

NAME

lbzip2 - parallel bzip2 filter

SYNOPSIS

lbzip2 [-d] [-n *WORKER-THREADS*] [-v] [-t]

lbzip2 -h

DESCRIPTION

Compress or decompress standard input to standard output, by calling Julian Seward's **libbz2** from multiple threads. **Lbzip2** utilizes multiple threads and an input-bound splitter even when decompressing bz2 files created by standard **bzip2** (but see **BUGS** below).

OPTIONS

-d Decompress. Default mode is compression.

-n *WORKER-THREADS*

Set the number of (de)compressor threads to *WORKER-THREADS*. If this option is not specified, the environment variable **LBZIP2_WORKER_THREADS** is consulted. If also **LBZIP2_WORKER_THREADS** is unset or empty, then **lbzip2** either queries the system for the number of online processors (if both the compilation environment and the execution environment support that) or exits with an error.

-v Print condition variable statistics to standard error in the end. If this option is not specified, the environment variable **LBZIP2_PRINT_STATS** is consulted: statistics will be printed if and only if **LBZIP2_PRINT_STATS** is set to a non-empty value.

-t Write memory allocation trace to standard error. If this option is not specified, the environment variable **LBZIP2_TRACE_ALLOC** is consulted: allocation trace will be printed if and only if **LBZIP2_TRACE_ALLOC** is set to a non-empty value. The trace can be checked with the **malloc_trace.pl** Perl script (part of the source distribution).

-h Print command line help to standard error and exit.

ENVIRONMENT

LBZIP2_WORKER_THREADS

See **-n**.

LBZIP2_PRINT_STATS

See **-v**.

LBZIP2_TRACE_ALLOC

See **-t**.

EXIT STATUS

0 if **lbzip2** finishes successfully. This presumes that whenever it tries, **lbzip2** never fails to write to standard error.

EXIT_FAILURE

if **lbzip2** encounters any fatal error. (**EXIT_FAILURE** corresponds to a platform-dependent integer constant in [1..255], usually **1**.)

SIGABRT

(via **assert()**) if a runtime assertion fails (ie. **lbzip2** detects a bug in itself). Hopefully whoever compiled your binary wasn't bold enough to **#define NDEBUG**.

BUGS

The compressed output is a sequence of single-block bzip2 streams, instead of being a single multi-block bzip2 stream. Thus no output-global combined CRC can be written, and the resulting bz2 file will return multiple **BZ_STREAM_ENDS** when decompressed with **libbz2**.

The multiple-workers decompressor (**lbzip2 -d -n WORKER-THREADS**, where *WORKER-THREADS* is greater than *1*, either by default or by manual selection) doesn't check stream-level (combined) CRCs, only block-level CRCs. It can also fail to decompress some rare valid bz2 files. (See the Bugs section of the

README in the source distribution for more on this.) Since the author cannot rule out that such a failure may get reported as corrupt input instead of the dedicated error message, **lbzip2** doesn't reserve a separate exit status for corrupt input -- it could be misleading.

Neither the input block size nor the bzip2 block size can be specified for compression, they are both fixed at the maximum bzip2 block size of 899,981 bytes.

AUTHOR

Written by Laszlo Ersek <lacos@elte.hu>, <http://phptest11.atw.hu/>.

COPYRIGHT

Copyright (C) 2008 Laszlo Ersek.

This is free software. You may redistribute copies of it under the terms of the GNU General Public License, version 2:

<http://www.gnu.org/licenses/gpl-2.0.html>

or -- at your opinion -- any later version, eg.

<http://www.gnu.org/licenses/gpl.html>

THANKS

Adam Maulis at ELTE IIG; Julian Seward; Paul Sladen; Michael Thomas from Caltech HEP; Department of Electrical and Information Engineering at the University of Oulu.

SEE ALSO

bzip2(1)

<http://www.bzip.org/>

pbzip2(1)

<http://compression.ca/pbzip2/>

bzip2smp(1)

<http://bzip2smp.sourceforge.net/>

smpbzip2(1)

<http://home.student.utwente.nl/n.werensteijn/smpbzip2/>

dbzip2(1)

<http://www.mediawiki.org/wiki/Dbzip2>