only want to see synonyms for the MS SQL Server tables. It also lets you restrict display to files with specified names (using wildcards) or descriptions, files modified in a specific time period, or those of a specified size.
On the Home tab, from the Filter group, click Advanced. The Filter Applications Tree dialog box opens, as shown in the following image.
Simple Data Flow at a Glance

You can use DataMigrator to build data flows ranging from the simple, where data is moved from one source into one target, to the complex, which employs multiple sources and targets and a variety of data extraction techniques.

This simple data flow consists of a data source, an SQL object, and a data target, as shown in the following image.

To create a simple data flow:

1. Open a new flow.
   The Data Flow tab opens by default, with an SQL object in the center of the workspace.

2. To add a data source object, drag a data source object from the list of synonyms in the navigation pane into the workspace, to the left of the SQL object.

3. Select the columns that you want to extract from the source and copy into the target.
   **Note:** If **Automatically select all columns** is selected in the Data Flow Designer user preferences tree from the Options dialog box, you do not need to manually select columns.

4. To add a data target object to the data flow, right-click the workspace to the right of the SQL object, select **Add Target**, and then click **New**.

5. Save and submit the data flow.
   **Note:** Flow names can be up to 64 characters.

6. View the log file for data flow details.
Complex Data Flow at a Glance

You can define a complex data flow that extracts information from multiple data sources, and copies it into multiple data targets using extract transformations, joins, filters, and sorts to refine the results.

This data flow consists of two data sources, an extract transformation, a join, a filter, a sort, an SQL object, and two data targets, as shown in the following image.

To create a complex data flow:

1. Open a new flow.
   
   The Data Flow tab opens by default, with an SQL object in the center of the workspace.

2. To add a data source, drag a data source object from the navigation pane into the workspace, to the left of the SQL object.

3. Add a second data source to the flow.

   The data sources are connected by a join object, which is linked to the SQL object.

   **Note:** The default setting of the DMC automatically adds a join object to the workspace. If the DMC is not automatically adding join objects, you can change the setting by selecting *Add Join Object if needed* in the Data Flow Designer user preferences found in the Options dialog box.

4. Build an expression that creates a new column for extraction in the second data source.
5. Specify join properties.

An inner join is the default. If you are using the default setting for the DMC, an inner join is created automatically for columns that are identically named and formatted, if one of them is a key column.

**Note:** If an inner join is not automatically created, you can change the setting by selecting *Add Join Object if needed* in the Data Flow Designer user preferences found in the Options dialog box.

6. Create an expression that filters the data prior to extraction.

7. Select the columns you want to extract from the source and copy into the target.

**Tip:** You do not need to manually select columns if *Automatically select all columns* is selected in the Data Flow Designer user preferences found in the Options dialog box.

8. To add two data target objects to the data flow, drag one from the navigation pane and create the other.

9. Map the columns from the joined data sources to the existing data target.

   The columns in the new data target are mapped automatically.

   **Tip:** You do not need to map columns manually if the source and target column names are the same if *Automatically map all columns* is selected in the Data Flow Designer user preferences found in the Options dialog box.

10. Save and run the data flow.

   **Note:** Flow names can be up to 64 characters.

11. View the log file for data flow details.

   **Tip:** You can create this data flow by completing *Exercise 1: Creating the Data Flow* on page 438. In addition, the exercises in *Advanced DataMigrator Tips and Techniques* on page 555, include several complex data flows that illustrate a range of useful options.

**Creating a Data Flow**

The first step in building a data flow is to create an empty flow. You will then be able to add objects to this flow, and set up mappings and relationships.

**Procedure:** How to Create a Data Flow

1. Right-click the application directory where you want the new data flow.
2. Point to *New* and then click *Flow*.

   or

   Click the New button on the Quick Access Toolbar.
An empty data flow appears in the data flow workspace. The Data Flow tab opens by default, with an SQL object in the center of the workspace.

**Adding and Deleting Objects in Your Data Flow**

You can add or delete objects in the data flow workspace. Objects on the left of the SQL object are sources, and objects on the right are targets.

**Note:** When adding objects, there is no limit to the area they can take up. If the number and arrangement of objects exceeds the viewable area of the workspace, a scroll bar will appear.

**Procedure:** **How to Add Objects to Your Data Flow**

- Drag objects from the navigation pane into the data flow workspace.
  
  or
  
  - Right-click the data flow workspace. On the left side of the SQL object, you can click *Add Source, Add Join, Add Select,* or *Add Union.* On the right side of the SQL object, you can click *Add Target.*
  
  or
  
  - Right-click a *Join, SQL,* or *Union* object and click *Add.* Select the object from the available list.
  
  or
  
  - On the Flow tab, in the Insert group, click one of the commands, and drag it into the workspace.
    
    If you are adding a source or an existing target, the Select Synonym dialog box opens.
    
    Select a synonym and click the *Select* button.

**Procedure:** **How to Search for Existing Synonyms to Use as Sources or Targets**

The Select Synonym dialog box has a search feature that allows you to find synonyms by searching their names and descriptions.

1. Right-click the workspace while working with a data flow, point to *Add Source* or *Add Target,* and then click *Existing.*
   
   The Select Synonym dialog box opens.

2. Click the *Find* button at the upper right.
   
   The Find dialog box opens.
3. Type a string in the Find What field, and click the Find Next button. You can refine your search with match whole word only and match case. You can also search in either direction.

   The first synonym name, or description containing the string, is highlighted in the Select Synonym dialog box.

4. Click the Find Next button to find the next match.

   **Note:** After closing the Find dialog box, you can still search for the next match by pressing the F3 key. Shift + F3 will find the next match searching upward.

**Reference:** Select Synonym Dialog Box

To open the Select Synonym dialog box while working in the Data Flow tab, right-click the workspace and click Add Source or point to Add Target, and then click Existing.

The Select Synonym dialog box has five default columns, but this can be customized. The following fields and options are available:

**Look in**

Is the name of the application directory where the synonyms are located. Select the server name to see all synonyms in all directories.

**Name**

Is the name of the synonym.

**Type**

Is the type of synonym.
**Description**

Is the description of the synonym.

The following two columns are available if you are adding a source to a Join object that already has a synonym selected:

**Join Conditions**

Are the join conditions that can be used to join the already selected synonym to a new synonym.

**Join Type**

Displays how the join conditions were found. You can find join conditions using:

- **Matching Key Names.** Identically named fields were found in both synonyms.
- **Matching Primary Keys.** The already selected synonym has a foreign key relationship to a primary key in the new table.
- **Matching Foreign Keys.** The new table has a foreign key relationship to a primary key in the already selected table.

**Note:** Clicking any of the column headings will sort the list on that field. A column sorted ascending will display an up arrow, while a column sorted descending will display a down arrow.

**New**

Allows you to create a new synonym.

**Find**

Allows you to search the application directories for a particular synonym.

**Filter**

This button is enabled when you already have one table connected to a join condition, and you right-click a Join object, point to Add, and then click Source.

When this option is selected, only tables in the selected application directory, or server with columns that match key columns, are displayed.

When this option is not selected, all tables in the selected application directory or server are displayed.
**Name text box**

Allows you to enter a synonym name, or shows the name you have selected.

**Type**

Allows you to select the synonyms by their adapter type. For example, you can display only synonyms for MS SQL Server tables.

**Select button**

Adds the highlighted synonym to the workspace.

**Title Bar**

Left-click a column header in the title bar to sort the list of synonyms on that column. Do it again to reverse the order. Right-click a column header in the title bar to open the context menu with the following options:

- **Sort Ascending/Descending.** Sorts the list of synonyms on the selected column.
- **Find.** Opens the search dialog.
- **Tree View/List View.** Switches views of the synonyms.
- **Best Fit.** Optimizes the column widths for best viewing.
- **Customize.** Allows you to select the columns that are displayed.

**Procedure:** How to Delete Objects From Your Data Flow

Right-click the object you want to remove and click *Delete*.

**Connecting and Disconnecting Objects in Your Data Flow**

The connections between objects in the data flow workspace allow information to flow between them.

The DMC default User Preferences automatically connects objects.

- Objects dragged onto the left (source) side of the workspace are connected to the SQL object, or, if more than one data source is dragged onto the workspace, to each other through a Join object.

- Objects dragged to the right (target) side of the workspace are connected to the SQL object.
**Note:** The Options dialog box provides access to the DMC User Preferences. If the *Add Join Object if needed* setting on the Data Flow Designer page is selected, Source objects are automatically connected to a Join object, if more than one is dragged onto the workspace. For details, see *How to Automatically Add Join Objects* on page 235.

**Procedure:** How to Connect Objects

When objects are deleted, you may need to recreate the connections between some of the remaining objects.

1. Right-click an object, hold the mouse button down, and drag to the object you want connected.
2. Release the mouse button.

An arrow will appear indicating that a connection has been made.

You may also need to create connections if the DMC default settings have been changed.

**Procedure:** How to Disconnect Objects

Right-click the arrow connecting two objects, and click *Delete*.

or

Right-click an object, hold the mouse button down, and drag an arrow to a different object.

**Saving a Data Flow**

Data flows are saved on the DataMigrator Server. The first time you save a data flow, you will select an application directory and assign a name to it. After that, you can save it with the same name, or assign a new one.

The following restrictions apply to flow names:

- Names can contain only letters, numbers, and underscores.
- Uppercase letters are converted to lowercase when the flow is saved.
- Names can be up to 64 characters long, except on z/OS with PDS deployment, where names are limited to 8 characters.

**Procedure:** How to Save a Data Flow

1. Click the Save button on the Quick Access Toolbar.

   **Note:** A flow can only be saved once changes have been made to it. If no changes have been made, both the button and Save option in the main menu will be inactive (grey).
If this is the first time a flow is being saved, the Save As dialog box will open. See *How to Save a Data Flow with a New Name* on page 213.

2. Check the console window for a message indicating that the flow was saved successfully.

**Procedure:** *How to Save a Data Flow with a New Name*

1. Click the Save As button on the Quick Access Toolbar.

   The Save Procedure As dialog box opens.

2. Select the application directory where you want to save the flow.

3. In the Name box, enter the name you wish to use for the flow.

   Click Save.

4. Check the console window for a message indicating that the flow was saved successfully.

**Specifying Data Sources**

Once you have opened a new data flow, you need to identify the data sources you want to use, and add them to the flow. Data sources are represented by synonyms in the navigation pane, which are under an application directory folder when the Synonyms button is selected. The options available when you right-click a synonym are described in *Navigation Pane: Synonyms* on page 200. For guidelines on synonyms used as data sources, see *Using Synonyms for Data Sources and Data Targets* on page 108.

When creating a data flow, information about the data source, such as the column names, is required and supplied when the DataMigrator Server is connected.

When creating or editing a data flow, in order to test your flow, data sources referenced in the flow must be available.

- The server with a relational data source must be running to use that source.
- A data set must be allocated for data in a flat file or VSAM structure.
- The source servers must be running. and they must be reachable from the Hub server in a typical configuration with a Hub server and one or more remote servers.

**Procedure:** *How to Determine What is in a Data Source*

Before you add a data source to a flow, view the columns to verify that it contains the data that you want and that the data is available.
To view a sample of the data that a synonym will produce, right-click the synonym in the navigation pane and click Sample Data. The sample data will appear in the data flow workspace, as shown in the following image.

![Sample Data Image]

**Note:** After a synonym has been dragged into the data flow workspace, you can view its columns, sample data, and transformations. Right-click the synonym and click Expand. You can view the data by clicking either the Columns, Sample Data, or Test Transforms tab.

**Procedure:** How to Add a Data Source to a Data Flow

1. If you have not already created a new data flow, right-click the application directory where you want the new data flow, point to New, and then click Flow.

   An empty data flow appears in the data flow workspace. The Data Flow tab opens by default, with an SQL object in the center of the workspace.

2. Drag a data source object from the navigation pane into the data flow workspace, to the left of the SQL object.

   or

   Right-click in the workspace to the left of the SQL object, and click Add Source.

   The Select Synonym dialog box opens.

3. Select a synonym and click the Select button.

**Procedure:** How to Change a Data Source in a Data Flow

With the data flow open in the workspace:

1. Right-click the Source object that you want to change and click Change Source.

   The Select Synonym dialog box opens.

2. Select the replacement synonym and click the Select button.
**Note:** If you change a data source that has source transformations associated with it, the transformations are preserved, but they will appear red if the new source does not contain all of the columns used in the transformation.

**Procedure:** How to Use a Relational Stored Procedure as a Data Source

You can use a relational stored procedure as a data source in a flow if you have created a synonym for that stored procedure. For information on creating a synonym, see *How to Create a Synonym for a Relational Stored Procedure* on page 76.

One use for this would be to perform a lookup, for example, passing parameter values to the stored procedure and retrieving an answer set based on the supplied parameter.

This is supported for:

- Db2 (using CLI)
- Oracle (using CLI)
- SQL Server (using OLE DB and JDBC)
- Sybase IQ (using ODBC).
- Teradata (using ODBC)
- MySQL (using JDBC)

If you have created a synonym for a relational stored procedure, it will appear in the navigation pane, along with all other synonyms.

As an example, we are using an Oracle stored procedure that has `empno` as a parameter, and that retrieves information associated with that employee number. The Oracle stored procedure goes against the EMP Oracle sample table. We will join the stored procedure to a flat file that contains valid employee numbers.

The sample stored procedure is shown below.

```sql
/* create get_emp routine */

CREATE OR REPLACE PACKAGE get_emp AUTHID CURRENT_USER AS
PROCEDURE get_emp_proc (pempno IN scott.emp.empno%TYPE,
                         pname   OUT scott.emp.ename%TYPE,
                         pjob    OUT scott.emp.job%TYPE,
                         pMgr    OUT scott.emp.mgr%TYPE)
);
END get_emp;
```
CREATE OR REPLACE PACKAGE BODY get_emp AS
PROCEDURE get_emp_proc (pempno IN scott.emp.empno%TYPE,
pename OUT scott.emp.ename%TYPE,
pjob OUT scott.emp.job%TYPE,
pMgr OUT scott.emp.mgr%TYPE)
IS
CURSOR gRec_cur IS
SELECT ename, job, mgr
FROM scott.emp
WHERE EMPNO = pempno;
gRec_rec gRec_cur%ROWTYPE;
BEGIN
OPEN gRec_cur;
FETCH gRec_cur INTO gRec_rec;
IF (gRec_cur%FOUND) THEN
    pname := gRec_rec.ename;
pjob := gRec_rec.job;
pMgr := gRec_rec.mgr;
END IF;
CLOSE gRec_cur;
EXCEPTION
    WHEN OTHERS THEN
        ROLLBACK;
        RAISE;
END get_emp_proc;
END get_emp;
/
show error
GRANT EXECUTE ON get_emp TO public;

To use a relational stored procedure as a data source:

1. With a data flow open in the workspace, drag the synonym into the workspace, to the left of the SQL object.

   Note: The flow we are using joins a flat file named ora_emps with the get_emp_get_emp_proc Oracle stored procedure. The flat file has only one column, empno, which is joined to the stored procedure. For each empno in the flat file, the stored procedure runs and retrieves the name, job, and manager information for that empno.
2. Right-click the synonym, point to Operations, and then click Sample Data. The Sample Data dialog box opens.

3. Select the parameter check boxes, enter their values, and click Sample Data.
   
   **Note:** If the stored procedure can have more than one answer set, you need to select the appropriate one.

   The sample data appears in the workspace.

   Whenever you test sample data, you will need to supply the values for the parameters.

   ![](image)

   ![](image)

   Reference: Data Source Object Options

   After a synonym has been dragged into the workspace, the following options are available when you right-click the synonym:

   **Expand/Collapse**

   Opens or closes a window that allows you to view the columns, sample data, and transformations of the synonym.

   **Refresh Columns**

   Updates the list of columns in a synonym.

   **Source Transformations**

   Opens the Source Transformations window. For details, see Creating a Source Transformation in a Data Source Object on page 219

   **Operations**

   **Open**

   Opens the synonym in the workspace. The file, segments, and columns appear on the left and their attributes and values appear on the right.

   **Edit as Text**

   Opens the synonym syntax in the workspace, where it can be edited.

   **Sample Data**

   Shows a sample of the data associated with the synonym.
Data Profiling

Provides the data characteristics for synonym columns.

Change Source

Opens the Select Synonym dialog box, where a different source can be selected.

Delete

 Deletes the synonym from the data flow.

Properties

Opens the general attributes of the synonym. The following properties are available:

- **Adapter.** The name of the adapter.
- **Application Directory.** The application directory for the synonym.
- **Synonym Name.** The name of the synonym.
- **Description.** A description of the selected synonym.
- **Table Name.** The name of the source table or data file.
- **Connection.** The name of the connection.
- **Tag.** The correlation name of the synonym. You can change the default name (T1, T2, and so on) to one that is more descriptive.
- **Display Name.** An alternate name you can give to the data source.
- **Notes.** An annotation describing the synonym.

When more than one data flow object is selected, the following additional options are available:

Collapse Selected

Collapses the selected objects. This option is only available when one or more of the selected objects is expanded.

Delete

Deletes the selected objects. This option is not available when the SQL object is selected.

Updating the List of Data Sources

You can update the list of data sources by refreshing the tree. You can update the list of columns in a data source by refreshing the source.
This is useful when data sources are added to the server, or when columns are added to a data source while you are building a data flow, and you want these changes reflected in the data source or column list.

**Procedure:** How to Update the List of Data Sources

In the navigation pane, right-click the server, the Applications Directories folder, or subfolder, and click Refresh.

**Procedure:** How to Update the List of Columns

In the navigation pane, right-click a data source, and click Refresh.

**Note:** After you drag the data source onto the data flow workspace, you can refresh the column list by right-clicking the data source object and clicking Refresh Columns.

**Creating a Source Transformation in a Data Source Object**

You can create a source transformation in a data source to add intermediate calculations to the request. Source transformations are performed on a data source object as the records are read, before any filtering or aggregation is performed. This is often useful for complex calculations that are composed of multiple expressions. Creating a source transformation is also helpful in performing data type conversions when multiple steps are required, and if you want to aggregate or filter a calculated column.

Source transformations are only performed for a particular data flow.

A source transformation can contain an expression or a constant.

- If the source transformation contains an expression, you can create the expression with the Transforms Calculator.

- You can also enter a value or expression in the expression area. The format is automatically determined and displayed.

**Note:**

- If you are adding a source transformation to a hierarchical file (for example, FOCUS/FDS), the WITH attribute may be required. The WITH attribute may also be required if you add a source transformation that references a constant.

- Use of some source transformations disables Automatic Passthru. For more information, see Using Automatic Passthru on page 494.

- You can create a virtual column that appears whenever a data source is used by adding it to the data source itself via the navigation pane. For more information, see Adding Virtual Columns (DEFINE) in a Synonym on page 540.
When re-using a synonym in a data flow, such as for a recursive join, the transformations must be the same in all instances of the synonym in the flow.

If different transformations are required in each instance, you should create multiple synonyms with different synonym names for these sources and use them in the data flow.

In addition to source transformations, you can use the Relational Expression tab of the Transformation Calculator to create a relational expression transformation that can be used in a filter.

Procedure: How to Create a Source Transformation in a Data Source Object

1. From the data flow workspace, double-click a data source object, or right-click it and click Source Transformations.

   The Source Transformations dialog box displays all of the columns in the data source.

   ![Source Transformations dialog box](image)

   **Note:** The columns of the synonym are shaded because they cannot be edited here. Only source transformations can be edited in this dialog box.

2. Click the Insert Transforms button.

   The Transformation Calculator opens.

3. Type a name and format for the virtual column.

   **Note:** If you do not supply a format, the DMC will automatically provide the format based on the expression. Similarly, if you supply the wrong format, the DMC will replace it with the correct format using a pop-up notification to advise you of the change.
4. Use the Columns/Variables and Functions tabs, and the calculator buttons, to build the expression for the virtual column. For information about using functions, see *Using Functions* on page 347. For information about using variables, see *Using Variables in a Flow* on page 336.

**Note:** When you create expressions, keep in mind their order of evaluation. For details, see *Order of Evaluation* on page 233.

5. Click **OK**.

The new column appears in the Columns list.

6. To verify that the transformation is working properly, click the Test Transforms button.

The Test Transformations dialog box opens, showing the result of the source transformation.

7. Click **OK** to close the Source Transformations dialog box.

**Procedure:**  **How to Create a Transformation That Uses a Variable**

From the Source Transformations dialog box:

1. Click the Insert Transforms button.

   The Transformation Calculator opens.

2. Type a name and type or select a format for the virtual column.

   **Note:** If you do not supply a format, the DMC will automatically provide the format based on the expression. Similarly, if you supply the wrong format, the DMC will replace it with the correct format using a pop-up notification to advise you of the change.
3. Double-click the variable you want assigned to the column from the Columns/Variables tab. You can also type in the variable, which must be enclosed in single quotation marks (’), as shown in the following image.

![Transformation Calculator Image]

4. Click OK.
The variable appears in the Expression column in the Source Transformations dialog box, as shown in the following image.

![Source Transformations](image)

For more information about using variables, see Using Variables in a Flow on page 336.

**Procedure:** How to Create a Relational Expression in the Transformation Calculator

The Relational Expression tab of the Transformation Calculator allows you to build an expression that can be used as a filter.

From the Source Transformations dialog box:

1. Click the Insert Transforms button.
   
   The Transformation Calculator opens.

2. Click the **Relational Expression** tab.

3. Type in a name for the column.

   **Note:** Since the transformation will always evaluate to either 1 (true) or 0 (false), the format will be automatically set to I1.

4. Optionally, type in a title for the column.

5. Click the Add New Row button above the Expression field.
   
   A blank line is added to the Expression field.
6. Click the drop-down menu in the Field column, as shown in the following image.

7. Select a Column from the drop-down menu.

8. Select a Relation from the drop-down menu.

9. Select a Type from the drop-down menu.

   The following options are available:

   - Value
   - Field
   - Variable

   **Note:** This field is not available if the Relation is set to IS MISSING or IS-NOT MISSING. If the Relation is set to DB_FILE or NOT DB_FILE, Type must be a value.

10. Type in a value or click the ellipsis (…) to select an available value.

    The Value(s) Selection dialog box opens.

11. Select one (or more) Available Values and move it to the Selected Value(s) pane using the arrow. Click OK.
The Transformation Calculator should look like the image below.

![Transformation Calculator Image]

12. Click OK to close the Transformation Calculator.

The relational expression appears as a column in the Source Transformations dialog box, as shown in the following image.

![Source Transformations Image]

**Note:** Using a source transformation column in a filter or as a selected column may disable Automatic Passthru, so it should be used with care.
Reference: **Source Transformations Dialog Box**

To access the Source Transformations dialog box from the data flow workspace, double-click a source, or right-click a source and click **Source Transformations**.

The Source Transformations dialog box has the following fields and options:

**Source Columns**

Lists the columns in the data source. The first section contains the column order.

The columns symbols are:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Symbol]</td>
<td>Key column</td>
</tr>
<tr>
<td>![Symbol]</td>
<td>Regular column</td>
</tr>
<tr>
<td>![Symbol]</td>
<td>Index column</td>
</tr>
<tr>
<td>![Symbol]</td>
<td>Virtual column</td>
</tr>
</tbody>
</table>

**Note:** The symbol for columns in the synonym is gray, while the background for a source transformation (virtual column in the flow) is white.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Symbol" /></td>
<td>Source Transformation</td>
</tr>
</tbody>
</table>

**Display Name (Title)**

Is the name of the column, and also shows a symbol that indicates regular columns, key columns, index columns, and virtual columns.

**Format**

Is the format of the column.

**Expression**

Is the expression for the column.

**Description**

Is the description of the column.

**Find**

Allows you to search for a column.

**Refresh**

Refreshes the columns in the data source.

**Sample data**

Shows a sample of the data from the data source.

**Insert transforms**

Opens the Transformation Calculator to create a transformation.

**Delete transforms**

Deletes the selected transformation.

**Edit transforms**

Opens the Transformation Calculator to edit a transformation.
**Test transforms**

Tests the transformations.

**Up and down arrows**

Moves the selected transformation up or down within the Source Transformations list.

**Reference:** Column Header Shortcut Menu

To access the column header shortcut menu, right-click any column header.

The column header shortcut menu has the following options:

- **Find**
  
  Allows you to search for a column

- **Best Fit**
  
  Resizes the column to fit its data.

- **Customize**
  
  Allows you to select what information is displayed for each column.

- **Paste Transforms**
  
  Pastes transformations from memory into the transformation list.

**Reference:** Source Transformations Shortcut Menu

To access the source transformations shortcut menu, right-click any column of a transformation.

The source transformations shortcut menu has the following options:

- **Edit Transformation (Advanced)**
  
  Opens the Transformation Calculator for the selected transformation. This is the default action.

- **Insert Transformation**
  
  Displays a shortcut menu with a list of available functions that can be added as a transformation. Selecting one of the functions opens the Function Assist dialog box, which allows you to specify parameters for that function.

- **Delete**
  
  Deletes the transformation.
**Test Transformations**

Tests all of the transformations. This can also be done by clicking the *Test Transforms* button.

**Error Details**

This option is only available if the selected transformation contains an error, which is displayed in red text. Selecting this option opens a dialog box with the error code and a short explanation of the error.

**Copy Transformations**

Copies the selected transformations to memory, so that they can be pasted into this, or another, flow.

**Paste Transformations**

Pastes transformations from memory into the transformation list.

**Note:** Copying and pasting transformations does not use the Windows clipboard.

**Reference:**  Transformation Calculator

To access the Transformation Calculator, click the *Insert* button in the Transformations dialog box.

![Transformation Calculator](image)
The Transformation Calculator has the following fields and options:

**Name**

Is the name of the column.

**Format**

Is the format of the column.

*Note:* If you do not supply a format, the DMC will automatically provide the format based on the expression. Similarly, if you supply the wrong format, the DMC will replace it with the correct format using a notification to advise you of the change.

**Title**

Is the title of a new target table created by a data flow. The title will be included in the synonym that is created.

**Missing**

Indicates if missing (null) values are allowed in this field. The following options are available:

- **OFF.** The result is never missing.
- **ON.** The result is missing only if all missable fields are missing.
- **ON ALL.** The result is missing if any missable fields are missing.

**Expression tab**

Displays the expression dialog box.

**Relational Expression tab**

Displays the Relational Expression Builder dialog box.

**Columns/Variables tab**

Displays available columns and variables in a hierarchical tree or in a grid.

**Functions tab**

Displays a list of functions that are available for your transformations.

**SQL Optimization button**

Opens the SQL APT Status report.
Check Expression button

Verifies the validity of the expression.

Sample Data button

Produces sample data for the expression.

Find button

Allows you to search for a function or variable by name.

Calculator buttons

Insert numbers and operators.

Function Assist button

Allows you to specify parameters for the function through a dialog box when creating or editing a transformation.

The following operators are available:

| (single concatenation bar)

Concatenates two values, retaining any trailing blanks after the first one. For example, if FIRST_NAME and LAST_NAME were both in A15 format, the expression

\[ \text{FULL_NAME} = \text{FIRST_NAME} \ | \ \text{LAST_NAME} \]

would produce a column like the following:

| MICHAEL | SMITHSON |
| JANE | JONES |

|| (double concatenation bar)

Concatenates two values, suppressing any trailing blanks in the first. For example, to construct the full name and insert a comma, the syntax

\[ \text{FULL_NAME} = \text{LAST_NAME} \ | \ (\ , \ ) \ | \ \text{FIRST_NAME} \]

would produce a column like the following:
SMITHSON, MICHAEL
JONES, JANE

The concatenation in the parentheses is done first (preserving the blank space after the comma), and the result is then concatenated to LAST_NAME, suppressing the trailing blanks of LAST_NAME.

**IF**

Establishes a conditional test.

**THEN**

Specifies the action to perform if the result of a conditional test is TRUE.

**ELSE**

Specifies the action to perform if the result of a conditional test is FALSE.

**LT**

Returns the value TRUE if the value on the left is less than the value on the right.

**NOT**

Returns the value TRUE if the operand is false.

**LE**

Returns the value TRUE if the value on the left is less than, or equal to the value on the right.

**EQ**

Returns the value TRUE if the value on the left is equal to the value on the right.

**AND**

Returns the value TRUE if both operands are true.

**GT**

Returns the value TRUE if the value on the left is greater than the value on the right.

**GE**

Returns the value TRUE if the value on the left is greater than, or equal to the value on the right.
NE
Returns the value TRUE if the value on the left is not equal to the value on the right.

OR
Returns the value TRUE if either operand is true.

**
Raises a value to the specified power.

()
Adds parentheses.

""
Inserts two single quotation marks. Enter alphanumeric test values between these.

a->A
Converts selected text to uppercase.

A->a
Converts selected text to lowercase.

Date
Opens the Date Editor dialog box, which lets you use the current date or specify a date from the calendar.

Datetime
Opens the Date Editor dialog box, which lets you use the current date and time or specify a date and time from the calendar.

Order of Evaluation
The order of evaluation can affect the result of an expression in a transformation. DataMigrator evaluates numeric and Boolean operators from left to right in the following order:

- Operations enclosed in parentheses
- Exponentiation
- Multiplication and division
- Addition and subtraction
- EQ, NE, LE, LT, GE, GT, NOT
Changing the Order of Evaluation

The following expression gives an incorrect result because UNIT_SOLD is first multiplied by UNIT_COST, and then the result is subtracted from RETAIL_PRICE:

\[ \text{RETAIL\_PRICE} - \text{UNIT\_COST} \times \text{UNIT\_SOLD} \]

You can change the order of evaluation and improve readability by enclosing expressions in parentheses. An expression in parentheses is evaluated first. The correct expression for the preceding example is:

\[ (\text{RETAIL\_PRICE} - \text{UNIT\_COST}) \times \text{UNIT\_SOLD} \]

Joining Data Sources

If you select more than one data source for your data flow, you must use a Join object to connect them.

You need to specify the common column or columns that will link the tables, as well as the type of join.

You can perform an inner, left outer, right outer, full outer, or cross join.

- An inner join extracts those rows that appear in both tables. This is the default.

- A left outer join includes all of the rows from the table listed in the Left Source Columns grid, and the columns from the Right Source Columns grid that match one of these columns.

- A right outer join includes all of the rows from the table listed in the Right Source Columns grid, and the columns from the Left Source Columns grid that match one of these columns.

- A cross join is a Cartesian product of two tables. It consists of all possible pairs of rows from the two tables.

- A full outer join includes all of the rows that are extracted from both data sources.

**Note:** You can create a recursive join by selecting the same data source twice.

When you link data sources, DataMigrator identifies each table with \( T_n \), where \( n \) is the order in which the table was chosen. For example, the first source chosen would be identified as \( T_1 \), the second as \( T_2 \), and so on.
DataMigrator creates join syntax in ANSI 92 SQL. For example, when selecting the tables Product and Vendor, and columns VENDORNUM and VENDOR_NUM, DataMigrator generates the SQL syntax:

```
SELECT...FROM Product t1 INNER JOIN VENDORS T2 on t1.vendornum=T2.vendor_num
```

**Reference:** Considerations for Join Performance

iWay lets you create joins across similar data sources, different data sources, and across servers (also known as cross-platform joins).

In general, cross-platform joins, and joins without key fields, can create many intermediate files, depending on the number of tables joined. This issue may be less significant for some iWay platforms. For example, the Server for z/OS can use hyperspaces and alternate sort utilities. If performance is a concern, you should carefully evaluate the impact of your joins.

To improve performance, try any of the following:

- Minimize the number of tables used in a join.
- Join tables using key columns and indexes, when possible.
- Make sure that the smaller of the two tables is selected first, and the larger one is selected second.

**Procedure:** How to Automatically Add Join Objects

The default settings of the DMC automatically create Join objects. If Join objects are not being created, you can change your settings so that Join objects are created automatically.

1. On the Home tab, in the Tools group, click Options.
2. Expand Data Flow and select Data Flow Designer from the tree.
3. Select the Add Join Object if needed check box in the On Add Source section.
4. Click OK.

**Procedure:** How to Join Data Sources

If the DMC is using the default settings, Join objects are created automatically when you drag additional data sources into the data flow workspace.
You can also add a join by right-clicking in the data flow workspace and selecting Add Join. In that case, you must connect the data source to the join by right-clicking it and dragging it to the Join object.

1. To complete the join, double-click the Join object, or right-click a Join object and select Join Editor.

   The Join Editor dialog box opens, listing the columns from the two data sources in the Left and Right Source Columns lists. For more information, see Join Editor Dialog Box on page 238.

   If the two data sources have a column with the same name and format, and that column is a key column in one of the sources, an inner join will automatically be created for those columns. This default join can be changed, if necessary.

2. If you want to join columns with different names, select one column from each list, and click the equal sign to create the join.

   The equality condition will appear in the Expression box under Join Conditions.

   DataMigrator supports both equi and non-equi (conditional) joins. In an equi-join, the condition is one of equality, while a non-equi join uses any other condition, such as greater than, less than, and so on.

   **Note:** You can delete the condition from the Expression box by selecting it and clicking the Delete Condition button.

   You can edit the condition by selecting it in the Expression box and clicking the Edit Condition button.

3. Select the type of join you want from the Join Type drop-down menu, by clicking the join diagram, or right-clicking anywhere in the diagram and selecting a join type from the menu.

   - Click the area where the circles intersect for an inner join. This is the default.
   - Click the left circle for a left outer join.
   - Click outside the circles for a cross join or Cartesian product.
   - Click the right circle for a right outer join.
   - Click an inner circle a second time for a full outer join.

   The type of join you create will be graphically displayed in the workspace.
4. To add a calculation or SQL function to the condition, select it and click the *Edit Condition* button. When the Join Calculator opens, enter your calculation in the Conditions box by typing it, or by using the Columns/Variables and Functions tabs and the calculator buttons. For details, see *Join Calculator* on page 240. For information about using functions, see *Using an SQL Function* on page 258. For information about using variables, see *Using Variables in a Flow* on page 336.

5. To define a new condition, click the *Insert Condition* button, which will launch the Join Calculator. This is where you can create the condition.

6. Click the *Test Join* button to test the join.

7. Click *OK* to save the join.

8. Drag another synonym into the workspace, to the left of the SQL object.

9. Double-click the second Join object.

10. Select one column from each list, and click the equal sign to create the join. The equality condition will appear in the Expression box under Join Conditions.

11. Click *OK*.

An example of a data flow with three joined sources is shown in the following image.

![Data Flow Diagram](image)

**Procedure:** How to Create a Join Condition With the Join Calculator

1. From the Join Editor dialog box, click the *Add new row* button, and then click the *Calculator* button.

   The Join Calculator opens.

2. Enter your calculation in the Expressions box by typing it, or by using the Columns/Variables and Functions tabs and the calculator buttons. For information about using functions, see *Using an SQL Function* on page 258. For information on using variables, see *Using Variables in a Flow* on page 336.
3. Click OK.

**Procedure: How to Edit a Join Condition With the Join Calculator**

1. From the Join Editor dialog box, select the condition and click the *Edit Condition* button.
   
   The Join Calculator opens.

2. Edit your calculation in the Expressions box by typing it, or by using the Columns/Variables and Functions tabs and the calculator buttons.

3. Click OK.

**Reference: Join Editor Dialog Box**

To access the Join Editor dialog box from the workspace, double-click a Join object, or right-click a Join object and select *Join Editor*.
Next to each column name, there is a symbol that indicates regular columns, key columns, index columns, or virtual columns.

**Note:** Clicking a column heading will sort the column grid on that heading.

**Right Source**
Displays the available columns from the second data source.

The columns symbols are:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>⌨</td>
<td>Key column</td>
</tr>
<tr>
<td>⬅️</td>
<td>Regular column</td>
</tr>
<tr>
<td>➘️</td>
<td>Index column</td>
</tr>
<tr>
<td>✗️</td>
<td>Virtual column</td>
</tr>
</tbody>
</table>

=  
Establishes the relationship between the selected columns as equal.

**Find**
Allows you to search for a column.

**Insert condition**
Opens the Join Calculator to create a condition.

**Delete Condition**
Deletes the selected condition.

**Edit condition**
Opens the Join Calculator to edit a condition.
Join Analysis

Opens the Join Analysis window, where the join condition is analyzed to show the number of rows and distinct rows that would result from a left outer, inner, and right outer join.

Test join

Tests the join conditions.

Note: If you have not selected columns, test join displays the keys columns from the tables in the join. If you have selected columns, it displays the columns selected. If neither table has keys, and no columns were selected, test join displays an empty answer set.

Join Conditions

Lists the conditions by which the two data sources are connected.

Join Type

Enables you to select the join type from the drop-down menu or by clicking in the join diagram, which graphically represents the type of join being used.

Reference: Join Calculator

To access the Join Calculator from the Join Editor dialog box, select a row, and click the Edit Condition button.
The Join Calculator contains the following fields and options:

**Conditions window**
Displays the expression.

**Columns/Variables tab**
Displays available columns and variables.

**Functions tab**
Displays a list of SQL functions that are available for your transformations.

**Function Assist button**
Allows you to specify parameters for the function through a dialog box when creating or editing a transformation.

**Calculator buttons**
Insert numbers and operators.

The following operators are available in both the Join and Filter calculators:

**BETWEEN**
Specifies data will be returned if the value of the field falls between the values entered on either side of the word AND. For example:

\[\text{RETAIL BETWEEN 20 AND 30}\]

**IN**
For alphanumeric data, this option specifies data will be returned if the field value is a substring of the right operand.

**IS NULL**
Specifies that records with a null value for the specified column are returned.

**LIKE**
Returns data if the value of the field matches the character string on the right operand. In the character string, use the underscore character (_) to represent any single character, and the percent character (%) to represent any sequence of characters. For example, the following string returns the value TRUE if LAST_NAME contains the string SMITH anywhere inside it (for example, SMITH, SMITHSON, SILVERSMITH):

\[\text{LAST\_NAME LIKE ’SMITH’}\]

The following string returns the value TRUE if LAST_NAME is exactly five characters long and the first letter is S.
LAST_NAME LIKE 'S_ _ _ _'

<

Returns data if the value of the field is less than the right operand. For example, the following string restricts the query to those employees who earn less than $30,000:

SALARY < 30000

<=

Returns if the value of the field is less than, or equal to the right operand.

=

Specifies data if the field value equals the right operand.

AND

Is used in combination with the BETWEEN operator.

NOT

Is used in combination with other relational operators to negate them. For example, the following string returns all records where the value of LAST_NAME is not null:

LAST_NAME IS NOT NULL

>

Specifies data will be returned if the value of the field is greater than the right operand.

>=

Specifies data will be returned if the value of the field is greater than, or equal to the right operand.

<>

Specifies data will be returned if the left operand is not equal to the right operand.

OR

Inserts the operator OR.

( )

Adds parentheses.

""

Inserts two single quotation marks ('). Enter alphanumeric test values between these.
a->A

Converts selected text to uppercase.

A->a

Converts selected text to lowercase.

**Date**

Opens the Date Editor window, which lets you use the current date or specify a date from the calendar.

**Datetime**

Opens the Date Editor window, which lets you use the current date and time or specify a date and time from the calendar.

**Reference:** Join Analysis Dialog Box

To access the Join Analysis dialog box from the Join Editor dialog box, click the *Join Analysis* button.

<table>
<thead>
<tr>
<th></th>
<th>Left</th>
<th>Middle</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95</td>
<td>43</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>95</td>
<td>103</td>
<td>1</td>
</tr>
</tbody>
</table>

The Middle column shows the number of rows (Count) and unique rows (Distinct Count) that the entered join criteria would return for an Inner join.

The Left and Right columns show the number of additional values and rows that would be returned for their respective join type.

From the example shown in the above image, a Left outer join would return 95 additional rows with 95 unique values, in addition to the rows returned by the inner join.
Note that all of the numbers in each row or column are also hotlinks. For example, if you double-click 43, a Test Results window opens, as shown in the image below. This window shows the 43 values for Store ID that were in the left source, but that were not found in the right source.

Reference: Join Object Options

After a join has been added to the workspace, the following options are available in the shortcut menu:

Expand/Collapse

Opens or closes a window that allows you to view the join expression and sample data.

Join Editor

Opens the Join Editor which allows you to create or edit the join conditions.

Add

Opens a submenu with the following options. Note that these options are only available when the Join object does not already have two inputs.

- **Source.** Opens the Select Synonym panel to add a new synonym to the data flow connected to this Join object. When adding the second synonym the Select Synonym opens with Show synonyms with matching keys filter already enabled.

- **Join.** Adds a new Join object as input to this Join object.

- **Select.** Adds a new Select object as input to this object.

- **Union.** Adds a new Union object as input to this object.

Delete

Deletes the selected Join object.
Properties

Opens the Properties panel for the selected Join object.

Error Info

Displays a dialog box with an error message. This option is only available if there is an error. For more information, click Details.

Reference: Test Individual Join Window

To access the Test Individual Join window from the Join Editor window, select a join and click the Test Join button. Sample data from the join will be shown, as shown in the following image.

Adding a SELECT to a Data Flow

Adding a SELECT to a data flow lets you add a subselect, a select within a select. A subselect, also called a subquery, allows you to use the results of one select as part of another select. While subselects can be used in different places in the SQL language, adding a select to a data flow lets you use a select statement as a source, for example, to join to an aggregated or filtered selection from a table. By using a subselect, the need for intermediate staging tables can be avoided.

Procedure: How to Add a SELECT to a Data Flow

From the data flow workspace:

1. Right-click to the left of the SQL object and click Add Select.
   
   or
   
   Right-click a SQL, Join, or Union object, point to Add, and then click Select.

2. A new SQL object appears in the data flow. Right-click the object, point to Add, and then click another object type.
3. Continue to build the data flow.

**Adding a UNION to a Data Flow**

Adding a union to a data flow lets you combine the results of two select statements into a single result set.

There are some restrictions on the select statements that can be used:

- The two select statements must return the same number of columns.
- The data type of each column in the first table must be the same as the data type of the corresponding column in the second table.
- Neither of the select statements can include a sort (ORDER BY). However, the combined results may be sorted.

**Procedure:** How to Add a UNION to a Data Flow

From the data flow workspace:

1. Right-click to the left of a SQL object and click *Add Union*.
   
   or
   
   Right-click a SQL, Join, or Union object, point to *Add*, and then click *Union*.
   
   or
   
   Drag the *Union* object from the Insert group of the Flow tab, into the workspace.

2. A new Union object appears in the data flow. Right-click the object, select *Add*, and then click *Union*.

3. Continue by adding objects to the select as needed.

4. Right-click the Union object, point to *Add*, and then click *Select* (each union requires two select statements).

5. Continue by adding objects to the second select as needed.

   **Note:** You can test the union by double-clicking the object and clicking the *Test SQL* tab to see a sample of the data.

6. Right-click the Union object and click *Properties*.

7. The Properties pane opens. Select the Union type. See *Union Types* on page 247 for details.
Reference: **Union Types**

The Union object can be used to combine answer sets in different ways, depending on the type selected in the Properties pane. The different union types are:

<table>
<thead>
<tr>
<th>Object</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Union" /></td>
<td><strong>Union</strong>&lt;br&gt;Returns a result that contains all rows from both sets with duplicate rows removed.</td>
</tr>
<tr>
<td><img src="image2" alt="Union All" /></td>
<td><strong>Union All</strong>&lt;br&gt;Returns a result that contains all rows from both sets, including duplicates.</td>
</tr>
<tr>
<td><img src="image3" alt="Except" /></td>
<td><strong>Except</strong>&lt;br&gt;Returns a result set that contains all rows from the first set, excluding common rows from the second set.</td>
</tr>
<tr>
<td><img src="image4" alt="Except All" /></td>
<td><strong>Except All</strong>&lt;br&gt;Returns a result set that contains all rows from the first set, excluding common rows from the second set, including duplicates.</td>
</tr>
<tr>
<td><img src="image5" alt="Intersect" /></td>
<td><strong>Intersect</strong>&lt;br&gt;Returns a result set that contains the common rows from two result sets.</td>
</tr>
<tr>
<td><img src="image6" alt="Intersect All" /></td>
<td><strong>Intersect All</strong>&lt;br&gt;Returns a result set that contains the common rows from two result sets, including duplicates.</td>
</tr>
</tbody>
</table>

**Selecting Columns**

The columns you select define the input to your data target. If you are using the DMC to format a new data target, the order of the columns will reflect the order that the columns are specified in the SQL command.
The DataMigrator Server, which DataMigrator uses to process its extraction logic, can have up to approximately 400 columns in a single data flow. If you intend to move more than 400 columns, several data flows may be linked together by creating a process flow. For information on process flows, see *Designing a Process Flow* on page 385.

**Procedure:** How to Automatically Select All Columns

The default settings of the DMC require you to manually select the columns you want to extract. To change your settings so that all columns are selected automatically:

1. On the Home tab, in the Tools group, click *Options*.
2. Expand *Data Flow* and select *Data Flow Designer* from the tree.
3. Select the *Automatically select all columns* check box in the On Add Source section.
4. Click OK.

**Procedure:** How to Select Columns

1. Right-click the SQL object in the data flow workspace and click *Column Selection*.

   The Column Selection dialog box opens.

2. Double-click the columns you want to select from the Available Columns grid on the left.  
   or  
   Select one or more columns and click the right arrow.

   **Tip:** To select all columns, click anywhere in the column list, and press Ctrl + A.

   The columns appear in the Selected Columns grid on the right.

3. You can change the order of the columns by selecting them and clicking the up and down arrows.
4. Optionally, you can select the *Distinct* check box to remove duplicate rows from the results of your query.
5. You can rename a selected column by entering a name in the Sql Alias field.
6. To remove a column from the Selected Columns grid, select it and click the left arrow, or click the *Delete* button.
7. Click OK.

**Procedure:** How to Edit Selected Columns

1. Right-click the SQL object in the data flow workspace and click *Column Selection*.

   The Column Selection dialog box opens.
2. Select the column you want to edit.

3. Click the *Edit column* button, or double-click the number next to the column in the Selected Columns grid.

   The SQL Calculator opens.

4. Using the Columns/Variables and Functions tabs and the calculator buttons, edit the expression for the column. For information about using functions, see *Using an SQL Function* on page 258. For information about using variables, see *Using Variables in a Flow* on page 336.

5. Click OK.

For information about creating a calculated column in the Selected Columns grid, see *Creating an SQL Calculation* on page 254.

**Procedure:** How to Add or Change Column Descriptions From the Column Selection Dialog Box

If a column in a source synonym has a description, that description will be carried over with the column into any new target synonyms. You can also use the Column Selection dialog box to add descriptions for columns that do not have them, or to change the descriptions for columns that do have them. These descriptions will also be carried over into any new target synonyms in the current flow.

1. From the data flow workspace, double-click the SQL object, or right-click it and click *Column Selection*.

   The Column Selection dialog box opens.

2. Double-click the *Description* field for the column in the Selected Columns grid.

3. Type the description.

4. Click OK.
Reference: Column Selection Dialog Box

To access the Column Selection dialog box from the data flow workspace, double-click the SQL object, or right-click it and click Column Selection.

The Column Selection dialog box contains the following fields and options by default:

**Distinct**
Removes duplicate rows from the results of your query.

**Available Columns**
Lists the columns in each data source. Right-clicking a column header opens a shortcut menu with the following options:

- **Sort Ascending.** Sorts the list of items in the selected column in ascending order.
- **Sort Descending.** Sorts the list of items in the selected column in descending order.
- **Find.** Opens the search dialog box.
- **Field View.** Switches views of the synonyms.
- **List View.** Switches views of the synonyms.
- **Best Fit.** Optimizes the column widths for best viewing.
- **Customize.** Allows you to select the columns that are displayed,
- **Column naming display strategy.** Controls the information that appears in trees and grids. The available options are Name, Title, and Description. If no title or description exists, the display will default to the alias.
Note: When in List View, clicking a column heading will sort the column grid on that heading.

The columns symbols are:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Key column</td>
</tr>
<tr>
<td></td>
<td>Regular column</td>
</tr>
<tr>
<td></td>
<td>Index column</td>
</tr>
<tr>
<td></td>
<td>Virtual column</td>
</tr>
</tbody>
</table>

Note: The symbol for a virtual column in the synonym is blue, while the symbol for a transformation (virtual column in the flow) is yellow.

**Display Name (Title)**

Is the name of the column. Next to each column name, there is a symbol that indicates regular columns, key columns, index columns, virtual columns, and transformations.

**Table**

Is the name of the data source.

**Format**

Is the format of the column.

**Description**

Is the description of the column.

**Nulls**

Indicates whether the column can contain NULL values.

**Right arrow**

Adds columns to the Selected Columns grid.

**Left arrow**

Deletes columns from the Selected Columns grid.
Selected Columns

Lists the columns to be extracted from the data source(s).

Sql Expression

Is the column name, or an expression created in the SQL Calculator.

Sql Alias

Is an alternate name for referencing the column. If you add an alias and expression, the expression shows up where the column name would normally be.

Format

Is the format of the column.

Aggregate

Is the aggregation applied to the column. The available options are: Group By, Group By Not Selected, Sum, Min, Max, Count, Avg, Sum Distinct, Count Distinct, and Avg Distinct.

Description

Is the description of the column.

Nulls

Indicates whether the column can contain NULL values.

Find

Allows you to search for a column.

Insert columns

Opens the SQL Calculator to create a column.

Delete columns

Deletes the selected column.

Edit condition

Opens the SQL Calculator to edit a column.
Test SQL Statement

Tests the SQL statement and displays the result.

Note: The test button will be grayed out if you have not selected any columns.

Up and down arrows

Moves the selected column up or down within the Selected Columns grid.

If you right-click anywhere on the header bar, you get the following options:

Find

Allows you to search for a column.

Best Fit

Resizes the column to fit its data.

Customize

Allows you to select what information is displayed for each column.

If you right-click any column, you get the following options:

Edit

Opens the SQL Calculator where you can edit the properties of a column.

Delete

Deletes the selected column.

Aggregating Columns

You can aggregate your columns in the Column Selection dialog box.

Procedure: How toAggregate a Column

In the Column Selection dialog box:

1. Select the column or columns you want to aggregate. To select more than one column, hold the Shift or Ctrl keys while selecting columns.

2. Click the down arrow in the Aggregate field to display the drop-down menu containing the types of aggregation, and select an aggregation. Your options are:

   - Group By. Groups the data by each distinct value in the column.

   - Group By Not Selected. Groups the data on (or by) the column that you do not want to appear in the results.
- **Sum.** Adds the values within the column.
- **Min.** Returns the minimum value found in the column.
- **Max.** Returns the maximum value found in the column.
- **Count.** Returns a count of all non-null column values.
- **Avg.** Returns the average value for the column.
- **Sum Distinct.** Adds the distinct values within the column, excluding duplicates.
- **Count Distinct.** Returns a count of all distinct, non-null column values, excluding duplicates.
- **Avg Distinct.** Returns the average value for the column, excluding duplicates.

**Note:** If you apply an aggregate function to one of your selected columns, the rest of the selected columns must, in most cases, have aggregation functions applied, or be GroupBy columns. After an aggregation is applied to a column, all columns without an aggregation turn red until an aggregation is applied to them as well.

### Creating an SQL Calculation

You can create an SQL calculation for your data flow that:

- Renames a column.
- Performs a calculation. For details, see *How to Create a Calculation* on page 254.
- Designates a constant value for a column.
- Uses a variable to calculate a column. For details, see *How to Create a Column That Uses a Variable* on page 255.

**Procedure:** *How to Create a Calculation*

1. In the Column Selection dialog box, click the *Insert Columns* button.
   
   The SQL Calculator opens.

2. Enter the calculation you want to perform using the Columns/Variables and Functions tabs, and the calculator buttons. For details, see *SQL Calculator* on page 256. For information about using functions, see *Using an SQL Function* on page 258. For information about using variables, see *Using Variables in a Flow* on page 336.

3. In the Alias box, enter an alias for the column.
4. Click OK.

The calculation appears in the SQL Expressions field.

**Note:** When you create expressions, keep in mind their order of evaluation. For details, see *Order of Evaluation* on page 233.

5. Optionally, select an aggregation for the column in the **Aggregate** field. For details, see *Aggregating Columns* on page 253.

**Note:**

- You can type a calculation directly into the Selected Columns grid by double-clicking the **Sql Expressions** field.
- You can type an alias directly into the Selected Columns grid by double-clicking the **Sql Alias** field.

**Procedure:** **How to Create a Column That Uses a Variable**

From the Column Selection dialog box:

1. Click the **Insert Columns** button.

   The SQL Calculator opens.

2. Type an alias for the virtual column.

3. Double-click the variable you want assigned to the column from the Columns/Variables tab. You can also type in character-valued or date-valued variables, which must be enclosed in single quotation marks (').

4. Click **OK**.

   The variable and alias appear in the Selected Columns grid.

For more information about using variables, see *Using Variables in a Flow* on page 336.
Reference: SQL Calculator

To access the SQL Calculator from the Column Selection dialog box, click the Insert Columns button, or select a column and click the Edit Column button.

The SQL Calculator contains the following fields and options:

**Alias**

Is an alias for the column.

**Expression window**

Displays the expression.

**SQL Optimization button**

Opens the Filter SQL Optimization Report, which allows you to specify settings for, and run, the report.

**Check Expression button**

Verifies the validity of the expression.
Test SQL Statement
Tests the SQL statement and displays the result.

Find button
Allows you to search for a function or variable by name.

Columns/Variables tab
Displays available columns and variables in a hierarchical tree or in a grid.

Functions tab
Displays a list of SQL functions that are available for your transformations.

Function Assist button
Allows you to specify parameters for the function through a dialog box when creating or editing a transformation.

Calculator buttons
Insert numbers and operators.

() 
Adds parentheses.

Inserts two single quotation marks ('). Enter alphanumeric test values between these.

a->A 
Converts selected text to uppercase.

A->a 
Converts selected text to lowercase.

Date 
Opens the Date Editor dialog box, which lets you use the current date or specify a date from the calendar.

Datetime 
Opens the Date Editor dialog box, which lets you use the current date and time or specify a date and time from the calendar.
Reference: SQL Operators

The following table describes the result of an operator on an expression, where `expr` is an expression.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>expr + expr</code></td>
<td>Returns the sum of two expressions.</td>
</tr>
<tr>
<td><code>expr - expr</code></td>
<td>Returns the difference of two expressions.</td>
</tr>
<tr>
<td><code>- expr</code></td>
<td>Returns the negated value of an expression.</td>
</tr>
<tr>
<td><code>expr * expr</code></td>
<td>Returns the product of two expressions.</td>
</tr>
<tr>
<td><code>expr / expr</code></td>
<td>Returns the dividend of two expressions.</td>
</tr>
<tr>
<td>`expr</td>
<td></td>
</tr>
</tbody>
</table>

Using an SQL Function

You can use an SQL function in the SQL Calculator to create a complex expression. In addition, most ANSI 92 SQL functions can be used in a SELECT command, depending on the data source type and platform.

Procedure: How to Use an SQL Function

1. In the SQL Calculator, click the Functions tab.
2. Double-click the folder for the type of function you want to use.
3. Double-click the function you want to use.
   The Function Assist dialog box opens.
4. Use the Function Assist dialog box to set the parameters for the function and click OK.
5. The function appears in the Expressions dialog box. Click OK.
6. The function is added to the Selected Columns grid of the Column Selection dialog box.

For more information about the Function Assist dialog box, see How to Use Function Assist on page 364.
Reference: SQL Functions

You can view the functions in the SQL Calculator Functions tab by category (Tree View) or alphabetically (List View). Right-click the Name header bar to select between the two views, as shown in the following image. For more information about SQL functions, see the SQL function chapters in the Functions Reference manual.

<table>
<thead>
<tr>
<th>SQL Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Character Functions</strong></td>
<td></td>
</tr>
<tr>
<td>CHAR_LENGTH(arg)</td>
<td>Returns the length of arg in a character string.</td>
</tr>
<tr>
<td>CONCAT(arg1, arg2)</td>
<td>Concatenates the values of two arguments.</td>
</tr>
<tr>
<td>EDIT(arg, mask)</td>
<td>Edits a numeric or character value according to a format specified by a mask.</td>
</tr>
<tr>
<td>LCASE(arg)</td>
<td>Converts a character string value to lowercase. Identical to LOWER and LOWERCASE.</td>
</tr>
<tr>
<td>LTRIM(arg)</td>
<td>Removes leading spaces from a character string.</td>
</tr>
<tr>
<td>SQL Function</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td><code>POSITION(substring IN arg)</code></td>
<td>Returns the location of <code>substring</code> in <code>arg</code>, or returns a zero if <code>substring</code> is not found.</td>
</tr>
<tr>
<td><code>REPLACE(string, pattern, [replacement])</code></td>
<td>Replaces the pattern in the pattern in the string with the replacement (if supplied) or removed.</td>
</tr>
<tr>
<td><code>RTRIM(arg)</code></td>
<td>Removes trailing spaces from a character string.</td>
</tr>
<tr>
<td><code>string RLIKE regular_expression</code></td>
<td>Matches a string to a Regular Expression.</td>
</tr>
<tr>
<td><code>SUBSTR(arg FROM start-pos[FOR length])</code></td>
<td>Returns a string from <code>arg</code> starting at <code>start-pos</code> to the end of the string, or if specified <code>length</code> characters.</td>
</tr>
<tr>
<td><code>SUBSTR(arg, start_pos[, length])</code></td>
<td>Returns a string from <code>arg</code> starting at <code>start-pos</code> to the end of the string, or if specified <code>length</code> characters.</td>
</tr>
<tr>
<td><code>TRIM(arg)</code></td>
<td>Removes leading and/or trailing spaces and other characters from a character string.</td>
</tr>
<tr>
<td><code>TRIM(trim_char FROM arg)</code></td>
<td>The character to be removed may be specified as <code>trim-char</code>. If it is not, the space character is assumed.</td>
</tr>
<tr>
<td><code>TRIM(trim_where[trim_char ] FROM arg)</code></td>
<td>Whether to remove leading and/or trailing characters may be specified as <code>trim-where</code>. If not, both leading and trailing spaces are removed.</td>
</tr>
<tr>
<td><code>UCASE(arg)</code></td>
<td>Converts a character string value to uppercase. Identical to UPPER and UPPERCASE.</td>
</tr>
</tbody>
</table>

**Current Date and Time Functions**

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>CURRENT_DATE</code></td>
<td>Returns the date in YYMD format.</td>
</tr>
<tr>
<td><code>CURRENT_TIME[(precision)]</code></td>
<td>Returns the current time of the operating system in the form, HHMMSS.</td>
</tr>
<tr>
<td><code>CURRENT_TIMESTAMP [(precision)]</code></td>
<td>Returns the current timestamp of the operating system (date and time) in the form, YYYYMMDDHHMMSS.</td>
</tr>
<tr>
<td>SQL Function</td>
<td>Description</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>CURRENT_TIMEZONE()</td>
<td>Returns the current time zone.</td>
</tr>
<tr>
<td><strong>Data Type Conversion</strong></td>
<td></td>
</tr>
</tbody>
</table>
| `CAST(expression AS data_type{[length]})` | Converts the value of its argument to a specified data type.  
  *Length* is an optional integer that specifies the length of the target data type. |
| `CHAR(arg[,date_format])` | Converts a number to a character string, or data value to a string in a specified format. |
| `DATE(arg)`            | Returns a date, time, or timestamp value that is computed from *arg*.         |
| `DECIMAL(arg, [length [dec_places]])` | Converts a number into a fixed-length decimal where  
  *dec_places* is the number of decimal places in the result. |
| `DIGITS(arg)`          | Converts a numeric expression to a fixed length character string. Db2 only.  |
| `FLOAT(arg)`           | Converts *arg* to a floating point value.                                     |
| `INT(arg)`             | Converts a number to an integer. Identical to INTEGER.                       |
| `SMALLINT(arg)`        | Converts a number to a small integer.                                         |
| `TIME(arg)`            | Converts the argument to a time.                                              |
| `TIMESTAMP(arg)`       | Converts the argument to a timestamp.                                         |
| `VARGRAPHIC(arg)`      | Converts the argument to a vargraphic.                                        |

**Date and Time Functions**

<table>
<thead>
<tr>
<th>SQL Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>DAY(arg)</code></td>
<td>Returns a number that contains the day of the month from a date or timestamp.</td>
</tr>
<tr>
<td>SQL Function</td>
<td>Description</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>DAYS(arg)</td>
<td>Returns the number of days since December 31, 1900.</td>
</tr>
<tr>
<td>DTIME</td>
<td>Returns the value of all of the time components up to and including the requested component. The remaining time components in the value are set to zero.</td>
</tr>
<tr>
<td>DTRUNC(date_or_timestamp, date_period)</td>
<td>Returns the date, which is the first day of the date period (YEAR, QUARTER, MONTH, DAY, etc.), encompassing the given date or timestamp.</td>
</tr>
<tr>
<td>EXTRACT(field FROM arg)</td>
<td>Extracts a numeric value from a date or time value arg, where field is either a year, month, day, hour, minute, or second.</td>
</tr>
<tr>
<td>HOUR(arg)</td>
<td>Returns the hour field from a time or timestamp value.</td>
</tr>
<tr>
<td>MICROSECOND(arg)</td>
<td>Returns the number of microseconds from a time or timestamp value.</td>
</tr>
<tr>
<td>MILLISECOND(arg)</td>
<td>Returns the number of milliseconds from a time or timestamp value.</td>
</tr>
<tr>
<td>MINUTE(arg)</td>
<td>Returns the number of minutes from a time or timestamp value.</td>
</tr>
<tr>
<td>MONTH(arg)</td>
<td>Returns the number of the month in which arg falls.</td>
</tr>
<tr>
<td>SECOND(arg)</td>
<td>Returns the second field from a time or timestamp value.</td>
</tr>
<tr>
<td>YEAR(arg)</td>
<td>Returns the year in which arg falls.</td>
</tr>
</tbody>
</table>

**Miscellaneous Functions**

<p>| COUNTBY(arg) | Produces a column whose values are incremented row-by-row by a specified amount. |</p>
<table>
<thead>
<tr>
<th>SQL Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>GREATEST(arg1, arg2[,...arn])</code></td>
<td>Finds the maximum value.</td>
</tr>
<tr>
<td><code>HEX(arg)</code></td>
<td>Converts the input value to hexadecimal.</td>
</tr>
<tr>
<td></td>
<td><strong>Note:</strong> This function is available only for Db2, Ingres, and Informix.</td>
</tr>
<tr>
<td><code>IF(test, val1, val2)</code></td>
<td>Tests a condition and returns a value based on whether the condition is true or false.</td>
</tr>
<tr>
<td><code>LEAST(arg1, arg2[,...arn])</code></td>
<td>Finds the minimum value.</td>
</tr>
<tr>
<td><code>LENGTH(arg)</code></td>
<td>Returns the number of bytes of storage used by <code>arg</code>.</td>
</tr>
<tr>
<td><code>USER()</code></td>
<td>Returns the user ID running the query.</td>
</tr>
</tbody>
</table>

### Numeric Functions

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ABS(arg)</code></td>
<td>Returns the absolute value of a number.</td>
</tr>
<tr>
<td><code>CEIL(number)</code></td>
<td>Returns the smallest integer that is greater than or equal to the number.</td>
</tr>
<tr>
<td><code>EXP(arg)</code></td>
<td>Returns e raised to the power of the argument.</td>
</tr>
<tr>
<td><code>FLOOR(number)</code></td>
<td>Returns the largest integer that is less than or equal to the number.</td>
</tr>
<tr>
<td><code>LOG(arg)</code></td>
<td>Returns the natural logarithm of the input value.</td>
</tr>
<tr>
<td><code>MOD(dividend, divisor)</code></td>
<td>Calculates the remainder from a division as an integer.</td>
</tr>
<tr>
<td><code>POWER(arg1, arg2)</code></td>
<td>Returns the first number raised to the power of the second.</td>
</tr>
<tr>
<td><code>SQRT(arg)</code></td>
<td>Returns the square root of the input value.</td>
</tr>
</tbody>
</table>

### SQL Operators
<table>
<thead>
<tr>
<th>SQL Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CASE</strong> test-expr</td>
<td>Allows a value to be computed depending on the values of expressions.</td>
</tr>
<tr>
<td>WHEN value-expr-1 THEN result-expr-1</td>
<td></td>
</tr>
<tr>
<td>. . .</td>
<td></td>
</tr>
<tr>
<td>WHEN value-expr-n THEN result-expr-n</td>
<td></td>
</tr>
<tr>
<td>[ ELSE else-expr ] END</td>
<td></td>
</tr>
<tr>
<td><strong>CASE</strong></td>
<td>Allows a value to be computed depending on the truth or falsity of conditions.</td>
</tr>
<tr>
<td>WHEN cond-1 THEN result-expr-1</td>
<td></td>
</tr>
<tr>
<td>. . .</td>
<td></td>
</tr>
<tr>
<td>WHEN cond-n THEN result-expr-n</td>
<td></td>
</tr>
<tr>
<td>[ ELSE else-expr ] END</td>
<td></td>
</tr>
<tr>
<td><strong>COALESCE</strong>(arg1, arg2, [ ... argn ])</td>
<td>Can take two or more arguments. The first argument that is not NULL, is returned. If all arguments are NULL, NULL is returned.</td>
</tr>
<tr>
<td><strong>EXISTS</strong>(SELECT * FROM LOOKUP_MFD SQ [WHERE condition] )</td>
<td>Tests if subquery returns one or more rows.</td>
</tr>
<tr>
<td><strong>NULLIF</strong>(arg1, arg2)</td>
<td>Returns NULL if its two arguments are equal. Otherwise, the first argument is returned.</td>
</tr>
<tr>
<td><strong>test_expr</strong> IN (exp1, exp2)</td>
<td>Determines whether a specified value matches any value in a list.</td>
</tr>
<tr>
<td>(SELECT lookup_result from LOOKUP_MFD [WHERE condition])</td>
<td>Used in a select list to return a column value.</td>
</tr>
<tr>
<td>(SELECT * FROM LOOKUP_MFD SQ [WHERE condition] )</td>
<td>Returns result set based on subquery.</td>
</tr>
</tbody>
</table>
### Testing the SQL Statement

Once you have created your SQL calculation, you should test it to see if you are selecting the data you need. The SQL object enables you to retrieve a limited set of rows and review them as part of the process of designing your data flow.

**Note:** You can set the number of rows retrieved from the Options dialog box in the Run Options pane. The default is 50 rows.

You can also edit the SQL code.

**Tip:** You can move SQL SELECT statements between applications by copying and pasting.

#### Procedure: How to Test the SQL

In the data flow workspace:

Double-click the SQL object and click the Test SQL Statement button in the upper-right corner of the Column Selection dialog box.

The Test SQL Statement Columns dialog box opens, which shows the sample data for the SQL object, as shown in the following image.

![Test SQL Statement Columns dialog box](image)

**Note:** You can also test the SQL by right-clicking the SQL Object and clicking Expand. Click the Test SQL tab to see a sample of the data.
Procedure: How to Test the SQL From the SQL Statement

1. Right-click the SQL object and click SQL Statement.

   The Select Statement dialog box opens, as shown in the following image.

   ![Select Statement Dialog Box](image-url)

   The Select Statement dialog box has the following buttons:

   - **Find**: Searches for a text string in the SELECT statement.
   - **Test SQL Statement**: Tests the SELECT statement.

   For more information on the shortcut menu options in this dialog box, see *Text Editor Shortcut Menu* on page 148.

2. Click the Test SQL Statement button in the upper-right corner.

   The Test SQL Statement dialog box shows a sample of the data.

   **Note:** You can also test the SQL by right-clicking the SQL object, selecting *Column Selection*, and clicking the Test SQL Statement button in the Column Selection dialog box.

   There is also a Test SQL Statement button in the Filter Calculator and the Sort Columns dialog box.
**Procedure:** How to Edit the SQL Statement

You can edit the generated SQL manually. If you do, the interface may redraw the SQL object based on the revised SQL when it is reloaded. If the interface is unable to redraw the SQL object, the SQL object will be labeled Custom SQL.

1. In the data flow workspace, right-click the SQL object and click SQL Statement.
   
   The Select Statement dialog box opens.

2. Make your changes to the SQL statement.

3. Click OK to save your changes.

**Note:** The SQL command must conform to the SQL supported by the iWay server. For more information on using a dialect SQL specific to a relational database system, see Creating a DBMS SQL Flow on page 379.

**Specifying Filters in a Data Flow**

You can create a filter to control the data included in your data flow. A filter creates WHERE criteria in the SQL of your flow. You can filter your data with a constant value or a variable.

**Note:** When a filter is saved, it is parsed and any superfluous parentheses are removed.

**Procedure:** How to Add a Filter to a Data Flow

From the data flow workspace:

1. Right-click the SQL object and click WHERE Filter.
   
   The WHERE Filter Calculator opens.

2. Enter an expression in the Conditions window using the Columns/Variables and Functions tabs and the calculator buttons. For information about using functions, see Using an SQL Function on page 258. For information about using variables, see Using Variables in a Flow on page 336.

3. If you create multiple expressions that must all be satisfied, separate them with AND.

4. If you create multiple expressions where any of the expressions must all be satisfied, separate them with OR.

5. Click OK.

**Procedure:** How to Create a Data Flow Filter That Uses a Date/Time Constant Value

From the data flow workspace:

1. Right-click the SQL object and click WHERE Filter.
   
   The WHERE Filter Calculator opens.
**Note:** If the Relational Expression tab is not selected, click on it to select it.

2. Double-click a column from the Columns/Variables tab. You can also type the column name in the Conditions field.

   The column is added to the Conditions field.

3. Use the calculator buttons to select a relationship.

4. Double-click the constant you want from the Date and Time Constants folder of the Columns/Variables tab, and replace the sample constant with the desired date/time. Character and date values must be enclosed in single quotation marks ("), as shown in the following image.

   ![Conditions Tab Example](image-url)

5. Click OK.

**Procedure:**  **How to Create a Data Flow Filter That Uses a Constant Value**

From the data flow workspace:

1. Right-click the SQL object and click *WHERE Filter*.

   The WHERE Filter Calculator opens.

   **Note:** If the Relational Expression tab is not selected, click on it to select it.

2. Double-click a column from the Columns/Variables tab. You can also type the column name in the Conditions field.

   The column is added to the Conditions field.

3. Use the calculator buttons to select a relationship.
4. Type in the constant value. Character and date values must be enclosed in single quotation marks (').

5. Click OK.

Procedure: How to Create a Data Flow Filter That Uses a Variable

From the data flow workspace:

1. Right-click the SQL object and click *WHERE Filter*.
   
   The WHERE Filter Calculator opens.

2. Double-click a column from the Columns/Variables tab. You can also type the column name in the Conditions field.

   The column is added to the Conditions field.

3. Use the calculator buttons to select a relationship.

4. Double-click the variable you want from the DM or System Variables folder on the Columns/Variables tab. You can also type in character-valued or date-valued variables, which must be enclosed in single quotation marks (').
5. Click OK.

For more information about variables, see *Using Variables in a Flow* on page 336.

**Procedure:** How to Create a Filter That Uses a Sub-Select

A filter that uses a sub-select can be typed into the filter calculator or created using the Function Assist.

1. Drag *ibisamp/dmord* into a data flow.
2. Right-click the SQL object and click *WHERE Filter*.

   The WHERE Filter Calculator opens.

   **Note:** If the Conditions tab is not selected, click on it to select it.

3. Click the Functions tab, and expand *SQL Operators*.
4. Select the *IN ... SELECT* function.
5. A prototype condition is filled in:

   
   
   (test_exp1, test_exp2) IN (SELECT exp1, exp2 FROM mfd WHERE condition)

6. Replace the text with the desired condition. For example, select rows where the value of PLANT is found in DMSALE.

   
   T1.PLANT IN (SELECT T2.PLANT FROM DMSALE T2 )

   You can also select rows where the combination of values of PLANT and ORDER_NUM are found in DMSALE.

   
   (T1.ORDER_NUM ,T1.PLANT ) IN (SELECT T2.ORDER_NUM , T2.PLANT FROM DMSALE T2 )

**Reference:** WHERE Filter Calculator

To access the WHERE Filter Calculator from the data flow workspace, right-click the SQL object and click *WHERE Filter.*
The calculator opens, as shown in the following image.

The WHERE Filter Calculator contains the following fields and options. For a full description of the operators available, see Join Calculator on page 240.

The following options are available if you click the Conditions tab:

**Conditions window**

Displays the expression.

**Function Assist button**

Allows you to specify parameters for the function through a dialog box when creating or editing a transformation.

**Calculator buttons**

Insert numbers and operators.

( )

Adds parentheses.

“”

Inserts two single quotation marks. Enter alphanumeric test values between these.
Specifying Filters in a Data Flow

a->A
Converting selected text to uppercase.

A->a
Converting selected text to lowercase.

Date
Opens the Date Editor window, which lets you use the current date or specify a date from the calendar.

Datetime
Opens the Date Editor window, which lets you use the current date and time or specify a date and time from the calendar.

SQL buttons
The following SQL buttons are available:

- BETWEEN
- IN
- IS NULL
- LIKE
- AND
- OR
- ANY
- NOT
- ALL

Functions tab
Displays a list of SQL functions that are available for your transformations.

The following options are available if you click the Relational Expression tab:

Relational Expression window
Displays the Relational Expressions builder.
Test SQL Statement

Tests the SQL statement and displays the result.

**Note:** The test button will be grayed out if you have not selected any columns.

**Columns/Variables tab**

Displays available columns and variables in a hierarchical tree view or in a list view.

**Columns**

When displayed in a tree view, expanding the column name shows a list of values.

For a synonym for a data source in a table or file, a query is done to the data source to get a list of unique values. For large data sources, this could take a while.

A list of acceptable values is retrieved from the synonym if the synonym is for any of the following: database stored procedure, web service, application adapter, any other procedure based adapter.

**Specifying Sorts**

You can specify a sort to determine the order in which new columns appear in the results. A sort orders your results according to the results of a column you choose. When sorting by multiple columns, the first sort field sets the major sort break, the second sets the second sort break, and so on.

A sort is useful only in some cases. Use the following guidelines to consider whether or not to include a sort in your data flow.

- **FOCUS/FDS or XFOCUS data target.** A sort is useful for speeding up the process of loading the data target. Since loading without a sort will cause the process to take more time, you will receive a warning message.

- **Flat file.** A sort is useful if you are loading your data into an application that respects sorting. If it does not, a sort is not useful with a flat file.

- **Relational data target.** A sort is generally not useful because it is not respected when loading a relational data target.

**Procedure: How to Specify a Sort**

1. From the data flow workspace, right-click the SQL object and click **Sort**.

   The Sort Columns dialog box opens.
2. Double-click the column you want to sort on.

   or

   Select the column you want to sort on and click the right arrow.

   The column is added to the Sort Columns box.

3. Optionally, change the order in which the results will be sorted by clicking the drop-down menu under the Order column and selecting an order.

   - *ASC (ASCENDING)* sorts the results from lowest to highest. This is the default.
   - *DESC (DESCENDING)* sorts the results from highest to lowest.

4. Optionally, if you select more than one column, you can change the dominant sort by using the up and down arrows to move the columns up and down in the list.

5. Click OK.

   **Note:** Clicking a column heading will sort the column grid on that heading.

**Reference:**  Sort Columns Dialog Box

To access the Sort Columns dialog box from the data flow workspace, right-click the SQL object and click Sort.
The Sort Columns dialog box contains the following fields and options:

**Available Columns**
Is a list of columns available in the data source.

**Source Column**
Is the name of the column.

**Format**
Is the format of the column.

**Description**
Is a description of the column.

*Note:* Clicking a column heading will sort the column grid on that heading.

**Nulls**
Indicates if the column can contain NULL values.

**Sort Columns**
Is a list of sort columns.

**Order By**
Is the column(s) that sorts the results.

**Order**
Specifies whether results are sorted in ascending or descending order.

**Aggregate**
Is the column aggregation, if one has been applied.

**Right and Left arrows**
Add or remove sort columns.

**Find**
Allows you to search for a column.

**Test SQL Statement**
Tests the SQL statement and displays the result.

*Note:* The test button will be grayed out if you have not selected any columns.
Selecting a Data Target

You need to specify the data target(s) to which your data will be copied. You can choose to use an existing table as your target, or create a new target.

- For details on creating a new target, see How to Create a New Data Target on page 277.

  Note: For new target tables, column descriptions are included in the synonym. The descriptions may come from the source table or from a source transform.

- For details on using an existing data target, see How to Select an Existing Data Target on page 283.

A data target can be a relational data target, non-relational data target, or a flat file. In some cases, relational targets can provide the ability to write to remote tables. For more information, see Reviewing Your Data Targets on page 115.

Note: This is only supported for existing targets. The name of the remote server must match the real server name of the subserver. This option should only be used for small data volumes. For large volumes, you should install the DataMigrator Server on the platform where the data will be loaded.

You can load your data into as many as fifteen tables in a single data flow.

After selecting a data target, you can do the following:

- Specify target options. For details, see Specifying Data Target Options on page 283.

- Specify mappings for existing tables. For details, see Mapping Transformations on page 324.

Reference: Data Target Types

The following data target types can be used:
Relational data targets may include the following formats:

- Access
- Db2/UDB
- EDA (for existing targets only)
- Greenplum
- Hive
- Hyperstage
- Informix
- MySQL
- Nucleus
- ODBC
- Oracle
- Postgres
- Red Brick
- SQL Server
- Sybase
- Teradata

Non-relational data targets may include the following formats:

- FOCUS/FDS
- XFOCUS

Other data targets may include the following:

- Delimited flat file (delimiter can be specified as a character string or hex digit)
- XML or JSON file
- Flat file
- Formatted file
- Hive Impala Phoenix/HBase MongoDB

Procedure: How to Create a New Data Target

1. Right-click in the space to the right of the SQL object, point to Add Target, and then click New.
   
   A target object appears in the workspace.

2. Right-click the new target and click Properties.
   
   The attributes of the target open in a new pane.

3. Select an Adapter from the drop-down menu to specify the target type.

   Note: Only data targets types that are configured on the server and supported by DataMigrator will appear.
4. Select a connection from the drop-down menu.
5. Select an Application Directory for the synonym that will be created for the target.
6. Optionally, enter a synonym name, which is the name of the data target.
   
   If you have chosen a FOCUS/FDS, XFOCUS, flat file, delimited flat file, or formatted file, when you press the Enter key, the data file name will automatically change to match the synonym name. This can be changed.

   **Note:** If you do not provide a synonym name, it will default to target01 for the first new target, target02 for the second, and so on.

   New targets will automatically contain the same columns that were selected in the SQL object. For information about managing columns, see *Selecting Columns* on page 247.

   **Note:** If the columns selected from the source have descriptions, these descriptions will be included with the corresponding columns in a new target.

   For relational database targets, you can specify connection (if applicable), application directory, synonym, and table name. For details, see *How to Set Options for Relational Targets* on page 312.

