УДК 004.051

Исследование эффективности внедрения библиотеки потокового сжатия данных ZStandart в систему ведения журналов работы ПО

К. А. Щербатов

Московский физико-технический институт (государственный университет)

kirill.shcherbatov@phystech.edu

Протоколирование (журнал, лог, англ. Log) - хронологическая запись с настраиваемой степенью детализации сведений о происходящих в системе событиях. Ведение лога является наиболее доступным, масштабируемым и достаточно эффективным способом отслеживать причины возникновения ошибок и «узкие места» производительности программ.

При подробном журналировании или продолжительном времени работы ПО в протоколе становится много записей, что приводит к проблемам чрезмерного потребления дискового пространства. Одним из способов решения этой проблемы является запись данных в сжатом представлении.

К алгоритму сжатия предъявляют взаимоисключающие требования высокой производительности и существенного уменьшения объема хранимых данных.

Работа посвящена исследованию эффективности замены алгоритмов gzip на zstd и экспериментальный поиск оптимального уровня сжатия данных, обеспечивающего наилучшее соотношение измеряемых величин. Целью исследования было наблюдаемое повышение эффективности системы логирования.

Литература

1. *P. Deutsch, Aladdin Enterprises*: GZIP file format specification version 4.3 — 1996
2. *Yann Collet*: Zstandard Compression Format — 2016 https://github.com/facebook/zstd/blob/master/zstd_compression_format.md