Informatica Fast Clone FAQs

Abstract

This article describes frequently asked questions about using Informatica Fast Clone for bulk data movement. It includes information about Fast Clone features and some common errors.

Supported Versions

- Informatica Fast Clone 6.6

Table of Contents

General Questions

What is Fast Clone?

Fast Clone is a high-performance cloning tool for moving bulk data from Oracle databases to heterogeneous destinations, including relational databases and flat files. The optional DataStreamer component can stream data to Greenplum and Teradata destinations, which avoids intermediate storage use and reduces I/O. You can use Fast Clone for any of the following purposes:

- Cloning Oracle databases.
- Moving Oracle data to another type of platform.
- Migrating production data into a test environment.
- Initially loading data into tables that will be the targets of Informatica Data Replication jobs, before starting transactional data replication. In this case, Fast Clone is a high-speed alternative to the Data Replication InitialSync component.

What is the difference between the direct path unload and conventional path unload methods?

Fast Clone has two methods of unloading data from an Oracle source: direct path unload and conventional path unload. The direct path unload method is much faster. It extracts source metadata to physical files and reads Oracle data files directly. Also, it can stream data to Greenplum and Teradata destinations. The conventional path unload method uses the standard Oracle Call Level Interface (OCI) to read data and metadata. With this method, you can extract data over the network and read data from additional types of Oracle objects. The following lists summarize key features of each method.

Direct path unload:

- Can unload data much faster.
- Supports Fast Clone DataStreamer, which streams data to Greenplum or Teradata destinations.
- Requires Fast Clone to be installed on the Oracle source system.
- Requires Fast Clone to have read access to the Oracle data files to read data directly.
- Requires the Oracle user to have special user permissions, such as SELECT FROM DICTIONARY.
- Does not extract Oracle views, cluster tables, and index-organized tables (IOTs)

Conventional path unload:

- Uses the Oracle Call Level Interface (OCI) to extract data.
- Allows you to run Fast Clone remotely and extract data over the network.
- Does not require the Oracle user to have special user permissions.
- CPU consumption is higher when Fast Clone runs on the Oracle server.
- Does not support DataStreamer.
- Extracts Oracle views, cluster tables, and index-organized tables (IOTs)
- Extracts user-defined SQL queries including table joins.

Can I use multiple Fast Clone Servers? Can Fast Clone Servers communicate with one another?

You can use multiple Fast Clone Servers, for example, one to initiate extraction requests and another to store output data and control files. A Fast Clone Server can communicate with another Fast Clone Server as well as with Fast Clone instances on the source and destination.

Does Fast Clone Manager have to run on the Oracle source system if I use the direct path unload method?

No. You can run the Fast Clone Manager GUI on another Linux, UNIX, or Windows system that has a supported operating system. The GUI uses the appropriate JDBC drivers to connect to the source to get metadata and optionally to connect to targets to verify connection and database owner information. Fast Clone provides JDBC drivers for most destinations in its *installation*\lib subdirectory. However, you must download JDBC drivers for MySQL 4.1 and later, Netezza, and Teradata. After you create a bulk data movement configuration file from the remote Fast Clone Manager, you can transfer the configuration file to the Oracle source system and then run Fast Clone engine from there.

What is QZip?

Fast Clone can compress data in an output file during unload processing to reduce the file size. In Fast Clone Manager, on the **Runtime Settings** tab > **Miscellaneous Conditions** view, select **Compress data on the fly** and then select the compression method. If you select QZIP compression, you can configure the compression ratio. To decompress the data later, use the QZip executable.

Can Fast Clone move data from a single Oracle instance to multiple destinations?

If you use DataStreamer, Fast Clone streams data from a single Oracle source to a single Greenplum or Teradata destination. If you do not use DataStreamer, Fast Clone generates a loader script for a single destination based on the configuration file that you create from Fast Clone Manager. By creating multiple bulk data movement configurations, each with the same source but a different destination, you can generate multiple loader scripts. You can then run the scripts to load the source data to multiple destinations.

How do I start Fast Clone Manager?

The Fast Clone Manager GUI executable file, gui.sh or gui.cmd, is in the top-level Fast Clone installation directory.

- On Windows, run gui.cmd. You can double-click gui.cmd, run gui.cmd in the Run window, or create a shortcut for it on your desktop.
- On Linux or UNIX, run gui.sh.

