

# **Technical Bulletin**

**Part No. 74-0127**

## **DataStage Informix XPS Load**

This technical bulletin describes Release 1.2 of the DataStage Informix XPS Load Plug-in, formerly known as the DataStage Informix XPS Bulk Load Plug-in. This stage loads data into an Informix Dynamic Server with Advanced Decision Support and Extended Parallel Options (Informix XPS) table.

# NOTICE

Copyright © 2003, 1999–2002  
Ascential Software Corporation  
All rights reserved.

© 2003 Ascential Software Corporation. All rights reserved. Ascential, Ascential Software, DataStage, MetaStage, MetaBroker, and Axielle are trademarks of Ascential Software Corporation or its affiliates and may be registered in the United States or other jurisdictions. Adobe Acrobat is a trademark of Adobe Systems, Inc. IBM and Informix are either registered trademarks or trademarks of IBM Corporation. Microsoft, Windows, Windows NT, and Windows Server are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries. UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company, Ltd. Other marks mentioned are the property of the owners of those marks.

This product may contain or utilize third party components subject to the user documentation previously provided by Ascential Software Corporation or contained herein.

## Printing History

First Edition (74-0127) for Release 1.0, February 1999  
Second Edition (74-0127) for Release 1.1, March 1999  
Third Edition (74-0127) for Release 1.2, September 2001  
Updated for Release 1.2, December 2001  
Updated for Release 1.2, August 2002  
Updated for Release 1.2, August 2003

## How to Order Technical Documents

To order copies of documents, contact your local Ascential subsidiary or distributor, or call our main office at (508) 366-3888.

Documentation Team: Marie E. Hedin

## Introduction

This technical bulletin describes the following for Release 1.2 of the Informix XPS Load stage, updated for DataStage Release 7.0:

- Functionality
- Configuration requirements
- Load methods
- Defining data types
- Properties

The Informix XPS Load stage is a passive stage that loads data into an Informix XPS table. This stage has one input link and no output links. The input link provides the rows of data to be loaded into the Informix XPS table.

## Functionality

Informix XPS Load has the following functionality:

- Support for data files that exceed the 2-GB file size limit for 64-bit file systems.
- Support for MetaStage. For more information, see *MetaStage User's Guide*.
- NLS (National Language Support). For information, see *DataStage NLS Guide*.
- Bulk loading from a stream input link to provide rows of data into the target table.
- Bulk loading using sequential files and named pipes.
- Immediate and delayed loading.
- Target table creation.
- Express and deluxe load modes.

The following functionality is not supported:

- Output links
- Meta data import

## Installing Informix XPS Load

For instructions and information supporting the installation, see *DataStage Plug-In Installation and Configuration Guide*.

## Load Methods

The two methods of loading data into an Informix XPS table are as follows:

- Sequential file method
- Named pipe method

The value for the Load Method property determines which method to use to load the data. The value for the sequential file method is DISK, and the value for the named pipe method is PIPE.

### Sequential File Method

When you load data using the sequential file method, rows from the input link are written in delimited format into an intermediate data file. The data file is loaded immediately or it can be deferred, depending on the value of the Automatic Load property.

**Immediate.** When you set Automatic Load to Yes, the DataStage job automatically loads the table. With automatic loading using a sequential file, a data file consisting of rows from the input link is constructed and is loaded into the target table after all rows have been read from the input link. Automatic loading using the sequential file method is slower than the named pipe method because the data cannot be loaded until this data file is constructed.

**Deferred.** When you set Automatic Load to No, a data file consisting of rows from the input link is constructed, but the data is not loaded into the target table. Additionally, a file containing SQL statements is generated. You can manually edit the data file and the SQL file for advanced customization of the load process and filtering of input data before loading. You can then move these files to any Informix XPS host system for execution.

The data file can be loaded manually later by using the SQL file as input to the *dbaccess* utility. Use this method if the DataStage server does not reside on a system hosting an Informix XPS server or coserver.

### Named Pipe Method

When you load data using the named pipe method, the rows from the input link are streamed continuously into the named pipe for loading the data until the end of the data. Use this method when you need to load the rows from the input link immediately as they are streamed into the pipe. There is no delay between the time during which the rows are received from the input link and the start of the actual load process.

## Loading a Target Table

Informix XPS Load supports the following three values of the Load Action property for loading a target table. For an explanation of how DataStage columns are translated into Informix columns, see “Creating the Target Table” on page 5.

- **TRUNCATE.** This default value deletes all existing rows from the table before starting the load.
- **CREATE.** Informix XPS Load creates the table on the target database. It supports table creation using properties that specify table type, fragmentation strategies, and storage allocation. For an explanation of how DataStage columns are translated to Informix columns, see “Creating the Target Table” on page 5.
- **APPEND.** This value loads its data regardless of whether the target table already contains data. Constraint violations may result. However, any rejected rows are written to a configurable reject file.

For more information about the Load Action property, see “Properties” on page 6.

