



System Integration Experience Across the Cray Product Line

Presented to the Cray User Group
Albuquerque, NM
May 18, 2005

Arthur Bland, Richard Alexander, Steven
Carter, Don Maxwell, Robert Silvia –
Oak Ridge National Laboratory
Nina Hathaway – University of Tennessee
Jeff Becklehimer, Cathy Willis – Cray, Inc.



This work was sponsored by the U.S. Department of Energy's Office of Advanced Scientific Computing Research and performed at the Oak Ridge National Laboratory, managed by UT-Battelle, LLC under contract number DE-AC05-00OR22725.

Outline

- Overview of National Center for Computational Sciences (NCCS)
- Cray X1 experiences
- Cray XT3 experiences
- Cray XD1 experiences

Overview of NCCS

- Founded in 1992 to evaluate parallel systems for DOE and determine how science applications work on these systems
- Systems include:
 - Intel Paragon XP/S 5, 35, and 150MP
 - IBM Power3, Power3+, Power4, Power4+
 - Compaq AlphaServer SC
 - Cray X1, XT3, XD1

NCCS Evaluation Systems



Paragon XP/S 150 MP - 1995



IBM Power3 - 1999



Paragon XP/S 35 - 1992



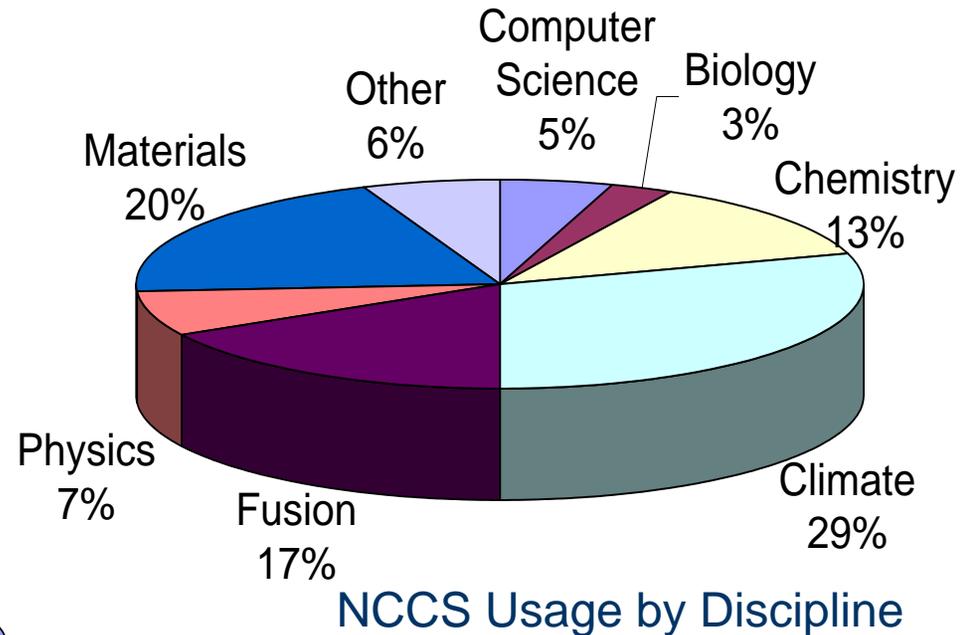
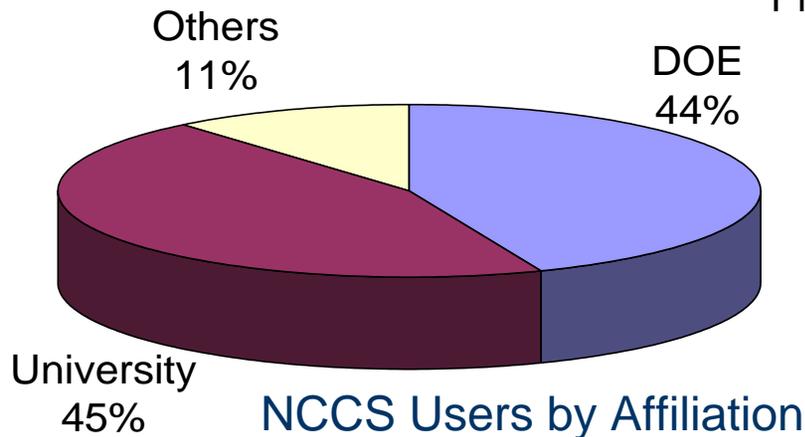
IBM Power4 - 2002



