Best Practices for High-Speed Data Transfer and Migration from Oracle to Teradata using FastReader

Overview

WisdomForce FastReader™ (www.wisdomforce.com) is high performance data transfer and migration software for high volumes of Oracle data. FastReader supports direct path and parallel extraction and unloads large Oracle databases in a fraction of the usual time into ASCII flat files or pipe.
Teradata RDBMS (www.teradata.com) is a massively parallel processing system running a shared nothing architecture. Teradata provides integrated, optimized and extensible technology for a single application-neutral repository of your current and historical data, forming the framework of the business intelligence architecture.

This document describes important points for unloading data from Oracle using FastReader and loading into Teradata. When Teradata is chosen as a destination in FastReader, load scripts for the Teradata loader utilities are generated during the extraction. FastReader is capable of producing output in Teradata binary format suitable for loading using the FastLoad and Multiload utilities.

Output Formats

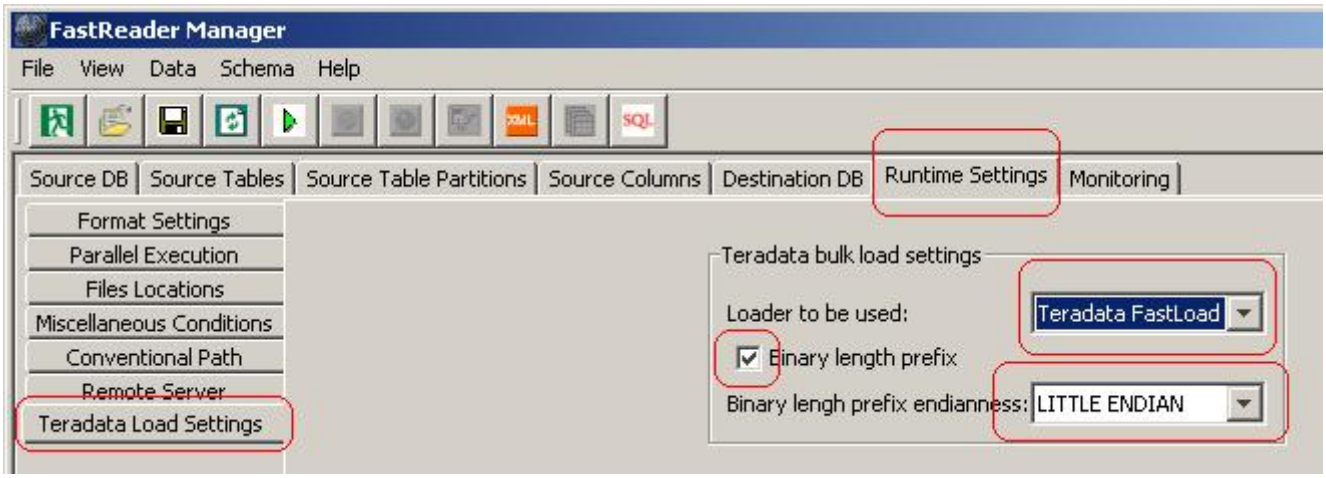
FastReader normally extracts data into text delimited files with customizable delimiters and formatting. An additionally output format is the Teradata binary format which can be used to efficiently load data into a Teradata database. In the binary format each record is prefixed by a 2 byte integer specifying the total length of the record minus 2 bytes, each field is prefixed by a 2 byte integer specifying the length of column data. FastReader extracts all columns, including numeric and date columns, in textual format.

Recommended Parameters

The following parameters control the format and the generation of control files for the extraction:

- The parameter **teradata_loader_to_use** determines whether the loader script will be generated for the Teradata FastLoad or MultiLoad load utility using the appropriate syntax. The loader scripts are generated for both text delimited and binary format files. **{teradata_fast_load|teradata_mload}**
- The parameter **binary_lengh_prefix** specifies if the output file format will be in Teradata binary format. (true for binary format/false for regular text delimited format) **{true|false}**
- The parameter **binary_lengh_prefix_little_endian** determines the format of the 2 byte integer specifying the length of the records and fields in binary formatted files. This might be required for loading on big-endian hardware platforms. **{true|false}**
- The parameter **suppress_trailing_nullcols** needs to be **false** to enable proper loading into Teradata for both file formats.
- The parameter **truncate_numeric_precision** should preferably set to 18 for Teradata.

- The **enclosed_by** delimiter parameter should be empty for Teradata (it has a double quote " by default).
- For text delimited format the end of record character must by "\n" (end of line). For binary format all the delimiter parameters except **enclosed_by** are ignored.
- For extraction into text delimited format if the column data contains "end of line" characters the parameter **replace_new_line_with_space** should be true. (This parameter is not required for binary format)



Extraction & Load Considerations

- *Record length* – in binary format, the total record length is limited by 65536 bytes. Take into consideration the fact that FastReader extracts integer and date values as text, which may further increase the record length. If for a particular record the record length is exceeded some columns of the record may be extracted as null in order to prevent failure of the entire load attempt.
- *Numbers and Precision* – the number data type in Oracle allows a maximum of 36 significant digits while in Teradata the decimal data type has a maximum of 18 significant digits. The **truncate_numeric_precision** parameter in FastReader allows to round extracted numbers up-to a specified amount of significant digits. For example if the parameter is set to 18, the following number 129.333333333333366733333333333334543457 will be extracted as 129.333333333333367. FastReader also allows specifying custom precision for each numeric column individually.
- *Handling of Nulls* – when extracting into binary format a null column will be extracted as a column of size 0. The control files generated for Teradata FastLoad and MultiLoad use the "nullif" clause to load Null values into Teradata. This works correctly because in Oracle a zero length column is by definition a Null column
- *Date and Timestamp columns* – in Oracle the date data type contains a time part, where as in Teradata the date data type only contains a date part. Consider importing Oracle date columns into Teradata timestamp columns in on the staging area or in the destination tables. You can also specify explicit date formatting for each column for extracting.
- CLOBs – Teradata load utilities limit loading of CLOB columns up to a maximal size of 32K.