ORACLE®

ORACLE®

Lustre Update for IDC HPC User Forum
Seattle – September 13, 2010

Peter Bojanic
Director, Lustre Group

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions.
The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.
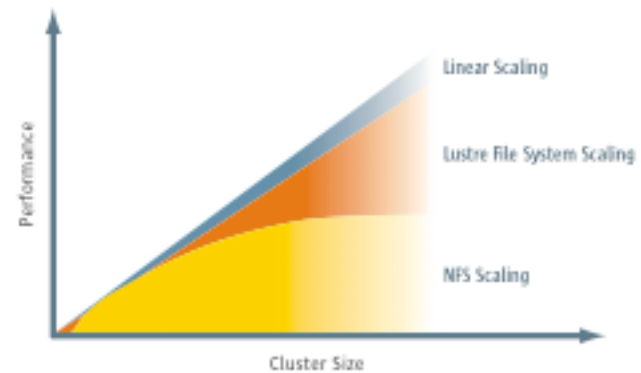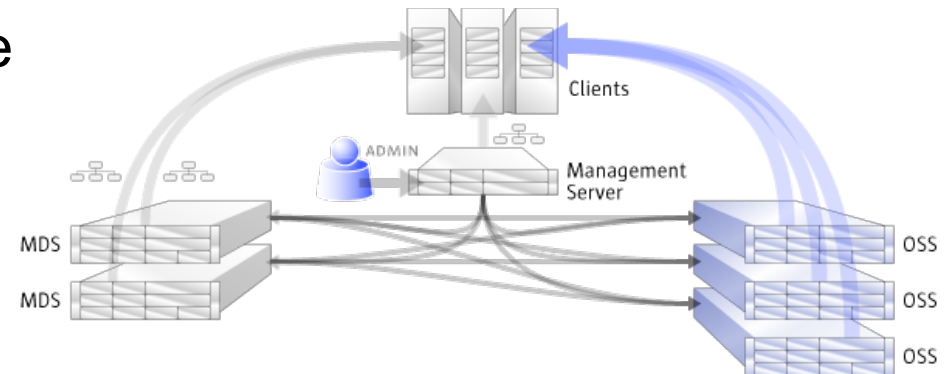
ORACLE

Here, now.
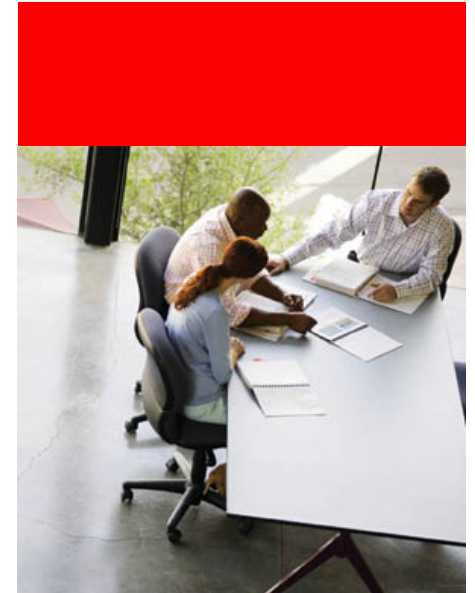
Lustre

Powers 7 of the top 10 Supercomputers in the World

- #1 in storage aggregate bandwidth
- #1 in storage scale
- #1 choice of supercomputing clusters

ORACLE

Module, Scalable, Highly Available

- Modular storage architecture consisting of metadata (MDS) and object storage server (OSS) modules

- File system IO throughput scales nearly linearly by adding incremental Object Storage Server modules

- High availability configurations offer transparent failover between redundant servers

Lustre Adoption in HPC

File Systems Used by Top 10 and Top 100 Supercomputers
(June 2010)

