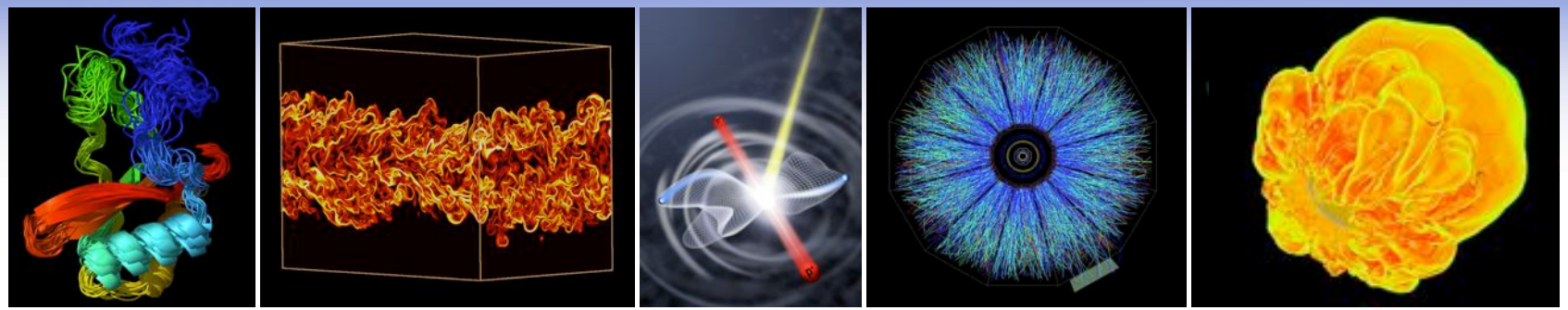


External Services on the NERSC Hopper System

Katie Antypas, Tina Butler, and Jonathan Carter

Cray User Group
May 27th, 2010

U.S. DEPARTMENT OF ENERGY | Office of Science

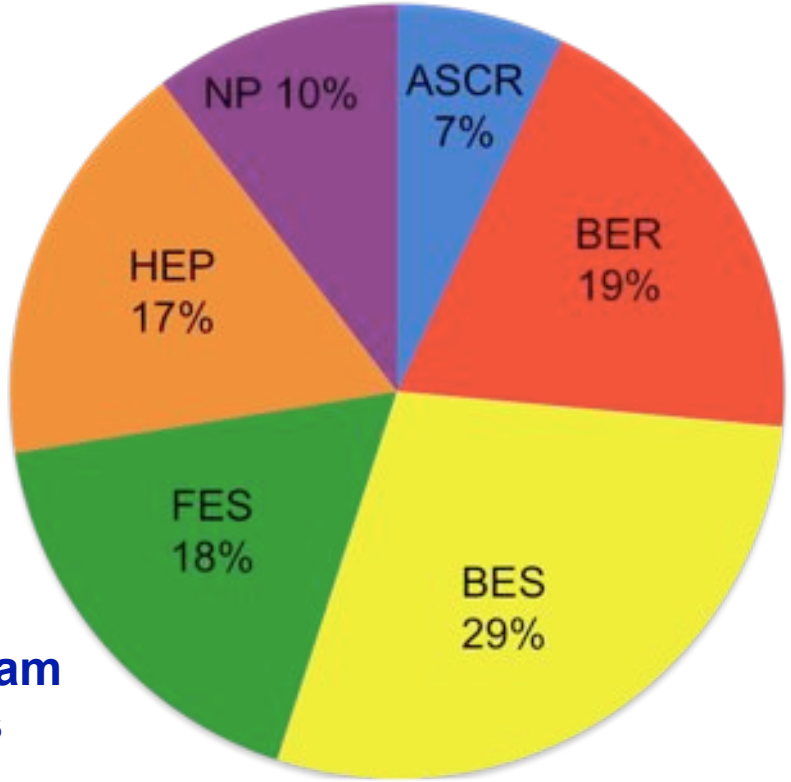
NERSC

BERKELEY LAB

NERSC is the Production Facility for DOE Office of Science

- **NERSC serves a large population**
  - **Approximately 3000 users,**
  - **400 projects, 500 code instances**
- **Focus on**
  - **Expert consulting and other services**
  - **High end computing systems**
  - **Global storage systems**
  - **Interface to high speed networking**
- **Science-driven**
  - **Machine procured competitively using application benchmarks from DOE/SC**
  - **Allocations controlled by DOE/SC Program Offices to couple with funding decisions**