   For FOCUS/FDS or XFOCUS files, you can specify synonym application directory, synonym, data application directory, and dataset (name of file). For details, see *How to Set Options for FOCUS/FDS or XFOCUS Targets* on page 315. Synonym names for FOCUS/FDS files cannot be longer than eight characters.

   For flat files, you can specify synonym application directory, synonym, data application directory, and dataset (name of flat file). For formatted files, you can specify format, synonym application directory, synonym, data application directory, and dataset (name of formatted file). For details, see *How to Set Options for Flat, Delimited Flat, Formatted, and XML File Targets* on page 322.

   **Note:** To create a file in HFS (hierarchical file system) on IBM z/OS, enter the name. To create a dataset enter the name as

   ```
   //'qualif.tablename.type'
   ```

   where:

   - **qualif**
     
     Is the fully qualified location.

   - **tablename**
     
     Should match the synonym name.
type

Is DATA or FOCUS, by convention.

7. Set the load options. For more information, see Specifying Data Target Options on page 283.
8. Click OK.
9. Right-click the target and click Target Transformations.
10. Select the key column(s) in the Target Columns grid.
11. Click OK.

Procedure: How to Create a New Data Target With a Variable Name

When creating a new target table, the name of the synonym that describes that table is normally entered as text and saved with the flow. This procedure should be used when you desire to specify the name of the target table dynamically when the flow is run.

1. Right-click in the space to the right of the SQL object, point to Add Target, and then click New.
   
   A target object appears in the workspace.

2. Right-click the new target and click Properties.
   
   The attributes of the target open in a new pane.

3. Select an Adapter from the drop-down menu to specify the target type.
   
   Note: Only data target types that are configured on the server and supported by DataMigrator will appear.

4. For the synonym name, enter a name starting with an ampersand (&) to indicate that it is a variable.
The Properties pane for a MS SQL Server table target is shown in the following image.
The Properties pane for a Delimited Flat File target is shown in the following image.

5. By default, the variable name you entered is automatically used as the name for the Table (relational database target) or Data File (flat file, delimited file, xml target). If needed, another name can be used.

6. To test the data flow, click the Run button and click Run.

A dialog box opens requesting a value for the variable.

7. Click Run to run the flow.

Procedure: How to Submit or Schedule a Data Flow With a Variable Name

To submit a data flow with a variable name, or to schedule such a flow, the value for the variable must be assigned. This can be done with a process flow that calls the data flow.

1. Right-click an application directory in the navigation pane, point to New, and then click Flow.

The data flow workspace opens.

2. Click the Process Flow tab to open the process flow workspace.

3. Drag the data flow with the variable target name into the process flow workspace.
Right-click the Start icon, drag a line connecting to the data flow icon, and release. Then double-click the data flow icon to open its properties, as shown in the following image.

4. Click the ellipsis in the Parameters box.
   The Parameter Editor opens.

5. Click the Add Parameter button.
   The Drill Down Parameter Editor dialog box opens, as shown in the following image.
6. Select the name of the variable from the Parameter Name drop-down menu and enter a name for the table. Click OK twice to close the Parameter Editor dialog box. The parameter now appears in the properties window, as shown in the following image.

7. Click the Save button to save the data flow. The process flow can now be submitted, or scheduled, with the specified table name.

**Procedure: How to Select an Existing Data Target**

In the DMC, drag a synonym object from the navigation pane to the right of the SQL object in the workspace.

or

Right-click in the workspace to the right of the SQL object, point to Add Target, and then click Existing. Open an Application Directory, select an object, and click Select.

The data target is automatically connected to the SQL object.

**Procedure: How to Change a Data Target in a Data Flow**

With the data flow open in the workspace:

1. Right-click the target object that you want to change and click Change Target.

   The Select Source dialog box opens.

2. Select the replacement synonym and click the Select button.

**Specifying Data Target Options**

After the data targets have been added to the flow, there are two types of options that govern the behavior of the data flow when copying data into the targets.

- **Data flow properties**, which are set for the entire flow and, therefore, effect every data target. For convenience these options are displayed with each target. They are marked with an asterisk in the Target Properties dialog box to indicate that they are set at the flow level.
Target-specific properties, which can be set on individual target objects.

Data flow properties include:

- **Load type**, which specifies the method DataMigrator uses to load the data. The options for load types depend on the target database type and may include: Insert/Update, Insert records from memory, IUD processing, Slowly Changing Dimension, Bulk load utility via disk file, and Extended Bulk Load Utility. If you are moving data to more than one data target, they will all use the same Load type. It will be displayed on all target objects and, if changed in one target, the change will be reflected in all the target objects. However, if you choose Insert/Update for the Load type, targets can have different settings for If the record exists/does not exist.

- **Record logging**, which allows you to create log files based on specific criteria. Each logging option can write information to a separate flat file. For example, if you choose Duplicate Rejection as a Record Logging option, a flat file called `procedurename_dupl.ftm` is created where duplicate records that are rejected will be recorded. You can use the log file as input for other processes. For details, see How to Set the Data Flow Record Logging Options on page 307.

- **Commit**, which specifies the number of rows to process before issuing a commit to a relational database target, or CHECK to a FOCUS or XFOCUS target. This option does not apply to file targets.

  A commit value of 0 indicates that a commit is only issued when the data flow completes.

  **Note:** A commit value of NONE can be used to indicate that DataMigrator should not issue any commit. This may be useful in cases where two or more flows should complete successfully before any data is committed to the target tables. A commit (or rollback) should be issued in a stored procedure after the data flows are complete. This option should be used with caution because if no commit or rollback is issued, the target table can become locked.

- **Blocksize**, which is the number of rows for loading in a block. This option only appears when insert records from memory is selected.

Target-specific properties include:

- **General Properties**, such as:

  - **Display Name**. The name that appears on the data flow under the target icon. By default, the synonym name and location (optionally) appear.
Notes. User entered notes. These notes only appear in the properties panel and now the data flow itself. To make the notes visible in the data flow, select the Show Notes option in the Data Flow Designer Options dialog box.

Prior to Load Options, which can be set for existing targets to clear the target before the flow runs. The available options will depend on the type of data target.

If the record exists/does not exist options, which specify behavior for Insert/Update.

Target Load Options

Load options specify the way DataMigrator loads data into the targets. If you are moving data to more than one target, they will all use the same load type. The load type is displayed on all target objects and, if changed in one target, the change will be reflected in all the target objects.

Load options vary depending on the type of target. They can include:

Insert/Update. Allows you to set a behavior when loading records.

If you select this option, you can set a behavior when duplicate records are found from the If the record exists drop-down menu. Insert/Update options can be set on a per target basis.

If the record exists:

Include the record. Allows the relational database target to handle duplicates.

Reject the record. Issues a SELECT command against the target table to see if a record exists with the key values. If there are no key columns, DataMigrator screens all columns. If the record is found, the record is rejected.

Update the existing record. Updates records if the key value on the input record, or entire record if there are no keys, are found in the table. All non-key values are updated in the target.

Delete the existing record. Deletes the record if the key on the input record, or entire record if no keys, is found in the table.

If the record does not exist:

Include the record. Includes the key in the data target.

Reject the record. Does not include the key in the data target. This option is only available if Update or Delete the existing record was selected.
Note: When loading relational data targets, the Reject, Update, and Delete options can adversely affect performance, because they determine the existence of a key value on an incoming record before performing the specified action. This is done by issuing a SELECT command against the target table, and then waiting for the response. If one of these actions is required, try to limit the volume of records in the incremental change. These actions will perform best if there are unique keys on the table.

If you select Update the existing record or Delete the existing record from the If the record exists drop-down menu, you can also set behavior when the record does not exist using the If the record does not exist drop-down menu:

- **Commit every row(s).** Specifies the row interval to commit or write transactions to the database.
- **Insert records from memory.** Speeds the loading of the data target by inserting a block of rows at once. You can set the row interval to commit or write transactions, and the number of records to load in a block.

**Note:** This option is only available for Db2/UDB CLI, MySQL, MS SQL Server, Oracle, Informix, Sybase ASE, and Teradata 13. For other databases, the number of records in a block will default to 1 and cannot be changed. For details about loading relational targets, see *How to Set Options for Relational Targets* on page 312, and *Target Properties Pane for Relational Targets* on page 309.

Insert records from memory:

- Loads much faster, since a block of rows are inserted at once.
- Requires input data to be clean. If any one row in the block causes a data source constraint violation, such as not null or unique index, the entire block is rejected for an MS SQL Server target. For other targets it depends on the database.
- Does not provide row counts (the number of records inserted or rejected) in the detail log or statistics.
- Does not provide record logging (used to write rejected records to a flat file for review).

**Bulk load utility via a disk file.** Uses database bulk loaders instead of iWay to insert data into a target. DataMigrator automates bulk loading for Greenplum, Hyperstage, Ingres, Informix, Microsoft SQL Server, IBM Db2, MySQL, Teradata, Nucleus, Oracle, Sybase Adaptive Server Enterprise, and Sybase Adaptive Server IQ. You can set format version, maximum number of errors, first and last row to copy, packet size, and row interval to commit or write transactions.

The options available will depend on the RDBMS of your target type.
Note that rows rejected by a DBMS vendor bulk load program are not counted as errors. Users of the bulk loader should look into the vendor specific documentation on how to handle the error handling of the bulk loader program.

**Note:**

- To use a database vendor bulk load program, such as `bcp` from MS SQL Server or Sybase, `sqlload` from Oracle, or `mysql` from MySQL, that program must be installed on the system where the DataMigrator Server is running and in the PATH.
- Does not provide row counts (the number of records inserted or rejected) in the detail log or statistics.

- **Extended Bulk Load Utility.** Uses a database loader program with additional capabilities. Currently available for Apache Hive, Cloudera Impala, Google BigQuery, EXASol, Greenplum, Hyperstage, Informix, Jethro, MongoDB, MS SQL Server ODBC, MySQL, Netezza, Oracle, PostgreSQL, Sybase ASE, Sybase IQ, Teradata, and 1010data.

The options available will depend on the RDBMS of your target type. For more information on the options available for each adapter, see *Extended Bulk Load Utility Options* on page 288.

**Note:** Does not provide row counts (the number of records inserted or rejected) in the detail log or statistics.

- **IUD Processing** enables you to load a data target with only the records that have changed. This feature is an optional, add-on component to DataMigrator available for selected relational databases.

  **Note:** This setting is not available if the Optimize Load option is enabled in the Data Flow properties.

- **Slowly Changing Dimensions.** Enables you to load a data target with column values that change over time. There are two column types. Type I simply overwrites the value. Type II allows you to track activity before and after the change. Type II changes are handled either using an activation flag or end date/begin date. When you update a Type II column, you do not actually change its value. Instead, you update the activation flag to inactive, or the end date to the current date. You then add a new record with the new value and the activation flag on by default or the begin date set.

  **Note:** This setting is not available if the Optimize Load option is enabled in the Data Flow properties.

- **IUD/Slowly Changing Dimensions** is a combination of the previous two load types. The Insert/Update/Delete indications in the CDC logs are not applied directly to the target. Instead, the SCD Types of the target are used.
Note: This setting is not available if the Optimize Load option is enabled in the Data Flow properties.

Extended Bulk Load Utility Options

This option uses a database loader program with additional capabilities and is currently available for specific adapters. When extended bulk load is used for large data volumes, faster throughput is achieved by creating multiple load sessions (or threads). It reads from the source in parallel, a block at a time, where the block size is set by the Commit every rows parameter. It also creates multiple sessions for loading if the database load utility supports that functionality, up to the number specified in Maximum number of load sessions parameter.

Note: The options available will depend on the RDBMS of your target type.

An example of the Extended Bulk Load Utility options for MS SQL Server can be seen in the following image:

The following option is available for every adapter that has an Extended Bulk Load Utility load type.

Commit every row(s)

Number of rows in a block. Use at least the default value of 1,000,000.
The following parameters depend on the database:

**Maximum Number of Load Sessions**

The maximum number of the concurrent load sessions logged on to the database. This parameter is applicable when the BULKLOAD parameter is set to ON. Its setting can be defined on the flow level and also on the Adapter level. The level setting of the adapter applies to the designs of the new Data Flows. The value set on a flow level takes precedence during the execution.

The default value is set to 10. However, the user should set the value carefully and consider current environment factors, such as the number of CPU cores, available memory, the parameters on the data access agents on a server, network capabilities, and the number of users running the requests.

**Maximum Number of Load Session Restarts**

The maximum number of restart attempts after recoverable load session error.

The above three options are the only Extended Bulk Load options available for the following adapters:

- EXASol
- Hyperstage
- MySQL
- MariaDB
- Netezza
- Oracle

**Note:**

- The Extended Bulk Load Utility option must not be used on a target synonym that is for an Oracle table containing a Sequence. This is due to the rejection of triggers by the Oracle API used in the Extended Bulk Load Utility.

- The "Bulk thread postponed: too many connections" information message means that the maximum amount of loading threads was already started, so the new available intermediate load file will be queued and started when one of the loading threads become available. This situation occurs when either the Maximum number of Sessions value is too low or Commit value is too low.

The following sections describe the additional options that are available for each adapter.
Reference: Apache Hive and Cloudera Impala

Apache Hive, Cloudera Impala, and Salesforce.com

Extended Bulk Load Options

Apache Hive and Cloudera Impala targets only have the default Extended Bulk Load Utility option, as shown in the following image.

Note: For the best performance when loading Apache Hive or Cloudera Impala, it is recommended that the DataMigrator Server be located on a data node or Edge Node (with the Hadoop client installed) of your Hadoop cluster.

However, it is possible to use DataMigrator to load data into Hadoop, even when the servers are not co-located. In this case, the DataMigrator Server will create a temporary file and use FTP to copy it to the Hadoop node. This requires that an FTP server is available on the Hadoop cluster, and in addition for use with Impala REXEC daemon.

Example: Configuring an Auxiliary Connection for Extended Bulk Load

Using Extended Bulk Load with several adapters requires that intermediate data files are transferred to the server where the database is running or a data node of a Hadoop cluster. These adapters include:

- Apache Hive
- Apache Drill
- Cloudera Impala
- EXASol
- Jethro
For these adapters, you must enable the Flat File adapter and create a connection using FTP or SFTP to the server. From the DMC or Web Console, perform the following steps:

1. Expand the **Adapters** folder or click the **Adapters** tab.
2. Expand the **Configured** folder.
3. Right-click **Flat File** and click **Add Connection**.
4. In the **Connection Name** box, type the name of the connection for your Hive, Impala, or database server. The connection names must match.
5. From the Protocol drop-down menu, select **FTP** or **SFTP**.
6. In the **Host Name** box, type the name of the host for your server.
   - **Note:** For Hive and Drill this must be a data node of your Hadoop cluster. For Impala, it must be a Data node or an Edge node. For Jethro, it must be the system where the Jethro server is running. For EXASol it should not be the system where it is running, rather it should be some other system that it can access using FTP.
7. In the **Initial Directory** box, type the name of a directory that you have write access to, such as your home directory.
8. In the **User** and **Password** boxes, type your user ID and password.
9. Click **Test**.
   - The following message should appear: *User has read access to the FTP server.*
   - **Note:** Although not tested at this time, you must also have write access. A future release will add testing for write access.
10. Click **OK** to close the dialog box.
11. If successful, click **Configure**.
12. Click **Close**.

**Reference:** Greenplum Extended Bulk Load Options
Greenplum targets have the following additional options, as shown in the following image.

**Column Delimiter**

Delimiter characters between fields in the intermediate flat file that is loaded into the database. The delimiter can be up to four characters and can be specified as the following:

- **TAB.** A tab character. This value is the default.
- **a - A.** A character string. For example, ~.
- **0x nn.** A hex code. For example, 0x44 (a comma), or 0x0D0A (a return and a linefeed). The hex code uses ASCII for Windows or UNIX systems and EBCDIC for IBM Mainframes.

**Escape Character**

Specifies the single character that is used for escape sequences, such as ,, or @;. Also used for escaping data characters that might otherwise be interpreted as row or column delimiters.

**Column Enclosure**

Characters used to enclose each alphanumeric value in the intermediate flat file that is loaded into the database. The enclosure consists of up to four printable characters. The most common enclosure is one double quotation mark. Numeric digits, symbols used in numbers, such as a period (.), plus sign (+), or minus sign (-), and 0x09 (Horizontal Tab), 0x0A (Line Feed), or 0x0D (Carriage return) cannot be used in the enclosure sequence. In order to specify a single quotation mark as the enclosure character, you must enter four consecutive single quotation marks.
**Note:** When a non-default value is used, the intermediate file is created as the Greenplum term CSV, instead of what is called TEXT.

**Global Settings**

These options are only available when BULKLOAD=ON and apply to all new flows using Greenplum as the target.

**Note:** ALIAS_CASE controls the case of column names when new tables are created. Greenplum users may want to select *Enforce lower case*.

**Reference:** Informix Extended Bulk Load Options

Informix targets have the following additional options, as shown in the following image.

**Column Delimiter**

Delimiter characters between fields in the intermediate flat file that is loaded into the database. The delimiter can be up to four characters and can be specified as the following:

- **TAB.** A tab character. This value is the default.
- **a - A.** A character string. For example, ~.
- **0x nn.** A hex code. For example, 0x44 (a comma), or 0x0D0A (a return and a linefeed). The hex code uses ASCII for Windows or UNIX systems and EBCDIC for IBM Mainframes.
Reference: Jethro Extended Bulk Load Options

Jethro targets have the following additional options, as shown in the following image.

Column Delimiter

Delimiter characters between fields in the intermediate flat file that is loaded into the database. The delimiter can be up to four characters and can be specified as the following:

- **TAB.** A tab character. This value is the default.
- **a - A.** A character string. For example, ~.
- **0x nn.** A hex code. For example, 0x44 (a comma), or 0x0D0A (a return and a linefeed).
  The hex code uses ASCII for Windows or UNIX systems and EBCDIC for IBM Mainframes.

Reference: MS SQL Server ODBC Extended Bulk Load Options
MS SQL Server ODBC targets have the following additional options, as shown in the following image.

**Packet size to send**

The packet size in bytes.

*Reference:* PostgreSQL Extended Bulk Load Options
PostgreSQL targets have the following additional options, as shown in the following image.

![Properties](image)

**Column Delimiter**

Delimiter characters between fields in the intermediate flat file that is loaded into the database. The delimiter can be up to four characters and can be specified as the following:

- **TAB.** A tab character. This value is the default.
- **a - A.** A character string. For example, ~.
- **0x nn.** A hex code. For example, 0x44 (a comma), or 0x0D0A (a return and a linefeed). The hex code uses ASCII for Windows or UNIX systems and EBCDIC for IBM Mainframes.

**Escape Character**

Specifies the single character that is used for escape sequences, such as ,, or @;. Also used for escaping data characters that might otherwise be interpreted as row or column delimiters.

**Column Enclosure**

Characters used to enclose each alphanumeric value in the intermediate flat file that is loaded into the database. The enclosure consists of up to four printable characters. The most common enclosure is one double quotation mark. Numeric digits; symbols used in numbers, such as a period (.), plus sign (+), or minus sign (-); and 0x09 (Horizontal Tab), 0x0A (Line Feed), or 0x0D (Carriage return) cannot be used in the enclosure sequence. In order to specify a single quotation mark as the enclosure character, you must enter four consecutive single quotation marks.
**Note:** When a non-default value is used, the intermediate file is created as the PostgreSQL term CSV, instead of what is called TEXT.

**Global Settings**

These options are only available when BULKLOAD=ON and apply to all new flows using PostgreSQL as the target.

**Note:** ALIAS_CASE controls the case of column names when new tables are created. PostgreSQL users may want to select *Enforce lower case*.

**Reference:**  **Sybase Extended Bulk Load Options**

Sybase ASE and Sybase IQ targets have the following additional options, as shown in the following image.

---

**Column Delimiter**

Is the delimiter character(s) used between fields. It is used by the intermediate flat file used to load the database.

The delimiter can be up to four characters and can be specified as:

- **TAB.** A tab character. This is the default value.
- **a - A.** A character string. For example, ~.
- **0x nn.** A hex code, for example, 0x44 (a comma) or 0x0D0A (a return and a linefeed).

**Row Delimiter**

Is the delimiter character used between records. It is used by the intermediate flat file to the load the database. The row delimiter can be specified in the same manner as the (field) delimiter, except that character comma (,) is not permitted.
**Activate Server Log**

This parameter is available for Sybase IQ only.

Activate log information about integrity constraint violations and the types of violations. With this option two types of LOG files could be written into a temporary space in the configuration directory: MESSAGE LOG file etlblk.msg and ROW LOG file etlblk.log.

For Sybase IQ:

- **Yes.** Writes both MESSAGE LOG and ROW LOG files: etlblk.msg and etlblk.log.
- **No.** The LOG files are not written. This is the default value.

**Note:**

- Extended Bulk Load supports all Sybase IQ native data types, except BINARY and VARBINARY.
- A user running a data flow using Extended Bulk Load to Sybase IQ must have the following permissions:

  ```
  allow_read_client_file = on
  allow_write_client_file = on
  ```

  This can be confirmed by using the following Sybase stored procedure:

  ```
  sp_iqcheckoptions
  ```

  If these permissions are not set, the flow will fail and display the following error:

  ```
  FOC1261 PHYSICAL INTERFACE NOT FOUND FOR SUFFIX SQLSYB. MODULE
  NAME : dbodbc12_r
  ```
Refernce: 1010 Extended Bulk Load Options

1010 targets have the following additional options, as shown in the following image.

![Properties](image)

**Column Delimiter**

Delimiter characters between fields in the intermediate flat file that is loaded into the database. The delimiter can be up to four characters and can be specified as the following:

- **TAB.** A tab character. This value is the default.
- **a - A.** A character string. For example, ~.
- **0x nn.** A hex code. For example, 0x44 (a comma), or 0x0D0A (a return and a linefeed).

The hex code uses ASCII for Windows or UNIX systems and EBCDIC for IBM Mainframes.

Refernce: Teradata Extended Bulk Load Options

The Teradata Extended Bulk Load Utility uses TPT (Teradata Parallel Transporter) to load data.
Teradata targets have the following additional options, as shown in the following image.

**Rejected records maximum number**

The maximum number of rejected records that can be stored in one of the error tables before the job is terminated. If unspecified, the default is unlimited.

**TDP Id**

The Teradata Director Program Id.

**Account Id**

The Account Id used to access the database.

**Data error table**

Name of table to contain records that were rejected because of data conversion errors, constraint violations, or AMP configuration changes. This must be a new table name. The default is the database default.

**Index violation table**

Name of table to contain records that violated the unique primary index constraint. This must be a new table name. The default is the database default.
Db2 Target Bulk Load Options

Db2 targets have the following options:

**Loading method**

- **INSERT.** Insert the rows only. This is the default.
- **REPLACE.** Drops the existing destination table, and creates a new table.

**Note:**
- The bulk load option for Db2 is available for z/OS and IBM i only when configured for CLI, not CAF.
- For bulk load on IBM i, the target table name for Db2 must not exceed 10 characters.
- For bulk load on IBM i, the target table name must be a qualified name. The current library name must be specified, for example, library1/table1.

Db2 Target Insert Records from Memory Options

The Bulk Insert API (FASTLOAD) could be in effect for Db2 targets with load option Insert records from memory. The command that changes the settings for the Db2 FASTLOAD is:

```sql
ENGINE DB2 SET FASTLOAD [ON | OFF]
```

If effective, the default value for FASTLOAD is ON.

A SET command for Db2 to control whether or not the FASTLOAD operations are recoverable is:

```sql
SQL DB2 SET FASTLOAD_OPTION RECOVERABLE / NONRECOVERABLE
```

The default value is RECOVERABLE.

MySQL Bulk Load Options

There are no additional bulk load options for MySQL.
**Reference:** Nucleus Target Bulk Load Options

Nucleus targets have the following options:

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Load Type</td>
<td>Bulk load utility via a disk file</td>
</tr>
<tr>
<td><strong>Version of ndl</strong></td>
<td></td>
</tr>
<tr>
<td>Single-user</td>
<td></td>
</tr>
<tr>
<td><strong>Multi-user</strong></td>
<td></td>
</tr>
<tr>
<td>Overwrite the error log file</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Commit loaded rows if error occurs</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>First input file bytes to ignore</td>
<td>&lt;default&gt;</td>
</tr>
<tr>
<td>Null indicator</td>
<td>&lt;default&gt;</td>
</tr>
<tr>
<td>First input file rows to skip</td>
<td>&lt;default&gt;</td>
</tr>
<tr>
<td>Number of input file rows to process</td>
<td>&lt;default&gt;</td>
</tr>
<tr>
<td>Disconnect all other connections</td>
<td>No</td>
</tr>
<tr>
<td>Skip rows w/ unprintable characters</td>
<td>No</td>
</tr>
<tr>
<td>Server password</td>
<td>&lt;default&gt;</td>
</tr>
</tbody>
</table>

**Version of ndl**

**Single-user**

Starts the server in exclusive mode. No other database sessions may be established.

**Multi-user**

The server must already be started in multi-user mode and listening on the default port 1285.

**Overwrite the error log file**

**Yes**

Creates a new file.

**No**

Appends to existing log.ndl file.

**Commit loaded rows if error occurs**

**Yes**

Issues a commit.

**No**

Issues a rollback.
First input file bytes to ignore
Number of bytes to skip before loading data.

Null indicator
Specifies the character to represent null.

First input file rows to skip
Number of rows to skip before loading data.

Number of input file rows to process
Number of rows to process before stopping.

Disconnect all other connections
Disconnects connections.

Skip rows w/ unprintable characters
Skips rows with unprintable characters.

Server password
NDL server password. Required if Single User (ndls) mode and a server password actually exists.

Reference: Red Brick Target Bulk Load Options
Red brick targets have the following options:

Start record number in input field
Row number to begin copying data to the target.
End record number in input field

Row number to end copying data to the target.

Locale

The combination of language and location.

Maximum number of discarded records

The maximum number of discarded records allowed before loading stops. The default is the target RDBMS default.

Discard file name

File where duplicate records are discarded. The records will be stored for possible reloading.

RI discard file name

File where discarded records based on referential integrity will be stored for possible reloading.

Loading mode

INSERT. Insert the rows only. If the table is not empty, the load operation ends. This is the default.

APPEND. Used to insert additional rows of data into an existing table. Each new row must have a primary-key value that does not already exist in the table. Otherwise, the record is discarded.

REPLACE. Replaces the entire contents of a table.

MODIFY. Used to insert additional rows or to update existing rows in a table. If the input row has the same primary-key value as an existing row, the new row replaces the existing row. Otherwise, it is added as a new row.

UPDATE. Updates existing rows in an existing table. Each new row must have a primary-key value that is already present in the table. Otherwise, the record is discarded.

MODIFY AGGREGATE. If the primary key of the input row matches an existing row in the table, the existing row is updated as defined for the specified aggregate operator. If the primary key of the input row does not match an existing row in the table, the row is inserted.

UPDATE AGGREGATE. If the primary key of the input row does not match the primary key of a row already in the table, the input row is discarded. If it does match an existing row, the existing row is updated as defined for the specified aggregate operator.
Optimize

**OFF.** Indexes are updated when each input row is inserted into the data file, which provides better performance when the data being loaded contains many duplicate rows. This is the default.

**ON.** Overrides the global optimize mode setting in the rbw.config file.

*Reference:* Teradata Target Bulk Load Options

Teradata targets have the following options:

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Existing</td>
</tr>
<tr>
<td>Adapter</td>
<td>Teradata</td>
</tr>
<tr>
<td>Connection</td>
<td>cardb</td>
</tr>
<tr>
<td>Synonym</td>
<td>ORIGIN</td>
</tr>
<tr>
<td>Table</td>
<td>No changes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Load Type</th>
<th>Bulk Load Utility via a Disk File</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rejected records maximum number</td>
<td>25</td>
</tr>
<tr>
<td>TDP Id</td>
<td>&lt;default&gt;</td>
</tr>
<tr>
<td>Account Id</td>
<td>&lt;default&gt;</td>
</tr>
<tr>
<td>Maximum number of sessions</td>
<td>2</td>
</tr>
<tr>
<td>Work table</td>
<td>&lt;default&gt;</td>
</tr>
<tr>
<td>Acquisition phase errors table</td>
<td>&lt;default&gt;</td>
</tr>
<tr>
<td>Application phase errors table</td>
<td>&lt;default&gt;</td>
</tr>
<tr>
<td>Loading method</td>
<td>INSERT</td>
</tr>
<tr>
<td>AMP Check</td>
<td>APPLY</td>
</tr>
<tr>
<td>Log Table</td>
<td>&lt;default&gt;</td>
</tr>
</tbody>
</table>

*Note:* Teradata has two bulk load programs that DataMigrator supports. FastLoad is used for new target tables, and MultiLoad is used for existing target tables. Not all options shown in the image above are available for new target tables.

**UNIX only**

Two environment variables, $FASTLOAD_EXE and $MLOAD_EXE, should be set for using bulk load.

**$FASTLOAD_EXE.** Specifies the location of the Teradata FastLoad utility.
$MLOAD_EXE. Specifies the location of the Teradata MultiLoad utility.

Two lines should be added to the profile, or server start up shell script (edastart), that are similar to the following:

```sh
export FASTLOAD_EXE = ~teradata/fastload
export MLOAD_EXE = ~teradata/mload
```

**Rejected records maximum number**

The maximum number of rejected records allowed before loading stops. The default is the target RDBMS default.

**TDP Id**

The Teradata Director Program Id.

**Account Id**

The account Id used to access the database.

**Maximum number of sessions**

The maximum number of MultiLoad or FastLoad sessions logged on to the Teradata database.

**Start record number in source**

Row number to begin copying data to the target.

**End record number in source**

Row number to end copying data to the target.

**Work table**

Name of the work table.

**Acquisition phase errors table**

This table provides information about all errors that occur during the acquisition phase of your Update operator job, as well as some errors that occur during the application phase if the Teradata RDBMS cannot build a valid primary index.

**Application phase errors table**

This table provides information about uniqueness violations, field overflow on columns other than primary index fields, and constraint errors.

**Loading Method**

**INSERT.** Inserts the rows only. This is the default.

**UPSERT.** Does inserts for missing update rows.
Note: This option is available for existing target tables with keys.

AMP Check
The MultiLoad response to a down Access Module Processor (AMP) condition.

ALL. Pauses the MultiLoad job when AMP is down.

NONE. Allows the MultiLoad job to start, restart, or continue as long as no more than one AMP is down in a cluster.

APPLY. Inhibits the MultiLoad job from entering or exiting the application phase when an AMP is down.

Note: This option is available for existing targets only.

Log Table
Restarts the log table for the MultiLoad checkpoint information. This option is available for existing target tables only.

Notes
Enter an annotation for this target.

Procedure: How to Set the Data Flow Record Logging Options
Setting the data flow record logging options allows you to write particular types of transactions to log files.

Although Record Logging options appear in every data target, they are set on a per flow basis. Changing the options in one target will reset them for all targets.

   The Properties pane opens.
2. Expand the Record Logging attribute.
3. Select the check box next to the options that you want to log.
   Record logging options are optional and are not available when the data target is a formatted file.
4. Click Save in the Quick Access Toolbar.
   For more information on record logging options for a data flow, see How to Set the Flow Record Logging Properties on page 167.
Target Properties Pane

To access the Target Properties pane, right-click a target object and click Properties. The following image shows the Properties pane for an existing Relational target.

![Properties pane](image)

All targets have the following properties:

**Display Name**

By default, shows the application name, if enabled, and synonym name. Controls the name that is shown under the object in the flow.

**Notes**

Enter an annotation for this target.

**Type**

Is the target type.

**Adapter**

Is the target adapter.

**Synonym**

Is the synonym name.

**Note:** Synonym names cannot be longer than eight characters.
The following sections describe the Target Options available for each different target.

**Reference:**  **Target Properties Pane for Relational Targets**

The Target Options section of the Properties pane for relational targets contains the following additional options:

### Connection

Is the connection for the data target. For a relational data target, this is a database server. For ODBC, this is a data source.

### Table

Is the name of data target table.

### Keys

Is the names of key columns in order. This option is only available for new targets.
Prior to Load Options:

These options are only available for existing targets.

No changes

Does not delete the records already in a data target.

Delete all rows from table

Deletes all rows and creates a database log.

Truncate table

Deletes all rows from the table, but does not generate a database log. This is a faster option than using Delete all rows from table.

The Target Load Options include the following load options:

Load Type:

Note: Although Load Type appears in every data target, it is set on a per flow basis. Changing the type in one target will reset it for all targets.

Insert/Update

Specifies the behavior of DataMigrator while loading data.

Insert/Update options can be set on a per target basis.

If you select this option, you can set a behavior when duplicate records are found from the If the record exists drop-down menu.

☐ Include the record. Allows the relational database target to handle duplicates.

Note: If you select Include the record, the record is passed directly to the relational database, which determines whether to accept it or not. If inserting a record would result in a duplicate key, the RDBMS will reject it due to a Unique Index constraint violation and return an error. Processing continues even if such errors occur, up to the number of errors specified under Stop processing after ___ DBMS errors in the General properties of the flow.

☐ Reject the record. Issues a SELECT command against the target table to see if a record exists with the key values. If there are no key columns, DataMigrator screens all columns. If the record is found, the record is rejected.

☐ Update the existing record. Updates records if the key value on the input record, or entire record if there are no keys, are found in the table. All non-key values are updated in the target.
☐ **Delete the existing record.** Deletes the record if the key on the input record (or entire record if no keys) is found in the table.

If you select *Update the existing record* or *Delete the existing record* from the If the record exists drop-down menu, you can also set behavior when the record does not exist using the If the record does not exist drop-down menu:

☐ **Include the record.** Includes the key in the data target.

☐ **Reject the record.** Does not include the key in the data target.

*Note:* The Reject, Update, and Delete options can adversely affect performance because they determine the existence of a key value on an incoming record before performing the specified action. This is done by issuing a SELECT command against the target table, then waiting for the response. If one of these actions is required, try to limit the volume of records in the incremental change. These actions will perform best if there are unique keys on the table.

☐ **Commit every row(s).** Specifies the row interval to commit or write transactions to the database.

**Insert records from memory**

When you specify this load type, you need to specify:

☐ **Commit every row(s).** Specifies the row interval to commit or write transactions to the database.

☐ **Block size.** Specifies how many records you want to process at a time.

*Note:* The value for Commit every rows should be a multiple of the Block size, for example, 10000 and 2000.

**Bulk load utility via a disk file**

These options will depend on your target type.

**Extended bulk load utility**

Uses a database loader program with additional capabilities. For more information, see *Extended Bulk Load Utility Options* on page 288.

**Slowly Changing Dimensions**

This option is only available for existing relational targets.
Enables you to load a data target with column values that change over time. There are two column types. Type I simply overwrites the value. Type II allows you to track activity before and after the change. Type II changes are handled either using an activation flag or an end date/begin date. When you update a Type II column, you do not actually change its value. Instead, you update the activation flag to inactive, or the end date to the current date. You then add a new record with the new value, and the activation flag on by default or the begin date set.

- **Updates for Type I.** Enables you to specify how updates will affect Type I fields. The options are Change all rows or Change only active rows. Change all rows is the default value.

- **Commit every row(s).** Specifies the row interval to commit or write transactions to the database.

**IUD Processing**
Enables you to load a data target with only the records that have changed. This feature is an optional, add-on component to DataMigrator. When this option is selected, the Prior to Load option is eliminated.

**IUD/Slowly Changing Dimensions**
This load type is a combination of the previous two load types. The Insert/Update/Delete indications in the CDC logs are not applied directly to the target. Instead, the SCD Types of the target are used and deleted records in the source are marked as inactive in the target.

**Procedure: How to Set Options for Relational Targets**

1. In the data flow workspace, right-click the data target, and click Properties.
   The target Properties pane opens.

2. For existing targets, select whether to remove data prior to loading the data target in the Prior to Load Option drop-down menu. The options are:
   - **No Changes.** Does not delete the records already in a data target.
   - **Delete all rows from table.** Deletes all rows and creates a database log.
   - **Truncate table.** Deletes all rows from the table but does not generate a database log. This is a faster option than Delete all rows from table.

   **Note:** Truncate table is not supported by Db2 on IBM i.
3. Select a Load Type. The options are:

- **Insert/Update.** Allows you to set a behavior when loading records.

- **Insert records from memory.** Speeds the loading of the data target by inserting a block of rows at once. You can set the row interval to commit or write transactions and the number of records to load in a block. This option is only available for relational databases that support it, including Db2 on i V6R1 for CLI, Db2 on z, Informix, MS SQL Server, MySQL, Oracle, Sybase ASE, Teradata 13, and UDB.

  This option:

  - Requires clean input data. If any row in the block causes a data source constraint violation, such as not null or unique index, the entire block is rejected.

  - Does not provide row counts (the number of records inserted or rejected) in the detail log or statistics. NA (not available) will appear instead.

  - Does not provide record logging. For example, rejected records cannot be written to a flat file for review.

- **Bulk load utility via a disk file** and **Extended Bulk Load Utility.** Use database bulk loaders instead of iWay to insert data into a target. DataMigrator automates bulk loading for Hyperstage, Ingres, Informix, Microsoft SQL Server, IBM Db2, Teradata, Nucleus, Oracle, Sybase Adaptive Server Enterprise, and Sybase Adaptive Server IQ. You can set format version, maximum number of errors, first and last row to copy, packet size, and row interval to commit or write transactions.

- **Slowly Changing Dimensions.** Enables you to load a data target with column values that change over time. There are two column types. Type I simply overwrites the value. Type II allows you to track activity before and after the change. Type II changes are handled using either an activation flag or an end date/begin date. When you update a Type II column, you do not actually change its value. Instead, you update the activation flag to inactive, or the end date to the current date by default. You then add a new record with the new value and the activation flag on by default or the begin date set.

- **IUD Processing.** Enables you to load a data target with only the records that have changed. This feature is an optional, add-on component to DataMigrator.

  **Note:** The IUD Processing option is only available for existing relational targets.

- **IUD/Slowly Changing Dimension.** Is a combination of the previous two load types. The Insert/Update/Delete indications in the CDC logs are not applied directly to the target. Instead, the SCD Types of the target are used.
**Reference:** Target Properties Pane for FOCUS/FDS or XFOCUS Targets

The Target Options section of the Properties pane for FOCUS/FDS or XFOCUS targets contains the following additional options:

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td></td>
</tr>
<tr>
<td>Display Name</td>
<td>baseapp/dmsrep</td>
</tr>
<tr>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td><strong>Target Options</strong></td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>New</td>
</tr>
<tr>
<td>Adapter</td>
<td>Focus/FDS</td>
</tr>
<tr>
<td>Synonym</td>
<td>baseapp/dmsrep</td>
</tr>
<tr>
<td>Data File</td>
<td>baseapp/dmsrep.foc</td>
</tr>
<tr>
<td>Keys</td>
<td>1</td>
</tr>
</tbody>
</table>

**Data File**

Is the name of the data file pointed to by the synonym.

On IBM z/OS, to create a file in HFS (hierarchical file system), enter the name. To create a dataset, enter the name as

`//qualif.tablename.FOCUS`

where:

- **qualif**
  - Is a fully qualified location.

- **tablename**
  - Should match the synonym name.

**Keys**

Is the number of key columns. This option is only available for new targets.
Prior to Load Options:
These options are only available for existing targets.

No changes
Does not drop (delete) the data target.

Drop Table
Drops and recreates the data target.

Load Type
Specifies the method DataMigrator uses to load data.

Insert/Update
Since DataMigrator uses Insert/Update to load FOCUS/FDS or XFOCUS targets, you can set a behavior when duplicate records are found from the If the record exists drop-down menu. These options can be set on a per target basis.

- Include the record. Includes the duplicate record in the data target.
- Reject the record. Issues a SELECT command against the target table to see if a record exists with the key values. If there are no key columns, DataMigrator screens all columns. If the record is found, the record is rejected.
- Update the existing record. Updates records if the key value on the input record, or entire record if there are no keys, are found in the table. All non-key values are updated in the target.
- Delete the existing record. Deletes the record if the key on the input record (or entire record if no keys) is found in the table.

If you select Update the existing record or Delete the existing record from the If the record exists drop-down menu, you can also set behavior when the record does not exist using the If the record does not exist drop-down menu:

- Include the record. Includes the key in the data target.
- Reject the record. Does not include the key in the data target.
- Commit every row(s). Specifies the row interval to commit or write transactions to the database.

Procedure: How to Set Options for FOCUS/FDS or XFOCUS Targets
1. In the data flow workspace, right-click the data target and click Properties.

The Target Properties pane opens.
2. For existing targets, select whether to remove data prior to loading the data target in the Prior to Load Options section. The options are:

- **No changes.** Does not delete the records already in a data target.
- **Drop Table.** Drops and recreates the data target.

3. Select a behavior for loading records using the If the record exists and If the record does not exist drop-down menus. These options can be set on a per target basis. They are:

- **Include the record.** Includes the duplicate record in the data target.
- **Reject the record.** Issues a SELECT command against the target table to see if a record exists with the key values. If there are no key columns, DataMigrator screens all columns. If the record is found, the record is rejected.
- **Update the existing record.** Updates records if the key value on the input record, or entire record if there are no keys, are found in the table. All non-key values are updated in the target.
- **Delete the existing record.** Deletes the record if the key on the input record (or entire record if no keys) is found in the table.

4. If you select **Update the existing record** or **Delete the existing record** from the If the record exists drop-down menu, you can also set behavior when the record does not exist using the If the record does not exist drop-down menu:

- **Include the record** includes the key in the data target.

5. Click **OK**.

**Note:**

- For FOCUS/FDS or XFOCUS targets, the only Load Type available is Insert/Update.
- Synonym names for FOCUS/FDS files cannot be longer than eight characters.

**Reference:** Target Properties for New XML Targets

DataMigrator can create a new XML document with name/value pairs. For more information on any other structures, see **Creating a Data Flow Using a Target Based on a Predefined XML Schema** on page 607.

**Note:** In order to create an XML document, a Server Administrator must first configure an adapter for XML.
The Target Options section of the Properties pane for XML targets contains the following additional options:

**Data File**

Is the name of the XML document described by the synonym. DataMigrator will also create an XML Schema Definition with the same name as the data file and an extension of .xsd.

**Load Type**

Is set to Insert/Update for XML documents. Note that no updates are currently performed.

**Records**

Is the name of the top level element.

**Record**

Is the name of the row level element.

**Column**

Is the name of the column level element.

**Example:** **New XML Targets**

The following example uses the same table ibisamp/dminv as a source with all the actual columns, and then creates a new XML target called dminvx. The element names are entered as shown:

Records - Inventory
Specifying Data Target Options

Record - Item

Column - Detail

The first three rows of the resulting XML document would look like the following example when viewed in a browser:

```xml
<?xml version="1.0" encoding="UTF-8" ?>
<Inventory xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <item>
    <detail name="PROD_NUM">1004</detail>
    <detail name="PRODNAME">2 Hd VCR LCD Menu</detail>
    <detail name="QTY_IN_STOCK">43068</detail>
    <detail name="PRICE">179.00</detail>
    <detail name="COST">129.00</detail>
  </item>
  <item>
    <detail name="PROD_NUM">1006</detail>
    <detail name="PRODNAME">Combo Player - 4 Hd VCR + DVD</detail>
    <detail name="QTY_IN_STOCK">13527</detail>
    <detail name="PRICE">399.00</detail>
    <detail name="COST">289.00</detail>
  </item>
  <item>
    <detail name="PROD_NUM">1008</detail>
    <detail name="PRODNAME">DVD Upgrade Unit for Cent. VCR</detail>
    <detail name="QTY_IN_STOCK">199</detail>
    <detail name="PRICE">199.00</detail>
    <detail name="COST">139.00</detail>
  </item>
</Inventory>
```

**Reference:** Target Properties Pane for Delimited Flat File Targets

The Target Options section of the Properties pane for delimited flat file targets contains the following additional options:
**Note:** Delimited Flat File will only appear as a target type if it is configured on the server.

### Connection
Select the connection, FTP or SFTP server, where the data should be written.

### Data File
Is the name of the data file pointed to by the synonym.

On IBM z/OS, to create a file in HFS (hierarchical file system), enter the name. To create a dataset, enter the name as

```
//'qualiftablename.DAT
```

where:

- `qualif`

  Is a fully qualified location.

- `tablename`

  Should match the synonym name.
**Code Page**

Specifies a code page for the target data.

**Field Delimiter**

Is the delimiter character(s) used between fields. The delimiter can be up to 30 characters. It can be specified as:

- **TAB.** A tab character. This is the default.
- **a.** A character string, for example ~.
- **0xnn.** A hex code, for example, 0x44 (a comma), or 0x0D0A (a return and a linefeed). The hex code uses ASCII for Windows or UNIX systems and EBCDIC for IBM Mainframes.

**Header**

Inserts column headings as the first row of data and surrounds the column names with the character specified in the Enclosure field. When set to Yes, the FIELDNAME is used for the header.

**Enclosure**

This character is used to surround the column headings when Header is set to Yes.
Reference: Target Properties Pane for Flat File Targets

The Target Options section of the Properties pane for flat file targets contains the following additional options:

**Connection**

The connection configured for the selected adapter.

**Data File**

Is the name of the data file pointed to by the synonym.

On IBM z/OS, to create a file in HFS (hierarchical file system), enter the name. To create a dataset, enter the name as

```
'//qualif.tablename.DATA'
```

where:

- **qualif**
  - Is a fully qualified location.

- **tablename**
  - Should match the synonym name.

**Code Page**

Specifies a code page for the target data.
Specifying Data Target Options

**Prior to Load Options:**
This option is only available for existing targets.

**No changes**
Does not delete the data target.

**Delete File**
Deletes and recreates the data target.

**Reference:**  Target Properties Dialog Box for Formatted File Targets
The Target Options section of the Properties pane for formatted file targets contains the following additional options:

![Properties Dialog Box](image)

**Format**
Is the format of the target data.

**Data File**
Is the name of the data file pointed to by the synonym.

**Procedure:**  How to Set Options for Flat, Delimited Flat, Formatted, and XML File Targets
1. In the data flow workspace, right-click the data target and click *Properties*.
   The Target Properties pane opens.
2. For existing targets, select whether to remove data prior to loading the data target in the *Prior to Load Options* section. The options are:
   - *No changes* does not delete the data target. New records are appended.

322  Information Builders
Delete File drops and recreates the data target.

3. Click OK.

**Note:**

- For flat files and XML files, the Load Type cannot be changed from Insert/Update, and the data is always appended.
- For formatted file type EXL2K, the only Load Type available is Loading Flat File using HOLD.

**Loading Data to Salesforce.com**

In order to insert or update rows in a Salesforce object, you can use a custom field identified as an External ID field.

**Procedure:** How to Create a Custom Field as an External ID

1. Log in to salesforce.com.
2. Click your name and select Setup.
3. Under App Setup, expand Customize. Select an object, for example, Account.
4. Click Add a field to on the object page.
5. Under Custom fields and relationships, click New.
6. Select a data type appropriate to an ID field, such as Text or Number. Click OK.
7. Enter a name, for example, Account ID. Fill in the other fields as desired.
8. Select the External ID checkbox.
9. Click Next twice and then click Save.

The Account (or other object) Fields page opens. The new field is shown with its AP name, for example, Account_ID__C.

10. Create, or update, a synonym for Accounts (or another object).

**Procedure:** How to Identify an External ID Field In a Synonym

The synonyms created for Salesforce.com objects identify the ID field as the key. To use the External ID field as the key, you must edit the generated synonym and save a copy of it.

1. In the DMC, double-click the synonym for Account (or another object).
2. Right-click the ACCOUNT segment and click Properties.
3. For the attribute KEY, type the name of the External ID field, such as ACCOUNT_ID__C.

4. In the DMC, click Save As.

5. Enter a new name, such as accountx. Click Save.

Reference: Using an External ID Field to Insert or UpdateRows in a Salesforce.com Object

Using the alternative synonym you can now use the External ID field to identify rows to be inserted or updated as you would with a table in a relational database. Salesforce.com will automatically do an *upsert*; if a value of the External ID field matches a value already in the object being loaded, an update will occur. If not, it will do an insert.

Reference: Deleting Rows from a Salesforce.com Object

The internal ID field is used to delete rows from a Salesforce.com object. To delete rows using the External ID, create a data flow that performs a DB_LOOKUP. You can also create a data flow that joins a table with a list of External ID values to be deleted to the Salesforce.com object to obtain the values of the corresponding ID field.

Mapping Transformations

After selecting a data source and target, you can create mappings to transform your data in the following ways:

- Map the columns of the source, or combinations of columns in the source, to columns in the data target.
- Convert the extract data into the form necessary for the data target.
- Create a temporary column for mapping.

*Note:* You can create a temporary column within the mapping step to add intermediate calculations that are not part of the original SQL command or target data. This is often useful for complex calculations that are actually composed of multiple expressions. Creating a temporary column makes the transformation easier to understand and support. It is also helpful in performing data type conversions when multiple steps are required.

There are three methods for mapping columns:

- **Automapping.** Automatically maps columns with the same name.
- **Simple mapping.** Maps one column in the source to one in the target.
- **Complex mapping.** Builds expressions.
**Note:** The target column format should be at least as large as the source column format, but you can adjust the transformation to accommodate different format sizes.

**Reference:** Considerations for Mapping to an Existing Flat File

While you can map to an existing flat file, there are some considerations that should be taken into account.

- Flat files do not have all of the options for loading data into the target that other data types do (for example, key matching logic).

- DataMigrator respects the data types described in the synonym of the flat file. For example, if the synonym describes numeric columns in integer format, data will be written to the file in that format.

- If a file already exists with the specified name, DataMigrator will append data to that file, unless you select the Prior to Load option of Delete File. Existing data in a flat file cannot be updated.

**Procedure:** How to Automatically Map All Columns

1. On the Home tab, in the Tools group, click Options.
2. Expand Data Flow and select Data Flow Designer from the User Preferences tree.
3. Select the Automatically map all columns check box in the On Add Target section.
4. Click OK.
**Procedure:** How to Map Transformations

1. From the data flow workspace, double-click a target, or right-click it and click Target Transformations.

The Transformations dialog box opens, as shown in the following image.

2. To map columns with identical names, click the **AutoMap** button.

The mapping will appear in the grid under the Expressions tab.

3. To map columns with different names, select one column from each grid and click the equal sign.

The mapping will appear in the grid under the Expressions tab, with the Selected Column appearing in the Transform Expression box.

4. To create a complex mapping for a target column, double-click a column to move it to the Expressions tab.

5. Click the **Edit Transforms** button.

The Transformation Calculator opens.
6. Create an expression using the Columns/Variables and Functions tabs, and the calculator buttons. For details on using the Transformation Calculator, see Transformation Calculator on page 229. For information about using functions, see Using Functions on page 347. For information about using variables, see Using Variables in a Flow on page 336.

   **Note:** When you create expressions, keep in mind their order of evaluation. For details, see Order of Evaluation on page 233.

7. Click **OK**.

   **Note:** Clicking a column heading will sort the column grid on that heading.

The new mapping will appear in the grid under the Expressions tab.

   **Note:** The transformations will occur in the order in which they appear in the Expressions tab. If you need to change the order, select a transformation and use the up and down arrows.

**Procedure:** How to Create a Temporary Column for Mapping

1. In the data flow workspace, double-click a target, or right-click it and click Target Transformations.

   The Transformations dialog box opens.

2. Click the Insert Transforms button.

   The Transformation Calculator opens.

3. Type a name and type or select a format for the column.

4. Create an expression using the Columns/Variables and Functions tabs and the calculator buttons. For details on using the Transformation Calculator, see Transformation Calculator on page 229. For information about using functions, see Using Functions on page 347. For information about using variables, see Using Variables in a Flow on page 336.

   **Note:** When you create expressions, keep in mind their order of evaluation. For details, see Order of Evaluation on page 233.

5. Click **OK**.

The new mapping will appear in the grid under the Expressions tab.

   The new mapping is now available as a column in the Transformation Calculator, and can be used to create a complex expression.

**Tip:** Since calculations are performed in the order they appear, make sure that your temporary column appears *before* any column in the Transform Rules box that uses it to calculate a value.
Procedure:  How to Create a Target Transformation Using Relational Lookups

The DB_LOOKUP function allows you to build data flows that perform lookups into multiple tables using target transformations instead of explicit joins. This approach can simplify flows that require lookup values from numerous tables. For information about using the DB_LOOKUP function to create source transformations, see Using Lookups to Create Multi-Table Flows Without Joins on page 594.

1. From the data flow workspace, double-click a target object, or right-click it and click Target Transformations.
   
   The Transformations dialog box opens.

2. Double-click the column that will use the lookup. The column is moved to the grid under the Expressions tab.

3. With the column selected in the Expressions grid, click the Edit Transforms button.
   
   The Transformation Calculator opens.

4. Click the Functions tab, expand the Data Source and Decoding - Simplified folder, and double-click the DB_LOOKUP function.
   
   The prototype for DB_LOOKUP appears in the Expression window and the DB_LOOKUP dialog box opens.

5. Select a synonym to use for the lookup by clicking the ellipsis (...) after the Lookup Synonym field.
   
   The Select Synonym dialog box opens.

6. Select a synonym and click Select.

7. From the Lookup dialog box, click the Add/Delete button.

8. A lookup is similar to a join in that it uses columns from two different tables. Select a column from the Source and Lookup Fields lists, and click the equal sign (=).
   
   Note: If you have Automatically add join conditions checked in the Data Management section of the Options dialog box, the Lookup is added automatically, since both data sources have a column with the same name and format, and that column is a key column in the lookup table.
When you have selected the columns, the DB_LOOKUP dialog box will look like this:

![DB_LOOKUP dialog box](image)

9. Select the value to return from the lookup from the Return Field drop-down menu.
10. Select the original column name and click OK. Note that in this case, the format of the column is A20.
11. Click OK.

The expression is completed in the Transformation Calculator dialog box.
The Transformation Calculator should now look like this.

![Transformation Calculator](image)

12. Click **OK**.

The target transformation is added to the Transformations dialog box.

13. Click **OK** to close the Transformations dialog box.

**Procedure:** How to Edit a Transformation Mapping

From the Transformations dialog box:

1. Double-click the column number or name in the Expressions tab.

   or

   Select the column name in the Expressions tab and click the *Edit transforms* button.

   The Transformation Calculator opens.

2. Modify the expression.

   **Note:** You can use the *Function Assist* button to change the parameters of a transformation that uses the DB_LOOKUP function.

3. When you are finished with your changes click **OK**.

   You can also double-click the *Expression* field in the Expressions tab, and enter the expression directly.
Procedure: How to Adjust a Transformation Mapping

If the format of the target column is smaller than the format of the source column, the transformation mapping will fail. For example, moving a column formatted as A24 into a column formatted as A10 will fail because you cannot move 24 characters into a 10-character field. To solve this problem, you would need to use edit or substring in a transform expression.

The Transformations dialog box provides an option for automatically adjusting the transformation to reconcile different format sizes. Adjust will also convert integer, alphanumeric, and datetime source columns so that they match the target column.

In the Transformations dialog box:
1. Map the source and target columns by selecting one from each list and clicking the equal sign.
   The mapping appears in the Expressions list.

2. With the expression highlighted, click the Adjust Expression button.

Procedure: How to Add Comments to a Transformation

1. Right-click the title bar under the Expression tab in the Transformations dialog box.
2. Point to Customize from the shortcut menu and then click Description.

The Transformations dialog box now shows a Description column. Here you can enter a comment for the transformation that is stored in the flow.

Procedure: How to Test Mappings

From the Transformations dialog box:

1. Click the Test Transforms button.
   The Test Transformations dialog box opens with sample data for the transformations.

2. Click the Close button.
Reference: Transformations Dialog Box

To access the Transformations dialog box from the data flow workspace, double-click a target, or right-click a target and click Target Transformations.

The Transformations dialog box contains the following fields and options:

**Find**

Allows you to search for a column.

**Refresh**

Refreshes the columns in the target.

**Automap**

Maps all columns with identical names.

**Target Columns**

Lists the columns in the data target.
Note: Clicking a column heading will sort the column grid on that heading.

**Selected Columns**

Lists the columns in the data source. Next to each column name, there is a symbol that indicates regular columns, key columns, index columns, and virtual columns.

In both Target and Selected columns, the symbols are:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Key column</td>
</tr>
<tr>
<td></td>
<td>Regular column</td>
</tr>
<tr>
<td></td>
<td>Index column</td>
</tr>
<tr>
<td></td>
<td>Virtual columns (in Target columns only). These are grayed and cannot be selected for mapping.</td>
</tr>
</tbody>
</table>

=

Establishes the relationship between the selected columns as equal.

**Insert Transforms**

Inserts a new mapping and opens the Transformation Calculator.

**Delete**

Deletes a selected mapping.

**Edit Transforms**

Opens the Transformation Calculator to edit the selected mapping.

**Adjust Expressions**

Adjusts the transformation expression to reconcile different format sizes.
Expressions tab

Lists the transformation mappings. To access the context menu with available options for
the transformations, right-click a transformation in this tab. For more information, see
Source Transformations Shortcut Menu on page 228.

On Match Expressions tab

Lists the expressions that are performed for rows that are updated.

Validates tab

Lists transformations that screen records as they are loaded into the data target. If a
record does not meet the criteria, it is rejected.

Test transforms

Retrieves, by default, 50 records and 4096 characters per row from the source, and
applies all available transformations. This option enables you to review the data being
moved and to ensure that the syntax of each transformation is correct. No duplicate
processing is performed. For example, if a key value is duplicated, a unique constraint
violation may occur. Similarly, if a value is missing for a column described as not null, a
constraint violation will occur at run time.

Note: You can set the number of rows retrieved from the Options dialog box under the Run
Options branch of the User Preferences tree. The default is 50 rows.

If there are any errors, the Server Messages window displays an error message and marks
the incorrect transformation in red.

Up and Down arrows

Move the selected mapping to change the order of the transformations.

Procedure: How to View Error Details

Error details for transformations can be viewed in the DMC. If an error is made, the transform
will appear in red.

1. Right-click the error and select Error Details.
2. The Error(s) Details dialog box opens, as seen in the following image.

![Error(s) Details Image]

To see the error message details, click the error message number. A window opens explaining the error, as seen in the following image.

![Error Message Image]

3. Click OK.

**Note:** This method of viewing error details also applies to target transformations.

**Validating Records**

You can create a business rule that screens records as they are loaded into the data target. If a record does not meet the criteria, it is rejected.

If a value meets the criteria of the rule, a value of 1 is assigned to the record and the record is loaded. If the value does not meet the criteria of the record, the value of 0 is assigned to the record and the record is rejected.
**Procedure: How to Validate Records**

1. In the Transformations dialog box, select the Validates tab.

2. Click the Insert Intermediate Transforms button.
   
   The Transformation Calculator opens with a name and format already provided.

3. Optionally, enter a name for the validation rule in the Column field.

4. Create an expression for the validation rule using the Columns, Functions, and Variables tabs, and the calculator buttons.
   
   If the expression evaluates to TRUE then the row is loaded in the target table. If the expression evaluates to FALSE then the row is rejected as invalid.

   Note that the validation is performed after any filters or aggregation. For example, if you enter QUANTITY GT 10, only records where the sum of the values of QUANTITY for the group by fields is greater than 10 will be loaded.

   The count of rejected rows is reported in the DataMigrator log as invalid records. These records can also be logged to a file for review. For more information, see Transformation Calculator on page 229.

5. Click OK.

   The validation rule appears in the grid below the Validates tab.

**Using Variables in a Flow**

You can use a variable in your data flow to represent a value in a transformation. For process flows, variables are typically used in stored procedures or to control the flow.

The following types of variables can be used in a flow.

- DataMigrator variables are supplied with the DMC.
- System variables are supplied with the DMC.
- User-defined variables are created by a user.

   **Note:** You can set a value for a variable in the set variables objects on the process flow tab. You can also specify run-time parameters directly from a batch file, shell script, or JCL. For information on setting values when executing flows from CMASAP or CMRUN, see Event-Based Scheduling on page 695.

- Statistical and Control variables can be referenced in stored procedures. They are primarily used in process flows.
Reference: DataMigrator Variables

DataMigrator variables are available from the Variables tab in the calculators. DataMigrator variables start with &&CM__ (double underscores). They include:

&&CM__AUTHOR

Is the user ID that originally saved the flow.

Note: &&CM__AUTHOR is the user ID that originally saved the flow, which is not always the user ID that ran the flow. To find the user ID that ran the flow, assign a value of GETUSER('A32') to a variable in the set variables object on the Process Flow tab.

&&CM__REQUEST

Is the name of the flow.

&&CM__TARGET

Is the name of the target table for the data flow. Note that for flows with multiple target tables, the variable CM__TARGET has a number appended to the name.

Reference: Control Variables

Control variables are typically used in process flows.

&&KILL_RPC

Shuts down all flow processing if the value is set to Y. The flow immediately branches to the end of its logic, skipping all code that follows the stored procedure where this setting is used. Any data flows or procedures are not executed. This may be useful in cases where a site-dependent variable (program result, special file allocation, and so on) fails, and the entire extract is to be stopped. The default is N.

To reset, assign a value of 'Y' to the global variable KILL_RPC in a set variables object on the process flow tab.

Reference: System Variables

System variables are identified by a single ampersand and an alphanumeric name, up to ten characters long. These variables can be referenced in stored procedures. Each date shown below represents May 25, 2005.

&DATE

Is the current date (MM/DD/YY) as a character string (A8).

Example: 05/25/05
&DATEDMY
Is the current date (DD/MM/YY) as a character string (A8).

Example: 25/05/05

&DATEMDYY
Is the current date as month, day, with a four-digit year.

Example: 05/25/2005

&DATEMDY
Is the current date as month, day, year.

Example: 052505

&DATEMYMDY
Is the current date (YY/MM/DD) as a character string (A8).

Example: 05/05/25

&DATEMYMDY
Is the current date (CCYY/MM/DD) as a character string (A10).

Example: 2005/05/25

&DMY
Is the current date as day, month, year. Useful for numerical comparisons.

Example: 250505

&DMYY
Is the current date as day, month, with a four-digit year.

Example: 25052005

&MDY
Is the current date as month, day, year.

Example: 052505

&MDYY
Is the current date as month, day, with a four-digit year.
Example: 05252005

&TOD
Is the current time. (HH.MM.SS)

Example: 17:00:00

&YMD
Is the current date as year, month, day.

Example: 050525

&YYMD
Is the current date as a four-digit year, month, day.

Example: 20050525

Reference: Statistical Variables
DataMigrator supports the following statistical variables that can be referenced in stored procedures. Statistical variables are recorded up to a value of 2,147,483,647. They are primarily used in process flows.

&ACCEPTS
Is the number of transactions accepted.

&CHNGD
Is the number of rows updated.

&DBMSERR
Is the number of rows rejected due to database constraint violations, or for flat file targets due to a format error.

&DELTd
Is the number of rows deleted.

&DUPLS
Is the number of rows rejected as a result of duplicate values in the data source.

&EXITRC
Is return code value from the execution of an operating system command.

&FORMAT
Is the number of rows rejected as a result of a format error, or when a field is truncated.
&INPUT
Is the number of rows added to the data source.

&INVALID
Is the number of rows rejected as a result of an invalid condition.

&NOMATCH
Is the number of rows rejected as a result of not matching a value in the data source.

&REJECTS
Is the number of rows rejected for reasons other than the ones specifically tracked by other statistical variables.

&TRANS
Is the number of rows processed.

&LINES
Is the number lines returned in the last answer set.

&RECORDS
Is the number of records retrieved in last answer set.

Procedure: How to Use DataMigrator or System Variables

1. From a calculator, select the Columns/Variables tab.
   
   **Note:** The variables are grouped by category. To view all variables, right-click in the variables workspace and click View Alphabetical.
   
2. Double-click a variable to add it to an expression or condition. You can also type in the name of the variable. DataMigrator and System variables must be enclosed in single quotation marks (‘).
**Note:** Match the format (alphanumeric or numeric) to the system variable in a Source or Target Transformation. &TOD and system variables that start with &DATE are alphanumeric. All others are numeric.

3. When the expression is complete, click **OK**.
The variable appears in the Expression column in the Transformations window, as shown in the following page.

The Variables tab is also available in the SQL and Filter calculators.

**Using a Parameter With a Flow**

Variables that are used as parameters are called local variables and are identified within a flow by a single ampersand and an alphanumeric name, up to ten characters long. Local variables can be used throughout the flow, for example, in transformations or filters. In order to use a local variable it must have a value.

- **Establishing local variables when designing a flow.** If a flow contains a local variable and you run any test, you will be prompted to enter a value for the variable.

- **Setting a value when running a flow.** Local variables are set in a process flow that calls a data flow in the flow properties.
Procedure: How to Use a Local Variable as a Passed Parameter

1. You can use a local variable, a variable known only to a particular procedure, as a passed parameter.

2. Type the variable name into an expression or condition in one of the calculators. The variable name must be enclosed in single quotes if it contains alphanumeric or data values, as shown in the following image.

![Image of WHERE Filter Calculator with an example expression: `T2.ORDER_DATE = '20230101'`]

3. Click OK.

Procedure: How to Test a Local Variable

1. Open the expression in a calculator.

2. Click the Test SQL Statement button.

3. DataMigrator prompts you to enter a value for the variable.
Note that when entering a date value, no quotes are required, and the date must be entered in the format used in a where condition, as shown in the following image.

4. Click Run.

**Procedure:** How to Pass a Parameter Using a Local Variable

1. Create a process flow.
2. In the Process Flow tab, drag the data flow that uses the local variable into the workspace, to the right of the Start object.
3. Double-click the Data Flow object (ibisamp/dmvar), or right-click the object and click Properties.
   
   The Properties panel opens.
4. Enter the value for the parameter in the Parameters field, as shown in the following image.

Note that you do not need to manually enter the parameter name and value. Clicking the ellipsis button to the right will bring up a dialog box where you can enter values for the parameters. If you choose to enter a date value manually, no quotes are required, and the date must be entered in the format used in a where condition. Multiple parameters can be separated by commas. Also note that checking Run as Stored Procedure is not required.

5. Click OK.

6. Save and run the flow.

*Procedure:* How to Pass a Local Variable as a Parameter in a Parallel Group

1. Create a process flow.
2. Add a Parallel Group.
3. Drag the Data Flows that you want to run inside the Parallel Group.

4. Double-click the data flow that requires a parameter, or right-click it and click *Properties*.

5. Enter the value for the parameter in the Parameters field, as shown in the following image.

Note that you do not need to manually enter the parameter name and value. Clicking the ellipsis button to the right will bring up a dialog box where you can enter values for the parameters. If you choose to enter a date value manually, no quotes are required and the date must be entered in the format used in a where condition. Multiple parameters can be separated by commas.
6. Alternatively, the value for the local variable can be a (global) variable set in a stored procedure, as shown in the following image.

![Image of Properties window for a stored procedure]

7. Click OK.

8. Save and run the flow.

**Using Functions**

When you click the *Functions* tab in the Transformation, Filter, or SQL calculators, a list of available functions appears. Right-click the Functions tab to display them alphabetically or by category. For a list of available SQL functions, see *SQL Functions* on page 259.

When using functions in a calculator, the Function Assist dialog box can help you set the parameters.

