Whamcloud

Lustre Client Encryption

Lustre User Group 2022

sbuisson@whamcloud.com

Lustre Client Encryption

Whamcloud

- What is encryption for Lustre and features available with new Lustre 2.15:
  - File name encryption
- Current encryption limitations
- Upcoming encryption improvements
  - Performance optimizations
  - Compatibility with native fscrypt in newer kernels

What is Lustre Client Encryption?

Whamcloud

- Kernel side
  - in-kernel fscrypt (5.4)
  - embedded *llcrypt* (CentOS/RHEL 8.1+, Ubuntu 18.04+, SLES 15 SP2+)
- User-space side
  - fscrypt userspace tool: works out of the box, thanks to fscrypt API support
- Landed in 2.15: filename encryption – LU-13717
  - Convert between plain and cipher text names
  - Access with and without the key

Encrypted file name length

Whamcloud

- Ciphertext names are binary data
  - Client: encode binary names and send to server
    - Use custom encoding, to limit overhead to strictly necessary
- Server with ldiskfs backend
  - Support binary names, so decode in osd-ldiskfs and pass along

  ⇒ *For Lustre encrypted files: name length up to 255 chars (NAME_MAX)*
- Server with ZFS backend
  - Binary dir entries would require special immutable ZAP flag on directory, so keep encoded binary names in osd-zfs, and pass to ZFS

  ⇒ *For Lustre encrypted files: name length up to 224 chars almost guaranteed*

Encrypted file names – e2fsprogs support

- Client encryption compatible with ext4
  - Encryption context stored in **`encryption.c`** xattr
  - LDISKFS_ENCRYPT_FL flag for on-disk inodes (ldiskfs)
- Benefits
  - Compatible with lfsck/fsck/e2fsck/debugfs

```
[root@lnode-vm3 ~]# debugfs -R "ls ROOT/vault" /var/loop/mdt1
debugfs 1.46.2.wc3 (18-Jun-2021)
 259  (12) .    31512  (28) ..    266  (60) <encrypted (32)>
 267  (3996) <encrypted (32)>
[root@lnode-vm3 ~]# debugfs -R "stat <266>" /var/loop/mdt1
debugfs 1.46.2.wc3 (18-Jun-2021)
Inode: 266   Type: regular    Mode:  0644   Flags: 0x800
Generation: 3099768797    Version: 0x00000001:0000001f
User:     0   Group:     0   Project:     0   Size: 0
File ACL: 0
Links: 1   Blockcount: 0
Fragment:  Address: 0    Number: 0    Size: 0
 ctime: 0x6256d783:00000000 -- Wed Apr 13 16:00:35 2022
 atime: 0x62553431:00000000 -- Tue Apr 12 10:11:29 2022
 mtime: 0x6256d783:00000000 -- Wed Apr 13 16:00:35 2022
crtime: 0x62553431:832cb4a4 -- Tue Apr 12 10:11:29 2022
Size of extra inode fields: 32
Extended attributes:
  lma: fid=[0x200000402:0x5:0x0] compat=0 incompat=20
  trusted.lov (56) = d0 0b d1 0b 01 00 00 00 05 00 00 00 00 00
  c (40) = 02 01 04 03 00 00 00 00 ae 55 1b db 60 5d a4 ad 17 e
  linkea: idx=0 parent=[0x200000402:0x1:0x0] name='\xa1\x14\xa8
  trusted.som (24) = 04 00 00 00 00 00 00 00 00 34 00 00 00 00
```

Client encryption – upgrade from 2.14 and filename encryption

Whamcloud

- In 2.15, name encryption for new files and directories
  - if under a parent encrypted directory created with 2.15
- New files and directories under a parent encrypted directory created with 2.14
  - will **not** have their names encrypted
- Files created with 2.14 do not have their names encrypted
  - they will remain so after upgrade to 2.15
- How to get name encryption for files created with 2.14
  - upgrade to 2.15
  - create new encrypted directory
  - unlock old directory
  - copy files from old directory to new
  - remove old directory

