

14kettle

5000 + 30%1500

kettlegploadGreenplum 5.Xkettle7.X

kettleFineReportetlfrkettlekettle

kettlegpload

1gpload kettlegploadyamlFORMATtextNULL_AS\textNULL_AS\N

2gploadgploadpgsql:1kettlekettle

3gploadgreenplum4.Xgreenplum5.X

43greenplum4.Xgreenplum5.Xkettlegploadgp5.x

gpload http://gpdb.docs.pivotal.io/500/utility_guide/admin_utilities/gpload.html

5.0.0gpccjdbch<http://pan.baidu.com/s/1nvBIZdR> cecq

redhat 6.x redhat 7.x centos 6.xcentos 7.x

yaml

--- VERSION: 1.0.0.1 --schema

DATABASE: db_name --,\$PGDATABASEgpload-d

USER: db_username --gp_external_grant_privilegeson

HOST: master_hostname --mastergpload-h\$PGHOST

PORT: master_port --master5432,gpload-p\$PGPORT

GPLOAD: --INPUT:OUTPUT:

INPUT: ---

SOURCE: --sourcesource

LOCAL_HOSTNAME: --gploadip -hostname_or_ip

PORT: http_port --gpdfistport | PORT_RANGE: [start_port_range, end_port_range]

FILE: --gpzipbzip2*C - /path/to/input_file -

COLUMNS: --source - field_name: data_type -

FORMAT: text | csv --textcsv -

DELIMITER: 'delimiter_character' --| -

ESCAPE: 'escape_character' | 'OFF' --texttext\textoff -

NULL_AS: 'null_string' --text\Ncsv -

FORCE_NOT_NULL: true | false --csv""0 -

QUOTE: 'csv_quote_character' --csv" - HEADER: true | false -- -

ENCODING: database_encoding -- -

ERROR_LIMIT: integer --gpload -

ERROR_TABLE: schema.table_name --

OUTPUT: - TABLE: schema.table_name -- -

MODE: insert | update | merge --insertmerge -

MATCH_COLUMNS: --updatemerge -

target_column_name - UPDATE_COLUMNS: --updatemerge -

target_column_name - UPDATE_CONDITION: 'boolean_condition' --wheremergeinsert) -

MAPPING: --source

target_column_name: source_column_name | 'expression' PRELOAD: --load -

TRUNCATE: true | false --true, -

REUSE_TABLES: true | false --true

SQL: - BEFORE: "sql_command" --SQL - AFTER: "sql_command" --SQL