2009 Allocations



NERSC Systems for Science

Large-Scale Computing System

Franklin (NERSC-5): Cray XT4
- 9,532 compute nodes; 38,128 cores
- ~25 Tflop/s on applications; 356 Tflop/s peak

Hopper (NERSC-6): Cray XT
- Phase 1: Cray XT5, 668 nodes, 5344 cores
- Phase 2: > 1 Pflop/s peak (late 2010 delivery)

Clusters

Carver
- IBM iDataplex cluster

PDSF (HEP/NP)
- Linux cluster (~1K cores)

Cloud testbed
- IBM iDataplex cluster

NERSC Global Filesystem (NGF)

Uses IBM's GPFS
1.5 PB; 5.5 GB/s

HPSS Archival Storage

- 59 PB capacity
- 11 Tape libraries
- 140 TB disk cache

Analytics / Visualization

- Euclid large memory machine (512 GB shared memory)
- GPU testbed

~40 nodes

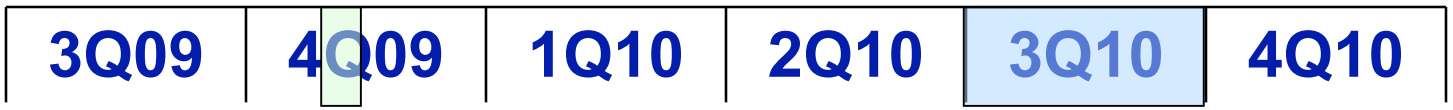
Hopper System

Phase 1 - XT5

- 668 nodes, 5,344 cores
- 2.4 GHz AMD Opteron (Shanghai, 4-core)
- 50 Tflop/s peak
- 5 Tflop/s SSP
- 11 TB DDR2 memory total
- Seastar2+ Interconnect
- 2 PB disk, 25 GB/s
- Air cooled

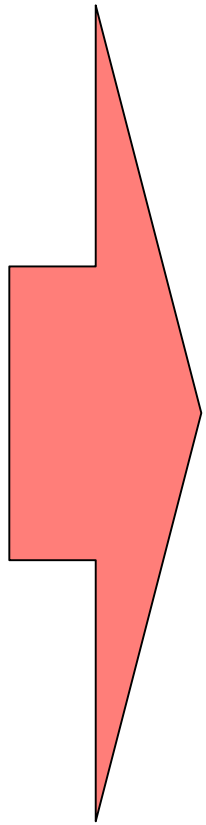
Phase 2

- ~6400 nodes, ~150,000 cores
- 1.9+ GHz AMD Opteron (Magny-Cours, 12-core)
- ~1.0 Pflop/s peak
- ~100 Tflop/s SSP
- ~200 TB DDR3 memory total
- Gemini Interconnect
- 2 PB disk, ~70 GB/s
- Liquid cooled



Feedback from NERSC Users was crucial to designing Hopper

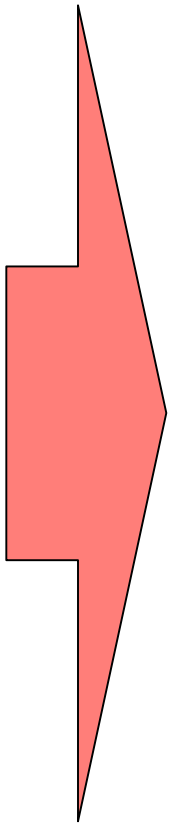
User Feedback from Franklin	*Hoppper Enhancement*
Login nodes need more memory	8 external login nodes with 128 GB of memory (with swap space)
Connect NERSC Global FileSystem to compute nodes	Global file system will be available to compute nodes
Workflow models are limited by memory on MOM (host) nodes	•Increased # and amount of memory on MOM nodes •Phase II compute nodes can be repartitioned as MOM nodes



Feedback from NERSC users was crucial to designing Hopper

User Feedback from Franklin

Improve Stability and Reliability



Hopper Enhancement

- External login nodes will allow users to login, compile and submit jobs even when computational portion of the machine is down
- External file system will allow users to access files if the compute system is unavailable and will also give administrators more flexibility during system maintenances
- For Phase 2, Gemini interconnect has redundancy and adaptive routing.