Compaq AlphaServer - 2000

NCCS Users

- Users are equally split between universities and DOE
- Utilization across many different science disciplines



NCCS File Systems

Three level structure:

- Each computer has its own high-speed local scratch file system for running jobs. These files are subject to purging.
- A single home directory file system is mounted on all of the computers for low-bandwidth needs. Today, this is NFS. We are working on changing this to Lustre.
- Very large, high-bandwidth archival storage using HPSS. Today we have 560 TB on tape.

lustre

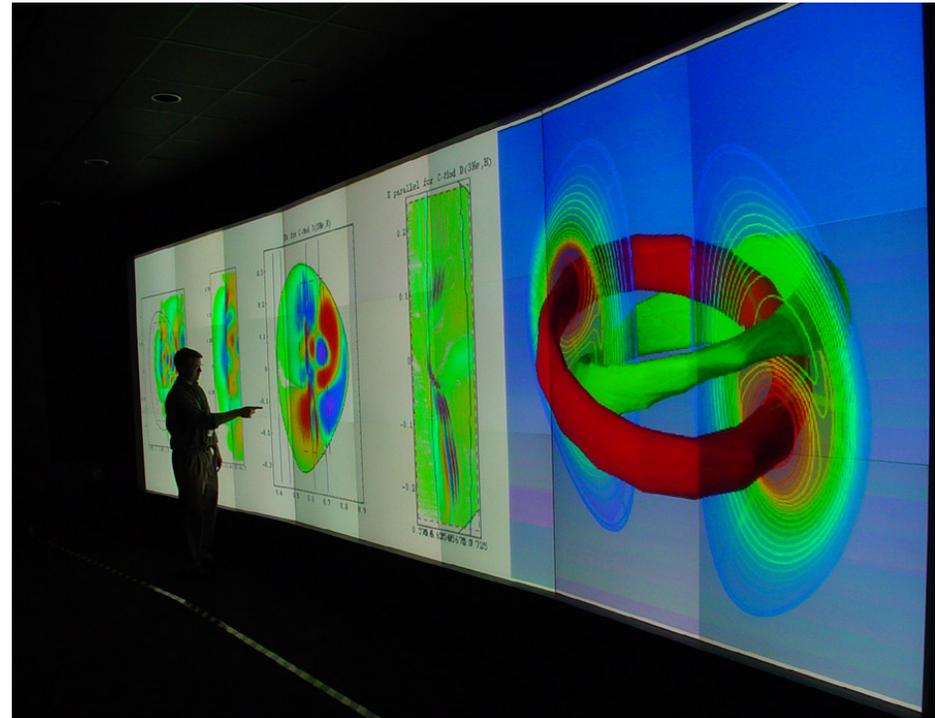


Networks

- Internal networks are GigE, upgrading to 10GigE this year.
- External connections to:
 - ESnet – OC192
 - Internet2 – OC192
 - National Lambda Rail – 2 x OC192
 - UltraScienceNet Testbed – up to 16 x OC192

Visualization

- 35 megapixel power wall – 30' x 8' – for visualizing large datasets
- Cluster based visualization tools
- Supporting both local and remote users





Facility

- 40,000 ft² raised floor computer room
- 4 megawatts of power for computers + 4 MW for utilities. Expandable to 12 MW
- 500 KVA UPS – Networks, disks, servers, but not the supercomputers.
- 3,600 tons of cooling, expandable to 5,100 tons
- Office space for 500 staff



Outline

- Overview of National Center for Computational Sciences (NCCS)
- Cray X1 experiences
- Cray XT3 experiences
- Cray XD1 experiences

Disclaimer

At this type of conference, we talk a lot about what has gone right and what has gone wrong. Lots of people have talked about what is going well from a systems and science perspective. I will talk more about what has not gone so well. In doing so, I want it to be understood that the purpose of this is to help other sites learn from our experiences, not to criticize Cray or others.

