Lustre on Red Sky

LUG 2010

April 15, 2010

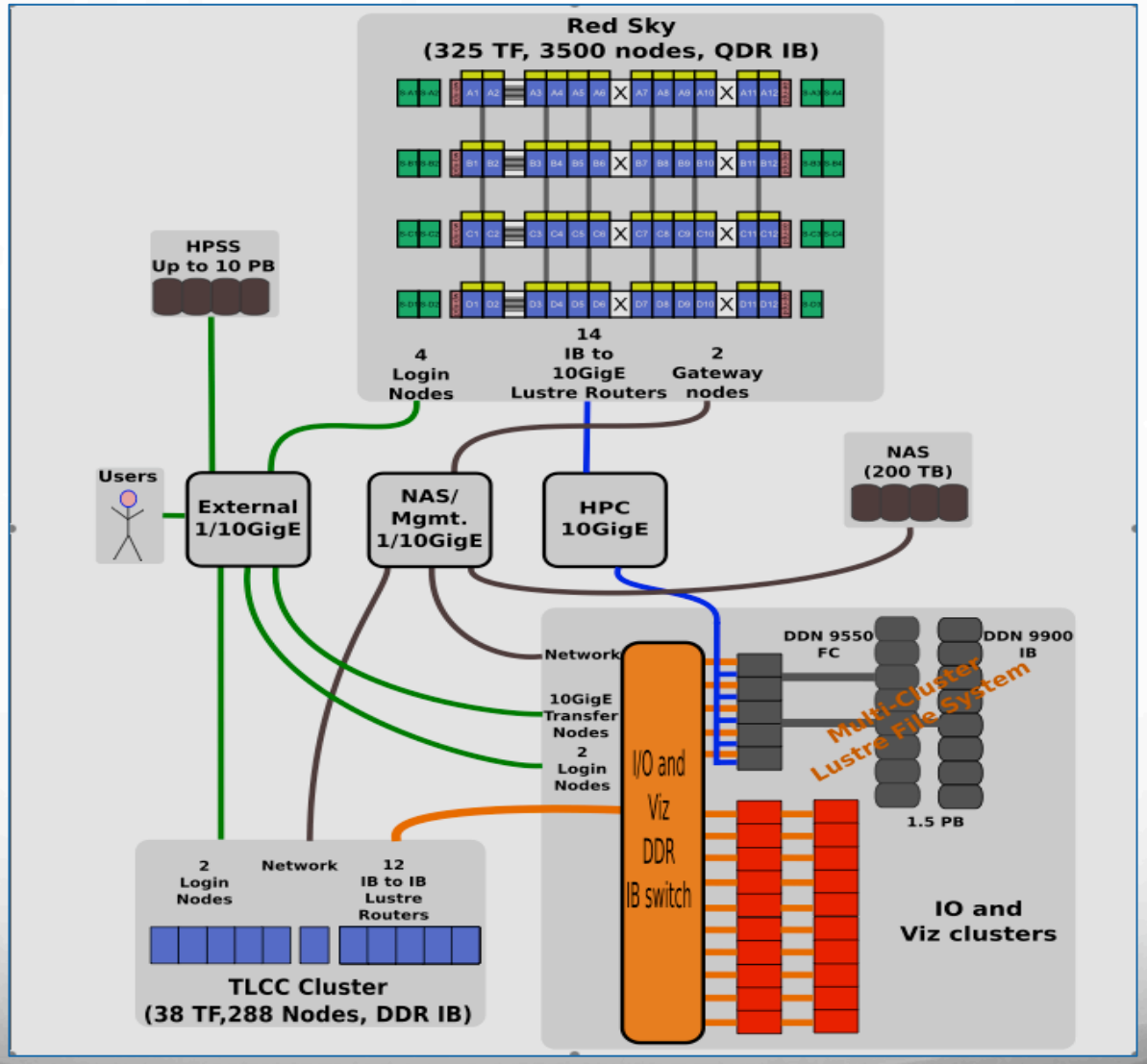
Steve Monk and Joe Mervini

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND 5281788

HPC at Sandia

- Capability Computing
  - Designed for scaling of single large runs
  - Usually proprietary for maximum performance
  - Red Storm is Sandia's current capability machine
- Capacity Computing
  - Computing for the masses
  - 100s of jobs and 100s of users
  - Extreme reliability required
  - Flexibility for changing workload
  - Thunderbird will be decommissioned this quarter
  - Red Sky is our future capacity computing platform
  - Red Mesa machine for National Renewable Energy Lab