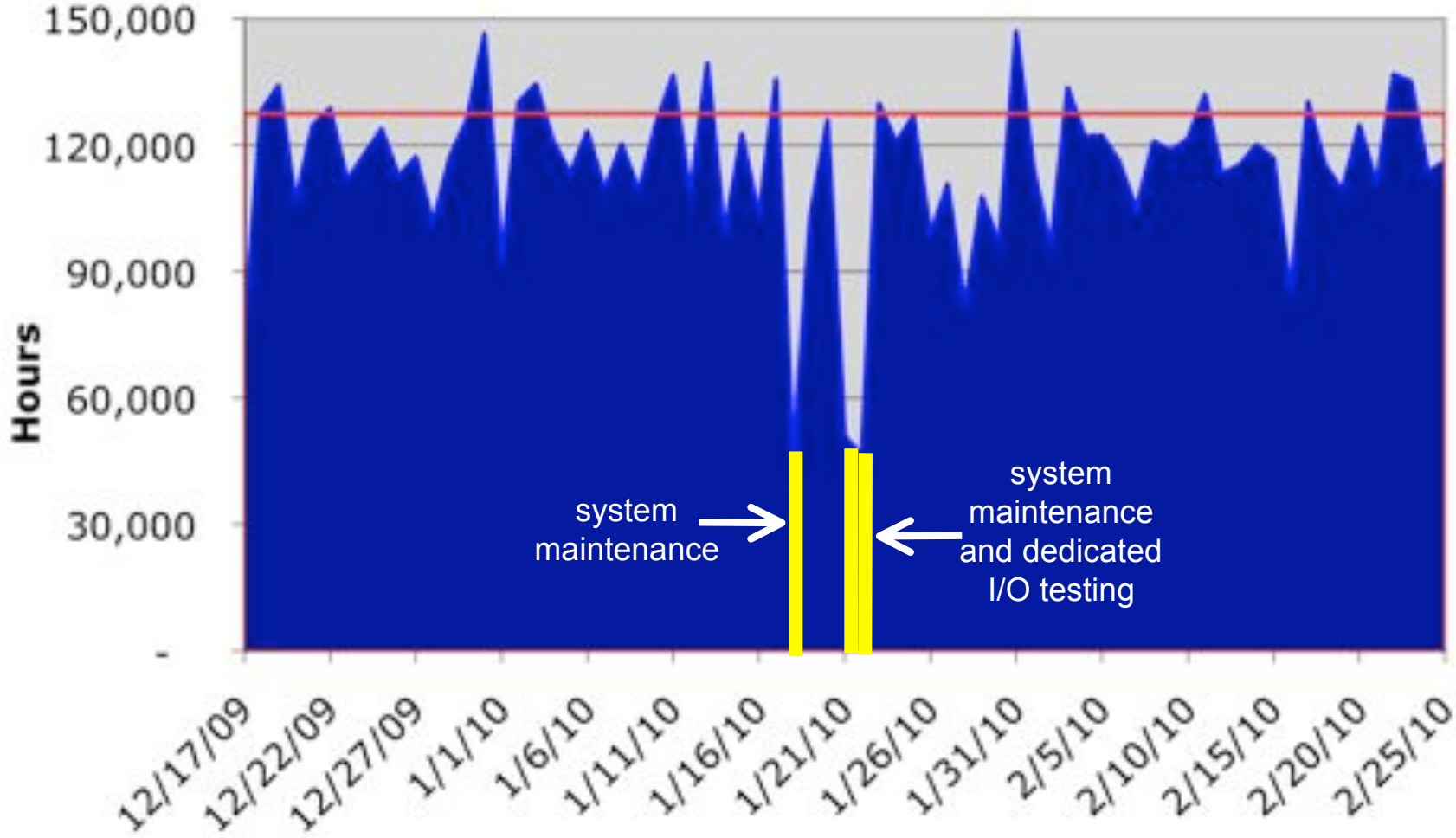
Hopper Phase 1 - Key Dates

• Phase 1 system arrives	Oct 12, 2009
• Integration complete	Nov 18, 2009
• Earliest users on system	Nov 18, 2009
• All user accounts enabled	Dec 15, 2009
• System Accepted	Feb 2, 2010
• Account charging begins	Mar 01, 2010

