

Enhanced Adaptive Compression in Lustre

Status Report

Anna Fuchs, Michael Kuhn

anna.fuchs@informatik.uni-hamburg.de

Research Group Scientific Computing
Department of Informatics, Universität Hamburg

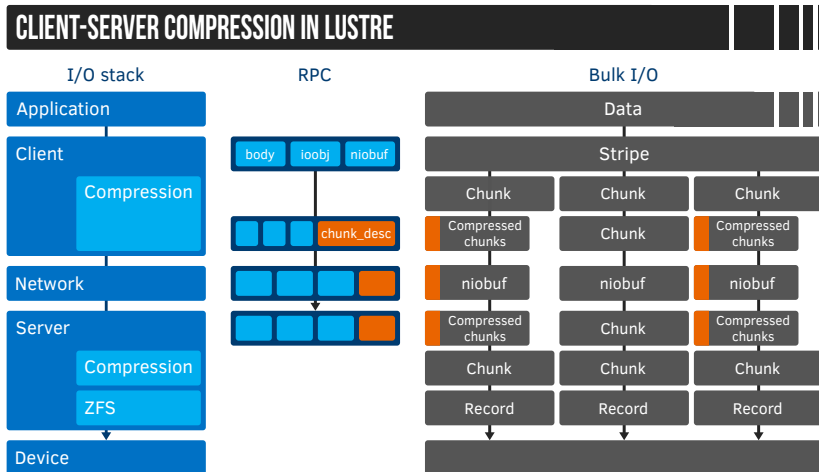
07 December 2017



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG



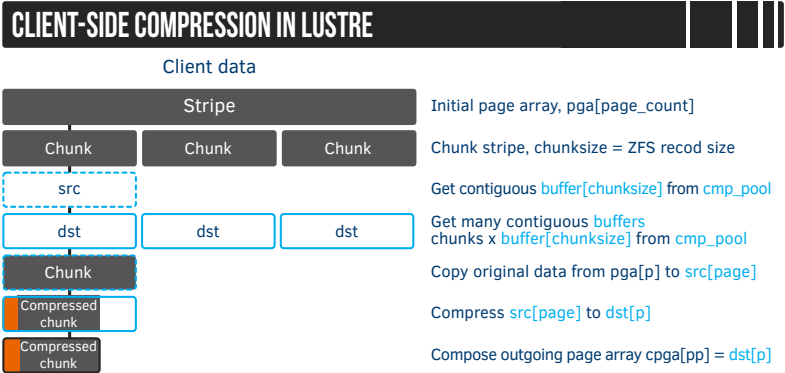
Client-server



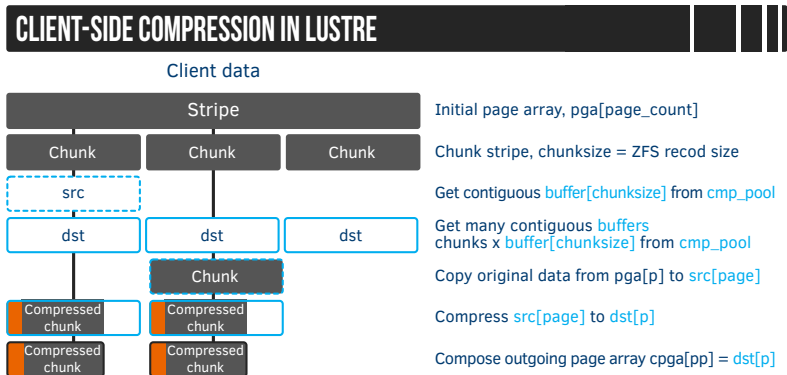
Client internals



