XT7? Integrating and Operating a Conjoined XT3+XT4 System

NATIONAL CENTER
FOR COMPUTATIONAL SCIENCES

presented by

R. Shane Canon

and Don Maxwell

Leadership Computing Facility

Oak Ridge National Laboratory

ORNL XT7?

- XT History

- Specs

- Facilities

- Data Transfer

- Transition from PBSPro to TORQUE

- Lustre

- Problems

- Successes

- Plans

XT History

- January 2005
  - **XT3 Development cabinet**
    - **DevHarness**

- March 2005
  - **10 Cabinet XT3**
    - **2.4 GHz**
    - **RSMS/CRMS**

- April 2005
  - **10 more cabinets of XT3 at beginning of month**
    - **Routing**
  - **20 more cabinets at end of month**

- June 2005
  - **Final 16 cabinets for total of 56 XT3 cabinets**

- July 2006
  - **XT3 Dual Core 2.6 GHz**
    - **Many test shots for 1.5 prior to upgrade**

- November 2006
  - **XT4 Dual Core 2.6GHz**
    - **32 and 36 cabinets**

- March 2007
  - **XT3 and XT4 Combined for total of 124 cabinets**

Specs

- 11508 2.6 GHz Dual-Core AMD Opteron Computes

- 45 TB Memory (2 GB/core)

- 4 rows/3D Torus

- 20 cabinets of DDN 9550
  - **720 TB**
  - **3.06d firmware**

- Mix of SeaStars

- XT4 Dual Core / XT3 Single Core Service Nodes

- XT4 4Gb HBAs / XT3 2Gb HBAs

- Boot/SDB/PBS/SVC
  - **1GigE nodes**
  - **XT4 Dual Core**

- 8 logins/8 yods
  - **10GigE nodes**
  - **XT4 Dual Core**

- 80 oss/mds nodes
  - **2 4Gb HBAs**

- 1.5.31 UNICOS/lc + Vector14 (SPR 736879) + Lustre Timeout (738021)

Facilities

Starting on June 26, 2006 and finishing on November 11, 2006

1. Salvaged over 2 tractor tailor loads of equipment and furniture

2. De-installed 60 feet of cubicle partition walls and 150 ft. of demountable walls including all electrical outlets and data jacks

3. Packed and moved 100+ boxes of books, 10 book cases, conference room table and chairs, A/V equipment

4. Relocated the Ultra Science network (2 racks) to JICS computer room

5. Relocated 3 individual servers and two racks of servers to the south side of the computer room

6. Relocated 4 racks of a Dell cluster (128 nodes) 15 individual servers and various racks from the Cluster lab to JICS computer room

7. Moved 11 cabinets containing 60K+ CDs

8. Removed A/V equipment from conference room B228

9. Demolished the sheetrock walls and ceiling of conference B228

10. Completely removed 8000 sq. ft. of 2' raised floor tiles, pedestals, and stringers and reinstalled a 3' raised floor

11. Removed 5 access doors and installed 2 access doors. Relocated and installed door cameras and proximity card access readers.

11. Completely removed the under floor wet pipe sprinkler system and reinstalled in a new configuration

12. Completely rearranged the over head wet pipe sprinkler system into a new configuration

13. Completely removed the under floor VESDA smoke detection system and reinstalled in a new configuration

14. Completely removed all of the ceiling tiles and rearranged 72 four foot light fixtures

15. Removed three under floor HVAC cooling systems

16. Performed four 12" hot taps on two 18" chilled water pipes on an operating system

17. Installed and insulated 175" of 16" chilled water supply and return pipes from the Central Energy Plant through the electrical room and into the 1st floor computer room.

18. Punched two 24" holes through to the 2nd floor computer room for the chilled water pipes

19. Installed and insulated 225' of chilled water pipe on the 2nd floor computer room to form a loop around the 8000' area

20. Built a 12' X 20' enclosure around chilled water pipes in 1st floor ceiling.

21. Purchased and installed 34 30T Liebert HVAC units

22. Installed 14 400A breakers in a 4000A switchboard

23. Installed a massive wire chase way in the 1st floor electrical room

24. Purchased and installed 11 300kVA power distribution units

25. Installed 1800' (14 separate runs) of 4" conduit from the 1st floor electrical room through the 1st floor computer room and through the slab to the 2nd floor computer room

26. Pulled 14 runs 500 MCM power cable (14 runs X 5 power cables each)

27. Ordered and installed an 8000 sq. ft. of expanded metal ceiling

28. Installed two 400A power panels for the new Liebert HVAC units

29. Installed Johnson Controls monitoring devices on the Lieberts

30. Built and installed 175' of overhead and under floor air dams to contain the cooled air in the 8000 sq. ft. area

31. Installed 175' of hand rails

32. Raised three existing power panels, all wall outlets, and data jacks 1 foot to meet code requirements for new 3' raised floor

33. Installed new emergency exit signs and fire alarms

Data Transfer Options

- Considered several options and techniques
    - **Have users transfers their data**
    - **GridFTP, scp**
    - **Cross mounting XT3 on XT4 (or vice versa)**
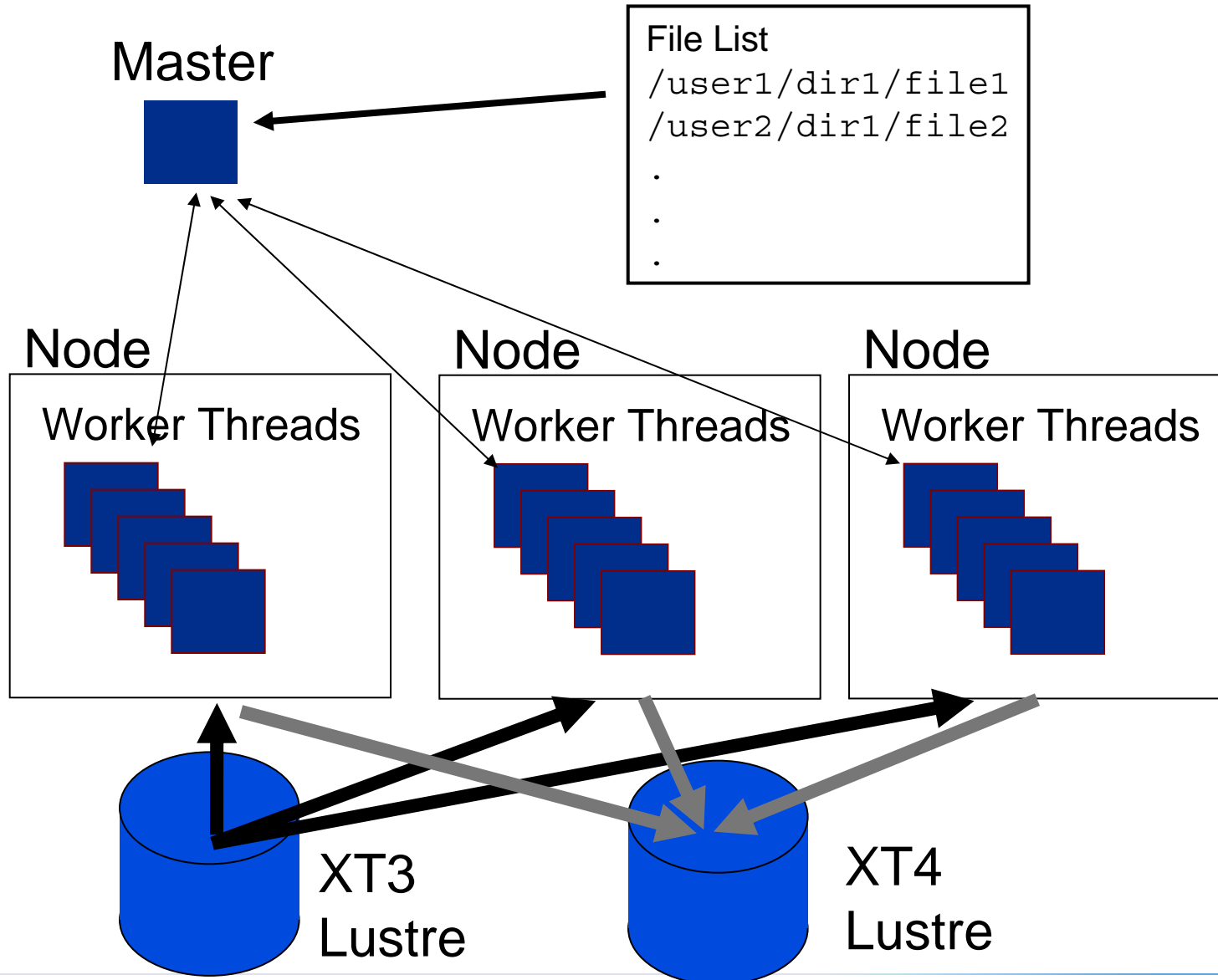
Data Transfer Approach

- First ask users to backup data to HPSS and delete what they can

- LCF Staff performed the transfer

- Transferred data using external nodes
  - **Work around for Lustre client performance bug in Cray stack**
  - **Patch for bug applied to client on external nodes**

- Verify data after copy

Mounting XT File System on External Hosts

- **Utilized Lustre Routing**
  - **ORNL already testing this capability for *Spider* file system**
  - **Used multiple LNET routers on XT system**
  - **Routers and external nodes connected to 10G Ethernet network**

```
options lnet ip2nets="\
  ptl1        192.168.*.*            # XT3 linux nodes;\
  ptl0        192.168.*.*            # XT4 linux router;\
  tcp0(eth0)  160.91.205.[215,218]  # XT4 router;\
  tcp0(eth0)  160.91.205.[210,211]  # XT3 router;\
  tcp0(eth2)  160.91.195.[60-80]    # transfer nodes"\
 routes="\
  ptl0  1      160.91.205.[215,218]@tcp0 # XT4 <- IP;\
  ptl1  1      160.91.205.[210,211]@tcp0 # XT3 <- IP;\
  tcp0  1      [8,12]@ptl0       # IP <- XT4;\
  tcp0  1      [24,12824]@ptl1  # IP <- XT3 ;"
```

Transfer Method (and Verification)

Master

File List
```
/user1/dir1/file1
/user2/dir1/file2
.
.
.
```

Node
Worker Threads

Node
Worker Threads

Node
Worker Threads

XT3
Lustre

XT4
Lustre

Data Transfer

- Transferred around 20 TB of data

- Peak transfer rates around 750 MB/s
  - **Limited by two routers on XT3**
  - **Used 5 transfers nodes each running 5 threads**

- No mis-compares

- Working on utility to perform parallel copy
  - **Uses MPI**
  - **Appropriate for users**

Transition from PBSPro to TORQUE

- Why TORQUE?
  - **Need for specifying XT3 or XT4**
  - **PBSPro MOM throws away nodes chosen by MOAB**
    - **Early on, problems for reservations**
    - **Continuous synchronization**
  - **Tighter integration with MOAB**
  - **Open Source**
    - **Added spool to PBS_O_WORKDIR**
    - **qstat**
      - **Changed Tasks to display size resource**
      - **Tasks to optional 5-digit field**
  - **MOAB creates CPA partition when using TORQUE**
    - **Placement algorithm**

Node Attribute Implementation

- Use MOAB – not TORQUE – to define attributes (features in MOAB terms)

- #INCLUDE of files containing features in moab.cfg
  - **NODECFG[0] FEATURES+=xt3**

- User asks for feature
  - **qsub –l feature=xt3|xt4 job.pbs**

- Span feature defined
  - **Spans XT3 and XT4**
  - **Request for count of features**
    - **-l feature=xt3:1000,xt4:1000**

- Reservations based on features
  - **troubleshooting and testing hardware**

Moab/TORQUE Configuration

- Two queues
  - **Debug**
    - **Standing reservation for 10% of machine**
    - **Wallclock limit**
  - **Production**
    - **Step-wise wallclock limit based on jobsize**

- Scheduler
  - **Priorities normalized to minutes since QueueTime is based on minutes**
  - **Multi-dimensional Fairshare**
    - **Allocation**
    - **Jobsize**
  - **Account Priority**
    - **Allocation**
    - **No Allocation**
    - **Exhausted Allocation**
  - **QueueTime**
  - **Quality of Service for the chosen ones**

Lustre

- 72 OSSs / 8 MDSs (3 MDSs currently used)

- 3 filesystems
  - **2 filesystems that span all OSSs but only half of DDN couplets**
    - **150 TB**
    - **Each has 72 OSTs**
    - **Default scratch for users on one or the other**
    - **Risk mitigation based on DDNs**
      - **Hopeful that OSS failover will eventually work**
  - **Large filesystem spans all OSSs (both HBAs) and all DDN couplets**
    - **300 TB**
    - **144 OSTs**
    - **Each user has access**
    - **OSTs use one HBA on each OSS then loop back and use next HBA on each OSS**

- Small piece of each tier available for Cray/ORNL test filesystems

DDN 9550s

- 36 TB per couplet

- 1 MB cache size

- 2TB LUNs

- 2 Tiers/LUN

- Initial problems with Task Exceptions
  - **Problem resolved by dual power feeds**
  - **3.06d firmware fix to reboot controllers when problem encountered**

2 Tiers/LUN A Win

	1 Tier/LUN	2 Tiers/LUN
FPP Write	30 GB/s	22 GB/s
FPP Read	29 GB/s	44 GB/s
SSF Write	28 GB/s	27 GB/s
SSF Read	27 GB/s	45 GB/s

File per process (FPP) writing 2.8 TB:
Striping: size 8m, start -1, count 1
processors: 576, 4:1 processors to stripes ratio
transfer size: 8MB

Single shared file (SSF) writing 2 TB:
Striping: size 4m, start -1, count -1
processors: 432, 3:1 processors to stripes ratio
transfer size: 128MB

Problems

- Hardware
  - **Samsung DDR2 memory**
  - **XT4 Mezz cards**
  - **VRMs on XT4 mezz cards**
    - **Change to fan speed based on pressure**

- Lustre
  - **MDS overwhelmed during acceptance**
    - **23000 file creations at once (FPP)**
    - **Simpleio code caused timeouts**
      - **SMP kernel didn't help**
      - **Workaround to increase timeouts from 300 to 600**
  - **OSSs completely disappear after large GTC runs**
    - **Could not recreate with either GTC or MADbench from NERSC**

- Certain jobs cause many compute nodes to be marked down
  - **Warmboot works**

Successes

- "The all-orders spectral algorithms (AORSA) code is a full-wave model for radio frequency heating of plasmas in fusion energy devices such as ITER, the international fusion energy project. AORSA just completed its biggest problem ever on jaguar."

- "Jaguar has now beaten the previous GTC performance record."

Plans

- XT4 Quad-core in Fall 2007 for theoretical peak of 250TF

- CNL and N-way Catamount testing in support of quad-core

- Center-wide Lustre filesystem

- Petaflop system in late 2008

Acknowledgments

ORNL

- Josh Lothian (XT Admin)

- Ken Matney (Lustre Tools)

- Makia Minich (Lustre Routing and Testing)

- Sarp Oral (Lustre Testing)

- Trey White (Titles)

Cray

- Jeff Becklehimer (Princ. Engineer)

- Cathy Willis (Site Analyst)

Cluster File Systems, Inc.

- Eric Barton

Cluster Resources Inc.

- Dave Jackson

- Scott Jackson