Capacity Computing



National Nuclear Security Administration

Sandia National Laboratories

Red Sky main themes

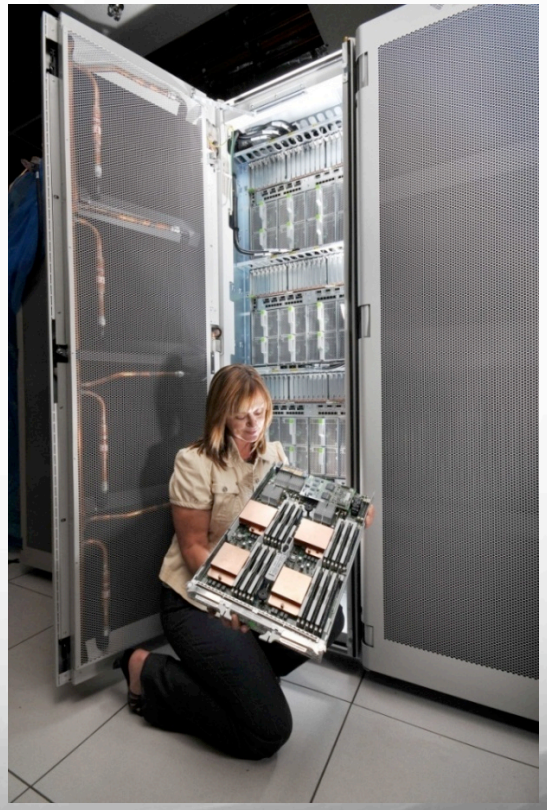
- **Cheaper**
  - Over 5X the capacity of Tbird at 2/3 the cost
  - Substantially cheaper per flop than previous large capacity machine purchases
- **Leaner**
  - Lower operational costs
  - Three security environments via modular fabric
  - Expandable, upgradeable, extensible
  - Designed for 6yr. life cycle
- **Greener**
  - 15% less power … 1/6th power per flop
  - 40% less water … 5M gallons saved annually
  - 10X better cooling efficiency … from 70% to 97%
  - 4x denser footprint