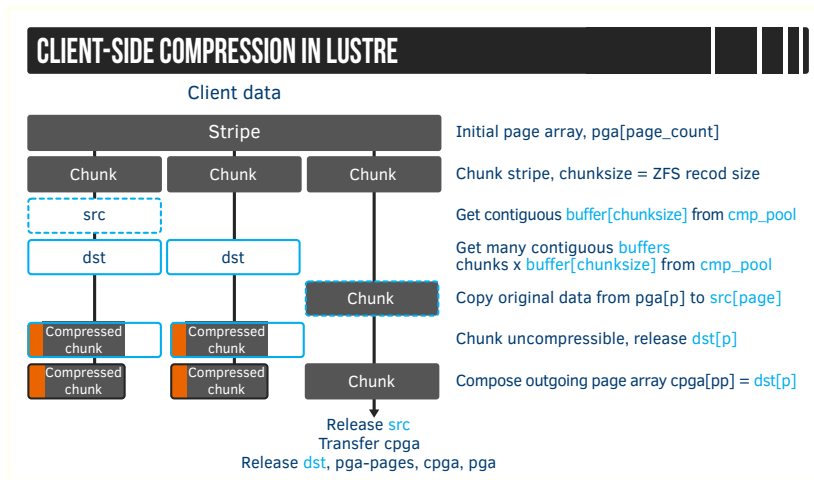
Client internals



Client internals

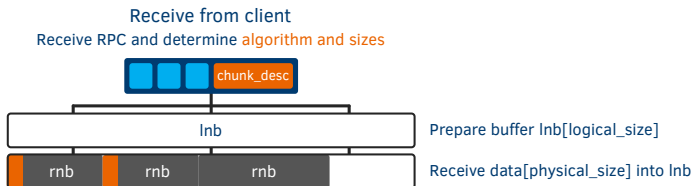


Client internals



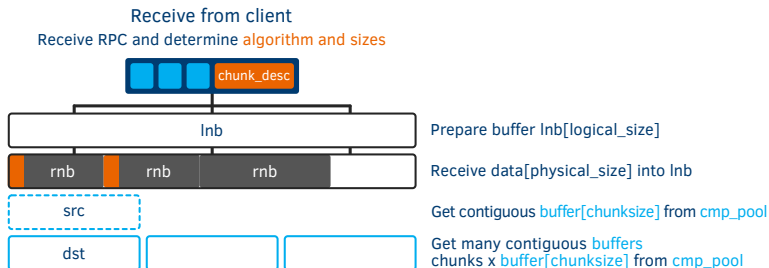
Server internals (client-server compression)

SERVER-SIDE DECOMPRESSION IN LUSTRE



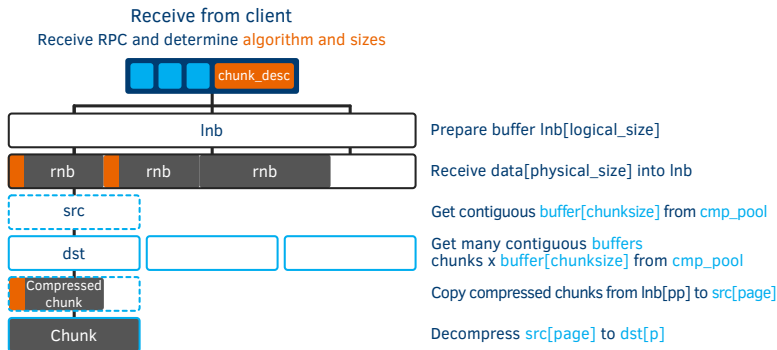
Server internals (client-server compression)

SERVER-SIDE DECOMPRESSION IN LUSTRE



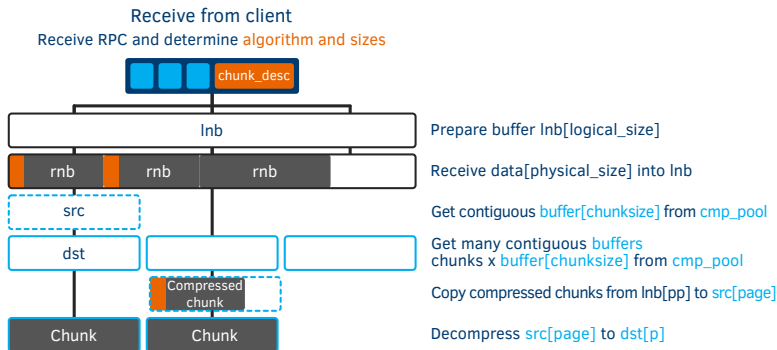
Server internals (client-server compression)