## Load Modes

The Load Modes property allows tables to be loaded in express or deluxe mode. If this property is not specified, the default load mode is based on the table type.

Consult *Informix XPS Administrator's Guide* for further information regarding the various actions for loading XPS tables.

## Defining Data Types

The DATETIME and DATE data types need special consideration.

### DATETIME Considerations

In Informix, a valid DATETIME format can be either *yyyy-mo-dd hh:mi:ss.ff* or any contiguous subset of this string, for example, *mo-dd hh*. Any column defined to contain such a subset would not accept a full ANSI DATETIME.

Therefore, the DATETIME type generated within the data file should correspond to the DATETIME type defined for the actual column in a table. To be consistent with the Informix DDL, you should use the Datetime Format property to specify a default format for all DATETIME columns. Then specify a DATETIME format in the Description field of the column definition grid for every DATETIME column

that differs from the default. The following table shows some examples for specifying this property:

<b>DATETIME Column Description Expression</b>	<b>Description</b>
Year to Fraction(3)	The default value (does not have to be entered). Full ANSI DATETIME with a precision of 3.
Hour to Second	A partial ANSI DATETIME type, starting from hours and ending at seconds.
Month to Minute	A partial ANSI DATETIME type, starting from months and ending at minutes.
Hour to Fraction(5)	A partial ANSI DATETIME type, starting from hours and ending at fractions with a precision of 5.

## DATE Considerations

Informix uses the DBDATE or GL\_DATE environment variables for interpreting and displaying the string representation of the DATE data type. For example, if DBDATE is set to DMY4, then the *dd-mm-yyyy* string is interpreted as *dd* day of *mm* month of *yyyy* year.

Reading DBDATE or GL\_DATE environment variables automatically for date formatting by Informix XPS Load is not an option, since you can run DataStage jobs on a physical machine different from the one where the Informix XPS server is installed.

The default DBDATE Informix setting is MDY4. Informix understands the date *yyyy-mm-dd* correctly only if the DBDATE setting is Y4MD. Otherwise, an error occurs when converting a string to a date. Independent of the DBDATE setting, Informix correctly interprets the following:

- Date separator as - or /
- Two-digit as compared to four-digit year (MDY4 is equivalent to MDY2 dates such as *mm-dd-yy*)

Use the Date Format property to specify the format of date strings that the Informix XPS Load stage sends to Informix. The value of this property should correspond to the value of DBDATE or GLDATE for the target database.

## Creating the Target Table

If the value of the Load Action property is CREATE, Informix XPS Load creates the table on the target database if it does not already exist.

**Note:** Informix XPS Load does not drop the table if it already exists. The job will abort.

If you do not want automatic loading, the CREATE TABLE SQL statement is written to the SQL file.

The table is created using the table name specified in the Table Name property and the column names specified for the column definitions. If the table has a primary key, it is created using the column order indicated by the column definitions. Manually add any additional table attributes such as constraints following the load. Consult *Informix XPS Administrator's Guide* for additional information.

The following table describes how DataStage column data types are translated to Informix column data types:

DataStage	Informix
BigInt	INTEGER
Binary	BYTE
Bit	Not supported
Char	CHAR
Date	DATE
Decimal	DECIMAL
Double	FLOAT
Float	FLOAT
Integer	INTEGER
LongVarBinary	BYTE
LongVarChar	VARCHAR
Numeric	DECIMAL
Real	REAL
SmallInt	SMALLINT
Time	DATETIME hour to second
Timestamp	DATETIME year to fraction (3)
TinyInt	SMALLINT

DataStage	Informix
VarBinary	BYTE
VarChar	VARCHAR

## Properties

The following table includes these column heads:

- **Prompt** is the text that the job designer sees in the stage editor user interface.
- **Type** is the set of values that are valid for the property.
- **Default** is the text used if the job designer does not supply any value.
- **Description** describes the properties.

Each property is described in the order in which you might use it. You can also access this information from the stage property page in the DataStage Designer.

Prompt	Type	Default	Description
Load Mode	String List	DEFAULT	The method used to load the data into the target file. If set to DEFAULT, "Create Table Type" determines the loading mode when "Load Action" is set to CREATE. Otherwise, it is determined by the type of existing table. If "Load Action" is set to CREATE, set "Create Table Type" to a value compatible with the desired loading mode. Use express-mode loading for faster loads. Express-mode loading is only possible when "Create Table Type" is set to RAW or OPERATIONAL and has no indexes. Use deluxe-mode loading when "Create Table Type" is set to OPERATIONAL or STANDARD. (DEFAULT/EXPRESS/DELUXE)
Server Name	String		The name of the Informix XPS server on which the target table resides.
Database Name	String		The name of the database on which the target table resides.