Hopper Installation

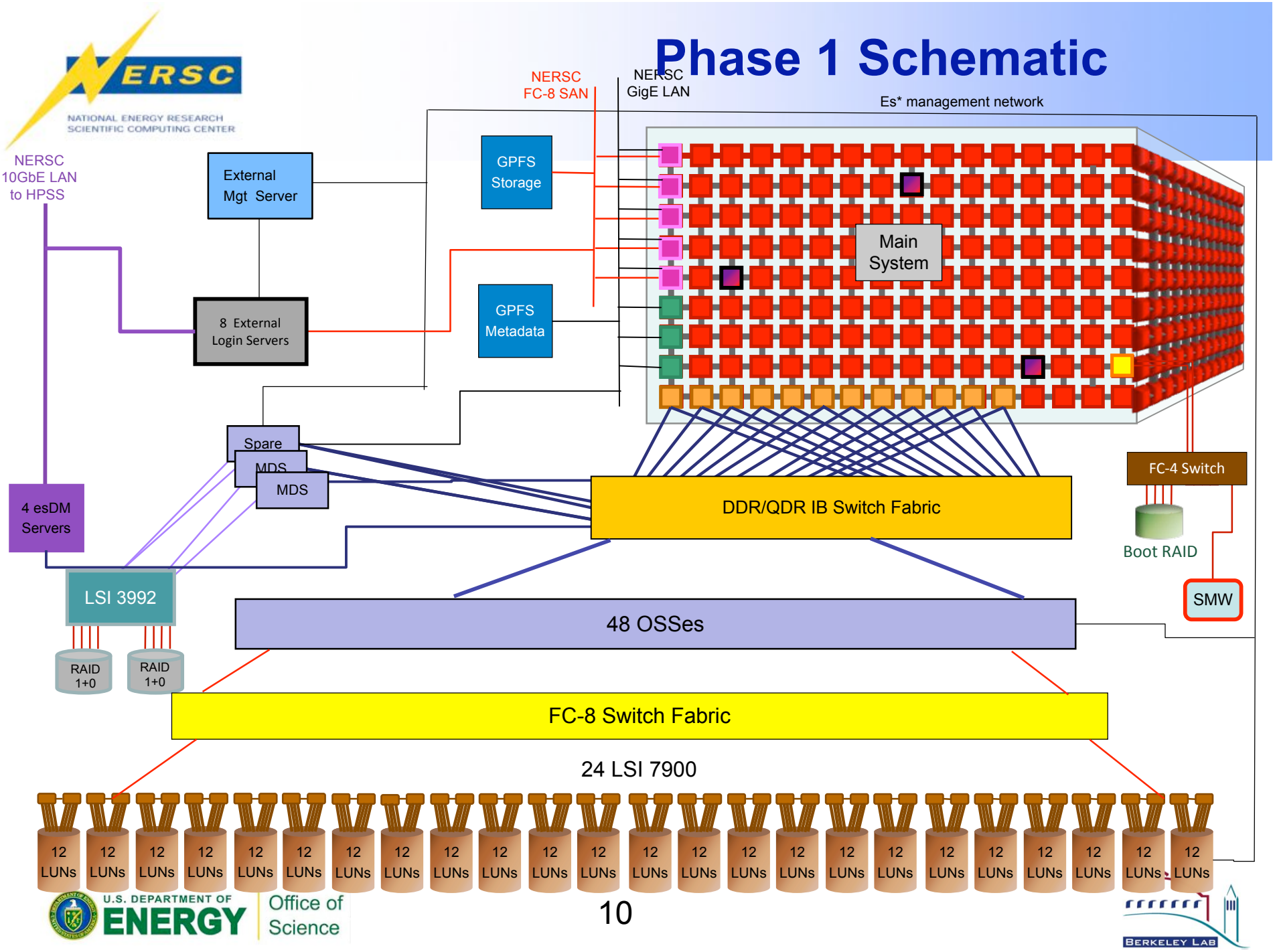
Delivery ***Unwrap*** ***Install***

Hopper Phase I Utilization



- *Users were able to immediately utilize the Hopper system*
- *Even with dedicated testing and maintenance times, Hopper utilization from Dec 15th- March 1st reached 90%*