Top 10: Lustre 7, GPFS 2, PanFS 1, Other 0

Top 100: Lustre 61, GPFS 22, PanFS 3, Other 14

Legend:
- Lustre
- GPFS
- PanFS
- Other

ORACLE
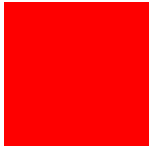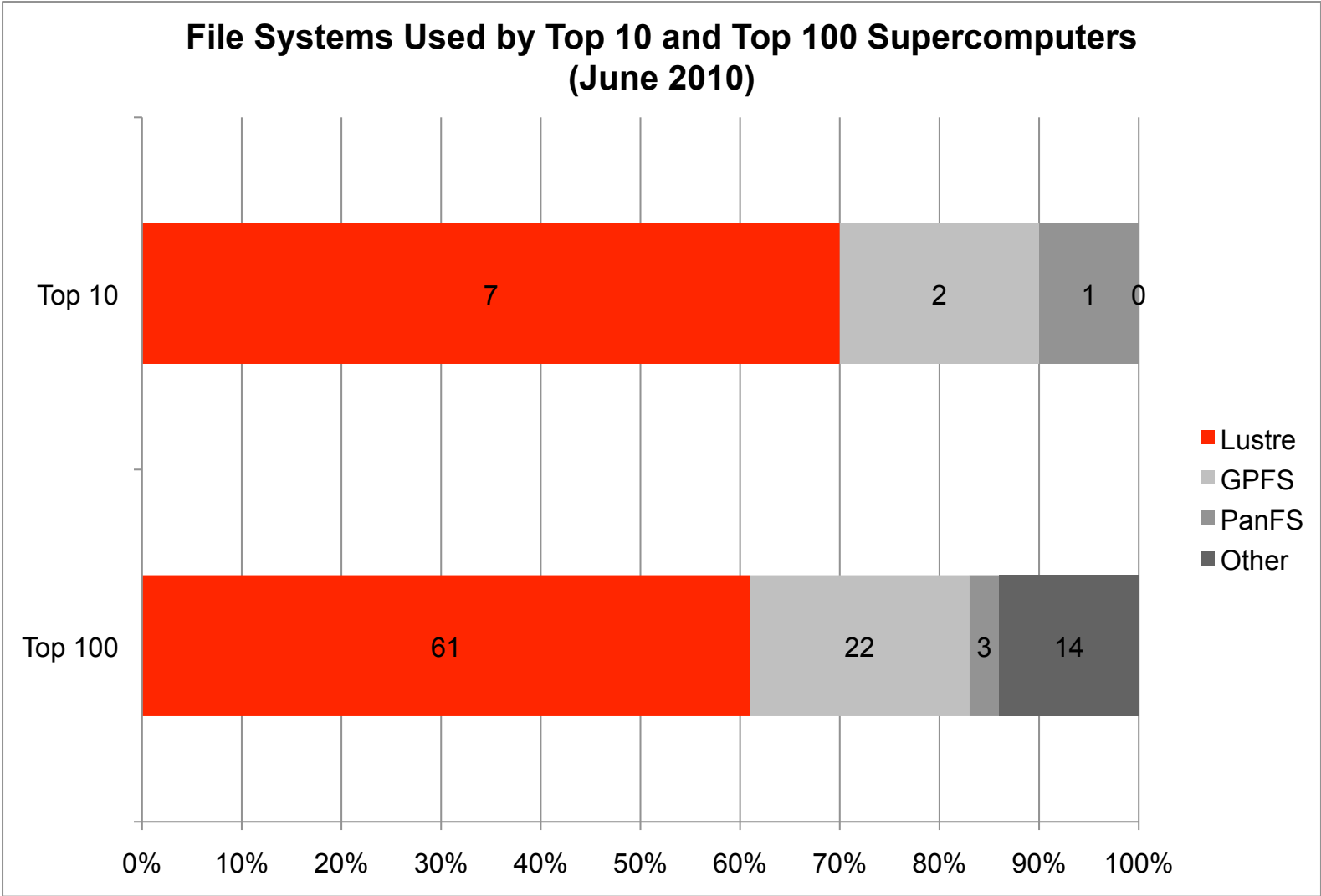
Lustre in Action
Jaguar + Spider Cluster at ORNL

- Storage solution based on Lustre 1.8
- 27,000+ clients / 265,708 processor cores
- 10.5PB storage
  - 96 Object Storage Servers
  - 6720 disk spindles
- 240 GB/s IO achieved

ORACLE

What's New in Lustre 2

- New architecture: Lustre 2.0 introduces a rewritten metadata server and client IO stack
  - This architecture establishes a stable foundation for platform portability and sustainable feature innovation
  - Metadata server is cluster capable with further development
- Changelogs allows selective monitoring of data and metadata changes without having to scan the entire filesystem
  - Lustre_rsync tool leverages Changelogs to provide namespace and data replication to an external (remote) backup system
- Recovery system improvements with Commit On Share
- Size-on-MDS (preview) improves stat operations by caching file object attributes on the MDS

ORACLE

Development & Testing Highlights

- 7 alpha releases
- 1 beta releases
- 2 release candidates
- >12,000 test runs since September 2009
- 527 bugs fixed against the "Lustre 2 Master" code base
- Many hundreds more bugs fixed against the "Hendrix 3" project source tree

- 0 blockers (S1/S2 issues)
- Consistent successful Hyperion runs
- Hyperion scale at 1000 clients for 72 hour continuous test
- Hyperion Master scale as good as b1_8 scale
- 1 billion file creates on Hyperion
- Lustre 1.8 client interoperability

Early Adopters

- Active alpha and beta testing by CEA & Bull
  - They have adopted every alpha and beta throughout the release cycle
  - 44 total bugs filed (32 fixed, 12 deferred)
  - CEA are in a semi-production rollout of 2.0 ahead of a planned rollout later this summer
- Pittsburgh Super Computing (PSC)
  - Focused on general and security testing
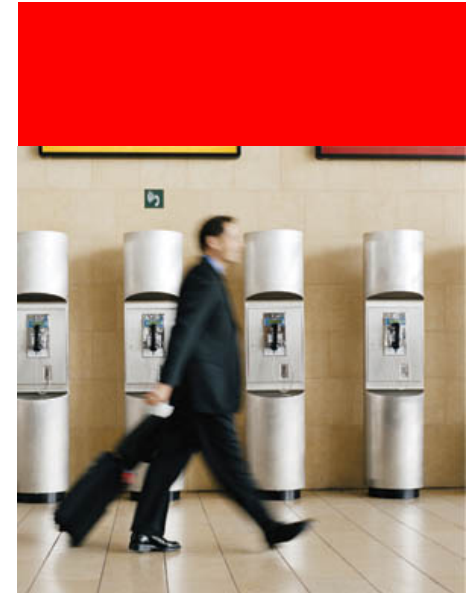 - 6 total bugs filed (4 fixed, 2 deferred)

Release Availability

- Available for early evaluators as open source software
  - Run servers on Oracle Enterprise Linux or RHEL
  - Run on Oracle Storage or other existing storage
  - Fully integrated and supported storage solutions will be available from Oracle in the future
- Release is available from a couple of sources
  - Check out from lustre.org open source git repository
  - Downloadable RPMs from Oracle's Download site
- Documentation
  - Operations Manual: http://dlc.sun.com/pdf/821-2076-10/821-2076-10.pdf
  - Release Notes: http://dlc.sun.com/pdf/821-2077-10/821-2077-10.pdf
- Users can file bugs through bugzilla.lustre.org

ORACLE

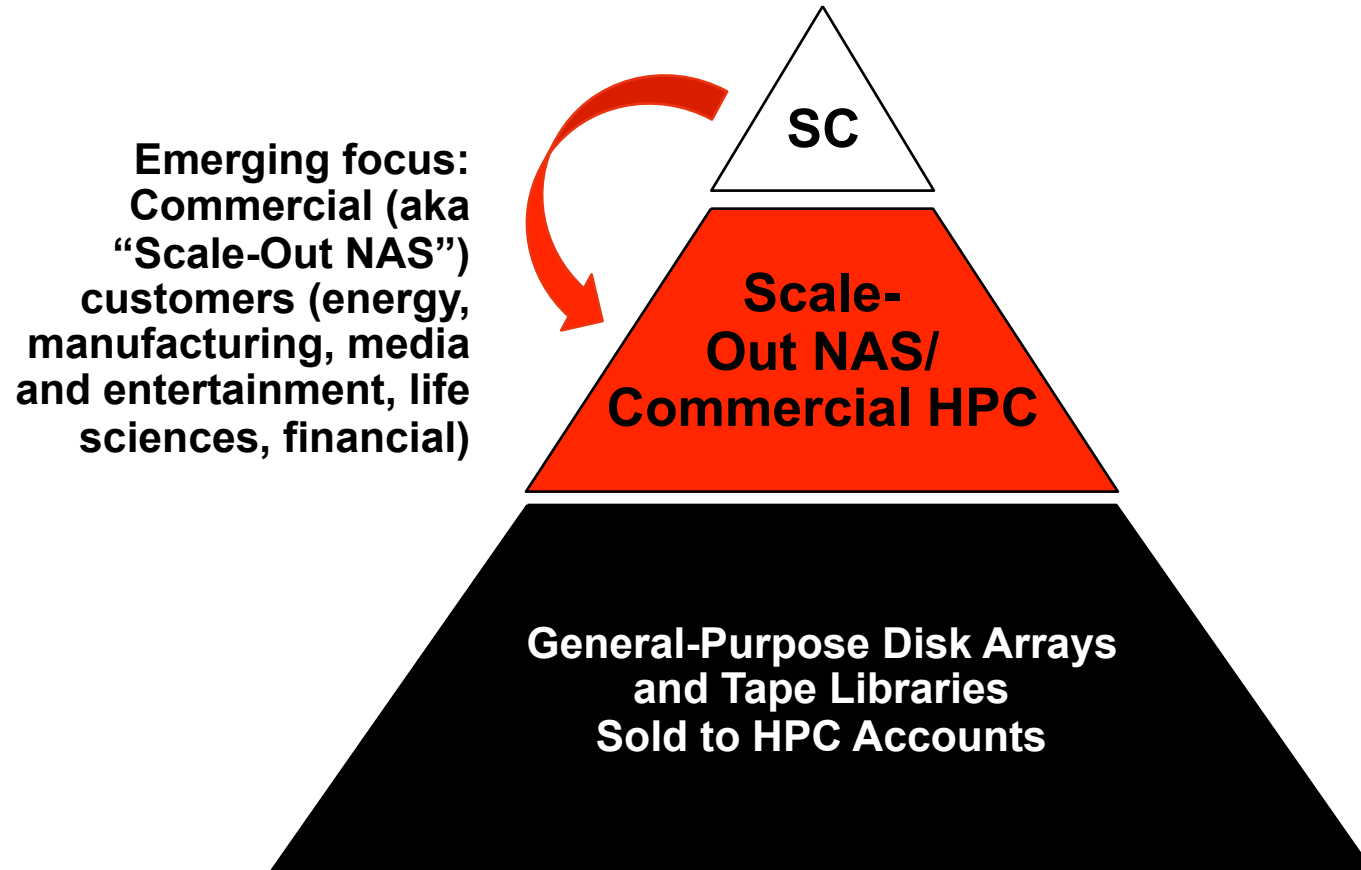# Where to?

ORACLE

Lustre Solutions from Oracle

Current focus: Supercomputer customers
(education, government, scientific)

SC

Emerging focus:
Commercial (aka
"Scale-Out NAS")
customers (energy,
manufacturing, media
and entertainment, life
sciences, financial)

Scale-Out NAS/ Commercial HPC

**General-Purpose Disk Arrays
and Tape Libraries
Sold to HPC Accounts**

ORACLE

New Best Practices for Enterprise Applications

- Typical applications: energy, manufacturing, media and entertainment, life sciences, and financial

- Users require both performance and usability

- Requirements for emerging data-hungry applications

  - Linearly-scalable global namespace

  - Easy deployment/ease of use

  - Reliability: data preservation and maximum uptime

  - Support: 24/7/365 worldwide support

  - Flexible protocol support: CIFS, NFS, HTTP attach, and native clients for Windows and Mac

  - Data services: snapshot, replication, de-duplication

ORACLE

Lustre 2 Qualification Program

- Oracle is developing a program to allow resellers to qualify their own reference configurations to run Oracle's Oracle Enterprise Linux software stack featuring Lustre 2

- This will enable Oracle to maintain good quality and a positive reputation for Lustre 2 and assures customers that the one canonical version meets the highest standards in the industry

- Under the program, resellers can qualify systems with a verification test plan, working closely with Lustre Engineering to configure, test, and tune the configuration

- Lustre 2 partners qualified under the program are entitled to use Oracle's officially supported Oracle Enterprise Linux software stack and to purchase an Oracle Premiere Support contract for each Lustre server (OSS and MDS), based on the qualified reference configuration

- The reseller remains responsible for the hardware and servers, including servicing and diagnostics

ORACLE

Feature Planning Underway…

These feature areas are being planned and do not *yet* have deliverables committed to a particular software update.
Depending on the output of our planning process, features *may or may not* be delivered within CY10 or CY11. Watch future revisions of the roadmap for updates.

Better performance through SMP scaling

Support for Lustre with ZFS on Solaris

End-to-end data integrity protection

Faster (imperative) recovery

Basic HSM support

Security (Kerberos authentication)

Clustered Metadata (CMD)

Support for OEL/RHEL 6 clients and servers

Prioritization feedback? Email lustre-interest@oracle.com

ORACLE

Summary

- Oracle is committed to long-term success for Lustre in HPC

- Lustre 2 raises the bar and establishes a solid baseline for future innovation

- Oracle's Lustre business strategy is focused on Enterprise HPC and Scalable NAS

- Oracle will work with HPC Partners to make Lustre 2 available through a Qualification Program

ORACLE

SOFTWARE.
HARDWARE.
COMPLETE.