NNSA National Nuclear Security Administration

Sandia National Laboratories

Red Sky major innovations

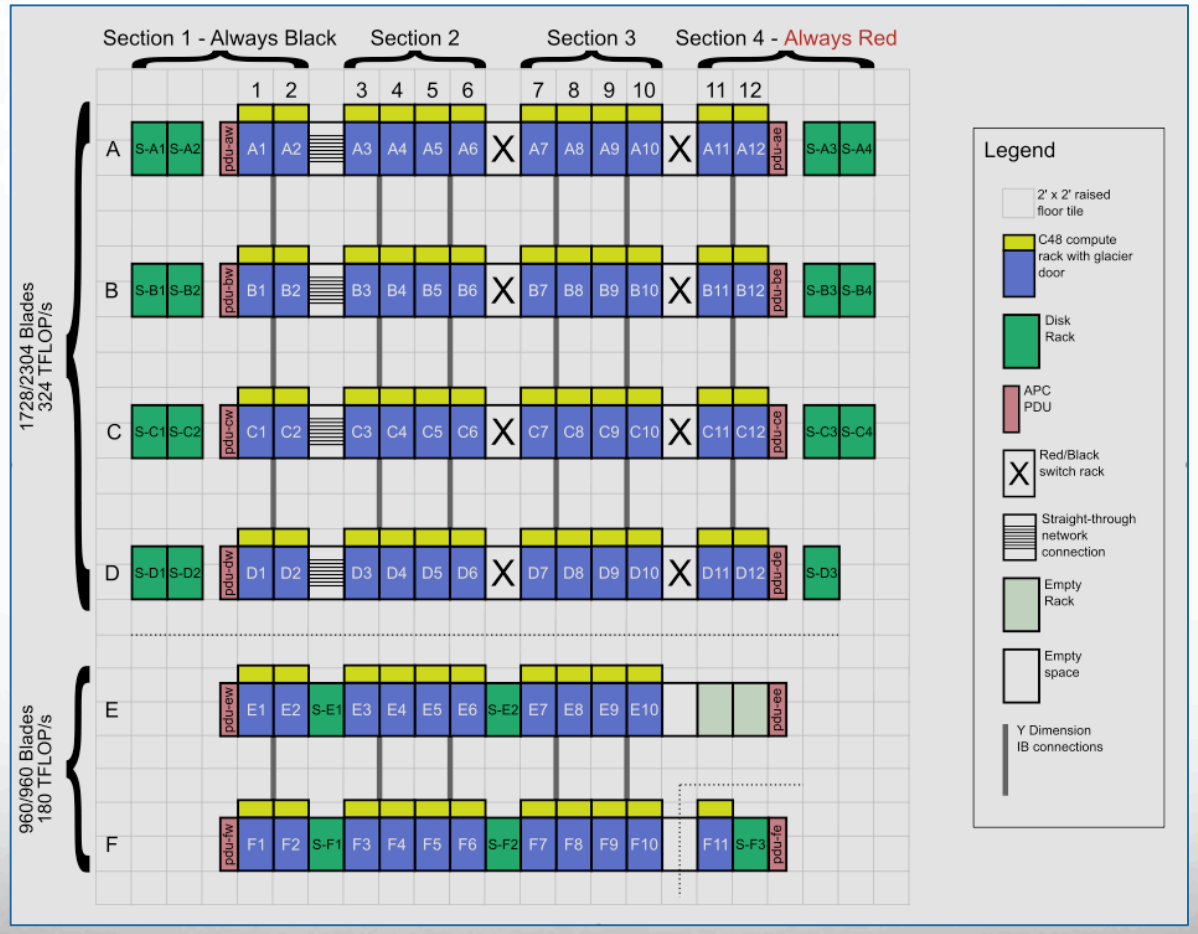
- **Bridging from capacity to capability**
  - Many “Red Storm” characteristics (scaling) at commodity price
  - 2-3X *faster* than Red Storm in mid range
  - 1/3 operational costs
- **Top Red Sky innovations**
  - Petascale midrange architecture
  - Intel Nehalem processor
  - QDR Infiniband
  - 3D mesh/torus
  - Optical cabling
  - Optical Red/Black switching
  - Refrigerant cooling/glacier doors
  - Power distribution
  - Routing & Interconnect resiliency
  - Minimal Ethernet (RAS & mgmt. only)
  - Boot over IB
  - Software RAID at scale



NNSA National Nuclear Security Administration

Sandia National Laboratories

Red Sky: Floor Plan



National Nuclear Security Administration

Sandia National Laboratories

Cheat sheet

RED SKY

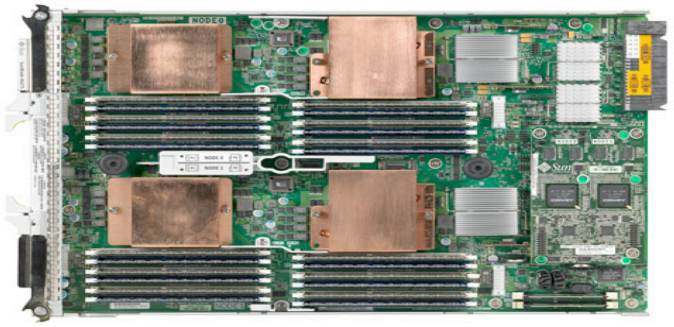
Hardware facts:

- 505 TF Peak
  - Red Sky: 325 TF
  - Red Mesa: 180 TF
- 5,386 nodes (2,693 Sun X6275 blades)
  - Total for Rows A-F (Red Sky + Red Mesa)
- 2.93 GHz quad core, Nehalem X5570 processor
  - 43,088 cores total
- 12 GB RAM per node (1.5 per core)
  - 64 TB RAM total
- 3D torus InfiniBand
  - QDR via Mellanox ConnectX on motherboard and InfiniScale IV in QNEM
  - 1,440 IB cables =9.1 miles (220 miles of optical strands)
- 124- J4400 JBOD storage enclosures providing ~ 6 PB (raw) for scratch, home and projects file systems
  - 5952 1TB Seagate SATA disks
- R134a based cooling doors
- 1.7 MW power (PUE of 1.27)
- 1,848 square feet of space in 6 rows
  - 68 Sun C48 cabinets
  - up to 96 nodes per rack
  - up to 768 cores per rack

Software facts:

- CentOS 5.3
- OFED 1.4.1
- SNL modified OpenSM (Subnet manager) with custom routing engine for 3D Torus (Torus-2QoS)
- Diskless boot over IB using a custom isolinux boot strap or GPXE
- SNL developed system management toolset
- SNL developed RAS system
- Lustre 1.8.x with patchless clients

X6275 Blade (2 nodes)



4/9/10

Red Sky near future

- IB Torus Quality of Service testing
  - So far we've not seen Lustre and MPI traffic "collide", but as we scale out we are anticipating this to be a potential issue
- Phases of Red Sky and Lustre
  - Phase 1: get the local disks into production first
    - Storage issues have caused significant delays
  - Phase 2: bring the 1 PB + site Lustre file systems to the machine using IB to 10GigE routers
    - Infrastructure is in place and we should have this completed by the end of this month

Lustre on Red Sky

Goals:

- Provide home/projects/scratch Lustre file systems
- Adhere to the Sun HPC stack
- Implement software raid on Sun provided JBODs
- Design for easy administration

