Whamcloud

Lustre Client Encryption

09/2019

sbuisson@whamcloud.com

DDN® STORAGE

Lustre Client Encryption

Whamcloud

- Optimization for encryption context
- Parallelization of encryption/decryption
- Handling size of encrypted files
- Impact of/on other Lustre features
- Security related code reviews

Lustre Client Encryption – encryption context optimization

- Per-file encryption context is stored in an xattr
  - Hopefully not changed after file creation

- Getting/setting xattrs impacts performance by generating additional requests

- Insecure window with file created but no encryption context set

Lustre Client Encryption – encryption context optimization

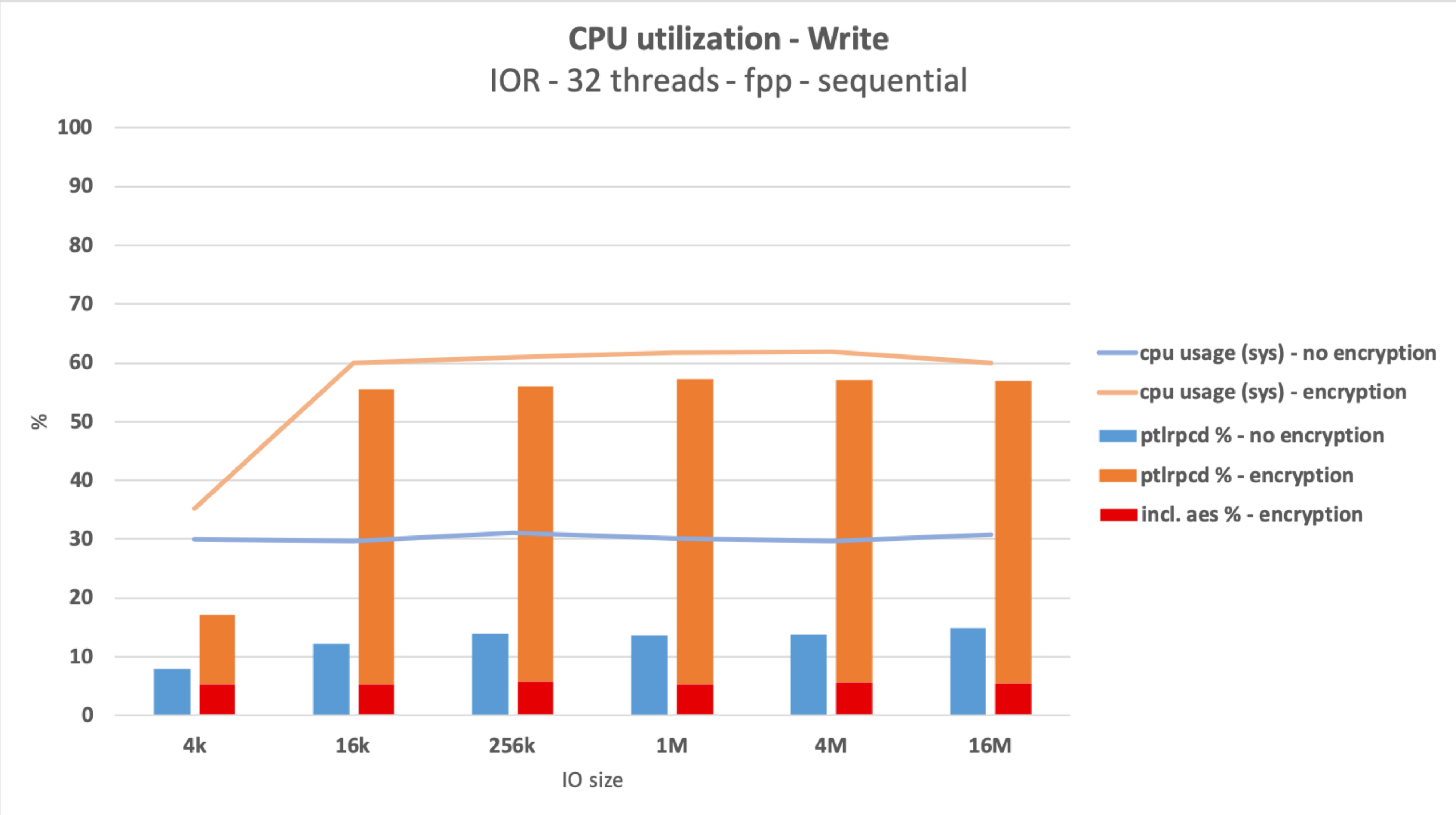
- Lustre must be able to
  - Set encryption context directly with create request
  - Fetch encryption context directly with open/lookup request
- Similar to SELinux optimizations
  - LU-5560 security: send file security context for creates
    - https://review.whamcloud.com/19971
  - LU-9193 security: return security context for metadata ops
    - https://review.whamcloud.com/26831
- Lustre has xattr cache
  - But only filled at getxattr, not setxattr
  - Need additional primitives to explicitly set xattr in cache (file create case)