Phase 1 Schematic



System Configuration

Nodes	Chip	Freq	Memory
664 Compute	2 x Opteron QC	2.4 GHz	16 GB
36 (10 DVS + 24 Lustre + 2 Network)	1 x Opteron DC	2.6 GHz	8 GB
4 Service	1 x Opteron DC	2.6 GHz	8 GB
12 DVS (Shared root)	2 x Opteron QC	2.4 GHz	16 GB
6 MOM	1 x Opteron DC	2.6 GHz	8 GB

ES System Configuration

Nodes	Sever	Chip	Freq	Memory
8 Login	Dell R905	4 x Opteron QC	2.4 GHz	128 GB
48 OSS + 3 MDS	Dell R805	4 x Opteron QC	2.6 GHz	16 GB
4 DM	Dell R805	4 x Opteron QC	2.6 GHz	16 GB
MS	Dell R710	4 x Xeon QC	2.67 GHz	48 GB

- 24 LSI 7900 controllers
- 120TB configured as 12 RAID6 LUNs per controller

esLogin

- **Goals**
  - Ability to run post-processing and other small applications directly on login nodes without interfering with other users
  - Faster compilations
  - Ability to access data and submit jobs if system goes down
- **Challenges**
  - New for Cray; one of first sites
  - Creating a consistent environment between external and internal nodes
  - Configuring batch environment with external login nodes
  - Provisioning and configuration management
- **Solutions**
  - *Cray packaged software updates both internal and external nodes*
  - Run local batch servers transparently
  - Configuration management software, e.g. SystemImager
- **Results**
  - Users report more responsive login nodes
  - “The login nodes are much more responsive, I haven't had any of the issues I had with Franklin in the early days.” Martin White
  - No complete cluster mgt system yet

esFS

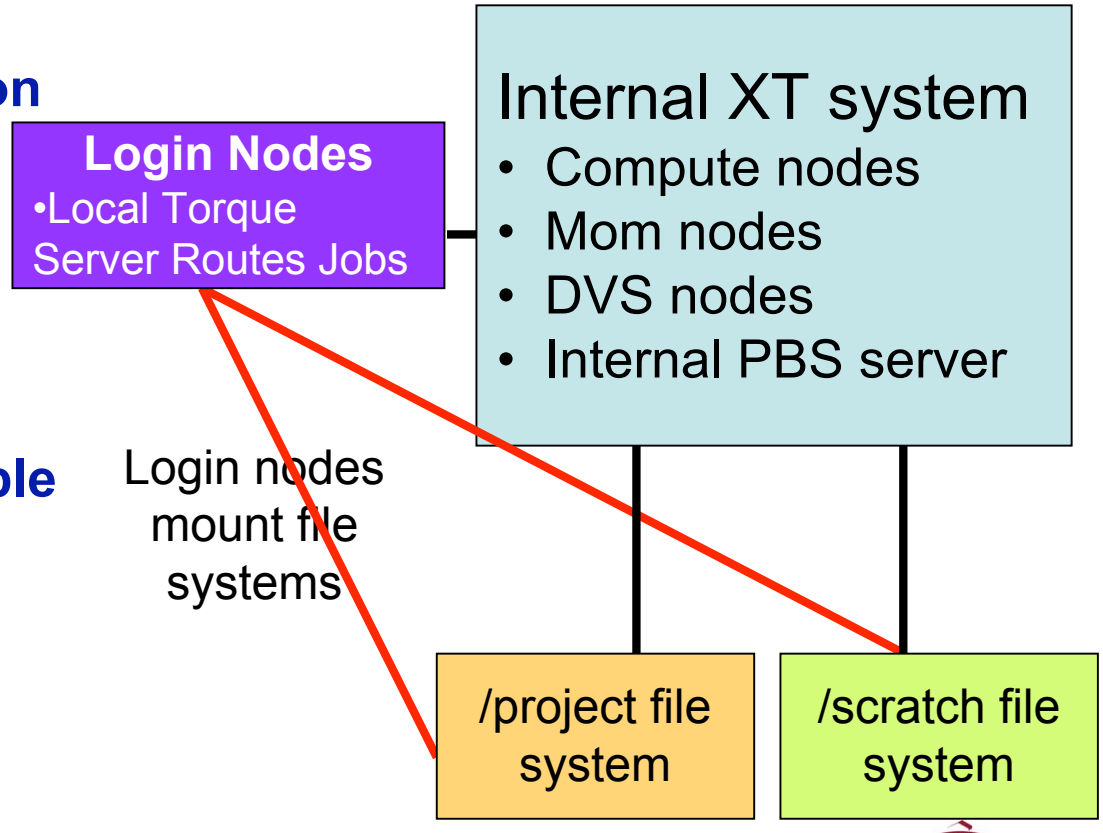
- **Goals**
  - Highly available filesystem
  - Ability to access data when system is unavailable
- **Challenges**
  - Different support model
  - Oracle-supported Lustre 1.8 GA server, Cray-supported 1.6 clients
  - Automatic failover, assuring that if one OSS or MDS fails the spare picks up
  - Provisioning and configuration management
- **Solutions**
  - With *manual* failover, servers can be updated via a rolling upgrade reducing downtime
  - Configuration management software, e.g. SystemImager
- **Results**
  - Users report a stable reliable system
  - "I have had no problems compiling etc, and my jobs have had a very high success rate." Andrew Aspen
  - No complete cluster mgt system yet
  - No automatic failover yet