The GUI requires the Java Runtime Environment (JRE) 1.3.7 or later. On Linux and Windows, the GUI also requires an X Window environment. If you cannot configure X Window on your system, contact Informatica Global Customer Support for a workaround.

How do I start the Fast Clone engine to unload and move data?

After you configure a bulk data movement job, you can start the Fast Clone engine to move data, in any of the following ways:

- From a Windows command prompt, run unload.cmd.

- From a Linux or UNIX command line, enter xunload.sh to run Fast Clone in the foreground or enter unload.sh to run Fast Clone silently in background.
- In Fast Clone Manager, click the green arrowhead button on the toolbar.

For which destination-database loaders does Fast Clone generate input?

Fast Clone generates load control scripts for the following destination-specific loaders:

Destination	Loader
EMC Greenplum Database	gpload or gpfdist utility
IBM DB2 for Linux, UNIX, and Windows	db2 command on Linux or UNIX or the db2cmd command on Windows
Microsoft SQL Server	SQL Server bcp load utility
MySQL RDBMS	mysql command-line tool
Netezza	nzload utility **Note:** You must specify the nzload utility in the PATH environment variable.
Oracle	SQL*Loader (sqlldr) utility
PostgreSQL	psql too
Sybase ASE	Sybase bcp utility
Teradata Database	Teradata FastLoad or Multiload utility
Vertica	SQL COPY command

These load utilities are usually part of the destination database software and do not need to be installed separately. However, if you want to run the load operation from the source, you must install the loader there.

If you use Fast Clone DataStreamer, you need the following external loaders on the destinations:

Destination	Loader
Greenplum	gpfdist utility
Teradata	Teradata Parallel Transporter (TPT) libraries with the load and update operators that are equivalent to the FastLoad and MultiLoad utilities, respectively

Does Fast Clone work with Informatica Data Replication?

You can use Fast Clone instead of the Data Replication InitialSync component to perform an initial load of target tables from an Oracle source, before starting change data replication the first time. Fast Clone is faster than InitialSync, particularly if you use the direct path unload method.

In the following situations, Data Replication can augment Fast Clone functionality:

- Fast Clone assumes the source and destination have the same structure. With Data Replication, the source and destination structure can be different, and you can map individual source columns to target columns.

- Fast Clone exports source schema to Oracle and MySQL destinations only. Data Replication can export source schema to all supported destination types.

For more information about using Data Replication, see the *Informatica Data Replication User Guide*.

Does Fast Clone support versioning of the extracted data across copy operations?

No. Versioning is not supported.

Configuration Questions

How can I improve Fast Clone performance when cloning large amounts of data, such as 400 GB?

Run Fast Clone on the Oracle source system and use the direct path unload method, if possible.

Also, avoid defining many selection criteria for the source. Fast Clone processing of WHERE clauses can slow processing. As a general rule, if the selection criteria is for 20 percent or less of the source data, a conventional path unload is faster than a direct path unload.

If the system has multiple CPUs, Fast Clone can perform parallel extraction processing. To tune parallel extraction processing, set the following parameters in Fast Clone Manager, on the **Runtime Settings** tab > **Parallel Execution** view:

- **Threads Number**. The maximum number of tables or partitions from which data can be extracted in parallel.
- **Threads per Segment**. The maximum number of threads that extract data from a single table or partition.

By multiplying these values, you get the total number of extraction threads that can be used. If you need to extract data from many tables of similar size, set **Threads Number** to the appropriate concurrency and set **Threads per Segment** to 1. If you need to extract data from a few large tables, set **Threads Number** to the number of tables and set **Threads per Segment** to the appropriate concurrency. Usually, for direct path unload, the appropriate concurrency is the number of CPUs or cores x 2, and for conventional path unload, it is the number of CPUs or cores. However, if I/O bandwidth or the number of CPUs that can be allocated is restricted, you might need to lower the value.

To extract source data into ASCII flat files, do I need to define a destination on the Destination DB tab in Fast Clone Manager?

No. In Fast Clone Manager, click the **Destination DB** tab and then click **No Destination, only FLAT FILE or PIPE as Output**. Fast Clone then does not generate a script for a database loader.

Does Fast Clone have any limitation on LOB data extraction?

If you use the conventional unload method, Fast Clone has no limitation on the amount of data that it can extract from a column with a LOB datatype. If you use the direct path unload method, Fast Clone extracts up to 96 KB of data from a LOB column and then switches to the conventional path unload method to complete data extraction.