**Note:** All functions may not be available on all platforms.

<table>
<thead>
<tr>
<th>Function</th>
<th>Arguments</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Character Functions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ARGLEN</strong></td>
<td><em>(length, source_string, output_format)</em></td>
<td>Measures the length of a character string within a field, excluding trailing blanks.</td>
</tr>
</tbody>
</table>
| **ASIS** | *(argument)* | Distinguishes between a space and a zero.  
**Note:** This function is only available for variables and appears only in the User Variables calculator. |
<table>
<thead>
<tr>
<th>Function</th>
<th>Arguments</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHKFMT</td>
<td>(numchar, source_string, 'mask', output_format)</td>
<td>Checks numeric and alphanumeric fields for invalid character types.</td>
</tr>
<tr>
<td>CTRAN</td>
<td>(length, source_string, decimal, decvalue, output_format)</td>
<td>Replaces characters in a string, taking decimal ASCII values as arguments for the target.</td>
</tr>
<tr>
<td>CTRFLD</td>
<td>(source_string, length, output_format)</td>
<td>Centers a character string within a field.</td>
</tr>
<tr>
<td>EDIT</td>
<td>(source_string)</td>
<td>Converts between alphanumeric and integer data.</td>
</tr>
<tr>
<td>EDIT</td>
<td>(source_string, 'mask')</td>
<td>Using a mask, extracts characters from an alphanumeric string and inserts new characters.</td>
</tr>
<tr>
<td>GETTOK</td>
<td>(source_string, length, token_number, 'delim', outlen, output_format)</td>
<td>Divides a string into tokens and returns the specified token.</td>
</tr>
<tr>
<td>LCWORD</td>
<td>(length, source_string, output_format)</td>
<td>Converts the letters in a given string to mixed-case: the first letter of each new word and the first letter after a single or double quotation mark are converted to uppercase. The rest are converted to lowercase.</td>
</tr>
<tr>
<td>LCWORD2</td>
<td>(length, source_string, output_format)</td>
<td>Converts a character string to mixed-case. The first letter of every word is converted to uppercase.</td>
</tr>
<tr>
<td>Function</td>
<td>Arguments</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>LCWORD3</td>
<td>(length, source_string, output_format)</td>
<td>Converts a character string to mixed-case; the first letter after every apostrophe remains capitalized and every other letter is converted to lowercase, unless it is the last letter of the word. If it is the last letter, it will be converted to lowercase.</td>
</tr>
<tr>
<td>LJUST</td>
<td>(length, source_string, output_format)</td>
<td>Left-justifies a character string within its field.</td>
</tr>
<tr>
<td>LOCASE</td>
<td>(length, source_string, output_format)</td>
<td>Returns the infield in lowercase.</td>
</tr>
<tr>
<td>OVRLAY</td>
<td>(source_string, length, substring, sublen, position, output_format)</td>
<td>Overlays a substring to a specified place on another character string.</td>
</tr>
<tr>
<td>PARAG</td>
<td>(length, source_string, 'delim', max_token_size, output_format)</td>
<td>Divides lines of text into smaller lines by marking them off with a delimiter character.</td>
</tr>
<tr>
<td>PATTERN</td>
<td>(length, source_string, output_format)</td>
<td>Converts text to its pattern.</td>
</tr>
<tr>
<td>PATTERNS</td>
<td>(string)</td>
<td>Returns a character string that represents the pattern of the input string. Letters appear as 'A' or 'a', digits appear as '9', and special characters appear as is. This function is available starting in release 7708.</td>
</tr>
<tr>
<td>Function</td>
<td>Arguments</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-----------</td>
<td>-------------</td>
</tr>
<tr>
<td><strong>POSIT</strong></td>
<td>$(source_string, length, substring, sublength, output_format)$</td>
<td>Finds the starting position of a substring within a larger string.</td>
</tr>
<tr>
<td><strong>REPLACE</strong></td>
<td>$(input_string, search_string, replacement)$</td>
<td>Replaces all instances of a search string in an input string with the given replacement string.</td>
</tr>
<tr>
<td><strong>RJUST</strong></td>
<td>$(length, source_string, output_format)$</td>
<td>Right-justifies a character string within a field.</td>
</tr>
<tr>
<td><strong>SOUNDEX</strong></td>
<td>$(length, source_string, output_format)$</td>
<td>Converts character strings to 4-letter codes, enabling you to search for character strings phonetically without knowing how they are spelled.</td>
</tr>
<tr>
<td><strong>SPELLNM</strong></td>
<td>$(outlength, number, output_format)$</td>
<td>Takes a number with two decimal spaces and spells it out with dollars and cents.</td>
</tr>
<tr>
<td><strong>SPLIT</strong></td>
<td>$(element, string)$</td>
<td>Returns a specific type of element from a string. The output is returned as variable length alphanumeric.</td>
</tr>
<tr>
<td><strong>SQUEEZ</strong></td>
<td>$(length, source_string, output_format)$</td>
<td>Reduces multiple contiguous blank characters within an input string to a single blank character.</td>
</tr>
<tr>
<td><strong>STRIP</strong></td>
<td>$(length, source_string, char, output_format)$</td>
<td>Removes all occurrences of a specific character from an input string.</td>
</tr>
<tr>
<td>Function</td>
<td>Arguments</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>STRREP</td>
<td>(inlength, instring, searchlength, searchString, replength, repstring, outlength, output_format)</td>
<td>Replaces character strings.</td>
</tr>
<tr>
<td>SUBSTR</td>
<td>(length, source_string, start, end, sublength, output_format)</td>
<td>Extracts a substring of a given length and starting position from a larger string.</td>
</tr>
<tr>
<td>TRIM</td>
<td>(trim_where, source_string, length, pattern, sublength, output_format)</td>
<td>Removes leading and/or trailing occurrences of a pattern within a string.</td>
</tr>
<tr>
<td>TRUNCATE</td>
<td>(ARGUMENT)</td>
<td>Removes trailing blanks from Dialogue Manager amper variables, and adjusts the length accordingly. <strong>Note:</strong> This function is only available for variables and appears only in the User Variables calculator.</td>
</tr>
<tr>
<td>UPCASE</td>
<td>(length, source_string, output_format)</td>
<td>Converts all characters in a string to uppercase.</td>
</tr>
</tbody>
</table>

### Character Functions - DBCS Code Pages

<table>
<thead>
<tr>
<th>Function</th>
<th>Arguments</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCTRAN</td>
<td>(length, source_string, decimal, decvalue, output_format)</td>
<td>Translates a single-byte or double-byte character within a character string to another character based on its decimal value.</td>
</tr>
<tr>
<td>Function</td>
<td>Arguments</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>DEDIT</td>
<td>(length, source_string, mask_length, 'mask', output_format)</td>
<td>Extracts characters from or adds characters to a string when server is configured to use a DBCS code page.</td>
</tr>
<tr>
<td>DPART</td>
<td>(date, 'component', output_format)</td>
<td>Retrieves a date component as a numeric value.</td>
</tr>
<tr>
<td>DSTRIP</td>
<td>(length, source_string, char, output_format)</td>
<td>Removes all occurrences of a specific single-byte or double-byte character from a string.</td>
</tr>
<tr>
<td>DSUBSTR</td>
<td>(length, source_string, start, end, sublength, output_format)</td>
<td>Extracts a substring based on its length and position in the parent string when server is configured to use a DBCS code page.</td>
</tr>
</tbody>
</table>

### Character Functions - Variable Length

<table>
<thead>
<tr>
<th>Function</th>
<th>Arguments</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LENV</td>
<td>(source_string, output_format)</td>
<td>Returns the actual length of an AnV input field or the size of an An field.</td>
</tr>
<tr>
<td>LOCASV</td>
<td>(upper_limit, source_string, output_format)</td>
<td>Converts alphabetic characters to lowercase and is similar to LOCASE.</td>
</tr>
<tr>
<td>POSITV</td>
<td>(source_string, upper_limit, substring, sub_limit, output_format)</td>
<td>Finds the starting position of a substring within a larger string and is similar to POSIT.</td>
</tr>
<tr>
<td>SUBSTV</td>
<td>(upper_limit, source_string, start, sub_limit, output_format)</td>
<td>Extracts a substring from a string and is similar to SUBSTR.</td>
</tr>
<tr>
<td>Function</td>
<td>Arguments</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-----------</td>
<td>-------------</td>
</tr>
<tr>
<td>TRIMV</td>
<td>(trim_where, source_string, upper_limit, pattern, pattern_limit, output_format)</td>
<td>Removes a pattern from a string and is similar to TRIM.</td>
</tr>
<tr>
<td>UPCASV</td>
<td>(upper_limit, source_string, output_format)</td>
<td>Converts alphabetic characters to uppercase and is similar to UPCASE.</td>
</tr>
</tbody>
</table>

### Data Source and Decoding Functions

<table>
<thead>
<tr>
<th>Function</th>
<th>Arguments</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>COALESCE</td>
<td>(expression1, expression2 [,,...])</td>
<td>Evaluates the arguments in order, and returns the first argument that is not MISSING.</td>
</tr>
<tr>
<td>DB_LOOKUP</td>
<td>(lookupsynonym, lookupcolumn1, sourcecolumn1, lookupcolumn2, sourcecolumn2 ... returncolumn)</td>
<td>Looks up values in a table and returns the value from another column of the row where the first value was found. For more information on improving lookup performance, see <em>Improving Lookup Performance</em> on page 498.</td>
</tr>
<tr>
<td>DB_EXPR</td>
<td>(native_SQL_expression)</td>
<td>Inserts a native SQL expression exactly as entered into the native SQL generated for a FOCUS or SQL language request.</td>
</tr>
<tr>
<td>DB_INFILE</td>
<td>(target_file, s1, t1, ... sn, tn)</td>
<td>Compares one or more field values in a source file to values in a target file.</td>
</tr>
<tr>
<td>DECODE</td>
<td>fieldname(codel 'result1'... [ELSE default])</td>
<td>Changes a variable to an associated name.</td>
</tr>
<tr>
<td>LAST</td>
<td>fieldname</td>
<td>Retrieves the previous value for a field.</td>
</tr>
<tr>
<td>Function</td>
<td>Arguments</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>NULLIF</td>
<td>(expression1, expression2)</td>
<td>Returns MISSING if the value of the two arguments are equal. Otherwise it returns the first value.</td>
</tr>
</tbody>
</table>

**Date Functions - Legacy**

<table>
<thead>
<tr>
<th>Function</th>
<th>Arguments</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AYM</td>
<td>(indate, months, output_format)</td>
<td>Adds or subtracts months from dates stored in alphanumeric or integer format.</td>
</tr>
<tr>
<td>AYMD</td>
<td>(indate, days, output_format)</td>
<td>Adds or subtracts a given number of days to a date in [YY]YYMMDD format.</td>
</tr>
<tr>
<td>CHGDAT</td>
<td>('in_display_options', 'out_display_options', date_string, output_format)</td>
<td>Rearranges the year, month, and day portions of an input character string representing a date. It may also convert the input string from long to short or short to long date representation.</td>
</tr>
<tr>
<td>DADMY</td>
<td>(indate, output_format)</td>
<td>Given the date in day-month-year format, calculates the number of days since 1/1/1900.</td>
</tr>
<tr>
<td>DADYM</td>
<td>(indate, output_format)</td>
<td>Calculates the number of days since 1/1/1900 given the date in day-year-month format.</td>
</tr>
<tr>
<td>DAMDY</td>
<td>(indate, output_format)</td>
<td>Calculates the number of days since 1/1/1900 given the date in month-day-year format.</td>
</tr>
<tr>
<td>DAMYD</td>
<td>(indate, output_format)</td>
<td>Calculates the number of days since 1/1/1900 given the date in month-year-day format.</td>
</tr>
<tr>
<td>DAYDM</td>
<td>(indate, output_format)</td>
<td>Calculates the number of days since 1/1/1900 given the date in year-day-month format.</td>
</tr>
<tr>
<td>Function</td>
<td>Arguments</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>DAYMD</td>
<td>(indate, output_format)</td>
<td>Calculates the number of days since 1/1/1900 given the date in year-month-day format.</td>
</tr>
<tr>
<td>DMY</td>
<td>(from_date, to_date)</td>
<td>Returns the difference between two dates in day-month-year order.</td>
</tr>
<tr>
<td>DOWK</td>
<td>(indate, output_format)</td>
<td>Returns the day of the week in A4 format.</td>
</tr>
<tr>
<td>DOWKL</td>
<td>(indate, output_format)</td>
<td>Returns the day of the week in A12 format.</td>
</tr>
<tr>
<td>DTDMY</td>
<td>(number, output_format)</td>
<td>Returns date in day-month-year format given the number of days since 1/1/1900.</td>
</tr>
<tr>
<td>DTDY</td>
<td>(number, output_format)</td>
<td>Given the number of days since 1/1/1900, returns date in day-year-month format.</td>
</tr>
<tr>
<td>DIMDY</td>
<td>(number, output_format)</td>
<td>Returns date in month-day-year format given the number of days since 1/1/1900.</td>
</tr>
<tr>
<td>DIMYD</td>
<td>(number, output_format)</td>
<td>Returns date in month-year-day format given the number of days since 1/1/1900.</td>
</tr>
<tr>
<td>DTYDM</td>
<td>(number, output_format)</td>
<td>Returns date in year-day-month format given the number of days since 1/1/1900.</td>
</tr>
<tr>
<td>DTYMD</td>
<td>(number, output_format)</td>
<td>Returns date in year-month-day format given the number of days since 1/1/1900.</td>
</tr>
<tr>
<td>GREGDT</td>
<td>(indate, output_format)</td>
<td>Converts a Julian date to a Gregorian date.</td>
</tr>
<tr>
<td>Function</td>
<td>Arguments</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-----------</td>
<td>-------------</td>
</tr>
<tr>
<td>JULDAT</td>
<td>(indate, output_format)</td>
<td>Converts a Gregorian date to a Julian date.</td>
</tr>
<tr>
<td>MDY</td>
<td>(from_date, to_date)</td>
<td>Returns the difference between two dates in month-day-year order.</td>
</tr>
<tr>
<td>TODAY</td>
<td>(output_format)</td>
<td>Returns the current date in alphanumeric format (MM/DD/YYYY).</td>
</tr>
<tr>
<td>YM</td>
<td>(from_date, to_date, output_format)</td>
<td>Returns the number of days between two dates.</td>
</tr>
<tr>
<td>YMD</td>
<td>(from_date, to_date)</td>
<td>Returns the difference between two dates in year-month-day order.</td>
</tr>
</tbody>
</table>

**Date Functions - Standard**

<table>
<thead>
<tr>
<th>Function</th>
<th>Arguments</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATEADD</td>
<td>(date, component_code, increment)</td>
<td>Adds or subtracts years, months, or days to or from a date.</td>
</tr>
<tr>
<td>DATECVT</td>
<td>(date, in_format, out_format)</td>
<td>Converts dates from one date format to another.</td>
</tr>
<tr>
<td>DATEDIF</td>
<td>(from_date, to_date, component_code)</td>
<td>Calculates the difference between two dates, expressed as years, months, or days.</td>
</tr>
<tr>
<td>DATEMOV</td>
<td>(date, 'move-point')</td>
<td>Moves a date to a significant point on the calendar.</td>
</tr>
<tr>
<td>DATETRAN</td>
<td>(indate, '(intype)', '(formatops)', 'lang', outlen, output)</td>
<td>Formats dates in international formats.</td>
</tr>
</tbody>
</table>

**Date-Time Functions**
<table>
<thead>
<tr>
<th>Function</th>
<th>Arguments</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTADD</td>
<td>(date, component, increment)</td>
<td>Adds or subtracts the specified number of calendar or time components to the input date or date-time fields, and returns the new date or date-time value.</td>
</tr>
<tr>
<td>DTDIFF</td>
<td>(end_date, start_date, component)</td>
<td>Given two dates in standard date or date-time formats, returns the number of a given component boundaries between the two dates.</td>
</tr>
<tr>
<td>HADD</td>
<td>(timestamp, component, increment, length, output_format)</td>
<td>Increments a date-time field by a given number of units.</td>
</tr>
<tr>
<td>HCNVRT</td>
<td>(timestamp, (format), length, output_format)</td>
<td>Converts a date-time field to alphanumeric format for use with operators, such as EDIT, CONTAINS, and LIKE.</td>
</tr>
<tr>
<td>HDATE</td>
<td>(timestamp, output_format)</td>
<td>Extracts the date portion of a date-time field and converts it to a date format.</td>
</tr>
<tr>
<td>HDIFF</td>
<td>(end_timestamp, start_timestamp, component, output_format)</td>
<td>Finds the number of boundaries of a given type crossed going from date 2 to date 1.</td>
</tr>
<tr>
<td>HDTIM</td>
<td>(date, length, output_format)</td>
<td>Converts a date field to a date-time field. The time portion is set to midnight.</td>
</tr>
<tr>
<td>HEXTR</td>
<td>(source, 'component_string', length, output_format)</td>
<td>Extracts components of a date-time value and sets remaining components to zero.</td>
</tr>
<tr>
<td>HGETC</td>
<td>(length, output_format)</td>
<td>Stores the current date and time in a date-time field.</td>
</tr>
<tr>
<td>Function</td>
<td>Arguments</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>HHMMSS</td>
<td>(output_format)</td>
<td>Returns the current time in alphanumeric format.</td>
</tr>
<tr>
<td>HINPUT</td>
<td>(source_length, source_string, timestamp_length, output_format)</td>
<td>Converts an alphanumeric string to a date-time value.</td>
</tr>
<tr>
<td>HMASK</td>
<td>(source, 'componentstring', input, length, output_format)</td>
<td>Extracts components of a date-time value and preserves remaining components.</td>
</tr>
<tr>
<td>HMINDN</td>
<td>(timestamp, length, output_format)</td>
<td>Changes the time portion of a date-time field to midnight (all zeroes).</td>
</tr>
<tr>
<td>HNAME</td>
<td>(timestamp, component, output_format)</td>
<td>Extracts a specified component from a date-time field and returns it in alphanumeric format.</td>
</tr>
<tr>
<td>HPART</td>
<td>(value, component, output_format)</td>
<td>Extracts a specified component from a date-time field and returns it in numeric format.</td>
</tr>
<tr>
<td>HSETPT</td>
<td>(timestamp, component, value, length, output_format)</td>
<td>Inserts the numeric value of a specified component into a date-time field.</td>
</tr>
<tr>
<td>HTIME</td>
<td>(length, timestamp, output_format)</td>
<td>Converts the time portion of a date-time field to a numeric number of milliseconds (if the first argument is 8) or microseconds (if the first argument is 10).</td>
</tr>
<tr>
<td>HTMTOTS</td>
<td>(time, length, output_format)</td>
<td>Converts a time to a timestamp.</td>
</tr>
</tbody>
</table>
## Format Conversion Functions

<table>
<thead>
<tr>
<th>Function</th>
<th>Arguments</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EDIT</strong></td>
<td><em>(fieldname)</em></td>
<td>Converts between alphanumeric and integer data. Used with a mask, extracts characters from an alphanumeric string and inserts new characters.</td>
</tr>
<tr>
<td><strong>FTOA</strong></td>
<td><em>(number, (format)', output_format)</em></td>
<td>Converts numbers from decimal format to alphanumeric format.</td>
</tr>
<tr>
<td><strong>ITONUM</strong></td>
<td><em>(maxbytes, infield, output_format)</em></td>
<td>Converts large binary integers in non-FOCUS files to double-precision format.</td>
</tr>
<tr>
<td><strong>ITOPACK</strong></td>
<td><em>(maxbytes, infield, output_format)</em></td>
<td>Converts large binary integers in non-FOCUS files to packed-decimal format.</td>
</tr>
<tr>
<td><strong>ITOZ</strong></td>
<td><em>(length, in_value, output_format)</em></td>
<td>Converts numbers from numeric format to zoned format for extract files.</td>
</tr>
<tr>
<td><strong>PTOA</strong></td>
<td><em>(number, (format)', output_format)</em></td>
<td>Converts a number from numeric format to alphanumeric format.</td>
</tr>
<tr>
<td><strong>TO_INTEGER</strong></td>
<td><em>(string)</em></td>
<td>Converts a character string that contains a valid number consisting of digits and an optional decimal point to an integer value.</td>
</tr>
<tr>
<td><strong>TO_NUMBER</strong></td>
<td><em>(string)</em></td>
<td>Converts a character string that contains a valid number consisting of digits and an optional decimal point to the numeric format most appropriate to the context.</td>
</tr>
<tr>
<td>Function</td>
<td>Arguments</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>TSTOPACK</td>
<td><code>(timestamp, 'P2')</code></td>
<td>Converts an MS SQL Server or Sybase column described as &quot;timestamp&quot;, which contains an increasing counter that represents a unique value in the database in hex notation with a format of A16 into a packed decimal number with a format of P21.</td>
</tr>
</tbody>
</table>

**Legacy Functions**

<table>
<thead>
<tr>
<th>Function</th>
<th>Arguments</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASIS</td>
<td><code>(argument)</code></td>
<td>Distinguishes between a space and a zero.</td>
</tr>
<tr>
<td>ATODBL</td>
<td><code>(source_string, length, output_format)</code></td>
<td>Converts a number in alphanumeric format to decimal format.</td>
</tr>
<tr>
<td>BAR</td>
<td><code>(barlength, infield, maxvalue, 'char', output_format)</code></td>
<td>Produces a bar chart.</td>
</tr>
<tr>
<td>BITSON</td>
<td><code>(bitnumber, source_string, output_format)</code></td>
<td>Evaluates an individual bit within a character string to determine whether it is on or off.</td>
</tr>
<tr>
<td>BITVAL</td>
<td><code>(source_string, startbit, number, output_format)</code></td>
<td>Evaluates a string of bits within character strings and returns its binary value.</td>
</tr>
<tr>
<td>BYTVAL</td>
<td><code>(character, output_format)</code></td>
<td>Translates a character to its corresponding ASCII code.</td>
</tr>
<tr>
<td>CHKPCK</td>
<td><code>(length, in_value, error, output_format)</code></td>
<td>Validates the data in a field described as packed format.</td>
</tr>
<tr>
<td>Function</td>
<td>Arguments</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>FIND</td>
<td>(fieldname [AS dbfield] IN file);</td>
<td>Verifies if a value exists in an indexed field in another file.</td>
</tr>
<tr>
<td>HEXBYT</td>
<td>(decimal_value, output)</td>
<td>Converts a numeric value to its corresponding ASCII character.</td>
</tr>
<tr>
<td>LOOKUP</td>
<td>field</td>
<td>Retrieves a data value from a cross-referenced FOCUS data source in a MODIFY request.</td>
</tr>
<tr>
<td>PCKOUT</td>
<td>(in_value, length, output_format)</td>
<td>Writes packed numbers of varying lengths (between one and 16 bytes) to extract files.</td>
</tr>
<tr>
<td>REVERSE</td>
<td>(length, source_string, output_format)</td>
<td>Reverses the input characters.</td>
</tr>
<tr>
<td>UFMT</td>
<td>(source_string, length, output_format)</td>
<td>Converts input field values to hexadecimal (hex) notation.</td>
</tr>
</tbody>
</table>

### Numeric Functions

<table>
<thead>
<tr>
<th>Function</th>
<th>Arguments</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td>(in_value)</td>
<td>Returns the absolute value of its argument.</td>
</tr>
<tr>
<td>CEIL</td>
<td>(number)</td>
<td>Returns the smallest integer that is greater than or equal to the number.</td>
</tr>
<tr>
<td>DMOD</td>
<td>(dividend, divisor, output_format)</td>
<td>Returns the remainder from a division as a number in decimal format.</td>
</tr>
<tr>
<td>EXP</td>
<td>(power, output_format)</td>
<td>Raises the number e to a power you specify.</td>
</tr>
<tr>
<td>Function</td>
<td>Arguments</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>EXPN</td>
<td>((n.nn \ (E</td>
<td>D) \ (+</td>
</tr>
<tr>
<td>FLOOR</td>
<td>((number))</td>
<td>Returns the largest integer that is less than or equal to the number.</td>
</tr>
<tr>
<td>FMOD</td>
<td>((dividend, divisor, output_format))</td>
<td>Returns the remainder from a division as a number in floating-point format.</td>
</tr>
<tr>
<td>IMOD</td>
<td>((dividend, divisor, output_format))</td>
<td>Returns the remainder from a division as a number in integer format.</td>
</tr>
<tr>
<td>INT</td>
<td>((in_value))</td>
<td>Returns the integer part of its argument.</td>
</tr>
<tr>
<td>LOG</td>
<td>((in_value))</td>
<td>Returns the natural logarithm of its argument.</td>
</tr>
<tr>
<td>MAX</td>
<td>((value1, value2, \ldots))</td>
<td>Returns the maximum value from its list of arguments.</td>
</tr>
<tr>
<td>MIN</td>
<td>((value1, value2, \ldots))</td>
<td>Returns the minimum value from its list of arguments.</td>
</tr>
<tr>
<td>MOD</td>
<td>((dividend, divisor))</td>
<td>Calculates the remainder from a division as an integer.</td>
</tr>
<tr>
<td>NORMSDST</td>
<td>((value, output_format))</td>
<td>Calculates the cumulative normal standard distribution.</td>
</tr>
<tr>
<td>NORMSINV</td>
<td>((value, output_format))</td>
<td>Calculates the inverse cumulative normal standard distribution.</td>
</tr>
<tr>
<td>PRDNOR</td>
<td>((seed, output_format))</td>
<td>Generates reproducible random numbers.</td>
</tr>
<tr>
<td>PRDUNI</td>
<td>((seed, output_format))</td>
<td>Generates reproducible random numbers.</td>
</tr>
<tr>
<td>Function</td>
<td>Arguments</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>RDNORM</td>
<td>(output_format)</td>
<td>Generates random numbers.</td>
</tr>
<tr>
<td>RDUNIF</td>
<td>(output_format)</td>
<td>Generates random numbers.</td>
</tr>
<tr>
<td>SQRT</td>
<td>(in_value)</td>
<td>Returns the square root of its argument.</td>
</tr>
<tr>
<td><strong>System Functions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APINPATH</td>
<td>(length, appname, 'A1')</td>
<td>Returns Y if application appname is in the path, N if not.</td>
</tr>
<tr>
<td>CLSDDREC</td>
<td>(output_format)</td>
<td>Closes all files opened by the PUTDDREC function.</td>
</tr>
<tr>
<td>FEXERR</td>
<td>(error, 'A72')</td>
<td>Returns the first line of the error message.</td>
</tr>
<tr>
<td>FGETENV</td>
<td>(length, varname, outlen, output_format)</td>
<td>Gets the value of an environmental variable and returns it as an alphanumeric string.</td>
</tr>
<tr>
<td>FPUTENV</td>
<td>(varname_length, varname, value_length, value, output_format)</td>
<td>Assigns a character string to an environmental variable.</td>
</tr>
<tr>
<td>GETUSER</td>
<td>(output_format)</td>
<td>Retrieves the user ID from the system.</td>
</tr>
<tr>
<td>PUTDDREC</td>
<td>(ddname, dd_len, record_string, record_len, output_format)</td>
<td>Writes a character string as a record in a flat file.</td>
</tr>
</tbody>
</table>
Procedure: How to Use Function Assist

When you use a function in one of the calculators, the Assist dialog box can help you set the parameters.

1. Click the Functions tab in one of the calculators.
2. Double-click a function to add it to the Expression tab.

   **Note:** Right-click the Name heading bar to display the functions by Tree View or in List View.

   The Function Assist dialog box opens.

3. Enter alphanumeric parameters in their fields.
4. Enter parameters that require column names by selecting them from the drop-down menu, as seen in the following image.

   ![Function Assist Dialog Box](image)

   In addition to the column name, the format, description, and nullability are also displayed.

5. When all the parameters in the Function Assist dialog box have been set, click OK.
The function is added to the Expression tab of the calculator, as shown in the following image.

Note: The Function Assist button enables you to edit a function in the Expression tab using the Function Assist dialog box.

Column Formats and Mapping

When you add transformations using either the Source or Target Transformations window, you can assign formats to the columns that you are adding. There are four types of column formats:

- **Alphanumeric formats.** Composed of letters, numbers, or special characters. There are three alphanumeric formats:
  - **Fixed Alphanumeric.** Used for fixed-length columns. The maximum length is 256.
  - **Variable Alphanumeric.** Used for variable-length columns. The maximum length is 4,096.
  - **Text.** Used for text columns, such as MS SQL Server VARCHAR(MAX) or an ORACLE CLOB.
• **Numeric formats.** Composed of the digits 0 through 9 and, optionally, a minus sign. There are four numeric formats:
  - **Integer.** Used for whole numbers, using 4 bytes of storage and a maximum number of 9 positions. Numbers after the decimal are truncated.
  - **Float.** Used for whole numbers or fractions, using 4 bytes of storage and a maximum number of 9 positions.
  - **Double.** Used for whole numbers or fractions, using 8 bytes of storage and a maximum number of 15 positions.
  - **Decimal Packed.** Used for whole numbers, using one byte for each two digits, plus a sign and a maximum number of 15 positions. This type of column will automatically round any number with more than the specified number of decimal places to the number of decimal places in the format. This is recommended for monetary values.

For more information about numeric formats, see *Numeric Format Options* on page 369.

• **Date and time formats.** Composed of dates and times or date and time components. There are three formats:
  - **Date.** Used for dates and provides combinations of year, month, day, and quarter.
  - **Time.** Used for times and provides combinations of hour, minute, second, and millisecond.
  - **DateTime (Timestamp).** Used for both dates and times and provides combinations of year, month, and day for dates, and hour, minute, second, millisecond, and microsecond for time.

• **BLOB.** Represents a Binary Large Object. A BLOB is used to store an image, sound, or some other large, unstructured block of data. The corresponding database data types are:
  - ORACLE - BLOB
  - MS SQL Server - Varbinary (max)
  - Informix - Byte
  - Db2 - BLOB

A column described as BLOB can only be copied from source to target. No transformations can be performed.
**Reference:** Rules for Mapping and Column Formats

Columns that you map to each other must have the same format, or DataMigrator must be able to translate the source query format into the target format automatically. Details are shown in the following table:

<table>
<thead>
<tr>
<th>If the Target format is...</th>
<th>...and the Source query format is...</th>
<th>Can you map?</th>
</tr>
</thead>
<tbody>
<tr>
<td>A11</td>
<td>A10</td>
<td>Yes</td>
</tr>
<tr>
<td>A10</td>
<td>A11</td>
<td>No. Use the EDIT or SUBSTR functions to make this work.</td>
</tr>
<tr>
<td>D6.2</td>
<td>I5</td>
<td>Yes</td>
</tr>
<tr>
<td>I10</td>
<td>D6.2</td>
<td>Yes, but you will lose the numbers after the decimal.</td>
</tr>
<tr>
<td>A10</td>
<td>D6.2</td>
<td>No</td>
</tr>
</tbody>
</table>

**Procedure:** How to Format a Column as Alphanumeric or Numeric

If you are creating a new or virtual column, you need to assign a format to the column. You can assign a format in two ways:

- By typing the format into the Format field in the Transformation Calculator.
- By using the Define Properties dialog box.

To use the Define Properties dialog box:

1. From the Transformation Calculator, click the ellipsis (...) button to the right of the Format field.
   - The Define Properties dialog box opens.
2. Select a format from the Type drop-down menu.
3. Click OK.
Reference: Define Properties Dialog Box

To access the Define Properties dialog box from the Transformation Calculator, click the ellipsis (...) button to the right of the Format field.

The Define Properties dialog box has the following fields and options:

**DEFINE**
- An optional attribute that is used to create a virtual field for reporting.

**FORMAT**
- The format name of the column.

**Type**
- The format type of the column.

**Note:** The options available will depend on the type of format selected.
MISSING

Indicates if the results can be missing. The following options are available:

☐ OFF. The result is never missing.

☐ ON. The result is missing only if all missable fields are missing.

☐ ON ALL. The result is missing if any missable fields are missing.

REDEFINES

Lets you redefine or recompute a field whose name exists in more than one segment.

WITH

Lets you identify the logical home of the defined calculation, if no fields from the synonym are used in the expression or have been defined.

DESCRIPTION

An optional attribute that enables you to provide comments and other documentation for a field within the Master File. Comments can be up to 2048 characters in length.

Reference: Numeric Format Options

When using a numeric format (integer, float, double, or decimal packed), the following options are available:

Negative

These are the display options for negative numbers:

Default - Minus sign at the left

Positions the minus (-) sign at the left of the number.

- - Minus sign at the right

Positions the minus (-) sign at the right of the number.

B - Bracket negative

Encloses a negative number in parentheses.

C - Credit (CR) negative

Adds the characters CR after a negative number.
Comma

These are the display options for commas:

Default - Suppress

Suppresses the use of a comma in the number.

c - Suppress

Suppresses the use of a comma in the number.

C - Insert

Inserts a comma after every third significant figure. This option takes effect automatically for Double format.

L - Leading zeroes

Adds leading zeros to the full field length.

S - Print blank for zero

Suppresses leading zeros; displays a space if the value is 0.

% - Percent Sign

Displays a percent sign along with the numeric data.

E - Scientific notation

Displays the number in scientific notation.

Currency Symbol

These are the extended currency symbol display options, which allow you to select a currency symbol regardless of what is set as the default:

- N - Fixed $ sign
- M - Floating $ sign
- !e - Fixed euro symbol
- !E - Floating euro symbol
- !l - Fixed pound sterling sign
- !L - Floating pound sterling sign
- !y - Fixed Japanese yen symbol
- !Y - Floating Japanese yen symbol
Currency

Specifies a field as a currency denominated value.

The Define Properties dialog box for the Numeric format is shown in the image below:
**Procedure:** How to Format a Column as Date, Time, or DateTime (Timestamp)

1. In the Transformation Calculator, click the ellipsis (...) button to the right of the Format field.
   The Define Properties dialog box opens.
2. Select Date, Time, or Date and Time from the FORMAT drop-down menu.
   The Date format allows you to choose the display for dates (YYMD, MDY, and so on).
   The Time format allows you to choose the display for time (seconds, milliseconds, and so on) and am/pm.
   The Date and Time format allows you to choose the display for both date and time portions.

**Reference:** Date Type Format Options

The following choices are available from the Define Properties dialog box for the date format (the examples illustrate August 21, 2004):

<table>
<thead>
<tr>
<th>Format</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>YYMD</td>
<td>Four-digit year/month/day.</td>
<td>2004/08/21</td>
</tr>
<tr>
<td>MDYY</td>
<td>Month/day/four-digit year.</td>
<td>08/21/2004</td>
</tr>
<tr>
<td>DMYY</td>
<td>Day/month/four-digit year.</td>
<td>21/08/2004</td>
</tr>
<tr>
<td>YYQ</td>
<td>Four-digit year/quarter</td>
<td>2004 Q4</td>
</tr>
<tr>
<td>QYY</td>
<td>Quarter/four-digit year.</td>
<td>Q4 2004</td>
</tr>
<tr>
<td>YYM</td>
<td>Four-digit year/month.</td>
<td>2004/08</td>
</tr>
<tr>
<td>MYY</td>
<td>Month/four-digit year.</td>
<td>08/2004</td>
</tr>
<tr>
<td>YYJUL</td>
<td>Four-digit year with the month in Julian format. The last three digits are the number of days since January 1.</td>
<td>2004/234</td>
</tr>
<tr>
<td>YY</td>
<td>Four-digit year.</td>
<td>2004</td>
</tr>
<tr>
<td>Y</td>
<td>Two-digit year.</td>
<td>04</td>
</tr>
<tr>
<td>Format</td>
<td>Description</td>
<td>Example</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>Q</td>
<td>Quarter.</td>
<td>Q3</td>
</tr>
<tr>
<td>M</td>
<td>Month.</td>
<td>08</td>
</tr>
<tr>
<td>MT</td>
<td>Short month in uppercase.</td>
<td>AUG</td>
</tr>
<tr>
<td>Mt</td>
<td>Short month in mixed-case.</td>
<td>Aug</td>
</tr>
<tr>
<td>MTR</td>
<td>Full month in uppercase.</td>
<td>AUGUST</td>
</tr>
<tr>
<td>Mtr</td>
<td>Full month in mixed-case.</td>
<td>August</td>
</tr>
<tr>
<td>W</td>
<td>Three-letter day of the week abbreviation in uppercase.</td>
<td>SAT</td>
</tr>
<tr>
<td>w</td>
<td>Three-letter day of the week abbreviation in mixed-case.</td>
<td>Sat</td>
</tr>
<tr>
<td>WR</td>
<td>Day of the week in uppercase.</td>
<td>SATURDAY</td>
</tr>
<tr>
<td>wr</td>
<td>Day of the week in mixed-case.</td>
<td>Saturday</td>
</tr>
<tr>
<td>D</td>
<td>Day.</td>
<td>21</td>
</tr>
</tbody>
</table>

**Reference:**  
**Time Type Formats**

The following choices are available from the Define Properties dialog box for the time format (the example illustrates 2:05 A.M.):

<table>
<thead>
<tr>
<th>Format</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH</td>
<td>Hours.</td>
<td>02</td>
</tr>
<tr>
<td>HHI</td>
<td>Hours/Minutes.</td>
<td>02:05</td>
</tr>
<tr>
<td>HHIS</td>
<td>Hours/Minutes/Seconds.</td>
<td>02:05:27</td>
</tr>
<tr>
<td>HHISs</td>
<td>Hours/Minutes/Seconds/Milliseconds.</td>
<td>02:05:27:123</td>
</tr>
<tr>
<td>HHISsm</td>
<td>Hours/Minutes/Seconds/Milliseconds/Microseconds.</td>
<td>02:05:27:123456</td>
</tr>
<tr>
<td>Format</td>
<td>Description</td>
<td>Example</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>HI</td>
<td>Minutes.</td>
<td>05</td>
</tr>
<tr>
<td>HIS</td>
<td>Minutes/Seconds.</td>
<td>05:27</td>
</tr>
<tr>
<td>HISs</td>
<td>Minutes/Seconds/Milliseconds.</td>
<td>05:27:123</td>
</tr>
<tr>
<td>HISsm</td>
<td>Minutes/Seconds/Milliseconds/Microseconds.</td>
<td>05:27:123456</td>
</tr>
<tr>
<td>HS</td>
<td>Seconds</td>
<td>27</td>
</tr>
<tr>
<td>HSs</td>
<td>Seconds/Milliseconds.</td>
<td>27:123</td>
</tr>
<tr>
<td>HSsm</td>
<td>Seconds/Milliseconds/Microseconds.</td>
<td>27:123456</td>
</tr>
<tr>
<td>Hh</td>
<td>Hours without leading zeros.</td>
<td>2</td>
</tr>
<tr>
<td>HhI</td>
<td>Hours without leading zeros/Minutes.</td>
<td>2:05</td>
</tr>
<tr>
<td>HhIS</td>
<td>Hours without leading zeros/Minutes/Seconds.</td>
<td>2:05:27</td>
</tr>
<tr>
<td>HhISs</td>
<td>Hours without leading zeros/Minutes/Seconds/Milliseconds.</td>
<td>2:05:27:123</td>
</tr>
<tr>
<td>HhISsm</td>
<td>Hours without leading zeros/Minutes/Seconds/Milliseconds/Microseconds.</td>
<td>2:05:27:123456</td>
</tr>
<tr>
<td>Hi</td>
<td>Minutes without leading zeros.</td>
<td>5</td>
</tr>
<tr>
<td>HiS</td>
<td>Minutes without leading zeros/Seconds.</td>
<td>5:27</td>
</tr>
<tr>
<td>HiSs</td>
<td>Minutes without leading zeros/Seconds/Milliseconds.</td>
<td>5:27:123</td>
</tr>
<tr>
<td>HiSsm</td>
<td>Minutes without leading zeros/Seconds/Milliseconds/Microseconds.</td>
<td>5:27:123456</td>
</tr>
</tbody>
</table>

**Reference:** Date and Time Type Formats

The Date and Time type format combines date, time, and a separator, which are reflected in the format name. For example, HYYMD-H describes a four-digit year, month, and day separated by dashes plus hours.
The following choices are available from the Define Properties dialog box for the date component:

<table>
<thead>
<tr>
<th>Format</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>HYYMD</td>
<td>Four-digit Year/Month/Day.</td>
<td>2004/08/21</td>
</tr>
<tr>
<td>HYMD</td>
<td>Two-digit Year/Month.</td>
<td>04/08</td>
</tr>
<tr>
<td>HMD</td>
<td>Month/Day.</td>
<td>08/21</td>
</tr>
<tr>
<td>HD</td>
<td>Day</td>
<td>21</td>
</tr>
</tbody>
</table>

The time component is added to the date format name as follows:

<table>
<thead>
<tr>
<th>Format</th>
<th>Description</th>
<th>Format Name</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>Hours.</td>
<td>HYYMDH</td>
<td>2004/08/21 02</td>
</tr>
<tr>
<td>I</td>
<td>Hours/Minutes.</td>
<td>HYYMDI</td>
<td>2004/08/21 02:05</td>
</tr>
<tr>
<td>S</td>
<td>Hours/Minutes/Seconds.</td>
<td>HYYMDS</td>
<td>2004/08/21 02:05:27</td>
</tr>
<tr>
<td>s</td>
<td>Hours/Minutes/Seconds/Milliseconds.</td>
<td>HYYMDs</td>
<td>2004/08/21 02:05:27:123</td>
</tr>
<tr>
<td>m</td>
<td>Hours/Minutes/Seconds/Milliseconds/Microseconds.</td>
<td>HYYMDm</td>
<td>2004/08/21 02:05:27:123456</td>
</tr>
</tbody>
</table>

The following choices are available from the Define Properties dialog box for the separator, which is added to the format name.

<table>
<thead>
<tr>
<th>Format</th>
<th>Description</th>
<th>Format Name</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>/</td>
<td>Slash.</td>
<td>HYYMDH</td>
<td>2004/08/21 02</td>
</tr>
<tr>
<td>.</td>
<td>Period.</td>
<td>HYYMD.H</td>
<td>2004.08.21 02</td>
</tr>
<tr>
<td>-</td>
<td>Dash.</td>
<td>HYYMD-H</td>
<td>2004-08-21 02</td>
</tr>
</tbody>
</table>
### Setting Properties of a Flow

The properties of a flow allow you to view or set properties that apply to the entire flow. These include:

- **General properties**, including who created the flow, the last time it was updated, a description, and comments.

- **Execution properties**, including stop processing options and the scheduler status. (Scheduler status is used for process flows.)

- **Restart properties**, including the number of restart attempts and their start points. A restart attempt occurs if the flow execution fails. Restart attempts can be made from the first record or from the last commit point.

- **Email notification properties**, including setting different criteria for sending email to a specific address.

- **Comments**, which can be helpful as reminders, or to pass information among a group of users working on the same flows.

For more information on how to set logging properties for a flow, see *Setting Flow Properties* on page 161.

### Running a Data Flow

There are two main ways to run a data flow from the DMC: Run and Submit.

**Run** will run the data flow immediately. When you run a flow, processing information is displayed in the Console Log.

**Submit** will submit the data flow to the scheduler to run. When you submit a flow, the processing information is recorded in the DataMigrator Log and Statistics tables. Email messages are also sent if the flow properties are set to send email messages at the start or end of the job.
Procedure: How to Run a Data Flow

On the Flow tab, in the Run group, click Run, and select one of the run options available.

or

Right-click a data flow in the navigation pane and select one of the Run options available.

Note: When you submit the flow, processing information is recorded in the DM log and, if specified, email messages are sent. When you run the flow, the processing information is only displayed in the Console Log. It is not written to the DM log.

For more information on the Run options, see Navigation Pane: Procedures on page 195.

Creating a Direct Load Flow

A Direct Load flow is a simplified flow that can be used for a single source. It loads data directly from the file into an existing target, without the need to create an intermediate file. This occurs unless Bulk Load from Disk File is used, which always creates an intermediate file.

Note: The synonym that describes the source table for a Direct Load Flow must contain a unique ALIAS that matches the FIELD name for every column that is referenced.

A Direct Load flow can load from hierarchical data sources, such as VSAM. For example, a file with multiple record types can be used as a source with fields from each record type loaded into a separate table.

A Direct Load flow has no SQL object, so no aggregation (group by), filtering (where conditions), or SQL calculations can be performed.

Source transformations and virtual columns in the source synonym cannot be used.

Target transformations, functions, variables, validates, target properties, and record logging can be used. The procedures for using these capabilities, as well as for running the flow, are the same for Direct Load flows as for more typical data flows. For information, see the appropriate topics earlier in this chapter.

While other types of data flows discard records with format errors, when a direct load flow runs, any records rejected due to format errors are counted, and the number of rejected rows appears in the log. Optionally, the rejected records can be written to a log file.

For information, see How to Set the Data Flow Record Logging Options on page 307.
**Note:** Support for using Direct Load Flows with a source and target tables in the same relational database, and using insert records from memory, depends on the characteristics of the database. For MS SQL Server JDBC, ORACLE, and Teradata, there are no restrictions. For MS SQL Server OLE DB, Sybase, and Informix, use a Commit number greater than the number of rows to be loaded. For Db2, use separate connections for source and target.

**Procedure:** How to Create a Direct Load Flow

1. Right-click the application directory where you want the new data flow, point to New, and then click Direct Load Flow.
   
   An empty flow appears in the data flow workspace. The Direct Load Flow tab opens by default.
   
   **Note:** An SQL object does not appear in the center of the workspace.

2. Drag a data source into the workspace to the left of the center line.

3. Drag a data target into the workspace to the right of the center line. You can also right-click in the workspace to the right of the center line, point to Add Target, and then click Existing.

4. Set up target transformations. For information, see *Mapping Transformations* on page 324.

5. Right-click a target object and click Properties. Select the load type from the drop-down menu.

6. Click the Save button on the Quick Access Toolbar.

7. Enter a name for the flow in the Name field and click OK. You can also select a Server and Application Directory from the drop-down menus.

8. Run the flow.
An example of a Direct Load flow is shown in the following image.

Creating a DBMS SQL Flow

A DBMS SQL flow allows you to enter dialect-specific SQL SELECT statements for extracting data from a relational database. Dialect-specific SQL is the SQL specific to one particular relational database, such as ORACLE or MS SQL Server.

You can type in the SELECT statement, or import it from a pre-existing .sql file.

A DBMS SQL flow only contains a specialized SQL object where you select the DBMS adapter and connection, and one or more target objects. Synonyms for sources are neither required nor used. Tables must be identified using the same names as in the database.

While source transformations cannot be used, target transformations, functions, variables, validates, target properties, and record logging can be used. The procedures for using these capabilities are the same for DBMS SQL flows as for more typical data flows.

For information on these procedures, see the appropriate topics earlier in this chapter.

Procedure: How to Create a DBMS SQL Flow

The following procedure uses the Microsoft SQL Server pubs.dbo.authors sample table, and the Microsoft SQL Server difference scalar function, which is not available in ANSI SQL.

1. Right-click the application directory where you want the new data flow, point to New, and then click DBMS SQL Flow.
An empty flow appears in the data flow workspace. The DBMS SQL Flow tab opens by default.

A DBMS SQL Object appears in the center of the workspace.

2. Right-click the DBMS SQL object and click Properties.

3. Select an adapter and connection from the drop-down menus and close the Properties panel.

4. Right-click the DBMS SQL object and click SQL Statement.

   The Select Statement dialog box opens.

5. Type a dialect-specific SQL Statement, as shown in the following Select Statement dialog box.

   ```sql
   SELECT
     au_fkey, au_lkey
   FROM
     pubs.dbo.authors
   WHERE
     difference (au_lkey, 'delay') > 3
   ```
6. Test your SQL by clicking the Test SQL Statement button. The results are shown in the following image.

<table>
<thead>
<tr>
<th></th>
<th>au_fname</th>
<th>au_iname</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Innes</td>
<td>del Castillo</td>
</tr>
<tr>
<td>2</td>
<td>Ann</td>
<td>Dull</td>
</tr>
</tbody>
</table>

7. Click Close. Then click OK.

**Note:** The SQL statement must be a SELECT statement that conforms to the dialect-specific requirements of the adapter.

8. Optionally, you can import an SQL statement from an existing .sql file by clicking the Import button, pointing to an SQL file, and clicking the Select button. The SQL file must be in one of your Application Directories.

9. Optionally, you can export SQL statements as .sql files for later use by clicking the Export button. Select a Server and Application Directory from the drop-down menus, enter a File Name, and click OK.

10. Drag a data target into the workspace. You can also right-click in the workspace, point to Add Target, then Existing or Add Target, and then click New.

11. Click the Save button on the Quick Access Toolbar.

12. Enter a name for the flow in the File Name field and click OK. You can also select a Server and Application Directory from the drop-down menus.

13. Run the flow.

**Creating a Quick Copy**

The Quick Copy option generates a simple Data Flow without using the Data Flow designer. For the selected tables, all selected rows and columns are copied to new tables.

Although no transformations or where conditions are generated, the Data Flow can be opened as such from the Data Management Console, and additional criteria can be added. Columns identified as key columns in the source are created as key columns in the target.

**Procedure:** How to Create a Quick Copy for a Single Table

1. Right-click a synonym name and click Quick Copy.

The Quick Copy dialog box opens. The following options are available:

**Load Option**

- **New.** Any existing metadata or data is replaced.
Append Existing Data (Keep Metadata). The source data is appended to the target.

Adapter

Use the drop-down menu to select a database configured as an adapter.

Connection

Use the drop-down menu to select a connection for the selected adapter.

Synonym Application

The application directory where the target tables and generated flow will be stored.

Synonym

The synonym name for the target table. By default, this is the same name as the source table unless the Application directory is the same, in which case _t is added to the name.

Table Name

The name of the target table name in the database. By default, this is the same name as the source table, unless the Adapter and Connection are the same as the source, in which case _t is added to the name.

Bulk Load

Specifies the method of loading data. When selected, the database vendors bulk load program is used, instead of inserting records from memory.

2. In the Target Synonym Application box, select the application directory where to save the flow by either typing in the location of the directory, or clicking the ellipses and manually selecting it from the Select Application dialog box.

3. Click Save or Run. Two flows will be created, a process flow with the name you specified, and a data flow with the name followed by the table name, for example, flow01_dmhr. The process flow calls the data flow to copy the table.

Procedure: How to Create a Quick Copy for Multiple Tables

1. Right-click an application directory and click Quick Copy.
   The Quick Copy dialog box opens.

2. Select the tables you want to copy by selecting the check box in the first column. Doing so allows you to change the Name, which is the synonym, and the Table Name, which is the name of the table in the database.

   Note: If you select the Parallel Execution checkbox, the flows for each table will be run in parallel.
3. Click Save or Run. Two or more flows will be created, a process flow with the name you specified, and a data flow with the name followed by the table names for each synonym you have selected. For example, flow01_dmreps. The process flow calls all of the data flows.
Designing a Process Flow

A process flow defines a set of steps that execute a series of components or objects, including the data flows that define your extractions, transformation, and load logic. You can combine:

- A Start object with a schedule that defines when to run the process.
- Data Flow objects that indicate what data to extract and move from data sources to data targets.
- Email objects that notify users about the status of the process at specified points.
- Procedure objects that perform a variety of supplementary tasks.
- Connector objects, represented as arrows, that specify execution logic for the other objects included in the process flow.
- Variable objects to set the values for variables used in other objects.
- Group objects that specify the flow of a subset of objects within the total flow.

In this chapter:

- Process Flow Methodology
- Process Flow Interface
- Simple Process Flow At a Glance
- Complex Process Flow at a Glance
- Creating a Process Flow
- Adding a Saved Procedure to a Process Flow
- Assigning and Using Variables in a Process Flow
- Maintaining Continuous Processing in a Single Agent
- Executing a Process Flow That Contains Only Procedures
- Adding Email Messages to a Process Flow
- Setting Connection Logic
- Grouping and Controlling Simultaneous Execution of Flows
- Running a Process Flow
- Scheduling a Process Flow
Process Flow Methodology

You can create a process flow either with or without a data flow. You can open a data flow from within a process flow by right-clicking it and selecting Open.

Creating a Process Flow With a Data Flow

Although you can build a simple process flow as an extension to an existing data flow, it is recommended that you create and save each flow and procedure separately. With this modular approach, individual flows (both data and process flows) and procedures can be executed from more than one process flow, and a process flow can be modified as needed and used to incorporate different data flow and procedure objects. A change to a data flow or procedure automatically affects all of the process flows that use it.

If you click the Text View tab in a process flow created with this method, any data flows or procedures used in the process flow appear as EX procedure_name. If you want to view or edit the underlying code for an individual procedure, you can right-click the object and click Edit as Text.

However, if you want to add a few process flow objects, such as a schedule or an email, to a data flow, and do not require the advantages of a modular design, save the data flow and click the Process Flow tab, and then add the process flow components to the existing data flow.

Note: If you click the Text View tab in a process flow created with this method, the underlying data flow and process flow syntax appears fully, since it is part of the same procedure.

Creating a Process Flow Without a Data Flow

Normally, process flows manage extraction, transformation, and load processing. However, process flows are not limited to this functionality. Under some circumstances, you may wish to create process flows that do not contain any data flows.

Using the same set of design tools, you can, for example, create a process flow that launches a series of reporting procedures on a defined schedule, and then sends email messages to appropriate recipients when the procedures are complete.

Process Flow Interface

The Process Flow interface consists of:

- The Process Flow tab and workspace, where you build your process flows.
- The active Flow tab, which enables you to add objects to your process flow.
When you create a new flow and click the Process Flow tab in the workspace, a Start object appears in the workspace to anchor the flow, as shown in the following image.

You can add a variety of other objects from the Insert group. These may include data flows or other process flows, as single components or as a group, procedures, and email messages, in whatever combination and order suits your needs.

For more information on the Process Flow Interface, see *DataMigrator Workspace* on page 138.

The remainder of this chapter describes, in detail, all the options that are available to you as you design your process flows.

For more information on using the Procedures button, see *Navigation Pane: Procedures* on page 195.
Simple Process Flow At a Glance

You can use DataMigrator to design process flows ranging from simple, where you use only a few process flow components, to complex, where you take advantage of a wide range of components.

In the example shown in the following image, the process flow includes a data flow that is scheduled to run weekly, on Thursdays, starting at midnight. At the end of each run, an email message is sent to a designated recipient to indicate the successful completion of the data flow processing.

To create a simple process flow:

1. Open a new flow and click the Process Flow tab.
   The Start object appears by default.

2. To add a data flow object, drag the data flow from the navigation pane into the workspace.

3. To add an email object and create a message, ensure that your server is configured for email notification.

4. To add a Schedule, specify when and how often to run the process flow in the Start object properties. For more information, see Scheduling a Process Flow on page 423.

5. Connect the objects.
6. Save and run the process flow.
7. View the log file for flow processing details.

Complex Process Flow at a Glance

You can design a complex process flow by combining a wider range of objects in the workspace. You can run procedures before and after executing flows, execute flows concurrently or individually, and send email messages based on the outcome of the execution.

In the process flow shown in the following image, a setup procedure prepares the environment for data extraction. Two data flows are grouped to run simultaneously, followed by a Wait object. When the processing of both data flows is complete, one of two email messages is sent: one if the previous step is successful, the other if the previous step is unsuccessful. After a successful run, a cleanup procedure is executed to prepare for the next scheduled process flow run.

To create a complex process flow:
1. Open a new flow and click the Process Flow tab.

   The Start object appears by default.

2. Add a procedure to set up the environment prior to data extraction, and connect it to the Start object.

   This procedure runs at the start of the process flow.

3. Add a Parallel Group object for concurrent processing of two data flows, and connect it to the procedure.

   This enables you to run multiple data flows simultaneously.
**Tip:** Right-clicking the Parallel Group line opens a context menu, which allows you to delete the whole group, or select all objects in the group and move them. You can then either delete the frame or add other objects to the group.

4. Add the data flows to the Parallel Group object.
   These data flows run simultaneously.

5. Add a Wait object to ensure that the next step does not begin until all of the data flows in the Parallel group are complete.

6. Connect the Group object to the Wait object.

7. Add another Data Flow object to be executed when the Wait conditions are satisfied.

8. Connect the Data Flow object to the Wait object.

9. Create two email messages and connect both to the Data Flow object, forming two branches in the process flow.

10. Send the messages to appropriate recipients when the data run succeeds or fails.

11. Add a cleanup procedure that runs after the data flows, and that connects to the email object upon successful completion the flow.

12. Open the Start object and specify when and how often to run the flow.

13. Save and run the process flow.

14. View the log file for flow processing.
Creating a Process Flow

You can create a process flow either by starting a new flow (the preferred method), or by building on an existing data flow (which is suitable for some simple situations). In either case, you add objects to the process flow base, which varies slightly depending upon the method you choose. For more information about these approaches, see Process Flow Methodology on page 386.

If you create a new flow, the process flow workspace initially contains only the Start object, as shown in the following image. You can then begin to add other objects to the flow.

For details, see How to Create a Reusable Process Flow on page 392.
If you build upon an existing data flow, the process flow workspace initially contains the Start object connected to the embedded data flow object, as shown in the following image. You can then add others to the flow.

![Diagram showing a Start object connected to a Data Flow object.](image)

For details, see *How to Create a Process Flow From a Data Flow* on page 393.

With either method, the Start object can only have one object as output. It can be any object type except for a Wait object, which can only follow a Parallel Group object.

The connector arrow from the Start object to the next object in the flow will always be black (that is, unconditional) and cannot be changed.

**Procedure: How to Create a Reusable Process Flow**

This procedure describes the recommended modular process flow design methodology described in *Process Flow Methodology* on page 386. To create a reusable process flow:

1. Right-click an application directory in the navigation pane, point to New, and then click Flow.
2. Click the Process Flow tab. A new workspace containing a Start object opens in the workspace.
3. To add objects to the workspace:
   - Drag a data flow, another process flow, or a procedure into the workspace from the navigation pane. You can open another flow within the process flow by right-clicking it and clicking Open.
Click an object from the Insert group and drag it to the workspace. Then double-click the object to open a properties specification pane. See the detailed topics in this chapter for information on property specifications for each type of object.

4. To connect objects in the workspace, right-click the left object and, holding the mouse button down, drag from the left object to the right object, creating a connecting arrow. Release the mouse button.

5. When you have finished building the process flow, save and name it.
   a. Click the Save or Save As button from the Quick Access Toolbar. The Save Procedure As dialog box opens.
   b. Name the process flow.
   c. Save it to any application directory in your application path.
   d. Click OK to return to the Process Flow workspace.

Alternatively, you can save and name the process flow at any point during development, and then continue to click Save to update the flow as you build it.

**Procedure:** How to Create a Process Flow From a Data Flow

Use the following steps only when you wish to create a process flow that contains a limited number of objects, and when you do not intend to use the embedded data flow with other process flows.

1. After creating and saving a data flow, click the Process Flow tab. The Start object and the embedded Data Flow object are connected by default.

2. To add new objects to the workspace, you can do either or both of the following:
   - Drag a data flow, another process flow, or a procedure into the workspace from the navigation pane. A process flow can be comprised of data flows, process flows, and/or procedures from any application directory in the search path of the user.
   - Click an object in the Insert group and drag it into the workspace. Then double-click the object to open a properties specification pane. See the detailed topics in this chapter for information on property specifications for each type of object.

3. To connect objects in the workspace, right-click the left object and, holding the mouse button down, drag from the left object to the right object, creating a connecting arrow. Release the mouse button.

   **Tip:** If you want to add an object between the Start object and Data Flow object, right-click the connecting arrow and choose Delete. Then add the new object. Connect the Start object to the new object, and then connect the new object to the Data Flow object.
4. When you have finished building the process flow, save it under the data flow procedure name. Both the data flow and process flow logic are contained in the same procedure. Alternatively, you can save the process flow at any point in the development cycle. For an illustration of this technique, see Navigation Pane: Procedures on page 195.

Saving a Process Flow

Process flows are saved on the DataMigrator Server. The first time you save a process flow, you will select an application directory and assign a name to it. After that, you can save it with the same name or assign a new one.

The following restrictions apply to flow names:

- Names can contain only letters, numbers, and underscores.
- Uppercase letters are converted to lowercase when the flow is saved.
- Names can be up to 64 characters long, except on z/OS with PDS deployment, where names are limited to 8 characters.

**Procedure:** How to Save a Process Flow

1. Click the Save button from the Quick Access Toolbar.

   **Note:** A flow can only be saved once changes have been made to it. If no changes have been made, both the button and Save option in the main menu will be inactive (grey).

   If this is the first time a flow is being saved, the Save Procedure As dialog box will open. See How to Save a Data Flow with a New Name on page 213.

2. Check the console window for a message indicating that the flow was saved successfully.

**Procedure:** How to Save a Process Flow with a New Name

1. Click the Save As button from the Quick Access Toolbar.

   The Save Procedure As dialog box opens.

2. From the Save in: drop-down menu, select the application directory where you want to save the flow.

3. In the Name box, enter the name you wish to use for the flow.

   Click Save.

4. Check the console window for a message indicating that the flow was saved successfully.
Adding a Saved Procedure to a Process Flow

You can add any previously saved procedure to the process flow you are currently designing. The procedure may be:

- A data flow. For details about creating and saving data flows, see Designing a Data Flow on page 193.
- Another process flow. The current chapter describes all aspects of this task.
- A procedure. For details, see How to Create a Procedure on page 396.

**Procedure: How to Add a Flow to the Process Flow**

1. Right-click an application directory in the navigation pane, point to New, and then click Flow.
2. Click the Process Flow tab. A new workspace containing a Start object opens in the right pane.
3. From the navigation pane, select a flow and drag it into the process flow workspace. The flow may be a data flow or another process flow.
4. Right-click the object to the left of the flow and, holding the mouse button down, drag from the left to the right object, creating a connecting arrow. Release the mouse button.

   The procedure will be executed within the process flow at the point where it is encountered.
5. Click the Save button to update the process flow.

**Procedure: How to Add a Procedure to the Process Flow**

1. Right-click an application directory in the navigation pane, point to New, and then click Flow.
2. Click the Process Flow tab. A new workspace containing a Start object opens in the right pane.
3. From the navigation pane, select a procedure and drag it into the process flow workspace. For related information, see Procedures and Flows on page 396.
4. Right-click the object to the left of the procedure and, holding the mouse button down, drag from the left object to the right object, creating a connecting arrow. Release the mouse button.

   The procedure will be executed within the process flow at the point where it is encountered.
5. Click the Save button to update the process flow.
Adding a Saved Procedure to a Process Flow

Reference: Procedures and Flows

In the course of developing a flow (either a data flow or a process flow), you may need to perform a variety of specialized tasks, from setting up the environment parameters and variables, to copying files. The DMC provides an editing window in which you can type your 4GL code or SQL. When you save the file, it is stored in an application directory, identified by the icon.

Then you can:

- Add it as an object in a process flow.
  
  In fact, you can create a process flow composed of procedures and other objects, such as email messages, without any data flows. See Executing a Process Flow That Contains Only Procedures on page 409.

- Use it in the design and testing of a data flow. For details, see Using Variables in a Flow on page 336.

Procedure: How to Create a Procedure

To create a procedure:

1. Right-click an application directory in the navigation pane, point to New, and then click Procedure.

2. A text editing window opens. Type the procedure code, which may be either FOCUS or SQL.

   For example, create a procedure named PRODPARMS that has two parameters, &PTYPE and &CAT, as shown in the following image.

   ![Procedure code example]

   You are responsible for the accuracy of the syntax you enter. However, any errors that are generated will be displayed in the Server Messages tab below the workspace when you execute the procedure.
3. Click the Save or Save As button from the Quick Access Toolbar. Enter a procedure name in the Save Procedure As dialog box.

The procedure is listed in the navigation pane under an application directory, where it is preceded by the Procedure icon.

From here you can drag it into a process flow at any point. For details, see Adding a Saved Procedure to a Process Flow on page 395.

Assigning and Using Variables in a Process Flow

Variables allow you to control the execution of a process flow by substituting values at runtime. A Set Variables object lets you set the value for a variable to be used in other objects. These variables store a string of text or numbers that can be used in the process flow, passed as parameter to a data flow, or used in a procedure.

Some common uses of variables are in a SELECT statement to add a constant value to an answer set or in a WHERE condition to restrict the rows selected. Another use of a variable is to assign a name for a target table.

Procedure: How to Set a Value for a Variable

1. Right-click an application directory, point to New, and then click Flow.

A data flow window opens.

2. Click the Process Flow tab to switch to that view.

3. On the Flow tab, in the Insert group, drag the Set Variables button onto the workspace.

4. Right-click the Start icon, drag the cursor to the Set Variables icon, and release.

The workspace should resemble the following image.

5. Right-click the Set Variables object and click Variables Editor.

The Set Variables Editor opens.

6. Click the Add SET Parameter button.
The User Variables Calculator opens.

7. Enter a name and value for the variable, as shown in the following image.

8. Click OK.
The Set Variables Editor now displays the details for the parameter, as shown in the following image.

9. Click OK to close the window.

The global variable &&RUNYEAR can now be used in any procedure added to the process flow after the Set Variables object. It can also be used in an embedded (purple) data flow, or in a data flow that is added to the process flow, provided that the Run as procedure option is selected in the properties page for that data flow.

Reference: Set Variables Editor Toolbar

The Set Variables Editor allows you to assign values to variables.

The Set Variables Editor toolbar contains the following buttons:

<table>
<thead>
<tr>
<th>Button</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td>Add DEFAULT Parameter</td>
</tr>
<tr>
<td><img src="image2.png" alt="Image" /></td>
<td>Adds a DEFAULT parameter.</td>
</tr>
<tr>
<td><img src="image3.png" alt="Image" /></td>
<td>Add SET Parameter</td>
</tr>
<tr>
<td><img src="image4.png" alt="Image" /></td>
<td>Opens the User Variables Calculator to assign a value to a parameter.</td>
</tr>
<tr>
<td>Button</td>
<td>Function</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td><img src="image" alt="Delete Parameter" /></td>
<td><strong>Delete Parameter</strong></td>
</tr>
<tr>
<td></td>
<td>Deletes the currently selected parameter.</td>
</tr>
<tr>
<td><img src="image" alt="Edit Selected Parameter" /></td>
<td><strong>Edit Selected Parameter</strong></td>
</tr>
<tr>
<td></td>
<td>Opens the User Variables Calculator for the currently selected parameter.</td>
</tr>
<tr>
<td><img src="image" alt="Test Variables" /></td>
<td><strong>Test Variables</strong></td>
</tr>
<tr>
<td></td>
<td>Tests the variable assignments.</td>
</tr>
</tbody>
</table>

The Set Variables Editor also contains the following input areas:

**Name**

Name of the variable.

**Command**

Provides a drop-down menu with the options of SET or DEFAULT.

**Scope**

Provides a drop-down menu with the options of LOCAL or GLOBAL.

**Expression**

Value to be assigned to the variable.

**Reference:** [Variable Names, Types, and Assignments](#)

A variable name may be any combination of the characters A through Z, 0 through 9, and the underscore. The first character of the name should be a letter. Variable names are preceded by one or two ampersands, depending on the type.

There are two types of variables that can be assigned values in a Process Flow.

**LOCAL**

Assigns a value for a variable that is used throughout the current flow. It can be used in the current process flow in conditions, email objects, and as parameters to data flows and process flows.
GLOBAL

GLOBAL assigns a value for a variable that is used throughout the session. In addition to the places where a local variable can be used, global variables can also be used in the current process flow by any procedures it invokes, as well as any data or process flow that it invokes, provided that Run as procedure is selected in its properties.

There are two ways of assigning a value for a variable within a process flow:

DEFAULT

Supplies an initial (default) value for a variable that had no value before the command was processed. It ensures that values are passed to the variable, whether or not they are provided elsewhere. If the variable already has a value, the command is ignored.

SET

Assigns a value for a variable, which overrides any previously set values.

Procedure: How to Set a Global Variable

To set a global variable and its value:

1. Open a data flow.
2. Click the Process Flow tab.
   The Process Flow workspace opens.
3. Right-click the line joining Start and Data flow and click Delete.
4. On the Flow tab, in the Insert group, drag the Set Variables button into the workspace between Start and Data Flow.
5. Right-click Start, drag a joining line to Set Variables, and release. Do the same between Set Variables and Data Flow. Your workflow should look like the following image.

6. Double-click the Set Variables object. The Set Variables Editor opens.
7. Click the Add SET Parameter button to open the User Variables calculator. In the Expression box, enter a name for the variable and a value. Click OK.
   The Set Variables Editor now shows the name and value entered.
8. Click OK. The name and value appear on the Properties pane, as shown in the following image.

![Properties pane image]

**Procedure: How to Pass Parameters in a Process Flow**

When you are using a procedure in a process flow, you can change the value of its parameters when the flow runs.

From the Process Flow tab:

1. Double-click the procedure.
   
   or
   
   Right-click the procedure and click Properties.

   The Properties pane opens.

2. Select the Parameters ellipsis button to open the Parameter Editor. Then, select the Add Parameter button to open the Drill Down Parameter Editor.
Enter `PCAT` in the Parameter Name field. Leave Constant Value selected and enter `Cameras` as a parameter value, as shown in the following image.

![Drill Down Parameter Editor](image)

3. Click OK.

The value appears in the Parameter Editor. Repeat the process for the variable `PTYPE`, using a parameter value of `Digital`.

The parameters can also be entered directly in the Properties pane by following the guidelines below:

- Do not include the ampersand in the variable name.
- Use single quotes around character and date valued variables.
- Separate multiple variables with commas.

4. When done, click OK to close the Parameter Editor.

The parameters you entered appear in the Properties page.

5. Save and submit the flow.
**Reference:** DataMigrator Variables

Most DataMigrator variables start with `&&CM__` (double underscores). They set the following variables in each flow:

`&&CM__AUTHOR`

Is the user ID that originally saved the flow.

**Note:** `&&CM__AUTHOR` is the user ID that originally saved the flow, which is not always the user ID that ran the flow. To find the user ID that ran the flow, assign a value of `GETUSER('A32')` to a variable in the set variables object on the process flow tab.

`&&CM__REQUEST`

Is the name of the flow.

`&&CM__TARGET`

Is the name of the target table for the data flow. Note that for multi-table requests, the variable `CM__TARGET` has a number appended to the name.

When you run a flow, these variables are automatically set and can be used in a procedure that follows. For example, you can display them in logs, embed them in reports, and use them in email messages. You can reference these variables in procedures that run before or after data flow execution.

**Reference:** Control Variables

Control variables are typically used in process flows.

`&&KILL_RPC`

Shuts down all flow processing if the value is set to Y. The flow immediately branches to the end of its logic, skipping all code that follows the procedure where this setting is used. Any subsequent data flows or procedures are not executed. This may be useful in cases where a site-dependent variable (such as a program result or special file allocation) fails, and the entire extract is to be stopped. The default is N.

To reset, type:

```
-SET &&KILL_RPC = 'Y';
```
Reference: Statistical Variables

DataMigrator supports the following statistical variables that can be referenced in procedures. Statistical variables are primarily used in process flows.

&ACCEPTS
Is the number of transactions accepted.

&CHNGD
Is the number of rows updated.

&DBMSERR
Is the number of rows rejected due to database constraint violations, or for flat file targets, format conversion errors.

&DELTED
Is the number of rows deleted.

&DUPLS
Is the number of rows rejected as a result of duplicate values in the data source.

&FOCERRNUM
Is the last error number displayed. If the last command did not generate an error, the value of &FOCERRNUM is 0.

&FORMAT
Is the number of rows rejected as a result of a format error.

&INPUT
Is the number of rows added to the data source.

&INVALID
Is the number of rows rejected as a result of a VALIDATE condition.

&NOMATCH
Is the number of rows rejected as a result of not matching a value in the data source.

&REJECTS
Is the number of rows rejected for reasons other than the ones specifically tracked by other statistical variables.
&TRANS
Is the number of rows processed.

&LINES
Is the number of lines returned in the last answer set.

&READS
Is the physical number of reads from an external file.

&RECORDS
Is the number of records retrieved in the last answer set.

Reference: Slowly Changing Dimension Variables

Slowly Changing Dimension variables are used to control flows that have their target properties Load Type set to Slowly Changing Dimensions. They allow overriding of the default begin and end dates when rows with Type II changes are updated and inserted.

&&CM__SCDATE
Overrides the current date as the begin date and end date for Type II changes. This setting is typically used when a run was missed, and a rerun should be run as though it were run on the earlier date. For example, the following could be used so that an effective date, June 1, 2008, is used by including it in a procedure that is run before the data flow.

-SET &CM__SCDATE = 20080601;

&&CM__SCDEIDT
Overrides the current date as the end date for Type II changes for inactive rows. This setting is typically used so that the day prior to the run date becomes the end date. For example, if today is June 1, 2008, then either of the following could be included in a procedure run before the data flow, so that an end date of May 31, 2008 is used.

-SET &CM__SCDEIDT = 20080531;
-SET &CM__SCDEIDT = AYMD(&YYMD, -1, 'I8YYMD');

&&CM__SCDEADT
Overrides a NULL as the end date to use for Type II changes for active rows. This setting is typically used with a date far in the future. For example, the following could be used in a procedure before the data flow, so that January 1, 3000 is used as the end date.

-SET &CM__SCDEADT = 30000101;
Maintaining Continuous Processing in a Single Agent

By default, each time you run a flow, either standalone or as a component in a process flow, a new agent is started. However, you may want to run a flow in the agent that is currently running, rather than start a new agent. Suppose, for example, that you run a procedure that issues a FILEDEF, a SET command, or that creates a global (&&) variable, and you want to reference the setting or variable in the flow that follows the procedure. Since a flow that runs in a new agent will not recognize or be able to use these settings or variables, you need to maintain continuous processing in the same agent.

DataMigrator supports the runtime requirement. To ensure that the continuous connection is maintained at runtime, you can request that a flow that creates input required by another flow be marked to Run as Stored Procedure. See How to Maintain Continuous Processing in a Single Agent on page 407.

Procedure: How to Maintain Continuous Processing in a Single Agent

1. From the navigation pane, select a flow, and drag it into the process flow workspace. The flow may be a data flow or another process flow.

2. Right-click the object to the left of the flow and, holding the mouse button down, drag from the left object to the right object, creating a connecting arrow. Release the mouse button.

   The procedure will be executed within the process flow at the point where it is encountered.

3. Double-click the flow object to open the Properties pane.

4. Select the Run as Stored Procedure check box to execute the flow in the running agent. For details, see Properties Pane for Flows and Procedures on page 407.

5. Close the window.

6. Click the Save button to update the process flow.

   Tip: If you switch to Text View and examine the underlying execution code for a flow, you will see EX CMASAP as the flow name. CMASAP triggers a connection to a new agent. If you select Run as Stored Procedure in the Properties pane, the code shows EX flow name and keeps the same agent open for processing.

Reference: Properties Pane for Flows and Procedures

To access the Properties pane, double-click a flow object or procedure.

The Properties pane has the following fields and options:

Flow or Procedure Name

   Displays the name of both the object and the directory in which the flow or procedure is stored.
If you drag a Flow icon from the toolbar, you can enter:

- The name of an existing procedure and its application directory.
- The name of a new procedure that you must build and save.

**Note:** If you have already created and saved the procedure, the application directory and procedure name appear when you open the window.

**Get Parameters using Synonym**

Select this check box to obtain parameters using a synonym, instead of entering them in the Properties pane.

When this option is used, a variable is automatically generated with the name of each field in the selected synonym. If the associated data source of the synonym has more than one row, the flow is submitted once for each row, with values for the variables obtained from each row.

For an example of using this option, see *Iterating Through Multiple Tables* on page 631. If you select this option, the following fields are available:

- **File Name.** Enter or select the application directory and name of a synonym for DataMigrator to use for obtaining parameters. This option only appears when the Get Parameters using Synonym option is selected.

- **Parallel Run.** When this option is selected and the associated data source of the synonym has more than one row, flows are submitted asynchronously. This option only appears when the Get Parameters using Synonym option is selected and is only available for flows.

**Parameters**

Indicates any parameters that should be used when the flow or procedure is run.

Clicking the ellipsis opens the Parameter Editor for setting the parameter values. For more information, see *How to Pass Parameters in a Process Flow* on page 402.

**Run as Stored Procedure**

Select this check box if you wish to run the flow in the agent that is currently running, rather than by connecting to a new agent. This option is only available for flows. For example, you would use this method to maintain continuous processing:

- If one data flow creates a temporary flat file that is used as an input to the next flow.
- If a procedure creates a global variable, or issues a FILEDEF or SET command that you need to reference in the new flows.
**Do not allow multiple instances**
Select this option to prevent this flow from running if another instance is already running. This option is only available for flows.

**Display Name**
Displays the name of the flow or procedure.

**Notes**
Add notes describing the purpose of the flow or procedure.

**Executing a Process Flow That Contains Only Procedures**
You can create a process flow that contains one or more procedures and other objects, but no data flows. This may be useful for testing purposes, or for a particular warehouse scenario where the DataMigrator Scheduler is being used, but the only production code is the logic within the procedure.

For example, you can create and run a procedure that connects you to a DBMS, or that produces a report from a data file or a DataMigrator table, such as the DataMigrator log.

**Procedure: How to Execute a Process Flow That Contains Only Procedures**
1. In the DMC, right-click an application directory in the navigation pane, point to New, and then click Flow.
2. Click the Process Flow tab.
3. Drag a procedure from the navigation pane into the process flow workspace, following the Start object.
4. Connect the objects.
5. Optionally, add other procedures to the process flow.
6. Save the procedure.
7. On the Flow tab, in the Run group, click Run, and then click Run from the shortcut menu.
   or
   Right-click the saved process flow in the navigation pane and choose Run from the menu.
Output is displayed in the Console Log below the workspace.
Adding Email Messages to a Process Flow

Before you create email messages for a process flow, you must complete or verify your email notification setup in the Web Console. For details about accessing the E-Mail Notification tab and configuring SMTP services, see Configuring DataMigrator Server Operations on page 84.

Once SMTP services are properly configured, you can construct your messages within the process flow. You can include email messages at any point in the flow.

Procedure: How to Add an Email Object to the Process Flow

1. On the Flow tab, in the Insert group, drag the E-Mail icon from the ribbon into the process flow workspace.

2. Double-click the E-mail object to open the Properties pane, as shown in the following image.

3. In the Email To box, type the email addresses of one or more recipients. Separate multiple addresses with semi-colons (;).

4. In the Subject box, type a subject.

   Note: Optionally, select a file to attach. Files of any type can be attached to the email object.
5. Optionally, select the Attach File checkbox to add an attachment. Once selected, a File Name box becomes available. Here, you can either type in the file name, or use the ellipses button to browse for the file. Any file type or extension is supported.

6. In the Message Text box, type your message.

7. Close the pane.

8. Click the Save button to update the process flow.

**Note:** You can also add email properties from the Web Console.

### Setting Connection Logic

When you connect one object to another, DataMigrator draws an arrow between them. These arrows, or *connector objects*, specify the logic that governs what is executed next in a process flow. The connection logic may be unconditional or conditional, represented graphically by colored arrows. Conditional logic is set by default. However, you can edit the connection to change the nature of the condition or to make it unconditional.

- **Black arrows** indicate unconditional connection logic. This means that the next object will be processed immediately after its predecessor. For some pairs of objects, unconditional logic is required. The connections between the following types of objects and their successors are unconditional by default, and cannot be changed:
  - Start connects to any other object type
  - Parallel Group connects to Wait

- **Green arrows** indicate that the next object will be processed if the previous one runs successfully. This is the default when you draw an arrow from one object to the next in a single path, as shown in the following image.
Red arrows indicate that the next object will be processed if the previous one fails to run properly. This is the default for the second arrow when you draw two arrows from the same object to other objects, forming two branches in the flow. Since the first arrow (the green one) means run when the flow is successful, it follows that the second arrow (the red one) means run when the flow fails, as shown in the following image.

Purple arrows indicate that the next object will be run based on a custom condition you define. You can use statistical or global variables (**variables**) in custom conditions to test specific conditions, return codes, or values based upon which branching will occur. For example, you might want the flow to branch differently based on a particular error message number. For related information, see *Branching in a Process Flow* on page 418.

**Note:** When multiple conditional arrows are used, the first object where the condition is met is run.

The connections between the following objects are conditional by default. You can change the conditions or make them unconditional:

- Email connects to any other object type.
- Flow connects to Flow, Parallel Group, Procedure, Email. (The conditions you can apply are: OK, FAILURE, and a Custom condition.)
- Procedure connects to Flow, Parallel Group, Procedure, Email. (The conditions you can apply are: OK, FAILURE, and a Custom condition.)
- Wait connects to Flow, Parallel Group, Procedure, Email. (The conditions you can apply are: Wait satisfied and Timeout.)
Procedure: How to Edit a Connection Condition

Follow these steps to edit a connection condition between a flow or a procedure and one of the following object types: a flow, a Parallel Group, a Procedure, or an Email.

1. Drag the required objects into the process flow workspace.
2. Connect the objects.
3. To view or edit the connection condition, double-click the arrow connecting the two objects. The Condition dialog box opens.
   For a single connection, OK is the default. For a second connection, Failed is the default.
4. Click a radio button to change the condition.
5. If you wish to specify a Custom condition, click Custom. The Edit Custom Focus Condition box becomes active.
   Type the code for the condition you want to apply. Typically, you use a global (&&) variable, DataMigrator variable, or statistical variable in your custom conditions to test specific conditions, return codes, or values based upon which branching will occur.
6. Click OK to close the condition box.
7. Click the Save button to update the process flow.

Example: Creating a Custom Connection Condition

You can set a custom condition for a connection between objects in a process flow. In this case, a connection will become active if duplicate rows are rejected while a data flow is run. The number of duplicates will be included in an email that is sent if the condition is met.

Note: Your DataMigrator Server must be configured for email notification to complete this example. For details, see Configuring DataMigrator Server Operations on page 84.

1. In the DMC, right-click an application directory in the navigation pane, point to New, and then click Flow.
2. Click the Process Flow tab.
3. Drag the DMDUPLG data flow object from the ibisamp application directory in the navigation pane into the workspace, to the right of the Start object.
4. Create a connection by right-clicking the Start object and dragging to the DMDUPLG data flow object.
5. On the Flow tab, in the Insert group, drag an email object into the workspace, to the right of the DMDUPLG data flow object.
6. Create a connection by right-clicking the DMDUPLG data flow object and dragging to the email object.
The workspace should appear, as shown in the following image:

7. Double-click the arrow between the DMDUPLG data flow object and the email object.
   or
   Right-click the connection arrow and click Edit Condition.
   The Condition dialog box opens.
8. Click the Custom radio button.
   The Edit Custom Focus Condition field becomes active.
9. Type the following in the Edit Custom Focus Condition box and click OK, as shown in the following image.
Note: Ensure the Data Flow that was added to the Process Flow has the Run as Stored Procedure setting enabled in the Properties pane. This will maintain the continuous connection, and will enable the reference to the variables at run time.

10. Right-click the email object and click Properties.

   The Properties pane opens.

11. In the Email To box, type an email address. In the Subject box, type Duplicate Rows Rejected.

12. Type the following in the Message Text box, as shown in the following image.
&DUPLS rows rejected

This connection will become active if duplicate rows are rejected while a data flow is run, and an email notification will be sent with the number of duplicate rows.
**Reference: Condition Dialog Box**

To access the Condition dialog box shown in the following image, from the process flow workspace, double-click a connection arrow between objects, or right-click a connection arrow and click Edit Condition.

The Condition dialog box has the following fields and options:

**OK Default Condition (Return Code is 0)**

Executes the connection if the prior object executed successfully.

**Failed Default (Return Code is 0)**

Executes the connection if the prior object failed to execute.

**Unconditional**

Executes the connection regardless of the success or failure of the prior object. This option is only available when there is only one connection.

**Custom**

Allows you to set a custom condition for the connector.

**Edit Custom FOCUS Condition**

The code for a custom condition. Enter a condition using FOCUS Dialogue Manager syntax that evaluates to true or false.
Branching in a Process Flow

The ability to set connection logic enables you to create branches within a process flow. You simply connect multiple objects to the object from which you want the branching to begin.

The branching process flow shown in the following image is built in Exercise 2: Creating the Process Flow on page 460. It is strongly suggested that you complete these exercises in order to create the branches shown here. As you do so, you will become familiar with many aspects of data and process flow design in the DMC.

Procedure: How to Create Branches in a Process Flow

1. Drag an object into the workspace following the object from which you want the branching to begin.

2. To build the first branch, connect the object at the starting point for the branches to the object you just dragged in. A green arrow is created by default.

3. To build the second branch, drag another object into the workspace and connect it to the object at the start of the branch. A red arrow is created by default.

4. If you wish to change the connection conditions for any branch, double-click the arrow and edit the condition in the Conditions dialog box. Then click OK.

5. Click the Save button to update the process flow.

Note: You can also create branches using custom conditions. Note, however, that while you can use multiple custom conditions to specify different branches, you should not use custom conditions with OK and FAILED conditions.

When multiple conditional arrows are used, the first object where the condition is met is run.
Grouping and Controlling Simultaneous Execution of Flows

Often you need to define a process flow in which multiple steps occur simultaneously and where other steps cannot begin until all of the prior steps have been completed.

For example, suppose that you are creating an employee information data mart as a data target. However, the data mart requires data from two other independent data flows, one that loads data from a personnel database and one that loads data from an accounting database. You need to make sure that the data flow for the data mart does not run until the two feeder data flows have completed their runs.

By grouping the feeder data flows, you can enforce their simultaneous execution. Then, by adding a Wait object, you can ensure that the third data flow, the employee data mart, does not run until the other two have loaded the required data into their respective targets. This technique eliminates the need to estimate what time the first two flows will end in order to schedule the third flow to run at that time.

You can also pass parameters to the procedures in a parallel group using the Properties pane for the flow by entering a value into the Parameters box.

**Note:** The Scheduler must be running to execute process flows that require simultaneous execution or wait conditions. If the Scheduler is not running, the flow will not run, and a message will appear in the log.

**Procedure:** How to Execute a Group of Procedures and Wait

1. Right-click an application directory in the navigation pane, point to *New*, and then click *Flow*.
2. Click the *Process Flow* tab. A new workspace containing a Start object opens in the right pane.
3. On the Flow tab, in the Insert group, click *Parallel Group* and drag it into the workspace, where it appears as an empty box with a border of dashes.
4. Right-click the previous object in the workspace and draw an arrow connecting it to the Parallel group object.
5. Drag a flow (it can be a data flow or another process flow) from the navigation pane into the Parallel Group object.
6. Drag another flow from the navigation pane into the Parallel Group object, placing it below the previous object.

   **Note:** You can group as many flows or procedures as you require in a Parallel Group object.

7. On the Flow tab, in the Insert group, click *Wait*, and drag it into the workspace to the right of the Parallel Group object.
8. Right-click the Parallel Group object, which now contains the multiple flows, and draw an arrow connecting it to the Wait object. Note that this is an unconditional connection.

9. By default, the Wait object simply waits until the flows in the Parallel Group object complete their runs. If you want to limit the wait time, double-click the Wait object to open the Wait attribute pane, as shown in the following image.

![Wait attribute pane image]

In the Interval text box, specify the maximum number of seconds to wait before executing the next component in the flow.

10. From the navigation pane or the Insert group, drag the next object into the workspace, to the right of the Wait object.

11. Connect the Wait object to the object that follows it in the flow. By default, the processing will continue to the next object when the default condition is satisfied.

12. If you wish to change the default condition, double-click the connecting arrow to open the Wait Condition Type window. See *Wait Condition Type Dialog Box* on page 422.

   Change the condition to *Timeout or Failure* and click *OK*.

13. Click the *Save* button to update the process flow.

**Procedure:**  How to Edit a Wait Connection Condition

1. Drag a Flow, Procedure, or E-mail object into the process flow workspace, following a Wait object.

2. Connect the objects.

3. To view or edit the connection condition, double-click the arrow connecting the objects. The Wait Condition Type window opens.

   For a single connection, the default condition is Wait satisfied, represented by a green arrow.
For a second connection, the default condition is Timeout or Failure, which shows that either the Wait is not satisfied OR one of the procedures in the group failed. This is represented as a red arrow.

4. Click an option button to change the condition, and then click OK to close the window.
5. Click the Save button to update the process flow.

**Reference:** Wait Attribute Pane

To access the Wait attribute pane shown in the following image, from the process flow workspace, double-click a Wait object.

![Properties pane for a Wait object](image.png)

The Properties pane for a Wait object has the following fields and options:

**Interval**

- Number of seconds to wait for a parallel group to finish executing. The default value is 0.
- An interval of 100 causes the process flow to wait up to 100 seconds or until all flows in the group have been completed.

**Display name**

- The alternative display name of the object.

**Notes**

- An annotation describing the object.
Reference: Wait Condition Type Dialog Box

To access the Wait Condition Type dialog box shown in the following image, from the process flow workspace, double-click the arrow coming from a Wait object, or right-click it and select Edit Condition.

The Wait Condition Type dialog box has the following fields and options:

**OK Default Condition (Wait satisfied)**

If the grouped objects complete their runs successfully within the set time limit, the wait condition is satisfied and the next upon-successful-completion step in the flow runs.

**Timeout or Failure**

If the grouped objects do not complete their runs successfully within the set time limit, or one of the grouped objects fails, the wait condition is not satisfied and the next Timeout or Failure step in the flow is run.

**Unconditional**

The next steps run regardless of any conditions.

Running a Process Flow

There are two ways to run a process flow from the DMC: Run and Submit.

Run will run the process flow immediately. When you run a flow, processing information is displayed in the Console Log.

Submit will submit the process flow to the scheduler to run. When you submit a flow, the processing information is recorded in the DataMigrator Log and Statistics tables. Email messages are also sent if the flow properties are set to send email messages at the start or end of a job.
**Note:** You can also schedule the flow based on a variety of conditions. For information about scheduling options, see *Scheduling a Process Flow* on page 423.

**Procedure:**  **How to Run a Process Flow**

There are two ways in which you can run a process flow:

- On the Flow tab, in the Run group, click *Run*, and then click *Run*.
- Right-click a process flow in the navigation pane. Select one of the run options available from the menu.

For more information regarding the run options, see *Navigation Pane: Procedures* on page 195.

**Scheduling a Process Flow**

DataMigrator provides several ways to run data and process flows. In addition to executing flows immediately, you can use the DataMigrator internal scheduler to run requests off-hours, on a recurring basis, or on specified days of the week or month.

Alternatively, you can schedule or execute a flow from an application other than DataMigrator. For details, see *Event-Based Scheduling* on page 695.