SERVER-SIDE DECOMPRESSION IN LUSTRE



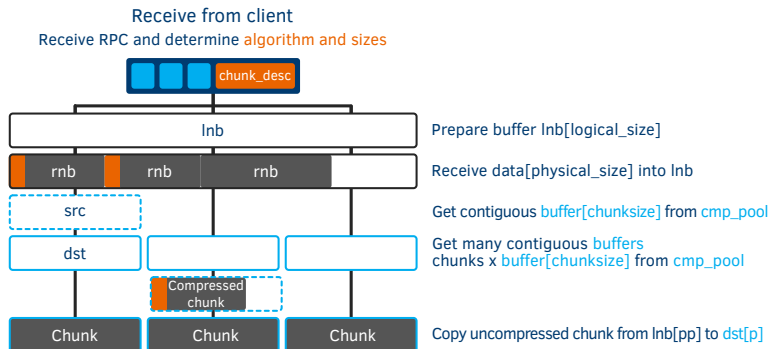
Server internals (client-server compression)

SERVER-SIDE DECOMPRESSION IN LUSTRE



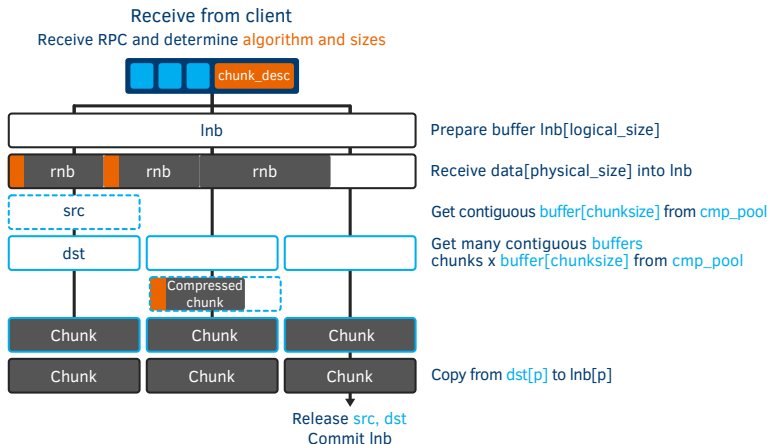
Server internals (client-server compression)

SERVER-SIDE DECOMPRESSION IN LUSTRE



Server internals (client-server compression)

SERVER-SIDE DECOMPRESSION IN LUSTRE



TODOs - I

- Memory
 - Client: correct release of pages
 - Server: correct summary of used memory
- Compression pool
 - Atomic allocation of many blocks
 - Clean up forgotten blocks when freeing
 - Threshold for less memory than required
- Configuration
 - Compatibility version checks
 - Dynamic enabling/disabling of compression
 - Set correct page pool size
 - Get/set ZFS record size for chunk size
 - Compression threshold
 - Algorithms and policies

TODOs - II

- Algorithms
 - Add more; also page-based
 - Define policy of mixing algos per file/RPC/stripe
 - Decision step
- Style, refactoring, etc.
- Add read path
- Testing
 - Build tests
 - Correctness tests
 - Memory tests
- RMW – fix offsets (do not assume page aligned data)
- Error handling – fallbacks, recovery, crash

Store compressed data

- Requires ZFS changes
 - Test our changes
 - Submit to ZFS
- Consider ZFS' and Lustre's community reviews
- Client mechanisms remain the same for client-server or end-to-end compression
- Continue implementing support for OSS, MDS and ZFS