Can Fast Clone extract source data into files in EBCDIC or COMP-3 format for z/OS targets?

Yes. Fast Clone can extract data in EBCIDIC and COMP-3 format, in addition to ASCII and Teradata or Vertica binary format, with either the direct path unload or conventional path unload method. In Fast Clone Manager, on the **Runtime Settings** tab > **Format Settings** view, you can use the following options to specify the EBCDIC or COMP-3 output format:

- **Convert numbers into packed decimal COMP3**. Select this option if you need to read the output file with COBOL or another language that has built-in COMP-3 support.
- **COMP3 implied decimal**. The COMP-3 implied decimal value. Default is 0.
- **Convert output ASCII into EBCIDIC**. Select this option if you need the output in IBM EDCIDIC format.

To load data to a DB2 for Linux, UNIX destination, are there any special configuration considerations?

For DB2 destinations, consider the following configuration issues:

- DB2 LOAD requires that the line separator be the new line character (\n). If you have this character in your Oracle data, the load fails. To prevent this problem, specify the Fast Clone command-line parameter REPLACE_NEWLINE_WITH_SPACE=Y. This parameter replaces new line characters in data with spaces.
- DB2 LOAD requires the column delimiter and "enclosed by" character to be a single ASCII character. You can use one of the nonprintable ASCII characters that are documented in the *Informatica Fast Clone User Guide*.
- If you have LOB column data that is smaller than 64 KB and that does not contain row delimiter characters, you can extract this data to the main output file. In Fast Clone Manager, on the **Runtime Settings** tab > **Format Settings** view, clear the **Unload binary into separate files** option. If you have LOB column data that is greater than 64 KB or that includes delimiter characters, you must extract each LOB column into a different output file. In this case, select **Unload binary into separate files**.

To load data to a Netezza destination, does Informatica recommend any special configuration settings?

In Fast Clone Manager, on the **Runtime Settings** tab > **Format Settings** view, set the following options:

- In the **Column separator** field, enter a single character that does not occur in extracted text data from your database. You can use one of the nonprintable ASCII characters that are documented in the *Informatica Fast Clone User Guide*.
- In the **Record separator** field, enter /n as the end-of-line character.
- In the **Enclosed By** field, delete the default double quotation mark (") character. An empty field makes extraction and nzload processing faster.
- For extraction of data that includes the end-of-line character into text-delimited files, select the **Replace new line with space** option.
- For **Timestamp format**, enter YYYY-MM-DD HH24:MI:SS.FF6, and for **Date format**, enter YYYY-MM-DD HH24:MI:SS. The nzload utility supports these ISO date formats.
- If you extract data from CLOB or LONG columns, clear the **Unload binary into separate files** option. The nzload utility can load up to 32,768 bytes of data into VARCHAR type columns.

To load data to a Teradata destination, does Informatica recommend any special configuration settings?

In Fast Clone Manager, on the **Runtime Settings** tab > **Teradata Load Settings** view, set the following options:

- In the **Loader to be used** field, select the Teradata load utility that you want to use. If you select Teradata FastLoad or MultiLoad, Fast clone generates a loader script.
- If you want to the output data file in Teradata binary format instead of text-delimited format, select **Binary length prefix**.
- If you selected binary format, in the **Binary length prefix endianness** field, select the endianness of the destination platform. This value determines the format of the integer that specifies the length of records and fields in the binary files.

On the **Runtime Settings** tab > **Format Settings** view, set the following options:

- In the **Record separator** field, enter /n as the end-of-line character for text-delimited files. For binary files, this field is ignored.
- In the **Enclosed By** field, delete the default double quotation mark (") character so that the field is empty.
- For extraction of data that includes the end-of-line character into text-delimited files, select the **Replace new line with space** option.
- Clear the **Suppress trailing null columns** option to enable proper loading of either text-delimited or binary data files into Teradata.

- In the **Truncate numeric precision integer** field, enter 18. Oracle NUMBER columns can have a maximum of 36 significant digits. Teradata DECIMAL columns can have a maximum of 18 digits. Ths parameter rounds extracted numbers to 18 digits.

Questions about Oracle Sources

How do I configure Fast Clone to unload data from Oracle views?