You can also disable a scheduler without removing the Schedule object from the flow.

You can schedule a process flow to run in the following ways:

- **Run Once.** Runs your flow once.
- **Recurring.** Repeats your flow at an interval that you define.
- **Multi-Day.** Runs your flow on certain days of the month.
- **Run when server starts.** Runs your flow every time the DataMigrator Server starts.

When the scheduler checks, it executes all flows scheduled to run in the period of time since the scheduler last checked.

To access DataMigrator Scheduler parameters:

1. Expand the *Workspace* folder.
2. Expand the *Special Services and Listeners* folder, right-click *SCHEDULER*, and click *Properties*.

For scheduled flows to run, the scheduler must be running. On the Scheduler Configuration dialog box, the *sched_autostart* drop-down menu must be set to yes.

**Note:** You can also add or change schedule properties from the Web Console.
**Procedure:** How to Schedule a Process Flow to Run Once

2. Select Active from the Schedule Status drop-down menu.
3. Select Run Once from the Schedule Type drop-down menu.
4. Select a start date in the Start Date field by typing the date in the month/day/year areas, or by clicking the down-arrow to open a calendar in which you can choose a date. The default is the current day.
5. Select a start time in the Start Time field by typing values directly into the hour/minute/seconds/AM or PM areas, or by highlighting an area and using the Up and Down Arrow keys to select new values. The default is the current time.
6. Close the pane.
7. Click the Save button to update the process flow.

**Procedure:** How to Schedule a Process Flow to Run on a Recurring Interval

2. Select Active from the Schedule Status drop-down menu.
3. Select Recurring from the Schedule Type drop-down menu.
4. Select an interval type (such as daily, monthly) from the Interval Type drop-down menu.
5. Specify an interval number in the Interval Value field to indicate the frequency of execution. For example, if you select Weekly as your Interval Type, and 2 as your Interval Value, the flow will be executed every two weeks.
6. Select a start date in the Start Date field by typing the date in the month/day/year areas, or by clicking the down-arrow to open a calendar from which you can choose a date. The default is the current day.
7. Select a start time in the Start Time field by typing values directly into the hour/minute/seconds/AM or PM areas, or by highlighting an area and using the Up and Down Arrow keys to select new values. The default is the current time.
8. Optionally, select an end date and time by selecting the Specify End Date check box. Specify the end date by typing the date in the month/day/year areas, or by clicking the down-arrow to open a calendar from which you can choose a date. Specify the time by typing values directly into the hour/minute/seconds/AM or PM areas, or by highlighting an area and using the Up and Down Arrow keys to select new values.
9. Optionally, specify special date and time ranges. You can configure the scheduler to run on specific days of the week or month, on the last day of the month, or specify the beginning and ending times of the daily time span.

10. Close the pane.

11. Click the Save button to update the process flow.

**Tip:** If you wish to stop a query from executing regularly, you do not need to change the schedule. Instead, set Schedule Status to *Inactive* in the Schedule attribute pane.

**Procedure:** How to Schedule a Process Flow to Run Multi-Day Intervals

1. Right-click the Start object in the Process Flow tab and click Properties.
   
   The Properties pane opens.

2. Select *Active* from the Schedule Status drop-down menu.

3. Select *Multi-Day* from the Schedule Type drop-down menu.

4. Select a start date in the Start Date field by typing the date in the month/day/year areas, or by clicking the down-arrow to open a calendar from which you can choose a date. The default is the current day.

5. Select a start time in the Start Time field by typing values directly into the hour/minute/seconds/AM or PM areas, or by highlighting an area and using the up and down arrows to select new values. The default is the current time.

6. Optionally, select an end date and time by selecting the Specify End Date check box. Specify the end date by typing the date in the month/day/year areas, or by clicking the down-arrow to open a calendar from which you can choose a date. Specify the time by typing values directly into the hour/minute/seconds/AM or PM areas, or by highlighting an area and using the up and down arrows to select new values.

7. Configure the scheduler to run on specific days of the week or month, on the last day of the month, or specify the beginning and ending times of the daily time span.

8. Close the pane.

9. Click the Save button to update the process flow.

**Tip:** If you wish to stop a query from executing regularly, you do not need to change the schedule. Instead, set Schedule Status to *Inactive* in the Schedule attribute pane. For details about setting flow properties, see *Setting Flow Properties* on page 161.

**Procedure:** How to Schedule a Process Flow to Run Holidays

You can schedule a process flow to execute or not execute on holidays. The Holiday option is available when Recurring or Multi-Day are selected in the Schedule Type drop-down menu. For more information, see *Specifying Holiday File* on page 430.
**Note:** You can only schedule the flow to run on holidays after you create a holiday file.

1. Right-click the **Start** object in the Process Flow tab and click **Properties**.
   
   The Properties pane opens.

2. Select **Active** from the Schedule Status drop-down menu.

3. Select **Recurring** or **Multi-Day** from the Schedule Type drop-down menu.

4. Select an option from the Run on Holidays drop-down menu. The following options are available:

   **Skip**
   
   Excludes a run on desired holidays.

   **Only**
   
   Only run on desired holidays.

   **Run**
   
   Only available if the schedule type is set to **Multi-Day**. Runs on all holidays.

5. Click the **Save** button to update the process flow.

To verify that the holiday schedule is set, in the navigation pane, expand the **Workspace** folder. Then expand the **Special Services and Listeners** folder. Right-click **SCHEDULER** and click **Scheduled Events**. The Scheduler/Scheduler Events report should list scheduled flows by the holiday set date.

**Procedure:** **How to Schedule a Process Flow to Run When the Server Starts**

You can schedule a process flow to run whenever the DataMigrator server is started. This option is useful that need to run all of the time, for example, when using Change Data Capture or File Listener capabilities.

1. Right-click the **Start** object in the Process Flow tab and click **Properties**.
   
   The Properties pane opens.

2. Select **Active** from the Schedule Status drop-down menu.

3. Select **Run when server starts** from the Schedule Type drop-down menu.

4. Click the **Save** button to update the process flow.

**Procedure:** **How to Disable a Schedule Object**

You can disable a the schedule of a flow from the Properties pane for a schedule.

1. Right-click the **Start** object in the Process Flow tab and click **Properties**. 
The Properties pane opens.

2. Select *Inactive* from the Schedule Status drop-down menu.
3. Close the pane.

**Reference:** **Schedule Attribute Pane**

To access the Schedule attribute pane shown in the following image, right-click the *Start* object and click *Properties.*
The Schedule attribute pane contains the following fields and options:

**Schedule Status**

Determines if there is a schedule and if so, whether the schedule is active or inactive.

**Schedule Type**

Determines whether a process flow will execute once, on a recurring basis, on several specific days (Multi-Day), or when the server starts.

**Interval Type**

Specifies the interval in which a process flow will be executed. For example, if you select Month as your Interval Type, and select 2 as your Interval Value, the flow will execute every two months. This option is only available if the schedule type is set to Recurring.

**Interval Value**

Specifies the frequency at which a process flow will be executed. For example, if you select Week as your Interval Type, and select 2 as your Interval Value, the flow will execute every two weeks. This option is only available if the schedule type is set to Recurring.

**Start Date**

Specifies the start date for the execution of the process flow.

**Start Time**

Specifies the start time for the execution of the process flow.

**Specify Stop Time**

Allows you to specify an end time for the execution of the process flow. This option is only available if the schedule type is set to Run Once or Multi-Day.

**Specify End Date**

Allows you to specify an end date and time for the execution of the process flow. This option is only available if the schedule type is set to Recurring or Multi-Day.

**End Date**

Specifies the end date for the execution of the process flow. This option is available if Specify End Date is selected.

**End Time**

Specifies the end time for the execution of the process flow. This option is available if Specify End Date is selected.
**Days of the Week**

Specifies the day of the week on which the process flow will be executed. This option is only available if the schedule type is set to *Recurring* or *Multi-Day*.

**Days of the Month**

Specifies the day of the month on which the process flow will be executed. This option is only available if the schedule type is set to *Recurring* or *Multi-Day*.

**Intraday Start**

Specifies the beginning of the daily time span of the process flow. This option is only available if the schedule type is set to *Recurring*.

**Intraday End**

Specifies the end of the daily time span of the process flow. This option is only available if the schedule type is set to *Recurring*.

**Run on Holidays**

Specifies the option to execute the process flow on desired holidays.

The following options are available:

- **Skip.** Excludes run on desired holidays.
- **Only.** Only run on desired holidays.

This option is only available if the schedule type is set to *Recurring* or *Multi-Day*.

**Log and output Destinations**

This option controls where the log information from the Maintenance Job should be sent when the Job finishes.

The following options are available:

- **ETLLOG.** The log is sent to the DataMigrator log.
- **EDAPRINT.** The log is sent to the Server log.
- **E-Mail.** The log is sent as an attachment to an email message if specified in Job Properties.

Multiple options can be selected. The default values for this option are ETLLOG and E-Mail.

**Note:** If the job is run from the DMC or Web Console, information will go to the console log regardless of this setting.
Send HTML output as inline message. When selected, if any stored procedures called by the flow generate an output, the message is included as message text, instead of as an attachment.

Note: This option for a flow schedule works in conjunction with the Merge formatted output into log option that is available for a Scheduler Special Service. The following considerations should be taken into account:

- When the Send HTML output as inline message and Merge formatted output into log options are selected, the submitted flow emails the attached ETLLOG file and the HTML output is embedded into the body of the email.
- When the Send HTML output as inline message and Merge formatted output into log options are not selected, the submitted flow emails the attached ETLLOG file and the HTML output is embedded into the attached ETLLOG file.

Reference: Specifying Holiday File

You can specify a list of dates that are designated as holidays in your company. These dates can be scheduled to be excluded, only included, or included to run with regular days in the process flow. To define a list of holidays, you must create a holiday file using a standard text editor and select the holiday file by issuing the SET command with the HDAY parameter.

1. In a text editor, create a list of dates designated as holidays using the following rules for creating a holiday file:

- Dates must be in YYMD format.
- Dates must be in ascending order.
- Each date must be on its own line.
- You may include an optional description of the holiday, separated from the date by a space.

For example: 20160315 Ides of March

2. Save the file into the dm\etc directory:

- In Windows and UNIX the file must be HDAYxxxx.err.
- In OS/390, the file must be a member of ERRORS named HDAYxxxx, where xxxx is a string of text four characters long.

3. From the navigation pane, expand the Workspace folder. Expand the Configuration Files folder and right-click Server Profile - edasprof.prf. Click Edit and type:

```plaintext
SET HDAY = xxxx
```
where:

xxxx

Is the part of the name of the holiday file after HDAY. This string must be four characters long. For example, if your file name is hdaymyh1.err, then the setting will be SET HDAY = myh1.

**Note:** Alternatively, instead of the server profile, the setting can be saved in the user profile for the scheduler as set in sched_scan_id.

4. Save the new Server Profile.

5. Test that holiday is correctly set up by creating a sample hdayfile synonym.

FILENAME=HDAYFILE, SUFFIX=FIX, DATASET=C:\ibi\srv77\dm\etc\hdaymyh1.err, $ SEGMENT=HDAYFILE, SEGTYPE=S0, $ FIELDNAME=HDAY_DATE, ALIAS=hddate, USAGE=YYMD, $ FIELDNAME=HDAY_DESC, ALIAS=hddesc, USAGE=A66, ACTUAL=A68, $

**Note:** DATASET should be set to your appropriate directory path.

6. Run Sample Data for hdayfile synonym. An example of the sample data is shown in the following image.

<table>
<thead>
<tr>
<th>HDAY_DATE</th>
<th>HDAY_DESC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2010/01/01</td>
<td>New Year's Day</td>
</tr>
<tr>
<td>2 2010/01/18</td>
<td>Martin Luther King, Jr. Day</td>
</tr>
<tr>
<td>3 2010/02/15</td>
<td>Presidents' Day</td>
</tr>
<tr>
<td>4 2010/05/31</td>
<td>Memorial Day</td>
</tr>
<tr>
<td>5 2010/07/05</td>
<td>Independence Day</td>
</tr>
<tr>
<td>6 2010/09/06</td>
<td>Labor Day</td>
</tr>
<tr>
<td>7 2010/10/11</td>
<td>Columbus Day</td>
</tr>
<tr>
<td>8 2010/11/25</td>
<td>Thanksgiving Day</td>
</tr>
<tr>
<td>9 2010/11/26</td>
<td>Day after Thanksgiving</td>
</tr>
<tr>
<td>10 2010/12/24</td>
<td>Christmas Eve</td>
</tr>
</tbody>
</table>
Your company is trying to empower both its sales force and its line managers with timely and focused information to support needed decisions:

- Sales agents need information that will enable them to report quickly on the most and least popular products, and to track regional trends.

- Line managers need to assess strengths and weaknesses in their sales approaches using comparative analysis of competing stores to determine, for example, if the other stores are outselling them on one type of product. If so, they may need to initiate a sales promotion or provide more sales staff training.

Both groups need current and customized sales data to create self-service reports and perform analyses. It is your job to provide that data.

To accomplish these goals, you will create a flow that moves data out of an operational data store and that builds a denormalized table that can be used to run sales and analytic reports.

**Note:** Refer to the dmdemo sample data flow in ibisamp for the complete example. However, note that dmdemo has only one data target, while the flow you build will contain two.

**In this chapter:**

- What You Are Going To Do
- Startup and Setup Tasks
- Exercise 1: Creating the Data Flow
- Selecting and Joining Data Sources
- Selecting and Managing Columns
- Adding Data Targets
- Submitting the Data Flow
- Data Flow Review
- Exercise 2: Creating the Process Flow
- Starting the Process Flow
- Adding a Data Flow Object to the Process Flow
- Branching in the Process Flow
- Submitting the Process Flow
What You Are Going To Do

The DataMigrator tutorial guides you through the creation of two flows:

- **In Exercise 1, you create and save a Data Flow** that extracts data from two data sources, and then transforms and loads the extracted data into two data targets. You will save the data flow and use it again in Exercise 2.

- **In Exercise 2, you create and save a Process Flow** that uses the data flow from exercise 1 as a component, along with two stored procedures that log the success or failure of this processing. The flow manages this sequence of events.

Startup and Setup Tasks

This tutorial assumes that your DataMigrator Server and DataMigrator Client software have been installed and configured.

You will need to complete the following startup and setup tasks before you begin to design your data flow:

- **Start your DataMigrator Server.**
- **Start the Data Management Console.**
- **Add a server node** that points to the DataMigrator Server (optional). You can use this node instead of the one that was configured during the installation process.
- **Connect to the server.**
- **Create the sample data and metadata** you will be using as data sources and data targets in the tutorial.
- **Set up the sample application directory** you will use in the tutorial. You create the new application and configure its path from the Web Console. You will be storing the flows and associated files you create in this application.

Start Your DataMigrator Server and Data Management Console

Using whatever method is appropriate for your operating system:

1. Start your DataMigrator Server.
2. Start the DMC.
3. If you are not going to configure an additional server node, right-click the default LOOPBACK node in the navigation pane and click Connect.
Configure a Server Node From the Data Management Console (Optional)

For the tutorial, we will create an additional node and name it DEMO. We will store and work with flows and related files in this location.

**Note:** Creating an additional node is only required if the server is remote.

Since many users prefer, or need, to establish additional server nodes that point to the DataMigrator Server they will be using in the course of their work, we have added this setup task.

However, you may wish to use the DataMigrator Server node that you configured during the DataMigrator Client installation. The node is simply a name that, by way of its properties, points to the DataMigrator Server.

If you wish to use the node name you set up during the DataMigrator Client installation skip to *Create Sample Data and Metadata and Create an Application Directory* on page 437. Keep in mind that when an instruction in the tutorial refers to DEMO, you must apply the instruction to the name of the server node you are using.

1. If the DMC is not already open, select *iWay Software\iWay 77 Client or Server*, and then click *Data Management Console* from the Windows Programs menu. The DMC opens, with the server node you configured during installation listed under Servers in the navigation pane at the left.

   The black circle that appears on the server icon indicates that the DMC is not currently connected to the DataMigrator Server. You will connect shortly.

   Notice that the Console Log at the bottom of the screen has tracked your startup steps. It will continue to track each step in your work process.

2. For the tutorial, we will name the new node DEMO. To make the DEMO server available, you will define the server node.
Right-click Servers and click Add Server Node. The Server Node Configuration dialog box opens, as shown in the following image.

3. Enter the following settings:

- In the Host field, enter the machine name or IP address where the DataMigrator Server is running.
- In the HTTP Listener Port field, enter the Port number of the DataMigrator Server. If you used the defaults when you installed your DataMigrator Server, the port number is 8117.
- In the Node Name field, enter DEMO as the name for the Server node.
- Optionally, enter a description of the node in the Node Description field.
- Leave the Security field set to Explicit.
Enter a valid user ID and password, and if required, a domain or provider name.

4. Click Save.

The DEMO node is now added to the Servers list.

5. Right-click the DEMO server node in the navigation pane, and click Connect.

Create Sample Data and Metadata and Create an Application Directory

Your first task is to create the sample data and metadata you will be using in the tutorial.

To create the sample data and metadata from the DMC, see How to Create Sample Procedures and Data on page 78.

After you have created your sample data, set up a separate application directory called salesdemo in the Application Directory tree. This application directory will contain the flows and other files you will be creating.

1. Connect to the DEMO server using a user ID with server administrator privileges.

2. Under the DEMO server, right-click Application Directories, point to New, and then click Application Directory.

   The Create New Application dialog box opens.

3. In the Application Name box, enter salesdemo as the new application directory name and click OK.

   The salesdemo application directory is now listed in the approot tree in the navigation pane. Notice that ibisamp, the directory in which sample data is installed by default, is already in the Application Path box, which means it is in the application path of the server.

   If the directory is not listed, right-click Application Directories and click Application Path. Check that the check box in the In Path column is selected for salesdemo. Click Save to save your changes.

Tip: In addition to salesdemo and ibisamp, the Application Directory tree contains a directory called baseapp. This is a default directory that is added automatically to your Application Path during installation. There is no need to add baseapp explicitly to your Application Path. The tree also contains a folder called foccache (Temporary) which contains temporary files for the duration of your session.

The salesdemo folder you created in the Web Console now appears in the Application Directories list.

Next step: Your setup is complete. The next step is to create a data flow.
Exercise 1: Creating the Data Flow

Your setup is done, and you are ready to start building the Data Flow.

Right-click the salesdemo application directory, point to New, and then click Flow.

A workspace opens in the right pane. The Data Flow tab is active and an SQL Select Columns object is displayed in the center of the pane.

Next step: You will start by selecting the data sources you will be using in this data flow.

Although you will be creating your flows in the salesdemo application directory, you will be using sample files from the ibisamp application directory (since this is where they are stored when you create the sample procedures).

Selecting and Joining Data Sources

Your first task is to specify the operational data sources from which you will be extracting and moving data. You will be adding two data sources to the data flow and joining them. In one of the data sources, you will create a virtual column, called a source transformation in DataMigrator terminology. This transformation will exist only in the data flow object where you build it.

1. In the DMC navigation pane, expand the ibisamp application directory.
2. On the Home tab, in the Filter group, click Synonyms.
3. Scroll down to the files that begin with the letters dm. All of these files were created in the data adapter of your choice when you chose Create Sample Procedures and Data (see Create Sample Data and Metadata and Create an Application Directory on page 437).

You will be using three synonyms from this group. Synonyms are used to describe both data sources and data targets. The first two are used for data sources and the last one for a data target:

- dminv
- dmord
- dmrpts

Tip: At many sites, data sources and data targets are on different servers and platforms. For this tutorial, we have chosen the simplest, most universal configuration. For information about different configuration options, see Remote Server Setup At a Glance on page 44 and Planning a DataMigrator Application on page 107.
Add Your First Data Source

You will start by selecting the synonym for the inventory table called `dminv` from the expanded ibisamp directory.

1. Drag `dminv` from the ibisamp application directory into the Data Flow workspace, to the left of the SQL object, as shown in the following image.

   ![Image of Synonym Selection](image1.png)

   A connecting arrow is automatically added between these objects.

   **Tip:** If you drag a synonym object to the left of the SQL object, it is understood to be a data source. If you drag it to the right of the SQL object, it is understood to be a data target.

2. Let us take a quick look at the data in the table you selected. By default, you can retrieve 50 rows from the server to verify that this is the correct table and that the data is available.

   Right-click the `dminv` object in the workspace, point to Operations, and then click Sample Data. The sample data for the table opens, as shown in the following image.

   ![Sample Data](image2.png)
The information has been retrieved, so you know the data source is available. Close the sample window by clicking the X in the tab above it.

**Add and Join a Second Data Source to the First**

To add another data source to the flow:

1. Right-click in the workspace to the left of the SQL object and click **Add Source**. The Select Synonym dialog box opens.
   
   If the ibisamp directory is not selected in the Look in: section, click it in the tree.

2. Scroll down and select the order table called **dmord**. Then click **Select**. (You can also drag dmord from the navigation pane, as you did earlier with dminv.)
   
   The dmord source object is now connected to the dminv source object by way of a join object, which is linked to the SQL object, as shown in the following image.

   **Tip:** The connecting arrows between these objects are added because the option *Automatically add join conditions* is selected by default. If these connections are not appearing on your screen, go to the Home tab and, in the Tools group, choose **Options**. Click **Data Flow Designer** under Data Flow in the navigation tree, and then click *Automatically add join conditions*.

3. We will verify the data in the orders table using a different method.

   Right-click the **dmord** source object and click **Expand**. The Columns tab shows the columns in the data source.
Click the Sample Data tab to verify that data is being retrieved from the orders table, as shown in the following image.

The first 50 rows are retrieved.

Tip: You can change the number of rows to retrieve from the Tools group. On the Home tab, in the Tools group, click Options. From the Options dialog box, click the Run Options link, and then change the number in Maximum number of rows for test reports.

4. Double-click the title bar or click the X in the upper-right corner to close the window and return to the object view.

Note: You can also double-click a source object and click the Sample Data button in the Source Transformations window. You can place sample data in the workspace by right-clicking a source object, pointing to Operations, and clicking Sample Data.

Create A Virtual Column in the Second Data Source

Shortly, we will take a closer look at the join you created. Before that, you will create a virtual column in the dmord source object (this is also called a source transformation). Data transformations in a source object are performed when the records are read, before any filtering or aggregation occurs.
Planning ahead a bit, you know that the time frame for the data target you will be creating is the year and month. To facilitate the mapping of source to target data, you decide to create a virtual column called YEARMONTH in the dmord data source, to be derived using the expression YEAR x 100 + MONTH. This column will represent both the year and the month.

1. Double-click the dmord source object. (You could also right-click the dmord source object and click Source Transformations.) The Source Transformations dialog box displays all of the columns in the data source.

2. To define the virtual column, click the Insert transforms button.

   The Transformation Calculator opens.

   a. In the Name box, type YEARMONTH.
   b. In the Format box, type I6 (I as in integer).
   c. In the Columns/Variables tab, double-click ORDYEAR.
   d. On the calculator keypad, click the multiplication sign (*).
   e. In the Expression box, type 100 +.
   f. Double-click ORDMONTH.

   The calculator should look like the image below:

   ![Source Transforms Calculator](image)

   **Tip:** Before you close the calculator, notice that the two columns you used in the expression are themselves virtual fields (or transformations) that were defined directly in the synonym.
Virtual columns are identified by the symbol \textcircled{V}. Adding a virtual column in a synonym is a useful strategy when you expect to use the same synonym with more than one flow. For details about adding virtual columns in a synonym, see \textit{Adding Virtual Columns (DEFINE) in a Synonym} on page 540.

3. Click OK to close the calculator and return to the Source Transformations dialog box.

Notice that YEARMONTH now appears in the Source Columns list, as shown in the following image.

4. To verify that the transformation is working properly, click the \textit{Test Transforms} button to view data.

5. You will now be able to use YEARMONTH as source data.

Close the Test Transformation dialog box, and then click OK to close the Source Transformations dialog box.

\textbf{Assign Join Properties}

Now let us consider the join object, which was added automatically when you selected the second data source. You will need to specify properties for the join.

By default, an inner join is created. An inner join extracts those rows that appear in both tables. You will base the join on an equality condition between two fields, one in each data source. The use of an equality condition is also called an equi-join.
**Tip:** Although not illustrated in this tutorial, DataMigrator supports multiple joins, joins based on conditions other than equalities, and joins that are modified by calculations, such as substrings or concatenations. A Join Calculator is available to assist you.

1. Right-click the join object and click **Join Editor**. The Join Editor dialog box opens.

2. The join must be based on columns in each of the joined data sources. Notice that `PROD_NUM` is in both Left and Right Source Columns lists. The join of `T1.PROD_NUM` and `T2.PROD_NUM` appears in the Expression field of the Join Conditions list. For the tutorial, the default join on `PROD_NUM` is sufficient.

**Tip:** Once again, the automatic join condition is in effect because the option *Automatically add Join conditions* is selected by default. If this is not the case, on the Home tab, in the Tools group, click **Options**. In the Options dialog box, click the **Data Flow Designer** link and select *Automatically add Join conditions*.

The inner join relationship is reflected in the Join Condition section and is represented graphically by the overlapping area in the Join Type diagram, as shown in the following image.
**Tip:** If you want to change the type to a right or left outer join, you can simply click the left or right circle. Try that now if you like, and then return to the inner join position. For details about join options, see *Joining Data Sources* on page 234.

3. Click **OK** to close the Join Editor.

**Next step:** You have now set up your source data. You are ready to select the columns from the source data to load into the data targets, which you will create shortly. To select and structure the columns, you will be using the SQL object mentioned earlier.

### Selecting and Managing Columns

You are now ready to select the columns of data you want to load into the data targets. You can perform a variety of operations on the selected columns.

Right-click the SQL object and notice the options on the menu.

- **Expand/Collapse.** Opens or closes an information window, as you saw earlier.

- **SQL Statement.** Displays the SQL code. Right now it simply reflects the join. It will be more interesting in a little while.

- **WHERE Filter.** Provides a calculator in which you can create an expression that limits record selection. For example, you might only want to retrieve records for a certain year.

- **HAVING Filter.** Provides a calculator in which you can create an expression that limits retrieval based on aggregated values after a GROUP BY. This option only appears when the Column Selection includes a GROUP BY.

- **Sort.** Provides a dialog box in which you can control the order in which data is retrieved.

- **Column Selection...** Opens a window in which you select the columns you want to include in your data target, and specify a variety of data retrieval requirements, as you will do in the following steps.

- **Add.** Allows you to add additional sources, selects, joins, and union objects.

- **Properties.** Opens a property panel that shows statistics for the selected statement.

- **Error Details.** This option only appears if there is an error in the columns or filters. When selected, a dialog box opens to explain the error.

### Select and Create Columns

1. Right-click the SQL column object and click **Column Selection.**

The Column Selection dialog box opens. You can identify column types at a glance:
2. You want the sales force to be able to track and evaluate product and store information, so double-click the columns PROD_NUM from INVINFO, and STORE_CODE from DMORD in the Available Columns list. This moves them into the Selected Columns list.

3. Next, you want to specify a time period. In this case, you want to select YEAR and MONTH, to provide the sales force with the ability to analyze sales on a monthly basis.

   Double-click the column YEARMONTH in the Available Columns list to move it into the Selected Columns list. (This is the column you created earlier as a Source Transformation.)

4. Now, add the columns you will use to measure sales. Double-click QUANTITY and LINEPRICE in the Available Columns list to add them to the Selected Columns list. The Column Selection window should now look like the image below:

   ![Column Selection Window](image)

5. Next, you want to calculate the line item cost of goods. This transformation can be done using an SQL calculation to create the column.

   To open the SQL Calculator, click the Insert Columns button above the Selected Columns list.

   a. In the Alias box, enter LINECOGS.
b. Ensure that the Columns/Variables tab is selected, and then double-click QUANTITY in the tree. The column appears in the Expression box.

c. Click the multiplication sign (*) on the calculator keypad.

d. Double-click COST in the tree to complete the expression.

The SQL Calculator should look like the image below:

6. Click OK to close the SQL Calculator and return to the Column Selection dialog box.

Notice that the expression has been added at the bottom of the Selected Columns list. The expression is in the first column and the alias you assigned is in the second column.

7. Since you want to generate month level detail, you must aggregate on the key columns, and then on the year/month column. This is done by adding the Group attribute to each column.

Under Selected Columns, click T1.PROD_NUM, then select Group By from the drop-down menu in the Aggregate column.

Repeat this step for T2.STORE_CODE and for T2.YEARMONTH.

8. Finally, let us sum the three measures: T2.QUANTITY, T2.LINEPRICE, and LINECOGS. You can multi-select them.
Under Selected Columns, press the Ctrl key and click T2.QUANTITY, T2.LINEPRICE, and LINECOGS. Choose Sum from the drop-down menu in the Aggregate column. Sum is applied to all three columns, as shown in the following image.

9. Click OK to close the Column Selection window and return to the object view in the Data Flow tab.

Filter the Data

You have identified the columns you want to extract from the source data, but the data in the source goes back several years, and only the last few years are really relevant to the salespeople who need to make product decisions for the coming year. Therefore, you are going to define a selection criterion to limit the retrieval to the appropriate years.

1. Right-click the SQL object again, and this time click WHERE Filter. The WHERE Filter Calculator opens to assist you in constructing the expression. Notice that this calculator is suitable for creating a wide range of selection criteria.

As you can see, the columns here are represented by the same icons used in the Selected Columns dialog box. Notice that YEARMONTH is a virtual column, rather than a real database column. You can filter on real or virtual columns.

2. To limit record retrieval to the time period beginning in January, 2006:
   a. Double-click YEARMONTH in the Column list. YEARMONTH is displayed in the Expression box.
b. Click \( \geq \) (greater than or equal to) in the calculator pad below the Expression window.

Click the numbers 200601 on the pad, or type 200601 directly into the Expression box to specify the year and month.

The expression in the WHERE Filter Calculator should look like the following image:

![WHERE Filter Calculator](image)

**Tip:** Though it is not required for this tutorial, you can build an SQL calculation using any ANSI SQL function. Simply click the *Functions* tab to see the available functions and arguments.

3. For now, click OK to complete the filtering expression.
4. Right-click the SQL object and click SQL Statement. The window should look like the following image:

![SQL Select Statement]

```sql
SELECT T1.PROD_NUM, --Product Number
       T2.STORE_CODE, --Company ID
       T2.YEARMONTH,
       SUM(T2.QUANTITY), --Quantity
       SUM(T2.LINEPRICE), --Line Total
       SUM(T2.QUANTITY * T1.COST) AS LINECOGS
FROM (dminv T1 INNER JOIN dmord T2
      ON T1.PROD_NUM = T2.PROD_NUM )
WHERE T2.YEARMONTH >= 200601
GROUP BY T1.PROD_NUM
       T2.STORE_CODE,
       T2.YEARMONTH
```

OK  Cancel
5. Click the Test SQL Statement button to see the results, as shown in the following image.

Note: The number of records retrieved will depend on the Run Options set from Tools and Options, and the database you choose for your sample data.

6. Click the red X to close the Test SQL Statement window, and click OK to close the Select Statement window.

Next Step: You are ready to create the data targets into which your source data will be moved, based on the mapping and rules you defined in the SQL columns object.

Adding Data Targets

You are going to add two data targets and specify the options you want to use when loading data into them.

- The first data target is a pre-existing table for sales reports that we designed just for this purpose. You are going to add data from last year to it.

- The second is a target table that you will create to provide additional analytic data that will be useful to the line managers.

The newly created target will be updated by other flows.

Map Columns in an Existing Data Target

To specify the first data target:

1. From the navigation pane, drag the synonym dmrpts from the ibisamp directory into the workspace, to the right of the SQL object. (The position to the right of the SQL object makes it a data target.)
2. Once you have added the data target to the data flow, you can specify how incoming data should be handled during the loading process.

Right-click the *dmrpts* target object and click *Properties*. For this exercise, you will accept the default Load Type, Insert/Update, as shown in the following image.

![Properties Window]

**Tip:** If you intend to run a flow more than once, it is advisable to clear the target before starting the run. This technique generally enhances performance in subsequent runs. In the *Prior to Load Option* in the Target Properties window, select the clearing option that suits the adapter type being used for your data targets. For example, for a relational table, the option is *Truncate table*. For a FOCUS/FDS or XFOCUS table, the option is *Drop table*.

3. Click X to close the Properties window.

The next step is to map the data source columns you are extracting to the data target columns.
4. Right-click dmrpts and this time, click Target Transformations. The Transformations dialog box opens.

5. Click the Automap button.

The five columns with identical names and data types are mapped and moved to the Expression tab. Note that the mapped columns are also selected in green in the Target Columns list, as shown in the following image.

6. Two target column names are deselected, so you need to create mappings or transformations for them. In the Target Columns list, YRMTH contains year and month data. You will map it to the YEARMONTH column in the Selected Columns list.
   a. Under Target Columns, click YRMTH.
   b. Under Selected Columns, click YEARMONTH.
   c. Click the equal sign (=) to move the mapping into the grid in the Expressions tab.

7. For the PROFIT column, you will build a transformation based on the definition of PROFIT as the difference between LINEPRICE and LINECOGS. As you can see, the data target is another place where we can apply transformations.
Under Target Columns, double-click **PROFIT** to move it to the Target Columns list in the Expressions tab, as shown in the following image.

8. To open the Transformations Calculator, double-click **PROFIT** in the Expressions tab. To create the expression in the Expression box of the calculator:
   a. In the tree, double-click **LINEPRICE**.
   b. On the calculator pad, click the minus (-) sign.
   c. In the Source Columns list, double-click **LINECOGS**.
9. Click OK to complete the expression and close the calculator.

The Expressions tab of the Transformations dialog box now looks like the following image:

10. Scroll to the right to see the last column. It may be useful to validate the incoming data. To do that, load only those records with a quantity greater than or equal to ten. Records that do not meet this validation criterion can be logged to a file for additional processing or for review at a later date.

Click the Validates tab in the Transformations dialog box. Then click the Insert Intermediate Transforms button to open the Validate Calculator.
11. To build the validation expression, double-click QUANTITY in the tree, and then select GE from the Relation drop-down menu. Input 10 into the Value field. Click OK. Your goal is to see which rows pass the validation test.

12. Now, let us test the validation and transformations to ensure that they are syntactically correct and performing the desired calculations. The test retrieves some rows from the server, applies the transformations, and displays the results.

Click the Test Transforms button in the upper-right corner, above the Validates tab.

The result of the transformation test appears, as shown in the following image.

![Test Transformations](image)

Note: The number of records retrieved will depend on the Run Options set from Options, in the Tools group on the Home tab, and on the database you choose for your sample data.

In the VALIDATE1 column, the number 1 represents rows that will be accepted based on the validation test. The 0s represent rows that will be rejected because QUANTITY is less than 10.

Tip: Since you have been working in the Transformations dialog box, testing from there is the simplest method, but you can also test the transformation by right-clicking the target object, clicking Expand, and clicking the Test Transforms tab.

13. Notice that Profits are positive in the answer set, which means that the transformation is working. It also satisfies the test condition, so you can move on to the next data target, which you will create from scratch.

14. Close the Test Transformation dialog box, and click OK to close the Transformations dialog box.
Create a Data Target

DataMigrator can load multiple data targets in a single data flow. Let us add another target object into the data flow. This time, the data target you need does not exist, so you will create it using the columns in the SQL SELECT statement that you defined for the SQL Select Columns object.

This target table will be used by line managers for competitive analysis. You will create the base table here, but other flows can update it with additional information.

1. Right-click in the workspace to the right of the SQL object, point to Add Target, and then click New. A new data target appears to the right of the SQL object.
2. Right-click the new target and click Properties.

   The Target Properties pane opens.
3. To define target properties:
   a. Choose the same type of adapter you used for the data sources in our example. In this case, it was MS SQL Server ODBC.
   b. Change the value for Synonym from baseapp/target01 to salesdemo/linerpts.
   c. Retain the default Load Type of Insert/Update.

   The target Properties pane should look like the following image:
d. Close the pane.

4. Right-click the new data target and click *Target Transformations*.

5. In the table you are creating, PROD_NUM, STORE_CODE, and YEARMONTH are all keys. Click each one to identify them, and those above it, as keys. The Transformations dialog box looks like the following image:
6. Click OK to return to the Data Flow tab, which now contains two data sources and two data targets, as shown in the following image.

![Data Flow Diagram]

**Next Step:** You are ready to save the data flow. Both data targets, dmrpts and linerpts, will be loaded based on your specifications.

**Save and Name the Data Flow**

Click the Save button in the Quick Access Toolbar.

The Save Procedure As window opens. Select which application directory to save the flow to using the Save In drop-down menu, and enter a flow name in the Name field. For this example, we will save our flow in the salesdemo application directory, and name it *salesdb*.

**Next Step:** You are ready to run the data flow.

**Submitting the Data Flow**

Now that you have completed the data flow, you can run it.

1. On the Home tab, in the Filter group, click *Procedures* to see all the procedures available. In the salesdemo application directory, right-click *salesdb* and click *Submit* from the menu.
   or

   On the Flow tab, in the Run group, click *Run* and click *Submit* from the drop-down menu.
2. When the execution completes, you can review the log, which records the process and identifies any problems that may have occurred.

From the navigation pane, right-click salesdb in the salesdemo application directory, point to Logs, and then click Last Log.

The Log report opens in the workspace as shown in the following image.

3. Close the log window to return to the Data Flow tab.

Notice that the Console Log indicates that the data flow has run successfully.

Data Flow Review

You have now completed Exercise 1, in which you built a data flow to move data out of the operational data store and into a denormalized table, so that the sales force can write meaningful sales reports, and the line managers can perform competitive analysis.

Next Step: When you are ready, proceed to Exercise 2 where you will create a process flow that uses the data flow you just built.

Exercise 2: Creating the Process Flow

A Process Flow controls how one or more data flows will be processed, by positioning a set of objects in the workspace and defining their interactions. You can combine:

- A Start object that defines when to run the process.
- Data Flow objects that indicate what data to extract, and copy from data sources to data targets.
- Email objects that notify users about the status of the process at specified points.
- Stored Procedure objects that perform a variety of supplementary tasks before or after the extraction, and load steps defined in a data flow.
- Connector objects, represented as arrows, that specify execution logic for the other objects included in the process flow. For example, what to do next when a step in the process succeeds or fails.
Group objects that specify the flow of a subset of objects within the total flow. For example, two objects in a group that must be processed simultaneously, rather than sequentially, with the object that follows next. This could be an email notification that is waiting until processing has been completed for both of the grouped objects.

**Scenario:** You will create a process flow that uses the data flow you created in the last exercise, and then branches to one of two procedures based on the success or failure of the execution of that data flow. The appropriate results will be recorded in a log to facilitate quick trouble-shooting, if necessary, or immediate communication to sales agents and line managers that the latest data is available for reporting.

**Tip:** Although not demonstrated in the tutorial, you can extend the success and failure branches of the process flow to trigger the distribution of email messages, either to those in charge of correcting errors (upon failure), or to those who need the current data (upon success).

**Next step:** The simplest way to proceed would be to continue to work with salesdb. You could click the Process Flow tab to see the salesdb data flow object in the workspace, and then add other elements to create the process flow.

However, this is not necessarily the best way to proceed. Instead, you will use a method that is more modular and flexible. You will start a new flow, then click the Process Flow tab and drag the Data Flow object into the Process Flow workspace in the correct position in the flow. With this method, each flow is saved separately; a Data Flow object can then be used in more than one Process Flow, and a Process Flow can be modified as needed and used to incorporate different data flow objects.

**Starting the Process Flow**

Start the process flow just as you started the data flow:

1. Right-click the salesdemo directory, point to New, and then click Flow.

   A workspace opens in the right pane. The Data Flow tab is active.

2. Click the Process Flow tab to change your design mode.

   To anchor the Process flow, the Start object is automatically added to the workspace.

   **Next step:** You are going to add other objects to the flow and connect them to one another.

**Adding a Data Flow Object to the Process Flow**

To add the data flow salesdb to the process flow:

1. Drag the salesdb data flow from the navigation pane into the Process Flow workspace to the right of the Start object.
2. Right-click the Start object in the workspace and drag toward the Data Flow object, forming an arrow that connects the two objects.

The Process Flow tab now looks like the following image:

![Diagram of the Process Flow tab](image)

**Save and Name the Process Flow**

The first time you save the process flow, you will assign a name to it.

1. Click the Save button. The Save Procedure As window opens.
2. Name the Process Flow salesflow.
3. Click Save to return to the Process Flow tab.

You will continue to save after each step.

**Next step:** Your next task is to create two stored procedures. You will connect these procedures to the data flow object so that one will run when the data flow executes successfully, and the other one will run when the data flow fails to execute properly.

**Branching in the Process Flow**

You want to be able to track success or failure when the data flow runs. You will do this by creating two procedures and including them in your process flow. Then you will supply the logic that invokes the appropriate one for the condition that is encountered.

Following the same modular model we have been using, you will create each of the stored procedures as a separate file that can be used in this process flow and others.
Create the Procedures

First, you will create a procedure that will appear in the log when the data flow executes correctly.

1. Right-click the salesdemo application directory, point to New, and then click Procedure.
2. A text editing window opens. Enter the following line of code:
   
   -TYPE SUCCESS!

3. Click the Save As button in the Quick Access Toolbar, and name this procedure success. Click Save.
4. Close the editing window.

Now you will create a stored procedure that will appear in the log when the data flow fails to execute properly.

5. Right-click the salesdemo application directory again, point to New, and then click Procedure.
6. The text editing window opens. Enter the code:

   -TYPE FAILURE!

7. Click the Save As button in the Quick Access Toolbar, and name this procedure failure. Click Save.

   If applicable, this entry will be displayed in the DataMigrator log. Close the editing window.

Both procedures are now listed in the salesdemo application directory, as shown in the following image.
Connect the Procedures to the Data Flow

The Process Flow tab should look like the following image:

![Diagram showing a process flow with branches to success and failure]

Next step: You have now created two branches from the data flow object to the procedures. Now you need to define the behavior associated with each branch of the flow. You will do that from the connector arrows.

Specify Your Connection Logic

Follow-up actions based on the success or failure of the data flow execution are initiated by connection logic, in the form of return codes.

In the tutorial, we will use the default connection logic associated with the green and red arrows that branch to the stored procedures you created. Since these procedures write a line of code in the DataMigrator log, you will quickly be able to see which path the execution followed.

Tip: The implementation we have chosen to demonstrate can be built quickly and offers easily demonstrable results. However, there are many ways you can use branching within a process flow. For example:

- You might wish to send one email message in response to a return code associated with success, and a different message to different recipients in response to a return code associated with failure.

  or
You might wish to initiate another data flow or process flow in response to a return code associated with success, but send an email message to someone charged with correcting processing errors in response to a return code associated with failure.

Your options are many and varied.

Let us now review the default logic that will trigger the running of your two stored procedures.

1. Double-click the green arrow connecting the data flow object to success. The Condition dialog box opens.
   
   OK Default Condition (Return Code is 0) is selected by default.

2. Click OK.

3. Now, double-click the red arrow that connects to the failed procedure. This time, FAILED Default (Return Code is not 0) is selected by default.

4. Click OK.

**Next step:** It is time to run the Process Flow to see the results as reflected in the DataMigrator log.

### Submitting the Process Flow

To see if the process flow is successful, and if the proper messages are displayed, follow the steps below.

1. On the Flow tab, in the Run group, click Run, and choose Submit from the menu to test the flow.

2. From the navigation pane, right-click salesflow in the salesdemo application directory, point to Logs, and then click Last Log.

   The Log report opens in the workspace, as shown in the following image.

   ![Log Report](image)

   The log shows the message

   **SUCCESS!**

   which was generated by the stored procedure called success.
3. If you click the highlighted line (Link to log), you will see the log of that step.

4. Close the log windows.

Congratulations. You have successfully completed Exercise 2.

You are now ready to create your own flows. Refer to *Designing a Data Flow* on page 193, and *Designing a Process Flow* on page 385, for detailed information about the options illustrated in this tutorial and many others that are available to you in the DMC.
DataMigrator enables you to view reports on how your flows run and what error messages the Server generates. Impact Analysis reports allow you to assess component usage and evaluate the effect of changes. You can print a graphical representation of a flow from the workspace.

You can also run reports and view logs from the iWay Web Console. For details, see the Server Administration manual or the Web Console online help.

In addition to the reports described in this topic, Data Profiling information is available from the Synonym Editor. Data Lineage information is also available from the Synonym Editor. For information, see Viewing Data Profiling Characteristics on page 543, and Viewing Data Lineage for a Column on page 552.

In this chapter:

- DataMigrator Log, Statistics, and Summary Flow Reports
- Scheduler Reports
- Impact Analysis Reports on Synonyms, Procedures and Columns
- Working With a Report
- Viewing a Detailed Flow Report and Log File
- Printing a Graphical Representation of a Data or Process Flow

DataMigrator Log, Statistics, and Summary Flow Reports

The DataMigrator logs and traces folder contains a link that opens the Logs and Traces dialog box. This provides access to viewing two reports:

- The Log report contains the message output generated from running flows. This information includes the steps taken during data extraction, table creation, and data load, and any errors that occur. The report can include the entire log, or the log for a specific time period. For process flows, the Log report provides drill-down capabilities, which allow you to select entries that display details of the logs of the individual flows in the process flow.
The Statistics report contains statistical information about flows, such as how many rows were written, when the flow ran, and how long it took. The report can include all statistics, or statistics over a specific time period, for all flows, or for different types of completed flows.

The Statistics report is a drill-down report, which allows you to select entries that display details about the individual flows.

**Note:** Statistics are not available for XML files described by multi-segment synonyms, or for flows that use load type Insert records from memory.

The Summary Flow Report flow contains information about all of the flows in the server path or in a specific application directory. It includes the flow type, application directory, description, modify date, and creator.

**Note:**

- DataMigrator reports support flow names of up to 64 characters.
- The DM log and statistic tables are not updated with messages when flows are initiated using Run, from the Run group, or context menu.

**Reference:** **Log Report Columns**

Log reports provide the information listed and described in the following table:

<table>
<thead>
<tr>
<th>Column</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Datetime</td>
<td>Date and time of log entry.</td>
</tr>
<tr>
<td>Message Code</td>
<td>Message code number.</td>
</tr>
<tr>
<td>Log Messages</td>
<td>Message strings in the log.</td>
</tr>
<tr>
<td>Application</td>
<td>Flow application directory.</td>
</tr>
<tr>
<td>Name</td>
<td>Flow name.</td>
</tr>
<tr>
<td>Job ID</td>
<td>Unique internal job name.</td>
</tr>
</tbody>
</table>

**Note:** If a log message contains a non-zero return code, the line will be shown in red.
**Reference: Statistics Report Columns**

Statistics reports provide the information listed and described in the following table:

<table>
<thead>
<tr>
<th>Column</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>User ID</td>
<td>User ID which executed a request.</td>
</tr>
<tr>
<td>Application</td>
<td>Flow application directory</td>
</tr>
<tr>
<td>Name</td>
<td>Flow name.</td>
</tr>
<tr>
<td>Log</td>
<td>Job ID as a link to the flow log.</td>
</tr>
<tr>
<td>Parent Log</td>
<td>Log for parent flow.</td>
</tr>
<tr>
<td>Start Datetime</td>
<td>Start date and time of execution.</td>
</tr>
<tr>
<td>End Datetime</td>
<td>End date and time of execution.</td>
</tr>
<tr>
<td>Total Runtime</td>
<td>Total run time for the flow.</td>
</tr>
<tr>
<td>Return Code</td>
<td>Internal return code for the flow.</td>
</tr>
<tr>
<td>Target Name</td>
<td>Target name.</td>
</tr>
</tbody>
</table>

**Note:** This, and the remaining columns, apply on a per target basis.

<table>
<thead>
<tr>
<th>Transactions</th>
<th>Number of rows processed.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accept</td>
<td>Number of rows accepted.</td>
</tr>
<tr>
<td>Nomatch</td>
<td>Number of rows rejected.</td>
</tr>
<tr>
<td>DBMS Error</td>
<td>Number of rows rejected by the database for a constraint violation, such as a Unique Index constraint, a Not Null constraint, or a user-specified constraint. For a file target, the number of rows rejected due to format conversion errors.</td>
</tr>
<tr>
<td>Format</td>
<td>Number of format rejections.</td>
</tr>
<tr>
<td>Invalid</td>
<td>Number of rows rejected for failing validation tests.</td>
</tr>
</tbody>
</table>
### Column | Function
---|---
Duplicates | Number of rejected duplicate records.
Input | Number of rows input.
Changed | Number of rows updated.
Deleted | Number of rows deleted.
Original Request | Original flow where dependency started.
Original User | Original user ID running dependent flow.
Dependency Count | Dependency number.
Message | Text explanation of error messages.
CPU Time | Total CPU time in milliseconds.

**Procedure:**  **How to Run a Log or Statistics Report**

You can access Log and Statistics reports to view information about executed flows. The Log report displays the messages generated by the flows that have been run from the DMC, and the Statistics report provides information about the flows that have been scheduled or submitted from the DMC. The Statistics report also provides access to the message log of each flow.

1. From the navigation pane, expand the **Workspace** folder. Expand the **Logs and Traces** folder.
2. Double-click **DataMigrator Log and Statistics** or right-click it and click **View**.
   The Log and Statistics dialog box opens.
3. From the **Report Type** drop-down menu, select **Log** or **Statistics**.
4. Select values for the following fields and options to specify the information shown in the Log or Statistics report.

**Report Range**

Allows you to select what time range you want to run the report for. The options are Last, All, or Specify Time. If Specify Time is selected, an area opens where you can specify the start and end date and time. If Statistics is selected, the default value is All. If Log is selected, the default value is Last.
**Completion Status**

Allows you to view all reports, completed reports, reports completed successfully, or reports completed with errors. This option is only available when running a Statistics report.

**Run by**

Allows you to view reports for all flows, and those initiated by the scheduler or submit. This includes flows initiated from the DMC, WC, and CMRUN. This option is only available when running a Statistics report.

**Application or Flow Name**

Allows you to specify a specific application, flow, or procedure to run the report on.

5. Click View Report.

The report opens in the workspace.

An example of a Log report is shown in the following image.

![Log report example](image)

**Note:** For process flows, you can drill down to details of the flow components by double-clicking the blue Job ID line in a Log report.

An example of a Statistics report is shown in the following image.

![Statistics report example](image)

To view the message log for an individual flow, double-click its link or right-click the link and click Show Log in a Statistics report. If the log has output, right-click the log, and click Output. Failed flows will be shown in red.
To delete a log and its output, select its corresponding check box in a Statistics report, and on the Report tab, in the Tools group, click Delete Selected.

**Procedure: How to Run a Summary Flow Report**

A Summary Flow Report displays all of the flows on the server or in a specific application directory.

1. In the navigation pane, right-click Application Directories or a specific application directory folder. Point to Impact Analysis, then Flow Report, and then click Summary.

2. If you select to view the report from Application Directories, all of the flows will be included. If you selected a specific application directory, the flows in that particular application directory are displayed.

The report opens in the workspace, as shown in the following image.

![Image of a Summary Flow Report](image)

**Scheduler Reports**

The Scheduler reports folder include two reports:

- The Scheduler Log report displays the date/time when scheduled flows were submitted, as well as their Job IDs. The report can include the entire log, or the log for a specific time period.

  The Scheduler Log report is a drill-down report, which allows you to click entries to display details about the individual flows.

- The Scheduled Events report shows what flows are scheduled to run on a specified day. The default is to display information about the current date.
**Procedure:** How to Run a Scheduler Log Report

The Scheduler Log report displays the date/time when scheduled flows were submitted, as well as their Job IDs.

1. From the navigation pane, expand the *Workspace* folder. Expand the *Logs and Traces* folder.
2. Right-click *DataMigrator Log and Statistics* and click View.

   The Log and Statistics dialog box opens.
3. For Report Type, select *Statistics*, for Report Range, select *All*, and for Run by, select *Scheduler*.

The report of scheduled flows opens.

**Note:** Double-clicking a Link to log row will display the log for that flow.

**Reference:** Scheduler Log Report Columns

Scheduler Log reports provide the information listed and described in the following table:

<table>
<thead>
<tr>
<th>Column</th>
<th>Format</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUBMIT_DATE_TIME</td>
<td>Alpha 20</td>
<td>Submit date and time for flow.</td>
</tr>
<tr>
<td>APPNAME</td>
<td>Alpha 64</td>
<td>Flow application directory</td>
</tr>
<tr>
<td>REQUEST</td>
<td>Alpha 64</td>
<td>Flow name.</td>
</tr>
<tr>
<td>LOG</td>
<td>Alpha 64</td>
<td>Link to the flow log.</td>
</tr>
</tbody>
</table>

**Procedure:** How to Run a Scheduled Events Report

The Scheduled Events reports provides information about what time flows are scheduled to run on a particular day.

1. In the navigation pane, right-click an application directory, point to *Schedule and E-mail*, and then click *Scheduled Events*.

   The Scheduled Events window opens.
2. The report defaults to today. Optionally, you can select a date and time in the future.
3. Click View Report.

The report opens in the workspace, as shown in the following image.

**Procedure:** How to Run a Scheduler Agents Report

The Scheduler Agents report provides information about every scheduled flow and the next and last runs.

In the navigation pane, right-click an application directory, point to *Schedule and E-mail*, and then click *Scheduler Agents*.

The Scheduler Agents report opens, as shown in the following image.

**Reference:** Scheduler Agents Report Columns

Scheduled Events reports provide the information listed and described in the following table:

<table>
<thead>
<tr>
<th>Column</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>The identification number associated with the scheduler agent.</td>
</tr>
<tr>
<td>Column</td>
<td>Function</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>State</td>
<td>The current state of the scheduler agent.</td>
</tr>
<tr>
<td>Application</td>
<td>The application name of the scheduled procedure.</td>
</tr>
<tr>
<td>Procedure</td>
<td>The name of the scheduled procedure.</td>
</tr>
<tr>
<td>Report</td>
<td>An optional report name.</td>
</tr>
<tr>
<td>User ID</td>
<td>The user credentials used to run the scheduled procedure.</td>
</tr>
<tr>
<td>Schedule Type</td>
<td>How often the procedure is run.</td>
</tr>
<tr>
<td>Last Run Time</td>
<td>The last time the procedure ran.</td>
</tr>
<tr>
<td>Next Run Time</td>
<td>The time of the next scheduled run.</td>
</tr>
</tbody>
</table>

**Reference:** **Scheduled Events Dialog Box**

To access the Scheduled Events dialog box from the DMC, expand the *Workspace* foldertab. Then, expand the *Special Services and Listeners* folder, right-click *Scheduler* and click *Scheduled Events*. You can also right-click an application directory, point to *Schedule and E-Mail*, and then click *Scheduled Events*.

The DM Scheduled Events window has the following fields and options:

**Start**

Start of the specific time period for generating this report.

**Note:** The date-field drop-down menu launches a calendar for selecting dates. The time-field scroll buttons allow you change the highlighted section of the time. For example, you can change the hour, minute, or second.

**End**

End of the specific time period for generating this report.

**Note:** The date-field drop-down menu launches a calendar for selecting dates. The time-field scroll buttons allow you change the highlighted section of the time. For example, you can change the hour, minute, or second.

**Impact Analysis Reports on Synonyms, Procedures and Columns**

**Note:** Both flows and stored procedures are referred to generically as procedures.
Impact Analysis Reports on procedures, synonyms, and columns provide information on the flows and stored procedures available on a particular server, and the synonyms and columns they use. They allow you to evaluate changes before they are made by showing which components will be affected. Impact Analysis Reports provide quick analysis of synonym and column usage, and allow you to get procedure information without opening them individually. These reports are drill-down reports, which allow you to double-click entries to open the procedures in the workspace. Impact Analysis reports are also available for individual synonyms and columns.

There are six types of reports:

- **Synonyms by Procedure report.** Lists your procedures and the synonyms that they use. This list is sorted by procedure. The report includes whether the synonym is used as a source, a target, or iterator, the adapter type, the table name and description, load type and prior to load options, and directory locations.

- **Procedures by Synonym report.** Lists your synonyms and the procedures that use them. This list is sorted by synonym. The report includes whether the synonym is used as a source, a target, or iterator, the adapter type, the table name and description, load type and prior to load options, and directory locations.

- **Columns by Procedure report.** Lists the columns used in your procedures and their synonyms. This list is sorted by data flow. The report includes how the column is used, as well as directory locations.

- **Procedures by Column report.** Lists your procedures and their synonyms. This list is sorted by column. The report includes how the column is used, as well as directory locations.

- **Synonyms by Procedure - Enterprise report.** Includes the same information as a Synonyms by Procedure report, but for the server and all its subservers in a hub or subserver environment. This report does not appear if the server you are reporting on does not have any remote servers configured.

- **Procedures by Synonym - Enterprise report.** Includes the same information as a Procedure by Synonym report, but for the server and all its subservers in a hub or subserver environment. This report does not appear if the server you are reporting on does not have any remote servers configured.

**Note:** Enterprise reports do not appear if you do not have any subservers.
Procedure: How to Run an Impact Analysis Report on Synonyms, Procedures, and Columns

Note: Both flows and stored procedures are referred to generically as procedures.

Impact Analysis reports provide information on the procedures, synonyms, and columns available on a particular server. They allow you to evaluate changes before they are made, by showing which components will be affected. The reports can be sorted by procedure, by synonym, or by column. The results can be exported to Microsoft Excel, copied as text, or printed. For details, see Working With a Report on page 481.

1. In the navigation pane, right-click Application Directories, or a specific application directory, and point to Impact Analysis.

2. From the shortcut menu, click Synonyms by Procedure, Procedures by Synonym, Columns by Procedure, or Procedures by Column. You can also point to Flow Report and click a Summary or Detailed report.

The report opens in the workspace.

These reports include the procedure and synonym directories, columns and column types, synonym usage (source or target), adapter type, table name, description, and prior to load and load type options. These reports are drill-down reports, which allow you to double-click the procedure name entries to open the flows or stored procedures in the workspace.

The Synonyms by Procedure report shown in the following image lists each procedure and the synonyms it uses.

The Procedures by Synonym report shown in the following image lists each synonym and the procedures that use it.
The Columns by Procedure report shown in the following image lists each procedure and the synonyms and columns it uses.

![Columns by Procedure](image1)

The Procedures by Column report shown in the following image lists each column by synonym and by the procedures that use it.

![Procedures by Column](image2)

**Procedure: How to Run an Impact Analysis Report on a Particular Synonym**

An Impact Analysis report on a particular synonym provides information on the procedures and reports that use it and how it is used. It allows you to evaluate changes to synonyms before they are made, by showing which components will be affected. The results can be exported to Microsoft Excel, copied as text, or printed.

These reports are drill-down reports, which allow you to double-click entries to open procedures in the workspace to make the necessary changes.

1. Right-click a synonym in the navigation pane and click *Impact Analysis*.
   
   The Impact Analysis report opens in the workspace.

2. Double-click a procedure name to open it in the workspace.

**Procedure: How to Run an Impact Analysis Report on a Particular Column**

An Impact Analysis report on a particular column provides information on the procedures and reports that use it and how it is used. It allows you to evaluate changes to the column before they are made, by showing which components will be affected. The results can be exported to Microsoft Excel, copied as text, or printed.
These reports are drill-down reports, which allow you to double-click entries to open procedures in the workspace to make the necessary changes.

1. Double-click a synonym in the navigation pane to open it in the Synonym Editor.

2. Right-click a column in the Synonym Editor and click Impact Analysis, as shown in the following image.

![Impact Analysis Image]

The Impact Analysis report opens in the workspace.

3. Double-click a procedure name to open it in the workspace.

**Procedure: How to Run an Impact Analysis Report on a Particular Procedure**

An Impact Analysis report on a particular procedure provides information on that procedure, reports that use it, and how it is used. It allows you to evaluate changes to the procedure before they are made, by showing which components will be affected. The results can be exported to Microsoft Excel, copied as text, or printed.

These reports are drill-down reports, which allow you to double-click entries to open procedures in the workspace to make the necessary changes.

1. Right-click a procedure in the navigation pane, point to Analysis Reports, and click Impact Analysis.

The Impact Analysis report opens.
2. Double-click a procedure name to open it in the workspace.

**Note:** The Analysis Reports folder also gives you access to the Dependency Analysis and Data Lineage reports.

**Reference:** Synonym, Data Flow, and Column Impact Analysis Report Columns

Synonym, Procedure, and Column reports provide the information listed and described in the following table:

<table>
<thead>
<tr>
<th>Column</th>
<th>Format</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Found In</td>
<td>Alpha 32</td>
<td>Flow or stored procedure name.</td>
</tr>
<tr>
<td>Found In Application</td>
<td>Alpha 18</td>
<td>Flow or stored procedure application directory.</td>
</tr>
<tr>
<td>Found In Type</td>
<td>Alpha 18</td>
<td>Indicates if the object was found in a stored procedure or flow.</td>
</tr>
<tr>
<td>Synonym</td>
<td>Alpha 64</td>
<td>Synonym name.</td>
</tr>
<tr>
<td>Synonym App</td>
<td>Alpha 64</td>
<td>Synonym application directory.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Note:</strong> This value will only appear if you are using application directories.</td>
</tr>
<tr>
<td>Usage</td>
<td>Alpha 36</td>
<td>How the synonym or column is used, for example, the source, target, or transformation.</td>
</tr>
<tr>
<td>Line</td>
<td>Alpha 8</td>
<td>The line number where the item appears.</td>
</tr>
<tr>
<td>Adapter</td>
<td>Alpha 64</td>
<td>Synonym format.</td>
</tr>
<tr>
<td>Table Name or Dataset</td>
<td>Alpha 133</td>
<td>Data target table name.</td>
</tr>
<tr>
<td>Synonym Remarks</td>
<td>Alpha 78</td>
<td>Synonym description.</td>
</tr>
<tr>
<td>Prior to Load</td>
<td>Alpha 13</td>
<td>Prior to load option for targets.</td>
</tr>
<tr>
<td>Load Type</td>
<td>Alpha 35</td>
<td>Load type for targets.</td>
</tr>
<tr>
<td>Procedure Description</td>
<td>Alpha 150</td>
<td>Flow or stored procedure description.</td>
</tr>
<tr>
<td>Column</td>
<td>Alpha 66</td>
<td>Column name.</td>
</tr>
</tbody>
</table>
## Working With a Report

You can use report data by copying it into another application, or by exporting it into Microsoft Excel. You can also manipulate the display of a report and print it.

For information about the Reports toolbar, which becomes active when you run a report, see *Reports Tab* on page 158.

### Procedure: How to Copy Report Output to the Clipboard

You can copy all or part of a report to the clipboard and paste it into a text editor or other application.

After running a report:

1. Select the part of the report that you want to copy. To copy the entire report, click the upper left hand cell, which selects all rows and columns.
   
   **Note:** Select a row by clicking its row number. Select a column by clicking its title. Use the Shift and Ctrl keys to make multiple row or column selections.

2. On the Report tab, in the Tools group, click *Copy as Text*, or right-click the selected text and click *Copy as Text*.

3. In another application, select *Paste* from the Edit menu or use Ctrl+V to copy the selected data.

### Procedure: How to Export a Report to Microsoft Excel

You can export all or part of a report to Microsoft Excel.

**Note:** This option is only available with the Alpha or Original output format.

After running a report:

1. Select the part of the report that you want to export. If you do not select a part of the report, the entire report will be exported.

   **Note:** Select a row by clicking its row number. Select a column by clicking its title. Use the Shift and Ctrl keys to make multiple row or column selections.
2. On the Report tab, in the Tools group, click Export to Excel, or right-click the selected text and click Export to Excel.

The report opens in Microsoft Excel, as shown in the following image.

![Microsoft Excel report](image)

**Procedure:** How to Change the Report Font

You can change the font used in reports.

1. On the Home tab, in the Tools group, click Options.

   The Options dialog box opens.

2. Click the Format link on the User Preferences tree.

   The Format settings page opens in the right pane.

3. Click DataMigrator Reports in the Fonts text box.

4. Click the Font button.

   The Font dialog box opens.

5. Select the font attributes, and click OK.

   The changes will be applied the next time you run a report.
Procedure: How to Change the Report Format

You can change the format for reports.

1. On the Home tab, in the Tools group, click Options.
   The Options dialog box opens.
2. Click the Run Options link on the User Preferences tree.
   The Run Options settings page appears in the right pane.
3. From the Test reports default format drop-down menu, select which format you would like the report to be in.
   The format options are the following:

   - HTML

<table>
<thead>
<tr>
<th>User ID</th>
<th>Application</th>
<th>Name</th>
<th>LOG</th>
<th>Parent Log</th>
<th>Start Datetime</th>
</tr>
</thead>
<tbody>
<tr>
<td>dm</td>
<td>salesdemo</td>
<td>salesdb</td>
<td>20090616170046-25010b34</td>
<td>20090616170046-2ef59b76</td>
<td>2009/06/16 17:00:48</td>
</tr>
<tr>
<td>dm</td>
<td>salesdemo</td>
<td>salesdb</td>
<td>20090616170046-25010b34</td>
<td>20090616170046-2ef59b76</td>
<td>2009/06/16 17:00:48</td>
</tr>
</tbody>
</table>

   - HTML - plain text

<table>
<thead>
<tr>
<th>User ID</th>
<th>Application</th>
<th>Name</th>
<th>LOG</th>
<th>Parent Log</th>
<th>Start Datetime</th>
</tr>
</thead>
<tbody>
<tr>
<td>dm</td>
<td>salesdemo</td>
<td>salesdb</td>
<td>20090616170046-25010b34</td>
<td>20090616170046-2ef59b76</td>
<td>2009/06/16 17:00:48</td>
</tr>
<tr>
<td>dm</td>
<td>salesdemo</td>
<td>salesdb</td>
<td>20090616170046-25010b34</td>
<td>20090616170046-2ef59b76</td>
<td>2009/06/16 17:00:48</td>
</tr>
</tbody>
</table>

   - active report
Working With a Report

Excel

Sheet1

PDF

DM Server Statistics

<table>
<thead>
<tr>
<th>User ID</th>
<th>Application</th>
<th>Name</th>
<th>LOG</th>
<th>Parent Log</th>
<th>Start Datetime</th>
</tr>
</thead>
<tbody>
<tr>
<td>dm</td>
<td>salesdemo</td>
<td>salesdb</td>
<td>20090616120090616170046-22009/06/16 17:00</td>
<td>20090616170046-2ef59b7e</td>
<td></td>
</tr>
<tr>
<td>dm</td>
<td>salesdemo</td>
<td>salesdb</td>
<td>20090616120090616170046-22009/06/16 17:00</td>
<td>20090616170046-2ef59b7e</td>
<td></td>
</tr>
</tbody>
</table>
Procedure: How to View the Top Level Procedure

When you schedule or submit a process flow that calls other flows, it can be useful to see both the name of the flow that is currently running, and the one that started the process while viewing the Data Services Agents report.

To view a top level procedure:

1. From the DMC, with an Application or Server Administrator user ID, expand the Workspace folder.
2. Right-click Data Services and click Agents.
3. On the Report tab, in the Tools group, click View and then click Choose Columns.
   The Select the columns dialog box opens.
4. Select the Top Level Procedure check box from the Category/Title column.
5. Click OK.
6. Click the Refresh button.

As shown in the following image, the dsloadfact procedure was called by the Top Level Procedure ibisamp/dsflows.
Procedure: How to Print a Report

When viewing a report on the screen, you can print the report by using any of the following methods:

- Clicking the Print button from the Quick Access Toolbar.
- Right-clicking anywhere in the report and selecting Print from the context menu.
- Pressing Ctrl+P.

The standard Print dialog box opens, where you can select a printer destination, the number of copies, and other print properties.

Note: If you selected a range of cells in the report, you can click the Selection radio button to only print that range.

Procedure: How to Print Preview a Report

When viewing a report on the screen, you can preview the report by using any of the following methods:

- Clicking the Print Preview button from the Quick Access Toolbar.
- Right-clicking anywhere in the report and selecting Print Preview from the context menu.
- Pressing Ctrl+I.

The print preview window opens in the workspace.

Reference: Print Preview Tab

The Print Preview tab contains groups and commands that allow you to see a preview of the report that you are going to print. The Print Preview tab contains the following groups:

Print Group

The Print group consists of the following command.

Print

Prints the report.

Zoom Group

The Zoom group consists of the following commands.

Zoom In

Zooms in on the page.
Zoom Out
Zooms out of the page.

One Page/Two Page
Switches the view between one or two pages.

Preview Group
The Preview group consists of the following commands.

Next Page
Shows the next page of the report.

Previous Page
Shows the previous page of the report.

Close
Closes the print preview.

Procedure: How to Print a Report for Specific Pages
When viewing a report on the screen:
1. Select Print from the Quick Access Toolbar.
   The Print dialog box opens.
2. Select the Pages radio button in the Print Range section, and enter the page numbers that you want to print.
3. Click OK.

Procedure: How to Print a Report for Selected Rows or Columns
When viewing a report on the screen in a grid:
1. Select the columns you want to print by clicking on the title bar. You can also select specific rows and columns by pressing the Ctrl or Shift key.
2. Select Print from the Quick Access Toolbar.
   The Print dialog box opens.
3. Select the Selection radio button from the Print Range section.
4. Click OK.
Viewing a Detailed Flow Report and Log File

A detailed flow report displays general properties like name, description, DM release, creator, and modify date. It also provides specifics on SQL Select Statements, target information, load options, target transformations, and execution properties. These reports are useful for documentation purposes.

Detailed reports are available for individual flows from the DMC. You also can generate detailed reports for all of the flows in the path of the server, or in a specific application directory from the Web Console.

The log file for a flow is useful in determining problems encountered when running a flow.

Procedure: How to View a Detailed Flow Report

Right-click a flow in the navigation pane and click Flow Report.

The flow report shown in the following image appears in the workspace. It can be printed from that window.

```
Flow Type: Data Flow

Name of Flow: dmdemo
Description: Join with source and target transformations, sql calculations
iWay DataMigrator Release: 7705
Created by: dm
Last Modified Date: 2012/02/06 22.37.10

Source Information
Source Name: DMINV
Source Adapter: MS SQL Server
Connection: CON01
Synonym Application Directory: Sample - Inventory
Real Table Name: dminv

Source Name: DMORD
Source Adapter: MS SQL Server
Connection: CON01
Synonym Application Directory: Sample - Orders
Real Table Name: etoinf

Source Transformations
Source: DMORD
```
**Tip 1:** You can also copy the report to the Clipboard and paste it in Word or any other word processing program. To do this, right-click the report and click Select All. Then right-click and select Copy. You can now paste the report into another program.

**Tip 2:** If you need to export the flow information into an XML file, you can use the stored procedure flow2xml, which is located in the ibisamp application directory. The following steps show how you can use this procedure and send the complete flow information into an XML file.

1. Right-click the flow2xml flow in the ibisamp directory and click Run.
2. A dialog box opens. In the first line, enter ibisamp as the name of the application directory. In the second line, enter dmdemo as the name of the flow.
   **Note:** These values are case-sensitive.
3. Click the Run button.

The procedure will run. If it runs successfully, the following message will appear in the Console Log: The XML file fl_xml_*_.xml created; located in the same directory where your flow is.

The fl_xml_dmdemo.xml file will be inside the ibisamp directory.

**Procedure:** How to View and Print a Set of Detailed Flow Reports

You can generate detailed reports for all of the flows in the server path or in a specific application directory.

**Note:** If there are many flows in this report, it may take a while to generate.

1. Right-click the server or an application directory, select Impact Analysis, and then click Flow Report.
2. Select Detailed from the context menu.

For output in default format, the first report is displayed. On the Report tab, in the Run group, click Next Result to view the next report, and continue until you have viewed all the reports.

For pdf output, a single multi-page document is displayed.
The flow reports appear in the workspace, as shown in the following image.

Flow Type: Data Flow

Name of Flow: dmdcode

Description: Decode function used inline and from file
iWay DataMigrator Release: 7705
Created by: dm
Last Modified Date: 2012/02/06 23.17.24

Source Information
Name: DMORD
e Adapter: MS SQL Server
Connection: CON01
Synonym Application Directory: Sample - Orders
Real Table Name: dmoinf

SQL Select Statement
SELECT DISTINCT T1.PLANT --Location Of Manufacturing Plant FROM dmord T1

Target Information
Target Name: DMLOCS
Target Type: Existing
Target Adapter: Flat File
Connection: .
Format: 

Note: The Next Result common in the Run group, on the Report tab, lets you navigate to the next flow report.

Procedure: How to View the Log File

You can generate a log file for reports that have been run.

1. Right-click a flow in the navigation pane, select Logs, and then click Log and Statistics.
   The Log and Statistics dialog box opens.

2. From the Report Type drop-down menu, select Log.
   Specify what time range you want to run the report for. The options are Last, All, or Specify Time.

   The log report appears in the workspace.
Example: Sample DataMigrator Log File

This example contains sample lines from log files, and an explanation of the code.

The first lines of a log display DM job initialization.

- This request was submitted:

  ------- Start of Log Record for baseapp/one -------
  Request - one (Owner: dm) submitted.

The next lines display the data target or file type, and whether it is a new or existing data target.

- The following is a new table, and DataMigrator attempts to drop and recreate the target table:

  linertps type MS SQL Server New target
  PREPARE executed
  Create synonym executed for: linertps
  Issuing DROP TABLE for linertps
  Issuing CREATE TABLE for linertps

- The following is an existing table:

  dmrpts type MS SQL Server Existing target
  PREPARE executed
  ROWS AFFECTED BY PASSTHRU COMMAND : 222/DELETE

The log then displays summary data for the request. You can see the number of records that were processed.

- In this example, there were 200 rows processed, which DataMigrator attempted to insert into the target table. Of those, 20 were rejected due to duplicate keys, and 180 were accepted by the data target.

  Starting Load-- stats for newtab ----------------------------------------------------------
  184 : Row(s) processed by job
  144 : Row(s) accepted by target table
  144 : Row(s) inserted into target table
  0 : Row(s) updated in target table
  0 : Row(s) deleted from target table
  40 : Row(s) rejected due to validation errors
  0 : Row(s) rejected due to no match logic
  0 : Row(s) rejected because duplicate exists

The remaining lines contain summary data and the return code from the flow.

- In the following, a return code of zero (0) means the request completed successfully:
Printing a Graphical Representation of a Data or Process Flow

As a flow gets more complex, it can be helpful to print out a graphical representation. This printed version of the flow can facilitate analysis and assist in troubleshooting. Flows are printed from the application that is associated with bitmap files on your desktop.

Procedure: How to Print a Graphical Representation of a Flow

1. Open the flow in the workspace.
2. Click the Print button on the Quick Access Toolbar.
   An image of the flow opens in the application that is associated with bitmap files on your desktop.
3. Select Print from the application File menu.

Note: If no application is associated with bitmap files, the graphical representation will open in the default Microsoft Picture and Fax Viewer.
Optimizing Performance

The run-time performance of DataMigrator depends on many factors: host platform, network protocol, I/O devices, type of SQL command, and selected DBMS. Depending on your environment, one or more of these factors may be in your control. The following topics outline the options that have the greatest impact on performance.

The issues described here may or may not affect your particular configuration. After careful review and comparison, you may determine that your request will need a lot of internal processing and significant I/O. This may be fine if your nightly incremental updates are low in volume or if you are on a processor where disk access is not a bottleneck. However, if you have only one hour to process your data and the server is on a small machine with few resources, an alternate approach may be necessary.

In either case, review these performance considerations, and compare them to your environment. For most applications, the default functionality of DataMigrator is sufficient for your data management needs. If the performance objectives cannot be met, these guidelines can at least provide you with some insight into the location of potential bottlenecks and possible strategies to overcome them.

In this chapter:

- Improving Data Extraction Performance
- Using Automatic Passthru
- Improving Data Loading Performance
- Improving Lookup Performance

Improving Data Extraction Performance

DataMigrator extracts data from data sources using SQL. The process of applying SQL to heterogeneous data sources is very complex and too broad a subject to easily outline here. However, the following are issues which you control.

- **Providing qualifiers to the data source.** iWay uses index structures (if available) to retrieve rows in the most efficient manner. Therefore, it is advisable to supply screening conditions in your WHERE criteria for as many columns as possible.
Make sure that critical native DBMS optimization fields (primary keys, secondary indexes, and so on) are correctly defined in the synonym. For more information on defining synonyms, see Chapter 6, Managing Metadata, in the Server Administration manual.

- **Evaluating your joins.** iWay is flexible, enabling similar data source, cross-data source, and even cross-server (sometimes called cross-platform) joins. The trade off for this flexibility is the amount of work required of the server, and the number of intermediate files being created. To minimize the impact of joins, try the following:

  - Join your tables using key columns and indexes.
  - Minimize the number of tables used in a single SELECT command.
  - Specify the smaller of the two tables being joined first.

### Using Automatic Passthru

Automatic Passthru (APT) is the ability of a DataMigrator Server to pass or hand-off an intact SQL command to a remote server or RDBMS subsystem. In effect, it enables a user to speed up DataMigrator. If a request does not meet the requirements for APT, the DataMigrator Server needs to do additional work. SQL commands must be translated into a format (Data Manipulation Language) that is understood by the underlying (typically non-relational) file system(s), forcing the need for a special work file to be created on the DataMigrator Server platform. DataMigrator waits for this temporary work file to be written, and then continues processing by loading the RDBMS table.

If a request meets the requirements for APT, DataMigrator uses its streaming or pipeline methodology. This means no intermediate files are written by DataMigrator. Rows are fetched directly from the source (effectively off the wire for communication scenarios) and are then written to the target. Whether DataMigrator is writing a local flat file, or sending dynamic SQL to an RDBMS table, an SQL request that runs in APT mode guarantees the best default performance.

The following environmental and request-based factors most often result in requests failing the requirements of APT:

- Non-relational file structures accessed directly at the server.
- DEFINE commands in a SUFFIX=EDA synonym.
- Cross-platform joins.

For more information on passthru functionality, see the Adapter Administration manual.
Avoiding Non-Automatic Passthru

If non-APT processing is critical to your extraction process, but is determined to be a bottleneck, it may help to add a secondary server to your configuration. The critical definition for DataMigrator (and its need to use a HOLD file) is the definition of APT at the server where DataMigrator is installed. For example, if you are reading from a C-ISAM file on UNIX (non-relational), describe the file to a second server configured expressly to avoid a non-APT condition. Server-to-server shipment of rows may outperform the disk I/O resulting from a HOLD file. This is illustrated in the following figure:

This technique requires that the additional server have access to a synonym describing the C-ISAM structure, and that the primary server is configured correctly for communicating with the secondary server. An iWay synonym must be created to point to the correct destination. DataMigrator will use this new synonym for the SQL request, instead of the original C-ISAM Master definition (see the server manual for your platform for more information on configuring an iWay hub and remote servers). Once again, the configuration work needed to avoid non-APT may be over-engineering in a situation where row volume, disk I/O, or processing power is not an issue.