Lustre Client Encryption – encryption/decryption performance

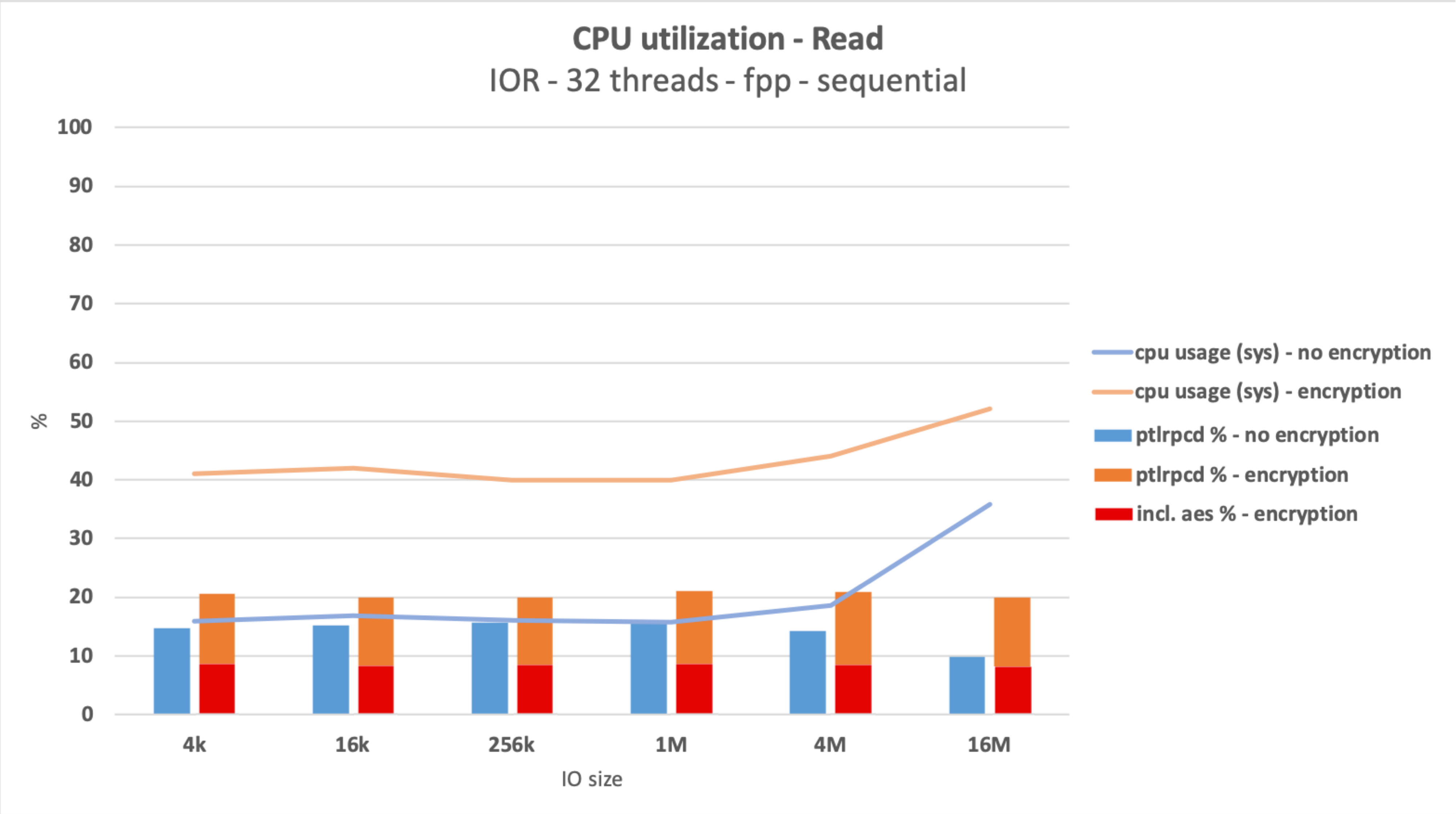
- Looking at time spent in aes routines, it seems we are not aggressive enough on encryption/decryption

Lustre Client Encryption – early performance evaluation

Whamcloud



Lustre Client Encryption – early performance evaluation



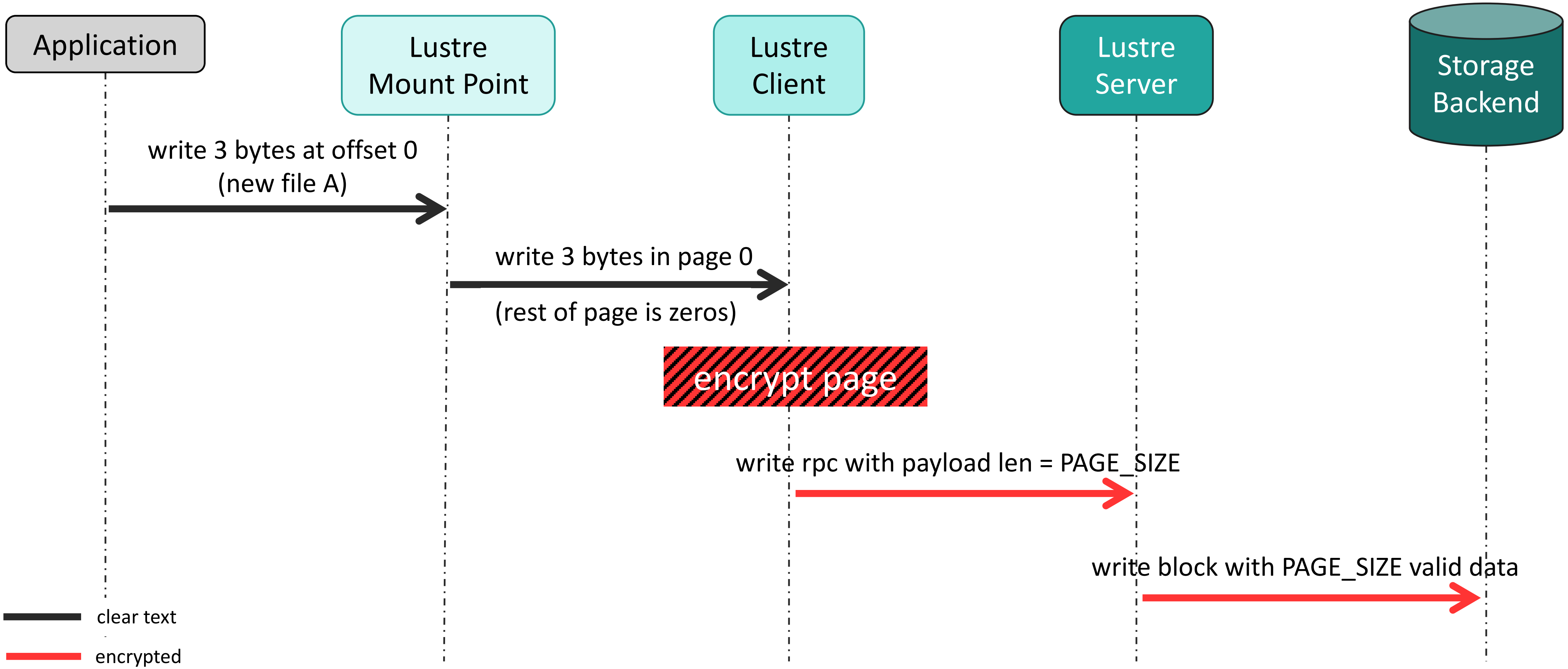
Lustre Client Encryption – encryption/decryption performance

- Pages are encrypted/decrypted sequentially
  - in osc_brw_prep_request(), for each page that will be added to the request, call fscrypt_encrypt_page()
    - allocates a bounce page
  - in osc_brw_fini_request(), for each page in request reply, call fscrypt_decrypt_page()
    - happens in place
- Encryption/decryption parallelization comes from parallel request processing
  - is that enough?
  - create new Lustre threads for encryption/decryption?

Lustre Client Encryption – file size handling

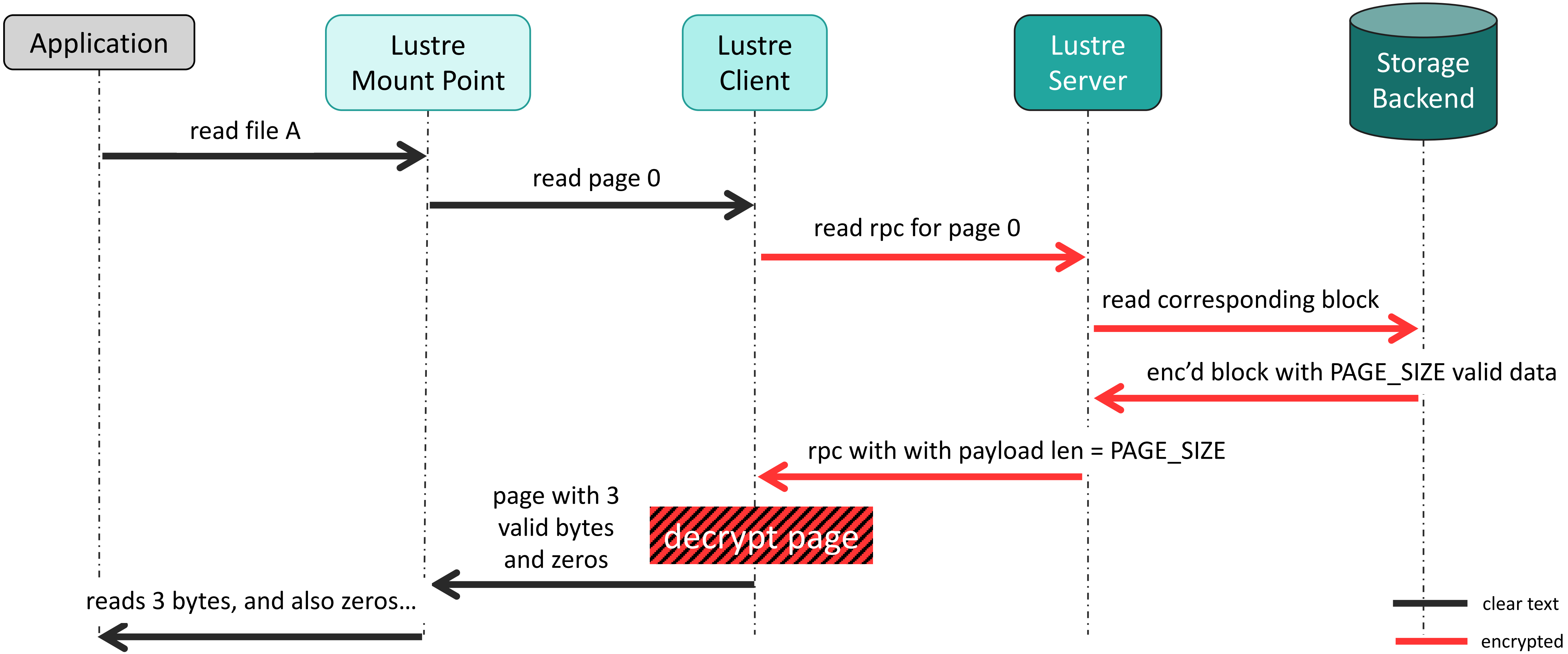
- Encryption chunk size is the page size
- Ciphertext page is always full of data… even if clear text only contains a few bytes
- But OSS assumes object size based on length of data received
  - Must carry on clear text length from client to server

Lustre Client Encryption – write case



Lustre Client Encryption – read case

Whamcloud



Lustre Client Encryption – file size handling

- Proposed solution on write path:
  - Client sends a value representing the length to be subtracted in order to get right file size (size diff)
    - Happens only for last page in write request
  - Server adjusts isize with this size diff
    - But still makes sure complete page is committed to disk
- Then on read path:
  - OSD behavior changed to always send complete pages
    - Data in page beyond isize will be finally discarded by client anyway
    - Should not be harmful to send at max PAGE_SIZE-1 more bytes for reads at end of file.

Lustre Client Encryption – Impact of/on other features

- Distributed Namespace (DNE)
  - make sure inodes on secondary MDTs have encrypted names
  - make sure encryption policy is correctly inherited on secondary MDTs
- File Level Redundancy (FLR)
- File migration
- Request replay
  - make sure file content is not encrypted twice
  - make sure replicate/migrated file has encrypted context
- Data-on-MDT (DoM)
  - make sure file content is encrypted when data sent to MDS instead of OSS

Security related code reviews

- Security features have impact on various Lustre components
  - maintainers for these components are affected…
  - but reviews also need a security sensitivity!
- Looking for people willing to participate in security oriented patch reviews
  - knowledge of
    - authentication
    - encryption
    - SELinux
  - implementation experience with other software, best practice, …
- Please add your name to the MAINTAINERS file

Whamcloud

Thank you!

sbuisson@whamcloud.com

DDN®
STORAGE

Lustre Client Encryption – remaining development

- Encryption of file, symlink and directory names
  - Measure metadata performance impact
- Ability to set encryption policies on directories
  - Support new IOCTLs from fscrypt userspace tool