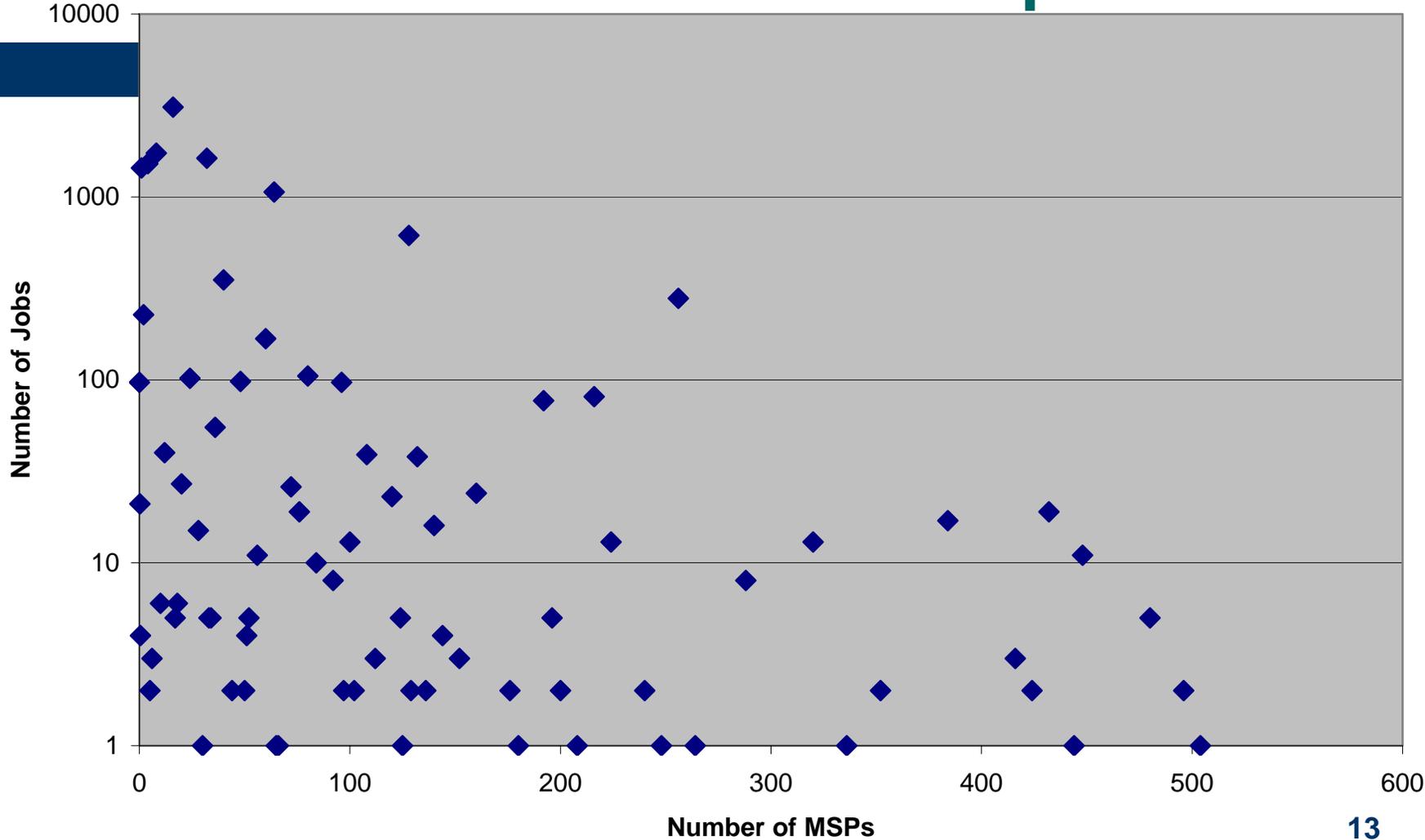
Cray X1 - Phoenix

- Largest X1 in the world today at 512 MSPs (6.4 TF)
- 4 GB of memory per MSP (2 TB total)
- 32 TB of directly attached disks