Another way to avoid non-APT processing is to place your transformations in the synonym of the remote server. This is accomplished by entering a DEFINE in the synonym to ensure that calculating and processing work is done directly on the platform where the data resides.
**Reference:** Determining the Use of Automatic Passthru

You can tell if a flow is processed using APT or non-APT by looking at the DataMigrator ETL Log. A message appears towards the top of the output for every flow, indicating its capacity for passthru processing.

**Reference:** Using Real Numbers With Direct Passthru

Real or floating-point numbers, also referred to as approximate numeric, are numbers stored in scientific notation.

When iWay retrieves numeric data, it is converted to an intermediate display format. For flat files, FOCUS/FDS files, and relational data sources, the display format is that found in the synonym. However, when direct passthru is used, for example in a DBMS SQL flow, then iWay does not use the synonym. In this case, iWay uses a default format of 20.2 (that is, 20 digits with two places to the right of the decimal). This format may not be appropriate for data warehousing needs.

You can change the format with an SQL command. You can include this command in the profile on the remote server (where the data source resides). If you do so, all real numbers will be rounded to four digits.

If different tables in different data sources have different precisions, set the precisions for each process flow with a pre-extract procedure. If the data source resides at the server, create a pre-extract procedure with this command.

**Syntax:** How to Change the Format of Real Numbers for Direct Passthru

```sql
ENGINE database SET CONVERSION {FLOAT|REAL} PRECISION nn [mm]
```

where:

- `database`
  - Is the type of data source to which you are connected.

- `nn`
  - Is the precision. This number must be greater than 1 and less than 20.

- `mm`
  - Is the scale (the number of places to the right of the decimal).

**Example:** Changing the Format of Real Numbers

For example, if your data source is Db2, and you want to set the number of digits to the right of the decimal to 4, issue the following command:
Reference: Direct Load Flows

Normally for flat file sources, DataMigrator extracts data from the files and creates an intermediate flat file. Then it invokes the load program to load from the flat file to target.

However, when the source for a DataMigrator flow is a single file, it is possible to load directly from that file using what is called a Direct Load Flow. This type of flow has no SQL object (there are no filters, sorts, or aggregations). There is a direct connection from source to target. Since there is no intermediate file created, this reduces disk space requirements, and can also reduce the time for the load.

Note that you cannot reference any DEFINEd columns in the synonym for the source. Also, each column in the source synonym must have an ALIAS.

The target table must be an existing table, and you can use Target Transformations.

Improving Data Loading Performance

You can improve performance when loading data into the data target. Keep the following suggestions in mind.

- **Prior to load options.** Before you load data into an existing table (for most RDBMS targets), DataMigrator gives you the option to Delete all rows from table or Truncate table. When deleting the rows, the relational database internally logs the rows deleted in case a ROLLBACK is issued, which for large tables can be slow. Truncating the table does not do any logging, so it is much quicker, but then a ROLLBACK cannot be issued.

- **Commit.** For relational database and FOCUS targets, DataMigrator lets you specify the number of rows before a COMMIT (or CHECK) is issued. The default value is 1000. A higher value generally improves throughput, but this depends on the memory available on the server, the database, the number of columns loaded, and their sizes. Increasing the commit value too much can slow performance. Try different values to see what works best.

- **Insert/Update.** This is the default load method. For a relational database target:
When include duplicates is selected each incoming row is sent to the target table without first checking to see if the key value is found. This generally provides the best loading time. If there is a duplicate key, the database will reject the row with a unique constraint violation. When reject or update duplicates is selected, the keys of each incoming row are matched against the target table. This additional step does slow processing, but allows for matched rows in the target table to be updated. In this case, care should be taken to use as few columns as keys as possible to uniquely identify the row. Note that if no columns are identified as key columns, the entire column is considered to be the key, which slows processing.

- **Using insert records from memory.** This load option uses a database vendor bulk or array insert API when available. This enables inserting a block of rows at a time. The default block size is 1000 rows, but a higher or lower value may provide better throughput, depending on the memory available on the server, the database, the number of columns loaded and their sizes. Try different values to see what works best.

  Note that incoming rows are not matched against the target table as with insert/update logic, so the number of rows inserted or rejected are not recorded. There are no flow statistics available.

  This method loads faster than the default load logic. For most databases, this is the fastest method, when the query can use automatic passthrough (so no intermediate flat file is created).

- **Using Bulk Load Utility using a disk file.** This load option uses a database vendor bulk load program. As the name implies, this method always creates an intermediate flat file. It is still faster than the default load logic. The options available depend on the database. For most databases, it is not quite as fast as Insert Records from Memory, when that option is available. However, for some databases designed for data warehousing (such as Sybase IQ and IBM Redbrick Warehouse), this is by far the fastest option. Additional tuning options are available depending on the database.

### Improving Lookup Performance

The function [DB_LOOKUP](#) provides an alternative to a SQL JOIN to look up a value in one (or more) columns of a table. When matching values are found, the value of another column is returned.

For example, a lookup table may contain a list of state codes and names. If you look up the state code NY, you get back the state name New York.

This process can be expedited. By loading the pairs of values into memory (called a cache) you can reduce the number of file or database read operations.
There are two types of cache, depending on the type of the lookup table.

- **Static.** For a flat file, all the lookup values are loaded into the cache.

- **Dynamic.** For a relational database table, the lookup values and results are added into the cache as they are used.

If each input value is only looked up once, there is no advantage to using a cache. If the same input values are looked up multiple times, throughput can be improved because the return value is read from the cache instead of from the database.

For database lookup tables, a system setting controls whether the values in the lookup table are loaded into the cache.

**Syntax:**

**How to Enable the Lookup Cache**

```sql
ENGINE INT CACHE {SET {ON [LIMIT \( nn \)] | OFF } | STAT}
```

where:

**ON**

Enables caching.

**LIMIT \( nn \)**

Limits the number of rows added from cache to \( nn \).

**OFF**

Disables caching.

**STAT**

Displays statistics for cache.

**Example:**

**Using the Lookup Cache**

To enable the lookup cache for a data flow, create a stored procedure with the following line:

```sql
ENGINE INT CACHE SET ON
```

From the Process Flow tab for the flow, drag the stored procedure into the workspace. Connect the Start object to the stored procedure, and connect that to the data flow.

To see how the cache was used, create another stored procedure with the following line:

```sql
ENGINE INT CACHE STAT
```
Drag it onto the process flow after the Data Flow. When you submit the flow, you will see lines, (like those shown in the following image), confirming that the cache was used.
The Data Management Console allows you to view, edit, and create synonyms. It includes a Synonym Editor, which provides a graphical interface for working with synonyms.

Editing synonyms can be particularly useful in DataMigrator for adding virtual columns that perform transformations.

The Synonym Editor also provides access to Data Profiling and Data Lineage information.

**In this chapter:**
- Synonym Editor Layout
- Viewing and Editing Synonym Attributes
- Using Variables in a Synonym
- Variables Attribute Summary
- Use a Variable in a Synonym
- Creating and Enhancing Synonyms Using Modeling View
- Adding Indexes to a Synonym
- Adding Virtual Columns (DEFINE) in a Synonym
- Viewing Data Profiling Characteristics
- Viewing Data Lineage for a Column

**Synonym Editor Layout**

The Synonym Editor occupies the DMC workspace and displays the Field View, Segment View, Modeling View, Text View, and Access File Text View tabs.

**Reference: Synonym Editor - Field View Tab**

The Field View tab shows a hierarchy of segments and columns on the left, with the attributes and values of the selected item on the right.

**Note:** The attributes available depend on the type of synonym.
The image below is an example of an SQL data source with a key column selected.

**Note:** Information about the attribute that has focus is displayed at the bottom of the attribute list. In this case, an explanation of the FIELDNAME attribute appears.

The following objects may appear on the Field View tab.

<table>
<thead>
<tr>
<th>Object</th>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synonym</td>
<td></td>
<td>General icon used to indicate synonyms, visible on top of the Field View tab. It provides information about the file name used and the application it resides in.</td>
</tr>
<tr>
<td>Segment</td>
<td></td>
<td>Indicates a segment in a synonym. The root or parent segment appears first in the tree.</td>
</tr>
<tr>
<td>Key Column</td>
<td></td>
<td>Indicates a key column.</td>
</tr>
<tr>
<td>Column</td>
<td></td>
<td>Indicates a general column.</td>
</tr>
<tr>
<td>Virtual Column (DEFINE)</td>
<td></td>
<td>Indicates a virtual or defined column.</td>
</tr>
<tr>
<td>Index Column</td>
<td></td>
<td>For FOCUS files, indicates the field index for quick retrieval of values for this field.</td>
</tr>
<tr>
<td>Object</td>
<td>Function</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Filter</td>
<td>Indicates a filter in the synonym.</td>
</tr>
<tr>
<td></td>
<td>Compute</td>
<td>Indicates a computed field.</td>
</tr>
<tr>
<td></td>
<td>Internal field</td>
<td>Indicates a column that does not appear in sample data or in the list of available columns.</td>
</tr>
<tr>
<td></td>
<td>Needs Value field</td>
<td>Indicates a column that requires a value to access the data.</td>
</tr>
</tbody>
</table>

**Reference:** Synonym Editor - Segment View Tab

The Segment View tab shows a hierarchy of segments on the left, with the attributes and values of the selected item on the right, as seen in the following image.
Reference:  Synonym Editor - List View Tab

The List View tab shows a list of objects on the left, with the attributes and values of the selected item on the right, as shown in the following image.

Reference:  Synonym Editor - Segment Shortcut Menu

When you right-click a segment in the Synonym Editor, the shortcut menu shown in the following image appears in both the Field View and Segment View tabs. This image shows a SQL data source with the root segment selected.
The following options are available:

**Properties**

Opens a pane showing the properties of the selected segment.

**Insert**

Enables you to insert one of the following:

**Reference to Existing Synonym as Child**

The reference is a pointer to the source synonym using the synonym as a child. If you subsequently make changes to the source synonym, reopening the current synonym will reflect those changes.

**Reference to Existing Synonym as Root**

The reference is a pointer to the source synonym using the synonym as a root. If you subsequently make changes to the source synonym, reopening the current synonym will reflect those changes.

**Segment via Metadata Import**

Enables you to create and add a new synonym to the current synonym using the Create Synonym tool.

**Segment Manually**

Adds a synonym that must be coded manually.

**Field**

Adds a general column to the segment.

**Define**

Adds a virtual or defined column to the segment.

**Filter**

Adds a filter to the segment.

**Compute**

Adds a calculated value to the file.

**Sort Object**

Adds a Sort Objects folder and a sort object. This option is available only if the Support extended options option is selected and is not used with DataMigrator.

**Variable**

Adds a Variables folder and a variable to the segment.
**Style**

Adds a Styles folder and a style object. This option is available only if the Support extended options option is selected and is not used with DataMigrator.

**Subquery**

Adds a subquery to the fields list.

**Index**

Adds an Index folder and an index.

**Data Profiling**

Provides the characteristics of the data for a segment. The following options are available:

**Statistics**

Shows statistic values for the selected segment only.

**Count**

Shows a count of the number of rows in the selected source table. For a multi-segment synonym, it shows the count for each segment.

**Statistics with parent segment**

This option is available only when viewing a lower segment of a hierarchy (cluster join).

**Key Analysis**

Allows you to select the columns that you would like to analyze as potential keys to the source table.

**Values**

Shows unique values and their percentages.

**Duplicate Values**

Shows identical values and their percentages.

**Sample Data**

Displays sample data in the workspace for the selected segment or synonym.

**Sample Data with parent key**

Displays sample data for the selected segment joined to the parent segment. This option is only available when a child segment is selected in a multi-segment synonym.

**Delete**

Deletes the segment.
**Rename**

Allows you to rename the segment.

**Edit All Properties**

Allows you to edit adapter specific properties, such as the delimiter value in a delimited flat file.

**Assign SCD Columns**

Allows you to assign columns to a SCD type by using drag-and-drop.

**Join Properties**

Opens the Join Editor so that you can specify how the selected segment is joined to the parent segment. This option is only available when a child segment is selected in a multi-segment synonym.

**Reference:** Synonym Editor - Column/Field Shortcut Menu

When you right-click a column in the Synonym Editor, a shortcut menu appears in both the Field View and List View tabs. The following image shows a SQL data source with a column selected.
**Note:** If you multi-select columns and right-click them, an abridged version of the shortcut menu opens.

The following options are available:

**Properties**
- Opens a pane showing the properties of the selected column.

**Insert**
- Enables you to insert one of the following:
  - **Field**
    - Adds a general column to the synonym.
  - **Date/Time Function**
    - Inserts a Date/Time function.
  - **Filter**
    - Adds a filter to the segment.
  - **Grouping**
    - Opens the Group values for dialog box, where you can configure group values.
  - **Aggregation Function**
    - Provides a choice of preconfigured aggregation functions that open in the Function Assist dialog box. The selected functions will be saved in the Constant Defines/Computes folder for the synonym.
  - **Analytic Function**
    - Provides a choice of preconfigured analytic functions that open in the Function Assist dialog box. The selected functions will be saved in the Constant Defines/Computes folder for the synonym.
  - **Sort Object**
    - Adds a Sort Objects folder and a sort object.
  - **Variable**
    - Adds a Variables folder, if needed, and a variable.
  - **Style**
    - Adds a Styles folder and a style object.
  - **Subquery**
    - Adds a subquery to the fields list.


**Pivot**

The following option is available:

**Multiple columns to rows**

Allows you to turn (or pivot) columns into rows or a single column into multiple columns, for retrieval.

**Impact Analysis**

Displays an Impact Analysis report for the particular column in the workspace. An Impact Analysis report identifies the procedures that access a Master File or field within a Master File.

**Data Lineage**

Provides the ability to see how the column was derived.
Data Profiling
Provides the characteristics of the data for a column. For more information, see Viewing Data Profiling Characteristics on page 543.

Statistics
Shows statistic values for the selected segment only.

Key Analysis
Allows you to select the columns that you would like to analyze as potential keys to the source table.

Hex View
Shows hexadecimal value of the selected field.

Values
Shows a values report for the selected column.

Values Graph
Displays a graph for both alpha and numeric field types.

Values Pie Graph
Displays a pie graph.

Duplicate Values
Shows identical values and their percentages.

Outliers
Shows outlier report for the selected column.

Sample Data
Displays sample data in the workspace.

Delete
Deletes the column.

Rename
Allows you to rename the column.

Decompose Date
**Reference:** Synonym Editor - Modeling View Tab

The Modeling View tab shows a graphical representation of the synonym, as shown in the following image. Use the Modeling View to define dimensions for OLAP analysis, view join properties, create cluster joins, and add or edit segments.
Reference: Synonym Editor - Text View Tab

The Text View tab shows the description of the underlying table.

You can print the synonym syntax by selecting Print from the Application menu.

The image below is an example of an MS SQL Server data source in the Text View tab.

Reference: Synonym Editor - Access File Text View Tab

The Access File Text View tab shows the description of the Access File for a synonym, which is used to access the database.

Note: You can print the synonym syntax by selecting Print from the Application menu.

The image below is an example of an SQL data source in the Access File Text View tab.

Note: While you can edit the text views, it is highly recommended that you do so with care. Editing the synonym incorrectly can invalidate the synonym.
Reference: Metadata Tab

The Metadata tab in the ribbon contains groups and commands that provide quick access to commonly performed functions while working in the Synonym Editor. The behavior of the command is determined by the selected object. Therefore, certain tab commands may be inactive.

The Metadata tab contains the following additional groups:

View Group

- **Field**
  Switches to the Field View in the workspace.

- **Segment**
  Switches to the Segment View in the workspace.

- **List**
  Switches to the List View in the workspace.

- **Modeling**
  Switches to the Modeling View in the workspace.

- **Text**
  Switches to the Text View in the workspace.

- **Access File Text**
  Switches to the Access File Text View in the workspace.

- **Aerial View**
  Opens the Aerial View window. This view is only available when in the Modeling View tab.

Insert Group

- **Insert**
  The Insert command allows you to insert one of the following:

  - Reference to Existing Synonym as Child
  - Reference to Existing Synonym as Root
  - Copy of Existing Synonym
  - Segment via Metadata Import
SELECT

Segment Manually
DBA
Field
Define
Filter
Compute
Group
Sort Object
Variable
Style

Reports Group

Sample Data
Enables you to view and refresh sample data for the object.

Data Profiling
Provides the data characteristics for synonym columns. The following options are available:

Statistics
Count
Key Analysis

Impact Analysis
Runs the impact analysis report for the selected column.

Tools Group

Properties
Displays the Properties panel for the selected object.

Quick Copy
This option is only available when the synonym name (root) is selected. For more information, see Creating a Quick Copy on page 381.
DBA

Adds field/value security DBA.

Business View

Enables you to create a Business View and a custom Master File that can use selected fields from the original synonym. In addition, you can customize field names, titles, and descriptions.

Viewing and Editing Synonym Attributes

The Synonym Editor enables you to view and edit the attributes of a synonym.

Procedure: How to View and Edit Synonym Attributes

To view and edit synonym attributes:

1. Double-click a synonym, or right-click it and click Open.

   The Synonym Editor opens to the Field View tab, which shows a hierarchy of segments and columns on the left, with the attributes and values of the selected item on the right.

   The attributes available depend on the type of synonym.

   The image below is an example of an SQL data source with a key column selected.

   The values for Usage Format, Expression, Description, and Nulls are viewable in the hierarchy with the columns. To edit these values, use the corresponding attribute fields in the Properties pane on the right-hand side of the Synonym Editor.

2. You can change the attribute values by typing in new values or by using the drop-down menus and check boxes.
The Synonym Editor does not let you make any changes that would render the synonym unusable. Therefore, you cannot edit any value field that is highlighted gray. In addition, if a change does not have proper syntax or format applied, the field may appear in red text. Messages and warnings appear if you try to save a file that contains an error.

3. Save changes by clicking Save from the Quick Access Toolbar.

   **Note:** You can multi-select fields to change attributes for multiple fields in a single operation. The display changes to show only those attributes that can be changed.

4. Close the Synonym Editor by clicking the X.

   **Note:** If you close the Synonym Editor without saving your changes, you are prompted to do so.

**Reference:** File Attributes Summary

The image below is an example of an SQL data source with the synonym file name selected.

```
<table>
<thead>
<tr>
<th>Display Name (Title)</th>
<th>Usage Format</th>
<th>Description</th>
<th>Nulls</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sample - Orders</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DMDORD</td>
<td></td>
<td>DMORD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Order... A5</td>
<td></td>
<td>Order Number</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Date... HYYMDI</td>
<td></td>
<td>Date of Order</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Shipp... HYYMDI</td>
<td></td>
<td>Date of Shipment</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Comp... A6</td>
<td></td>
<td>Company ID</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Plant,L... A3</td>
<td></td>
<td>Location Of M...</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Order... YY</td>
<td></td>
<td>Order Date Year</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Order... Q</td>
<td></td>
<td>Order Date Quarter</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Order... M</td>
<td></td>
<td>Order Date Month</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>DMPROD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

**Note:**

- Information about the attribute that has focus is displayed at the bottom of the attribute list.
- The attributes available depend on the type of synonym.

Synonyms can have the following file attributes:

**General**

**SUFFIX**

Identifies the type of synonym or data source.

**MFD_PROFILE**

The name of the foxexec that will be executed before the request containing MFD.
FDEFCENT
Defines the default century value, specifying a century number for handling cross-century dates.

Note: Use the default setting (0) unless you wish to retrieve data from a earlier century (for example, 19xx).

FYRTHRESH
Defines the base years, to represent the lowest year to which the century value applies (FDEFCENT).

Note: Use the default setting (0) unless you wish to retrieve data from a earlier century (for example, 19xx).

REMARKS
Enables you to include descriptive information at the file level and specify multiple language descriptions for the synonym. Remarks are displayed along with the file name during reporting.

DATASET
Identifies the location of the data source to be used in the file name, including the extension and the location of the data file. This option is only available for file data sources, such as flat files or XML and JSON documents.

DV_INCLUDE
Controls the base file DB info inheritance in a cluster.

BV_NAMESPACE
Controls the Business View run-time resolution strategy.

Reference: Segment Attributes Summary
If a synonym segment is selected, the attributes listed below are available.

Note:

■ Information about the attribute that has focus is displayed at the bottom of the attribute list.

■ The attributes available depend on the type of synonym.

General

SEGMENT
Is the name of the segment.
SEGTYPES

Specifies the type of relationship that a segment has to its parent, and indicates which of the segment fields are key fields and in what order they are sorted.

**Type.** Identify the segment type and sorting options from the Type drop-down list.

**KEY.** Enter the names of key fields that you want to use for sorting. For example, no two employees can have the same employee ID number, so you can use that field as the key. A segment instance can have more than one field that makes up the key. That is, two or more field values may be used to distinguish records.

SEGSUF

SEGSUF is used when part of the data source being described by the synonym is of a different data source type than that declared for the entire structure. This option is not available for the first segment in a synonym.

**Note:** SEGSUF is the data source type of a segment and any descendants it might have, where that type differs from the SUFFIX value.

**Miscellaneous**

**DESCRIPTION**

Contains a description or comments about the segment.

**CRFILENAME**

Is the name of the cross-referenced data source.

**CRSEGNAMEN**

Is the name of the cross-referenced segment.

**CRKEY**

Identifies the common join field for the cross-referenced segment.

**Note:** These cross-referenced values (CRFilename, CRSegname, CRKey) are available for FOCUS data sources.

**SEG_TITLE_PREFIX**

An optional attribute that enables you to provide the title prefix for the segment fields that appear in reports. You can split the text across up to five separate title lines by separating the lines with a comma.
Adapter Specific

Adapter specific fields are shown if an Access File component has been generated with the synonym.

CARDINALITY

Defines how many members of a dimension can be retrieved for a report.

TABLENAME

Identifies the table or view. It may contain the owner ID, as well as the table name. For some synonyms, it must also contain the data source name. This value may be a variable.

CONNECTION

Indicates the host server or data source for synonyms. This value may be a variable.

KEY

Identifies the names of the columns that constitute the primary key.

KEYORDER

Identifies the logical sort sequence of data by the primary key.

WRITE

Specifies whether write operations are allowed against the table.

DBSPACE

Identifies the storage area in which the table resides.

PERSISTENCE

Specifies the type of table persistence and related table properties. This is optional for database management systems that support volatile tables, and required otherwise.

Reference:  Column/Field Attribute Summary

If a column or field in a synonym is selected, the attributes listed below are available.

Note:

- Information about the attribute that has focus is displayed at the bottom of the attribute list.
The attributes available depend on the type of synonym.

General

**FIELDNAME**

Is the name of the column.

**TITLE**

Supplies a title to replace the column name that is normally used in reports and enables you to specify multiple language titles for the column or field.

**Usage Format**

Describes the data type and format for the column for usage or display.

**Note:** Additional attributes, DEFCENT and YRTHRESH, are available if the Usage field is set to Date, Time, or DateTime (Timestamp) format. Use these attributes to enter the century and year threshold values for the column or field.

**ALIAS**

Assigns an alternative name for a column, or the real column name for a DBMS synonym.

For file data sources, a special reserved ALIAS of INSTANCE means the associated FIELD will show the name of the file when it is read.

When a FIELDNAME of RECTYPE is used, the ALIAS contains the value that identifies the record type.

**Key Component**

Indicates whether the field is part of the primary key for the table.

**ACTUAL**

Describes the type and length of data as it is actually stored in the data source.

**MISSING**

Controls how null data is handled, that is, if no transaction value is supplied. This option is not available for KEY fields.

Miscellaneous

**DESCRIPTION**

Contains a description or comments about the column or field.
ACCEPT
Specifies criteria for validating data.

- **OR** enables you to specify an acceptable value.
- **SYNONYM** enables you to specify a synonym as the source of the accept list.
- **FOCEXEC** enables you to specify a FOCEXEC as the source of the accept list.
- **FROM-TO** enables you to specify a range of acceptable value fields.
- **FIND** enables you to supply file and field names to instruct where to search for a data source and for a list of acceptable values. You supply the field name of the data field for which the validation criteria are being assigned, the file name of the target FOCUS data source where the field can be found, and the field name of the target data field that contains the validation criteria.

FIND is only available for FOCUS data sources and does not apply to OLAP-enabled synonyms. Also note that, in the Maintain environment, FIND is not supported when developing a synonym.

WITHIN
Contains the name of a field to be included in a dimension.

PROPERTY
Describes which hierarchy property the column represents.

REFERENCE
Points to the field that is the unique identifier for the hierarchy.

FIELDTYPE
- **I - Index** identifies an indexed column in a FOCUS database.
- **R - Read only** indicates a read-only column. This setting is useful for columns that are automatically assigned a value by the RDBMS.

ACCESS_PROPERTY
Specifies access options for the column data.

- **INTERNAL** defines a column that does not appear in sample data or in the list of available columns. This restricts the field from showing in any of the field lists in the reporting tools.
- **NEED_VALUE** defines a column that requires a value to access the data.
Select By defines a column by value, range, or multi-values.

HELPMESSAGE

Appends a help message to a column.

This attribute is not used by WebFOCUS.

SCD Type

Sets slowly changing dimension attributes. This option is only available for existing relational targets and is only used by a flow when SCD is enabled in the target object.

A surrogate key is the KEY column in the table and has an SCD type of blank. Other columns with a blank SCD type have no SCD processing done to them. In a synonym, this column will always appear first, even if it is not the first column in the table.

Note: The surrogate key must be an integer. If the column is identified in the synonym as read-only and uses AUTOINCREMENT, then the database native processing is used to assign surrogate key values. If it is not so identified, then DataMigrator processing assigns values for this column, incrementing for each new row added.

Logical Key Field. Is the source database key. Multiple columns may comprise the key.

Begin Date/End Date for Type II changes. Indicates date range for the row values. A null end date indicates the row is current. If you use Begin and End Date, the End Date column must be created as nullable, unless you override the default End Date value of NULL with a value for &CM__SCDAEDATE.

The data type must be either a Date format or a Date and Time format, however, the formats of the two fields must be identical.

Begin/End Date values for Slowly Changing Dimensions can also be Datetime or Timestamp values.

Change Flag. For Type I changes, a value is assigned when a row is updated. If a value is not assigned in Target Transformations for the Date and Datetime columns, the current date or timestamp is assigned automatically.

Activation Flag. Indicates if the row is current.
This field must be an integer. Valid values are:

1

Indicates that a record is active.

0

Indicates that a record is inactive.

- **Type I** (overwriting history) designates columns whose database values are overwritten with new values.

- **Type II** (preserving history) designates columns whose database rows are flagged as inactive or assigned an end date. New rows are inserted with the new values.

- **blank** (non-key permanent columns) indicates that database values are not changed.

**TEMPORAL_PROPERTY**

Inserts virtual fields in the Master File to show how selected totals progressed through a time period for specified intervals.

**USESTYLE**

Are the styles (for example, font and color) to apply to the field.

The attributes available depend on the type of synonym.

**Adapter Specific**

**AUTOINCREMENT**

Indicates that the field is automatically incremented by the relational database when rows are inserted. This option is only available for selected relational databases. When selected, the field must also be identified with a field type of read-only.

The following attributes are only used if AUTOINCREMENT is checked.

- **START**. Indicates the starting value for this AUTOINCREMENT field. If not specified, 1 is used.

- **INCREMENT**. Indicates the increment value for this field. If not specified, 1 is used.

- **SEQUENCE**. Is the name of the corresponding database SEQUENCE object for this field. This option is only available for ORACLE tables.
Procedure: How to Assign SCD Columns on a Synonym

You can assign SCD (Slowly Changing Dimension) Types column attributes from a single dialog box in the Synonym Editor. This operation is performed on the segment level from the Field View, Segment View, or Modeling View.

1. In the Synonym Editor, right-click the segment name and click Assign SCD Columns from the shortcut menu, as shown in the following image.

![Assign SCD Columns Dialog Box](image)

The Assign SCD columns dialog box opens.
2. To assign the SCD Type to a column, click the column in the Available Field(s) section and the SCD type in the Selected SCD Type(s) section, followed by the right arrow. You can also drag the column from the left to the desired SCD Type on the right. An example of assigning SCD types to a column is shown in the following image.

![Assign SCD columns for DSDIMSAL](image)

3. Click Validate to confirm that all SCD Types are assigned to the synonym columns correctly.

   If the validation passes, the synonym can be used as a target synonym in a Data Flow with the Load Type Property set to *Slowly Changing Dimensions*. 
If the validation does not pass, the segment name appears in red. To view the explanation of the error, right-click the segment name and click *Error Details*, as shown in the following image.

**Reference:** File Listener Attributes

The image below is an example of a flat file data source with DATA_ORIGIN of LISTENER.
A flat file data source may have the following attributes:

**Data Allocation**

**CONNECTION**

The name of a connection for a flat file source that identifies an FTP or SFTP server. This field is optional. If not specified, the file resides locally.

**DATA_ORIGIN**

Indicates the origin of data described by the synonym.

- **FILE.** Indicates that a file (or collection of files if a wildcard is used in the dataset parameter) is read once. This is the default.
- **LISTENER.** Indicates that files are read as they appear, as specified in the remainder of the File Collection Parameters.
- **TAIL.** Is reserved for future use.

**DIRECTORY**

Defines the location of data files for polling. This can be an application directory, a mapped application name, or a directory location on your file system.

When using a directory on a file system for local files, the name of the physical directory can be used.

When using a directory on a file system for remote data files, a directory relative to the initial directory can be used. For example, if your remote initial directory is `/home/user1/apps`, using `sales` will set the data files directory location to `/home/user1/apps/sales`.

**NAME**

Indicates the name of the file. May include wildcard characters. The default is `*`.

**EXTENSION**

Indicates the file type that the listener is checking. The default is `.ftm`.

The parameters below only apply to Origin of Listener.

**Pre Processing Actions**

**PICKUP**

Indicates the mechanism used to pick up files from their directories. The options are:

- **IMMEDIATE.** Processes the file as soon as it is detected. This is the default.
**TRIGGER.** Processes the file when a trigger file is detected. This setting requires an additional setting for the trigger file extension.

A trigger is typically used when processing needs to be delayed. For example, when using FTP to copy files, if you wait until this secondary file appears while a large file is being copied into a directory, the trigger file indicates when processing can begin.

When a pickup strategy of trigger is used, indicates the trigger file extension. The trigger file extension is added to the full name of the file being listened for. For example, if the trigger file extension is .trg, the trigger file for a data file named input.dat would be input.dat.trg. There is no default value.

**Note:** For VMS only, use an underscore. For example, input_dat.trg.

**MONITOR.** Monitors the file as soon as a file change is detected.

The files themselves are not read, only their name and location are recorded. This processes files and sub-directories when they are added, changed, or deleted.
**Post Processing Actions**

**DISCARD**

Indicates what to do with the file after it has been processed by a file agent. The options are:

- **DELETE.** Deletes the file.
- **ARCHIVE.** Moves the file, while adding a timestamp to the name, to a specified directory. This is the default.
- **ARCHIVE_LOCAL.** Defines the location where the data files will be archived. This location can be an application directory, a mapped application name, or a directory on your local file system.

This option must be used only for a CONNECTION that is set with FTP or SFTP protocol.

This option allows archiving the files locally, where DataMigrator Server is running, while the listening directory is on a remote location.

- **KEEP.** Keeps the file. This option is only available for the Trigger pickup type.

**ARCHIVE**

Defines the location where the data files will be archived when DISCARD strategy of ARCHIVE is used. This can be an application directory, a mapped application name, or a directory location on your file system.

When using a directory on a file system for local files, the name of the physical directory can be used.

When using a directory on a file system for remote data files, a directory relative to the initial directory can be used. For example, if your remote initial directory is `/home/user1/apps`, using `sales` will set the data files directory location to `/home/user1/apps/sales.`
Listening Parameters

POLLING
Indicates the polling interval (in seconds) for new files arriving in the specified directory. For example, a polling interval of 10 means the directory is scanned every ten seconds for new files. The default is 10 seconds.

TIMEOUT
Indicates the timeout interval (in seconds) a flow or other read request should wait before terminating. For example, a timeout value of 100 means if no new files arrive in the directory after 100 seconds, the job should stop. The default is 10 seconds.

MAXERRORS
The maximum number of errors allowed.

MAXQSIZE
The maximum number of files in the processing queue.

QSORT
The sort order for files in the processing queue.

Read Limits

MAXFILES
Indicates the maximum number of data files processed by the request. The default is 99999999.

MAXRECS
Used only for the data origin of TAIL. Indicates the maximum number of records processed by the request. The default is 99999999.

Adapter Specific

MISSVAL_CHAR
Indicates a missing alpha value in the data source.

MISSVAL_NUMB
Indicates a missing numeric value in the data source.

MISSVAL_DATE
Indicates a missing date/time/timestamp value in the data source.
Reference: File Listener for Use With Excel Sheets

The File Listener can be used to process flat files as they appear in a specified directory. It's also possible to use it with Excel files with some additional preparation. These instructions describe the additional steps required to use an Excel worksheet as a source.

Use of the File Listener is documented in the DataMigrator User's Guide. For a complete example of using the File Listener see the section "Loading a Target Table Using the File Listener." For further explanation of the attributes described see the reference for "File Listener Attributes."

Procedure: How to Create a Synonym for Use With the File Listener and Excel Files

1. Right-click an application directory, point to New and click Synonym via Synonym Editor.
   A new synonym opens in the workspace.
   If the Properties panel is not visible, right-click the synonym name (for example baseapp/synonym01) and click Properties.
2. From the SUFFIX drop-down menu, select Flat File.
3. From the IOTYPE drop-down menu, select BINARY.
4. Right-click the synonym name, point to Insert, and click Segment Manually.
   A new segment named SEG1 and a new field named FIELD1 are added. SEG1 is selected.
5. In the Properties panel, from the Type drop-down menu, select Multiple.
6. From the DATA_ORIGIN drop-down menu, select LISTENER. Enter the following values:
   1. From the CONNECTION drop-down menu, select <local> or an [S]FTP server.
   2. In the DIRECTORY and NAME boxes, enter the location and name of the file(s) you want to listen for.
   3. In the EXTENSION box, enter xlsx.
   4. Enter values for PICKUP, DISCARD, POLLING, and TIMEOUT as required.
7. In the workspace, click FIELD1. Enter the following values in the Properties pane:
   1. In the FIELDNAME box, enter FILENAME_FULL.
   2. In the Usage Format box, enter A512.
   3. Expand the DBMS/Source Data Attributes section.
   4. In the ALIAS box, enter INSTANCE_FULL.
   4. In the ACTUAL box, enter A512.
**Procedure:** How to Add a Reference to an Existing Synonym for an Excel Worksheet

If you have a synonym for the Excel worksheet already in place, follow this procedure.

1. Right-click the **SEG1** segment, point to **Insert**, and click **Reference to Existing Synonym as Child**.

   The Insert Reference to Existing Synonym as Child dialog box opens.

2. Click the synonym that describes the Excel worksheets that you want to use as a source and click **Select**.

   The reference to the synonym is added and the segment name is selected.

3. In the Properties panel, from the POSITION drop-down menu, select **INSTANCE_FULL**.

4. From the Quick Access toolbar click the save icon. Enter a name for the synonym and click **Save**.

Now you can use the synonym you created as a source in a Data Flow or a Direct Load Flow.

**Reference:** Prototype File Monitoring Synonym

A prototype synonym is provided as filemntr.mas for a user with the File Listener pickup strategy Monitor. To create this synonym, first follow the instructions outlined in *How to Create Sample Procedures and Data for File Listener* on page 82. Then, copy it to an application directory in your application path and edit the File Listener parameters as needed. At a minimum, edit the value for DIRECTORY to point to the application directory (or mapped application name) that you want to monitor. Edit the NAME and EXTENSION values to show the file name (or * for all names) and extensions that you want to monitor.

**Note:** These virtual fields can also be used with pickup strategy other than Monitor, except as noted.

This synonym contains fields with special values for the ALIAS, as described below.

- **INSTANCE_FULL** - The full path and name of the file. Cannot be used in conjunction with INSTANCE. Also available with IMMEDIATE.

- **INSTANCE_SIZE** - The size of the file in bytes. Also available with IMMEDIATE.

- **INSTANCE_DATE** - The date that the file was last modified. Also available with IMMEDIATE.

- **INSTANCE_TIME** - The time that the file was last modified. Also available with IMMEDIATE.

- **INSTANCE_TYPE** - The value File to indicate a file, or when DATA_ORIGIN is FILE, the value `<DIR>` to indicate a directory. This directory is only available with MONITOR. Otherwise, the value is always File.
- INSTANCE_ACCS - The access mode for the file. Available on UNIX systems only.
- INSTANCE_RDWR - Indicates if the file is READONLY. Available on UNIX systems only.
- INSTANCE_USER - The user that owns the file. Available on UNIX systems only.
- INSTANCE_GROUP - Group identification. Available on UNIX systems only.
- INSTANCE_MONS - Monitoring status. Blank for the initial run, then ADDED, DELETED, MODIFIED, or RENAMED to reflect changes. Only available with MONITOR.

The properties of the prototype synonym are shown in the following image.

---

Using Variables in a Synonym

The Synonym Editor allows you to define variables in a synonym in order to specify synonym attributes, such as the connection name or the table name. For an example of how variables in a synonym can be used, see *Iterating Through Multiple Tables* on page 631.

**Procedure:** How to Add a Variable to a Synonym

1. Right-click an object in the Synonym Editor, point to Insert, and then click Variable.

---

iWay DataMigrator User’s Guide 533
2. A new folder named Variables is added to the object tree, with &VARIABLE1 as an entry. The Properties pane shows the attributes of the variable, as shown in the following image.

3. If desired, change the name of the variable and enter a default value in the appropriate fields.

Procedure: How to Add a Variable for an Attribute

Attributes in a synonym that can be specified with a variable are identified with an attribute name in blue, as shown in the following image.
To add variables for an attribute:

1. Right-click the attribute in the Properties pane and click Parameterize Value.
2. Expand the Variables folder. A variable with a name that combines the synonym name with the variable name has been added.
3. Click the variable name to assign values, as shown in the following image.

Variables Attribute Summary

Variables in a synonym can have the following attributes:

**Variable Name**

The name of the variable. A double ampersand (&&) is automatically inserted in the name to indicate a global variable.

**Prompt Caption**

A description of the variable that replaces the variable name in the prompt.

**Default Value**

The default value for the variable. If no value is set, the default value is used.

**Prompt Values**

Provides a drop-down menu of optional prompt values for the variable. The options are Input box, Static list of values, Static list of values/captions, Synonym based dynamic list of values, and Procedure based dynamic list of values.
Quoted
When selected, single quotation marks are added around the assigned string for the variable. A single quotation mark within the string is converted into two single quotation marks.

Usage Format
The data type. The following options are available:

- Integer
- Character (fixed)
- Decimal

Use a Variable in a Synonym
Variables in a synonym are commonly used to represent a table name or connection name.

1. In a synonym for a relational database where variables have been established, click a segment to open the properties page that shows the Adapter Specific attributes.
2. In the fields for TABLENAME and CONNECTION, enter the names of the variables to be used.
3. Click Save to save your changes.

Creating and Enhancing Synonyms Using Modeling View
The Synonym Editor Modeling View tab allows you to create and enhance synonyms graphically. You can also use this view to create a single view of multiple data sources.

Procedure: How to Create Synonyms Using Modeling View
1. Right-click an application directory, point to New, and then click Synonym via Synonym Editor.
2. Right-click the new synonym, point to Insert, and click Reference to Existing Synonym as Root.
3. Select the synonym to use as the root, and click Select.
   Note: If the synonym selected has more than one segment, a dialog box opens. Click the entry segment you wish to use and click OK.
4. Click the Modeling View tab.
5. Right-click the new synonym in the workspace and point to Insert.
6. Add segments through one of the four methods listed.
7. Optionally, if you have multiple joined segments you can double-click the arrowed lines between tables to open the Join Editor dialog box and determine how the tables are linked.

**Procedure:** How to Enhance Synonyms Using the Modeling View

Use the Modeling View to enhance a synonym by adding a segment.

1. Open the synonym in the Synonym Editor by double-clicking it from the navigation pane.
2. Click the *Modeling View* tab.
3. Right-click a segment in the workspace and point to *Insert*.
4. Add segments through one of the methods listed:
   - **Reference to Existing Synonym as Child.** Adds a segment to the existing synonym as a reference to another synonym. Only the name of the synonym is added. If the referenced synonym is changed, the new version is used.
   - **Reference to Existing Synonym as Root.** Inserts reference to selected table and any tables it references.
   - **Segment via Metadata Import.** Enables you to add segments by using the Create Synonym tool. This tool creates a synonym and includes it as a segment in the synonym from which the tool was launched.
   - **Segment Manually.** Enables you to add segments manually to the current synonym.
     - **Note:** You would only use this approach if you are coding a new synonym by hand, as you would for a FOCUS data source.
5. Optionally, you can double-click the arrowed lines between tables to open the Join Editor dialog box and determine how the tables are linked.

**Tip:** You may also right-click a file icon and click *Join Properties* from the context menu.

In addition, when creating a cluster join, there are two additional options for linking tables. Join All (default) is shown by a double arrow, and Join Unique is shown as a single arrow.

*Note:* Join options are not available for FOCUS files.

**Reference:** Modeling View - Segment Shortcut Menu

When you right-click the title bar of a segment in the Synonym Editor Modeling View tab, a shortcut menu opens.
The shortcut menu has the following options:

**Expand/Collapse**
Changes the view from a file icon to a table view, which enables you to see columns, sample data, and sample data parent keys.

**Tip:** You may also double-click a file icon to open the table view. Double-click the table to close the table view, or click the X button from the toggle toolbar to close.

**Properties**
Opens the Properties pane for the selected segment.

**Insert**
Has the following options:

- **Reference to Existing Synonym as Child.** Adds a segment to the existing synonym as a reference to another synonym. Only the name of the synonym is added. If the referenced synonym is changed, the new version is used.

- **Reference to Existing Synonym as Root.** Inserts reference to selected table and any tables it references.

- **Segment via Metadata Import.** Enables you to create a new synonym through the Create Synonym tool, and add it to the current synonym.

- **Segment Manually.** Adds a segment to the current synonym that must be coded manually.

**Data Profiling**
Provides the data characteristics for the columns of a synonym.

**Sample Data**
Shows a sample of the data associated with the synonym.

**Delete**
Deletes the segment.

**Edit All Properties**
Opens the Properties pane for the segment.

**Assign SCD Columns**
Allows you to assign columns to a SCD type by using drag-and-drop.
You can create indexes in a synonym by adding them in the Synonym Editor. When you create a table by right-clicking a synonym, pointing to Data Management, and clicking Recreate DBMS table, the indexes are created in the relational database.

When you point to New and click Synonym to create a synonym for a relational database table with more indexes than just the primary key, they will be included in the synonym.

Additional indexes are useful when you anticipate using a column in a JOIN operation, a WHERE condition, or as an argument to DB_LOOKUP. You can add multiple indexes. Each index can be on one or more columns. You can specify if the index is sorted, and whether it is a unique index.

Example: Adding an Index to a Synonym

This example uses the sample synonym ibisamp/dmreps that has a primary key on column REPNO and a unique index on that column. It shows how to add an index on the column EMPID, and another index on the FNAME and LNAME columns that does not have to be unique.

1. From the ibisamp application directory, double-click the dmreps synonym to open it in the Synonym Editor.
2. Right-click the segment name DMREPS, point to Insert, and then click Index.
   The Index Fields Editor dialog box opens. The field REPNO is already selected because it is the key column of the table.
3. Click OK.
4. A new folder named Indexes appears in the tree with Index1 defined for REPNO. Right-click the Indexes folder, point to Insert, and then click Index.
   The Index Fields Editor dialog box opens.
5. Double-click EMPID in the Available Fields pane to add it to the Selected Field(s) pane.
6. Click OK.
7. Right-click the Indexes folder, point to Insert, and then click Index.
   The Index Fields Editor dialog box opens.
8. Double-click FNAME and LNAME to add them to the Selected Field(s) pane.
9. Click OK.
10. In the Properties pane for Index3, select N – Index is not unique in the INDEX_UNIQUE drop-down menu.
Your synonym should look like the following image.

Adding Virtual Columns (DEFINE) in a Synonym

You can create transformations in a synonym by adding a virtual column, also known as a DEFINE. This creates a transformation that appears with the data source list of columns.

A virtual column in the synonym is useful in the following situations:

- **When you want to create static transformations that are constant through an organization.** For example, you might want to create a column of phone numbers with the dashes removed or a list of Department codes decoded (such as, 1 for Accounting, 2 for Marketing, and so on).

- **When you want all data flows to use a set of common business rules.** For example, you might want to set a definition like PROFIT = REVENUE - EXPENSES.

- **When you want to simplify data flow maintenance.** If the PROFIT expression noted above ever changes, it only needs to be edited at the source Server.

A virtual column in the synonym can also be used in the following situations, as an alternative to using a source transformation:

- **When you want to filter information based on the results of a transformation.** For example, you might want to filter based on the size of PROFIT.

- **When you want to group by the results of a calculation.** For example, the table has an order date column, and you want to group on the year extracted from that column.
When you want to create transformations based on user routines that are located on the source machine. Specialized in-house routines, tested over time, may be required for your warehousing application, or the Reporting Server and your data mart may be located on a UNIX platform. Business rules may dictate that a z/OS COBOL program be executed for each retrieved row. You can exploit this COBOL program, using a virtual column.

It is often useful to prototype the transformations with a limited number of rows, and then place them in the source definition after analyzing your performance.

**Note:** If you have a data flow that uses the synonym in which you are creating a virtual column, you should right-click the synonym’s application directory and select *Refresh* to use the new column.

A virtual column can contain an expression, a constant, or a column name.

- If the virtual column is a complex expression, you can create the expression with the Transformation Calculator or just type it into the Expression field.
- If the virtual column is a simple expression, such as a constant value, you can type the value in the Expression field.

After creating a virtual column, you can test it by right-clicking the synonym and clicking *Sample Data*. Sample data appears in a separate dialog box.

**Procedure:** How to Create a Virtual Column in a Synonym

To create a virtual column in a synonym:

1. From an application directory in the navigation pane, double-click a synonym, or right-click it and click *Open*.
   
The Synonym Editor opens.

2. Right-click a segment (or column), point to *Insert*, and then click *Define*.  

iWay DataMigrator User’s Guide 541
The Define Calculator opens, as shown in the following image.

3. From the Define Calculator, type a name for the column in the Name input field, or use the default define name.

4. You may enter a descriptive title for the virtual column (DEFINE) in the Title input field. **Tip:** From the Synonym Editor, click the browse (...) button at the right of the TITLE and DESCRIPTION value fields to specify multiple language titles.

5. Use the Expression tab and the calculator buttons to build the expression for the virtual column (DEFINE).

or

Use the Relational Expression tab to build the expression.

a. Click the Add New Row icon. From the Relational Expression tab, use the drop-down lists to select the filter Field (column), Relation, and Type. **Note:** Parameters are not supported.

b. Click the browse (...) button at the right of the Value input field.

The Value Selection dialog box opens.

c. Select from the available values and use the arrows to add or remove values.

d. Click OK to close the Value Selection dialog box and return to the Relational Expression tab.
The expression is added to the value field.

e. To add another filter, double-click a column or variable from the Columns/Variables tab on the right side of the Virtual Column Calculator. The filter is added to the Relational Expression tab, where you can add the expression value.

f. To delete an expression, click in the row number column of the expression that you are deleting, right-click, and select Delete selected row(s). The expression is removed from the Relational Expression tab.

6. You may select the Check Expression and Sample Data buttons, located on the top right of the Virtual Column Calculator, to verify that the expression is valid and to view sample data for the filter.

7. Click OK to close the Define Calculator and return to the Synonym Editor.

   **Note:** To edit the Define, Title, or Expression, you may do so directly from the Synonym Editor, or you may click the browse (...) button at the right of the EXPRESSION value field to relaunch the Define Calculator.

8. If no columns from the synonym are used in the expression or have been defined, you can use the WITH option to identify the logical home of the defined calculation. You can also use the WITH option to move the logical home for the virtual column to a lower segment than it would otherwise be assigned (for example, to count instances in a lower segment).

   **Tip:** You can click and drag the DEFINE field and move it to a different segment in the Modeling View tab, which also changes the segment association.

9. Specify the Missing Data options for columns that allow null data. You can allow all missing data.

10. Click Save from the Quick Access Toolbar to save the synonym.

11. To close the Synonym Editor, click the control button in the upper-right corner.

**Viewing Data Profiling Characteristics**

Data Profiling provides data characteristics for the columns in a synonym. You can display the characteristics for all the columns in a synonym or segment, or for an individual column.

For alphanumeric columns, Data Profiling provides the segment, format, count of distinct values, total count, patterns count, maximum, minimum, and average length, minimum and maximum values, and number of nulls. Patterns count shows the number of patterns found in each alphanumeric column.

For numeric columns, Data Profiling provides the segment, format, count of distinct values, total count, maximum, minimum, and average values, and number of nulls.
Data Profiling for an individual column provides access to Statistics, Patterns, Values, and Outliers reports.

**Data Profiling a Synonym or Segment**

Data Profiling provides information on all the columns in a synonym or segment. You can also drill down to the Values or Patterns reports for an individual column from a synonym or segment Data Profiling report.

**Note:** Data Profiling is also available from the navigation pane by right-clicking a synonym, pointing to *Data Profiling*, and then clicking *Statistics*.

**Procedure:** How to View Data Profiling for a Synonym or Segment

To view the Data Profiling information for a synonym or segment:

1. Open the synonym in the Synonym Editor by double-clicking it from the navigation pane or right-clicking it and clicking *Open*.
   
   The Synonym Editor opens to the Field View tab.

2. Right-click the synonym or segment name, point to *Data Profiling*, and then click *Statistics*.
   
   The Data Profiling information displays in the workspace. The last four columns are shown below the rest of the information for illustrative purposes only. The actual report runs across the workspace.

   You may use the Report tab in the ribbon to view server messages, print the report, copy data as text, and export the report.

3. Optionally, you can click a column name or patterns count (for alphanumeric columns) to drill down to the Values or Patterns reports, respectively.

   For pattern analysis, a 9 represents a digit, an A represents any uppercase letter, and an a represents any lowercase letter. All printable special characters are represented by themselves, and unprintable characters are represented by an X.

   **Note:** Data Profiling is also available from the navigation pane by right-clicking a synonym.
Procedure: How to View Key Analysis for a Synonym or Segment

Key Analysis provides a report that shows which columns in a data source can be used individually, or in combination, to uniquely identify a row. The columns identified in this report are candidates for key columns.

Note: Key Analysis is also available from the navigation pane by right-clicking a synonym, pointing to Data Profiling, and then clicking Key Analysis.

To view key analysis for a synonym or segment:

1. Open the synonym in the Synonym Editor.
2. Right-click the synonym or segment name, point to Data Profiling, and then click Key Analysis.

   The Key Analysis dialog box opens.
3. Select the columns that you would like to analyze as potential keys to the source table and click Analyze.

   The Key Analysis report for that synonym box opens.
4. If you selected a segment name, the report opens as shown in the following image. Note that all columns in the segment will be selected.

The report shows:

**Segment**

   The selected segment.

**Name**

   Name of the segment.

**Format**

   The format of each column.
Elements
The number of elements (columns) shown.

Count
The number of rows.

Distinct Count
The number of distinct rows.

Distinct Percent
The percentage of rows that are distinct. This value must be 100% for a combination of columns to be used as key.

Duplicate Count
The number of duplicate values.

Duplicate Percent
The percentage of duplicate values. This value must be 0% for a combination of columns to be used as key.

Example: Select Table DMHR and All of the Columns EXCEPT ID_NUM and Capture the Output
By default, the report is sorted by the number of elements, so the first rows in the report show one element each. This enables you to determine if any single column could be used by itself as a key. The report then shows all combinations of two columns, three columns, and so on.

To see the values in the report, right-click on any row.

The duplicate rows option shows all duplicate values, which prevent the desired column combination from being used as a key.

Data Profiling a Single Column
Data Profiling for an individual column provides access to nine reports:

☐ Statistics. Lists counts, percents, limits, distinct values, and information about null values in a field or a segment, as shown in the following image.

For alphanumeric columns, the Statistics report provides the segment, format, count of distinct values, total count, patterns count, maximum, minimum, and average length, minimum and maximum values, and number of nulls.
For numeric columns, the Statistics report provides the segment, format, count of distinct values, total count, maximum, minimum, and average values, and number of nulls.

- **Key Analysis.** Lists counts and distinct and duplicate counts and percents for a single column, as shown in the following image.

<table>
<thead>
<tr>
<th>Segment</th>
<th>Name</th>
<th>Format</th>
<th>Elements</th>
<th>Count</th>
<th>Distinct Count</th>
<th>Distinct Percent</th>
<th>Duplicate Count</th>
<th>Duplicate Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>INVINFO</td>
<td>PROCPNAME</td>
<td>A30V</td>
<td>1</td>
<td>17</td>
<td>17</td>
<td>100.00</td>
<td>0</td>
<td>0.00</td>
</tr>
</tbody>
</table>

- **Patterns** Shows patterns of letters, digits, and special characters, as well as counts and their percents, as shown in the following image. This report is only available for alphanumeric columns.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Count</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>999 9AA Aaaaaaaa 99 A</td>
<td>2</td>
<td>11.76</td>
</tr>
<tr>
<td>999 AAA- A Aaaaaaaa 99 A</td>
<td>2</td>
<td>11.76</td>
</tr>
<tr>
<td>999AA Aaaaaaa Aaaaaaaa 99 A</td>
<td>2</td>
<td>11.76</td>
</tr>
<tr>
<td>9 Aa AAA AAA Aaa</td>
<td>1</td>
<td>5.88</td>
</tr>
<tr>
<td>999AA Aaaaaaa Aaaaaa 9999A A</td>
<td>1</td>
<td>5.88</td>
</tr>
</tbody>
</table>

Optionally, you can click the pattern to drill down to the values for the rows containing those patterns.

For pattern analysis, a 9 represents a digit, an A represents any uppercase letter, and an a represents any lowercase letter. All printable special characters are represented by themselves, and unprintable characters are represented by an X.

- **Hex View.** Shows the hexadecimal value of any field, as shown in the following image.
Values. Shows unique values and their percents, as shown in the following image.

Values Graph. shows a horizontal bar graph of the field values, as shown in the following image. The tooltip for each bar also shows the count of instances of the field.
Values Pie Graph. Shows a pie chart of the field values, as shown in the following image.

The tooltip for each bar also shows the count of instances of the field and the percent of the total.

Duplicate Values. Lists the values that have multiple instances, the count of instances, and the corresponding percent, as shown in the following image.
Outliers. Shows high and low values with a count of instances that fall outside the bulk of the data for a field, as shown in the following image.

<table>
<thead>
<tr>
<th></th>
<th>High</th>
<th>Count</th>
<th>Low</th>
<th>Cour</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>ZT Digital PDA - Commercial</td>
<td>1</td>
<td>110 VHS-C Camcorder 20 X</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>ZC Digital PDA - Standard</td>
<td>1</td>
<td>120 VHS-C Camcorder 40 X</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>R5 Micro Digital Tape Recorder</td>
<td>1</td>
<td>150 8MM Camcorder 20 X</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>QX Portable CD Player</td>
<td>1</td>
<td>2 H VCR LCD Menu</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>DVD Upgrade Unit for Cent. VCR</td>
<td>1</td>
<td>250 8MM Camcorder 40 X</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Combo Player - 4 H VCR + DVD</td>
<td>1</td>
<td>330DX Digital Camera 1024K P</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>AR3 35MM Camera 10 X</td>
<td>1</td>
<td>340SX Digital Camera 65K P</td>
<td></td>
</tr>
</tbody>
</table>

Note: Outliers produce a maximum of 10 highest and lowest distinct values, if they exist.

These reports are available by right-clicking a column in the Synonym Editor and pointing to Data Profiling.

Procedure: How to View a Data Profiling Report

To view a Data Profiling report for a single column:

1. Open the synonym in the Synonym Editor by double-clicking it from the navigation pane, or right-clicking it and clicking Open.

   The Synonym Editor opens to the Field View tab.

2. Right-click a column, point to Data Profiling, and then click one of the available reports:
   - Statistics
   - Key Analysis
   - Patterns
   - Hex View
   - Values
   - Values Graph
   - Values Pie Graph
   - Duplicate Values
   - Outliers
Reference: Data Profiling Statistics Report Columns

The Statistical Data Profiling information displays in the workspace, as shown in the following image.

The report shows:

**Segment**

The name of the segment in the synonym. For a cluster synonym, each segment can refer to a separate table; a segment can also refer to a group of columns.

**Name**

The column name. If you click the name, another report opens for that column showing each unique value and a count of the number of times the value appears and the percentage.

**Format**

The FOCUS format of the column.

**Count**

The number of values for the field.

**Distinct Count**

The number of unique values for the column.

**Distinct Percent**

The percentage of values that are unique.

**Patterns Count**

The number of unique patterns (combinations of letters and digits) that are found.

**Average Value**

Only for numeric fields.

**Minimum, Maximum, and Average Length**

Only for character fields.
**Nulls Count**
The number of values that are missing.

**Nulls Percent**
The percentage of values that are missing.

**Duplicate Count**
The number of values that are not unique.

**Duplicate Percent**
The percentage of values that are not unique.

The list of columns displayed can be selected from the ribbon by clicking `Customize Columns`. Clear any that are not desired. The following optional columns can also be displayed:

**Median**
The value where half the values are less and half the values are more.

**Mode**
The most commonly occurring value.

Optionally, you can click a column name or patterns count (for alphanumeric columns) to drill down to the Values or Patterns reports, respectively.

**Note:** Data profiling requires reading every row and column of the source table. Use caution with large data sources.

---

**Viewing Data Lineage for a Column**

Data lineage provides the ability to see how a column was derived. It traces the column back to the data source(s) it came from, breaking down all of its components along the way.

Data lineage is available from the Synonym Editor. It can show, for example, that a target column is the result of an SQL calculation that uses source columns from other synonyms.

**Procedure:** **How to View the Data Lineage of a Column**

To view the data lineage information for a column:

1. Open the synonym in the Synonym Editor by double-clicking it in the navigation pane, or right-clicking it and clicking `Open`.
   The Synonym Editor opens to the Field View tab.

2. Right-click a column and click `Data Lineage`. 
The lineage information displays in the workspace, as shown in the following image.
Viewing Data Lineage for a Column
Chapter 11

Advanced DataMigrator Tips and Techniques

This chapter describes several advanced techniques for DataMigrator.

The examples used to illustrate these techniques use sample databases and data, which you must create. The synonyms and sample data are created in the ibisamp directory. This directory is automatically installed when you install the server, and is automatically added to the server path.

In this chapter:

- Overview of Examples
- Creating Sample Data and Metadata
- Horizontal Partitioning with Multi-Table Load
- Using the Decode Function
- Generating a Surrogate Key
- Generating a Surrogate Key Using Slowly Changing Dimensions
- Mapping to a Sequential File of a Specific Format
- Creating and Using Reusable Transformations
- Using Multiple Targets to Create a Single File With Multiple Records Types
- Creating a Target File With Multiple RECTYPES
- Combining Multiple Sources Rows in a Single Target Row
- Using Lookups to Create Multi-Table Flows Without Joins
- Record Logging
- Using Variables in Data Flows
- Creating a Data Flow Using a Target Based on a Predefined XML Schema
- Loading a Target Table Using the File Listener
- Sending a Target File to an FTP or SFTP Server
- Loading a Star Schema
- Iterating Through Multiple Tables
- Using Subqueries
- Loading Images Into a Database Table
- Updating Values in a Table
Overview of Examples

The following table lists and describes some of the advanced techniques available to you with DataMigrator. Step-by-step instructions are provided for each sample flow. The sample flows include:

<table>
<thead>
<tr>
<th>Flow</th>
<th>Applies to Topic...</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMSPLIT</td>
<td>Horizontal Partitioning with Multi-Table Load on page 558</td>
</tr>
<tr>
<td>DMCODE</td>
<td>Using the Decode Function on page 562</td>
</tr>
<tr>
<td>DMSURRG</td>
<td>Generating a Surrogate Key on page 566</td>
</tr>
<tr>
<td>DMSURSCD</td>
<td>Generating a Surrogate Key Using Slowly Changing Dimensions on page 571</td>
</tr>
<tr>
<td>DMSEQ</td>
<td>Mapping to a Sequential File of a Specific Format on page 575</td>
</tr>
<tr>
<td>DMREUSE</td>
<td>Creating and Using Reusable Transformations on page 578</td>
</tr>
<tr>
<td>DMRCTYP</td>
<td>Creating a Target File With Multiple RECTYPEs on page 586</td>
</tr>
<tr>
<td>DMPIVOT</td>
<td>Combining Multiple Sources Rows in a Single Target Row on page 590</td>
</tr>
<tr>
<td>DMLKUPS</td>
<td>Using Lookups to Create Multi-Table Flows Without Joins on page 594</td>
</tr>
<tr>
<td>DMDUPLG</td>
<td>Record Logging on page 599</td>
</tr>
<tr>
<td>DMVAR</td>
<td>Using Variables in Data Flows on page 602</td>
</tr>
<tr>
<td>DMORDXML</td>
<td>Creating a Data Flow Using a Target Based on a Predefined XML Schema on page 607</td>
</tr>
<tr>
<td>DMFLORD</td>
<td>Loading a Target Table Using the File Listener on page 610</td>
</tr>
</tbody>
</table>

The sample files listed and described in the following table are used in the examples.

<table>
<thead>
<tr>
<th>Sample File</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>dmcomp.mas</td>
<td>Sample - Stores</td>
</tr>
</tbody>
</table>
Creating Sample Data and Metadata

Before you create and run the data flows discussed in this chapter, you must create the sample data that you will be using.

To create the sample data from the Data Management Console, see How to Create Sample Procedures and Data on page 78.

To create the sample data from the Web Console, see the Server Administration manual or the Web Console online help.

Note: You should set your Run Options for Maximum number of rows for test reports to 100 in the Run Options section of the Options dialog box so that the sample flows will retrieve data using any database.
Tip: You can build your flows in any application directory in the servers search path, including the ibisamp directory. However, you may find it useful to create a dedicated sample application directory called mysamp (or another name) and add it to the server search path, following the directions in Create Sample Data and Metadata and Create an Application Directory on page 437. You can then build all of the flows in that directory, using sample data and metadata, as required, from the ibisamp directory.

**Horizontal Partitioning with Multi-Table Load**

You can use DataMigrator to build a flow that extracts data using a single pass of the data and loads multiple data targets. This can be useful if, for example, you need to create separate databases for each department or region in your enterprise and you can split the extracted data based on the value of a department code or region code. The term for splitting the data based on the value of a column in the extracted data is horizontal partitioning. Horizontal partitioning of the data is just one useful way to take advantage of the ability to load multiple data targets in DataMigrator.

Implement horizontal partitioning in DataMigrator by building a data flow using multiple data targets together with VALIDATE to extract data from a data source, and split the output to two different data targets based on the value of an input column. This can be accomplished in one pass of the data.

To do this, create a data flow with one or more data sources and multiple data targets. Each data target has its own VALIDATE statement based on the value of a particular column. When the data target is loaded, only records with the appropriate values are accepted; all other records are rejected.

In the following example, DMINV, an inventory table, is the data source, and PRODTYPE is the input column used to split the output into two data targets. Records containing a value of Analog for PRODTYPE are loaded into a data target named DMINVA. Records containing a value of Digital for PRODTYPE are loaded into a data target named DMINVD.

Refer to the sample data flow DMSPLIT for the complete example.

**Note:** DMSPLIT in the ibisamp directory writes to FOCUS/FDS data targets. If you installed the samples for a different adapter, you will need to open and resave DMSPLIT before running it.

**Example:** **Creating a Data Flow for Horizontal Partitioning**

In the DMC:

1. Right-click an application directory in the navigation pane, select New, and then click Flow. The Data Flow tab opens in the workspace with the SQL object displayed.
2. Drag the data source object **DMINV** from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.

3. Right-click the SQL object and click **Column Selection**.
   
   The Column Selection window opens.

4. Select each column in the Available Columns list and click the double arrow to move them into the Selected Columns list. Click **OK**.
   
   **Tip:** To select all columns, click any cell and press Ctrl + A.

5. Drag the data target object **DMINVA** from the ibisamp application directory into the workspace, to the right of the SQL object.

6. Right-click the target object **DMINVA**, and click **Properties**.
   
   The Properties dialog box opens.

7. If the flow may be rerun, you should clear the target, using the **Prior to Load Option**. For a relational target, select **Delete all rows from table** or **Truncate Table**. For a FOCUS/FDS target, select **Drop Table**. Click **X** to close the window.

8. Right-click the target object **DMINVA** and click **Target Transformations**.
   
   The Transformations window opens.

9. Click the **Automap** button.

10. Click the **Validates** tab, and then click the **Insert Transforms** button.

   The Transformation Calculator opens with the Relational Expression tab selected.

11. Double-click **PRODTYPE** from the Columns/Variables tab.

12. Enter ‘**Analog**’, with the quotes, under Value, to complete the calculation, and click **OK**.
When you have completed the validation, the Transformations window should appear as shown in the following image:

13. Click OK to close the Transformations window.

14. Drag a second target object DMINVD from the ibisamp application directory into the workspace, to the right of the SQL object.

15. Right-click the target object DMINVD and click Properties.

   The Properties dialog box opens.

16. If the flow may be rerun, you should clear the target, using the Prior to Load Option. For a relational target, select Delete all rows from table or Truncate Table. For a FOCUS/FDS target, select Drop Table. Click X to close the Properties page.

17. Right-click the DMINVD target object and click Target Transformations.

   The Transformations dialog box opens with the Relational Expression tab selected.

18. Click the Automap button.

19. Click the Validates tab, and then click the Insert transforms button.

   The Transformation Calculator opens.

20. Double-click PRODTYPE from the Columns/Variables tab.
21. Enter ‘Digital’, with the quotes, under Value to complete the calculation and click OK.

22. Click OK to close the Transformations window.

When you have completed these steps, your data flow should appear as shown in the following image:

23. Save the data flow as DMXSPLIT.

24. On the Flow tab, in the Run group, click Run and select Submit from the drop-down menu.

Two data targets are loaded, DMINVA with records containing a PRODTYPE of Analog and DMINVD with records containing a PRODTYPE of Digital.
To view the log, right-click DMXSPLIT in the navigation pane, select Logs, and then click Last Log. The log of the data flow, shown in the following image, indicates that 17 rows were processed, 7 rows loaded into DMINVA, and 10 rows loaded into DMINVD.

<table>
<thead>
<tr>
<th>Datetime</th>
<th>Message Code</th>
<th>Log Messages</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>2012/05/2...</td>
<td>ICM18516 : Row(s) processed by job</td>
</tr>
<tr>
<td>17</td>
<td>2012/05/2...</td>
<td>ICM18514 : Row(s) accepted by target table</td>
</tr>
<tr>
<td>18</td>
<td>2012/05/2...</td>
<td>ICM18515 : Row(s) inserted into target table</td>
</tr>
<tr>
<td>19</td>
<td>2012/05/2...</td>
<td>ICM18517 : Row(s) updated in target table</td>
</tr>
<tr>
<td>20</td>
<td>2012/05/2...</td>
<td>ICM18518 : Row(s) deleted from target table</td>
</tr>
<tr>
<td>21</td>
<td>2012/05/2...</td>
<td>ICM18520 : Row(s) rejected due to validation error</td>
</tr>
<tr>
<td>22</td>
<td>2012/05/2...</td>
<td>ICM18521 : Row(s) rejected due to no match</td>
</tr>
<tr>
<td>23</td>
<td>2012/05/2...</td>
<td>ICM18522 : Row(s) rejected because duplicate exists</td>
</tr>
<tr>
<td>24</td>
<td>2012/05/2...</td>
<td>ICM18800 : Row(s) rejected due to DBMS error</td>
</tr>
<tr>
<td>25</td>
<td>2012/05/2...</td>
<td>-- stats for target file: dminvd</td>
</tr>
<tr>
<td>26</td>
<td>2012/05/2...</td>
<td>ICM18516 : Row(s) processed by job</td>
</tr>
<tr>
<td>27</td>
<td>2012/05/2...</td>
<td>ICM18514 : Row(s) accepted by target table</td>
</tr>
<tr>
<td>28</td>
<td>2012/05/2...</td>
<td>ICM18515 : Row(s) inserted into target table</td>
</tr>
<tr>
<td>29</td>
<td>2012/05/2...</td>
<td>ICM18517 : Row(s) updated in target table</td>
</tr>
</tbody>
</table>

Using the Decode Function

DECODE functions are used to convert codes into values. For example, you use DECODE to convert a two-character state code into the complete state name. DECODE can be explicitly stated or you can store the decoded values in a separate file.

The primary use of a DECODE function is to convert codes to full names. To use DECODE in a data flow, build a data flow with transformations that use the function.

In the following example, DMLOCS, a location table, is the data target. Add three transformations to map the source data columns to the target data columns. These transformations are known only to this data flow.

- In the first transformation, STATE, you convert a city abbreviation to its corresponding state code.

- In the second transformation, PLANT_CITY, you turn the city abbreviation into the complete city name.

- In the final transformation, PLANT_STATE, you convert a two-character state code into the full state name, using a list of state codes stored in a separate flat file, identified as DMDCD and named DMDCD.FTM.
The location of this state code file must be specified in the data flow so that it can be found at execution time. The procedure DMSTDEF specifies the location of the file, and DMSTDEF should be designated as a pre-extract stored procedure. This means that when the data flow runs, prior to the extraction of data from the source file, the procedure DMSTDEF runs. The file location is now known, and when the PLANT_STATE transformation is calculated, the state file used in the decode is found.

In this example, DMSTDEF contains the following code:

```
FILEDEF DMDCD DISK ibisamp/dmdcd.ftm
-RUN
-TYPE (ICM1) Filedef for State decode file returned &RETCODE
```

This specifies the location of the file used in the DECODE.

**Note:** If you want to be able to test the DMSTDEF procedure from the DMC, you must first make it available for processing, as illustrated in the example below.

Refer to the sample data flow DMDCODE for the complete example.

**Example:** Using the DECODE Function in a Data Flow

In the DMC:

1. Right-click an application directory in the navigation pane, select **New**, and then click **Flow**. The Data Flow tab opens in the right pane, with the SQL object displayed.

   In order to test the query from DataMigrator, you must first specify the location of the DMDCD file. Do this by adding the procedure DMSTDEF to your profile to define its location, and test it before you begin to construct the flow.

2. Right-click a server, select **My Console**, and then click **Edit My profile**.

3. When the Edit User Profile editor opens, enter the line below at the end of the profile:

   ```
   EX DMSTDEF
   ```

   Click **X** to close the editor, then click **Yes** to save the changes.

4. Drag the data source object **DMORD** from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.

5. Right-click the SQL object and click **Column Selection**.

   The Column Selection dialog box opens.

6. Select **PLANT** from the Available Columns list, and click the arrow to move it to the Selected Columns list.
7. Click the Distinct check box, and click OK to close the Column Selection window.

8. Right-click in the workspace to the right of SQL object, select Add Target, and then click Existing.

9. Select DMLOCS from the ibisamp application directory, and click Select.
   The data target object DMLOCS is added to the workspace.

10. Right-click the target object DMLOCS, and click Properties.
    The Properties page opens.

11. If the flow may be rerun, you should clear the target using the Prior to Load Option. Since DMLOCS is a flat file, select Delete File. Click X to close.

12. Right-click the DMLOCS target object and click Target Transformations.
    The Transformations dialog box opens.

13. Click the Insert Transforms button.
    The Transformations Calculator opens.

14. Enter STATE in the Column field and A2 in the Format field.

15. In the expression field, type

   DECODE PLANT (BOS MA DAL TX LA CA ORL FL SEA WA STL MO)

   and click OK.

16. Double-click PLANT_CITY in the Target Columns list. This moves it to the Expressions list.
    Click Edit Transforms to open the Transformations Calculator.

17. Double-click the Transformation Expression field for PLANT_CITY and type the following as the transformation expression:

   DECODE PLANT (BOS 'Boston' DAL 'Dallas' LA 'Los Angeles' ORL 'Orlando'
   SEA 'Seattle' STL 'St. Louis')

18. Click OK to close the calculator.

19. Double-click PLANT_STATE in the target columns list. This moves it to the Expressions list.

20. Double-click the Transformation Expression field for PLANT_STATE and type the following as the transformation expression:

   DECODE STATE (DMDCD)

   Click OK to close the calculator.
When you have defined the three transformations, the Transformations dialog box should appear, as shown in the following image.

21. Click the Test transforms button to test the transformations you just built. Click Close. The following image shows the Test Transformations dialog box.

22. Click OK to close the Transformations window.

**Note:** If you leave EX DMSTDEF in your profile, then steps 22 to 26 are not needed.

If another user were going to use this flow, then you would need to perform the following steps.
23. Click the Process Flow tab.

24. Right-click the arrow between the Start object and the Data Flow object, and click Delete. You will be inserting the procedure between them.

25. Drag the procedure, DMSTDEF, from the ibisamp application directory in the navigation pane into the process flow workspace.

26. Drag from the Start object to the DMSTDEF object to connect them, then drag from the DMSTDEF object to the Data Flow object to connect them.

27. Save the flow as DMXDCODE.

28. On the Flow tab, in the Run group, click Run, and then select Run from the drop-down menu.

To verify that the target was loaded properly and that the transformations evaluated properly, click the Data Flow tab, right-click the DMLOCS data target, select Operations, and then click Sample Data.

Generating a Surrogate Key

Many data warehousing applications make use of a surrogate key instead of using the keys from an operational system, like an alphanumeric employee ID or an integer employee number.

A surrogate key is usually an incrementing integer with no intrinsic meaning. One reason applications use surrogate keys is to improve the performance of JOINs.

You can build a data flow to load a table that has surrogate keys by using a self-referential transformation to generate the values for the surrogate key field.

In the following example, DMSALE, a sales table, is the data source, and the extracted records are loaded into a table named DMREPS. Values for the surrogate key REPNO are generated when the data flow is run.

Refer to the sample data flow DMSURRG for the complete example.

Example: Generating a Surrogate Key in a Data Flow

In the DMC:

1. Right-click an application directory in the navigation pane, point to New, and then click Flow. The Data Flow tab opens in the right pane, with the SQL object displayed.

2. Drag the data source object DMSALE from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.

3. Right-click the SQL object and click Column Selection.

    The Column Selection dialog box opens.
4. Select EMPID, PLANT, and SALESREP in the Available Columns list and click the arrow to add them to the Selected Columns list.

5. With all three columns selected, select Group By from the Aggregate drop-down menu.

6. Click the Insert Columns button.

   The SQL Calculator opens.

7. In the Alias box, enter REPNO for the virtual column.

8. Click the Functions tab, expand the Miscellaneous folder, and double-click the COUNTBY function.

   The Function Assist dialog box opens.

9. Leave the attribute value set to 1 and click OK.

   The SQL Calculator should appear as shown in the following image.

10. Click OK to close the SQL Calculator.
The Column Selection dialog box should appear as the following image.

11. Click OK to close the Column Selector dialog box.

12. Drag the data target object DMREPS from the ibisamp application directory into the workspace, to the right of the SQL object.

13. Right-click the target object DMREPS, and click Properties.

   The Properties dialog box opens.

14. If the flow may be rerun, you should clear the target using the Prior to Load Option dropdown menu. For a relational target, select Delete all rows from table or Truncate Table. For a FOCUS/FDS target, select Drop Table. Close the Properties dialog box.

15. Right-click the target object and click Target Transformations.

   The Transformations dialog box opens.

16. Click the AutoMap button to map REPNO, EMPID, and PLANT. This moves them to the Expressions list.

   To map FNAME and LNAME, use substring to extract the necessary characters from the SALESREP source column, of which the first 16 characters are first name and the next 34 characters are last name.

17. Double-click FNAME in the Target Columns list.

   The Transformation Calculator opens.

18. Type the following expression under the Expression tab:

   \[
   \text{SUBSTR}(50, \text{SALESREP}, 1, 16, 16, \text{"A16"})
   \]

   Click OK.

19. Double-click LNAME in the Target Columns list.

   The Transformation Calculator opens.
20. Type the following expression under the Expression tab:

```sql
SUBSTR (50, SALESREP, 17, 50, 34, 'A34')
```

**Note:** While the creation of a surrogate key does not require the use of the SUBSTR function, it is a useful routine that can be used in a number of different situations.

Click OK.

21. To map STATE, double-click `STATE` in the Target Columns list.

The Transformation Calculator opens.

22. Type the following expression under the Expression tab:

```sql
DECODE PLANT (BOS MA DAL TX LA CA ORL FL SEA WA STL MO)
```

This expression will convert a three-character city name to its two-character state code.

Click OK.

23. Double-click `ACTIVE` in the Target Column list.

The Transformation Calculator opens.

24. Type 1 in the Expression tab.

25. Click OK.

When you have completed the transformations, the Transformations dialog box should appear, as shown in the following image.
26. Test the transformations by clicking the Test Transforms button.
   The Test Transformations dialog box opens, as shown in the following image.

![Test Transformations dialog box](image)

Note that REPNO is incremented by one for each record and that FNAME and LNAME are transformed correctly.

Close the Test Transformations dialog box.

27. Click OK to close the Transformations dialog box.

28. Save the data flow as DMXSURRG.

29. On the Flow tab, in the Run group, click Run, and then click Run from the drop-down menu.

When the data flow run is complete, verify that the target was loaded properly and that the surrogate key was generated by right-clicking the DMREPS object, clicking Expand, and then clicking the Sample Data tab in the window that opens. The following image shows the DMREPS window with the Sample Data tab selected.

![Sample Data window](image)

There are two ways to confirm that the job ran correctly:
On the Flow tab, in the Reports group, click View Last Log.

or

Right-click the flow from the navigation pane, point to Logs, and then click Last Log.

The following image shows the log of the most recently run procedure.

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Message Code</th>
<th>Log Messages</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18874)</td>
<td>Start of Log Record for ibisamp/dmxa...</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18742)</td>
<td>dmrep type MS SQL Server ODBC Existing target</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18429)</td>
<td>Issuing PREPARE</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(FOC1364)</td>
<td>ROWS AFFECTED BY PASSTHRU COMMAND: 148/DELETE</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18743)</td>
<td>Starting Load</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td></td>
<td>0 WARNING. ON MATCH INCLUDE INPUTS DUPLICATE SEGMENTS</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18041)</td>
<td>-- stats for source file</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18516)</td>
<td>148: Row(s) processed by job</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18519)</td>
<td>0: Row(s) rejected due to format error</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18372)</td>
<td>-- stats for target file dmrep</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18516)</td>
<td>148: Row(s) processed by job</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18514)</td>
<td>148: Row(s) accepted by target table</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18515)</td>
<td>148: Row(s) inserted into target table</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18517)</td>
<td>0: Row(s) updated in target table</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18518)</td>
<td>0: Row(s) deleted from target table</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM190520)</td>
<td>0: Row(s) rejected due to validation errors</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM190521)</td>
<td>0: Row(s) rejected due to no match</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM190522)</td>
<td>0: Row(s) rejected because duplicate exist</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18808)</td>
<td>0: Row(s) rejected due to DBMS error</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM19074)</td>
<td>Ending Load</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18040)</td>
<td>Return Code = 0</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18079)</td>
<td>Request: ibisamp/dmxa... finished processing</td>
</tr>
<tr>
<td>2017/08/23 12:58:08</td>
<td>(ICM18007)</td>
<td>CPU Time: 110</td>
</tr>
<tr>
<td>2017/08/23 12:58:09</td>
<td>(ICM18031)</td>
<td>Finished</td>
</tr>
</tbody>
</table>

**Generating a Surrogate Key Using Slowly Changing Dimensions**

DataMigrator includes support for automatically generating surrogate keys when used in a dimension table. This support is only available for relational databases and cannot be used with FOCUS or flat file targets.

For details about Slowly Changing Dimension support, see *Target Properties Pane for Relational Targets* on page 309.

Refer to the sample data flow DMSURSCD for the complete example.
**Example: Generating a Surrogate Key Using Slowly Changing Dimensions in a Data Flow**

In the DMC:

1. On the Home tab, in the Tools group, click Options, and then select Column Management in the Options dialog box.
2. Expand the Business View Editor folder and select SCDType to display it on the column grid. Click OK.
3. Right-click an application directory in the navigation pane, select New, and then click Flow. The Data Flow tab opens in the right pane, with the SQL object displayed.
4. Drag the data source object DMSALE from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.
5. Right-click the SQL object and click Column Selection.

   The Column Selection dialog box opens.
6. Select EMPID, PLANT, and SALESREP in the Available Columns list and click the arrow to add them to the Selected Columns list. Click OK.
7. Drag the data target object DMREPS from the ibisamp application directory into the workspace, to the right of the SQL object.
8. Right-click the target object DMREPS and click Properties.

   The Properties dialog box opens.
9. If the DMSURRG flow has been run previously, clear the target using the Prior to Load Option by selecting Delete all rows from table or Truncate Table.
10. To enable Slowly Changing Dimension support, for Load Options select Slowly Changing Dimensions, as shown in the following image.
11. Right-click the target object and click Target Transformations. The Transformations dialog box opens.

12. Click the Automap button to map EMPID and PLANT. This moves them to the Expressions list.

   To map FNAME and LNAME, use substring to extract the necessary characters from the SALESREP source column, of which the first 16 characters are first name and the next 34 characters are last name.

13. Double-click FNAME and LNAME in the target columns list. This moves the columns to the Expressions list.

14. Double-click the Transform Expression field for FNAME and type the following expression:
   
   \[ \text{SUBSTR (50, SALESREP, 1, 16, 16, 'A16')} \]

   Click OK.

15. Double-click the Transform Expression field for LNAME and type the following expression:
   
   \[ \text{SUBSTR (50, SALESREP, 17, 50, 34, 'A34')} \]

   **Note:** While the creation of a surrogate key does not require the use of the SUBSTR function, it is a useful routine that can be used in a number of different situations.

   Click OK.

16. To map STATE, double-click STATE in the Target Columns list. This moves the column to the Expressions list.

17. Double-click the Transform Expression field for STATE and type the following:
   
   \[ \text{DECODE PLANT (BOS MA DAL TX LA CA ORL FL SEA WA STL MO)} \]

   This converts a three character city name to its two character state code.
When you have completed the transformations, the Transformations dialog box should appear, as shown in the following image.

![Transformations dialog box](image)

The target table DMREPS has been created with slowly changing dimension attributes assigned for each of its columns. Since REPNO is a key column and the type is blank, it becomes the surrogate key and no transformation is required. The values for ACTIVE are also automatically generated when the data flow is run.

18. Test the transformations by clicking the Test transforms button. Click Close.

19. Click OK to close the Transformations dialog box.

20. Save the data flow as DMXSURSCD.

21. In the Flow tab, in the Run group, click Run, and then select Run from the drop-down menu.

When the data flow run is complete, verify that the target was loaded properly and that the surrogate key was generated by right-clicking the DMREPS object and selecting Toggle, and then Sample Data.
Note that the REPNO surrogate key values are assigned sequentially. However since PLANT and STATE are defined as SCD Type II, an employee is assigned a new surrogate key if the plant they are associated with changes. The ACTIVE flag is set to 1 for their current plant. The following image shows the DMREPS with the Sample Data tab selected.

There are two ways to confirm that the job ran correctly:

On the Flow tab, in the Reports group, click View Last Log.

or

Right-click the flow from the navigation pane, select Logs, and then click Last Log.

**Mapping to a Sequential File of a Specific Format**

Situations arise where you need to extract data from data sources and load a sequential table with a specific layout. For example, you may need to load a packaged application of a very specific format with a mix of both packed and alphanumeric columns. The data you load must be in the order the packaged application expects and the formats must match exactly. This means that when you build your data flow, you must make sure that each column is in the expected order and format.

To do this, build a data flow to load a sequential file with a previously defined synonym. This means that you create the synonym to match the format required by the application you are loading and you use that synonym as the target of the data flow.

In the following example, DMHR, a human resources table, is the data source, and extracted records are loaded into a data target sequential file named DMHRINP. A synonym describing the layout of this sequential table was created in the format of the sequential file that is being loaded. The sequential file includes two packed columns and two alphanumeric columns.
Refer to the sample data flow DMSEQ for the complete example.

**Example:**  
*Mapping to a Sequential File in a Data Flow*

In the DMC:

1. Right-click an application directory in the navigation pane, select *New*, and then click *Flow*. The Data Flow tab opens in the right pane, with the SQL object displayed.
2. Drag the data source object *DMHR* from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.
3. Right-click the SQL object and click *Column Selection*.

   The Column Selection window opens.

4. Select *ID_NUM*, *LNAME*, *FNAME*, and *PAYSCALE* in the Available Columns list and click the arrow to add them into the Selected Columns list. Click *OK*.
5. Drag the data target object *DMHRINP* from the ibisamp application directory into the workspace, to the right of the SQL object.
6. Right-click the target object *DMHRINP* and click *Properties*.

   The Properties dialog box opens.

7. If the flow may be rerun, you should clear the target using the *Prior to load* option. Since *dmhrinp* is a Flat File, select *Delete File*. Click *OK*.
8. Right-click the target object and click *Target Transformations*.

   The Transformation window opens.

9. Select *ID* in the Target Columns list and *ID_NUM* in the Selected Columns list, then click the equal sign (=). The relation is added to the Expressions list.
10. Select *SCALE* in the Target Columns list and *PAYSCALE* in the Selected Columns list, then click the equal sign (=). The relation is added to the Expressions list.
11. Click the *Automap* button to map *LNAME* and *FNAME*.
When you have completed the transformations, the Transformations window should appear, as shown in the following image:

12. Test the transformations by clicking the Test transforms button. Click Close.
13. Click OK to close the Transformations window.
14. Save the data flow as DMXSEQ.
15. On the Flow tab, in the Run group, click Run, and then select Run from the drop-down menu.
When the data flow run is complete, verify that the data was loaded properly in the sequential file by right-clicking the DMHRINP object, selecting Toggle, and clicking the Sample Data tab, as shown in the following image.

Creating and Using Reusable Transformations

There are many situations where you need to perform the same calculation on several different columns. It could be something as simple as converting a date to a different format, adding dashes to an ID number so that it displays as you would expect a social security number to display, or something much more complex that requires you to perform the same calculations on different columns of data. Instead of coding the calculation over and over for each column, it would be more efficient to create the calculation once, and then use it whenever you need that calculation. In essence, you create a reusable transformation.

Before you build the transformation, you must first create the function to perform the necessary transformations. Functions are created in the DMC as a server object with the type Define Function.

This technique is divided into two parts. In the first part, we build the function. In the second, we use the function in transformations within a data flow, then incorporate the data flow object into a process flow to make it available at execution time. In the following example, we build two functions:

- **DMPROPER** converts a name to last name, first name format, where the first letter of each name is in upper case and the rest of the name is in lowercase.

- **DMDATFMT** converts a date to 2009-JUL-04 format, which includes dashes between the year, month, and day.
These routines are stored in a stored procedure which is identified in the Data Management Console with a distinguished icon and a type of Userin Functions. In the following example, DMHR, a human resource table, is the data source, and the extracted records are loaded into a new data target named EMPDATES. Employee name (PROPERNAME), start date (START), and termination date (TERM) are calculated using the define functions.

Refer to the sample data flows in the ibisamp directory, DMREUSE, for the complete example. The User Functions procedure DMFNS contains the sample function DMPROPER and DMDATFMT.

**Example:** Creating the DEFINE Functions

The first function, which we call DMPROPER, takes two parameters, a last name and a first name. It returns a formatted proper name.

In the DMC:

1. Right-click an application directory in the navigation pane, select New, and then click User Functions.

   The Define Function View opens in the workspace in tree view.

2. Click FUN1 and change the name to DMPROPER.

3. Under the Parameters folder, click PARAM1.

   The Properties panel opens on the right.

4. In the PARAM field box, click the PARAM1 value and type the name of the first parameter LASTNAME. In the FORMAT field box, type in the value A17.

   Remove 1 from the DEFAULT field box in the Miscellaneous section. Select the Show Field List check box.

5. Right-click the Parameters folder, select Insert, and then click Parameter. In the PARAM field box in the properties panel, click the PARAM1 value and type the name of the first parameter FIRSTNAME. In the FORMAT field box, type in the value A14. Under Miscellaneous, select the Show Field List check box.

6. Under Defines folder, click DMPROPER. Its properties panel opens to the right. In the EXPRESSION field box, click the ellipsis button.

   The Define Calculator opens.

7. Enter A34V in the Format field.

   Type the following in the Expression field and click OK to close the calculator:

   \[
   \text{LCWORD}(17, \text{LASTNAME}, 'A17') || ('', ' | \text{LCWORD}(14, \text{FIRSTNAME}, 'A14'))
   \]
8. Click Save As in the Quick Access Toolbar and save the function file as \textit{dmfns}.

The second function converts a date to a common format.

1. Right-click \textit{DMFNS}, select \textit{Insert}, and then click \textit{Function}.

   A new function prototype is added.

2. Click \textit{FUN1} and change the name to \textit{DMDATFMT}.

3. Under the Parameters folder, click \textit{PARAM1}. Its properties panel opens to the right. In the \textit{PARAM} field box, click the \textit{PARAM1} value and type the name of the first parameter \textit{INDATE}. In the \textit{FORMAT} field box, type in the value \textit{YYMD} and in the DEFAULT field box under the Miscellaneous section, type in the value \texttt{&YYMD}. Select the \textit{Show field list} check box.

4. Right-click the \textit{Defines} folder, select \textit{Insert}, and then click \textit{Define}. In the \textit{DEFINE} field box in the properties pane, click the \textit{DEFINE1} value and type the name of the parameter \textit{CONVDATE}. In the \textit{EXPRESSION} field box, type in the value

\[
\text{DATECVT( INDATE, 'YYMD', 'A8YYMD' )}
\]

In the \textit{FORMAT} field box, type in the following value:

\textit{A8YYMD}

5. Click \textit{OK} to close the calculator.

6. Under the \textit{Defines} folder, click \textit{DMDATFMT}. Its properties panel opens to the right. In the \textit{EXPRESSION} field box, type in the following value:

\[
\text{CHGDAT( 'YYMD', 'YYMTD', CONVDATE, 'A11' )}
\]

In the \textit{FORMAT} field box, type in the following value:

\textit{A11}

7. Test the transform by selecting the function name \textit{DMDATFMT}. Click the \textit{Test Function} button.

   The Test Function window opens. Under Value, enter a numeric date, such as 20120202. The Report window opens, showing the default value of today’s date.

8. Click Save from the Quick Access Toolbar.

\textbf{Note:} When creating a define function, the following names must be the same:

- The name of the function.
- The name of the last define in the function.
Example: Using the DEFINE Functions

In the DMC:

1. Right-click an application directory in the navigation pane, select New, and then click Flow. The Data Flow tab opens in the workspace, with the SQL object displayed.

2. Drag the data source object DMHR from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.

3. Right-click the DMHR object and click Source Transformations.

   The Source Transformations window opens.

4. Click the Insert Transforms button.

   The Transformations Calculator opens.

5. Enter PROPERNAME in the Name field and A30 in the Format field.

6. Click the Functions tab and expand the User Defined folder, followed by the baseapp directory and the dmfns procedure, as shown in the following image.

7. Double-click DMPROPER.
The *Function Assist* window opens. Select *LNAME* and *FNAME* from the Field list drop-down menus for the parameters, as shown in the following image.

8. Click OK to close the Function Assist window and the Transformation Calculator.

9. Click the *Insert transforms* button.

10. Enter *START* in the Column field and *A12* in the Format field.

   Click the *Functions* tab. Expand the *User Defined* folder, the *baseapp* directory, and the *dmfns* procedure.

11. Double-click the *DMDATFMT* function. The function assistant opens. Select *START_DATE* as the parameter for the function.

   Click OK to close the Function Assist and the Transformations Calculator.

12. Click the *Insert Transforms* button. Enter *ENDDATE* in the Column field and *A12* in the Format field.

13. Click the ellipsis button in the Format field. Select the *Show Advanced Properties* check box and select the *Allow Missing Data* check box.

   Repeat the above procedure to select the *DMDTEFMT* function, using *TERM_DATE* as the parameter.
14. When you have completed the transformations, the Source Transformations window opens.

15. Click OK to close the Transformations window.

16. Right-click the SQL object and click Column Selection.

   The Column Selection window opens.

17. In the Available Columns list, select PROPERNAME, START, and ENDDATE and click the arrow to add them to the Selected Columns list.

18. To test the transformations, click the Test SQL Statement button. Note that the transformations are correct, with PROPERNAME correctly defined, and the correct START and ENDDATE dates.

19. Click X to close the Test SQL window, and OK to close the Column Selection window.

20. Right-click the workspace to the right of the SQL object, select Add Target, and then click New.

21. In the new target, right-click and click Properties.

22. In the Properties panel, select the Adapter you used for your sample procedures, the application directory you are using for your samples, and enter EMPDATES as the synonym. Click X to close the panel.

23. Save the flow as DMXREUSE.

24. On the Flow tab, in the Run group, click Run, and then select Submit from the drop-down menu.

   To verify that the transformations have worked properly, click the Data Flow tab and right-click the EMPDATES data target and select Toggle. Click the Sample Data tab to review the output.

Using Multiple Targets to Create a Single File With Multiple Records Types

This flow uses two targets that both point to the same flat file to write out two different record types from the HR file. There is a start record for all employees, but there is only a termination record for employees that have been terminated which is accomplished with a validate. The two different record types are designated by assigning a value to the RECTYPE field in the targets.

Refer to the sample data flow DMRCTYP for the complete example.

In the DMC:

1. Right-click an application directory, select New, and then click Flow. The Data Flow tab opens in the workspace, with the SQL object displayed.

2. Drag the DMHR source from the ibisamp application directory to the left of the SQL object.

3. Right-click the SQL object and click Column Selection.
4. Select ID_NUM, START_DATE, TERM_DATE, STATUS, POSITION and PAYSCALE in the available columns list, and click the arrow to add them to the Selected columns list. Click OK.

5. In the ibisamp application directory, Ctrl+click DMSTART and DMTERM to select them both. Next, drag them to the right of the SQL object, as shown in the following image.

6. Double-click DMTERM to open the Transformations window. Click the Automap button.

7. Click the Insert Transforms button. The Transformation Calculator opens. For the name type RECTYPE, for the format type A1, and for expression type 'T'.

8. Double-click SEVERENCE to add it to the Expressions. For the expression, enter PAYSCALE * 500.


10. Select the Expression tab and enter TERM_DATE NE MISSING. Click OK.

11. Click OK.

12. Double-click DMSTART to open the Transformations window. Select the Automap button.

13. Select the Insert Transforms button. The Transformation Calculator opens. For the name, type RECTYPE. For the format, type A1 and for expression, type 'S'.

14. Click OK.

15. On the Flow tab, in the Run group, click Run, and then select Submit from the drop-down menu.
16. On the Flow tab, in the Reports group, click View Last Log. Note that 223 Start records and 41 Termination records were loaded.

<table>
<thead>
<tr>
<th>Message Code</th>
<th>Log Messages</th>
</tr>
</thead>
<tbody>
<tr>
<td>(ICM18974)</td>
<td>------ Start of Log Record for ibisamp/dmrcotyp ------</td>
</tr>
<tr>
<td>(ICM18122)</td>
<td>Request - ibisamp/dmrcotyp (Owner: dm) submitted.</td>
</tr>
<tr>
<td>(ICM18742)</td>
<td>dmstart type Flat File Existing target</td>
</tr>
<tr>
<td>(ICM18372)</td>
<td>-- stats for target file: dmstart</td>
</tr>
<tr>
<td>(ICM18516)</td>
<td>223: Row(s) processed by job</td>
</tr>
<tr>
<td>(ICM18519)</td>
<td>0: Row(s) rejected due to format error</td>
</tr>
<tr>
<td>(ICM18520)</td>
<td>0: Row(s) rejected due to validation errors</td>
</tr>
<tr>
<td>(ICM18521)</td>
<td>0: Row(s) rejected due to no match</td>
</tr>
<tr>
<td>(ICM18522)</td>
<td>0: Row(s) rejected because duplicate exist</td>
</tr>
<tr>
<td>(ICM18518)</td>
<td>0: Row(s) rejected due to DBMS error</td>
</tr>
<tr>
<td>(ICM18372)</td>
<td>-- stats for target file: dmterm</td>
</tr>
<tr>
<td>(ICM18516)</td>
<td>223: Row(s) processed by job</td>
</tr>
<tr>
<td>(ICM18514)</td>
<td>41: Row(s) accepted by target table</td>
</tr>
<tr>
<td>(ICM18515)</td>
<td>41: Row(s) inserted into target table</td>
</tr>
<tr>
<td>(ICM18517)</td>
<td>0: Row(s) accepted by target table</td>
</tr>
<tr>
<td>(ICM18518)</td>
<td>0: Row(s) deleted from target table</td>
</tr>
<tr>
<td>(ICM18520)</td>
<td>182: Row(s) rejected due to validation errors</td>
</tr>
<tr>
<td>(ICM18521)</td>
<td>0: Row(s) rejected due to no match</td>
</tr>
<tr>
<td>(ICM18522)</td>
<td>0: Row(s) rejected because duplicate exist</td>
</tr>
<tr>
<td>(ICM18808)</td>
<td>0: Row(s) rejected due to DBMS error</td>
</tr>
<tr>
<td>(ICM18744)</td>
<td>Ending Load</td>
</tr>
<tr>
<td>(ICM18040)</td>
<td>Return Code = 0</td>
</tr>
<tr>
<td>(ICM18076)</td>
<td>Request: ibisamp/dmrcotyp - finished processing</td>
</tr>
<tr>
<td>(ICM18007)</td>
<td>CPU Time: 46</td>
</tr>
<tr>
<td>(ICM18031)</td>
<td>Finished</td>
</tr>
<tr>
<td>(ICM18072)</td>
<td>Elapsed run time 0:00:01</td>
</tr>
<tr>
<td>(ICM18975)</td>
<td>------ End of Log Record for ibisamp/dmrcotyp ------</td>
</tr>
</tbody>
</table>

17. To see the output with the two different record types in the ibisamp directory with no filter, scroll down to dmrecs - Flat File. Next, right-click dmrecs and click Edit. Scroll down to see the termination records.
Creating a Target File With Multiple RECTYPEs

Instead of loading a target, you can also use DataMigrator to extract data from a data source and create input files for a packaged application that requires RECTYPEs indicating different record layouts.

DataMigrator can create a flat file containing the different record layouts. This requires a data flow that loads a sequential file with a previously defined layout. The synonym for the sequential file contains one segment for each RECTYPE, along with all the associated columns. The RECTYPE identifies the type of record in the ALIAS. Each segment has a unique value for RECTYPE.

The following example uses a single synonym with multiple segments to create a flat file containing two RECTYPEs. DMORD, an orders table, is the data source. Extracted records are loaded into one sequential file data target, DMORECS, which contains two segments.

The first segment contains the header information, indicated by a RECTYPE of 'H', and includes the columns ORDER_DATE, STORE_CODE, and PLANT, as shown in the following image.
The second segment contains the detail information, indicated by a RECTYPE of 'D', and includes the columns PROD_NUM, QUANTITY, and LINEPRICE, as shown in the following image.
Creating a Target File With Multiple RECTYPEs

The graphical view of the DMORECS synonym shows how the two segments are related. OINFO contains the header records and PINFO contains the detail records, as shown in the following image.

![Diagram showing the relationship between OINFO and PINFO segments]

Refer to the sample data flow ORECS for the complete example.

**Example:** Creating a Flat File With Different Record Layouts

In the DMC:

1. Right-click an application directory, select New, and then click Flow. The Data Flow tab opens in the workspace, with the SQL object displayed.
2. Drag the DMORD data source object from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.
3. Right-click the SQL object and click Column Selection.
   
   The Column Selection dialog box opens.
4. Select ORDER_NUM, ORDER_DATE, STORE_CODE, PLANT, PROD_NUM, QUANTITY, and LINEPRICE in the Available Columns list and click the arrow to add them to the Selected Columns list. Click OK.
5. Drag the DMORECS data target object from the ibisamp application directory into the workspace, to the right of the SQL object.

6. Right-click DMORECS and click Properties.
   
The Properties dialog box opens.

7. If the flow will be rerun, you should clear the target using the Prior to Load Option. Since DMORECS is a flat file, select Delete File. Click X to close the Properties panel.

8. Right-click DMORECS and click Target Transformations.
   
The Transformation dialog box opens.

9. To map the source columns to the data target, click the AutoMap button.
   
The Transformations dialog box opens, as shown in the following image.

![Transformations Dialog Box](image)

Note that the RECTYPE columns remain unmapped. The RECTYPE values are assigned automatically at load time based on the value given for the RECTYPE in the synonym, in this case, H or D.

10. Save the data flow as DMXRCTYP.

11. On the Flow tab, in the Run group, click Run and select Submit from the drop-down menu.
Combining Multiple Sources Rows in a Single Target Row

To view the output file created by the data, ensure that the No Filter button is selected in the Browser. Expand the baseapp directory and scroll down past the list of synonyms. There you will see a filename of dmorecs with a type of flat file. Double-click it to open it.

Note that the output file contains records with two different layouts.

Combining Multiple Sources Rows in a Single Target Row

Sometimes data from multiple input rows needs to be combined into a single row. For example, you might encounter a situation where three separate records from a data source map to one record in the data target.

In the following example, the data source contains transaction records with an ID number, action code, and date. The action code, with values of either S or T, indicates whether the date is a Start Date or Termination Date. For example, the data source contains:

<table>
<thead>
<tr>
<th>ID_NUM</th>
<th>ACTION</th>
<th>ACT_DATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>S</td>
<td>1997/03/26</td>
</tr>
<tr>
<td>24</td>
<td>S</td>
<td>1995/08/24</td>
</tr>
<tr>
<td>24</td>
<td>T</td>
<td>1999/02/16</td>
</tr>
<tr>
<td>25</td>
<td>S</td>
<td>1995/10/05</td>
</tr>
<tr>
<td>25</td>
<td>T</td>
<td>1999/01/15</td>
</tr>
</tbody>
</table>

The data flow takes the source data and flips, or pivots, it so that the data target contains one row for each ID_NUM, each with two date columns: one for the Start Date, and one for the Termination Date (if any).

This solution uses two different synonyms to describe the same source file. Each synonym recasts the ACTION column to a RECTYPE, allowing the file to be split into Start Date records and End Date records. A select against DMSTART automatically selects only those records with a RECTYPE of S and a select against DMTERM retrieves only those with a RECTYPE of T.
The synonym for DMSTART shows a RECTYPE of S, which would retrieve only Start Date records, as shown in the following image.

The synonym for DMTERM shows a RECTYPE of T, which would retrieve only End Date records, as shown in the following image.

When the job is run, the output table for the ID_NUM values shown above is:
### Combining Multiple Sources Rows in a Single Target Row

<table>
<thead>
<tr>
<th>ID_NUM</th>
<th>START_DATE</th>
<th>TERM_DATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>1997/03/26</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>1995/08/24</td>
<td>1999/02/16</td>
</tr>
<tr>
<td>25</td>
<td>1995/10/05</td>
<td>1999/01/15</td>
</tr>
</tbody>
</table>

Since there was no T row in the input file for ID_NUM value 23, there is no TERM_DATE value in the output.

**Note:**

- This example uses the output from the data flow DMRCTYP as a source. Therefore, you must execute DMRCTYP before you execute this request.
- This example can easily be extended to support more than two record types by adding a synonym for each record type.

Refer to the sample data flow DMPIVOT for the complete example.

**Example:** Combining Data Source Rows in a Single Data Target Row

In the DMC:

1. Right-click an application directory in the navigation pane, select New, and then click Flow. The Data Flow tab opens in the workspace, with the SQL object displayed.
2. Drag the data source object DMSTART from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.
3. Drag a second data source object, DMTERM, into the workspace.
4. To define the Join condition, right-click the Join object and click Join Editor. The Join Editor window opens.
   **Note:** Since the data sources are flat files without any key columns, no join conditions are defined automatically, even if that setting is in effect.
5. Select ID_NUM in the Left and Right Source Columns lists and click the equal (=) sign.
6. To make sure that you obtain those records that have a START_DATE but not an END_DATE, the JOIN type should be a left outer Join. Click the left circle in the Join Type diagram.
When you have completed the Join, it should appear, as shown in the following image:

7. Click OK to close the Join Editor window.

8. Right-click the SQL object and click Column Selection.

   The Column Selection window opens.

9. Select T1.ID_NUM, T1.START_DATE, and T2.TERM_DATE in the Available Columns list and click the arrow to add them into the Selected Columns list. Click OK.

10. Right-click the workspace to the right of the SQL object, select Add Target, and then click New.

11. In the Properties window, select the adapter you used for your sample procedures, the application directory you are using for your samples, and enter DMXEDATE as the synonym. Click OK.

12. Save the data flow as DMXPIVOT.

13. On the Flow tab, in the Run group, click Run, and then select Run from the drop-down menu.
**Note:** Run *ibisamp/dmrctyp* first.

When the data flow run is complete, verify that the data was loaded properly by right-clicking the *DMXEDATE* object, selecting *Toggle*, and clicking the *Sample Data* tab.

### Using Lookups to Create Multi-Table Flows Without Joins

DataMigrator now allows you to build a single data flow that performs lookups into multiple tables using the `DB_LOOKUP` function without requiring explicit joins.

In the following example, DMORD is the primary source containing orders data, while DMCOMP and DMINV are the lookup sources. While DMORD contains store codes and product numbers, DMCOMP contains the store names for those store codes, and DMINV contains the product names for each product code.

Refer to the sample data flow DMLKUPS for the complete example.

For more information on improving lookup performance, see *Improving Lookup Performance* on page 498.

**Example:** Creating a Data Flow Using Lookups

This example uses the `DB_LOOKUP` function to create two target transformations.

1. In the DMC, right-click an application directory in the navigation pane, point to New, and then click Flow. The Data Flow tab opens in the workspace, with the SQL object displayed.
2. Drag the data source object *DMORD* from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.
3. Double-click the SQL object to open the Column Selection dialog box.
4. Select the *STORE_CODE* and *PROD_NUM* columns by using Ctrl + click. Click the double arrows button to add them to the selected columns.
5. Select Group By from the Aggregate drop-down menu.
6. Select the *QUANTITY* and *LINEPRICE* columns by using Ctrl + click. Click the double arrows button to add them to the selected columns.
7. Select Sum from the Aggregate drop-down menu.
8. Click OK to close the Column Selection panel.
9. Right-click the workspace to the right of the SQL object, point to Add Target, and then click Existing.

The Select Synonym dialog box opens.
10. Select the *ibisamp* directory from the Look drop-down menu.
    Select the *dmstores* synonym and click the *Select* button.

11. Double-click the *dmstores* target object or right-click it and click *Target Transformations*.
    The Transformations window opens.

12. Click the Automap button to map the identically named columns.

13. Double-click *STORENAME* to add it to the transformation list, then click the calculator button.
    The Transformation Calculator opens.

14. Click the *Functions* tab, expand the *Data Source and Decoding* folder, and double-click the *DB_LOOKUP* function.
    The prototype for DB_LOOKUP appears in the Expression window and the Lookup window opens.

15. Select a synonym to use for the lookup by clicking the ellipsis button after the Lookup Synonym field.
    The Select Synonym dialog box opens.

16. Select the *DMCOMP* synonym and click *Select*.

17. From the Lookup window, click the *Add/Delete* button.

18. A lookup is similar to a join in that it uses columns from two different tables. Select *STORE_CODE* from the Source and Lookup Fields. Click the = button to add the lookup.
When you have selected the columns, the Lookup window should appear, as shown in the following image.

19. Click OK to close the Lookup window.

20. Click the ellipsis button after Return Field. The Lookup Field window opens.

21. Select STORENAME and click OK. Note that the column format is A20.
The Lookup dialog box opens, as shown in the following image.

22. Click OK.

The expression is completed in the Transformation Calculator window.
The Transformation Calculator should now appear, as shown in the following image.

23. Click OK.

The STORENAME source transformation is added to the Transformations dialog box.

24. Repeat the process starting in step 4 to add another lookup, using the parameters listed and described in the following table:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Column Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lookup Synonym</td>
<td>DMINV</td>
</tr>
<tr>
<td>Source Column</td>
<td>PROD_NUM</td>
</tr>
<tr>
<td>Lookup Column</td>
<td>PROD_NUM</td>
</tr>
<tr>
<td>Return Value</td>
<td>PRODNAME</td>
</tr>
</tbody>
</table>

25. Click OK to close the Source Transformations dialog box.

26. Save the data flow as **DMXLKUPS**.

27. On the Flow tab, in the Run group, click *Run*, and then select *Run* from the drop-down menu.
When the data flow run is complete, verify that the data was loaded properly by right-clicking the **DMSTORES** object, selecting **Expand**, and clicking the **Sample Data** tab.

**Note:**
- The first four rows have a blank store name because the DMORD table contains store codes that do not exist in DMCOMP.
- Your output may vary depending on your database.

**Record Logging**

As you extract records from a data source and load them into a data target, a variety of record logging options are available. For example, you can log records that are rejected as duplicates, or records that are rejected due to format errors or when a field is truncated. The rejected records are written to a log file which can be used as input for another process. As this technique demonstrates, you can also write reports against these log files.

This technique is divided into two parts. In the first part, you create the log file. In the second, you view the records in the log file.

In the following example, we retrieve records from DMSALE, which contains sales representative information. The data source contains duplicate records, and these are rejected and written to a log file.

The location for the rejected records file is specified in the Record Logging tab of the target object, and we use the default, which will be dmxduplg_duplic. This allows you to easily display the records in the log file, which you will do in the second part of the technique.

Refer to the sample data flow DMDUPLG for the complete example.

**Example:**  **Creating a Log File and View Records in the Log File**

In the DMC:

1. Right-click an application directory in the navigation pane, select **New**, and then click **Flow**.
   The Data Flow tab opens in the workspace, with the SQL object displayed.
2. Drag the data source object **DMSALE** from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.
3. Right-click the SQL object and click **Column Selection**.
   The Column Selection window opens.
4. Select **SALESREP** and click the arrow to add it to the Selected Columns list. Then add **PLANT** to the Selected Columns list.
5. Click the *Distinct* check box, and click *OK*.
   
   This will create one row for each unique combination of SALESREP and PLANT.

6. Save the data flow as *DMXDUPLG*.

   **Note:** When you save, you will get a message warning you that the flow is incomplete, since you have not defined a target. Save the flow anyway.

7. Right-click the workspace to the right of the SQL object, select *Add Target*, and then click *New*.

8. In the Properties window, select the adapter you used for your sample procedures, the application directory you are using for your samples, and enter *DMSREP* as the synonym.

9. Under Target Load Options, select *Reject the record* from the If the record exists drop-down menu.


11. Under Record Logging, select the *Duplicate Rejections* check box.

   Note that the rejected duplicates will be written to *dmxduplg_dupl.ftm* in the application directory you select for the target.

12. Click the *X* to close the Properties page.

13. Right-click the *DMSREP* target object and click *Target Transformations*.

   The Transformations window opens.

14. Check the *Key column* for *SALESREP* in the Target Columns list to designate a key column in the data target. Click *OK*.

15. Save the data flow.

16. On the Flow tab, in the Run group, click *Run* and select *Submit* from the drop-down menu.

17. Right-click *DMXDUPLG* in the navigation pane, select *Logs*, and then click *Last Log*. 
The log of the data flow shows that 148 rows were processed, 126 rows were accepted and inserted into the data target, and 22 rows were rejected because of duplicates. The following image shows the log.

<table>
<thead>
<tr>
<th>Datetime</th>
<th>Message Code</th>
<th>Log Messages</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012/05/22 15:40:07</td>
<td>ICM18974</td>
<td>----- Start of Log Record for ibsamp/dmxduplg -----</td>
</tr>
<tr>
<td>2012/05/22 15:40:07</td>
<td>ICM18122</td>
<td>Request - ibsamp/dmxduplg (Owner: dm) submitted.</td>
</tr>
<tr>
<td>2012/05/22 15:40:07</td>
<td>ICM18741</td>
<td>ibsamp/dmxduplg type MS SQL Server New target</td>
</tr>
<tr>
<td>2012/05/22 15:40:07</td>
<td>ICM18429</td>
<td>Issuing PREPARE</td>
</tr>
<tr>
<td>2012/05/22 15:40:14</td>
<td>ICM18054</td>
<td>Issuing CREATE and DROP TABLE for dmsep</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18743</td>
<td>Starting Load</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18745</td>
<td>Commit forced at: 148 for 148 row(s)</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18041</td>
<td>-- stats for source file</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18516</td>
<td>148 : Row(s) processed by job</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18519</td>
<td>0 : Row(s) rejected due to format error</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18372</td>
<td>-- stats for target file: ibsamp/dmxdupl</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18516</td>
<td>148 : Row(s) processed by job</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18514</td>
<td>126 : Row(s) accepted by target table</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18515</td>
<td>126 : Row(s) inserted into target table</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18517</td>
<td>0 : Row(s) updated in target table</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18518</td>
<td>0 : Row(s) deleted from target table</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18520</td>
<td>0 : Row(s) rejected due to validation errors</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18521</td>
<td>0 : Row(s) rejected due to no match</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18522</td>
<td>22 : Row(s) rejected because duplicate exist</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18808</td>
<td>0 : Row(s) rejected due to DBMS error</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18744</td>
<td>Ending Load</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18040</td>
<td>Return Code = 0</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18076</td>
<td>Request: ibsamp/dmxduplg - finished processing</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18007</td>
<td>CPU Time : 31</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18031</td>
<td>Finished</td>
</tr>
<tr>
<td>2012/05/22 15:40:15</td>
<td>ICM18072</td>
<td>Ebang run time 0:00:09</td>
</tr>
<tr>
<td>2012/05/22 15:40:17</td>
<td>ICM18975</td>
<td>----- End of Log Record for ibsamp/dmxduplg -----</td>
</tr>
</tbody>
</table>

**Example:** Viewing Records in the Log File

After the DMXDUPS data flow created in the last example has run:

1. Right-click the application directory that you specified in the Record Logging tab during the previous step, and click Refresh.
2. Right-click the *dmxduplg_dupl* object from the application directory in the navigation pane and click Sample Data.
The records that are displayed in the following image are the 22 duplicate records that were rejected and placed in the dmxduplg_dupl log file.

<table>
<thead>
<tr>
<th></th>
<th>SALESREP</th>
<th>PLANT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Patrick</td>
<td>LA</td>
</tr>
<tr>
<td>2</td>
<td>Ian</td>
<td>DAL</td>
</tr>
<tr>
<td>3</td>
<td>Gueenter</td>
<td>BOS</td>
</tr>
<tr>
<td>4</td>
<td>Gil</td>
<td>DAL</td>
</tr>
<tr>
<td>5</td>
<td>Karl-Heinz</td>
<td>BOS</td>
</tr>
<tr>
<td>6</td>
<td>Andreas</td>
<td>BOS</td>
</tr>
<tr>
<td>7</td>
<td>Christopher</td>
<td>BOS</td>
</tr>
<tr>
<td>8</td>
<td>Yafa</td>
<td>ORL</td>
</tr>
<tr>
<td>9</td>
<td>Goran</td>
<td>BOS</td>
</tr>
<tr>
<td>10</td>
<td>Heinz</td>
<td>STL</td>
</tr>
<tr>
<td>11</td>
<td>Daniel</td>
<td>STL</td>
</tr>
<tr>
<td>12</td>
<td>Thomas</td>
<td>STL</td>
</tr>
<tr>
<td>13</td>
<td>Eberhard</td>
<td>ORL</td>
</tr>
<tr>
<td>14</td>
<td>Deiter</td>
<td>BOS</td>
</tr>
<tr>
<td>15</td>
<td>Anthony</td>
<td>BOS</td>
</tr>
<tr>
<td>16</td>
<td>Wilson</td>
<td>DAL</td>
</tr>
<tr>
<td>17</td>
<td>Jacky</td>
<td>BOS</td>
</tr>
<tr>
<td>18</td>
<td>Niklas</td>
<td>STL</td>
</tr>
<tr>
<td>19</td>
<td>Deigo</td>
<td>BOS</td>
</tr>
<tr>
<td>20</td>
<td>Alex</td>
<td>STL</td>
</tr>
<tr>
<td>21</td>
<td>Bjorn</td>
<td>DAL</td>
</tr>
<tr>
<td>22</td>
<td>Per</td>
<td>STL</td>
</tr>
</tbody>
</table>

Note: Your order may differ.

Using Variables in Data Flows

There are times when you want to move data and vary something in the data flow each time that it runs. You might want to only pull records for a particular range of dates, for a particular department, or any criteria that changes from run to run. DataMigrator lets you use a variable in the data flow to accomplish this.

The most common use of variables is in a filter. When the data flow is executed, the variable must be set to an appropriate value.

A value could also be set by executing a pre-extract stored procedure. For example, if the value of the variable is read from a flat file or other external source.
The following data flow extracts data from the DMORD table. The filter contains a variable for date, &ODATE, so that each time the flow is run &ODATE can be passed a different value, and records with different dates can be extracted from the data source and loaded into the data target.

The process flow calls the data flow passing a value for the variable as a parameter.

Refer to the sample data flow DMVAR and the sample process flow DMVARPF for the complete example.

**Example: Using Variables in Data Flows**

In the DMC:

1. Right-click an application directory in the navigation pane, select *New*, and then click *Flow*. The Data Flow tab opens in the workspace, with the SQL object displayed.

2. Drag the data source object *DMORD* from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.

3. Right-click the SQL object and click *Column Selection*. The Column Selection window opens.

4. Select *PLANT*, *PROD_NUM*, and *QUANTITY* in the Available Columns list and click the arrow to move them into the Selected Columns list.

5. Select *T1.PLANT* and *T1.PRODNUM* in the Selected Columns list, and select *Group By* from the Aggregate drop-down menu.

6. Select *T1.QUANTITY* in the Selected Columns list, and select *Sum* from the Aggregate drop-down menu.

7. Click *OK* to close the Column Selection window.

8. Right-click the SQL object and click *WHERE Filter*. The Filter Calculator opens.

9. Double-click *ORDER_DATE* under Columns/Variables. This adds the column to the Conditions area.

10. Complete the condition by typing the following after *T1.ORDER_DATE* in the Conditions field:

    
    = '&ODATE'

---

The following data flow extracts data from the DMORD table. The filter contains a variable for date, &ODATE, so that each time the flow is run &ODATE can be passed a different value, and records with different dates can be extracted from the data source and loaded into the data target.

The process flow calls the data flow passing a value for the variable as a parameter.

Refer to the sample data flow DMVAR and the sample process flow DMVARPF for the complete example.

**Example: Using Variables in Data Flows**

In the DMC:

1. Right-click an application directory in the navigation pane, select *New*, and then click *Flow*. The Data Flow tab opens in the workspace, with the SQL object displayed.

2. Drag the data source object *DMORD* from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.

3. Right-click the SQL object and click *Column Selection*. The Column Selection window opens.

4. Select *PLANT*, *PROD_NUM*, and *QUANTITY* in the Available Columns list and click the arrow to move them into the Selected Columns list.

5. Select *T1.PLANT* and *T1.PRODNUM* in the Selected Columns list, and select *Group By* from the Aggregate drop-down menu.

6. Select *T1.QUANTITY* in the Selected Columns list, and select *Sum* from the Aggregate drop-down menu.

7. Click *OK* to close the Column Selection window.

8. Right-click the SQL object and click *WHERE Filter*. The Filter Calculator opens.

9. Double-click *ORDER_DATE* under Columns/Variables. This adds the column to the Conditions area.

10. Complete the condition by typing the following after *T1.ORDER_DATE* in the Conditions field:

    
    = '&ODATE'
The Filter Calculator should appear, as shown in the following image.

11. Click OK to close the Filter Calculator.

12. Right-click the workspace to the right of the SQL object, select Add Target, and then click New.

13. Right-click Target object and click Properties. Select the adapter you used for your sample procedures. Then enter the application directory you are using for your samples dmplant as the name, for example, baseapp/dmplant. Then click the small x to close the properties panel.

14. Right-click the DMPLANT target object and click Target Transformations.

The Transformations window opens.

15. Check the Key column for PROD_NUM and PLANT is automatically selected in the Target Columns list to designate them as key columns in the data target. Click OK.

16. Save the data flow as DMXVAR.

**Example: Prompting for a Variable Value**

In order to test the data flow you need to supply a value for the variable.

1. On the Flow tab, in the Run group, click Run, and then select Submit.

The prompting window opens. Enter a value to test, for example, 2006/02/14.
2. Click Run.

The data flow is tested using the supplied value. The output from the flow appears in the console log window, as shown in the following image.

![Console Log Image]

**Example:** Creating a Process Flow to Pass Parameter to a Data Flow

To create the process flow that runs both the stored procedure and the data flow:

1. Right-click an application directory in the navigation pane, select New, and then click Flow.
2. Click the Process Flow tab.
3. Drag the data flow, DMXVAR, from the navigation pane into the process flow workspace, to the right of the Start object.
4. Right-click the DMXVAR object and click Properties. Click the ellipsis button for Parameters.
   - The Parameters editor opens.
5. Click the Add Parameter button.
   - The Drill Down Parameter editor window opens.
6. Select ODATE from the Parameter Name drop-down menu. Enter a date value, such as 2005/03/06.

![Drill Down Parameter Editor](image)

7. Click OK.

The parameter appears in the Parameter Editor.

8. Click OK.

The parameter appears in the property pane, as shown in the following image.

![Property Pane](image)

9. Save the flow as DMVARPFX. You can submit the flow now.
Creating a Data Flow Using a Target Based on a Predefined XML Schema

When creating an XML file, the format of the XML file is usually predefined by an xml schema definition or .xsd file. DataMigrator lets you create a synonym from a schema definition and then map to that target in the same manner as a relational database table. An existing JSON document can also be used as a prototype for a synonym, which can then be used as a target.

Single-segment and multiple-segment schemas are supported.

This technique is divided into two parts. In the first part, you create a synonym for the XML file. In the second, you use that synonym as the target of a flow.

In the example shown in the following image, the data source is DMORD. The target is described by the schema dmxord.xsd.

Refer to the sample data flow DMORDXML, for the complete example.

This is the structure of the schema (dmxord.xsd), which will be used to create a synonym to demonstrate this technique.

Note: To write to an XML target, the XML adapter must be configured on your server.

Example: Creating a Synonym for an XML File

An XSD (XML Schema Definition) file contains a description of an XML file. This example shows how to create a synonym for the dmxord.xsd, which is installed with your server.
Note: For more information about creating synonyms for XML, please see *Using the Adapter for XML* in the *Adapter Administration* manual.

1. In the DMC, right-click the *baseapp* application directory in the navigation pane, point to *New*, and then click *Synonym (Create or Update)*.

   The Select adapter to configure or Select connection to create synonym dialog box opens.

2. Expand the *Available* folder and then expand the *XML Based* folder.

3. Select *XML* and click *Configure*.

4. Click *OK*.

   The Select Document Instance for XML Synonym dialog box opens.

5. Enter `ibisamp/dmxord.xml` as the location where the XML file is created, as shown in the following image.

   ![Select Document Instance for XML Synonym](image)

6. Click *Next*.

   **Note:** The Create Synonym based on Document button allows you to create a synonym using an XML document alone. This example is for an XML target, so the xml document does not exist yet. This option is used for an XML source where no XML Schema Definition is available.

   The Select Data Definition for XML Synonym dialog box opens.

7. Click the ellipsis to open Data Definition Selection. Expand the *Applications* folder and *ibisamp*. Select `dmxord.xsd` and click *OK*. The name `ibisamp/dmxord.xsd` appears as the location of the XML Schema Definition that describes the xml target.

8. Click *Next*.

   The Create Synonym for XML dialog box opens.

9. Click *Create Synonym*.
10. Click Close in the confirmation window.

**Example:** Creating an XML File

For a complete example, see the ibisamp/dmordxml flow.

In the DMC:

1. Right-click an application directory in the navigation pane, select New, and then click Direct Load Flow. The Direct Load tab opens.
2. Drag the DMORD object from the ibisamp application directory in the navigation pane into the workspace, to the left of the line.
3. Drag the DMXORD object from the baseapp application directory in the navigation pane into the workspace, to the right of the line.
4. Double-click the DMXORD object.
   The Transformations window opens.
5. Click the Automap button.
6. Select the first row in Expressions panel, ORDER_NUM, and click the Adjust Expressions button. In the XML target, the ORDER_NUM column is described as Decimal. This expression converts the character-valued column to a number.
7. Click OK.
8. On the Flow tab, in the Run group, click Run and select Submit. A Save Changes dialog box opens. Click Yes.
9. When the Save Procedure as dialog box opens, enter dmxord2xml in the File Name field and click OK. Wait for the request to complete.
10. On Windows, you can view c:\ibi\apps\baseapp\dmxord.xml in Internet Explorer or in an XML editor. In the example shown in the following image, note that the hierarchical structure of the source has been preserved. There is one ORDER_INFO element which contains multiple PROD_INFO elements.

```xml
<?xml version="1.0" encoding="UTF-8" ?>
<ORDER xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation="c:\ibi\apps\ibisamp\dmxord.xsd">
  <ORDER_INFO>
    <ORDER_NUM>93674</ORDER_NUM>
    <ORDER_DATE>2003-02-15</ORDER_DATE>
    <STORE_CODE>999999</STORE_CODE>
    <PLANT>BOS</PLANT>
    <PROD_INFO>
      <PROD_NUM>1004</PROD_NUM>
      <QUANTITY>1</QUANTITY>
      <LINEPRICE>166.47</LINEPRICE>
    </PROD_INFO>
    <PROD_INFO>
      <PROD_NUM>1004</PROD_NUM>
      <QUANTITY>1</QUANTITY>
      <LINEPRICE>166.47</LINEPRICE>
    </PROD_INFO>
    <PROD_INFO>
      <PROD_NUM>1004</PROD_NUM>
      <QUANTITY>1</QUANTITY>
      <LINEPRICE>166.47</LINEPRICE>
    </PROD_INFO>
  </ORDER_INFO>
</ORDER>
```

**Loading a Target Table Using the File Listener**

You can use DataMigrator to build a flow that automatically processes source files as they appear in a designated directory. This can be useful if, for example, source files are transferred by FTP from various locations throughout the day and need to be processed immediately.
Before you create and run the data flows discussed in this section, you must create the sample data and synonyms. For more information, see *How to Create Sample Procedures and Data for Iteration* on page 82. This will create source and target synonyms and four source data files used in the example.

To implement and run a real time flow:

1. Create an application directory to receive files.
2. Create a synonym with File Collection Parameters.
3. Create a Data Flow that uses the File Listener.
4. Run the flow and drop the files.

This flat file synonym corresponds to the layout of the source files that will appear in the designated directory. The flat file source is loaded to the target of your choice in real time. Refer to the sample flow DMFLORD in the ibisamp directory for the completed example.

**Example: Create Listen and Archive Directories**

For DataMigrator to automatically process files when they arrive in a directory, you need to designate a directory for that purpose. It could be any directory that your server can access. For this example, we will create an application directory called listen.

In addition, if you want to keep a copy of the files after they are processed you need to designate a directory for that purpose too. For this example, we will create an application directory called archive.

1. Right-click *Application Directories*, select *New*, and then click *Application Directory*.
   
   The Create New Application dialog box opens.

2. For Application name, type *listen*.
   
   De-select the Add directory to APPPATH check box.

   Click OK.

3. Repeat steps 1 and 2 to create a directory called archive.
**Example: Create a Synonym With File Collection Parameters**

Next you need a synonym with File Collection Parameters to control how the file listener process files. We will start with a synonym that describes just the record layout of the source files and add the parameters.

1. On the Home tab, in the Filter group, click *Synonyms* to ensure that synonyms are displayed. Then, in the navigation pane, expand the *Application Directories* folder and then expand the *ibisamp* folder.

2. In the ibisamp folder, double-click *dmordflx* to open the synonym. Note that the column named *SOURCEFILE* has a special ALIAS of *INSTANCE*. This field does not exist in the source file itself, but it will contain the name of the source file when it is read.

3. Expand the *Data Allocation and other related groups*. 
4. Fill in the synonym parameters, as shown in the following image.

This creates a synonym that listens on the listen directory for files with names that start with dmord and an extension of data polling every 5 seconds until 60 seconds have passed without a new file. As files are processed they are archived to a directory archive. Files are processed as they arrive until 5 files or 50 records are processed.

5. Click Save As in the Quick Access Toolbar to save the synonym in the application directory of your choice.
**Example:** **Create a Flow with a File Listener Source**

A flow with a File Listener source can be either a Data Flow or a Direct Load Flow. For this example, use a Direct Load Flow.

In the DMC:

1. Right-click an application directory, point to New, and then click Direct Load Flow. The Direct Load Flow opens.
2. From the browser window, drag the source object DMORDFL that you created from its application directory to the left of the center line in the workspace.
3. Drag the target object DMORDF from the ibisamp application directory to the right of the target line in the workspace.
4. Right-click the target object DMORDF and click Properties.
   The Properties dialog box opens.
5. For Prior to Load Option, select Truncate table, if available. If not, select Delete all rows from table or Drop table. Click X to close the Properties page.
6. Right-click the target object and click Target Transformations.
   The Transformations window opens.
   
   Click the Automap button and click OK.
7. Save the flow as DMXFLORD.

**Example:** **Create a Synonym for a File Listener with a Trigger**

You can use DataMigrator to build a flow that automatically processes source files as soon as a specific Trigger file is detected in an application directory. A trigger is typically used when processing needs to be delayed. For example, when using FTP to copy files, if you wait until this secondary file appears while a large file is being copied into a directory, the trigger file indicates when processing can begin.

The following example shows how to create such synonym in a baseapp application using an Adapter for Delimited Flat File. You can also create such synonym using an Adapter for a Flat File or in any other application directory.

1. Create Listen and Archive directories as described in the Create Listen and Archive Directories example.
2. In the DMC, right-click the baseapp application directory, point to New, and then click Synonym (Create or Update).
The Select adapter to configure or Select connection to create synonym dialog box opens.

3. Select the connection from the Delimited Flat File node and click OK.

   The Select Data File dialog box opens,

4. Disable the Create Multiple Synonyms checkbox.

5. Click the ellipses in the Data File field and select the data file from the Listen directory, as shown in the following image.

![Image showing the Select Data File dialog box]

DataMigrator creates the Master File for the File Listener synonym for the synonym later, based on this file.

6. The Select Processing Options dialog box opens, as shown in the following image.

   Enter the following values:

   a. Select Collection of files from the drop-down menu on the Select Processing Options screen.

   b. Select Trigger from the Data File Pickup Strategy drop-down menu.

   c. Enter the extension in the Trigger File Extension field. For example, if the file name that you listen to is myfile.csv, the trigger file name is myfile.csv.trg. The value for the Trigger File Extension will be trg.

   d. Select Archive from the Processed Data Files Discard Strategy drop-down menu.

   e. Enter the name of your Archive directory in the File Listener Archiving Directory field.
7. Click Next.

The Define Collection dialog box opens.

8. Enter the following values:
   
   a. In the Directory or Application field, enter the name of your Listen directory.
   
   b. Enter a name or mask in the Name field.
   
   c. Enter the extension of a file, or a mask, in the Extension field.
The options should appear as shown in the following image.

![Image of Select Synonym Candidates dialog box]

9. Click Next.

The Select Synonym Candidates dialog box opens.

10. Enter all attribute values that are needed for the Delimited Flat File.

11. Click Next.

12. Click Create Synonym.

**Example:** Running a Flow With a File Listener Source

In this example, we will start the flow and then copy source files to the listen directory.

In the DMC:

1. Right-click *DMXFLORD* and click *Submit*.

2. Copy the following files, one at a time, from the ibisamp application directory to the listen application directory:
   
   - dmordbos.dat
   - dmorddal.dat
   - dmordorl.dat
   - dmorstl.da

   **Tip:** For a Windows based server, you can use the Edit and Copy/Paste functions.

   **Note:** As the files are processed, you should see them disappear from the listen directory. This could take up to the timeout interval specified in the synonym.
3. Return to the DMC. The source synonym has a timeout of 60 seconds, so you will see the Processing dialogue until the time has elapsed.

4. When the flow completes, on the Flow tab, in the Reports group, click View Last Log. You should see a report, like the following image, that shows all total number of rows processed.

```
Log Messages
(IXM18974) ------ Start of Log Record for ibisamp/dmfld -----------------------
(IXM18122) Request - dmfld (Owner: dm) submitted.
(IXM18742) ibisamp/dmordf type MS SQL Server Existing target
(IXM18743) Starting Load
(IXM18745) Commit forced at: 47 for 47 row(s)
(IXM18041) -- stats for source file
(IXM18516) 47 : Row(s) processed by job
(IXM18519) 0 : Row(s) rejected due to format error
(IXM18372) -- stats for target file: ibisamp/dmordf
(IXM18516) 47 : Row(s) processed by job
(IXM18514) 47 : Row(s) accepted by target table
(IXM18515) 47 : Row(s) inserted into target table
(IXM18517) 0 : Row(s) updated in target table
(IXM18518) 0 : Row(s) deleted from target table
(IXM18520) 0 : Row(s) rejected due to validation errors
(IXM18521) 0 : Row(s) rejected due to no match
(IXM18522) 0 : Row(s) rejected because duplicate exist
(IXM18808) 0 : Row(s) rejected due to DBMS error
(IXM18744) Ending Load
(IXM18040) Return Code = 0
(IXM18076) Request: dmfld - finished processing
(IXM18007) CPU Time : 60
(IXM18975) ------ End of Log Record for ibisamp/dmfld -----}
```
With the Discard option set to Archive, after each file is processed, it is copied to the archive directory, as shown in the following image.

![Archive directory screenshot]

These files can be deleted, as they are no longer needed.

**Sending a Target File to an FTP or SFTP Server**

You can use DataMigrator to build a flow that automatically sends the target file to a directory on a FTP server. This can be useful if, for example, a target file has to be created on a remote FTP location. This feature is available for the synonyms of a Flat File and a Delimited Flat File. The synonym can be used in a Direct Load flow, as well as in a data flow.

The examples in this section use a synonym for a Flat File in a Direct Load flow.

Before you create and run the data flow discussed in this section, you must configure the Adapter for a Flat File. After the adapter is configured, you need to create the sample data and synonyms for DataMigrator - General. This will create the source and target synonyms that are used in the following examples.

To implement and run a flow that sends a target to a FTP server, you must:

1. Create a connection to a FTP location.
2. Create a synonym with FTP attributes.
3. Create a Direct Load flow that uses this synonym as a target.
Example: Creating a Connection to a FTP or SFTP Location

Note: If you are using SFTP, you must first set up an SSH connection without a password. For more information, see Using Secure Shell FTP (SFTP) on page 99.

To configure a connection to a FTP server:

1. Expand the Adapters folder, followed by the Configured folder.
2. Right-click the Flat File or Delimited Flat File folder and click Add Connection.
3. Enter the Connection Parameters, as shown in the following image.

**Host Name**

Name or IP address of the FTP or SFTP server.

**Initial Directory**

Directory on the FTP or SFTP server to connect to. This value is optional. If it is not specified, either of the following will be used:

- The default initial directory.
- The directory name entered on a synonym for a DATASET parameter value, entered with a file name, such as mydir/myfile.txt.

**Note:** On z/OS, to access files in zFS, enter a directory name. To access sequential datasets, enter two periods (..) or a prefix such as 'USERID.'.

**User**

User ID to connect to the FTP or SFTP server.

**Password**

Password to connect to the FTP or SFTP server.

**Protocol**

The file protocol that is being used, either FTP or SFTP (SSH FTP).
Example: Creating a Synonym with FTP Attributes

1. Open the synonym *dminva* in the Synonym Editor, as seen in the following image.

![Synonym Editor Image]

2. Enter the file name in the DATASET parameter.
3. Expand the *DB Log Parameters* group, as shown in the following image.

![DB Log Parameters Image]
4. Enter the Connection value that was created for the Flat File adapter. Leave the \textit{DATA\_ORIGIN} parameter blank.

5. Leave the \textit{DIRECTORY}, \textit{NAME}, and \textit{EXTENSION} parameters blank since these values have already been specified in the \textit{DATASET} parameter.

6. Save the synonym with the new name \textit{dminva\_ftp}.

7. Close the synonym.

Example: \textbf{Creating a Flow that Sends a File to an FTP Server}

1. Right-click an application directory, select \textit{New}, and then click \textit{Direct Load Flow}.
   
   The Direct Load flow opens.

2. From the browser window, drag the \textit{dminv} source object from the application directory to the left of the center line of the workspace.

3. Drag the \textit{dminv\_ftp} target object that was created from the ibisamp application directory to the right of the target line in the workspace, as shown in the following image.

4. Right-click the target object and click \textit{Target Transformations}.
   
   The Transformations window opens.

5. Click the \textit{Automap} button.

6. Save the flow as \textit{dminvftp}.

\textbf{Loading a Star Schema}

A database in a data warehouse is often organized into a star schema, consisting of a central fact table with the data to be analyzed and multiple dimension tables that describes the data.

Each dimension table has a single surrogate key, an arbitrary unique identifier for the row. The fact table has multiple keys, each joining to a different dimension table.
Before you create and run the data flows discussed in this section, in addition to creating the source tables, you must also create the target tables. See *How to Create Sample Procedures and Data for Star Schema* on page 81.

This example has detailed instructions for loading one of the four dimension tables. Refer to the sample flows, DLOADCUST, DSLOADPROD, DSLOADSALE, and DSLOADTIME, for other examples of loading a dimension table. Refer to DSLOADFACT for an example of loading a fact table and DSFLOWP for the complete example of loading a star schema.

**Example: Loading a Dimension Table**

This example has instructions for loading the customer dimension table.

1. In the DMC, right-click an application directory in the navigation pane, select *New*, and then click *Flow*. The data flow opens in the right hand pane, with the SQL object displayed.

2. Drag the data source object DMCOMP from the ibisamp application directory in the navigation pane into the workspace, located to the left of the SQL object.

3. Right-click the SQL object and click *Column Selection*.

   The Column Selection window opens.

4. Right-click any column name and *Select All*. Then click the arrow to add them to the selected columns and click *OK*.

5. Drag the target object DSDIMCUST from the ibisamp application directory into the workspace to the right of the SQL object.

6. Right-click the target object DSDIMCUST and click *Properties*.

   The Properties page opens.

   From the *Load Type* drop-down menu, select *Slowly Changing Dimensions*. Click *X* to close the Properties pane.

7. Right-click the target object DSDIMCUST and click *Target Transformations*.

   The Transformations window opens.

8. Click the *Automap* button.
The identically named source and target column names are added to the Expressions window. The Transformations window should now look like the following image.

![Transformations Window Image]

**Note:** The columns CKEY and ACTIVE are not mapped because these columns are handled automatically by the DataMigrator Slowly Changing Dimension processing. CKEY is the surrogate key, which automatically starts with a value of 1 and increases by one, for each row added. The Active flag is set to 1 for currently active rows.

Click OK to close the Transformations window.

9. Click Save from the Quick Access Toolbar. Enter dsxloadcust as the file name.

**Example:**  
**Loading a Fact Table**

A fact table load requires looking up the keys for each row in the corresponding dimension tables, and obtaining a surrogate key so that the fact table can be joined to each of the dimensions for subsequent reporting.

To create the fact table follow these steps.

1. In the DMC, right-click an application directory in the navigation pane, select New, and then click Flow. The data flow opens in the right hand pane, with the SQL object displayed.
2. Drag the data source object DMORD from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.

3. Drag the data source object DMSALE from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.

A JOIN object is automatically added connected to DMORD[MS] and DMSALE.

   **Note:** If not, delete the object. On the Home tab, in the Tools group, click Options. In the Options dialog box, click Data Flow and select the Add Join Object if needed check box. Click OK. Then add dmsale again.

4. Drag the data source object DMINV from the ibisamp application directory in the navigation pane into the workspace, to the left of the SQL object.

A second JOIN object is automatically added connected to the first JOIN object and DMINV.

5. Right-click the SQL object and click Column Selection.

   The Column Selection window opens.

6. Select the following columns, and click the arrow to add them to the Selected Columns list:

   - ORDER_DATE from DMORD(T1)
   - SHIP_DATE from DMORD(T1)
   - PROD_NUM from DMINV(T3)
   - STORE_CODE from DMORD(T1)
   - EMPID from DMSALE(T2)
   - LINEPRICE from DMORD(T1)
   - QUANTITY from DMORD(T1)

7. Click the Insert Columns button.

   The SQL Calculator opens.

8. For Alias, type LINCOGS and for calculation, expand DMPROD. Double-click QUANTITY, type*, then expand DMINV and INVINFO. Select COST so that the Expression window shows T1.QUANTITY * T3.COST.

9. Click OK to close the SQL calculator.
The Column Selection window should appear as in the following image.

10. Click OK to close the Column Selection window.
11. Drag the target object DSFACT from the ibisamp application directory to the workspace to the right of the SQL object.
12. Right-click DSFACT and click Target Transformations.
   The Transformations window opens.
13. Double-click ODKEY, adding it to the Expressions window. With the line selected, click the Edit Transform button.
   The Transformation Calculator window opens.
14. Click the Functions tab, expand the Data Source and Decoding folder, and double-click DB_LOOKUP.
   The Lookup function assist opens.
15. Click the ellipsis button after Lookup Synonym. Select the synonym DSDIMTIME from the ibisamp directory and click Select.
16. Under Lookup fields, click the Add/Delete button. Select ORDER_DATE from the left hand side and TDATE on the right hand side. Click OK to close the dialog box.
17. Click the ellipsis button after Return Field.
   The Lookup Field dialog box opens.
18. Double-click TKEY to select it.
   The Lookup window opens.
   Click OK to close the window.
19. Repeat steps 12-17 above for SDKEY, using SHIP_DATE as the source field, TDATE as the lookup field, and TKEY as the return field.
20. Similar transformations are required for each of the key columns.

Use the following tables:

CKEY - DSDIMCUST PKEY - DSDMINPROD SKEY - DSDIMSALE

For the rest of the key columns, the Source Field and Lookup Field are pre-selected, since the names are the same in the Fact and Dimension tables. In addition, transformations are required for the remaining columns.

When you have created the remaining transformations, the Transformations window should appear, as shown in the following image.

![Transformations Window]

21. Click OK to close the window.

22. Click Save as from the Quick Access Toolbar. Enter dsloadfact as the file name.

**Example:** Loading a Star Schema Using a Parallel Group

In order to load the fact table in this example, all the dimension tables have to be loaded first, so that the surrogate keys are available.

1. In the DMC, right-click an application directory in the navigation pane, select New, and then click Flow. The data flow opens in the right pane.

2. Click the Process Flow tab. The view switches to the process flow with a Start icon.
3. On the Flow tab, in the Insert group, click *Parallel Group* and drag it into the work area to the right of the Start icon.

A box appears on the screen. This is an empty parallel group.

4. Right-click the *Start* icon and drag a line towards the parallel group box.

Release the line when the arrow is touching the box.

5. From the navigation pane, drag the *DSXLOADCUST* flow *inside* the parallel group box. If you did not create this data flow, drag the *DSLOADCUST* flow from the ibisamp directory.

6. From the ibisamp directory, drag the *DSLOADPROD*, *DSLOADSALE*, and *DSLOADTIME* flows into the parallel group.

7. The *DSLOADTIME* flow requires a parameter of the first date to load. Double-click the flow to open the properties for the flow. For Parameters, enter `STARTDATE=20040101`.

8. On the Flow tab, in the Insert group, click *Wait* and drag it into the work area to the right of the parallel group.

9. Right-click the parallel group, drag a line to the *Wait* icon, and release.

10. Drag the *DSXLOADFACT* flow to the right of the *Wait* icon. If you did not create this data flow, drag the *DSLOADCUST* flow from the ibisamp directory.

11. Right-click the *Wait* icon, drag the line to the *DSLOADFACT* flow, and release it.

When you have finished, your process flow should appear as shown in the following image.
12. Click Save as from the Quick Access Toolbar. Enter dsxflowp as the file name.
13. On the Flow tab, in the Run group, click Run and select Submit to run the flow.
14. When the flow completes, on the Flow tab, in the Reports group, click View Last Log.

The process flow log opens, as shown in the following image.

The log shows the four dimension data flows run and the fact table load, which runs when the data flow runs are complete. The lines in blue indicate links to the detail logs for the individual flows.
15. Click the fact table log, which is the last blue line shown. The detail log opens, as shown in the following image.

<table>
<thead>
<tr>
<th>Log Messages</th>
</tr>
</thead>
<tbody>
<tr>
<td>(CM18974) ----- Start of Log Record for ibisamp/dloadfact -----</td>
</tr>
<tr>
<td>(CM18972) Request - dloadfact (Owner: dmuser) submitted.</td>
</tr>
<tr>
<td>(CM18742) dsfact type MS SQL Server Existing target</td>
</tr>
<tr>
<td>(CM18429) Issuing PREPARE</td>
</tr>
<tr>
<td>(CM18440) Request will process data via NON-Pass Through (NON-APT)</td>
</tr>
<tr>
<td>(CM18451) HOLD file will be created for output file named: SQLIN.</td>
</tr>
<tr>
<td>0 NUMBER OF RECORDS IN TABLE = 594  LINES = 594</td>
</tr>
</tbody>
</table>

(1CM18743) Starting Load
0 WARNING, ON MATCH INCLUDE INPUTS DUPLICATE SEGMENTS
(1CM18745) Commit forced at: 594 for 594 row(s)
(1CM18041) -- stats for source file
(1CM18516) 594 : Row(s) processed by job
(1CM18519) 0 : Row(s) rejected due to format error
(1CM18372) -- stats for target file: dsfact
(1CM18516) 594 : Row(s) processed by job
(1CM18514) 594 : Row(s) accepted by target table
(1CM18515) 594 : Row(s) inserted into target table
(1CM18517) 0 : Row(s) updated in target table
(1CM18518) 0 : Row(s) deleted from target table
(1CM18520) 0 : Row(s) rejected due to validation errors
(1CM18521) 0 : Row(s) rejected due to no match
(1CM18522) 0 : Row(s) rejected because duplicate exist
(1CM18608) 0 : Row(s) rejected due to DBMS error
(1CM18744) Ending Load
(1CM18040) Return Code = 0
(1CM18076) Request: dloadfact - finished processing
(1CM18007) CPU Time : 802
(1CM18975) ----- End of Log Record for ibisamp/dloadfact ----- |

Iterating Through Multiple Tables

It is often the case that an organization has multiple tables with the same structure, that is column names and data types, but with different names, in different databases or schemas or on different servers.

Now a single synonym can be used to describe a collection of tables with the same structure because the TABLENAME and CONNECTION parameters in the access file can be specified as a variable.
Before you create and run the data flows discussed in this section, in addition to the source data, you must also create the multiple source tables. See *How to Create Sample Procedures and Data for Iteration* on page 82.

This will create five identically structured tables named dmordbos, dmorddal, dmordla, dmordorl, and dmordstl.

**Example:** **Create a Synonym With Parameterized Table Names**

This example has instructions for creating a synonym that can be used to read or write multiple tables. You will create two variables, TTABLE and TCON, that will be used for the table name and connection name.

Refer to the sample synonym, dmortmpl, for the complete example that also has titles and descriptions for each of the columns.

1. In the DMC, right-click an application directory in the navigation pane, select New, and then click Synonym.
   
   The Select adapter to configure panel opens.

2. Open the folder for the adapter where you created the sample procedures and select the first connection. Click OK.

   The Select Synonym Candidates Step 1 of 2 panel opens.

3. Click Next.

   The Create Synonym Step 2 of 2 panel opens.
4. Select the check box to the right of dmordbos, double-click the table name, and change it to \textit{dmortmplx}, as shown in the following image.

![Create Synonym for MS SQL Server (CON01) Step 2 of 2](image)

This will create a synonym that you will use as a template to access any of the six tables.

5. Click \textit{Create Synonym}.

6. Click the \textit{Open First Synonym} button.

   The dmortmplx synonym opens.

7. If there is Business View window, click the small x to close it.

8. Right-click the Variables folder, select \textit{Insert and Variable}. This adds a variable called \textit{VARIABLE1}.

9. Double-click the name and change it to \textit{TTABLE}. This creates a global variable in the synonym.

10. For DEFAULT, enter the value \textit{dmordbos}.

11. Right-click the Variables folder, select \textit{Insert and Variable}.

12. Double-click the name and change it to TCON. For the DEFAULT, enter the name of your first database connection, which by the server default would be CON01.
13. Click in the left pane and the names of the two variables you created are shown.

14. Click the segment name `DMORTMPLX`. For `TABLENAME`, enter the value `&&TTABLE` and for `CONNECTION`, enter the value `&&TCON`.

For `DEFAULT`, enter the name of your connection, which by default is CON01.

The synonym is now set up for use with multiple tables.
15. Click Save from the Quick Access Toolbar to save the synonym.

**Example: View Parameter Table**

This example has instructions for viewing the parameter table used in the other examples here. The parameter table used is a delimited flat file, but any data source that can be described by a synonym could be used. The table has three columns: a connection name, a table name, and a plant name.

The parameter table is automatically created when you run Create Sample Procedures and Data for the iterator examples. The connection name is the name of the first connection on your server for the selected database. In this example, they are all the same, but different connection names could also be used.

1. In the DMC, expand the *ibisamp* directory in the navigation pane.
2. Right-click the synonym *dmplnts* and click Sample Data.

   **Tip:** If you do not see synonyms when you expand the ibisamp directory, select Synonyms from the Filter group on the Home tab.

The Sample Data Report opens and should look like the following image.

![Parameter Table](image)

**Example: Create a Data Flow Using a Parameterized Synonym**

This example has instructions for creating a data flow that uses a synonym as a source that can be used to read multiple identically formatted tables.

Refer to the sample flow dmorsum, for the complete example.

1. In the DMC, right-click an application directory in the navigation pane, select New, and then click Flow. A new flow opens.
2. Drag a data source object into the data flow. If you created the synonym, dmortmplx, you can use it. Otherwise, from the ibisamp application directory, drag the synonym *dmortmpl*. 
3. Right-click the SQL object and click *Column Selection*.
   
   The Column Selection window opens.

4. Select each column in the Available Columns list, and click the arrow to move them into the Selected Columns list. Click *OK*.

   **Tip:** To select all the columns, right-click one of them and click *Select All*.

5. Drag the target object *dmorsum* from the ibisamp directory into the workspace, to the right of the SQL object.

6. Right-click the *dmorsum* target object and click *Target Transformations*.
   
   The Transformations window opens.

7. Click the *Automap* button.

8. Under target columns, double-click *PLANT* to add it to the list of expressions below.

9. For the PLANT expression, enter ‘&LPLANT’ with quotes as shown. Click *OK*.

10. Click the *Process Flow* tab.

11. Right-click the arrow between Start and Data flow and click *Delete*.

12. On the Flow tab, in the Insert group, click *Set Variables* and drag it onto the workflow between the Start and Data Flow icons. Then, right-click the *Start* icon and drag it to Set Variables and release it. Finally, right-click *Set Variables* and drag the arrow to Data Flow and release.
   
   The flow will look like the following image.

   ![Flow Diagram](image)

13. Right-click *Set Variables* and click *Properties*. In the Value area for List, click the ellipsis button.
   
   The Global Parameter Editor window opens.

14. Click the *Add Parameter* button.

   The Global Variables calculator opens.

15. Type over the default name VAR1 with *TCON*. In the expression entry area, type &LCON and then click *OK*. 

   Iterating Through Multiple Tables
16. Repeat to set the global variable, TTABLE to value $LTABLE. When you are done, the Global Parameter Editor window should look like the following image.

Click OK.

17. Click Save from the Quick Access Toolbar to save the flow. Enter dmorsumx as the flow name.

Example: Create a Process Flow to Use Parameterized Data Flow

This example has instructions for creating a process flow that runs a data flow repeatedly using a source a collection of identically formatted tables.

Refer to the sample flow dmorsump for the complete example.

1. In the DMC, right-click an application directory in the navigation pane, select New, and then click Flow.
   A new flow opens.
2. Click the Process Flow tab.
3. From the directory where you saved dmorsumx, drag the flow into the workflow area.
4. Right-click the Start icon, drag the arrow to dmorsumx and release.
5. Right-click dmorsumx and click Properties.
6. Check the check box labeled Get Parameters using Synonym.
7. Click the ellipsis button in the File Name entry area. In the ibisamp directory, select *dmplnts*.

The properties for the flow should look like the following image.

8. Click Save on the Quick Access Toolbar to save the flow. Enter *dmorsumpx* as the flow name.

9. To create the target table, in the ibisamp directory located in the navigation pane, right-click the table named *dmorsum*, select *Data Management*, and then click *Recreate DBMS table*.

10. On the Flow tab, in the Run group, click *Run* and select *Submit*.

When the flow completes, you should see a message in the console log:

```
DM: (ICM18763) Request ibisamp/dmorsumpx complete
```

**Example:** View Logs for a Parameterized Data Flow

When a process flow calls a data flow using the iterator, multiple logs are generated. This example shows how to view the logs.

1. With the dmorsumpx process flow open, on the Flow tab, in the Reports group, click *View Last Log*.

   The process flow log opens.

2. Click the blue line with the Job ID.
The iterator log opens. Note that the Parameters line show the values of each parameter for each run.

3. Click one of the blue lines with Job IDs. The log for an individual run of the data flow opens.
Using Subqueries

A subquery or subselect can be used in several different places in an SQL SELECT statement.

Example: Using a Subquery in the Select List

This example shows how subqueries can be used in the list of columns in a select statement to include values from additional tables without using a JOIN. The product name PRODNAME is retrieved from the inventory table dminv, and the store name STORENAME is retrieved from the table of the company dmcomp.

For the complete example, run the dmsubsel flow that is located in the ibisamp directory (ibisamp/dmsubsel).

1. In the DMC, right-click on an application directory in the navigation pane, point to New, and then click Flow.
2. From the ibisamp directory drag dmord as a source to the left of the SQL object.
3. Double-click the SQL object to open the Column Selection dialog box.
4. Ctrl + click STORE_CODE, PROD_NUM, QUANTITY, and LINEPRICE and click the Move Columns Right button to add them to the Selected Columns list.
5. Click the Insert Columns button.
   The SQL Calculator opens.
6. Enter PRODNAME for alias type.
7. Select the Functions tab and expand SQL Operators. Double-click the function that starts with (SELECT lookup_result).
8. Click the ellipsis for LOOKUP_MFD and look in ibisamp for the table dminv. Click Select.
9. Enter T2 for the lookup_mfd_tag type.
10. Click the ellipsis for lookup_result and click PRODNAME.
11. For lookup_aggregation, use the default MIN().
12. Click the ellipsis for lookup_condition. The condition calculator opens. Enter T2.PROD_NUM = T1.PROD_NUM.
13. Click OK.
The Function Assist dialog box opens, and should resemble the following image.

14. Click Ok to close the Function Assist dialog box and click OK to close the SQL Calculator.

15. Repeat steps 5 to 14 using the following values:

- For LOOKUP_MFD, enter DMCOMP.
- For lookup_result, enter T3.STORENAME.
- For lookup_aggregation, enter MIN().
- For lookup_mfd_tag, enter T3.
- For lookup_condition, enter T1.STORE_CODE = T3.STORE_CODE.

16. Click OK to close the Column Selection dialog.

17. From the navigation pane in the ibisamp directory, drag the synonym dmstore to the right of the SQL object.

18. Right-click dmstores and click Properties.

The properties page opens.

19. For the Prior to load option, select Truncate table if available, or Delete all rows if not.

Click the X to close the page.

The Transformations window opens.

21. Click the Automap button. Next, click OK to close the Transformations window.

22. On the Flow tab, in the Run group, click Run and select Submit from the drop-down menu.

23. When the flow completes, right-click dmstores, select Operations, and then click sample data. Note that the STORENAME and PRODUCTNAME columns are filled in with values.

Loading Images Into a Database Table

DataMigrator has the capability to load image files into a database table by specifying their location in the input data.

The first step is to create a table to store the images. The table to load should have a column with a data type that can hold binary data. In the current release of MS SQL Server, that would be varbinary(max). In previous releases, it is image. In the synonym, the USAGE and ACTUAL fields for the column are set to BLOB (Binary Large OBject, as shown in the following image.

The following example outlines the remaining steps needed to load images into a database table.
Example: Loading Images Into a Database Table

The first step is to map an application directory. The location of the files containing the images is specified in the usual appdir/filename format. To map the application directory:

1. Right-click Application Directories, select New, and then click Application Directory. The Create New Application dialog box opens.
2. From the Application Type drop-down menu, select Application Mapping to Disk.
3. In the Application Name box, enter graph.
4. In the Physical Location box, enter c:\ibi\srv77\home\graph.
5. From the Profile drop-down menu, select the desired profile. The Create New Application dialog box should look like the following image.

6. Click OK.

Next, create a source file and synonym. The source data needs a list of the names of the image files. From the DMC:

7. Right-click baseapp.
8. Select New and then click File.
9. Enter names as a header line, followed by the names of some of the .gif files (without the extension).
10. Click the **Save** button.

11. Enter *tiles* as the name and click **OK**.

   The following image shows the result.

   ![Image showing the result](image.png)

   Then, create a synonym for the file.

12. Right-click **Application Directories**, select **New**, and then click **Synonym**.

13. Select `<local>` for the Under Delimited Flat File. Click **OK**.

14. Clear the Create multiple synonyms checkbox and select **Server Side** for File Location.

15. For Data File, enter baseapp/tiles.txt.

16. Click **Next** twice.

17. For Header, select **Yes**.

18. Click **Create Synonym**

19. Click **Close**.

   Next, create a data flow that loads the images into a database table. The source is the delimited flat file tiles that contains the names of the image files. The target is the existing database table textures.

20. Right-click **baseapp**. Select **New** and then click **Flow**.

21. Right-click the left side of the workspace, select **Add Source**, and then click **baseapp/tiles**.

22. Right-click the right side of the workspace, select **Add Target**, then **Existing**, and then click **baseapp/textures**.

23. Double-click the SQL object to open it.

24. Double-click **Name** to add it to selected columns. Enter **FNAME** for SQL Alias.

25. Double-click **Name** and click the calculator button.

26. For Alias, enter **PICTURE**. For Expression, enter `graph/' || T1.NAME || '.gif'`. 
27. Click the Test SQL button to see the values for the columns.

28. Close the Test SQL dialog box and click OK to close the SQL object.

29. Double-click the Target object.

The Transformations window opens, as shown in the following image.

30. Click the Automap button.

31. Click OK to close the Transformations window.

32. On the toolbar, click the Run button and click Submit.

33. Enter loadim for the procedure name and click Save.

   Wait for the completed message to appear in the console log and then click View Last Log to verify that the three rows were loaded.

   To see the images that were loaded into the table, you need to write a FOCUS report that sends its output to a PDF file.

34. Right-click baseapp, select New, and then click Procedure.

35. Enter the following lines:
TABLE FILE textures
PRINT PICTURE
BY FNAME
ON TABLE PCHOLD AS baseapp/texturpt FORMAT PDF
ON TABLE SET STYLE *
TYPE=REPORT,$
TYPE=DATA,COLUMN=PICTURE,IMAGE=(PICTURE),SIZE=(1 1),POSITION=(+0.5 +0.5), $
END

36. Click Run.

The PDF file will open in the workspace.

Updating Values in a Table

Normally, when using the Insert/Update load type when a row in the target table is updated, all of the values in the row of the target table except for the key column(s) are replaced with the new values.

However, there are circumstances where it is useful to add an incoming value to the value in the target table. For example, if the source table contains a count of units sold in some time period and you want the target table to contain the total count.

Consider a source file with one row, which is copied to a new target table, as shown in the following image.

<table>
<thead>
<tr>
<th>Source</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product Number</td>
<td>Product Number</td>
</tr>
<tr>
<td>Month</td>
<td>Count</td>
</tr>
<tr>
<td>1001</td>
<td>1</td>
</tr>
<tr>
<td>1001</td>
<td>23</td>
</tr>
</tbody>
</table>
Now consider a new source file with one row, and instead of replacing the existing count with the new count, the new count is added to the existing count, as shown in the following image.

<table>
<thead>
<tr>
<th>Source</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product Number</td>
<td>Month</td>
</tr>
<tr>
<td>1001</td>
<td>2</td>
</tr>
</tbody>
</table>

The following examples will walk you through the steps of creating a synonym for a source table, writing a flow to create and load one month of data to a target table, and then writing another flow to add a subsequent month of data to the target table.

**Example:** Creating a Synonym for a CSV (Comma Separated Values) File

The source table for this tutorial can be found in the application directory where you created the tutorial files. By default, this location is the ibisamp directory.

From the DMC:

1. Right-click *ibisamp*, point to New, and click Synonym (Create or Update).
   
The Select adapter to configure dialog box opens.
2. Under the Delimited Flat File folder, click <local> and click OK.
   
The Select Data File dialog box opens.
3. In the Data Files Location box, enter *ibisamp*.
4. In the Extension box, enter csv.
5. Click Next.
   
The Select Synonym candidates dialog box opens.
6. Select the checkbox next to the *sales_orders* synonym.
   
   To view the CSV file, right-click *sales_orders* and click Show File. Click the X in the upper right hand corner to close the view.
7. Set the following parameters to create the synonym:
   
a. In the Field Delimiter, leave comma selected.

   b. In the Field Enclosure box, select " (double quote) from the drop-down menu.
c. Select the Header row checkbox.

d. Expand the Customize Data type mappings section. In the Decompose Data fields into components box, select OFF - Turn Off from the drop-down menu.

e. In the Prefix box, enter dm.

8. Click Next.

The Status dialog box opens.

9. Ensure that the status shows "Created successfully" and no other messages. Click Close.

10. To view the data through the synonym, expand the ibisamp directory, right-click dmsales_orders, and click Sample Data.

**Example:** **Creating an Initial Load Data Flow**

The next step is to create a Data Flow that creates a new target table and loads the first month of data. In this example, we will be using mysamp as the application directory.

From the DMC:

1. Right-click mysamp, point to New, and click Flow.

   The Data Flow workspace opens.

2. Right-click the left hand side of the workspace and click Add Source.

   The Select Synonym dialog box opens.

3. Click the ibisamp directory and then click dmsales_orders. Click Select.

   The source synonym is added to the workspace.

4. Double-click the SQL object to open it.

5. Hold the Ctrl key to select and select the PRODUCT_NUMBER and PRODUCT_NAME columns. Click the >> button to add them to the Selected Columns section.

6. With both columns selected, select Group By from the Aggregates drop-down menu.

7. From the Available Columns section, hold the Ctrl key and select the REVENUE and QUANTITY columns. Click the >> button to add them to the Selected Columns section.
8. With both columns selected, select *Sum* from the Aggregates drop-down menu, as shown in the following image.

9. Click *OK* to close the Column Selection dialog box.

10. Right-click the SQL object and click *WHERE Filter*.

   The WHERE Filter calculator opens.

11. Click the *Conditions* tab.

12. Right-click the *ORDER_DATE* column, point to *Insert Function*, and click *Return part month of date* from the list of functions.

   The expression `MONTH(T1.ORDER_DATE)` is added.

13. Enter `=1` at the end of the expression, as shown in the following image.
14. Click OK to close the WHERE Filter calculator.

15. Right-click the right side of the workspace, point to Add Target, and click New.

16. Right-click the target object and click Properties.

   The Properties panel for the target opens.

17. Enter the following values:
   a. For Adapter, select any adapter for a relational database.
   b. For Synonym, enter the name of your application directory and a name for the target table For example, mysamp/dmsales_counts.
   c. For Load Type, leave it as Insert/Update.
   d. In the Key box, under the Adapter Specific section, enter or select PRODUCT_NUMBER.

18. On the Flow tab, from the Run group, click Run and then click Submit.

19. Click OK in the Save changes dialog box.

   The Save Procedure As dialog box opens.

20. In the Name box, enter create_dmsales_counts.

21. Click Save.

   You should see a message in the Console Log at the bottom of the DMC that says Request mysamp/create_dmsales_counts complete.

   **Note:** If the Console Log is not visible, on the Home tab, from the View group, click Console Log.

22. On the Flow tab, from the Reports group, click View Last Log to view the number of rows that were inserted. Click the X on the tab to close the log view and return to the flow.
23. To view the data that was loaded, right-click the target object, point to Operations, and click Sample Data. The report shows the product information that was loaded, specifically, the quantity of each. For example, Product Number 1005 shows a Quantity of 2, as shown in the following image.

Example: Creating an Update Data Flow

In this step, we will start with the initial load flow and edit it so that it updates values in the target table by adding new to existing values.

1. In the Data Flow that was just created, right-click the SQL object and click WHERE Filter.

   The WHERE Filter calculator opens.

2. Click the Conditions tab.

3. In the expression, change the number 1 to a 2.

4. Click OK.

5. Right-click the target object and click Delete.

6. Click Yes to confirm that you want to delete the target object.

7. Right-click the right hand side of the workspace, point to Add Target, and click Existing.

   The Select Synonym dialog box opens.

8. Select the application directory you used to create the target table. In this example, it is mysamp. Then, select the target table in the available list, for example, mysamp/dmsales_counts.
9. Click **Select**.

10. Right-click the new target object and click **Properties**.

    The Properties pane opens.

11. Select *Update the existing record* from the If the record exists drop-down menu.

12. Double-click the target object.

    The Transformations dialog box opens.

13. Click the AutoMap button to automatically map the source to the target columns. This defines the assignments that are done from source to target when inserting a new row, as shown in the following image.

![Transformations dialog box](image)

14. Click the **On Match Expressions** tab.

15. In the Target Columns section, double-click **PRODUCT_NUMBER**.

    The On Match Transformation calculator opens.

16. Double-click **REVENUE** to add it to the expression. This is the value from the source table.

17. Click the plus sign.

18. Expand **dmsales_counts** and double-click **REVENUE**. This will add D.REVENUE to the expression, which is the value from the target table.

19. Click **OK** to close the On Match Transformation calculator.

20. In the Target Columns section of the Transformations dialog box, double-click **QUANTITY**.

    The On Match Transformation calculator opens.
21. Double click **QUANTITY** to add it to the expression.

22. Click the plus sign.

23. Expand dmsales_counts and double click **QUANTITY**. This will add D.QUANTITY to the expression, which is the value from the target table.

24. Click **OK** to close the On Match Transformation calculator.

The Transformations dialog box should look like the following image.

![Transformations Dialog Box](image)

25. Click **OK** to close the Transformations dialog box.

26. On the Quick Access Toolbar, click the **Save As** button.

The Save Procedure As dialog box opens.

27. Enter a name for the procedure, for example, **update_dmsales_counts**.

28. Click **Save**.

29. On the Flow tab, from the Run group, click **Run** and then click **Submit**.

The message "(ICM18763) Request ibisamp/update_dmsales_counts complete" should appear in the console log, confirming that the flow was submitted successfully.
30. On the Flow tab, from the Reports group, click **View Last Log** to view the log. The log shows that 1 row was inserted, and 74 were updated, as shown in the following image.

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Message Code</th>
<th>Log Messages</th>
<th>Application</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017/06/28</td>
<td>CIM18743</td>
<td>Starting Load</td>
<td>mysamp</td>
<td>update_dmsales_counts</td>
</tr>
<tr>
<td>2017/06/28</td>
<td>CIM18745</td>
<td>Commit forced at 75 for 75 row(s)</td>
<td>mysamp</td>
<td>update_dmsales_counts</td>
</tr>
<tr>
<td>2017/06/28</td>
<td>CIM18041</td>
<td>-- stats for source file</td>
<td>mysamp</td>
<td>update_dmsales_counts</td>
</tr>
<tr>
<td>2017/06/28</td>
<td>CIM18316</td>
<td>75: Row(s) processed by job</td>
<td>mysamp</td>
<td>update_dmsales_counts</td>
</tr>
<tr>
<td>2017/06/28</td>
<td>CIM18519</td>
<td>0: Row(s) rejected due to format error</td>
<td>mysamp</td>
<td>update_dmsales_counts</td>
</tr>
<tr>
<td>2017/06/28</td>
<td>CIM18372</td>
<td>-- stats for target file dmsales_counts</td>
<td>mysamp</td>
<td>update_dmsales_counts</td>
</tr>
<tr>
<td>2017/06/28</td>
<td>CIM18516</td>
<td>75: Row(s) processed by job</td>
<td>mysamp</td>
<td>update_dmsales_counts</td>
</tr>
<tr>
<td>2017/06/28</td>
<td>CIM18514</td>
<td>75: Row(s) accepted by target table</td>
<td>mysamp</td>
<td>update_dmsales_counts</td>
</tr>
<tr>
<td>2017/06/28</td>
<td>CIM18515</td>
<td>1: Row(s) inserted into target table</td>
<td>mysamp</td>
<td>update_dmsales_counts</td>
</tr>
<tr>
<td>2017/06/28</td>
<td>CIM18517</td>
<td>74: Row(s) updated in target table</td>
<td>mysamp</td>
<td>update_dmsales_counts</td>
</tr>
<tr>
<td>2017/06/28</td>
<td>CIM18518</td>
<td>0: Row(s) deleted from target table</td>
<td>mysamp</td>
<td>update_dmsales_counts</td>
</tr>
<tr>
<td>2017/06/28</td>
<td>CIM18320</td>
<td>0: Row(s) rejected due to validation error</td>
<td>mysamp</td>
<td>update_dmsales_counts</td>
</tr>
<tr>
<td>2017/06/28</td>
<td>CIM18521</td>
<td>0: Row(s) rejected due to no match</td>
<td>mysamp</td>
<td>update_dmsales_counts</td>
</tr>
</tbody>
</table>

31. To view the data that was loaded, return to the flow, right-click the target object, point to **Operations**, and click **Sample Data**. The report shows that product 1005 now has a Quantity of 445, as shown in the following image.
Change Data Capture (CDC) enables reading from database logs to determine what rows in a table have changed. This approach is useful when loading from large source tables where it would take too long to process the entire table.

Databases currently supported are MySQL 5.7 and higher, Microsoft SQL Server 2008, 2012, and 2014 Enterprise Edition, UDB/Db2, and ORACLE 10g, 11g, and 12c Enterprise Edition. Additional databases will be supported in future releases.

To use Change Data Capture, the database must be configured to use database logging. In most organizations this must be done by a database administrator. A Table Log Records synonym is created for a log table, allowing reading from the database log records for a selected table. This CDC synonym can be used as a source in a data flow or in a direct load flow. The target synonym on such flows has a load type of IUD Processing. IUD is an abbreviation for Insert/Update/Delete.

In this chapter:

- Enabling the Database for Logging in Relational Databases
- Properties of a Table Log Record Synonym
- Tutorial: Using Change Data Capture with Relational Databases

Enabling the Database for Logging in Relational Databases

To use Change Data Capture, the source database must be configured to use database logging. In most organizations this must be done by a database administrator.

This example uses dminv as the source table. To create this table, see How to Create Sample Procedures and Data on page 78.

The database table must be enabled for logging, and the procedure depends on the database being used.

Configuring CDC in Microsoft SQL Server

Microsoft SQL Server supports Change Data Capture in the Enterprise, Evaluation, and Developer editions.
To use logging, the source database must be configured to use database logging. In most organizations, this must be done by a database administrator.

Logging must be enabled at the database level since it is disabled by default. The database table must be enabled for logging as well. To enable logging, you must be connected as a member of the sysadmin server role.

To enable logging on a database level, connect to the database you want to enable and issue the following commands:

```sql
exec sys.sp_cdc_enable_db
select name, is_cdc_enabled from sys.databases
```

You will see a list of all databases and a 1 if logging is enabled for the database, or a 0 otherwise.

To check if the SQL Server Agent is running, open the Microsoft SQL Server Management Studio and navigate to the Object Explorer window. Click the icon for SQL Server Agent. It should have a green arrow. If it does not, right-click SQL Server Agent and click Start.

The SQL Server Agent and two jobs must be running. The jobs are created automatically when the database is enabled. For example, if the database is called main, the jobs would be named `cdc.main_capture` and `cdc.main_cleanup`.

After the SQL Server job Cleanup completes, the content of the current SQL Server transaction log changes. If the checkpoint file created by DataMigrator contains the LSN value, which is not within a Validity Interval of the retained transaction log, the Adapter setting LOG_SWITCHED is taken into account.

The Adapter setting LOG_SWITCHED is located on the Change Setting screen:

```sql
ENGINE engine SET LOG_SWITCHED { STOP | PROCEED }
```

The default value is STOP.

For more information on the checkpoint file, see Creating a DataMigrator Direct Load Flow on page 677.

In some situations, the updates on the database occur and the Cleanup job runs, but the DataMigrator flow is not running after the Cleanup job. In this case, the Validity Interval on the database might not have the LSN value that is matching the LSN in the checkpoint file.

When the value of the LOG_SWITCHED is set to STOP and the LSN value in the checkpoint file is not within a Validity Interval of the retained transaction log, the CDC flow returns the following error:

```
(FOC45008) INVALID CHECKPOINT
```
The checkpoint file is not updated at this point.

**Note:** To continue the successful flow execution, delete the existing checkpoint file.

When the value of the LOG_SWITCHED setting is set to PROCEED and the LSN value in the checkpoint file is not within a Validity Interval on the database, the CDC flow proceeds and updates the checkpoint file.

**Note:** In this situation, the change data values could be not updated by a flow as missing in the Validity Interval.

Each table to be used must have CDC enabled. This can be done from Microsoft SQL Server Management Studio. You must be a member of the db_owner database role to run the stored procedure. Issue the following commands, where `tablename` is the name of the table.

```sql
exec sys.sp_cdc_enable_table @source_schema = 'dbo', @source_name = 'tablename', @role_name = NULL
select is_tracked_by_cdc FROM sys.tables WHERE name = 'tablename'
```

**Example:**

```sql
exec sys.sp_cdc_enable_table @source_schema = 'dbo', @source_name = 'dminv', @role_name = NULL
select is_tracked_by_cdc FROM sys.tables WHERE name = 'dminv'
```

You will see a 1 if the table is enabled, and a 0 otherwise.

### Configuring CDC in MySQL

MySQL supports Change Data Capture.

### Preparing the Environment for CDC with MySQL

In order to use CDC with MySQL, you must install Java and MySQL client utilities on the system where the DataMigrator Server is installed. Additionally, on that system, you must have the Environment Variable CLASSPATH pointing to the location where the Connector/J is installed. Note that the location of Connector/J itself can be anywhere on your system. For complete documentation on configuring data adapters for MySQL, see the *Adapter Administration* manual.

### Java

To use CDC with the MySQL adapter, you must have Java installed. The system Environment Variables JAVA_HOME or JDK_HOME must be specified, and the location of Java must be added to the PATH Environment Variable.

For example, if your DataMigrator Server is on Linux, add the following lines to your profile:
Connector/J

Connector/J is an interface that provides MySQL support for Java client programs that use JDBC connections. Connector/J is a JAR file with a standard name of mysql-connector-java-version-bin.jar. The file can be installed on any system on your environment.

The location of the Connector/J must be specified in either the Environment Variable CLASSPATH on the same system where DataMigrator is installed, or in a IBI_CLASSPATH variable of a DataMigrator Server. To specify the location of the Connector/J in the IBI_CLASSPATH variable of a DataMigrator Server, perform the following steps:

1. In the DMC, expand Workspace and expand the Java Services folder.
2. Right-click DEFAULT and select Properties.
   The Java Services Configuration pane opens.
3. Expand the Class path section.
4. In the IBI_CLASSPATH field, specify the full path file name to the MySQL Connector J JAR file.
5. Click Save and Restart Java Services.

Utilities for Processing Binary Log Files

While the MySQL database does not have to be installed on the same system where DataMigrator server installed, the MySQL client utilities for processing binary log files, mysql and mysqlbinlog, must be placed into a directory on the same system where DataMigrator is installed.

The version of these MySQL client utilities must be the same as the version of the MySQL database. The Environmental Variable PATH should list the path to the directory where the MySQL client utilities are placed.

The maximum size of the binary log files is set in the max_binlog_size system variable, and has a maximum possible value of 1 GB. The Binary Log is flushed when its size reaches its maximum, 1 GB. As a general rule, avoid using the LONGTEXT data type in a table that is used for CDC operations, as this data type has large storage requirements.
Configuring MySQL

MySQL enables logging at the database level by storing the history of database changes in the Binary Log Files. These Binary Log Files are used by DataMigrator for CDC operations and must use row-based logging. For example:

```bash
--binlog-format=ROW
```

Although it is possible to change the logging format at runtime, the best practice is to set it at a database server startup.

A user ID assigned on a connection to a MySQL database must have either a SUPER privilege or three global privileges: REPLICATION CLIENT, REPLICATION SLAVE, and SELECT. This task is typically performed by a database administrator. One method to assign such privileges is to issue the SQL statements shown in the following examples:

To create `user1` on any host:

```sql
CREATE USER 'user1'@'%' IDENTIFIED BY 'password1';
```

To grant `user1` all privileges:

```sql
GRANT ALL PRIVILEGES ON *.* TO 'user1'@'%' IDENTIFIED BY 'password1';
```

To enable `user1` for reading the binary logs from any subordinate host:

```sql
GRANT REPLICATION CLIENT ON *.* TO 'user1'@'%' IDENTIFIED BY 'password1';
```

To enable `user1` for reading the binary logs from any master host:

```sql
GRANT REPLICATION SLAVE ON *.* TO 'user1'@'%' IDENTIFIED BY 'password1';
```

To enable `user1` for selecting all objects in a database named `dbname1`:

```sql
GRANT SELECT ON dbname1.* TO 'user1'@'%' IDENTIFIED BY 'password1';
```

Configuring CDC in Oracle

Configuring ORACLE for Change Data Capture can be done from Oracle SQL Developer or SQL Plus. In most organizations, this must be done by a database administrator.

```sql
connect sys/* as sysdba
```

To use the Change Data Capture, the database must be configured to use archiving. Enter the command:

```sql
ALTER SYSTEM ARCHIVE LOG START;
```

When using ORACLE XE (Express Edition), enter the following commands instead:
SHUTDOWN IMMEDIATE
STARTUP MOUNT
ALTER DATABASE ARCHIVELOG;
ALTER DATABASE OPEN;

Confirm that archive log is enabled with the command:

```
select log_mode from v$database;
```

You should see the following result:

```
LOG_MODE
------------
ARCHIVELOG
```

To retrieve log information, a user ID needs to access certain system tables through public views. Issue these commands, where `userid` represents the user ID running CDC jobs.

```
grant select on "SYS"."V_$LOG" to userid;
grant select on "SYS"."V_LOGMNR_CONTENTS" to userid;
grant select on "SYS"."V_$DATABASE" to userid;
```

To enable logging, a user ID must be granted the following, where `userid` represents the user ID to be used.

```
GRANT EXECUTE ON dbms_logmnr TO userid;
GRANT SELECT ANY TRANSACTION TO userid;
```

The table being logged must have all data columns logged by issuing the following command, where `tablename` represents the table to be logged.

```
ALTER TABLE tablename ADD SUPPLEMENTAL LOG DATA (ALL) COLUMNS;
```

Add supplemental log information for the database by entering:

```
ALTER DATABASE ADD SUPPLEMENTAL LOG DATA (ALL) COLUMNS;
```

You can confirm that supplemental logging is enabled for a table with the following command. It should show YES.

```
select distinct
  b.owner, b.object_name, supplemental_log_data_min
from v$database a, sys.all_objects b
where b.owner = 'userid' and b.object_name = 'tablename'
```

**Note:** If you recreate a database table, you must re-issue this command.
Change Data Capture for Oracle uses the DBMS_LOGMNR package to extract values from the Oracle redo log files. This does not support the following data types:

- LONG
- LONG RAW
- CLOB
- BLOB
- NCLOB
- ADT
- COLLECTION

**Configuring CDC in IBM Db2 Universal Database on Windows and UNIX**

Note that CDC support for UDB and Db2 requires the CLI interface. To enable the database for logging, issue the following commands from the Db2 command center:

```
UPDATE DATABASE CFG FOR database USING DFT_SQLMATHWARN YES
UPDATE DATABASE CFG FOR database USING LOGARCHMETH1 LOGRETAIN
UPDATE DBM CFG USING DISCOVER DISABLE
UPDATE DBM CFG USING DISCOVER_INST DISABLE
GRANT DBADM ON database TO USER userid
```

The Db2 table must have capture mode enabled. This can be done from the Db2 Command Center or a stored procedure.

In order to use Change Data Capture, each table must have data changes written to the log in an expanded format. Enter the following command, where `tablename` is the name of the table:

```
ALTER TABLE tablename DATA CAPTURE CHANGES
```

To return a table back to normal logging, use the following command:

```
ALTER TABLE tablename DATA CAPTURE NONE
```

**Note:** If you recreate a database table, you must re-enable expanded logging. To see which tables have expanded logging, you can issue the following Db2 select statement.

```
SELECT CREATOR, NAME, TYPE
FROM SYSSIBM.SYSTABLES WHERE DATA_CAPTURE = 'Y'
```

**Configuring CDC in IBM Db2 Universal Database on System Z**

**Note:** CDC support for UDB and Db2 requires the CLI interface.
To enable the database for logging, issue the following Db2 commands:

```
UPDATE DATABASE CFG FOR database USING DFT_SQLMATHWARN YES
UPDATE DATABASE CFG FOR database USING LOGARCHMETH1 LOGRETAIIN
UPDATE DBM CFG USING DISCOVER DISABLE
UPDATE DBM CFG USING DISCOVER_INST DISABLE
GRANT DBADM ON database TO USER userid
```

In order to use Change Data Capture, each table must have data changes written to the log in an expanded format. Enter the following command, where `tablename` is the name of the table:

```
ALTER TABLE tablename DATA CAPTURE CHANGES
```

To return a table back to normal logging, use the following command:

```
ALTER TABLE tablename DATA CAPTURE NONE
```

**Note:** If you recreate a database table, you must re-enable expanded logging. To see which tables have expanded logging, you can issue the following Db2 select statement:

```
SELECT CREATOR, NAME, TYPE
FROM SYSIBM.SYSTABLES WHERE DATACAPTURE = 'Y'
```

**Configuring CDC in IBM Db2 Universal Database on iSeries**

**Note:**

- The prerequisite requirement for CDC with Db2 on iSeries (IBM i) is to assign the value *MAXOPT3 to attribute RCVSIZOPT (receiver size options) at the initial stage, when the journal is created. Example commands are:

```
CRTJRN JRN(LIBRARY1/JRNLB) JRNRCV(LIBRARY1/RCV01) RCVSIZOPT(*MAXOPT3 *RMVINTENT)
```

When the journal receiver is changed, the new receiver gets a new (LSN) sequence number. Assign the *CONT value to the SEQOPT attribute to allow your journal entry sequence number to continue in the subsequent receivers. For example:

```
CHGJRN JRN(LIB71M1/QSQJRN) JRNRCV(*GEN) RCVSIZOPT(*RMVINTENT *MAXOPT3) SEQOPT(*CONT)
```

- The name of the journaling Db2 table on iSeries (IBM i) must not exceed 10 characters. To process DELETE records, make sure that the *BOTH (before-images and after-images) value has been set for the journal IMAGES parameter.

**Note:** Your flow retrieves the journal entries that are presently available, regardless of the journal state. To ensure that your flow gets the most current updates from the journal entries, verify that your journal is in the *ACTIVE state.
Journaling for an iSeries (IBM i) library can happen in one of three ways:

1. The library is initially created by Db2 with the CREATE COLLECTION collection-name or CREATE SCHEMA schema-name command, which creates the library itself, catalogs resources with the library and standard journal resources within the library, and turns on the journaling. The CREATE SCHEMA and CREATE COLLECTION commands are synonyms for each other, and there are no differences in their actual functions.

2. Add journal resources to an existing not journaling library and turn them on, effectively turning it into a Journal Library. This option also allows for non-standard names to be used in the journaling.

3. Add a QDFTJRN Data Area to an existing not journaling library. This effectively redirects journaling to happen in an alternate library (that already has journaling enabled). Thus, the not journaling library becomes a Journal Library with journals elsewhere.

These three items are described in the following sections.

**Creating a Journal Library**

Start the System i Main Menu screen with the user ID that will be used to run the CDC flows.

To start Structured Query Language interface, type:

```sql
STRSQL
```

To create a collection, type:

```sql
CREATE COLLECTION LIBRARY1
```

To create the Db2 table TABLE1 in the library LIBRARY1, type:

```sql
CREATE TABLE LIBRARY1/TABLE1(F1 INT, F2 CHAR (10 ))
```

To insert a row into TABLE1, type:

```sql
INSERT INTO TABLE1 VALUES (1,'Rec 1 t1')
```

To exit the Structured Query Language interface, press the functional key [PF3].

**Turning a Library Into a Journaling Library**

**Note:** Minimized journal entries are not supported by DataMigrator. Therefore, the MINENTDTA parameter in the command CRTJRN (or CHGJRN) should not be specified. If it must be specified, this parameter should use an attribute of *NONE.

To turn a library into a journaling library, use the following commands:

To create a library called LIBRARY2, type:

```sql
CRTLIB LIBRARY2
```
To create a journal receiver called QSQJRN0001 for the library LIBRARY2, type:

```crtjrnrcv jrnrcv(library2/qsqjrn0001)```

To create a journal QSQJRN in the library LIBRARY2, with journal receiver QSQJRN0001, type:

```crtjrn jrn(library2/qsqjrn) jrnrcv(library2/qsqjrn0001)```

To start Commitment Control with a lock level for ALL with default journal QSQJRN in library LIBRARY2, type:

```strcmtctl lcklvl(*all) dftjrn(library2/qsqjrn)```

To change current library to LIBRARY2, type:

```chgcurlib library2```

To start Structured Query Language interface, type:

```strsql```

To create the Db2 table TABLE2 in the library LIBRARY2, type:

```create table library2/table2(f1 int, f2 char(10))```

To insert a row into TABLE2, type:

```insert into table2 values (1, 'Rec 1 t2')```

To exit the Structured Query Language interface, press the Functional key [PF3].

To ensure journaling (to turn it ON) by adding a not journaling, pre-existing Db2 table, TABLE2, to an existing Journal Library, type:

```strjrnpf file(library2/table2) jrn(library2/qsqjrn)```

As a result of this statement, the following message appears: “Object of type *FILE already being journaled”.

**Note:** Journaling can be turned on and off for individual files with the STRJRNPF and ENDJRNPF commands.

At this point, the table TABLE2 is *log enabled*, which means that journaling works for this table. You can proceed to DataMigrator, creating the Table Log Records synonym for this table and CDC flows.
Enabling Journaling on a Not Journaling Library

There might be situations when you need to keep a table or tables in one library, and the journaling in a different library. In order to get these results, the not journaling library should get connected to a Journal library.

The following names will be used for the objects in the examples below:

- LIBRARY1 is the Journal Library that was created in Creating a Journal Library, above.
- QDFTJRN is a data area, which must exist in the LIBRARY3.

To create a library called LIBRARY3, type:

`CRTLIB LIBRARY3`

To Create a QDFTJRN data area in the not journaling library and have LIBRARY3 point to the Journal Library LIBRARY1, type:

`CRTDTAARA DTAARA(LIBRARY3/QDFTJRN) TYPE(*CHAR) LEN(25)
VALUE('LIBRARY1  QSQJRN    *FILE')`

As a result of this, the following statement message appears: “Data area QDFTJRN created in library LIBRARY3”.

Note: VALUE() is a fixed position string with schema starting in position 1, journal name in position 11 and *FILE in position 21. In other words, VALUE('SSSSSSSSSSJJJJJJJJJJ*FILE') with blank padding for names less than 10.

To change the current library to the library LIBRARY2 and display the properties of that library, type:

`CHGCURLIB LIBRARY3`

To start the Structured Query Language interface, type:

`STRSQL`

To create the Db2 table TABLE1 in the library LIBRARY2, type:

`CREATE TABLE LIBRARY3/TABLE3(F1 INT, F2 CHAR (10 ))`

To insert a row into TABLE1, type:

`INSERT INTO TABLE3 VALUES (1, 'Rec 1 t3')`

To exit the Structured Query Language interface, press the Functional key [PF3].

List of Optional Commands
Change the current library to the schema you have created with following command:

CHGCURLIB LIBRARY1

To change user profile, type:

CHGPRF CURLIB(LIBRARY1)

All tables that are created within this COLLECTION are being journaled. In other words, all the tables in this library have journals and Change Data Capture enabled for them.

You can confirm the current schema by issuing the command:

DSPLIBL

The result of this command should resemble the following image:

<table>
<thead>
<tr>
<th>Opt</th>
<th>Library</th>
<th>Type</th>
<th>Device</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>_</td>
<td>LIBRARY1</td>
<td>CUR</td>
<td></td>
<td>COLLECTION - created by SQL</td>
</tr>
</tbody>
</table>

To review the current library, type:

DSPLIBL

Scroll down the pages to view the journaling information.

Please observe the attributes of the table TABLE3, and notice that it is located in the library LIBRARY3, yet its current journal is QSQJRN which belongs to the library LIBRARY1.

Configuring CDC in IBM Db2 Universal Database on z/OS

Note: CDC support for UDB and Db2 requires the CLI interface.
The configuration file DSNAOINI (*.ini file) should contain the proper value for the MVSATTACHTYPE attribute. This configuration file should contain either:

```
MVSATTACHTYPE=CAF
```

or

```
MVSATTACHTYPE=RRSAF
MULTICONTEXT=0
```

Review the user ID that is specified in the DataMigrator connection for the Db2 adapter. This user ID needs to have SYSADM authority and an authority for monitoring Db2 logs. The latter implies that the system table SYSIBM.SYSUSERAUTH should contain the values for the MON1AUTH and MON2AUTH fields set to Y.

To ensure that data changes are written to the log, enabling Change Data Capture for a table, enter the following Db2 commands, where `tablename` is the name of the table to be log enabled.

```
ALTER TABLE tablename DATA CAPTURE CHANGES
COMMIT WORK;
GRANT ALL ON tablename TO PUBLIC;
COMMIT WORK;
```

To verify that the tables have the logging enabled, issue the following Db2 select statement:

```
SELECT CREATOR, NAME, TYPE
FROM SYSIBM.SYSTABLES WHERE NAME LIKE '%tablename%' AND TYPE IN( 'T','P');
```

**Configuring CDC in MariaDB**

MariaDB supports Change Data Capture.

**Preparing the Environment for CDC with MariaDB**

In order to use CDC with MariaDB, you must install Java, Connector/J, and the MariaDB client utilities. All of these components must be installed on the same system where the DataMigrator Server is installed. For complete documentation on configuring data adapters for MariaDB, see the *Adapter Administration* manual.

**Java**

To use CDC with the MariaDB adapter, you must have Java installed. The system Environment Variables JAVA_HOME and JDK_HOME must be specified and the location of Java must be added to the PATH Environment Variable.

For example, if your DataMigrator Server is on Linux, add the following lines to your profile:
Connexor/J

Connexor/J is an interface that provides MariaDB support for Java client programs that use JDBC connections. Connexor/J is a JAR file. MariaDB Connexor/J .jar files are available at https://downloads.mariadb.org/connector-java/.

After you download the .jar file, the location of the Connexor/J must be specified in either the Environment Variable CLASSPATH on the same system where DataMigrator is installed, or in a IBI_CLASSPATH variable of a DataMigrator Server. To specify the location of the Connexor/J in the IBI_CLASSPATH variable of a DataMigrator Server, perform the following steps:

1. In the DMC, expand Workspace and expand the Java Services folder.
2. Right-click DEFAULT and select Properties.
   
   The Java Services Configuration pane opens.
3. Expand the Class path section.
4. In the IBI_CLASSPATH field, specify the full path file name to the MariaDB Connexor J JAR file.
5. Click Save and Restart Java Services.

Utilities for Processing Binary Log Files

While the MariaDB database does not have to be installed on the same system where the DataMigrator Server is installed, the MariaDB client utilities for processing binary log files, mysql and mysqlbinlog, must be placed into a directory on the same system where DataMigrator is installed.

The version of these MariaDB client utilities must be the same as the version of the MariaDB database. The Environmental Variable PATH should list the path to the directory where the MariaDB client utilities are placed.

The maximum size of the binary log files is set in the max_binlog_size system variable, and has a maximum possible value of 1 GB. The Binary Log is flushed when its size reaches its maximum, 1 GB. As a general rule, avoid using the LONGTEXT data type in a table that is used for CDC operations as this data type has large storage requirements.
Configuring MariaDB

MariaDB enables logging at the database level by storing the history of database changes in the Binary Log Files. These Binary Log Files are used by DataMigrator for CDC operations and must use row-based logging. For example:

```
--binlog-format=ROW
```

Although it is possible to change the logging format at runtime, the best practice is to set it at the database server startup.

A user ID assigned on a connection to a MariaDB database must have either a SUPER privilege or three global privileges: REPLICATION CLIENT, REPLICATION SLAVE, and SELECT. This task is typically performed by a database administrator. One method to assign such privileges is to issue the SQL statements shown in the following examples.

To create `user1` on any host:

```
CREATE USER 'user1'@'%' IDENTIFIED BY 'password1';
```

To grant `user1` all privileges:

```
GRANT ALL PRIVILEGES ON *.* TO 'user1'@'%' IDENTIFIED BY 'password1';
```

To enable `user1` for reading the binary logs from any subordinate host:

```
GRANT REPLICATION CLIENT ON *.* TO 'user1'@'%' IDENTIFIED BY 'password1';
```

To enable `user1` for reading the binary logs from any master host:

```
GRANT REPLICATION SLAVE ON *.* TO 'user1'@'%' IDENTIFIED BY 'password1';
```

To enable `user1` for selecting all objects in a database named `dbname1`:

```
GRANT SELECT ON dbname1.* TO 'user1'@'%' IDENTIFIED BY 'password1';
```

Properties of a Table Log Record Synonym

The Table Log Records synonym is the synonym for a log table. These first three columns in this synonym are added automatically and used for DataMigrator IUD processing. The columns have the following functions:

- **CDC_OPER.** is the operation type (Insert, Update, or Delete)
- **CDC_TID.** is the Transaction ID
- **CDC_TIMES.** is the time stamp
The Configuration section of the CDC synonym consists of special parameters, which properties could be customized by a user. Each CDC parameter in the configuration section of the CDC synonym has a default value of a global variable whose name begins with &&CDC_ and is followed by the parameter name.

The following options are available:

**DB Log Parameters**

**DATA_ORIGIN**

This is set to DBMSLOG to indicate the origin of the data is a DBMS table log.

**START**

Indicates the starting point for reading log records.

- **CHKPT.** After the last transaction retained in checkpoint file. This is the default.
- **CUR_TRAN.** The first transaction in the DBMS log after the job started.
- **CUR_LOG.** The first available transaction in the current active DBMS log.

**CHKPT_SAVE**

Indicates if the last processed transaction should be saved in the checkpoint file.

- **YES.** Retains the last processed transaction in the checkpoint file.
- **NO.** Does not retain the last processed transaction. This option facilitates testing, because use of Sample Data or Test Transforms does not reset the checkpoint.

**CHKPT_FILE**

- **Blank.** This is the default. If no file name is specified, a file is created in the same application directory as the synonym with the same name as the synonym and the extension chp.

Indicates the location and name of the file to use to store checkpoint information in a:

- **Physical location.** The full directory and file name in the format for the operating system used. This option should be used when there are multiple flows that will access the log for the same table. Each flow should have a unique synonym with a unique name for the checkpoint file.
- **Application directory.** The file is created in the same application directory as the synonym.
Listening Parameters

POLLING

Indicates the polling interval in seconds, or how often the database log is scanned. The default is 1 second.

TIMEOUT

Indicates the timeout interval in seconds. If there is no activity in this time interval, the processing will stop. A value of zero means there is no limit. The default value is 1 second.

Note that the two options above work together. For example, if POLLING is 2 and TIMEOUT is 10, the server polls the log every 2 seconds for new transactions. If there are no new transactions after 10 seconds, then polling stops.

Read Limits

MAXLUWS

Indicates the maximum number of database transactions to process before stopping the job. A value of zero (0) means all transactions. The default value is 1 transaction.

Considerations for Table Log Records Synonyms

Timestamp parameter

The Timestamp value of the CDC parameter START is available for Db2 synonyms only.

CUR_LOG

The CUR_LOG value of the CDC parameter START is not supported by IBM Db2 on Mainframe (IBM System z).

Additional iSeries parameters

The following three parameters, COMMIT_MODE, LOG_NAME, and LOG_LOCATION are available on iSeries platforms only, with the Db2 database.

The LOG_NAME and LOG_LOCATION parameters are used when you are required to use a remote journal, instead of the local journal.

COMMIT_MODE

Supports the transactional commitment control.

- **ON.** Normal transaction mode with a commit used. This is the default.

- **OFF.** No commit issued or DBMS uses auto-commit mode.
**LOG_NAME**
Indicates the name of the journal.

**LOG_LOCATION**
Indicates the location name.

**Current library on iSeries**
When using CDC on iSeries, consider changing current library.

**Example:**
Start the DataMigrator server and the Data Management Console. From a user ID that is a member of the Server Administrator role, expand the *Workspace* folder and double-click *Server Profile - edasprof.prf*. Look for a line that resembles the following:

```
-SYSTEM CHGCURLIB LIBRARY1
```

Ensure that the parameter, in this example LIBRARY1, is the name of your collection. If you made a change to file, click the *Save* icon to save the changes.

**Checkpoint file on z/OS**
On a mainframe the checkpoint file should be created first. This can be done by creating two focexec procedures and running them simultaneously: one procedure to report against a Table Log Records synonym and the other procedure will issue an update against the underlying actual table. For more information, see *Creating a Checkpoint File on z/OS* on page 677.

**Tutorial: Using Change Data Capture with Relational Databases**
Change Data Capture loads data targets using only the rows that have changed since the last load. This approach is useful when dealing with large databases that would take too much time and resources to reload completely.

The source table used in this tutorial is an inventory table called *dminv* that is created by running Create Sample Procedures and Data. The source table is in Db2 or UDB.

The tutorial has the following steps:

- Create a target table and perform an initial load.
- Create a synonym for table log records.
- Create a DataMigrator direct load flow.
Update the source table.
Run the flow and view results.
Schedule the flow.

**Example:** Creating a Target Table and Performing an Initial Load

The first step is to create a data flow that creates the target table and loads it with data from the current contents of the source table.

The user ID to run flows using CDC should have the role of Administrator on the DataMigrator Server.

1. Create a new data flow in an application directory that is in the server path, like baseapp.
2. Select `dminv` as a source.
3. Select all the columns from `PROD_NUM` through `COST`.
4. Add a new target, selecting a database type, entering `newinv` as the name, and specifying one key.
5. Save the flow as `dminv_init`.
6. Run the flow.
7. Open the `newinv` target and click Sample Data to confirm that the data was loaded correctly.

**Note:** Key columns are required for loading a table to uniquely identify each row so that Insert/Update/Delete processing can be performed.

**Example:** Creating a Synonym for a Log Table

A synonym for a log table allows reading from the database log records for a selected table. The synonym contains the same columns as the table itself plus three additional columns that identify changes to the table.

This synonym can be used as a source in a data flow with the special load type of IUD Processing. IUD is an abbreviation for Insert/Update/Delete.

**Note:** A synonym for a log table is not a substitute for a synonym for the table itself. A synonym is still needed for the table to be able to read or write from the table directly.

To create a synonym for a log table, in the DMC:

1. Right-click an application directory in the navigation pane, point to New, and click Synonym (Create or Update).
The Select Adapter window opens.

2. Select a connection and click OK.

The Select Synonym Candidates window opens.

3. From the Restrict object type to drop-down menu, select Table Log Records.

   Optionally, select the Filter by owner/schema check box and for object name enter dminv.

4. Click Next.

5. Select the check box in front of DMINV.

   Under the Default synonym name, click DMINV and change the name to DMINV_LOG, as seen in the following image.

6. Click Create synonym. The Create Synonym Status window opens.

7. Click Open First Synonym. The synonym opens by default to the Field View. If it does not, click the tab to switch to that view. Select the segment name if it is not already highlighted.
8. In the DB Log Parameters settings, input CHKPT for START, YES for CHKPT_SAVE, and leave CHKPT_FILE blank.

**Tip:** Another way to change these configuration settings to the synonym is as follows. Leave the variables for the parameters values as: &&CDC_START, &&CDC_CHKPT_SAVE, and &CDC_CHKPT_FILE. Then, expand the Variables folder, right-click a variable, and click Properties. Enter the DEFAULT value for the variable.

The window now looks like this:
**Tip:** The attributes in a synonym that can be specified with a variable are identified in blue color, as shown in the image above. To use a variable for the CDC attribute, right-click the attribute and click *Parameterize Value* in the shortcut menu. The Parameterize Value dialog box opens, as shown in the following image.

![Parameterize Value dialog](image)

The following options are available:

**Use variable to parameterize value**
Select this option if you want to use the standard CDC global variable. For example, global variable CDC_START.

**Use existing variable to parameterize value**
Select this option if you want to use an existing variable, which is listed and can be selected in the dialog box.
Creating a Checkpoint File on z/OS

In this example, we will create a checkpoint file that will be used by a CDC flow. The checkpoint file should be created on a z/OS platform in advance, prior to CDC flow runs.

1. Create a FOCUS stored procedure to report against a Table Log Records synonym. For example:

   ```
   -SET &&CDC_POLLING=1;
   -SET &&CDC_TIMEOUT=60;
   -SET &&CDC_START=CUR_TRAN;
   -SET &&CDC_CHKPT_SAVE=YES;
   -SET &&CDC_MAXLWUS=0;
   -SET &&CDC_COMMIT_MODE=OFF;
   -SET &&CDC_CHKPT_FILE='baseapp/mycheck1.chp';
   TABLE FILE LOG_MTB7
   PRINT CDC_OPER FIELD2 FIELD3
   END
   -RUN
   ```

2. Create a FOCUS stored procedure to update the underlying table. For example:

   ```
   SQL DB2
   UPDATE MYETL.MYTABLE SET FIELD3 = FIELD3 +1 WHERE FIELD2 =2 ;
   END
   -RUN
   SQL DB2
   COMMIT;
   END
   -RUN
   ```

3. Run the first stored procedure - to report against a Table Log Records synonym. At the same time, run the second stored procedure – to update the underlying table. The file mycheck1.chp will be created.

Creating a DataMigrator Direct Load Flow

In this step, we will create a DataMigrator direct load flow to copy data from the log table to the target table.

**Note:** Using a direct load flow with the checkpoint option ensures that all rows read from the log are written to the target table before the checkpoint file is updated. This ensures that the flow can be re-run from the original checkpoint if it cannot write to the target table (for example, if the database is not running).

While a data flow can also be used, it does not provide this assurance. The read and write operations are separate so that the checkpoint file is updated when the log is read.

1. Right-click an application directory, point to New, and click Direct Load Flow.
2. Drag `dminv_log` to the left side of the flow.
3. Drag `newinv` to the right side of the flow.
4. Right-click `newinv` and select *Properties*.
   The target properties page opens.
5. For the Load Type drop-down menu, select *IUD Processing*.
6. For the IUD Control column drop-down, select `CDC_OPER`, as shown in the following image.

Click the X to close the properties page.

7. Right-click `newinv` and select *Target Transformations*.
   The Transformations window opens.
8. Click the Automap button to map the source to target columns.
   Click OK to close the window.

9. Click the Save button and save the flow as dminv_cdc.

**Example:  Updating the Source Table**

To test the Change Data Capture process, make changes to the dminv source table. In this example, three rows are inserted, one row is updated, and one is deleted.

1. Start the client for your database (Db2, MySQL, Oracle, or SQL Server) and connect to the server.
   For SQL Server, enter the line:
   ```
   begin transaction;
   ```

2. Enter the following lines:
   ```
   insert into dminv values ('2001','Compact Flash',1000,50,25) ;
   insert into dminv values ('2002','Memory Stick',1000,80,40) ;
   insert into dminv values ('2003','Memory Stick Pro', 1000,200,100) ;
   update dminv set QTY_IN_STOCK=500 where prod_num = '2002' ;
   delete from dminv where prod_num = '2001' ;
   ```

3. For SQL Server, enter the command:
   ```
   commit transaction;
   ```
   For MySQL, Oracle or Db2, enter the command:
   ```
   commit work;
   ```

**Example:  Reviewing the Log**

Note that the CDC configuration saves a checkpoint of the last transaction processed. However, when you run a Sample Data report using the synonym for the log table, the checkpoint is not saved.
Example: Running the CDC Flow and Viewing the Results

1. Open the CDC flow you created. Click the Run button and select Submit.
2. Click the View Last Log button. The following image shows the log report for the flow.
Note that the statistics show the number of rows inserted, updated, and deleted.

Close the report.

3. In the open flow, $dminv\_cdc$, double-click the $newinv$ target table and select the tab Sample Data. The following image shows that the changes were applied correctly.

![Sample Data Image]

**Note:** To return the $dminv$ table back to its original state, enter the following lines in a stored procedure, where $db$ is the engine used, and run them:

```sql
SQL \$db\ delete from dminv where prod_num > '2001' ;
END
SQL \$db\ commit;
END
```

**Example:** Scheduling the Flow

The final step in setting up Change Data Capture processing is scheduling the flow so that it runs periodically and processes transactions.

1. Switch to the process flow view by clicking the Process Flow tab.
2. Drag the Schedule button into the workflow and double-click it to open the Schedule Properties window.
3. Select a Schedule type, Interval type, and Interval value to run the flow.
4. Save the flow.
Reference: Sample Data

When you request sample data for a log table, a dialog box opens so that you can supply values to override the parameters in the synonym.

For example, you can sample data from a log table without creating a checkpoint by checking the check box in front of &&CDC_CHKPT_SAVE and entering the value NO.

You can also specify where to start reading the log by specifying a value for &&CDC_START. The values are the same as those specified in the synonym.
This appendix provides information about the DataMigrator internal data store for ETL log and statistics.

**In this appendix:**
- Log and Statistics Tables
- FOCUS Logs and Statistics
- CSV Logs and Statistics
- Managing Log and Statistics
- Custom Procedures for Managing FOCUS Logs and Statistics
- Description of the Log Files
- Description of the Statistics Files
- List of Essential ETL LOG Message Codes

### Log and Statistics Tables

DataMigrator jobs can be either run immediately or submitted to the Scheduler. When a job is run, the message from the job appears in the Console Log. When a job is submitted, the messages and summary are written into the logs and statistics files by a DataMigrator service called the Scheduler. The log files contain flow run-time messages, while the statistics files contain run-time statistics for each flow. You can get access to the logs and statistics files in a read-only mode. The structure of these files is subject to change by Information Builders.

You can view or edit the configuration setting for the logs and statistics from the Workspace folder by right-clicking `SCHEDULER` and clicking `Properties`. The Scheduler Configuration dialog box opens.

**Note:** You must be logged in as a Server Administrator with workspace configuration privileges to make any configuration changes.

By default, DataMigrator creates the logs and statistics in a FOCUS file format. Starting from the release 7.7.07, another option is available. Logs and statistics files can be in either a FOCUS file or in a CSV (Comma Separated Values) file.
**FOCUS Logs and Statistics**

When the `sched_standalone_log` Scheduler Configuration parameter is not selected, the logs and statistics are stored in a FOCUS file. This is the default. The two FOCUS files, etllog.foc and etlstats.foc, are created by DataMigrator.

The default location for these files is in the `<edaconf>/catalog/` directory path.

To create a copy of these files, you can perform the Backup operation. By default, Backup creates a backup copy of the logs and statistics inside a `<edaconf>/catalog/` directory path.

To delete the etllog.foc and etlstats.foc files and to create a new pair of empty files, you can perform the Recreate operation.

**Note:** In release 7.6, the logs and statistics files are in FOCUS format, which is limited to 2 GB. Starting from release 7.7, these files are in XFOCUS format and the limit for this format is 32 GB.

**CSV Logs and Statistics**

When the `sched_standalone_log` Scheduler Configuration parameter is selected, the messages from each submitted job are written into an individual set of three standalone CSV files; one set for each submitted job.

The default location for these files is in the `<edaconf>/catalog/dmlog/` directory path.

To create a copy of these files, you can to perform the Backup operation. By default, Backup creates a backup copy of the logs and statistics inside a `<edaconf>/catalog/dmlogbk/` directory path.

To delete these files, you can perform the Recreate operation.

The set of the three files is: the log, the statistics, and the statistics for the Targets on a flow. The naming convention for CSV logs and statistics is:

`appname_jobname_yyyymmddhhmss_jobid_a.csv`

where:

- `appname`
  - Is the name of the application directory where the submitted job was located.

- `jobname`
  - Is the name of the submitted job.

- `yyyymmddhhmss`
  - Is the timestamp when the job was submitted.
jobid
Is the ID assigned to the submitted job.

a
Is an indicator of the file; l - log, s - statistics, t - statistics for targets on a submitted flow.

Example: Names for CSV Logs and Statistics
A job name dmdemo, located in the application directory ibisamp was submitted at 11:19:01 on May 4, 2017. Three CSV files were created for this job:

- ibisamp_dmdemo_20170504111901_d497a29b_l.csv
- ibisamp_dmdemo_20170504111901_d497a29b_s.csv
- ibisamp_dmdemo_20170504111901_d497a29b_t.csv

Note: The size and the number of these files can grow according to the allowed size on a system.

Managing Log and Statistics
Note: You must be logged in as a Server Administrator with workspace configuration privileges in order to make any configuration changes, or to follow any of the following procedures.

To manage the log and statistics from the DMC, expand the Workspace folder, followed by the Logs and Traces folder. Then, right-click DataMigrator Log and Statistics. A shortcut menu will open with your options.

The following options are available in the shortcut menu:

View
Allows you to view the log and statistics tables.

Recreate
Creates empty new log and statistics tables.

Clean
Allows you to specify a date and time.

All records prior to the date and time entered are deleted, and the log and statistics files are rebuilt. This action frees up the space used by the deleted records, so that new records can be added. For FOCUS logs and statistics, this operation does not reduce the size of the log and statistics tables on disk, and new records can still be accepted without growing in size until the newly freed space has been freed up.
Backup

Creates a backup copy of the log and statistics tables on the disk, and a synonym that can be used to retrieve data from them. These tables are created in the <edaconf>/catalog directory that is located on the server (for example, c:\ibi\sv77\dm\catalog). A message will appear in the console showing the location.

Procedure: How to Send an Email Message When There Is an Error Writing to the ETL Log

You can configure the server so that an email message is sent to an address you specify if there is an error writing to the ETL log. For example, even though a log table is automatically maintained so that normal operations prevent it from reaching its maximum size, it is still possible for a flow to generate so many errors that it fills up the log table. This issue can be addressed by having the server send you a message when the log table is full.

1. Configure the server to connect to a SMTP mail server, as described in How to Configure the Server for Email Notification on page 91.

2. There are two ways to customize event routing from the DMC:

   - Right-click the Workspace folder and click Event Routing.

   or

   - Expand the Workspace and then the Logs and Traces folder.

     Right-click Workspace Logs and click Customize Event Routing.

     The Event Routing dialog box opens.

     Click Customize New Event Routing.
3. Select the following options for each field:

- **Type** - *Error*
- **Filter** - *60003 ETLLOG write error*
- **EDAPRINT** - *Yes*

4. In the **E-mail** field, enter the address where the email message will be sent.

5. In the **Procedure** field, enter the name of the process flow, for example, baseapp/pf_backup_recreate.

6. Click **Add**.

7. Click **Save and Restart**.

### Custom Procedures for Managing FOCUS Logs and Statistics

**Note:** You must be logged in as a Server Administrator with workspace configuration privileges to make any configuration changes, or to following either of the following procedures.

**Procedure:** How to Automatically Delete Rows From the ETL Log and Statistics Table

It may be useful for a Server to automate deleting rows from the ETL log and statistics tables. For example, you can delete all rows more than a week old on a daily basis.

1. Create a stored procedure with the following lines:
2. Create a new process flow.
3. Drag the procedure into the flow and connect it to the Start object.
4. Drag a Schedule icon into the workspace.
5. Double-click the Schedule icon and set a schedule for running the flow.
6. Save the flow.

**Note:** The Clean utility rebuilds the ETLLOG. The FDS service has to be stopped to prevent possible file corruption during the rebuild process. Therefore, the following considerations should be taken into account:

- No other jobs should run during an execution of the Clean utility.
- The value of the Log Destinations property of the flow should not be set to ETLLOG.

**Procedure:** How to Automatically Backup and Recreate the ETL Log and Statistics Tables

The DataMigrator server can be configured so that if there is an error writing to the ETL log, one or more of the following events will occur:

- A specified procedure is executed that will backup and recreate ETL log and statistics.
- The message is written in to the server log (edaprint.log).
- The email notifications are sent to a named address with a specified frequency.

The following procedure shows how to backup and recreate ETL log and statistics.

1. Create a new process flow and name it `pf_backup_recreate`.
2. Add the following stored procedures to the process flow in this exact order: `st_backup`, `st_rename`, and `st_recreate`.

The following is an explanation of the stored procedures mentioned.

**st_backup**

This procedure will create the `etllogbk.foc` and `etlstabk.foc` files in the `\dm\catalog\` directory. These files are copies of the `etllog.foc` and `etlstats.foc` files. Example code for `st_backup` is:

```
SHH CRTBACKUP
END
```
**st_rename**

This step is optional. This procedure prevents these files from being overwritten during the next backup process. As a result of this procedure, the `etllogbk.foc` and `etlstabk.foc` files will be renamed, with a timestamp added to their suffix. For example, `etllogbk_20101026_1120258.foc`. Example code for st_rename is:

```
-SET &LOGBK = c:\ibi\srv77\dm\catalog\etllogbk.foc;
-SET &STABK = c:\ibi\srv77\dm\catalog\etlstabk.foc;
-SET &LOGBKREN = etllogbk || _ ' ' || &YYMD || _ ' ' || EDIT(&TOD,'99$99$99') || .foc';
-SET &STABKREN= etlstabk || _ ' ' || &YYMD || _ ' ' || EDIT(&TOD,'99$99$99') || .foc';
!REN &LOGBK &LOGBKREN
!REN &STABK &STABKREN
END
```

**st_recreate**

This procedure deletes the original etllog.foc and etlstats.foc files, and will replace them with the new initialized files. Example code for st_recreate is:

```
EX ETLCRTLS ALL, Y, N
END
```

The process flow should look like the following image.

3. Run the new process flow by clicking Run or Submit one time, or scheduling the flow. You can also use Event Routing, as described in *How to Send an Email Message When There Is an Error Writing to the ETL Log* on page 686.

**Description of the Log Files**

The log files contain the DataMigrator run-time message log. The number of lines that are written to this table are managed by the `sched_log_lines` and `sched_Log_commit_maxlines` Scheduler Configuration parameters.

<table>
<thead>
<tr>
<th>Column</th>
<th>Format</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATE</td>
<td>Alpha 20</td>
<td>Date and time of log entry.</td>
</tr>
</tbody>
</table>
### Description of the Statistics Files

The Statistics file contains the run-time statistics for each flow.

<table>
<thead>
<tr>
<th>Column</th>
<th>Format</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>USER</td>
<td>Alpha 32</td>
<td>User who created the flow.</td>
</tr>
<tr>
<td>JOBNAME</td>
<td>Alpha 23</td>
<td>Unique internal job name.</td>
</tr>
<tr>
<td>APPNAME</td>
<td>Alpha 256</td>
<td>Flow application directory.</td>
</tr>
<tr>
<td>REQ_NAME</td>
<td>Alpha 64</td>
<td>Flow name.</td>
</tr>
<tr>
<td>COUNTER</td>
<td>Integer 9</td>
<td>Line number of the message.</td>
</tr>
<tr>
<td>MESSAGE</td>
<td>Alpha 80</td>
<td>Message strings in the log.</td>
</tr>
<tr>
<td>MSGCODE</td>
<td>Alpha 12</td>
<td>Message code number.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Column</th>
<th>Format</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>S_REQ_ID</td>
<td>Alpha 23</td>
<td>Name of the flow that started this flow.</td>
</tr>
<tr>
<td>S_USERID</td>
<td>Alpha 32</td>
<td>Flow author.</td>
</tr>
<tr>
<td>S_APPNAME</td>
<td>Alpha 64</td>
<td>Flow application directory.</td>
</tr>
<tr>
<td>S_REQ_NAME</td>
<td>Alpha 64</td>
<td>Flow name.</td>
</tr>
<tr>
<td>STAT_SBM_T</td>
<td>Alpha 20</td>
<td>Submit date and time for a flow.</td>
</tr>
<tr>
<td>STAT_STRT_T</td>
<td>Alpha 20</td>
<td>Start date and time of execution.</td>
</tr>
<tr>
<td>STAT_END_T</td>
<td>Alpha 20</td>
<td>End date and time of execution.</td>
</tr>
<tr>
<td>STAT_RTRV_T</td>
<td>Alpha 20</td>
<td>Date and time of retrieval for a flow.</td>
</tr>
<tr>
<td>STAT_FORMAT</td>
<td>Alpha 15</td>
<td>Number of format rejections.</td>
</tr>
<tr>
<td>REQ_RC</td>
<td>Alpha 16</td>
<td>Internal return code for the flow.</td>
</tr>
<tr>
<td>TSCOMID</td>
<td>Alpha 15</td>
<td>TSCOMID for killing a job.</td>
</tr>
<tr>
<td>Column</td>
<td>Format</td>
<td>Function</td>
</tr>
<tr>
<td>----------------</td>
<td>--------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>SESID</td>
<td>Alpha 15</td>
<td>SESID from EDAPRINT.</td>
</tr>
<tr>
<td>ORIG_REQ</td>
<td>Alpha 32</td>
<td>Original flow where dependency started.</td>
</tr>
<tr>
<td>ORIG_USER</td>
<td>Alpha 32</td>
<td>Original user ID running dependent flow.</td>
</tr>
<tr>
<td>DEP_CNT</td>
<td>Integer</td>
<td>Dependency number.</td>
</tr>
<tr>
<td>ETL__ERRTXT</td>
<td>Alpha 80</td>
<td>Text explanation of error messages.</td>
</tr>
<tr>
<td>FILLER</td>
<td>Alpha 30</td>
<td>Reserved for internal use. There are two columns with this field name.</td>
</tr>
<tr>
<td>USER_FILE_L</td>
<td>Alpha 9</td>
<td>Reserved for internal use.</td>
</tr>
<tr>
<td>RUN_BY_SCHED</td>
<td>Alpha 1</td>
<td>Scheduled request.</td>
</tr>
<tr>
<td>CPU_TIME</td>
<td>Alpha 15</td>
<td>Total CPU time in milliseconds.</td>
</tr>
<tr>
<td>S_TARG_NAME</td>
<td>Alpha 64</td>
<td>Target name.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Note:</strong> This and the remaining columns apply on a per target basis.</td>
</tr>
<tr>
<td>STAT_TRANS</td>
<td>Alpha 15</td>
<td>Number of rows processed.</td>
</tr>
<tr>
<td>STAT_ACCEPTS</td>
<td>Alpha 15</td>
<td>Number of rows accepted.</td>
</tr>
<tr>
<td>STAT_NOMATCH</td>
<td>Alpha 15</td>
<td>Number of rows rejected.</td>
</tr>
<tr>
<td>STAT_DBMSERR</td>
<td>Alpha 15</td>
<td>Number of rows rejected by the database for a constraint violation, such as a Unique Index constraint, a Not Null constraint, or a user-specified constraint.</td>
</tr>
<tr>
<td>STAT_INVALID</td>
<td>Alpha 15</td>
<td>Number of rows rejected for failing validation tests.</td>
</tr>
<tr>
<td>STAT_DUPLS</td>
<td>Alpha 15</td>
<td>Number of rejected duplicate records.</td>
</tr>
<tr>
<td>STAT_INPUT</td>
<td>Alpha 15</td>
<td>Number of rows input.</td>
</tr>
<tr>
<td>STAT_CHNGD</td>
<td>Alpha 15</td>
<td>Number of rows updated.</td>
</tr>
</tbody>
</table>
List of Essential ETL LOG Message Codes

The following is a list of essential lines that will be imposed into the ETL LOG, even when the number of lines is limited to a positive integer.

<table>
<thead>
<tr>
<th>Message Code</th>
<th>Log Message</th>
<th>Example or Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICM18007</td>
<td>CPU Time: %1</td>
<td>CPU Time : 46</td>
</tr>
<tr>
<td>ICM18031</td>
<td>Finished</td>
<td></td>
</tr>
<tr>
<td>ICM18040</td>
<td>Return Code = xxx</td>
<td>Return Code = 0</td>
</tr>
<tr>
<td>ICM18041</td>
<td>– stats for source file</td>
<td>Statistics about processing of a source file</td>
</tr>
<tr>
<td>ICM18072</td>
<td>Elapsed run time hh:mm:ss</td>
<td>Elapsed run time 0:00:02</td>
</tr>
<tr>
<td>ICM18122</td>
<td>Request - %1 (Owner: %2) submitted.</td>
<td>Request that was saved by the owner is being submitted. The owner is saved in a flow; e.g., control variable in focexec: -*DM_USERID=user1</td>
</tr>
<tr>
<td>ICM18237</td>
<td>Ending bulk load procedure</td>
<td>Shown in the LOG when a load type on a target is 'Bulk load utility via a disk file'</td>
</tr>
<tr>
<td>ICM18372</td>
<td>– stats for target file: %1</td>
<td>Statistics about processing of a target file</td>
</tr>
</tbody>
</table>

Note: You must be logged in as a Server Administrator with workspace configuration privileges in order to make any configuration changes, or to follow any of the following procedures.

In Release 7.6, the etllog file is in FOCUS format, which is limited to 2 GB. In Release 7.7, the etllog file is in XFOCUS format and the limit for this format is 32 GB.
<table>
<thead>
<tr>
<th>Message Code</th>
<th>Log Message</th>
<th>Example or Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICM18514</td>
<td>%1 : %2 accepted by target table</td>
<td>18 : Row(s) accepted by target table</td>
</tr>
<tr>
<td>ICM18515</td>
<td>%1 : %2 inserted into target table</td>
<td>18 : Row(s) inserted into target table</td>
</tr>
<tr>
<td>ICM18516</td>
<td>%1 : %2 processed by job</td>
<td>18 : Row(s) processed by job</td>
</tr>
<tr>
<td>ICM18517</td>
<td>%1 : %2 updated in target table</td>
<td>0 : Row(s) updated in target table</td>
</tr>
<tr>
<td>ICM18518</td>
<td>%1 : %2 deleted from target table</td>
<td>0 : Row(s) deleted from target table</td>
</tr>
<tr>
<td>ICM18519</td>
<td>%1 : %2 rejected due to format error</td>
<td>0 : Row(s) rejected due to format error</td>
</tr>
<tr>
<td>ICM18520</td>
<td>%1 : %2 rejected due to validation errors</td>
<td>0 : Row(s) rejected due to validation errors</td>
</tr>
<tr>
<td>ICM18521</td>
<td>%1 : %2 rejected due to no match</td>
<td>0 : Row(s) rejected due to no match</td>
</tr>
<tr>
<td>ICM18522</td>
<td>%1 : %2 rejected because duplicate exist</td>
<td>0 : Row(s) rejected because duplicate exist</td>
</tr>
<tr>
<td>ICM18744</td>
<td>Ending Load</td>
<td></td>
</tr>
<tr>
<td>Message Code</td>
<td>Log Message</td>
<td>Example or Explanation</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>ICM18750</td>
<td>bulkload: DBMS error code at execution</td>
<td>The attempt to start the bulkload failed. Check configuration. It is possible that environmental variables that point to the DBMS bulkload facilities have not been configured.</td>
</tr>
<tr>
<td>ICM18751</td>
<td>bulkload: Error found in the bulkload log file %1</td>
<td>Review the log file for DBMS messages that indicate the specific error</td>
</tr>
<tr>
<td>ICM18808</td>
<td>%1 : %2 rejected due to DBMS error</td>
<td>2 : Row(s) rejected due to DBMS error</td>
</tr>
<tr>
<td>ICM18974</td>
<td>Start of Log Record for %1/%2</td>
<td>—— Start of Log Record for baseapp/flow1 ——</td>
</tr>
<tr>
<td>ICM18975</td>
<td>End of Log Record for %1/%2</td>
<td>—— End of Log Record for baseapp/flow1 ——</td>
</tr>
<tr>
<td>ICM18995</td>
<td>The configured maximum number of log lines (%1) has been reached.</td>
<td>The message appears to inform a user that the log lines will be reduced as per the value in the sched_log_lines was set</td>
</tr>
</tbody>
</table>

---

**List of Essential ETL LOG Message Codes**

694 Information Builders
Event-Based Scheduling

You can perform event-based scheduling, which executes a flow from an external application. It can be useful when you want a remote application to run a flow. For example, you might have an application that requires a flow to run. You can add a button to your application, which sends a request to the DM Server to start a flow.

Event-based scheduling can also be useful in a job that updates data. After the update, a message can be sent to the DM Server, instructing it to run a flow. For example, when running a batch job on USS, you can make the final step sending a message to the server telling it to execute the flow. In this instance, the flow would access the latest data.

DataMigrator supports event-based scheduling with the following:

- CMRUN is a stand-alone program that initiates flows. It is available on all supported platforms. Sample scripts (for example, shell, batch, DCL, or JCL depending on your platform) are provided for setting the environment and submitting the flow.

- A FOCUS client using specific FOCUS code can initiate a flow.

**Note:** CMRUN is the preferred method, and should be used whenever possible.

**In this appendix:**

- Starting Flows With CMRUN
- Starting Flows With CMASAP
- Starting Flows From FOCUS
- Determining the Status of an Event-Based Flow

**Starting Flows With CMRUN**

The CMRUN program is a stand-alone program for initiating flows. It also allows you to specify values for user-defined variables. It is usually called from an z/OS JCL job stream, a UNIX shell script, OpenVMS DCL, or a Windows shortcut or command file that sets up prerequisite environment values, as shown in the examples that follow. Samples of these scripts may also be found in the bin location of the server configuration as cmrun, (suffixed with the applicable platform extension for platforms that use directories). You can use these samples as the base to create your own scripts, as described below.
After a flow is run with CMRUN, a return code tells the calling program whether the flow executed successfully. For details on these return codes, see CMRUN Return Codes on page 703.

For example, on Windows, CMRUN is called from the DataMigrator installation directory, which is by default:

`drive:\ibi\srv77\home\bin\cmrun.exe`

**Note:** An iWay communication file (odin.cfg) must be configured for the DataMigrator Server.

**Syntax:** How to Start Flows With CMRUN

```
CMRUN Server=server, [User=userid [, Password=password] , x=filename],
Request=request [, Appname=appname] [, Wait=minutes] [, I=seconds]
   [, V=variable] [, B=begin] [, E=end] [, Multi= Y/N]
```

where:

**server**

Is the name of the DataMigrator Server. This value is required.

**Note:** The maximum length is 16 characters.

**userid**

Is a valid user ID for the server. It is required for a secured server.

**password**

Is the password associated with the user ID. Enclose a password in double quotes if it contains special characters.

**Note:** Password and filename parameters are mutually exclusive.

**filename**

Is a physical file name, full or relative pathed, of a file containing the server user ID password. On z/OS, the value is not the direct dataset name, but a DD name issued in the z/OS CMRUN JCL to point at the dataset name.

On other platforms, including Windows and UNIX, it is specified as a file name in the current directory, or as the relative or full path to a file name.

**request**

Is the name of the flow, which can be up to 32 characters long.

**appname**

Is the application directory where the flow is located.
**minutes**

Is the number of minutes the server waits for a flow to complete. The default is 0, which specifies not to wait.

**seconds**

Is the number of seconds the server waits before checking if the flow has completed. The default is 15 seconds.

**variable**

The syntax for the local variable parameter for the aforementioned operating systems:

```
variable="VAR1=value1, VAR2=value2..."
```

The syntax for the global variable parameter differs for different operating systems:

**For MVS:**

```
variable="&&VAR1=value1, &&VAR2=value2..."
```

**For Windows:**

Example 1:
```
variable="&&VAR1=value1, &&VAR2=value2..."
```

Example 2:
```
variable="^^VAR1=value1, ^^VAR2=value2..."
```

**Note:** The CALL command should not be for executing a batch file when this syntax used, as CALL does not support caret ^ escape character.

**For UNIX:**

```
variable="\&\&VAR1=value1, \&\&VAR2=value2..."
```

**For System i:**

```
variable="\&\&VAR1=value1, \&\&VAR2=value2..."
```

**For OpenVMS:**

```
variable="&&VAR1=value1, &&VAR2=value2..."
```
where:

&&VAR1, &&VAR2

Are the names of the variables.

value1, value2

Are the values assigned to the variables.

begin

Is the record where loading starts. This is not available for flat files.

end

Is the maximum number of records to retrieve from the server.

multi

Indicates if multiple instances of the flow can run at one time. The default is YES, which means multiple instances can be run at one time. NO means only one instance can run at a time. If the flow is running and an attempt is made to submit a second instance, it fails with an error message:

ICM (18997) Request app/name did not start: another instance already

**Note:** A parameter may be abbreviated as its initial letter. Parameters must be separated by a comma and space. Parameter pairs (for example, Server=server) should never have spaces before or after the equal sign. Syntax may vary for different operating systems. Check the examples for specifics.

**Example:** Sample CMRUN Batch File for Windows

Before calling the CMRUN executable on Windows, you must set up certain environment variables in order for the executable to run properly. This is typically done in batch files that in turn call the CMRUN executable. You can also set the variables at the machine level, or your DOS session, so they do not need to be directly included in any batch files that you create. Additionally, the PATH variable is often adjusted to avoid using the full path name of the CMRUN executable. The following is an example of how to set these variables, using the typical default file install and configure locations, in addition to setting the PATH variable:

```
set EDAHOME=C:\ibi\srv77\home
set EDACONF=C:\ibi\srv77\cln
set EDACS3=C:\ibi\srv77\cln\etc\odin.cfg
set PATH=C:\ibi\srv77\home\bin;%PATH%
```

The following is an example of a Windows batch file that calls CMRUN:
Example: Running a Flow Using CMRUN Without Waiting for Completion

In the following example, CMRUN is called and DataMigrator does not wait for the procedure to complete:

Run ( 1) Connecting to 'demo' as 'dm'
Run ( 2) DM Manager Server: Version 7, Release 7
ICM (18762) Job ID: 20171127113718_698b58d6

Example: Running a Flow Using CMRUN and Waiting for Completion

In the following example, CMRUN is called and DataMigrator waits for the procedure to complete. The procedure is executed successfully.

Run ( 1) Connecting to 'demo' as 'dm'
Run ( 2) DM Manager Server: Version 7, Release 7
(ICM18016) Request appdir/flowname submitted. Please, wait for request to complete.

At completion, the following messages display:

(ICM18762) Job ID: 20171127113718_698b58d6
(ICM18764) Request appdir/flowname complete
In the following example, CMRUN is called and DataMigrator waits for the procedure to complete. The request fails.

Run (1) Connecting to 'demo' as 'dm'
DataMigrator Server: Version 7, Release 7

When the request fails, the following messages display:

ICM (18764) Request decode failed; RC = 350

Example: **Sample CMRUN Shell Script for HP UNIX (HPUX)**

The following is an example of UNIX shell script that calls CMRUN:

```
#!/bin/ksh
export EDAHOME=/home/ibi/srv77/home
export EDACONF=/home/ibi/srv77/dm
export SHLIB_PATH=$EDAHOME/bin:/usr/lib:/lib
export EDACS3=$EDACONF/etc/odin.cfg
$EDAHOME/bin/cmrun.out s=EDASERV, r=runtest, u=edachk, password=pass,
w=10, b=2, e=5, v="\&\&v1=x,\&\&v2=y"
```

Example: **Sample CMRUN Shell Script for SUN Solaris**

The following is an example of SUN Solaris shell script that calls CMRUN:

```
#!/bin/ksh
export EDAHOME=/home/ibi/srv77/home
export EDACONF=/home/ibi/srv77/dm
export LD_LIBRARY_PATH=$EDAHOME/bin:/user/lib:/lib
export EDACS3=$EDACONF/etc/odin.cfg
$EDAHOME/bin/cmrun.out s=EDASERV, r=runtest, u=edachk, password=pass,
w=10, b=2, e=5, v="\&\&v1=x,\&\&v2=y"
```

Example: **Sample CMRUN Shell Script for AIX**

The following is an example of AIX shell script that calls CMRUN:

```
#!/bin/ksh
export EDAHOME=/home/ibi/srv77/home
export EDACONF=/home/ibi/srv77/dm
export LIB_PATH=$EDAHOME/bin:/user/lib:/lib
export EDACS3=$EDACONF/etc/odin.cfg
$EDAHOME/bin/cmrun.out s=EDASERV, r=runtest, u=edachk, password=pass,
w=10, b=2, e=5, v="\&\&v1=x,\&\&v2=y"
```

Example: **Sample CMRUN JCL Batch File for USS Deployment**

The following is an example of JCL that calls CMRUN under USS Deployment:
//CMRUN JOB (CMRUN),CLASS=A,MSGCLASS=Q,MSGLEVEL=(1,1),
// NOTIFY=DMUSER
//*
//* This JCL can be used to run DataMigrator jobs
//* You will need to provide the following information under
//* the EDAPARM DDname
//* U= Defaults to current userid
//* P= Password of above userid
//* R= Name of DataMigrator flow to run
//* W= Wait time
//******************************************************************************
//CMASAP EXEC PGM=CMRUN
//STEPLIB DD DISP=SHR,DSN=qualif.SRV77.HOME.LOAD
//EDAPARM DD *
S=LOOPBACK, USER=edachk, p=passwd, R=request1, APPNAME=baseapp, W=1,
V="&&VAR1=©abc©,&&VAR2=123"
//EDAENV DD *
EDACONF=/home/ibi/srv77/dm
EDAHOME=/home/ibi/srv77/home
EDACS3=/home/ibi/odin.cfg
PATH=/bin
//SYSIN DD DUMMY
//STDOUT DD SYSOUT=A
//STDERR DD SYSOUT=A
//SYSPRINT DD SYSOUT=A

Note: To enable traces for CMRUN, pass the -traceon parameter with EDAPARM ddname. For example:

//EDAPARM DD *
-traceon
S=LOOPBACK, USER=edachk,, p=passwd,, R=request1, APPNAME=baseapp, W=60

As a result, the cmrun.trc and tscom3.trc files will be created in your USS user directory.

Example: Sample CMRUN JCL Batch File for PDS Deployment

The following is an example of JCL that calls CMRUN under PDS Deployment:
**Example:** Sample CMRUN Code for OpenVMS

The stand-alone program for initiating requests can be used on OpenVMS. The following is an example of a command procedure that calls CMRUN:
Example: Sample CMRUN Code for IBM iSeries

The following is an example of a shell script that calls CMRUN on IBM iSeries:

```bash
#!/bin/qsh
export EDAHOME=/ibi/srv77/home
export EDACONF=/ibi/srv77/dm
export EDACS3=/ibi/odin.cfg
liblist -a srv77
$EDAHOME/bin/cmrun.out s=EDASERV, r=request1, u=edachk, p=passwd, APPNAME=baseapp, W=60

where:

srv77

Is the library name for edahome provided during installation

Reference: CMRUN Return Codes

When a flow is completed, CMRUN sets a return code that can be viewed in the calling environment. These codes are listed and described in the following table:

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>The value of the 0 code depends on the Wait parameter setting:</td>
</tr>
<tr>
<td></td>
<td>- If set to zero, the request was submitted successfully.</td>
</tr>
<tr>
<td></td>
<td>- If set to a number greater than zero, the request completed successfully.</td>
</tr>
<tr>
<td>4</td>
<td>Invalid input parameter.</td>
</tr>
<tr>
<td>8</td>
<td>iWay API returned an error.</td>
</tr>
<tr>
<td>12</td>
<td>Request submission failed because the flow was not found, was incomplete, or the user was not authorized.</td>
</tr>
</tbody>
</table>
### Code Description

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>The request failed during execution. This return code may occur only when the Wait parameter is set to a number greater than zero.</td>
</tr>
<tr>
<td>20</td>
<td>The request timed out. This return code may occur only when the Wait parameter is set to a number greater than zero.</td>
</tr>
</tbody>
</table>

A return code of 8 indicates that the iWay API returned an error. This error message is displayed by CMRUN. The most common errors are:

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
</table>
| -7   | Unable to initialize Connector.  
This status code indicates that the API was either unable to find or unable to read the communications configuration file. Check that the location of the correct configuration is the first one accessible using the PATH, and ensure that the contents are readable and correct |
| -9   | Unable to connect to server.  
This return code generally indicates either that the server or communications resources are unavailable, or that there is an error in the configuration. Ensure that the resources indicated in the Connector configuration file are correct and available. Ensure that the server name indicated is correct and that it matches the service name on the server (for non-UNIX servers). Ensure that you can ping the host if TCP/IP is the protocol. Check to see if a connection to the server is possible from an alternate client. If this is successful, either the client configuration is incorrect or the communications resources are not available from that location. If this is not successful, ensure that the server is up and running and that it is not in an error state. |
| -12  | Data receive failure.  
This status code indicates that a problem was encountered while the client was attempting to receive data from the server after an initial successful connection. Check the server status after receiving this error and check that the communications resources are intact. A -12 may occur for a specific request or for specific data retrieval. |
<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
</table>
| -17  | Invalid user ID or password.  
This status code indicates that a connect attempt was made to a secure server and that the user ID and/or password details supplied were not authorized for that server. A -17 may also occur if the user ID and password do not pass security exit checks set up on the server.  
**Note:** Some servers are case-sensitive. |
| -20  | Expired password.  
This status code is issued if a user attempts to connect to the server and the password is expired. |

A flow may fail during connection to the server or while being run. For details on the problems that may cause a failed flow, see *Problems, Errors, and Troubleshooting* on page 709.