Client encryption – migrate/mirror

Whamcloud

- Migrate encrypted directory across MDTs, encrypted file across OSTs
- Mirror full support, except:
  - ‘lfs mirror split’ without ‘-d’ not allowed on encrypted files

- Works with the encryption key
- Works without the encryption key
  - Needs to get access to raw encrypted data
  - **`O_FILE_ENC | O_DIRECT`** open flags for ‘lfs migrate/mirror’
    - reserved to applications that know what they are doing

Client encryption – subdir mount of encrypted directory

Whamcloud

- When mounting an encrypted sub-directory
  - Need to fetch encryption context for root inode
  - Need to present .fscrypt directory for userspace command-line
    - Virtually present .fscrypt directory at root of mount point
    - Internally, .fscrypt always stored at root of Lustre

⇒ Clients access encrypted files only

Lustre Client Encryption – current limitations

Whamcloud

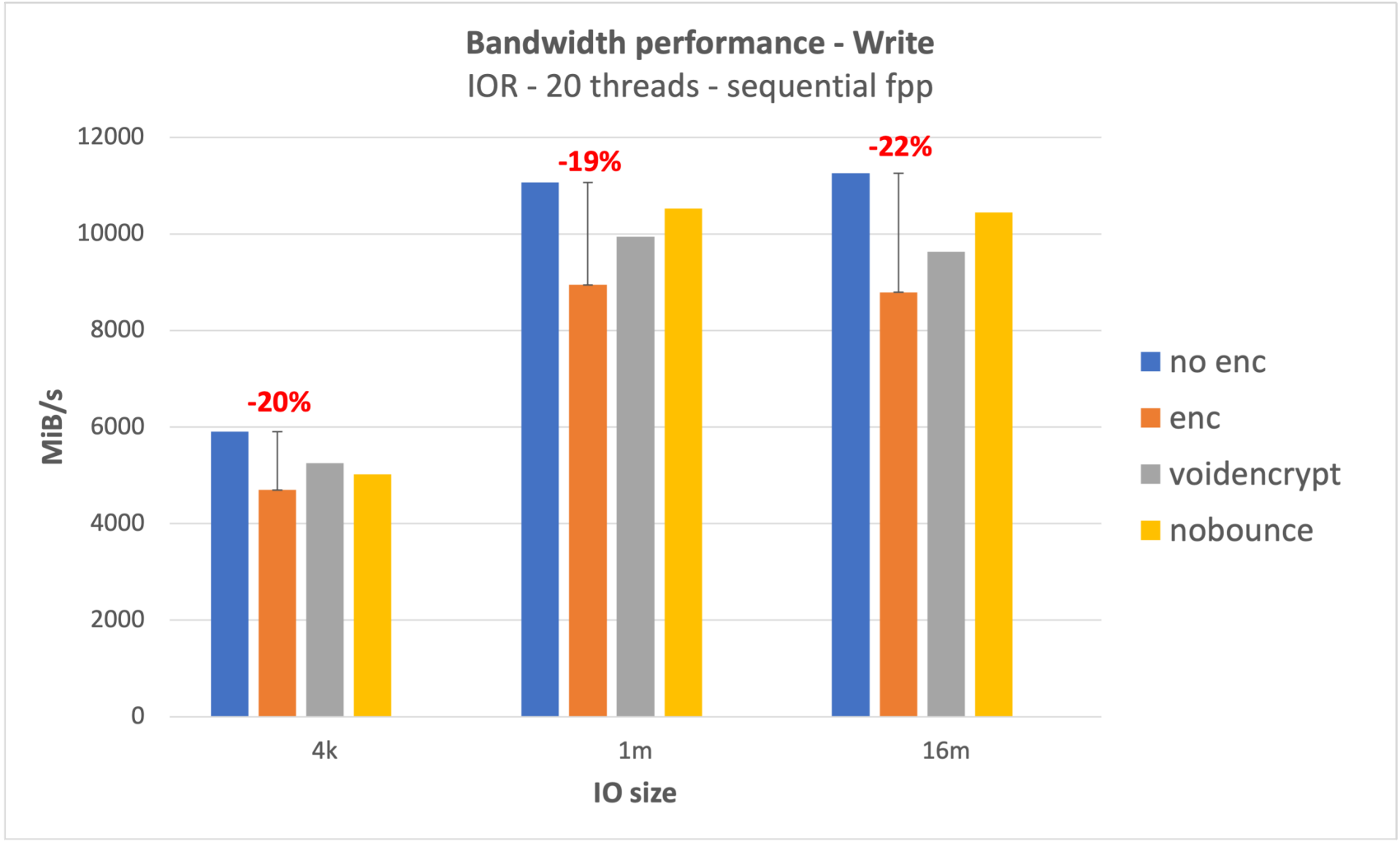
- What is currently incompatible with client encryption?
  - fid2path: full path built on server side
  - Lustre HSM, backup
    - Only possible to archive with the encryption key…
    - Without key, would need agents to use **`O_FILE_ENC | O_DIRECT`** flags
    - Also, would require special mechanism to retrieve clear text file size
- What is not client encryption?
  - Client encryption is not an ACL mechanism.
  - Client encryption does not protect file metadata
    - Size, permissions, timestamps, xattrs remain in clear text

Lustre Client Encryption – performance

Whamcloud

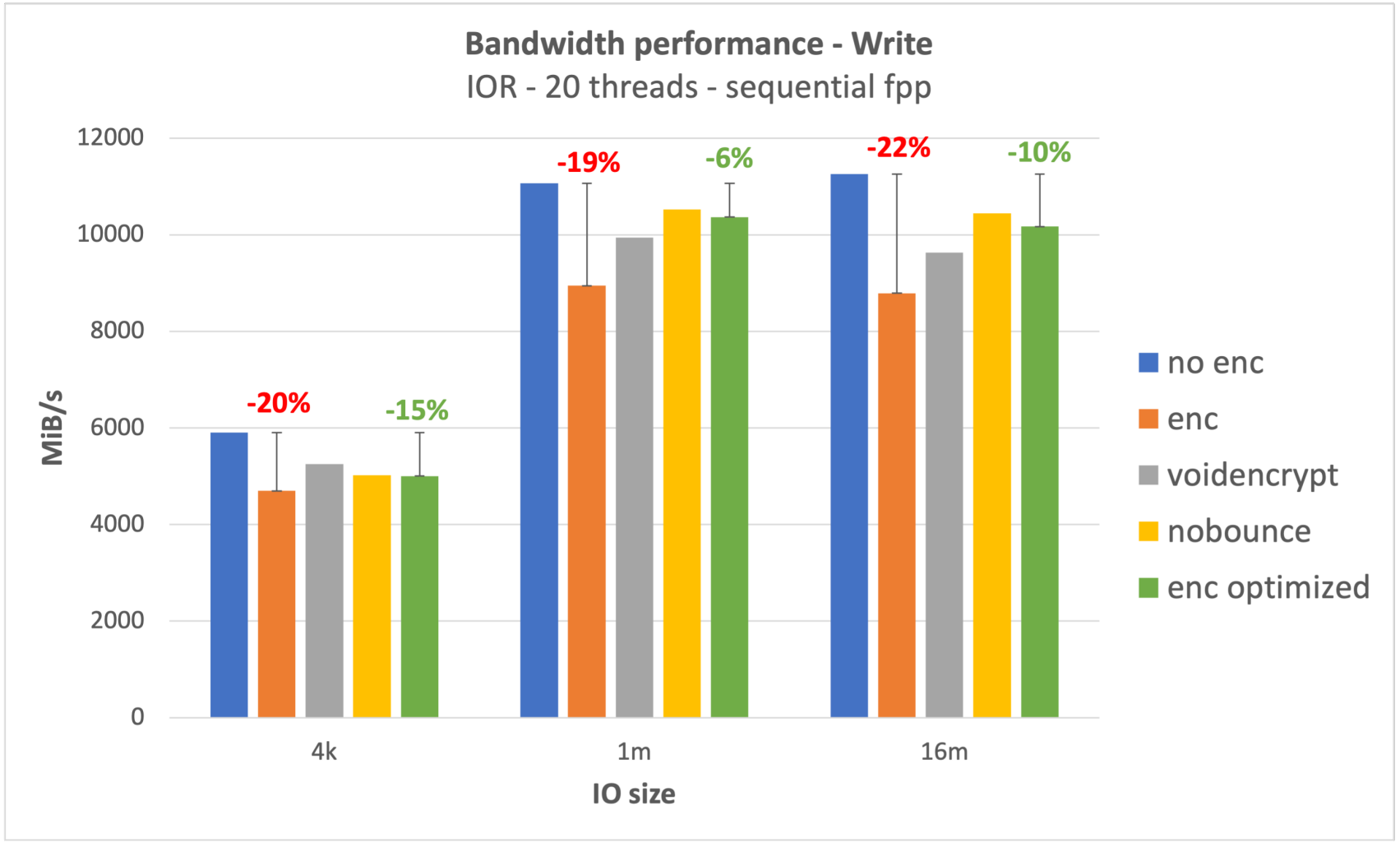
- Initial benchmarks
  - 30-35% drop in sequential write, 20-25% drop in sequential read
- Testbed
  - Client
    - Cascade Lake 20 cores, 6230 CPU @ 2.10GHz
    - 192 GB RAM
    - Infiniband adapter, EDR network
    - Ubuntu 20.04 kernel 5.4.0-107-generic
    - Lustre 2.15.0-RC3
  - Storage
    - ES400NVX
    - 20 x NVMe, 2 DCR 10 disks
    - 8 OSTs, 4 MDTs
    - CentOS 7.9 kernel 3.10.0-1160
    - Lustre 2.15.0-RC3
- Methodology
  - fscrypt with AES-256-XTS for file content, AES-256-CTS for file names

Lustre Client Encryption – performance



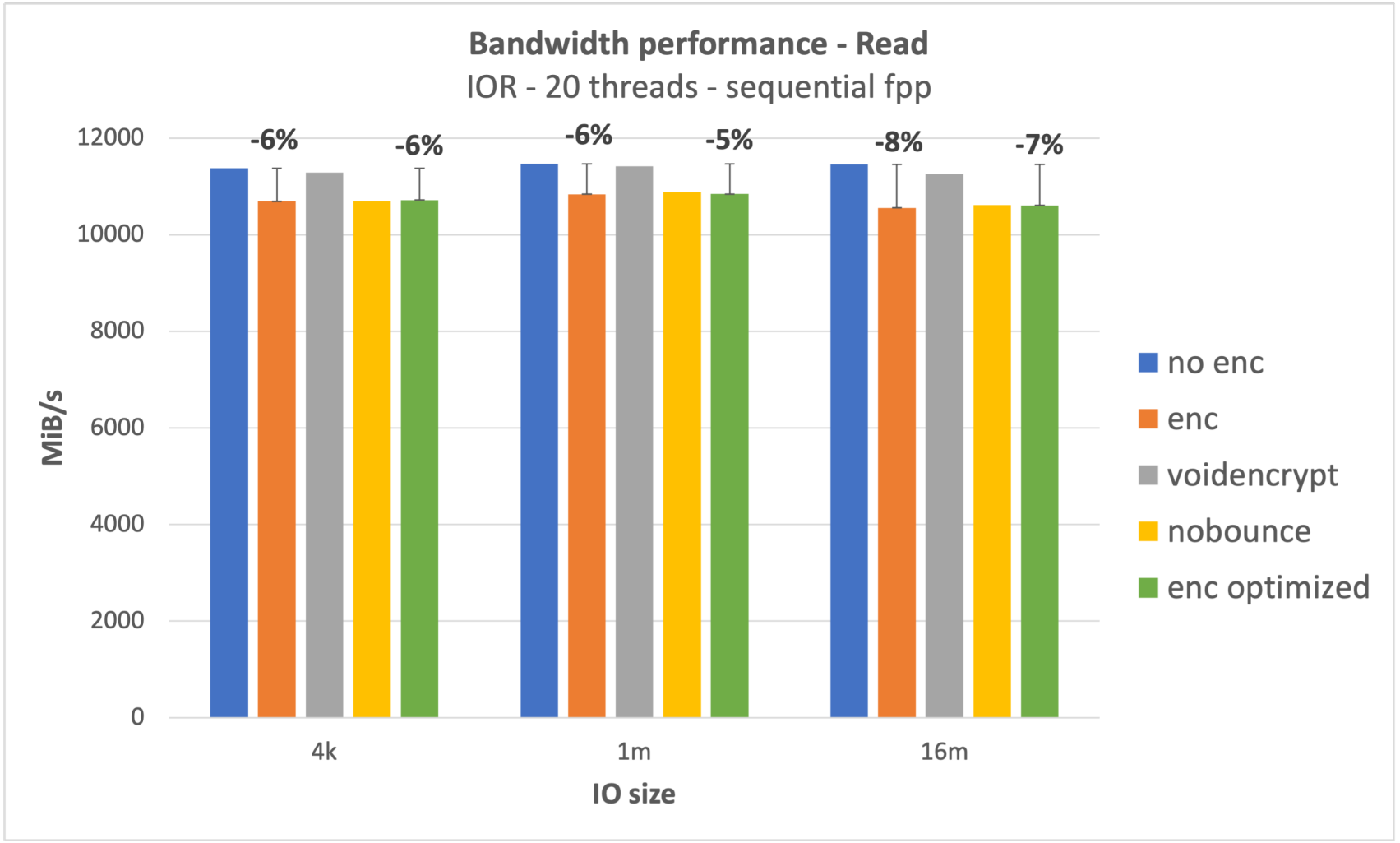
Performance drop for normal encryption code: 20%

Lustre Client Encryption – performance



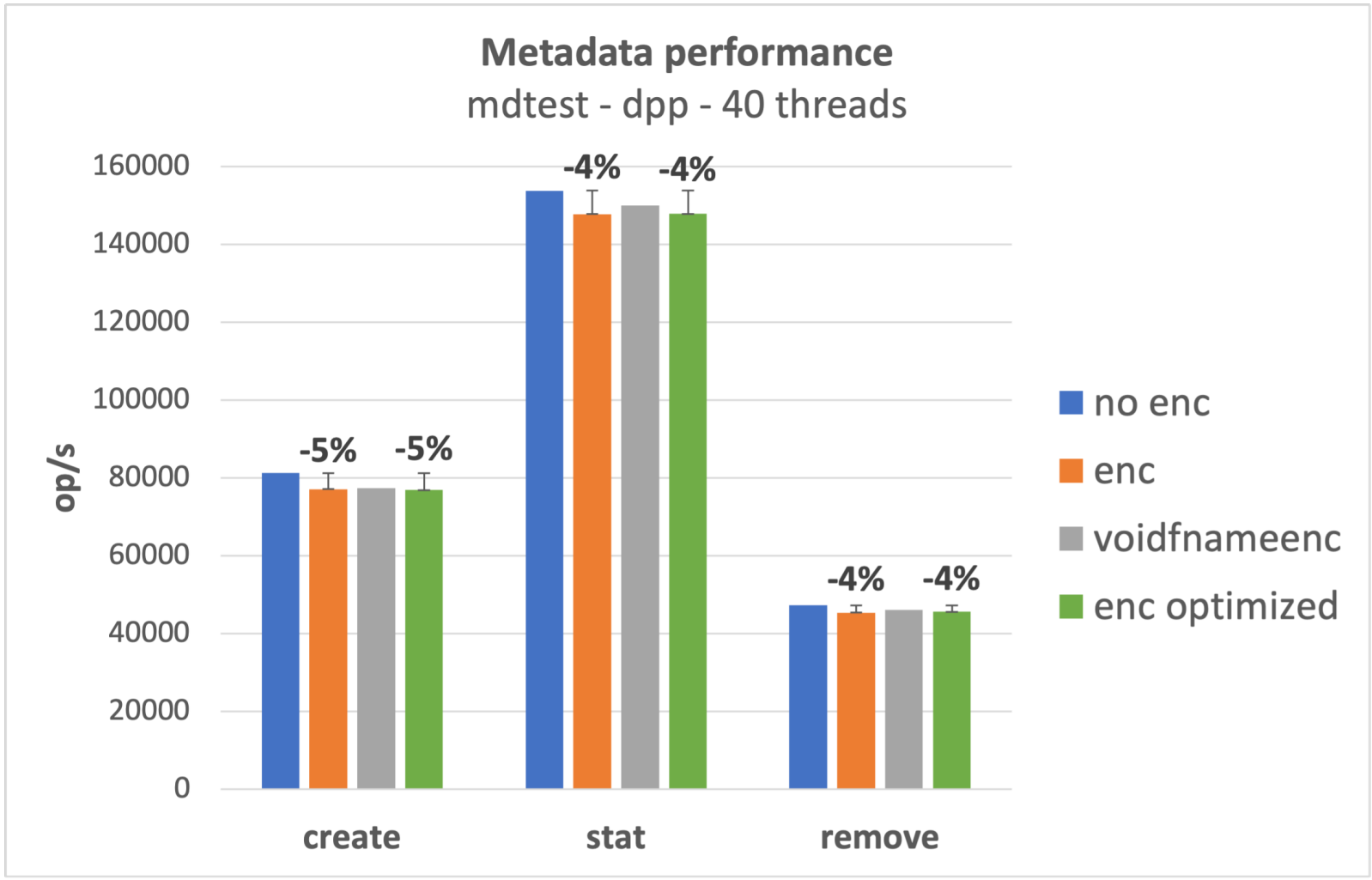
Performance drop for optimized encryption: 5-10% (>= 1MB IO)

Lustre Client Encryption – performance



Performance drop for all encryption versions: < 10%

Lustre Client Encryption – performance



Performance drop for all encryption versions: 5%

Lustre Client Encryption – performance

Whamcloud

- Code for bounce page optimization
  - LU-15003 sec: use enc pool for bounce pages
    - patch #47149
  - only available with embedded llcrypt
  - need to push kernel patch to improve fscrypt API
    - thanks to James Simmons for his help on this

Lustre Client Encryption – compatibility with newer kernels

Whamcloud

- Currently, Lustre compatible with in-kernel fscrypt from Linux 5.4
  - Ubuntu 20.04 initial kernel
- But Ubuntu kernels change fast
  - Now base kernel for 20.04 is 5.8
    - HWE kernel is 5.13
  - And Ubuntu 22.04 ships with 5.15
- Upcoming patches to support these kernels
  - LU-13783 sec: support of native Ubuntu 20.04 HWE 5.8 kernel

Lustre Client Encryption – wrap-up

Whamcloud

- Lustre 2.15 fully compatible with fscrypt
  - encryption of file content
  - encryption of file name
- But remember
  - Client encryption is not an ACL mechanism, and does not protect metadata
  - Current limitations: archive, backup
- With upcoming performance optimizations

	Performance penalty
Bandwidth – write	5%-10% for large IOs, 15% for small IOs
Bandwidth – read	less than 10%
Metadata – create, stat, remove	5%

Whamcloud

Thank you!

sbuisson@whamcloud.com