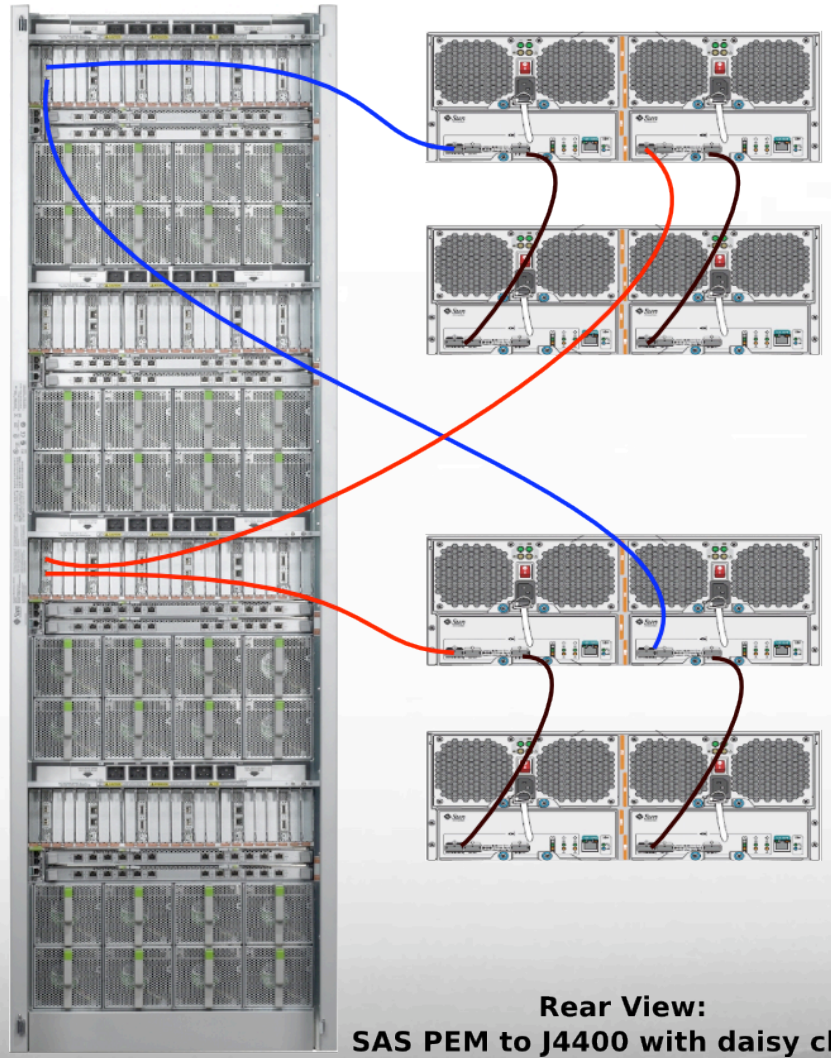
Storage Layout

- Three security domains require three sets of storage
- ~3 PB SATA, 68 TB SAS storage
- home/projects ~19 TB available/domain
- Scratch ~700 TB available/domain
- MDTs ~6.5 TB/file system

Software Raid

- MDTs configured as RAID10 arrays
- Home/projects OSTs configured as RAID10 arrays with external journal/bitmap
- Scratch OSTs configured as RAID6 8+2 arrays with external journal/bitmap
- Scripts developed to automate array creation
- Scripts developed to automate array assembly via UUID vs. mdadm.conf