**Specifying an External Password**

On non z/OS platforms, the X= parameter is used to point to a physical file, full or relative pathed, that contains the password for the CMRUN execution, which is protected by operating system security (read only to owner). On the z/OS platform, JCL and the X= parameter are used to point to a DD name that points to an external dataset that contains the password, which in turn can be RACF or Top Secret protected.

The file, or z/OS dataset, containing the password should contain a single line with the server password of the user ID. Passwords are case-sensitive and terminated by the first blank space or the end of the line.  
**Note:** A password cannot contain blank spaces.

On z/OS, the dataset may be variable or fixed length, but must be of a sufficient LRECL for the actual password. Releases prior to 7706M required the z/OS dataset to be specifically LRECL=80 and RECFM=FB, but this restriction has been lifted in a way that prior password datasets also remain usable as is.

**Example:**  
**Specifying an External Password With SYSIN**

The following is a snip of the z/OS JCL change that is required for using an external password.
Starting Flows With CMASAP

The CMASAP procedure enables event-oriented execution from an API program or FOCUS, and is provided with DataMigrator. It also allows you to specify values for user-defined variables.

If you need to pass parameters into an && variable, the variable must be set up with a -DEFAULT before you can use it in the execution line. For example:

```
-DEFAULT &&ODATE='&&ODATE'
EX CMASAP REQ_NAME=DMVAR, CM_ASYNC=OFF, PARMS="&&ODATE=ODATE"
```

**Syntax:** How to Start a Flow With CMASAP

```
EX CMASAP_REQ_NAME=name[,APPDIR=appdir][,STOPAT=stopat][,STARTAT=startat]
    [,CM_SETERR={ON|OFF}] [,CM_ASYNC={ON|OFF}]
    [,PARMS="&VAR1=value1, &VAR2=value2..."]
    [,CM_MULTI={ON|OFF}]
```

where:

**appid**

Is the name of the application directory, which can be up to 64 characters long.

**name**

Is the name of the flow, up to 64 characters long.

**startat**

Is the record to start loading at. This is not available for flat files.

**stopat**

Is the maximum number of records (rows) to retrieve from the Server. RECORDLIMIT and READLIMIT are passed to the Server.
CM_SETERR

Determines the types of messages returned to the calling program. Valid values are:

ON

CMASAP sends a confirmation message and administrative messages that are produced as CMASAP prepares a flow for execution. This is the default.

OFF

CMASAP sends only the confirmation or failure message.

When putting an event-oriented program into production, set CM_SETERR to OFF, so that the calling program only needs to check one message when the server processes the flow.

CM_ASYNC

Determines if the server waits for the flow to complete. Valid values are:

ON

CMASAP relinquishes control to the server and terminates when the request is submitted. This is the default.

OFF

CMASAP waits for the flow to complete. This setting is useful for prototyping or for anytime you want to be held in a wait state until the entire flow has finished.

&&VAR1, &&VAR2

Are the names of the variables.

value1, value2

Are the values assigned to the variables.

CM_MULTI

Indicates if multiple instances of the flow can run at one time. Valid values are:

ON

Multiple instances can be run. This is the default.

OFF

Only one instance can be run at a time. If a second instances is attempted, it fails with the following error:

ICM (18997) Request app/name did not start: another instance already
**Note:** When calling this procedure from within a WebFOCUS procedure, a single pipe symbol (|) should be placed before each pair of ampersands. For example:

```
[,PARMS="|&VAR1=value1, &VAR2=value2..."]
```

### Starting Flows From FOCUS

A FOCUS or WebFOCUS client on any platform is able to start flows. Notification of errors in execution will be returned to your application as error messages.

#### Syntax: How to Start Flows With a FOCUS or WebFOCUS Client

```
EX CMASAP parameters
```

where:

- **parameters**

  Are CMASAP parameters. For details, see *How to Start a Flow With CMASAP* on page 706.

### Determining the Status of an Event-Based Flow

When you schedule an event-based execution, the flow is scheduled for immediate execution. If resources are available, the flow will receive its run-time parameters from the DataMigrator internal data store and begin its extraction and load as designed.

Submit the request with CMRUN. It sets a return code to the calling environment indicating if the request was submitted or completed successfully, depending on the setting for the Wait parameter.

You can also determine the status of a flow after it has been submitted. This can be done in the following ways.

- **Run the DataMigrator ETL Log (details) report.** This allows you to view the ETL Log file, which contains detailed information concerning the status of any given request. For details on the DataMigrator ETL Log (details) report, see *Generating Reports, Viewing the Log File, and Printing Flows* on page 467.

- **Run the Statistics report and check the Server statistics.**

- **Check the Web Console Workspace.** Start the Web Console and select *Agents* from the Monitor/Tune menu.
This topic assists you in resolving problems that may occur during the development of your DataMigrator application.

**In this appendix:**

- DataMigrator Status Codes and Server Messages
- Server Run-Time Error Codes
- DataMigrator Management Problem Resolution
- Including Run-Time Messages in the ETL Log
- Performing a Trace to Diagnose Server Problems

---

**DataMigrator Status Codes and Server Messages**

Error and server messages are displayed in the Message window.

The following types of messages are displayed:

- **iWay Connector communication errors.** These error conditions usually indicate a problem communicating with the server or network.

- **Internal Data Management Console errors.** If these errors are self-explanatory, you are given the option of choosing OK or Cancel. If not, the error may indicate problems in memory or with the environment in which the DMC is operating.

- **Messages from the Server.** The most common messages are those that are sent by the Server. These will display in the message window of the DMC. Along with descriptive text, there will be a message origin and a message code.

The message typically begins with the following:

\[(XXXnnnn)\]

where:

- **XXX**

  Is the message origin.
**nnnn**

Is the message code.

Messages of this type may come directly from the iWay subserver during the retrieval of catalog information or rows from a source table, or from the Server as administrative messages.

Common codes of the above three types are described in the following topics. Also, see the appropriate iWay Server documentation for your platform and the Server codes at the end of this topic.

**Procedure: How to Get More Information About Messages**

The *iWay Messages and Codes* manual and online help document contains a list of all message numbers with a corresponding an explanation. You can also get this information directly from the DMC.

In the DMC:

1. Right-click on a server name and click *Test Server*.
2. In the section Choose Query Type, select the radio button for Focus.
3. In the procedure, type a question mark (?) followed by the message number. Then click *Test*.

The message text and explanation appear in the Console Log area.

**Server Messages**

Some common server messages are listed below:

*(FOC054)  EDA Security error*

There is security on the synonym, and the password has not been set correctly. This should be done in the EDASPROF or the user profile for the server.

*(FOC14010) -column-name- IS NOT A VALID COLUMN NAME*

DMC issues a PREPARE for the SQL command, and the column name does not match the synonym. If you have edited the synonym at the source, you may need to create the Synonym with Automatic Passthru turned off.

*(FOC1006)  THE DDNAME REQUESTED IS NOT ALLOCATED OR DLBL'ED : table*
*(FOC226)  FILEDEF OR ALLOC MISSING FOR EXTERNAL FILE: table*
*(FOC351)  INPUT FILE NOT ALLOWED: file*
*(FOC1400)  SQLCODE IS : 1006/000003EE*
*(FOC1405)  SQL PREPARE ERROR*
These messages, or some combination of them, can display when an Synonym for a source file exists but the file it describes is not allocated on the server. These errors can occur upon exiting from the Query Builder or upon an attempt to run a Test Query. They can also occur when using Metadata Manager or some other tool to create a synonym.

(FOC1538) 4GL procedure doesn't exist: proc

One of the procedures required by the Server could not be found. This may indicate that a new version of the server is needed for full compatibility with the client version being used.

(ICM18986) Request application/name could not be found or accessed

This error occurs when a flow is submitted or scheduled and it can not be found in the application directory specified. Check that it exists, that your user ID can see it, and that the name is in all lower case.

(ICM18708) 0 rows retrieved. Stop processing.

This message means that the flow execution properties Stop if 0 rows selected is checked and zero rows were returned from the query or loaded to the target. This could indicate an error if no data was available or it could be that there were no rows that met a WHERE condition.

(FOC35804) BAD STATUS CODE (-9) RECEIVED FROM SERVER subserve

The Server was unable to connect to a subserver listed in the server configuration file (ODIN.CFG).

(FOC541) SU. CENTRAL DATABASE MACHINE ERROR: FOCSU01

The FDS service is down and the Scheduler is attempting to write in the ETTLOG.

Starting from Release 7.7.07, the maximum length allowed for the Application Path increased from 64 characters to 256 characters.

An example of an application path is: baseapp/app01etl/appwarehouse/history/staging.

The length of the path should not exceed 256 characters.

To enable the increased length of the Application Path, users upgrading DataMigrator to the Release 7.7.07 must perform a Recreate operation for the Log and Statistics tables.

The following errors may occur when working with the application path:

(FOC32505) WRONG SEGMENT LENGTH DETECTED: STATS

The Log and Statistics tables are inherited from the older release and were not recreated.

(FOC015) THE TEST VALUE IS LONGER THAN THE FIELD FORMAT LENGTH
Server Log Messages for Agent Termination

Server log messages for agent termination are listed below:

(ICM18175) Agent terminated by operator.

The agent was killed by an administrator or operator.

(ICM18174) Agent has been forced to terminate.

The agent was hard killed or crashed.

Server Log Message for a Missed Scheduled Request

The following is an example of a server message that is written if a request with a recurring schedule misses a run:

01/03/2010 18:04:37 E DM: procedure "baseapp/flow02" missed a scheduled run

To detect the run of the missed request, the Scheduler must be active and a custom routing event must be created. There is no filter needed for this event. Based on the Event Routing configuration, the server log message is written into a server log (edaprint.log) or an email message.

iWay Connector Communications Errors

If there is a problem connecting to a Server, a message is displayed in the message window. The following are some of the more common errors:

Error (12029) – A connection with the server could not be established.

Server is not running or wrong port number specified. This error commonly occurs if the server has not started up yet. It could also occur if the host name or port number was incorrect.

Error (12152) – The server returned an invalid or unrecognized response.

This error occurs when something other than a DataMigrator http server is listening on the specified port number. For example, if a port number for TCP or FTP server was specified.

Error (32034) – Authentication failed, password is invalid.

This error could occur if either the user ID or the password specified were incorrect.

Error (32033) -- user ID is invalid Disconnected from server

The user ID or password that you entered was invalid for the selected server.
error <NGS> failed to connect to IP address:port, 10061Connection refused

Error (32007) communication error detected

The selected server was unavailable; it may not be running. The iWay Connector configuration (ODIN.CFG) may be incorrect.

Server Run-Time Error Codes

The errors listed here will usually display in the ETL Log unless otherwise noted. Several of these are formal Server error codes. You can see a complete list in *Server Messages* on page 710. Other errors are related to the iWay environment, operating system, or specific data source system being accessed. Additional references from these resources may be necessary.

FOC282 -or EDA282 RESULT OF EXPRESSION IS NOT COMPATIBLE WITH...

This error generally occurs on a transformation string or on a DEFINE in a source synonym. It means that the answer for a calculated field cannot be placed into the column name on the left-hand side of the equal sign. Typically, this is caused by an attempt to move an alpha field directly into a numeric field (or vice versa), or a character string into a column whose format is less than the length of the desired string. For details on this and other errors in this category, see the appropriate FOCUS documentation for your server platform.

FOC14010 -column-name- IS NOT A VALID COLUMN NAME

DataMigrator issued a prepare for the SQL command, and the column name does not match the synonym. The synonym may have changed since the flow was designed. For Server Transfer Files, another possibility is that the Synonym at the Server does not match the synonym at the source. For more information, see *Designing a Data Flow* on page 193. If you have edited the synonym at the source, you may need to create the Synonym with Automatic Passthru turned off.

FOC35804 BAD STATUS CODE (-12) RECEIVED FROM SERVER server
FOC35862 EDA STATUS (-12): ERROR RECEIVING DATA FROM SERVER

These errors indicate that something has gone wrong at the server identified in the message.

One common reason for this error is that the server is running out of disk space. Check the location of the NT or UNIX temporary directory, CMS temporary minidisk, or z/OS temporary data set allocation.

On z/OS, another possibility is that the data sets allocated to EDASYNA and EDASYNM are full. If so, delete unnecessary Master and Access numbers, and compress the data sets.
DataMigrator Management Problem Resolution

Situational errors sometimes occur in environments with complex iWay configurations or multiple users. The list below is meant to provide you with error messages similar to what you may encounter, and offer you resolutions such as you may find in DataMigrator. Each message is characterized by a single symptom, followed by an explanation or references to other materials.

**Unique Constraint Error... Duplicate keys not permitted... or similar error for any of the relational database targets.** DataMigrator has duplicate values in its extracted answer set, and there is a unique key on the table. For new targets, a unique key value may need to be specified in the Target Setup window. For existing targets, another column may need to be added to a concatenated key.

This error may also be caused by the retrieval mechanism: the use of the DISTINCT operator in the SQL command may really be the desired result (more rows than expected are being retrieved), or the source data may be a complex hierarchy or join that is creating duplicate rows. See the usage notes for your specific data target.

**Format errors caused by dirty or invalid data.** Legacy data is not always consistent in format and type. Older systems may have stored standard internal packed numbers for revenue figures but handled special logic conditions by inserting characters or low-values as needed by the application program. These issues often require handling with DEFINEs in the synonym at the source. There is a complete set of subroutines to handle various bit and byte oddities in your host information. For more information, see *Designing a Data Flow* on page 193, and the appropriate FOCUS documentation for your platform. Problems of this type usually make themselves known through data exceptions and other error conditions (these would display in the ETL log). Depending on the data type combinations, they may also result in rejected records or a misinterpretation of missing data.

**Confusing answer set; too many rows as the result of a Cartesian product.** The rules of SQL control the join process for the source data to DataMigrator. Many-to-many joins result in a Cartesian product. This may be what the warehouse requires, but very often it is confusing because it results in additional rows being generated. The Cartesian product may also result from the extraction of columns that cross two paths of a multi-path legacy hierarchy. If this occurs, reduce the number of joins, or access only one path of a complex hierarchy. For details on the Cartesian product, see the appropriate iWay Server documentation for your platform.
Including Run-Time Messages in the ETL Log

The Dialogue Manager command -TYPE can be used to display messages. This command can be included in a procedure. If this command is used in your procedure, the text will display in the ETL Log File along with all other status and administrative messages for a flow. The text will receive a time stamp as recorded by the previous formal administrative message that displayed in the log.

**Syntax:** How to Generate a Message Including a Time Stamp

Use the following syntax inside your procedure every time you want a message with a unique time ID:

```
-TYPE Messages from procedure
```

Performing a Trace to Diagnose Server Problems

When you receive a server error message, you can perform a trace to find the source of the problem.

**Procedure:** How to Create a Server Diagnostic Package

1. In the *iWay Software, iWay 77 DataMigrator Server* program folder, click Start Security ON/OFF With Traces from the Diagnostic Functions folder to start the server with traces.
2. Reproduce the problem you encountered.
3. In the *iWay Software, iWay 77 DataMigrator Server* program folder, click Stop.
4. In the *iWay Software, iWay 77 DataMigrator Server* program folder, select Diagnostic Functions, and then click Save Diagnostic. Reply to the prompts to specify a location to save diagnostics and generate a list of the application directories that contain the files used to reproduce the problem.
   
   A diagnostics archive is created with a default name of sdyymmdd_hhmms.zip.
5. Upload the log file to Information Builders.

**Procedure:** How to Enable or Disable Traces on a Running Server

For reporting issues to Information Builders, the complete diagnostics package described in the previous procedure is necessary. However, for diagnosing your own problem, running a single trace may be useful.

1. Using an Administrator User ID, expand Servers and then your server.
2. Expand the *Workspace* folder and then expand the *Logs and Traces* folder.
3. Right-click Traces and select Enable Traces or Disable Traces.

Procedure: How to View Traces

1. Using an Administrator User ID, expand Servers and then your server.
2. Expand the Workspace folder and then expand the Logs and Traces folder.
3. Right-click Traces and select View.
   The Traces window opens in the workspace.
4. Double-click the desired trace to open it.
   The trace opens in a new tab.

To view another trace, click the Traces tab and select the trace you wish to view.
Customizing Your Environment

You can control your environment by using the SET command with parameters.

The SET command is typically included in a user profile for the server or in the global profile, EDASPROF.

In this appendix:

- DataMigrator Parameters

DataMigrator Parameters

The following are some of the parameters that affect DataMigrator:

**MSG={ON|OFF}**

Turns on iWay messaging for the log.

**JOINTYPE={SORTMERGE|NESTEDLOOP}**

Affects SORTMERGE or NESTEDLOOP logic.

**JOINOPT={NEW|OLD|GNTINT}**

Affects how SQL is handed to remote servers.

**CDN={ON|OFF}**

Turns on continental decimal notation for flat file copies.

**MISSINGTEST={SPECIAL|OLD}**

Controls whether NULL values in a comparison affect the nullability of the result.

Syntax: How to Set Parameters in a Profile or Procedure

```
SET parameter = option
```

where:

- **parameter**
  
  Is the setting you wish to change.

- **option**
  
  Is a valid value for the parameter.
Example: Setting a Parameter in a Profile or Procedure

The following sets the CDN SET parameter to ON.

```
SET CDN = ON
```

Syntax: How to Set Target Columns as NULLABLE

When new target tables are created, by default, the resulting target columns inherit the NULLABLE setting of their corresponding source columns.

To set the nullability of target columns, use the following syntax:

```
SET TARGET_ATTR=ON|PROPAGATE
```

where:

ON

Makes all non-key target columns NULLABLE (MISSING=ON).

PROPAGATE

All non-key target columns inherit the nullability of their source columns. This is the default.

This setting can be used in a profile, where it will affect all new target tables.

Syntax: How to Limit the Number of DBMS Errors That Appear in the Log Table

When DataMigrator attempts to insert rows into a target relational database table, they may be rejected due to a unique index, not null, or other constraint violations. When rows are rejected, the error messages are written to the Log Table. This can fill up the log, and it may not be necessary to see the message for every rejected row to determine the problem.

The maximum number of errors written to the log can be restricted with the following command:

```
SET DBMSMSGLIMIT=nnnn
```

where

```
nnnn
```

Is the maximum number of errors.

**Note:** The processing of the job continues, although further errors are not logged.
This setting can be used in a profile, where it will affect all new target tables, or in a stored procedure that is executed before the data flow loads the target table.

**Syntax:** How to Set the Source for Oracle Timestamps

The DATETIME_PROCESS setting allows you to specify the source for the current date and time in SQL SELECT statements from Oracle tables using CURRENT_TIMESTAMP(0).

The syntax is:

```
SQL SET DATETIME_PROCESS=SERVER|DBMS
END
```

where:

**SERVER**

Obtains the current date and time from the iWay server. This is the default.

**DBMS**

Obtains the current date and time information from the relational database. This is only effective for queries that can use automatic pass through and only those that use the function CURRENT_TIMESTAMP(0).

This syntax should be used even if it is included in a profile, which affects all queries run on the server.

If the query cannot use APT (for example, if the source is a flat file) the following message is generated:

```
(FOC14070) Query cannot go APT to evaluate CURRENT_TIMESTAMP
```

**Procedure:** How to Use Core Engine Settings

Selected server settings can be set from the Data Management Console Workspace Configuration.

1. In the DMC, using an Administrator user ID, right-click the Workspace folder, point to Settings, and click Core Engine Settings.
The Change Core Engine Settings dialog box opens, as shown in the following image.

2. Select edasprof from the Select profile drop-down menu to make changes for all users. You can also select a single user ID.
3. Enter or select new values for the rest of the applicable fields.
4. Click Save.
Source Management

You can use source control programs to manage flows, stored procedures, and synonyms, typically providing file backup, change tracking, and version control. They also allow multiple users to work on the same files without overwriting changes from others. A source control program typically manages access to the source code and documents all code changes. It also typically maintains a separate version of your file each time it is checked in.

In this appendix:

- Overview of Source Management
- Source Control Security Server and Client Setup
- Configuring Source Management and Setting Up Source Control
- Using a Source Control System to Manage Procedures and Synonyms

Overview of Source Management

DataMigrator supports third-party source management products that use the industry-standard SCC API and a database for source control. The Data Management Console provides a direct connection to these source control programs. DataMigrator is compatible with any source control product that is SCC-compliant, as long as it has a source control database. This manual describes the procedures for configuring seven popular source management programs: Serena ChangeMan DS, IBM Rational ClearCase, Concurrent Versions System (CVS), Perforce SCM, Subversion, Team Foundation Server, and Microsoft Visual SourceSafe (VSS) and Team Foundation Services.

DataMigrator provides access to basic source management functions through right-click menu options in the navigation pane which only appear if you have a source control system configured.
Source control systems consist of a repository that holds a master copy of each revision. The local working copies of these files are stored on the DM server inside the user home directory. These files are the files changed by developers, before the file is committed or checked-in to the source control repository, which creates the next revision. DataMigrator can add synonyms, procedures, and flows to source control by exporting ACCESS, MASTER, FCEXEC, and ETG files, respectively. Intermediate files are created in the user home directory of the developer. It is the role of the manager to create a user home directory for each developer that will contain the local working copies. When a developer checks out a flow, DataMigrator sends a command through SCC API plug-in or by invoking client command line utility to source control to check out the file. The source control system then puts the latest version of the file into the user home directory of the developer. All other functions, such as add, check in, undo check out, view history, and get latest version functions, are invoked the same way.

The intermediate work space of the manager is the actual central application directory. Only the manager has write privileges to central application directory. The manager can deploy the latest files from source control repository to the central application directory by invoking get latest version. The following diagram shows the model of the workflow.
Source Control Security Server and Client Setup

For source management solution, DataMigrator allows SCC control at the client side and stores home directories on the server side. Each developer will have a user home directory, which serves as a storage location for ongoing work. User home directories will be read-write for their owner with no access by any other developer, which prevents overwriting code among multiple developers. In addition, to prevent circumventing the source code, central application directories allow read-only privileges for all developers, while write access is granted only to the manager. The manager will create user home directories and deploy latest files from source control management system to the central application directories.

The descriptions and roles of these two positions are:

**Developer**

Create and edit flows. Their role is that of Application Administrator. By default, developers are assigned to Application Administrator roles. However, more or less privileges can be granted by the Server Administrator.

**Manager**

Create user home directories for all developers and deploy latest files from source control management system to the central application directories. Their role is that of Server Administrator. By default, the manager is assigned to Server Administrator. However, you can create different roles so that manager can deploy files, but cannot stop and start the server or change adapters.

**Note:** In the procedures below, Configure Server Security, Create and Configure User Home Directory, Register Users, and Assign Access Control are completed by the Manager. Log In Data Management Console with Security On and Associate User Home Directory with Source Control Project are completed by all Developers.

It is critical that you follow the procedures carefully and in their proper sequence.

**Procedure:** How to Configure Server Security

Before starting, the manager must set up server security so that developers are required to login.

1. From the DMC, expand the Access Control folder.
2. Right-click the Security Providers folder and click Change Provider.
   
   The Change Provider dialog box opens.
3. Set the Security Mode drop-down menu to the desired authentication type.
4. Enter your password and click Apply and Stop Server.
5. Restart the server.

**Procedure: How to Create and Configure User Home Directory**

Once the manager logs in to the DMC with DM Security On, the next step is to create user home directory, also referred to as homeapps. Each developer will possess write access to their home directory inside homeapps.

1. From the DMC, right-click the Workspace folder, point to Workspace SET, and then click Application Settings.
   
   The Application Settings dialog box opens.

2. Enable nested_app by setting it to the Y value.

3. Enter the appropriate homeapps path. The homeapps path should be created under the DataMigrator root, ibi directory, as shown in the following image.

   ![Application Settings Dialog Box](image)

4. Click Save and Restart Server.

5. Right-click Users Home and select Create Disk.
Procedure: **How to Register Users**

The manager registers new developers and assigns access roles from the Web Console.

1. From the DMC, expand the *Access Control* folder.
2. Expand the *Roles* folder, right-click *Application Administrator*, and then click *Register User*.
3. Enter the appropriate user or developer name and their credentials in the User Registration dialog box.
   
   **Note:** A user or developer name must be defined in Windows Local Users and Groups if using OPSYS authentication type on Windows. A similar rule would apply to another authentication type.
4. Click *Register*.
   
   An Option dialog box will appear to indicate that the new user will be registered. Click *OK*.

Procedure: **How to Assign Access Control**

The default role setting for developers is that of Application Administrator. Each developer will possess write access to their home directory and read/list access to central application directories.

In the DMC:

1. Right-click an application directory and click *Privileges*.
   
   The Manage Privileges dialog box opens.
2. Right-click the role you wish to set control for, and click *Edit Privileges*.
   
   The File/Directory Privileges window opens.
3. Limit File Access options to Read files for AREAD and List files for ALIST, as shown in the following image.

![Image of File/Directory Privileges]

4. Click Apply.

Procedure: How to Log Onto Data Management Console With Security On

Developers log onto the Data Management Console with proper authentication.

1. Launch the Data Management Console.
2. Right-click a Server node and click Properties.
3. Log in with the appropriate User ID and Password that was created and registered by the manager. Click OK.
4. Right-click the Server node and click Disconnect and then Connect to connect through the new security setting.

Procedure: How to Associate User Home Directory With Source Control Project

Each developer will have a user home directory, which serves as a storage location for ongoing work. All developers must copy and paste the application directories into their user home directory.

1. In the DMC, right-click Application Directories and click Create My Home. My Home [developer name] directory appears.
2. Right-click the application directory you want to add to the source control, click Copy, and then paste it in the user home directory.

Repeat this step for all application directories you want to add to source control.

**Configuring Source Management and Setting Up Source Control**

If you have a source control system configured, you can use it to source manage DataMigrator procedures, synonyms, and flows.

To implement source control, all developers and the manager must:

1. Set up source control folders in your source management system. These folders are called Projects in ChangeMan DS, Perforce, and Visual SourceSafe, Views in ClearCase, and Modules in CVS.

   **Note:** Source control folders work essentially like DataMigrator application directories. You should create folders in your source control system that mimic your node/application directory structure. Creating folders using your server node names will help you distinguish application directories with the same names on different servers.

2. Configure DataMigrator to work with your source control system.

3. Create associations between DataMigrator application directories and the source control folders you created.

You can then add files to your source control system, check them out or in, and see file history, versions, and differences from the Data Management Console.

For detailed information on installing and configuring your source control system, and setting up source control databases and projects/views/modules, see your source control system documentation.

**Setup, Configure, and Associate DataMigrator to Work With ChangeMan DS**

Creating a ChangeMan DS project provides a way to logically associate files with a project. Files are associated to the project by attaching them when the project is created.

**Note:** Source control folders, called projects in ChangeMan DS, work essentially like DataMigrator application directories. ChangeMan DS projects can be created during configuration or when you create associations with application directories.

**Procedure:** How to Create ChangeMan DS Projects

1. Log into ChangeMan DS.

2. From the Tools menu, click **System Configuration**.

   The Global Process Control window opens.
3. From the Edit menu, click Users.
   Select New User or New Group from the bottom left of the Global User Configuration window, and create and edit user and group profiles. Provide privileges to the users.

4. From the Edit menu of the Global Process Control window, click Areas.
   Select New from the bottom left of the Global Area Configuration window. Create and edit areas, such as development, production, QA, and end users. The development area usually represents a directory on the local computer where the development work takes place. Files may be checked in and out of this directory.

5. From the Edit menu of the Global Process Control window, click Projects.
   Select New from the bottom left of the Project Configuration window to create and edit projects. ChangeMan DS projects help to manage logically related files. Attach files to a project during configuration.
   **Note:** Currently, identical sources cannot be used for the same projects. This feature is planned for a future release.

6. From the Process Control Window of the new project that was created, click the New Rule icon located right under the File menu.
   Create a rule by adding the appropriate nodes and connecting the Development Area with the Production Area.

**Procedure:**  **How to Configure DataMigrator to Work With ChangeMan DS**

To use a source control system to source manage your DataMigrator flows, stored procedures, and synonyms, you must have a source control program and database configured. The source control program will then appear as an SCC Provider.

In the Data Management Console:

   **Note:** If you have not properly configured ChangeMan DS Client, this option will not appear in the Tools group.
   
   The Configure Source Control dialog box opens.

2. Select ChangeMan DS as the SSC Provider from the drop-down menu.
   **Note:** Selecting none removes source control from your system.

3. Optionally, click Test to make sure you can connect to the source control system from the DMC.
   **Note:** If you do not have a license for the Source Control, you might be prompted to get one at this time.
4. Click OK.
   The Set Source Control Provider dialog box opens.
5. Click Yes.
   Once a SSC Provider has been set, source control options appear on the right-click menus.

**Procedure:** How to Create Associations Between Application Directories and ChangeMan DS Projects

Creating a source control project provides a place to manage your DataMigrator files. Creating an association between your application directory and that project tells DataMigrator where to manage those files.

**Note:** Developers must associate application directories inside their user home directory, while only the manager associates to the central application directory.

In the navigation pane of the Data Management Console:

1. Right-click an application directory and select Source Control, and then click Create Association with Source Control.
   The Associate Application with Source Control Project window opens.
2. Click Associate.
   The ChangeMan DS Login dialog box opens. Type your User Name, Password, and MainServer for your ChangeMan DS environment.
3. Click OK.
   The SCC Project Path from ChangeMan DS dialog box opens.
4. Enter Local Area and Project and click OK.
5. Click OK to close the Associate Application with Source Control Project Dialog box.
   The Application Directory icon displays a lock to indicate the source control association.

At this point, you have associated your application in DataMigrator with a project in ChangeMan DS.

**Setup, Configure, and Associate DataMigrator to Work With ClearCase**

Source control folders, called views in ClearCase, work essentially like DataMigrator application directories. You should create folders in your source control system that mimic your node/application directory structure. Creating folders using your server node names will help you distinguish application directories with the same names on different servers.
Procedure: How to Create ClearCase Views

1. From the Start Menu, select Rational Software, Rational ClearCase, and then click Create View.
   The View Creation Wizard’s Choose a Project dialog box opens.

2. Click Next.
   The View Creation Wizard’s Choose Snapshot View or Dynamic View dialog box opens.

3. Select Dynamic and click Next.
   The View Creation Wizard’s Choose Name and Drive dialog box opens.

4. Use the default view name or enter an alternate name.
5. Select a drive. In this example, the default drive, Z, is used.
6. Enter a comment for the view.
7. Click Advanced Options.
   The Advanced View Options dialog box opens.
8. Enter the name of your ClearCase storage share and click OK.
9. When the Choose Name and Drive dialog box reopens, click Finish.
   The Confirm dialog box opens.
10. Click OK.
    The ClearCase View Tool dialog box opens.
11. Click OK.
    Once you have created a ClearCase view, you should navigate to its drive/folder and create subfolders that match the application directories that you want to source manage.

Procedure: How to Configure DataMigrator to Work With ClearCase

To use a source control system to source manage your DataMigrator flows, stored procedures, and synonyms, you must have a source control program and database configured. The source control program will then appear as an SCC Provider.

In the Data Management Console:
   Note: If you have not properly configured ClearCase Client, this option will not appear in the Tools group.
   The Configure Source Control dialog box opens.
2. Select ClearCase as the SSC Provider from the drop-down menu.
   **Note:** Selecting none removes source control from your system.

3. Optionally, click Test to make sure you can connect to the source control system from the DMC.
   **Note:** If you do not have a license for the Source Control, you might be prompted to get one at this time.

4. Click OK.
   The Set Source Control Provider dialog box opens.

5. Click Yes.
   Once a SSC Provider has been set, source control options appear on the right-click menus.

**Procedure:** **How to Create Associations Between Application Directories and ClearCase View Folders**

Creating a source control folder provides a place to manage your DataMigrator files. Creating an association between your application directory and that folder tells DataMigrator where to manage those files.

**Note:** Developers must associate application directories inside their user home directory, while only the manager associates to the central application directory.

1. Right-click an application directory, select Source Control, and then click Create Association with Source Control.
   The Associate Application with Source Control Project dialog box opens.

2. Click Associate.
   The Browse for Folder dialog box opens.

3. Select the folder that you want to associate with this application directory and click OK.

4. Click OK to close the Associate Application with Source Control Project dialog box.
   The Application Directory icon displays a lock to indicate the source control association.

You are now ready to use ClearCase for source management.

**Setup, Configure, and Associate DataMigrator to Work With Concurrent Versions System (CVS)**

CVS is an open source system widely used in the UNIX world. It has a Windows implementation called CVSNT. Since CVS is a command-line system, it cannot be used directly by DataMigrator (or any other application that uses the SCC API). To use CVSNT, the PushOK CVS SCC Proxy is needed to convert the SCC API calls into CVS commands. Source control folders, called modules in CVS, work essentially like DataMigrator application directories.
**Note:** The PushOK proxy is available from http://www.pushok.com as the CVS SCC Plug-in. Use the Windows version, not the x64 version, even if you are running 64-bit Windows.

**Procedure: How to Create Concurrent Versions System (CVS) Modules**

Before you can create CVS Modules, you need to create a repository and add users. A repository needs a root directory where both the files under version control and auxiliary files for use by CVS server are stored. After creating the repository, create users that can access (store and retrieve) files from the repository. Once this is complete, there are two ways to create modules:

1. You can add a new module through the command line. For instance, to add a new module with the name appl:

   ```
   mkdir $CVSROOT/app1  # makes an empty module.
   cvs checkout -d ./app1 # checks it out to the current directory
   ```

2. Use a CVS client tool, such as TortoiseCVS. This tool allows you to work with files under CVS directly from Windows Explorer, thus eliminating the need to code. Other CVS client tools include CVSGui, SmartCVS, and many others.

**Note:** CVS modules can also be created when you create associations with application directories. For details, see *How to Create Associations Between Application Directories and CVS Modules* on page 733.

**Procedure: How to Configure DataMigrator to Work With CVS**

To use a source control system to source manage your DataMigrator flows, stored procedures, and synonyms, you must have a source control program and database configured. The source control program will then appear as an SCC Provider.

In the Data Management Console:

1. On the Home tab, in the Tools group, click *Configure Source Control*.

   **Note:** If you have not properly configured PushOK CVSSCC NT Proxy, this option will not appear in the Tools group.

   The Configure Source Control dialog box opens.

2. Select *PushOk CVSSCC NT* as the SCC Provider from the drop-down menu.

   **Note:** Selecting *none* removes source control from your system.

3. Optionally, click *Test* to make sure you can connect to the source control system from the DMC.

   **Note:** If you do not have a license for the Source Control, you might be prompted to get one at this time.
4. Click OK.
   The Set Source Control Provider dialog box opens.

5. Click Yes.
   Once a SSC Provider has been set, source control options appear on the right-click menus.

**Procedure:** How to Create Associations Between Application Directories and CVS Modules

Creating a source control folder provides a place to manage your DataMigrator files. Creating an association between your application directory and that folder tells DataMigrator where to manage those files.

**Note:** Developers must associate application directories inside their user home directory, while only the manager associates to the central application directory.

In the navigation pane of the Data Management Console:

1. Right-click an application directory, select *Source Control*, and then click *Create Association with Source Control*.
   The Associate Application with Source Control Project window opens.

2. Click Associate.
The Please select a CVSROOT, module and local path dialog box opens, as shown in the following image.

3. Enter the repository information in the CVSROOT field. CVSROOT describes the physical location of repository, the directory that stores the master copies of the files.

The syntax for the CVSROOT is:

`:protocol:pcname/path`

where:

- **protocol**
  
  Is for a Windows server, the recommended protocol is sspi.

- **pcname**
  
  Is the name of the machine running CVS.

- **path**
  
  Is the location of the CVS repository.
An example of a complete syntax is:

:sspi:PC1/cvs

4. Click Check.

If CVSROOT is valid, the following message should appear:

The CVSROOT is OK.

5. Click Properties.

The PushOK CVS Proxy properties window opens.

6. Click the CVS Options tab, deselect the Enable file remove monitoring check box, and click OK.

7. Create a new module or specify the existing module name. Module is a specific directory (tree of directories) in the main repository, CVSROOT.

8. Click Check.

If Module is valid, the following message should appear:

The MODULE is OK.

9. Set the local path to c:\ibi\srv77\dm\etlgtemp\server. The local path is where the source managed files temporarily reside.

10. Click OK to close the select a CVSROOT, module and local path dialog box.

11. Click OK to close the Associate Application with Source Control Project dialog box.

The Application Directory icon displays a lock to indicate the source control association.

You are now ready to use CVS for source management.

**Setup, Configure, and Associate DataMigrator to Work With Perforce**

Perforce does not store directories. Directories are simply components of the filespec which uniquely identify each file. However, it is recommended that you create folders in your source control system that mimic your node/application directory structure. Creating folders using your server node names will help you distinguish application directories with the same names on different servers.

In order to create an empty directory or structure of empty directories, you need to anchor those directories using a place-holder file. The place-holder file is added to Perforce and then deleted, creating a deleted revision that anchors the directory.

You can do this by using P4V or using the command line.
Procedure: How to Create Perforce Projects

Create Perforce Projects Using P4V:

1. Launch P4V and connect to the server using proper connection settings.

2. Select a workspace whose view includes the area of the depot in which you wish to create directories.

3. Create the new directory by right-clicking (in your system file tree) the folder where you want to create the directory, and clicking New Folder. When prompted, enter the directory name Empty.

4. Open a text editor of your choice and create a new file in the Empty directory named .place-holder. Using a period (.) in front of the name makes this a hidden file on Linux, UNIX, or Mac operating systems, though it remains visible in Windows workspaces. The file can be empty, it just has to be present.

5. In P4V, right-click the .place-holder file and click Mark for Add. Right-click .place-holder again and click Submit. A submit dialog opens. Enter something in the description (such as Create Empty) and click Submit.

   Note: If the .place-holder file is visible, right-click Empty and click Refresh Empty.

6. Click Depot tab.

   In the depot, you see the Empty directory and the .place-holder file it contains.

7. Right-click .place-holder and click Mark for Delete. Right-click it again and click Submit.

   A submit dialog opens.

8. Enter a description in the appropriate field and click Submit.

   The .place-holder file disappears from the directory, leaving an empty directory in the depot.

9. If you do not see the directory in P4V, confirm that Show Deleted Files is selected from the Filter Depot menu at the top of the depot tree.
The following image shows an example of what the tree structure should look like.

Create Perforce Projects Using the Command Line:

The following commands reproduce the procedure described above. Issue these commands from the directory where you wish to create the new directory.

```bash
mkdir Empty
cd Empty
vi .place-holder (on Windows, try edit .place-holder)
p4 add .place-holder
p4 submit -d "Create Empty"
p4 delete .place-holder
p4 submit -d "Create Empty"
```

Procedure: How to Configure DataMigrator to Work With Perforce

To use a source control system to source manage your DataMigrator flows, stored procedures, and synonyms, you must have a source control program and database configured. The source control program will then appear as an SCC Provider.
In the Data Management Console:

   
   **Note:** If you have not properly configured Perforce Client, this option will not appear in the Tools group.

   The Configure Source Control dialog box opens.

2. Select Perforce SCM as SSC Provider from the drop-down menu.
   
   **Note:** Selecting none removes source control from your system.

3. Optionally, click Test to make sure you can connect to the source control system from the DMC.
   
   **Note:** If you do not have a license for the Source Control, you might be prompted to get one at this time.

4. Click OK.

   The Set Source Control Provider dialog box opens.

5. Click Yes.

   Once a SSC Provider has been set, source control options appear on the right-click menus.

**Procedure:** How to Create Associations Between Application Directories and Perforce Project

**Note:** Developers must associate application directories inside their user home directory, while only the manager associates to the central application directory.

In the navigation pane of the Data Management Console:

1. Right-click an application directory, select Source Control, and then click Create Association with Source Control.

   The Associate Application with Source Control Project window opens.

2. Click Associate.
The Open Connection dialog box appears, as shown in the following image.

3. Enter the Server, User, and Workspace connection information:
   - **Server.** The name of the computer where your Perforce server is running.
   - **User.** Your Perforce user name.
   - **Workspace.** Folders or directories on the client computer where you work on revisions of files that are managed by Perforce.

   **Note:** To display a list of existing users and workspaces, click *Browse*. The server and port must be specified before clicking *Browse*. Click *New* to create new users and workspaces.

4. Click *OK* to close the Associate Application with Source Control Project dialog box.

   The Application Directory icon displays a lock to indicate the source control association.

   You are now ready to use Perforce for source management.

**Setup, Configure, and Associate DataMigrator With Subversion**

Subversion is an open source revision control system and successor to Concurrent Versions System (CVS) that can run on cross-platform operating systems. This section will discuss the Windows implementation. Subversion is a command-line system and cannot be used directly by DataMigrator. To use Subversion, PushOk SVN SCC Proxy is needed to convert the SCC API calls into Subversion commands. It is recommended that the source control folders, also known as modules, are structured the same as the DataMigrator application directories.

**Note:** The PushOK proxy is available from [http://www.pushok.com](http://www.pushok.com) as the SVN SCC Plug-in. Use the Windows version, not the x64 version, even if you running 64-bit Windows.
**Procedure:** How to Create Subversion Modules

Before you can create Subversion modules, you need to create a repository and add users. Once this is complete, you can create the modules through the Subversion command line, Subversion client, or at the DataMigrator configuration. The following procedure shows how to create a module with a Subversion client, TortoiseSVN. For details on how to create modules at DataMigrator configuration, see *How to Create Associations Between Application Directories and Subversion Modules* on page 743.

1. Install TortoiseSVN.
2. From Windows Explorer, right-click the Subversion Repository folder, point to TortoiseSVN, and then click Repo-browser.
3. Right-click on top of the tree path and click *Create folder*, as shown in the following image.

![Image of TortoiseSVN interface showing Create folder option](image)

4. Enter the folder name. This should be the same as the DataMigrator Application Directory, such as ibisamp.

**Procedure:** How to Configure DataMigrator to Work With Subversion

To use Subversion from client, you must install the client program. In Windows environment, this would be PushOk SVN SCC Proxy. Once this is installed and configured, the PushOk SVNSCC provider will show as an SCC Provider.

In the Data Management Console:

1. On the Home tab, in the Tools menu, click *Configure Source Control*.

   **Note:** If you have not properly configured PushOk SVNSCC Proxy, this option will not appear in the Tools group.

   The Configure Source Control dialog box opens.

2. Select *PushOk SVNSCC* as the SCC Provider from the drop-down menu.
Note: Selecting none removes source control from your system.

3. Optionally, click Test to make sure you can connect to the source control system from the DMC.

Note: If you do not have a license for the Source Control, you might be prompted to get one at this time.

4. Click OK.

The Set Source Control Provider dialog box opens.

5. Click Yes.

Once an SCC Provider has been set, source control options appear on the right-click menus of the application directories.

In PushOk SVNSCC configuration plug-in:

1. Open the SVNSCC plug-in by going to Start, All Programs, PushOk Software, SVNSCC, and Configure Plugin.
2. Click the SVN Options tab and verify that the options are set the same, as shown in the following image.

![SVN Options Tab](image)

**Procedure:** How to Create Associations Between Application Directories and Subversion Modules

Creating a source control folder provides a place to manage your DataMigrator files. Creating an association between your application directory and that folder tells DataMigrator where to manage those files.
Note: Developers must associate application directories inside their user home directory, while only the manager associates to the central application directory.

In the navigation pane of the Data Management Console:

1. Right-click an application directory, select Source Control, and then click Create Association with Source Control.

   The Associate Application with Source Control Project window opens.

2. Click Associate.

   The Select SVNURL, module and local path dialog box opens, as shown in the following image.

3. Enter the repository information in the SVNURL field. SVNURL describes the physical location of the repository. The repository is the directory that stores the revision copies of all the files.
The syntax for the SVNURL is:

`:protocol:pcname/path`

where:

`protocol`

Is the recommended protocol for a Windows server is `svn`.

`pcname`

Is the name of the machine running Subversion.

`path`

Is the location of the SVN repository. An example of a complete syntax is:

`svn://user1@PC1`.

Note: Each developer must connect with unique user name and password set in Subversion.

4. Click **Check**.

If SVNURL is valid, the following message should appear:

**The SVNURL is OK.**

5. Create a new module or specify the existing module name. Module is a specific directory (tree of directories) in the main repository, SVNURL. This should mimic the name of the DataMigrator application directory for consistency.

6. Click **Check**.

If Module is invalid and does not exist, click **Create**. The module needs to be created only once by the developer, who is the first to associate the application directory with Subversion module. The rest of the developers need to connect to, and work with, the created or pre-existing module created by the first developer.

If Module is valid, the following message should appear:

**The MODULE is OK.**

7. The local path can be set to any local directory, such as `C:\temp`, but a different local path for each module must be specified. It is recommended that you create and keep the same module name inside a single local path, such as `C:\temp\ibisamp`.

8. Click **OK** to close the Select SVNURL, module and local path dialog box.

9. Click **OK** to close the Associate Application with Source Control Project dialog box. If association is successful, the Application Directory icon should display a lock to indicate the source control association.
You are now ready to use Subversion for source management.

**Note:** When you disconnect and reconnect with a different user to the DMC under source control, PushOk SVN Proxy might prompt you again for a user name and password to Subversion repository. If this occurs, login again with the appropriate Subversion user credentials.

**Setup, Configure, and Associate DataMigrator to Work With Team Foundation Server (TFS)**

Before you can create TFS Team Projects, you have to create a TFS Collection and setup users. A TFS Collection contains Team Projects, which can be associated to DataMigrator Application Directories. It is recommended that you use TFS Collections to differentiate various stages of the product life cycle, such as development, quality assurance, and production. This gives you the option to track different revisions across various deployment stages. However, the setup can vary depending on the requirements of the organization.

To use TFS for source control solution with a 32-bit version of DataMigrator, Team Foundation Server MSSCCI Provider 2010 (or 2012) 32-bit is needed to convert the SCC API calls into TFS commands.

To use TFS for source control solution with a 64-bit version of DataMigrator, Team Foundation Server MSSCCI Provider 2012 (or higher) 64-bit is needed to convert the SCC API calls into TFS commands.

<table>
<thead>
<tr>
<th>TFS MSSCCI Provider</th>
<th>Link Available:</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012 32-bit</td>
<td><a href="http://visualstudiogallery.msdn.microsoft.com/b5b5053e-af34-4fa3-9098-aaa3f3f007cd">http://visualstudiogallery.msdn.microsoft.com/b5b5053e-af34-4fa3-9098-aaa3f3f007cd</a></td>
</tr>
</tbody>
</table>
Procedure: How to Create TFS Collection and Team Projects

TFS Collections are created from the TFS Administration Console. Please contact your TFS administrator to have a collection created if you do not have one already set up.

TFS Team Projects are created through Visual Studio. You can create a new Team Project by right-clicking TFS Collection from the Team Explorer tab and clicking New Team Project, as shown in the following image.

A wizard guides you in specifying the Team Project settings. The key setting is to name the team projects that correspond to the DataMigrator application directories. You can then click Finish. However, you can continue to the next steps if you want to specify the process template settings.
Use the information in the Properties pane to associate the TFS team project with DataMigrator application directories, as shown in the following image.
**Procedure:** How to Set an Exclusive Check-out Configuration

TFS allows concurrent check-out by default. It is required that you switch to an exclusive check-out method technique. To do this from Visual Studio, go to Team Explorer, right-click the Team Project, select Team Project Settings, and then click Source Control. Deselect the Enable multiple check-out check box and click OK, as shown in the following image.

![Source Control Settings - baseapp](image)

**Procedure:** How to Configure DataMigrator to Work With TFS

To use a source control system to source manage your DataMigrator flows, stored procedures, and synonyms, you must have TFS Collection and Team Projects set up and TFS MSSCCI Provider 2010 (or 2012) 32-bit installed on client machines.

In the Data Management Console:


   **Note:** If you have not properly configured Team Foundation Server MSSCCI Provider, this option will not appear in the Tools group.

   The Configure Source Control dialog box opens.
2. Select **Team Foundation Server MSSCCI Provider** from the SSC Provider drop-down menu, as shown in the following image.

![Configure Source Control](image.png)

**Note:** Selecting *none* removes source control from your system.

3. Click **OK**.

   The Set Source Control Provider dialog box opens.

4. Click **Yes**.

   Once a SSC Provider has been set, source control options will appear in the context menus.

**Procedure:** How to Create Associations Between Application Directories and TFS Collection Projects

Creating a source control folder provides a place to manage your DataMigrator files. Creating an association between your application directory and that folder tells DataMigrator where to manage those files.

**Note:** Developers must associate application directories inside their user home directory, while only the manager associates to the central application directory.
In the Data Management Console:

1. Right-click an application directory in the navigation pane, select Source Control, and then click Create Association with Source Control. In this procedure, baseapp will be used as an example. The Associate Application with Source Control Project dialog box opens, as shown in the following image.

![Associate Application with Source Control Project dialog box](image)

2. Click Associate.

3. Select the appropriate TFS Collection, as shown in the following image.

![Connect to a Team Foundation Server](image)

If the desired Team Foundation Server name is not listed in the drop-down list of the available Server names, you can perform the following steps to add the name:

a. Click the Servers button.

b. Click the Add button on the following screen and enter the full URL of the Team Foundation Server name (for example, http://mydomain:8080/tfs/DefaultCollection).

c. Click the OK button.
4. Click OK.

5. You can now specify the local and server locations for the project. Specify the server path to match Team Projects. Specify the local path to the designated directory, which should be etlgtemp\tmp directory within your DataMigrator Server install path. For example, C:\ibi\svn77\DataMigrator\etlgtemp\server. The local path is where the source managed files temporarily reside.
6. You should use the default value for local workspace, which should be generated for you, as shown in the following image.

![Add Workspace dialog box](image)

**Note:** The name of a temporary directory that is set for the Workspace on the Team Foundation Server is also used by DataMigrator as a ‘Local path’.

7. Click OK to close the Associate Application with Source Control Project dialog box. The Application Directory icon displays a lock to indicate the source control association.

You are now ready to use TFS for source management.

**Setup, Configure, and Associate DataMigrator to Work With Visual SourceSafe**

Source control folders, called projects in Visual SourceSafe, work essentially like DataMigrator application directories. You should create folders in your source control system that mimic your node/application directory structure. Creating folders using your server node names will help you distinguish application directories with the same names on different servers.

**Note:** Visual SourceSafe projects can be also be created when you create associations with application directories. For details, see *How to Create Associations Between Application Directories and Visual SourceSafe Projects* on page 755.

**Procedure: How to Create Visual SourceSafe Projects**

2. On the File menu, click Create Project.
3. In the Create Project in dialog box, type the project name in the Project field.
4. Optionally, enter a brief description of the project in the Comment box.
5. Click OK to create a single empty project.

Normally, SourceSafe requires you to set a working folder, but since DataMigrator is a client/server application, that is not required. DataMigrator sets an internal working directory for files under source control and pushes those files to the server when they are changed.

**Procedure:**  How to Configure DataMigrator to Work With Visual SourceSafe

To use a source control system to source manage your DataMigrator flows, stored procedures, and synonyms, you must have a source control program and database configured. The source control program will then appear as an SCC Provider.

In the Data Management Console:

1. From the Tools menu, click *Configure Source Control.*

   **Note:** If you have not properly configured Visual SourceSafe Client, this option will not appear on the Tools menu.

   The Configure Source Control dialog box opens.
2. Select *Microsoft Visual SourceSafe* as the SSC Provider from the drop-down menu.

   **Note:** Selecting *none* removes source control from your system.

3. Optionally, click Test to make sure you can connect to the source control system from the DMC.

   **Note:** If you do not have a license for the Source Control, you might be prompted to get one at this time.

4. Click *OK*.

   The Set Source Control Provider dialog box opens.

5. Click *Yes*.

   Once a SSC Provider has been set, source control options appear on the right-click menus.

**Procedure:** How to Create Associations Between Application Directories and Visual SourceSafe Projects

Creating a source control folder provides a place to manage your DataMigrator files. Creating an association between your application directory and that folder tells DataMigrator where to manage those files.

**Note:** Developers must associate application directories inside their user home directory, while only the manager associates to the central application directory.

In the navigation pane of the Data Management Console:

1. Right-click an application directory, select *Source Control*, and then click *Create Association with Source Control*.

   The Associate Application with Source Control Project dialog box opens.

2. Click *Associate*.

   The Visual SourceSafe Login dialog box opens.

3. Type your User ID, Password, and Path to the srcsafe.ini file for your SourceSafe database.

   **Tip:** For convenience, it is recommended that you use the same user ID and password for both DataMigrator and SourceSafe.

4. Click *OK*. 
The Add to SourceSafe Project dialog box opens, as shown in the following image.

![Add to SourceSafe Project dialog box](image)

5. Select a project folder and click **OK**. If you have not created a project through Visual SourceSafe Explorer, you can do so here by specifying project name and clicking **Create**.

6. Click **OK** to close the Associate Application with Source Control Project dialog box.

The Application Directory icon displays a lock to indicate the source control association.

**Note:** While the Data Management Console lists application directories according to their position in your Application Path, SourceSafe lists project folders alphabetically.

Visual SourceSafe offers Share function, which enables files to be shared among multiple projects. If an item is modified in one project, the changes will be reflected in other projects simultaneously. This can be beneficial in development environments that require projects to be synchronized when there are changes made to file. DataMigrator can work with shared projects as long as you associate and configure it with VSS.

**Using a Source Control System to Manage Procedures and Synonyms**

Once you have created source control folders and created associations in the Data Management Console, the next step is to put the files under source control. In other words, you need to add them to a source control folder. You can add individual files or select a group of files from an application directory.
Files that have been added to a source control system need to be checked out before changes can be made. Checking files back in creates a history of their development.

**Procedure: How to Add Files to a Source Control System**

To track and manage changes, DataMigrator files must first be added to a source control system. This can be performed by developers or the manager. In either case, each new file needs to be added to the source control system just one time. Once the file has been added to the source control system, the rest of the developers can attain the copy from the source control system by invoking Get Latest Version.

**Note:** Files that are added or checked into a source control system become read-only until they are checked out.

In the navigation pane of the Data Management Console:

1. Right-click an application directory, multiselected objects, an individual flow, stored procedure, or synonym, select *Source Control*, and then click *Add/Check In*.

   The Add/Check In Files dialog box opens.

2. Select the files you want to add to source control. Files that are added are automatically checked in.

   **Tip:** You can use the *Select All* check box to both select and deselect all the files.

3. Optionally, you can check the files out by selecting the *Keep files checked out* check box.

4. Optionally, you can add comments by typing them into the *Comment* text box. If you are adding multiple files, they will all have the same comment.

5. Click *OK*.

   **Note:** The default listing for files in the Add/Check In dialog box is by Type. To list them by name, click the *Files* column heading.

   Icons display a lock to indicate that the files are under source control.

If you open your source control program, you will see that the files from the DataMigrator application directory have been added to the corresponding project folder. The synonym (dminv) has a .mas extension and the stored procedure (dmdefs) has a .fex extension. The flow (dmdemo) has two files stored in source control. The source code is stored in the file with the .fex extension. The file with the .etg extension holds the information required for the DMC interface.
**Reference: Add/Check In Files Dialog Box**

The Add/Check In Files dialog box is available by right-clicking an application directory, multi selected objects, an individual flow, stored procedure, or synonym, selecting Source Control, and then Add/Check In.

The Add/Check In Files dialog box has the following fields options:

- **Select All**
  
  Allows you to select or deselect all of the files in the application directory.

- **Files/Type/Action**
  
  Shows all of the files in the application directory. The files can be sorted by file name, type, or action.

- **Keep files checked out**
  
  Allows you to keep a file checked out that is being added or checked in. A source managed file must be checked out before you can make changes to it.

- **Comment**
  
  Allows you to add comments for the files.

**Procedure: How to Get Latest Version From Source Control System**

Once files have been added to source control, you can invoke Get Latest Version to attain a copy of the latest files from the source control system.

In the navigation pane of the Data Management Console:

1. Right-click an application directory, multi selected objects, an individual flow, stored procedure, or synonym, select Source Control, and then click Get Latest Version.

   The Get Latest Version window will appear.

2. Click OK.

   **Note:** If you invoke the Get Latest Version from application directory, the Get Latest Version dialog box will list all the files that have been added to the source control system. There is a Select All option to get all the files from source control system, or you can optionally select the individual files that you want.

**Procedure: How to Check Files Out of Source Control**

For files that are being source managed, you must check them out before you can make and save changes. Until the files are checked out, they remain read-only.
In the navigation pane of the Data Management Console:

1. Right-click an application directory, multi selected objects, an individual flow, stored procedure, or synonym, select Source Control, and then click Check Out.

   The Check Out dialog box opens.

2. Select the files you want to check out.

   **Tip:** You can use the Select All check box to both select and deselect all the files.

3. Optionally, add comments by typing them into the Comment text box. If you are checking out multiple files, they will all have the same comment.

4. Click OK.

Icons in the application directories display a red check mark to indicate that the files have been checked out.

**Note:** The default listing for files in the Check Out dialog box is by Type. To list them by name, click the Files column heading.

If you open the source control program, you will see that the files that you checked out from the Data Management Console have been checked out of the source control program. These files are no longer read-only, which allows you to make changes before checking them back in to source control.

If a data source has an Access File associated with it, the Access File will also be automatically checked out, in the same way that an .etg file is checked in or out with its associated flow.

**Note:**

- If you try to edit a file that has been added to source management without checking it out, the following message appears:
If you try to check out a source-managed file that has already been checked out by another user, the following message appears:

Reference: Check Out Files Dialog Box

The Check Out Files dialog box is available by right-clicking an application directory, multi selected objects, an individual flow, stored procedure, or synonym, selecting Source Control, and then clicking Check Out.

The Check Out Files dialog box has the following fields and options:

Select All

Allows you to select or deselect all of the files that have been checked in and/or added to the application directory.

Files/Type/Action

Shows all of the files that have been checked in or added to the application directory. The files can be sorted by file name, type, or action.

Comment

Allows you to add comments for the files.

Procedure: How to Check Files Into Source Control

After you have made changes to files, you must check them back into your source control system to maintain a record of those changes. Files that are checked in are copied back to the server.

In the navigation pane of the Data Management Console:

1. Right-click an application directory, multi selected objects, an individual flow, stored procedure, or synonym, select Source Control, and then click Add/Check In.

   The Add/Check In dialog box opens.

2. Select the files you want to check in. Only files that have been checked out have Check In as the possible action. The action for files that have not been added to source control is Add File.
**Tip:** You can use the Select All check box to both select and deselect all the files.

3. Optionally, you can keep files checked out by selecting the *Keep files checked out* check box.

4. Optionally, you can add comments by typing them into the Comment text box. If you are checking in multiple files, they will all have the same comment.

5. Click OK.

**Note:** The default listing for files in the Add/Check In dialog box is by Type. To list them by name or action, click the *Files* or *Action* column headings.

**Procedure:** How to Undo a Check Out From Source Control

If you decide that you do not want to make changes to a file, you can undo its check out from the source control system. The file becomes read-only again, and source control maintains the version as it existed before the file was checked out.

In the navigation pane of the Data Management Console:

1. Right-click an application directory, multi selected objects, an individual flow, stored procedure, or synonym, select *Source Control*, and then click *Undo Check Out*.

   The Undo Check Out dialog box opens.

2. Select the files whose check out you want to undo.

   **Tip:** You can use the Select All check box to both select and deselect all the files.

3. Click OK.

   **Note:** The default listing for files in the Undo Check Out dialog box is by file name. To list them by Type, click the *Type* column heading.

**Reference:** Undo Check Out Dialog Box

The Undo Check Out Files dialog box is available by right-clicking an application directory, multi selected objects, an individual flow, stored procedure, or synonym, selecting *Source Control*, and then clicking *Undo Check Out*.

The Undo Check Out Files dialog box has the following fields and options:

**Select All**

Allows you to select or deselect all of the files that have been checked out.
**Files/Type/Action**

Shows all of the files that have been checked out from the application directory. The files can be sorted by file name or type.

**Procedure: How to View the History, Versions, and Differences of a File**

The source management history records of a file each date and time it was checked in, as well as the user involved.

**Note:** This procedure uses Visual SourceSafe, but other source control programs should work analogously.

In the navigation pane of the Data Management Console:

1. Right-click an individual flow, stored procedure, or synonym, select **Source Control**, and then click **History**.
   
   The History dialog box opens, showing the flow, stored procedure, or synonym.

2. Click **OK**.

   The History Options dialog box opens.

3. Optionally, use the **Include Labels** or **Labels Only** check boxes to show labeled activity.
   
   **Note:** Labels can be useful to distinguish versions of a particular file. You can assign labels from the SourceSafe File menu, or from the History Details dialog box that is available from the Data Management Console History of dialog box.

4. Optionally, use the From and To fields to specify a date range for the history.

5. Optionally, use the User field to specify the history for a specific user's activity.

   The History of dialog box opens, showing the user, date, and time for each check in.

In Visual SourceSafe, the History of dialog box allows you to view a version of the file, see or set details like labels and comments, get the latest version, compare the difference between two versions, pin or unpin the file, rollback the file to an earlier version and discard later versions, or generate a report on the file activity. Different source control programs may provide a different set of options.

**Procedure: How to Deploy Latest Files From Source Control to Central Application Directory**

The manager can deploy the latest files from source control to central application directory. Since the working directory of a manager is the central application directory, all they need to do is invoke Get Latest Version from each application directory or an individual flow, stored procedure, or synonym.
In the navigation pane of the Data Management Console:

1. Right-click an application directory, multi selected objects, an individual flow, stored procedure, or synonym, select Source Control, and then click Get Latest Version.

   The Get Latest Version dialog box will open.

2. Click OK.

   **Note:** If you invoke the Get Latest Version from application directory, the Get Latest Version dialog box will list all the files that have been added to the source control system. There is also a Select All option which gets all the files from the source control system.
Adding email in a process flow 410
Adding IDs 53
Adding nodes 435

Aerial View 150
Aggregating columns 253
Alphanumeric fields 367, 372

APP PATH 62
  Configuring 62
Application button 127
Application directories 59, 60, 62
  Adding 61
  Creating 59, 437
  Deleting 60, 62
Application menu 127
Application paths 57, 62
  APP PATH 62
  Application path 57
  Configuring 57
  Restricting access 57, 119, 120
Applications 107

Assist button 328, 595

Automapping 324, 326, 327

Automatic passthru (APT) 494, 496
  Creating pre-extract procedures 496
  iWay and 496
  Real numbers and 496
  Requirements 494

Access File Text View tab 512
Adding data adapters 65
columns 247
  changing display in the Synonym Editor 143
  extract transformations 254
  information displayed 143
  managing 445
  mapping 451
  naming 109, 118
  Selected Columns window 248, 250
  selecting 247, 248, 445
  text 108
  variable character length 108
  virtual 441
commands 265
  -TYPE 715
complex data flows 206
complex mapping 324, 330
components 208
  adding 208
  connecting 211, 212
  deleting 208, 211
  disconnecting 211, 212
Concurrent Versions System (CVS) 731
configuring application paths 57
configuring DataMigrator 41
configuring email notification 84
configuring server nodes 435
configuring SMTP services 84
connecting objects 211, 411, 413, 420
connection logic 411, 413, 420, 464
  specifying 464

columns 247
  changing display in the Synonym Editor 143
  extract transformations 254
  information displayed 143
  managing 445
  mapping 451
  naming 109, 118
  Selected Columns window 248, 250
  selecting 247, 248, 445
  text 108
  variable character length 108
  virtual 441
commands 265
  -TYPE 715
complex data flows 206
complex mapping 324, 330
components 208
  adding 208
  connecting 211, 212
  deleting 208, 211
  disconnecting 211, 212
Concurrent Versions System (CVS) 731
configuring application paths 57
configuring DataMigrator 41
configuring email notification 84
configuring server nodes 435
configuring SMTP services 84
connecting objects 211, 411, 413, 420
connection logic 411, 413, 420, 464
  specifying 464
connector objects 146, 411
Console Log 159
copying 137
    procedures 137
    synonyms 137
Core Engine Settings 719
creating application directories 437
creating branches 418
creating joins 238
creating metadata 72
creating objects 386
creating pre-extract procedures 496
creating process flows 391, 392
    from data flows 393
creating stored procedures
    in Data Management Console 396
creating synonyms 72
CVS (Concurrent Versions System) 733

D
data adapters 65
    adding 65
data extraction 493
Data Flow object 145
Data Flow tab 35, 139, 438
data flows 35, 193, 194, 336, 438
    adding objects 208
    complex 206
    creating 207, 433, 438
    DataMigrator variables 336
data flows 35, 193, 194, 336, 438
    deleting objects 208
designing 35
direct load 377
interface 194
objects 140
running 459
simple 205
storing 195
system variables 336
using variables 336
validating 139
data lineage 552
data loading 497
data management 28
Data Management Console 29, 31–33, 46, 107, 213
    adding application directories 61
deleting application directories 60, 62
designing process flows 35
filtering 267, 340
functions 31
interface 125
launching Web Console 53
navigation pane 32, 33, 132
ribbon 129
server administration tasks 53
sessions 191
shortcuts 131
Data Management Console 29, 31–33, 46, 107, 213
  sorting 273
  source component 213
SQLAnswerSet component 265
style menu 130
targets 276, 283, 324, 326, 327, 330
testing connections 52
testing servers 52
Text View tab 35
workspace 138
data marts 28
data migration 28
data objects 461
  adding to process flows 461
data profiling 200, 543, 546
  in synonyms 543
  outliers 546
  patterns 546
  single column 546
  statistics 546, 550, 551
  values 546
data source columns 219
  updating 219
data sources 28, 108, 113, 209, 213
  accessing 112
  column names 109
  columns 247
  joining 234, 438
  naming 109
  performance 112
  reviewing 112
  selecting 213, 438
  updating 219
data store tables 683
ETLSTATS 690
data targets 28, 108, 115
  accessing 119
  adding 276, 451
  column names 109, 118
  creating 457
  existing 276, 283
  flat files 276
guidelines 118
local relational 116
naming 118
new 276
non-relational 117, 276
performance 115
relational 276
remote destinations 118
reviewing 115
security 119
selecting 276, 283
Server Transfer Files 276
data warehousing 28
designing applications 107
data
  data sources 113
data
  data targets 115
  filtering 448
  validating 455
DataMigrator 27, 28
  architecture 29
  components 29
  configuring 41
  creating data flows 433
  creating process flows 433
  help 40
  installing 41
  processing 29
  security 119
  upgrading 102
DB log parameters 670
DB_LOOKUP 328
Db2 301
  tables 679
DBMS errors
  limit 718
DBMS
  SQL flows 379
DEFINE 540
deleting
  procedures 138
  synonyms 138
delimited flat file targets 277
designing data flows 35
designing process flows 35
detailed flow reports 488, 489
Dialogue Manager commands
  -TYPE 715
direct load flows 377, 677
  creating 377

E
e-mail objects 146
editing conditions 413, 420
editing SQL commands 265
editing synonyms 200
email notification 84, 91
  properties 165
email objects 410
error messages 709
  resolutions 714
  run-time 713
  Server 710
errors 709
ETLLOG table
  Server Log 687
ETLSTATS table 690
  Server Statistics table 687
event-based scheduling 695
  CMASAP procedure 706
  CMRUN procedure 695, 696, 698–702
  FOCUS and 708
  status 708
  WebFOCUS and 708
executing flows 396
  procedures 396
  simultaneously 419
expressions 233
  functions and 347
  order of evaluation 233
extract transformations 219, 220, 441
  calculations 221, 223, 254, 255
  constant values 254
  creating 219–221, 223, 254, 255, 441
date values 254
  in data source object 220
SQL Calculator 221, 223, 254–256
SQL functions 258
time values 254
extracting data 493

F

field formats 365, 367
  alphanumeric fields 367, 372
  formatting new fields 367, 372
Field View tab 502
Filter Calculator 271
filtering 267, 340, 448
  data 448
  expressions 449
  Filter Calculator 267, 271
flat files 276
floating-point numbers 496
flow object 145
flow properties 376
flow reports 488
  detailed 488, 489
  summary 468, 472
Flow tab 153
flows 28, 35, 119, 492
  execution properties 162
  building 433, 434
  controlling 409
  copying 137
  creating 433, 434, 438
data sources 213, 438
data targets 276, 283
e-mail notification 165
field formats 365, 367
filtering 267, 340
general properties 162
graphical representation 492
grouping 419
mappings 324
printing 492
procedures 409
properties 161, 376
record logging properties 167
report 488
restart properties 164
running 119
running remotely 695
scheduling 424, 425, 695
sorting 273
flows 28, 35, 119, 492
  status 708
  testing SQL 265
FOCUS 708
  executing flows 708
functions 347
  changing display 143

G

global variables 404
  graphical representation of flows 492

H

help button 131
Home tab 151

I

ibisamp directory 57
Impact Analysis 202
inner joins 443
inserting sample data 201
installing DataMigrator 41, 46
installing iWay Client
  Install Wizard 47
iWay Connector messages 712
iWay Web Console application path 57
iWay Web Console
  Scheduler configuration page 87

J

Join Calculator 234, 238, 241
Join Properties window 238, 444
joining data sources 234, 438, 443
  Join Calculator 234
  Join Properties window 238
  performance considerations 234, 235
joins 234, 438
  assigning properties 443
  inner 443
  types 444

L

List View tab 504
listening parameters 671
load options 285, 309, 310
  targets 309, 310, 314, 321, 322
local relational data targets 116
log file 467, 491
  viewing 488, 490
log messages 389
log reports 467, 470
  columns 468
logs 195
  sample data 682
lookup cache 499
lookups 594

M

managing columns 445
mappings 324, 326, 327, 330, 332
  automapping 324, 326, 327
  columns 451
  complex mapping 324, 330
  field formats 365, 367
  simple mapping 324
  temporary columns 324
  testing 330
Message window 33
messages 159
  run-time 715
metadata 71, 72
  creating 71
  testing 71, 77
Modeling View tab 511, 536
multiple data targets 457

N

navigation pane 32, 386
non-automatic passthru 495
non-relational data targets 117, 276
Nucleus 302
NULLABLE 718

O

objects 208, 386
  process flows 144
Oracle 289
Oracle tables 679
order of evaluation 233

P

Parallel Group object 145, 419
parameters 717
  setting 717
passwords 89, 705
  supplying with procedures 705
Perforce 735
performance
  data loading 497
  extracting data 493
Print Preview tab 486
printing flows 492
printing
  synonyms 512
procedures 195, 395, 407
  copying 137
  deleting 138
  executing 409
  flows 409
  process flows 395
  supplying passwords 705
Process Flow tab 35, 144, 386
process flows 35, 385, 386, 460
  adding components 389
  adding data objects 461
  adding email objects 410
  adding procedures 395, 407
  adding stored procedures 395
  branching 418, 462
  complex 389
process flows 35, 385, 386, 460
  creating 391, 433
  creating from data flows 393
  creating new 392
  designing 35
  objects 144, 385
  running 423
  scheduling 423–425
  simple 388
  standalone procedures 392
  storing 195
properties
  Application Directory 171
  email notification 165
  execution 162
  flows 161
  general 162
  record logging 167
  restart 164
prototype synonym 532

Q
Quick Access Toolbar 128

R
read limits 671
real numbers 496, 499
  automatic passthru and 496, 499
  changing formats 496, 499
record retrieval 449
recreating DBMS tables 201
Red Brick 303
relational data targets 276
relational lookups 328
relational targets 283, 309, 312
remote destination data targets 118
remote servers 44
  adding 495
remote table access 118
reports 195, 467, 472
  copying 481
  flow 468, 472, 488, 489
  fonts 482
  log 467, 470
  Report tab 158
  running 473, 474, 477
  statistics 468
return codes 695
  CMRUN procedure 695, 703
ribbon 129, 150
  Metadata tab 513
  tabs 130
running on recurring schedule 423, 424
running on specific days 423, 425
running once 423, 424
running process flows 389, 423

S
sample data 77, 200, 214, 217
  inserting 201
Index

Scheduler 86, 428
Scheduler configuration page 87
Scheduler window 428
scheduling flows 423–425, 695
  event-based scheduling 695, 696, 698–702, 706, 708
  Scheduler 424, 425
scheduling process flows 389, 424
security 119
Segment View tab 503
Select Synonym dialog box 209
Selected Columns window 248, 250
selecting columns 247, 248, 250, 445
server administration 53
Server Log
  ETLLOG table 687
server messages 709
  iWay Connector communications 712
  run-time 713, 715
  troubleshooting 710
server nodes 435
Server statistics table 687
  ETLSTATS table 687
Server Transfer Files 276
Server Web Console 29, 34
  functions 34
Server
  properties 168
  run-time errors 713
  Server Node Configuration window 169
SET command 717
Set variables object 147
setting Wait conditions 420
simple data flows 205
simple mapping 324
slowly changing dimensions 522
SMTP services 84
Sort Columns window 275
sorting 273, 275
source component 213
source control 756
  adding files 757, 758
  checking in files 760
  checking out files 759, 760
  ClearCase 731
  CVS 733
  history 762
  SourceSafe 755
  undo check out 761
source management 721
  ChangeMan DS 727
  ClearCase 729
  Concurrent Versions System (CVS) 731
  configuring 727
  Perforce 735
  SCC control 723
  source control 756
  Subversion 739
  Team Foundation Server (TFS) 746
  Visual SourceSafe 753
source window 213
SourceSafe 755
SQL Calculator 221, 223, 254–256
SQL commands 265
  editing 265
  testing 265
SQL functions 258, 259
SQL operators 258
SQL Server
  tables 657, 679
SQLAnswerSet component 265
standalone process flows 392
standalone stored procedures 396
Start object 145
statistical variables 339, 405
statistics reports 468
  columns 469
Stored procedure object 146
stored procedures 463
  adding to process flows 395
  creating 396, 463
  executing 389
  storing 195
Subversion 739
summary flow reports 468, 472
Synonym Editor 501
  Segment shortcut menu 504
  Access File Text View tab 512
  changing column display 143
  Column shortcut menu 507
  Field View tab 502
  file attributes summary 516
  File Listener attributes 526
  List View tab 504
  Metadata tab 513
  Modeling View tab 511, 536
  segment attributes summary 517
  Segment View tab 503
  synonym attributes 515
  Text View tab 512
  virtual columns 540, 541
synonyms 33, 71, 72, 438
  copying 137
  creating 71
  creating using Modeling View 536
  data lineage 552
  data profiling 543
  data sources and 108
  deleting 138
  editing 200
  guidelines 108, 109
  indexes 539
  key analysis 545
  log tables 673
  migrating 102
  printing 512
  prototype file 532
  testing 71, 77
  using variables 533
synonyms 33, 71, 72, 438  
virtual columns 540, 541  
system variables 337

T  
tabs 130  
Metadata 513  
Data Flow 139  
flow 153  
Home 151  
print preview 486  
Process Flow 144  
report 158  
text editor 159  
Text View 147  
user functions 157  
targets 276, 283, 309, 310, 318, 324  
flat files 318, 321  
FOCUS files 310, 314  
formatted files 322  
Fusion files 310, 314  
relational 276, 283, 309, 312  
Transformations dialog box 332  
Team Foundation Server (TFS) 746  
temporary columns 324  
Teradata 305  
Test Transforms tool 331, 332, 443  
testing flows 83  
testing metadata 77  
testing SQL 265  
testing synonyms 77  
text columns 108, 112  
text editor 147  
Text Editor tab 159  
Text View tab 147, 387, 512  
trace 715  
tracking results 462  
Transformation Calculator 233, 442  
functions 347  
order of evaluation 233  
Transformations dialog box 332  
troubleshooting 709  
server messages 710

U  
updating passwords 89  
upgrading DataMigrator 102  
User Functions tab 157  
user preferences 172  
user profiles 57  
using the Transformation Calculator 442

V  
validating data 455  
VARCHAR 108  
variable length character columns 108  
variables 336, 337, 340, 404  
changing display 142  
statistical 339, 405  
system 337, 340
verifying data 440

version control
  adding files 757, 758
  checking in files 760
  checking out files 759, 760
  ClearCase 731
  CVS 733
  history 762
  SourceSafe 755
  undo check out 761

viewing code 387

viewing flow reports 488, 489

virtual columns 441
  adding 540

virtual columns 441
  creating 541

Visual SourceSafe 753, 755

W

Wait conditions 420

Wait object 146, 419

Web Console 53
  accessing 53
  launching 53

WebFOCUS 708
  executing flows 708
Feedback

Customer success is our top priority. Connect with us today!

Information Builders Technical Content Management team is comprised of many talented individuals who work together to design and deliver quality technical documentation products. Your feedback supports our ongoing efforts!

You can also preview new innovations to get an early look at new content products and services. Your participation helps us create great experiences for every customer.

To send us feedback or make a connection, contact Sarah Buccellato, Technical Editor, Technical Content Management at Sarah_Buccellato@ibi.com.

To request permission to repurpose copyrighted material, please contact Frances Gambino, Vice President, Technical Content Management at Frances_Gambino@ibi.com.