Prompt	Type	Default	Description
User Name	String		The user name if you connect to the database as a user other than the user who executes the DataStage server process. If "User Name" is not set, the user name for the DataStage server process is used.
Password	String		If "User Name" is set, the password for this user name. If not set, "Password" is ignored.
Table Name	String		The name of the Informix table to load.
Automatic Load	String List	Yes	Specifies whether tables are automatically loaded by the DataStage job. This occurs only when the DataStage server resides on the same system as an Informix coserver. If set to No, the stage generates an external SQL file containing a sequence of SQL commands that can be manually edited and executed on any Informix host system. (Yes/No)
Load Method	String List	PIPE	The method of loading data into an Informix XPS table. Named pipe loading (set to PIPE) can be faster because data is loaded as it is written to the output stream. Set to PIPE only if the DataStage server is running on the same system as an Informix coserver. Otherwise, set to DISK to retain a copy of the loaded data in a sequential file or if you do not want automatic loading. (PIPE/DISK)
Load Action	String List	TRUNCATE	Specifies how data should be loaded to the target table. Set to CREATE if the target table does not exist. The load fails if the target table already exists. The XPS Load stage will not drop any existing tables on the target database. Set to TRUNCATE to delete all existing rows before starting the load. Set to APPEND to add rows to the table. Unique or primary key constraint violations can result. (CREATE/TRUNCATE/APPEND)

Prompt	Type	Default	Description
Create Table Type	String List	STANDARD	Specifies the type of table to create if "Load Action" is set to CREATE. The default load mode depends on the type of table being loaded. Express mode loads can occur only if set to RAW or OPERATIONAL. This property is ignored if "Load Action" is not set to CREATE. (STANDARD/RAW/OPERATIONAL/STATIC)
Fragmentation Strategy	String List	NONE	The fragmentation strategy to use if "Load Mode" is set to CREATE. NONE specifies that no strategy is used. NONE/ROUNDROBIN/EXPRESSION/HASH/HYBRIDS)
Fragmentation Rule	String		The clause that corresponds to the value for "Fragmentation Strategy." For example, for a value of EXPRESSION, enter a clause like the following: <i>col_id &lt; 100 IN dbbsp1, col_id &gt;= 100 IN dbbsp2</i> If "Fragmentation Strategy" is set to NONE, you can still specify a value for this property to direct table creation in the specified <i>dbspace</i> or <i>dbslice</i> .
Extent Size	Long	16	The minimum length is four times the disk page size for your system. The maximum length equals the chunk size.
Table Lock Mode	String List	EXCLUSIVE	The type of locking for user access to the table during loading. Set to EXCLUSIVE so no users can access the table while it is being loaded. Set to SHARED so other users can continue to read from the table while it is being loaded. (EXCLUSIVE/SHARED)
Directory Path	String		The pathname of the directory where the data file, reject file, and SQL file will be created. If not set, it defaults to the current DataStage project directory.

Prompt	Type	Default	Description
SQL File	String		The file generated when "Automatic Load" is set to No. It contains the sequence of SQL commands required to load an Informix AD/XP table using high-performance loading from external tables. You can modify and execute this file on any Informix AD/XP host system. If not set, a value is generated as <code>&lt;server_name&gt;_&lt;database_name&gt;_&lt;table_name&gt;.sql</code> . It is created in the directory specified by "Directory Path."
Data File	String		Identifies the sequential file created when "Load Mode" is set to DISK, or when automatic loading is not desired. If not set, a value is generated as "SQL File" with an extension of <code>.dat</code> .
Reject File	String		Identifies the reject file created if any rows are rejected during the loading process. These rows are written to this file. If not set, a value is generated the same way as for "SQL File" with an extension of <code>.rej</code> .
Delimiter	String	 (vertical bar)	The character to be used to delimit fields in the loader input data. This value is also used in the CREATE EXTERNAL TABLE statement.
Date Format	Long	1	Use one of the following date formats: 1 – <code>mm-dd-yyyy</code> (default) 2 – <code>dd-mm-yyyy</code> 3 – <code>yyyy-mm-dd</code> The Date type should correspond to the type defined in the DBDATE environment variable.
Datetime Format	String	Year to Fraction(3)	Valid DATETIME formats mirror Informix DDL syntax for a DATETIME type, for example, Hour to Fraction(2). The DATETIME type should correspond to the DATETIME defined for the actual column in a table. You can override the value defined in the present stage property within a description column option of a column. The default is Year to Fraction(3).
Maximum Bad Rows	Long	10	Sets the number of errors allowed per coserver before the database server stops the load.

<b>Prompt</b>	<b>Type</b>	<b>Default</b>	<b>Description</b>
Hex Escapes	String List	No	Directs the database server to recognize ASCII special characters embedded in ASCII text-based data files.
Before-link Routine	String		A command to be executed by the host system. It can be an executable, a UNIX shell script, or a DOS batch file. Its format must be understood by the host operating system. The stage does not perform any parsing or syntax checks.
After-link Routine	String		A command to be executed by the host system. It can be an executable, a UNIX shell script, or a DOS batch file. Its format must be understood by the host operating system. The stage does not perform any parsing or syntax checks.