ZFS Day 2012

Oct 2, 2012

Brian Behlendorf, Sequoia's 55PB Lustre+ZFS Filesystem

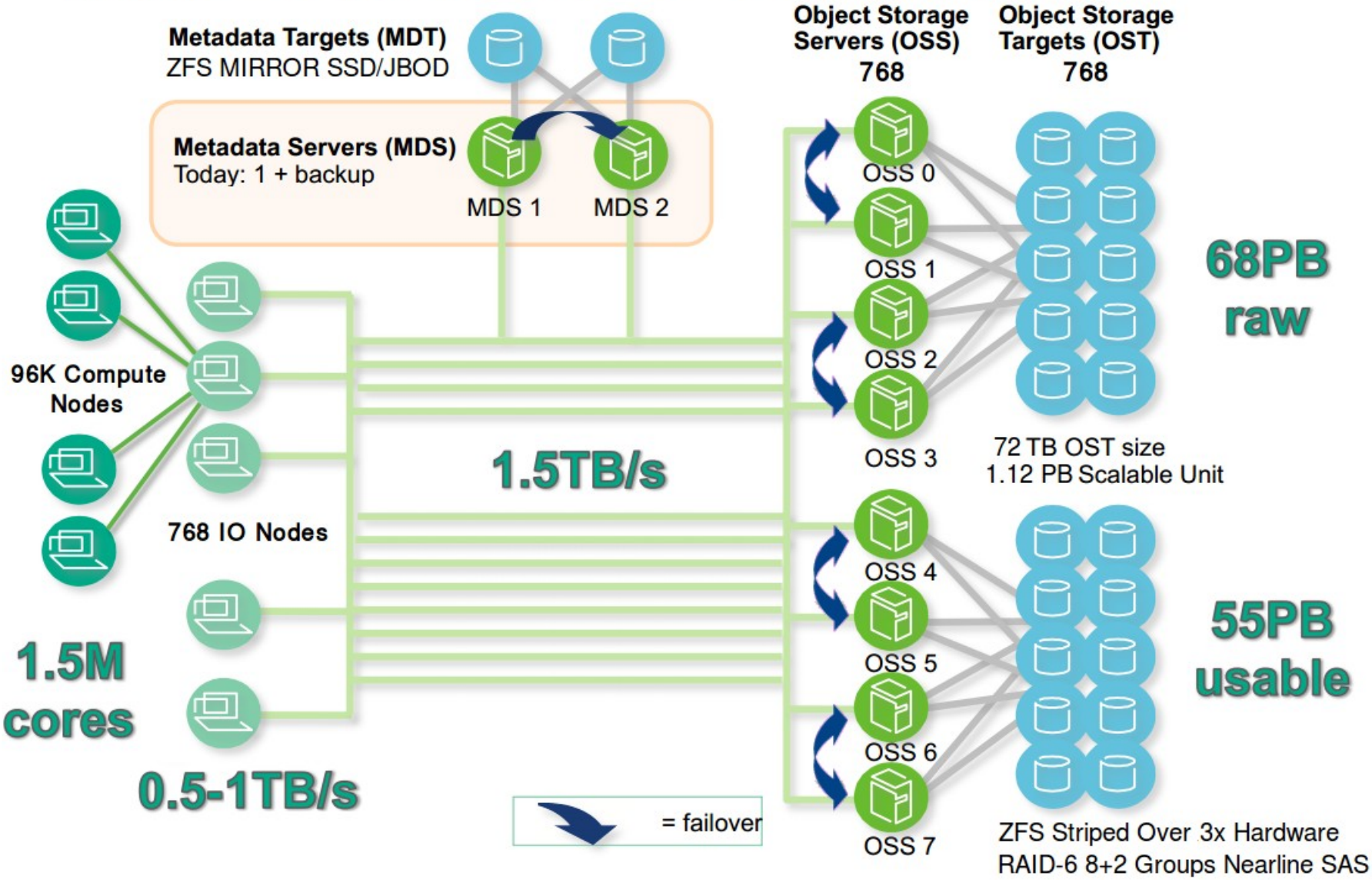
Lawrence Livermore National Laboratory

LLNL-PRES-551671

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract

DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

LLNL Sequoia Lustre Architecture

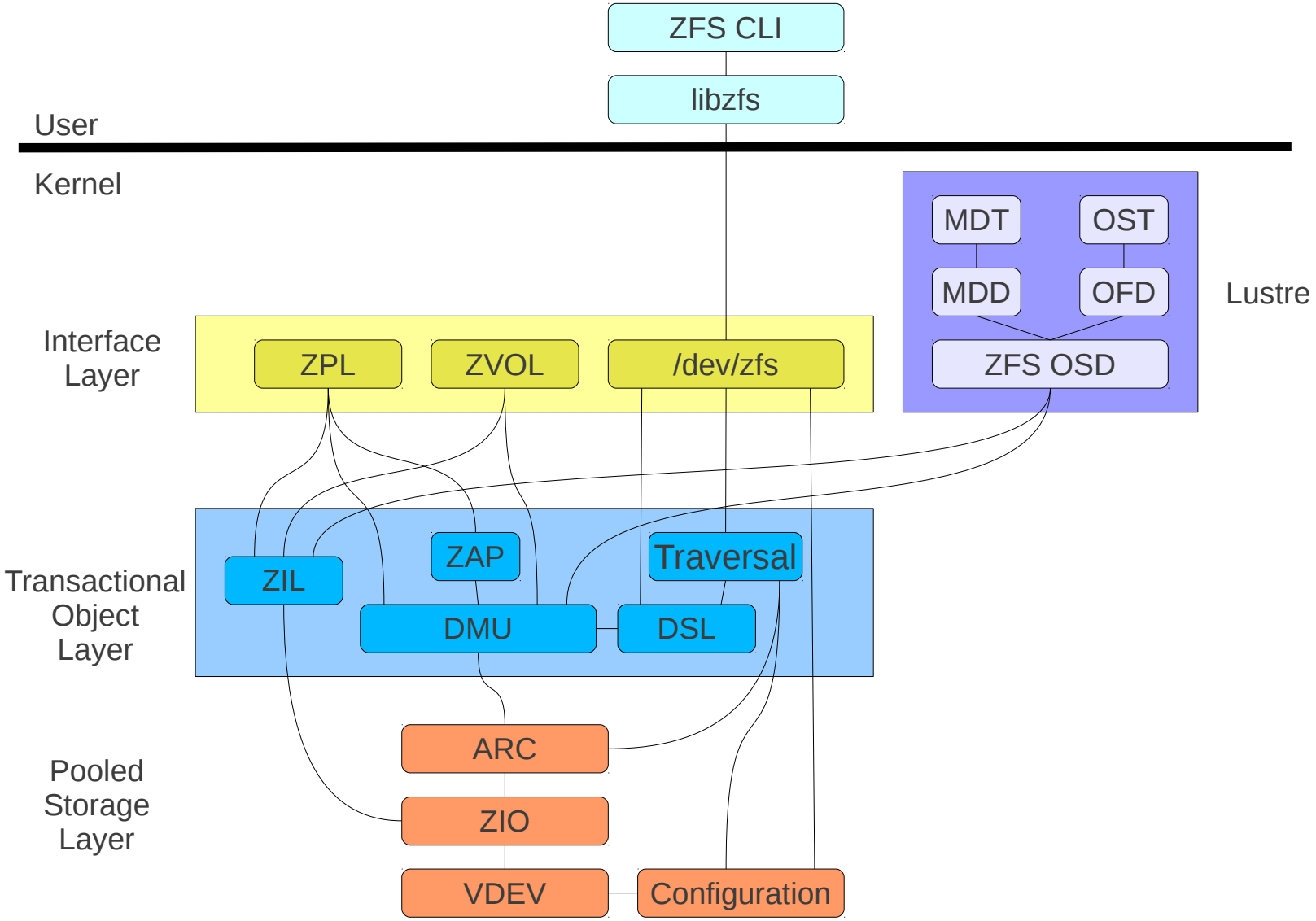



APPRO
APPRO
APPRO

Why ZFS

- Scalability
  - Massive storage capacity
    - 2^64 bytes per object
    - 2^78 bytes per pool
  - Dynamic Striping
  - Single OST per OSS
- Cost
  - Combined RAID+LVM+FS
  - Built for inexpensive disk
  - No vendor lock in
  - All open source
- Data Integrity
  - Copy-on-Write
  - Checksums
    - Meta data and block data
    - Checksums verified on read
    - Automatically repairs damage
  - Multiple copies of meta data
    - Small amount of storage
    - Spread over different disks
  - Ditto Blocks
  - Redundancy – Stripes, Mirrors, RAIDZ1, RAIDZ2, RAIDZ3

ZFS and Lustre Components



ZFS on Linux

- Stable Release Candidate
  - ZFS 0.6.0-rc11
- Community
  - Maintained packages:
    - Ubuntu PPA
    - Gentoo / Sabayon
    - Generic RPMs
  - Website
    - http://zfsonlinux.org
  - Mailing Lists
    - zfs-discuss@zfsonlinux.org
    - zfs-devel@zfsonlinux.org