



# Lustre\* Releases

Peter Jones

April 16, 2013

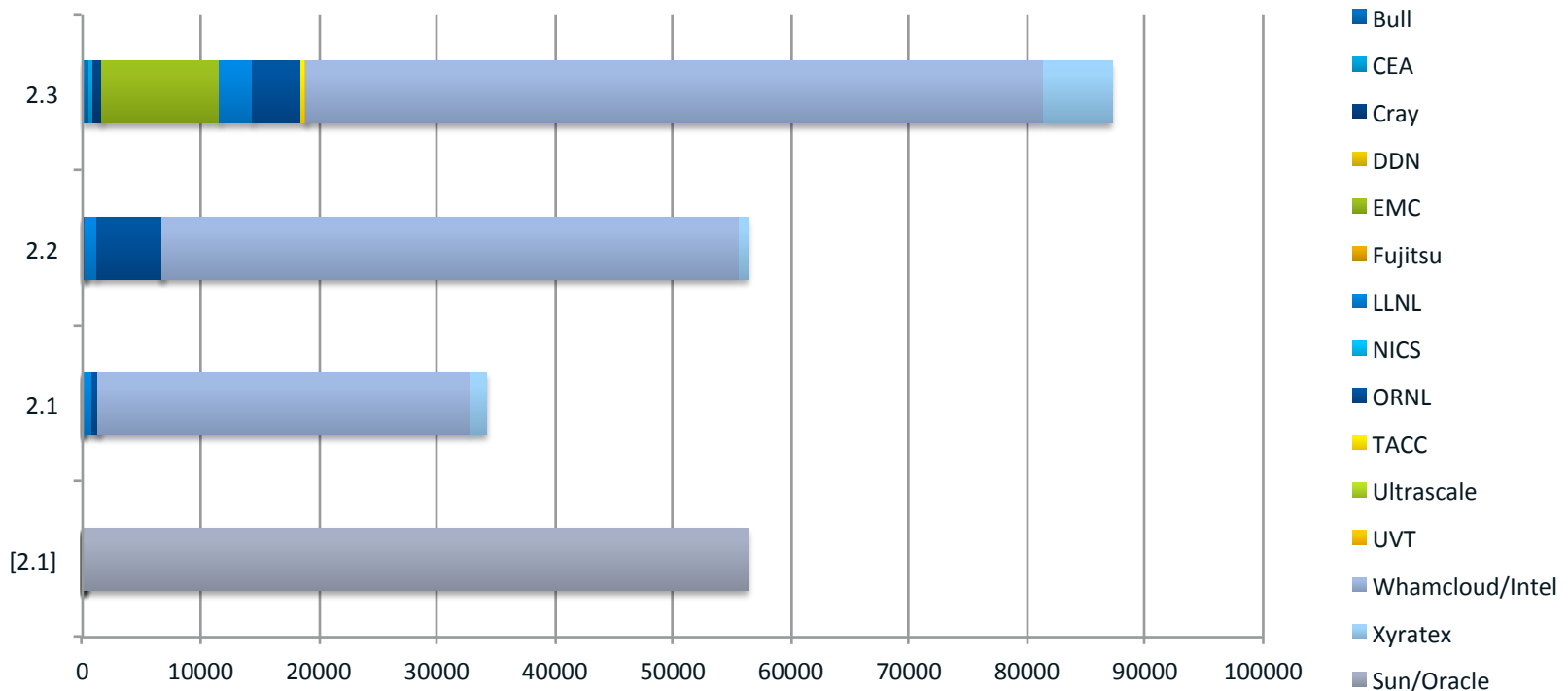
\* Other names and brands may be claimed as the property of others.

# Lustre Releases

- Evolving Lustre Community
- Active codelines
- Lustre 2.4.x
- Lustre 2.5
- Release quality

# Evolving Lustre Community

At CFS/SUN/Oracle third party contributions were rare



Now an increasing number of organizations contribute

# Evolving Lustre Community - CDWG

OpenSFS CDWG serves as a forum to discuss Lustre development matters

- Agreeing priorities for test matrix and roadmap
- Sharing test results
- Warning others of known issues
- Reminders of release deadlines
- CDWG wiki repository for development in progress

# Active Codelines – Lustre 1.8.x

Majority of Lustre production sites still running this codeline

- New deployments still occurring using 1.8.x releases
- OpenSFS survey showed 49% running recent Whamcloud/Intel 1.8.x (compared to 67% in 2012)

Support life extended

- EOL originally planned for June 2012
- Lustre 1.8.9-wc1 released in February 2013
- 1.8.x client interop still planned for 2.4 (but deprecated in 2.5)

Xyratex plans to complete Oracle 1.8.8 release

# Active Codelines – Lustre 2.1.x

Increasing number of sites in production on 2.1.x

- LLNL, NASA and CEA all running in production
- OpenSFS survey showed 39% running in production (compared to 21% in 2012)

Will remain maintenance release stream until 2.4 GA

- 2.1.5 released in March
- 2.1.6 scheduled for release in coming weeks
- Will only have ad hoc releases after 2.1.6

RHEL servers and larger LUN size main attraction

# Active Codelines – Lustre 2.2

Feature release only – no maintenance releases scheduled

- Went GA in March 2012
- OpenSFS survey showed 8% using this in production (consistent with 2012 survey)

## Feature Highlights:

- Asynchronous Glimpse Lock/Statahead (LU-925, LU-389)
  - Improved performance for ls -l/find and accessing object attributes (file sizes/\*time, etc)
- Imperative Recovery (LU-19)
  - Faster recovery
- Large Xattrs (aka ‘Wide Striping’) (LU-80)
  - Maximum stripe size expanded from 160 to 2,000
  - Maximum file size grows from 320 TB to 64 PB

# Active Codelines – Lustre 2.3

Feature release only; no maintenance releases planned

- Went GA in October 2012
- OpenSFS survey showed 14% of sites running this in production

## Feature Highlights

- Server Stack SMP (LU-56, 1315 and 1316)
  - Performance improvements for running with larger number of cores
- OI scrub (LU-957)
  - Background consistency check



# Lustre 2.4.x

Will become the next maintenance release stream

- Lustre 2.4.0 scheduled for April 2013 GA
- Followed by 2.4.1 in Q3 and 2.4.2 in Q4; quarterly through 2014
- Drops support for RHEL5 and SLES 11 clients
- Adds support for SLES 11 SP2 and Fedora 18 clients
- Many organizations have delayed 2.x rebase until the release
- Complete details can be found at <https://wiki.hpdd.intel.com/display/PUB/Lustre+2.4>

# Lustre 2.4.x - Features

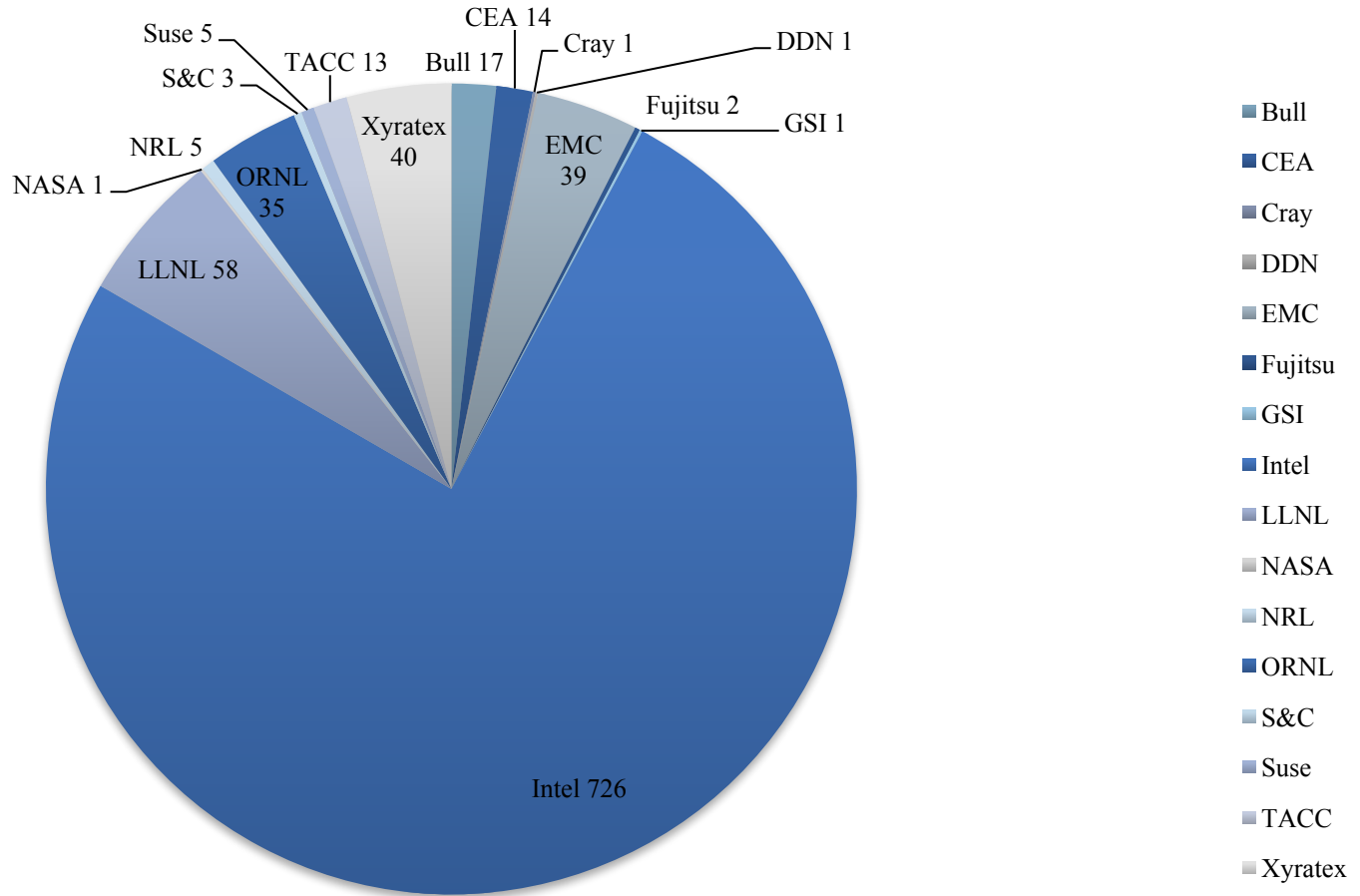
## Release Highlights:

- **4 MB I/O (LU-1431)**
  - Development by Xyratex – increases RPC size and improves performance with large I/O transfers to back-end disk
- **Disable Pinging (LU-2467)**
  - Upstreaming of Fujitsu FEFS feature; reduces ping overheads for large numbers of clients
- **DNE Phase I (LU-1187)**
  - Funded by OpenSFS
  - Multiple MDS/MDTs in a single file system, partitioned by subdir
- **Network Request Scheduler (LU-398)**
  - Mechanism to apply policies to how RPC requests are handled
  - Development by Xyratex based on Whamcloud prototype

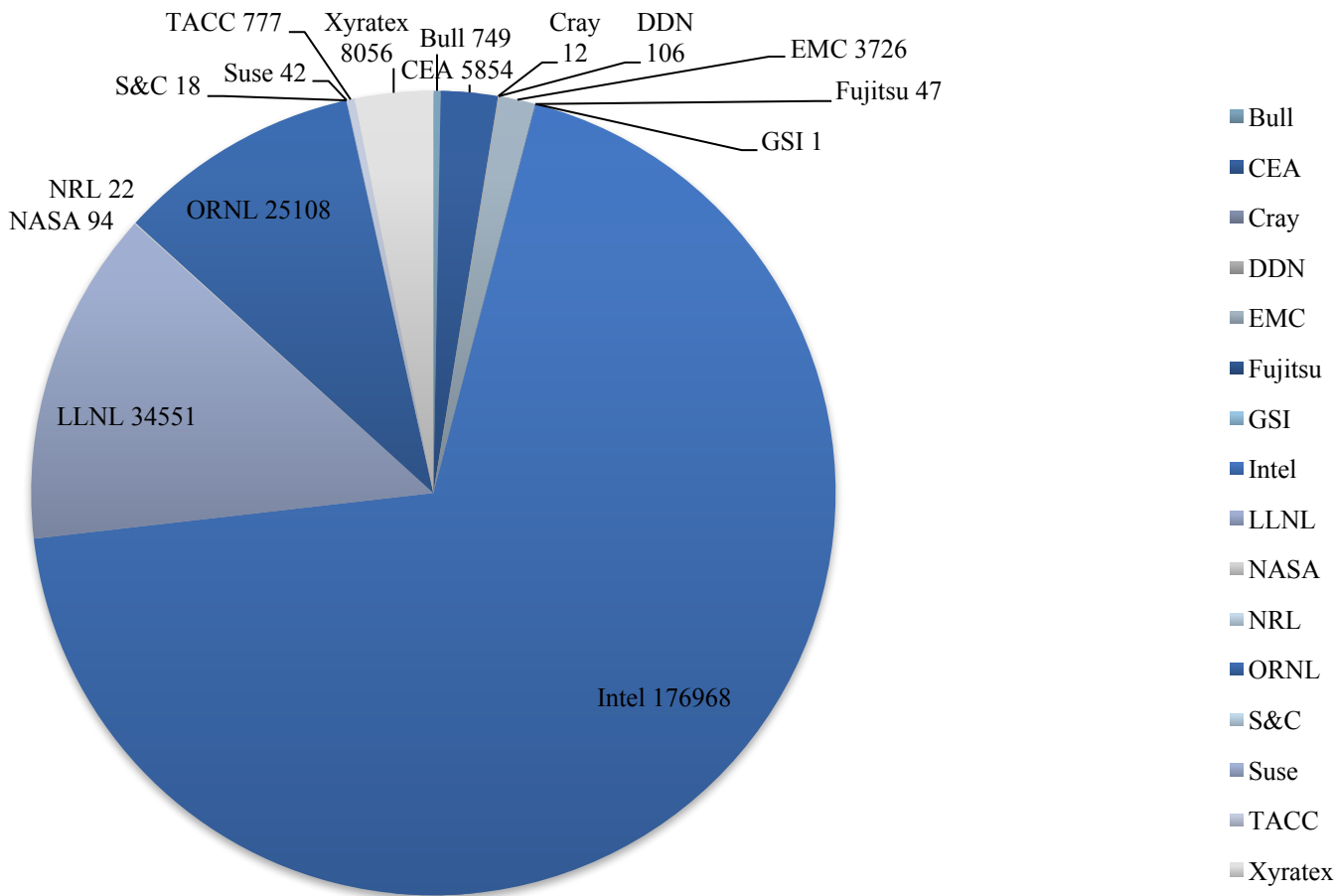
# Lustre 2.4.x – Features (continued)

- Layout Lock (LU-1876)
  - Prerequisite for HSM; allows clients to ensure they are doing I/O to correct OST Object
- LNET Networks Hashing (LU-2466)
  - Upstreaming of Fujitsu FEFS feature; mechanism to configure remote networks hash table size
- Single Client Performance (LU-744)
  - Addressing CLIO performance regression compared to 1.8.x
- Wireshark Support (LU-1434)
  - Enabled ability to use the Wireshark network packet inspection tool (see <http://www.wireshark.org>)
- ZFS Support (LU-1305)
  - Work funded by LLNL; alternative to traditional ldiskfs backend

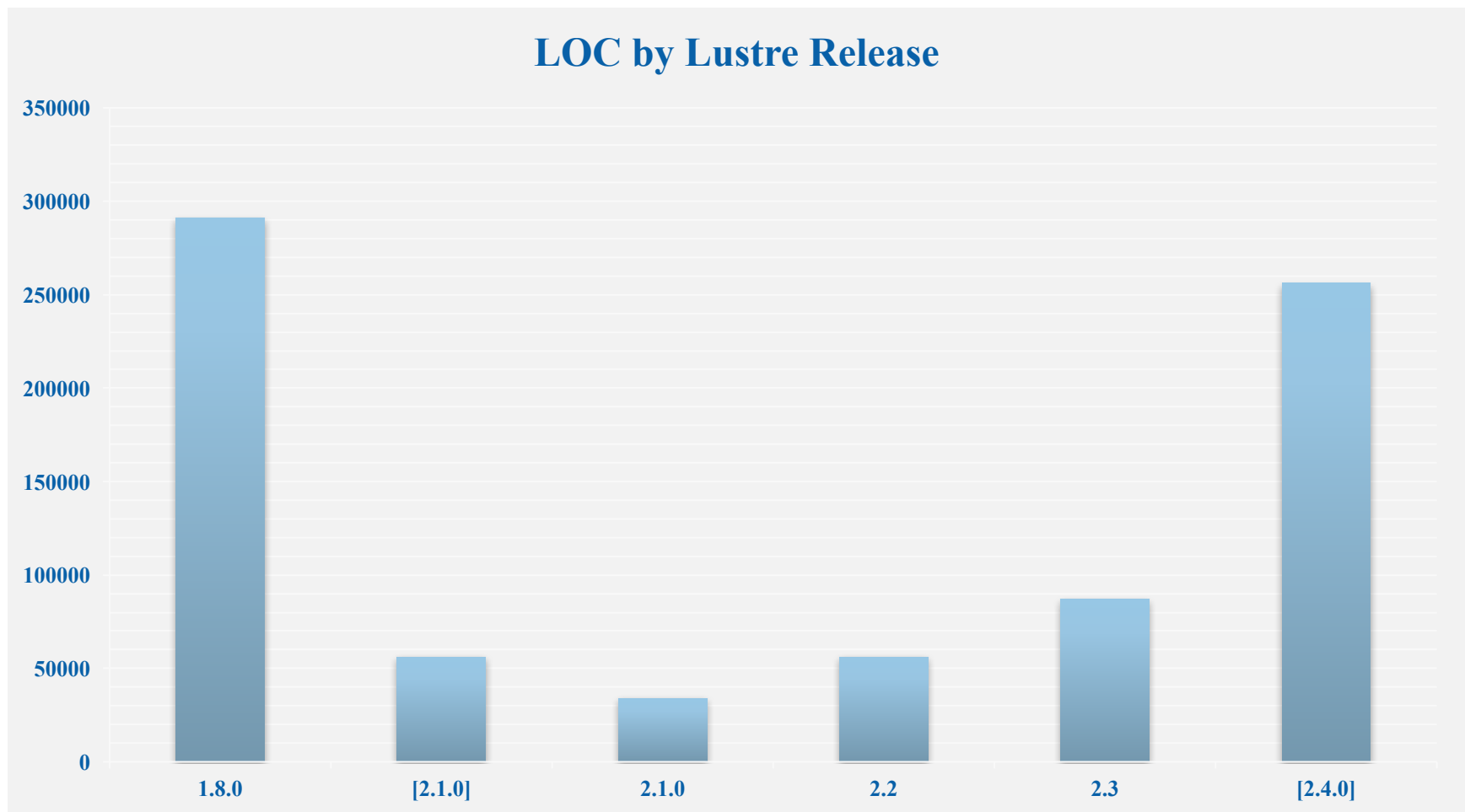
# Lustre 2.4.x – Contributions (up to April 9<sup>th</sup>)



# Lustre 2.4.x – LOC Contributions (up to Apr 9<sup>th</sup>)



# Lustre 2.4 – LOC Comparison



# Lustre 2.4.x – the Latest on 2.4.0

You can keep in touch on progress towards 2.4.0 being GA

- Updates biweekly to CDWG and hpdd-discuss mailing list
  - <http://lists.opensfs.org/listinfo.cgi/cdwg-opensfs.org>
  - <https://lists.01.org/mailman/listinfo/hpdd-discuss>
- Discussions in CDWG calls
  - If you cannot make it live read the minutes
  - [http://wiki.opensfs.org/Community\\_Development\\_Working\\_Group](http://wiki.opensfs.org/Community_Development_Working_Group)
- Watch the blockers list in JIRA
  - <https://jira.hpdd.intel.com/issues/?filter=10292>
- Direct from me now....

# Lustre 2.5

Feature release only, no maintenance releases planned

- Targeted for October 2013 GA
- Test matrix to be discussed at upcoming CDWG meeting
- Call for features already went out on mailing lists

## Feature Highlights

- HSM (LU-169, 827, 941, 1333 and 1338)
  - Enables data to be transferred between different storage types



# Release Quality

## Ongoing drive to raise the bar on quality

- Tightening landing requirements

## Continuous testing

- Automated testing on every proposed patch and landing
  - Reducing the delta between review and full tests runs
- Fuller manual testing on every tag (every 1-2 weeks)
- Hyperion dedicated time every month
- Community testing

## Extending testing capabilities

- Keeping current with test matrix (interop/distros)
- Jitter testing; virtual IB; DNE/ZFS