esDM

- **Goals**
  - Offload traffic to/from mass storage system from login nodes
- **Challenges**
  - Consistent user interface to mass storage system
- **Solutions**
  - Client modified for third-party transfers
- **Results**
  - Expect main benefits for Phase 2
  - Porting client to internal login nodes

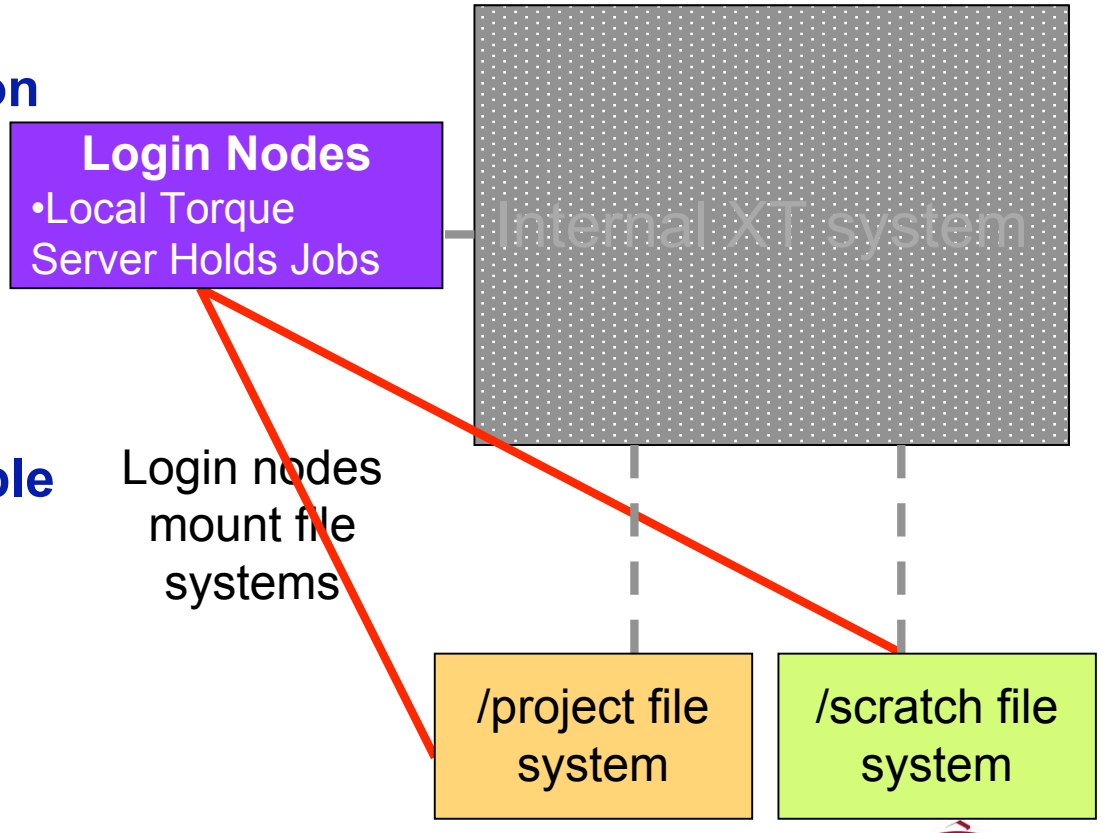
Data and Batch Access

- **Prepare and submit jobs when XT down**
  - **Compile applications and prepare input**
  - **Local Torque servers on login nodes provide routing queues**
  - **Holds jobs while XT is down**
  - **Jobs forwarded to internal XT Torque server when XT available**
  - **Batch command wrappers hide complexity of multiple servers and ensure consistent view**



Data and Batch Access

- **Prepare and submit jobs when XT down**
  - **Compile applications and prepare input**
  - **Local Torque servers on login nodes provide routing queues**
  - **Holds jobs while XT is down**
  - **Jobs forwarded to internal XT Torque server when XT available**