SAS cable layout

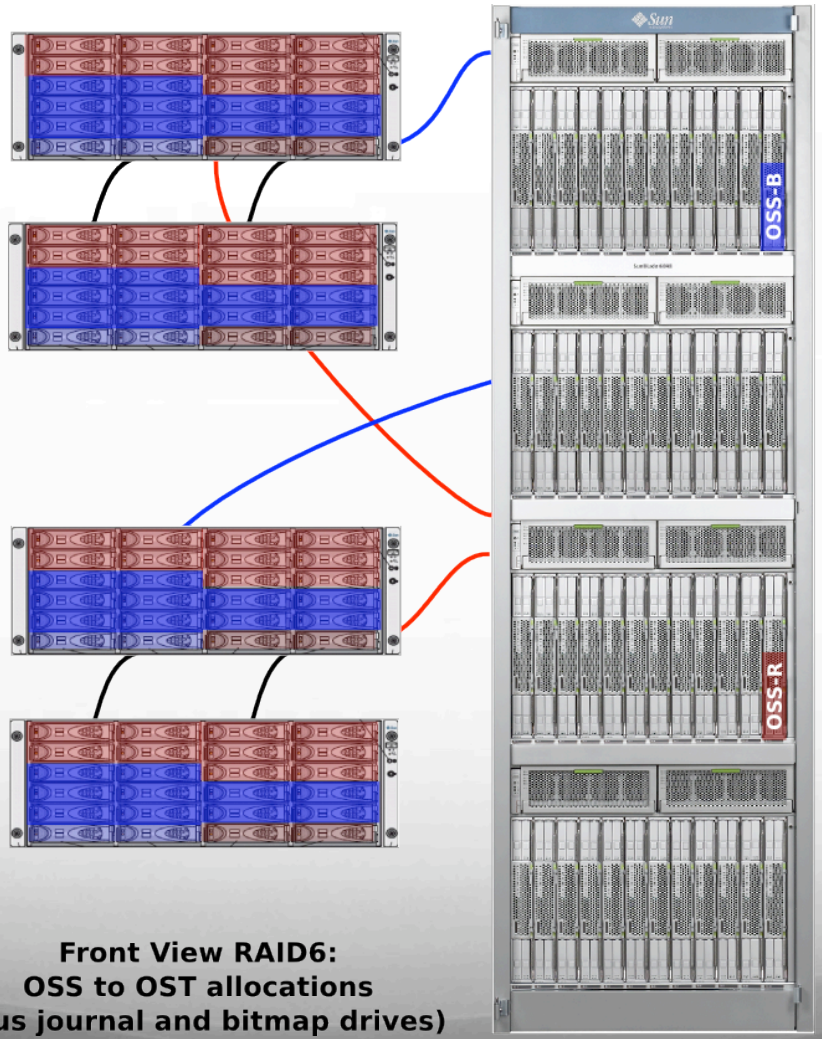


Rear View:
SAS PEM to J4400 with daisy chain

National Nuclear Security Administration

Sandia National Laboratories

RAID6 Physical Layout

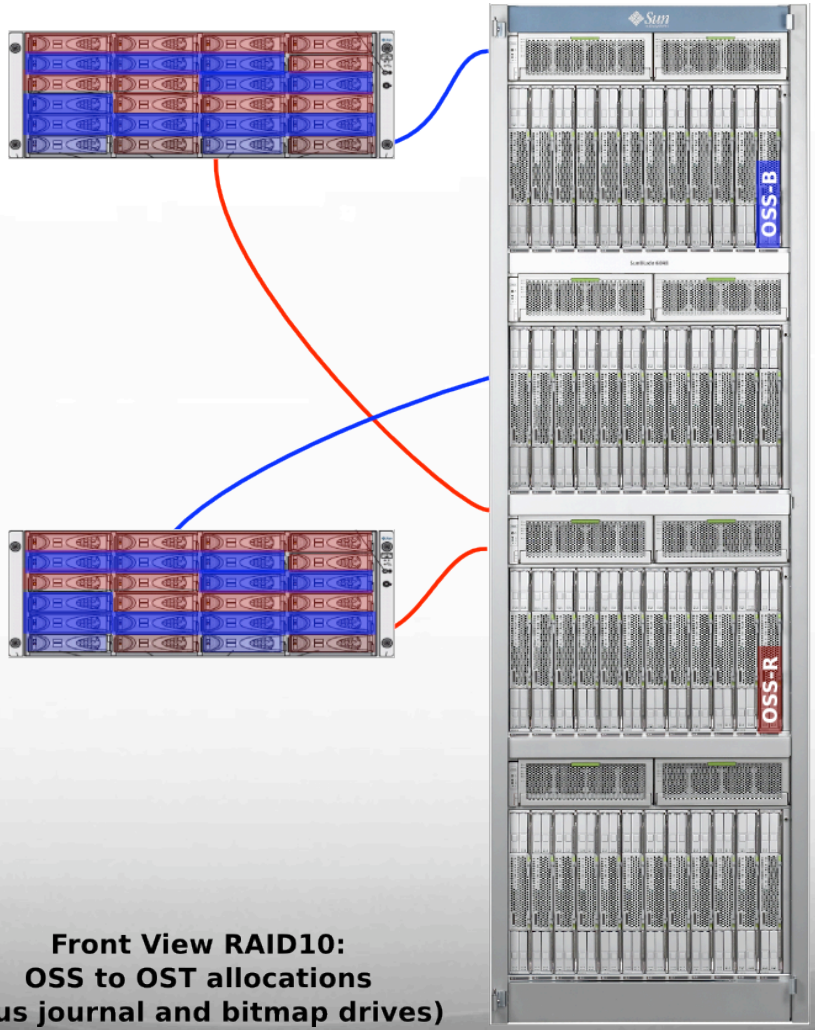


Front View RAID6:
OSS to OST allocations
(plus journal and bitmap drives)

National Nuclear Security Administration

Sandia National Laboratories

RAID10 Physical Layout



Front View RAID10:
OSS to OST allocations
(plus journal and bitmap drives)

National Nuclear Security Administration

Sandia National Laboratories

Startup

- Custom init script used to assemble and mount lustre OSTs/MDTs
- mdadm devices identified via common configuration file
- Includes switches from simple assembly to failover
- Additional automated processes employed

Configuration File Example

```
# mds-home1
#
mds-home1     data     /dev/md1     1ef0d052:bd33be71:5afa4482:ff4bfb87     mds-home2
## oss-home1
#
oss-home1     journal  /dev/md11    0469e61a:29d36257:62f61c6b:ad30a7f2     oss-home2
oss-home1     bitmap   /dev/md21    244f9e60:f06a8d98:40233d08:f14f4412     oss-home2
oss-home1     journal  /dev/md12    bb6248c1:140f335d:18e3b3b5:9c807b97     oss-home2
oss-home1     bitmap   /dev/md22    72a77611:68577fb9:63df278d:9dc6889e     oss-home2
oss-home1     data     /dev/md1     fe1afd7e:e61cb8b6:eb7a5af3:288168ea     oss-home2
oss-home1     data     /dev/md2     246a84a9:cadf0408:936a8e53:5082721d     oss-home2
#
# oss-home2
#
oss-home2     journal  /dev/md13    818c0f32:df64d3c2:b30e5074:bfcdf5c1     oss-home1
oss-home2     bitmap   /dev/md23    f8eeae5c:ae27c1f2:cc622350:240a687e     oss-home1
oss-home2     journal  /dev/md14    06bbf143:d1dab445:a0e3deca:77bd0d2b     oss-home1
oss-home2     bitmap   /dev/md24    3e4a5713:2ae474bd:4312b1c0:71e9f41d     oss-home1
oss-home2     data     /dev/md3     06b6eae3:6f881bec:6aa581bf:34f646fc     oss-home1
oss-home2     data     /dev/md4     b5552266:4f8ab4e1:d0a53a60:2be36bf0     oss-home1
```

Failover

- MDS and OSS servers configured in failover pairs
- Failover handled by init script
- Manual failover is an acceptable solution
- Automated failover will be tested and implemented in the future.
  - Linux-HA and STONITH
  - Hardened scripts to prevent false positives

Disk and Lustre Performance

- Individual disk performance on 1TB SATA drives average > 50 MiB/sec
- 1.2 GiB/sec theoretical per JBOD
- IOR used a primary test environment for Lustre
- Actual performance closer to 500 MiB/JBOD on RAID6 OSTs.
- Observed a 21% boost in performance when tested with RAID10 OSTs but at a cost of 37% reduction in capacity.

Difficulties

- Lack of any substantial configuration/delivery support of storage system from Sun PS group
- Difficulty getting the right people to address on-going hardware problems
- Hardware problems with hard drives/JBODs exposed during our load testing have made storage reliability questionable
  - Single disk failure causes file system hang
  - Abnormal addressing of drives during server reboot causing multiple host paths to a single drive
- Other machine specific hardware issues delayed storage testing (e.g. IB cable problems)
- Complexity of software RAID had a steep learning curve
- Red Sky storage issues delayed other projects (e.g. DDN 9900 based site file system)

Positives

- Lustre support has been exceptional
- Very few problems with Lustre 1.8.1.1
- Performance is respectable
- Early problems reported via Lustre support were resolved quickly with patches, tuning
- Very few disk failures compared with the total number of drives

Conclusion

- Software RAID includes additional risks and administration vs. hardware RAID solutions
- Limited testing of hardware in these configurations make it ill-suited for rapid deployment in a production environment
- Lustre has been a shining star on this machine
  - Red Sky users are pleased with its performance