To unload data from Oracle views, you must use the conventional path unload method. You can specify a source view in one of the following ways:

- After you define the source and destination in Fast Clone Manager GUI, edit the generated unload.ini configuration file. Under the [SOURCE_INDIRECT_TABLES] section, add an entry for the view. For example:

  ```
  TABLE1=view1
  ```

  To define a filter that selects a subset of columns, also enter the TABLE1_COLUMN_LIST statement under [SOURCE_INDIRECT_TABLES_COLUMNS_LIST] section.
- In Fast Clone Manager, manually define a SQL query that extracts data from a named view. On the menu bar, click **Data** > **Add Conventional SQL** and then enter the query.

For more information, see the *Informatica Fast Clone User Guide*.

Can Fast Clone read data from Oracle index-organized tables (IOTs)?

If you use the conventional path unload method, Fast Clone can extract data from IOTs. However, if you use the direct path unload method, Fast Clone cannot extract data from IOTs.

Does Fast Clone support Oracle Real Application Clusters (RACs)?

Yes. Fast Clone can extract data from an Oracle instance in a RAC environment with either the direct path unload or conventional path unload method. In the Fast Clone Manager GUI, on the **Source DB** tab > **Source Database** view, select the **RAC connect support** option.

Can the Fast Clone direct path unload method extract data from Oracle data files that are managed by Oracle Automatic Storage Management (ASM)?

Yes. Fast Clone can connect to the ASM instance to extract Oracle data, in a RAC or non-RAC environment. With the direct path unload method, you must define ASM connection information. In Fast Clone Manager, on the **Source DB** tab > **ASM Settings** view, enter the ASM instance name, host name, port number, user ID, and password.

Can Fast Clone extract Oracle compressed data?

Yes. Fast Clone can extract Oracle-compressed data with either the direct path unload or conventional path unload method. Fast Clone supports the compression capabilities of the Oracle Advanced Compression Option, including table compression and LOB compression.

Can Fast Clone extract data from Oracle materialized views?

Yes. Fast Clone can extract data from Oracle materialized views.

Can Fast Clone extract data from an Oracle Data Guard logical standby databases?

Yes. Fast Clone can extract data from an Oracle Data Guard logical standby database.

Does Fast Clone support Oracle TDE or TSE encryption?

No. Fast Clone does not extract data from tablespaces or columns that are encrypted with Oracle TDE or TSE.

What user privileges does Fast Clone require to connect to the Oracle source database to read data?

For Fast Clone extraction of Oracle data, configure user permissions based on whether you use the conventional unload or direct unload method:

- If you use the conventional unload method, which leverages the OCI, you do not need to configure any special user permissions.
- If you use the direct unload method, you must run extractions under a user in the DBA group or create a special Fast Clone user that has the required Oracle system and object permissions. If you create a special Fast Clone, grant the following system privileges to that user:

```
create role fastclone_direct_role;
grant alter system to fastclone_direct_role;
grant execute on dbms_flashback to fastclone_direct_role;
grant select any dictionary to fastclone_direct_role;
```

  Also, for the Fast Clone Manager GUI to connect to Oracle, log in under a user account that is in the DBA group or that has the following Oracle system privileges:

  ♦select any table

  ♦select any dictionary

Troubleshooting

Why does the Oracle error ORA-12528 occur when Fast Clone attempts to do a direct path unload of Oracle data in ASM-managed storage?

You need to add the ASM instance to your listener.ora configuration file, or equivalent, and then restart the Oracle listener.

In the listener.ora configuration file, add the ASM instance name to the SID_LIST. For example:

```
SID_LIST_LISTENER =
  (SID_LIST =
...
OTHER INSTANCES HERE
....
    (SID_DESC =
     (SID_NAME = +ASM)
     (ORACLE_HOME = oracle_path)
    )
  )
```

Then restart the Oracle listener and run FastReader again. From this point on, the ASM instance will be accessible from the Oracle listener.

Why does the Fast Clone Manager GUI stop responding on Linux?

Fast Clone Manager is a Java application that requires the JRE 1.3.7 or later. If the GUI hangs, verify that a supported JRE version is installed. You can download the JRE from Oracle. Then perform one of the following actions:

- Before starting the GUI, set the JAVA_HOME environment variable:

  - For a tcsh shell:

    ```
    setenv JAVA_HOME $ORACLE_HOME/jre/1.5/
    ```

  - For a bash shell:

    ```
    export JAVA_HOME=$ORACLE_HOME/jre/1.5/
    ```

- Edit the gui.sh file to add the full path to the java executable.

Author

Virginia Pfeifle
Principal Technical Writer