  - **Batch command wrappers hide complexity of multiple servers and ensure consistent view**

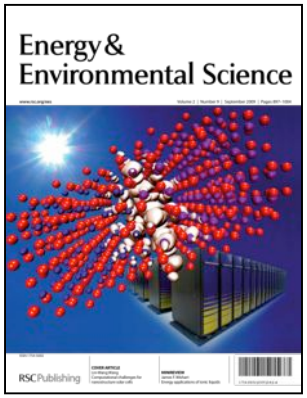


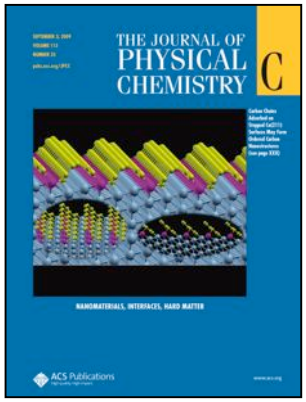
Summary

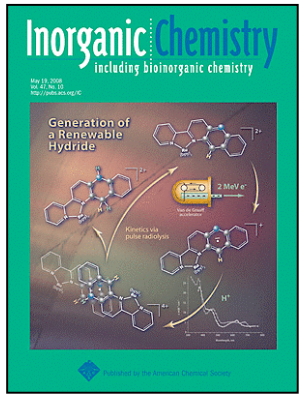
- **Benefits**
  - Improved reliability and usability
- **Challenges**
  - Not a standardized offering
    - One-of-a-kind systems by Custom Engineering
    - Software levels different from Cray products
  - Synchronization & Consistency
    - Lack of complete cluster management system
    - Software packaging
- **Recommendations**
  - *A product based on external services*

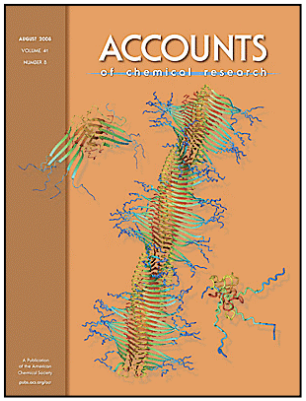
NATIONAL ENERGY RESEARCH SCIENTIFIC COMPUTING CENTER

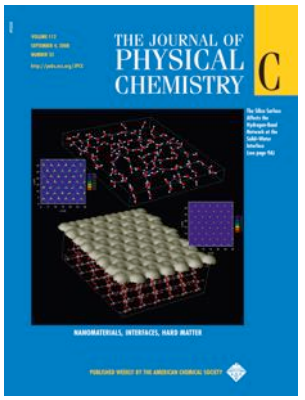
Enabling New Science

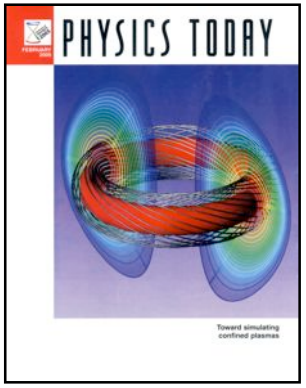












This work was supported by the Director, Office of Science, Office of Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No. DE- AC02-05CH11231.

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB