

NAME

db5.3_load – Load data from standard input

SYNOPSIS

db5.3_load [-nTV] [-c name=value] [-f file] [-h home] [-P password] [-t btree | hash | queue | recno] file

db5.3_load [-r lsn | fileid] [-h home] [-P password] file

DESCRIPTION

The db5.3_load utility reads from the standard input and loads it into the database file. The database file is created if it does not already exist.

The input to db5.3_load must be in the output format specified by the db5.0_dump utility, utilities, or as specified for the -T below.

OPTIONS

- c** Specify configuration options ignoring any value they may have based on the input. The command-line format is **name=value**. See the Supported Keywords section below for a list of keywords supported by the **-c** option.
- f** Read from the specified **input** file instead of from the standard input.
- h** Specify a home directory for the database environment.

If a home directory is specified, the database environment is opened using the DB_INIT_LOCK, DB_INIT_LOG, DB_INIT_MPOOL, DB_INIT_TXN, and DB_USE_ENVIRON flags to DB_ENV->open. (This means that db5.3_load can be used to load data into databases while they are in use by other processes.) If the DB_ENV->open call fails, or if no home directory is specified, the database is still updated, but the environment is ignored; for example, no locking is done.

- n** Do not overwrite existing keys in the database when loading into an already existing database. If a key/data pair cannot be loaded into the database for this reason, a warning message is displayed on the standard error output, and the key/data pair are skipped.
- P** Specify an environment password. Although Berkeley DB utilities overwrite password strings as soon as possible, be aware there may be a window of vulnerability on systems where unprivileged users can see command-line arguments or where utilities are not able to overwrite the memory containing the command-line arguments.
- r** Reset the database's file ID or log sequence numbers (LSNs).

All database pages in transactional environments contain references to the environment's log records. In order to copy a database into a different database environment, database page references to the old environment's log records must be reset, otherwise data corruption can occur when the database is modified in the new environment. The **-r lsn** option resets a database's log sequence numbers.

All databases contain an ID string used to identify the database in the database environment cache. If a database is copied, and used in the same environment as another file with the same ID string, corruption can occur. The **-r fileid** option resets a database's file ID to a new value.

In both cases, the physical file specified by the file argument is modified in-place.

- T** The **-T** option allows non-Berkeley DB applications to easily load text files into databases.

If the database to be created is of type Btree or Hash, or the keyword **keys** is specified as set, the input must be paired lines of text, where the first line of the pair is the key item, and the second line of the pair is its corresponding data item. If the database to be created is of type Queue or Recno and the keyword **keys** is not set, the input must be lines of text, where each line is a new data item for the database.

A simple escape mechanism, where newline and backslash (\) characters are special, is applied to the text input. Newline characters are interpreted as record separators. Backslash characters in the text will be interpreted in one of two ways: If the backslash character precedes another backslash character, the pair will be interpreted as a literal backslash. If the backslash character precedes any other character, the two characters following the backslash will be interpreted as a hexadecimal specification of a single character; for example, \0a is a newline character in the ASCII character set.

For this reason, any backslash or newline characters that naturally occur in the text input must be escaped to avoid misinterpretation by `db5.3_load`.

If the **-T** option is specified, the underlying access method type must be specified using the **-t** option.

- t** Specify the underlying access method. If no **-t** option is specified, the database will be loaded into a database of the same type as was dumped; for example, a Hash database will be created if a Hash database was dumped.

Btree and Hash databases may be converted from one to the other. Queue and Recno databases may be converted from one to the other. If the **-k** option was specified on the call to `db5.3_dump` then Queue and Recno databases may be converted to Btree or Hash, with the key being the integer record number.

- V** Write the library version number to the standard output, and exit.

The `db5.3_load` utility may be used with a Berkeley DB environment (as described for the **-h** option, the environment variable **DB_HOME**, or because the utility was run in a directory containing a Berkeley DB environment). In order to avoid environment corruption when using a Berkeley DB environment, `db5.3_load` should always be given the chance to detach from the environment and exit gracefully. To cause `db5.3_load` to release all environment resources and exit cleanly, send it an interrupt signal (SIGINT).

The `db5.3_load` utility exits 0 on success, 1 if one or more key/data pairs were not loaded into the database because the key already existed, and >1 if an error occurs.

EXAMPLES

The `db5.3_load` utility can be used to load text files into databases. For example, the following command loads the standard UNIX `/etc/passwd` file into a database, with the login name as the key item and the entire password entry as the data item:

```
awk -F: '{print $1; print $0}' < /etc/passwd |
sed 's/\\|\\|\\|g' | db5.3_load -T -t hash passwd.db
```

Note that backslash characters naturally occurring in the text are escaped to avoid interpretation as escape characters by `db5.3_load`.

ENVIRONMENT

DB_HOME

If the **-h** option is not specified and the environment variable **DB_HOME** is set, it is used as the path of the database home, as described in `DB_ENV->open`.

SUPPORTED KEYWORDS

The following keywords are supported for the **-c** command-line option to the `db5.3_load` utility. See `DB->open` for further discussion of these keywords and what values should be specified.

The parenthetical listing specifies how the value part of the **name=value** pair is interpreted. Items listed as (boolean) expect value to be **1** (set) or **0** (unset). Items listed as (number) convert value to a number. Items listed as (string) use the string value without modification.

bt_minkey (number)

The minimum number of keys per page.

`chksum` (boolean)
Enable page checksums.

`database` (string)
The database to load.

`db_lorder` (number)
The byte order for integers in the stored database metadata.

`db_pagesize` (number)
The size of database pages, in bytes.

`duplicates` (boolean)
The value of the DB_DUP flag.

`dupsort` (boolean)
The value of the DB_DUPSORT flag.

`extentsize` (number)
The size of database extents, in pages, for Queue databases configured to use extents.

`h_ffactor` (number)
The density within the Hash database.

`h_nelem` (number)
The size of the Hash database.

`keys` (boolean)
Specify whether keys are present for Queue or Recno databases.

`re_len` (number)
Specify fixed-length records of the specified length.

`re_pad` (string)
Specify the fixed-length record pad character.

`recnum` (boolean)
The value of the DB_RECNUM flag.

`renumber` (boolean)
The value of the DB_RENUMBER flag.

`subdatabase` (string)
The subdatabase to load.

AUTHORS

Sleepycat Software, Inc. This manual page was created based on the HTML documentation for `db_load` from Sleepycat, by Thijs Kinkhorst <thijs@kinkhorst.com>, for the Debian system (but may be used by others).