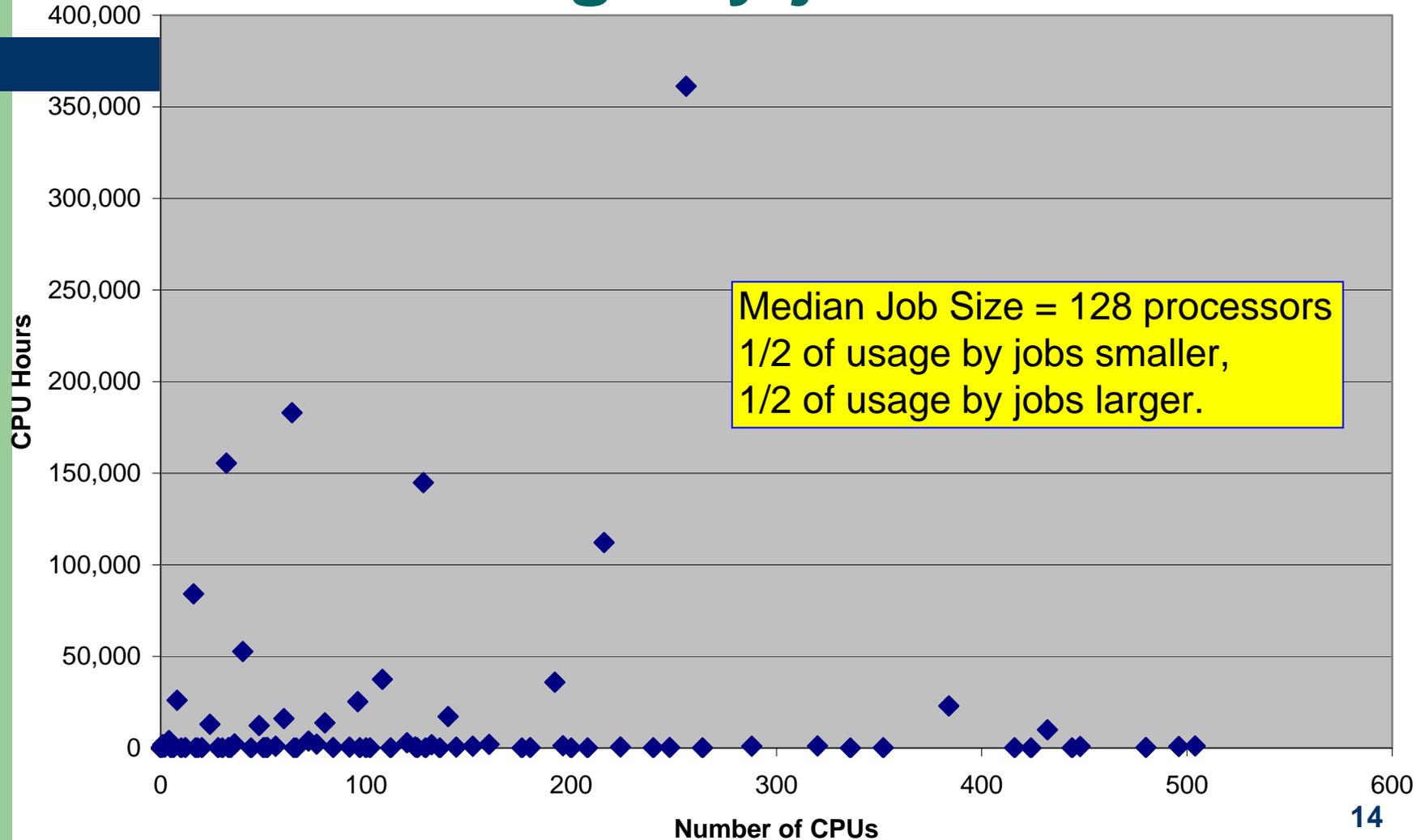


- System will be upgraded to X1E over the summer
- 1,024 MSPs
- 18.5 TF

Phoenix Job Mix – Nov '04 – Apr '05

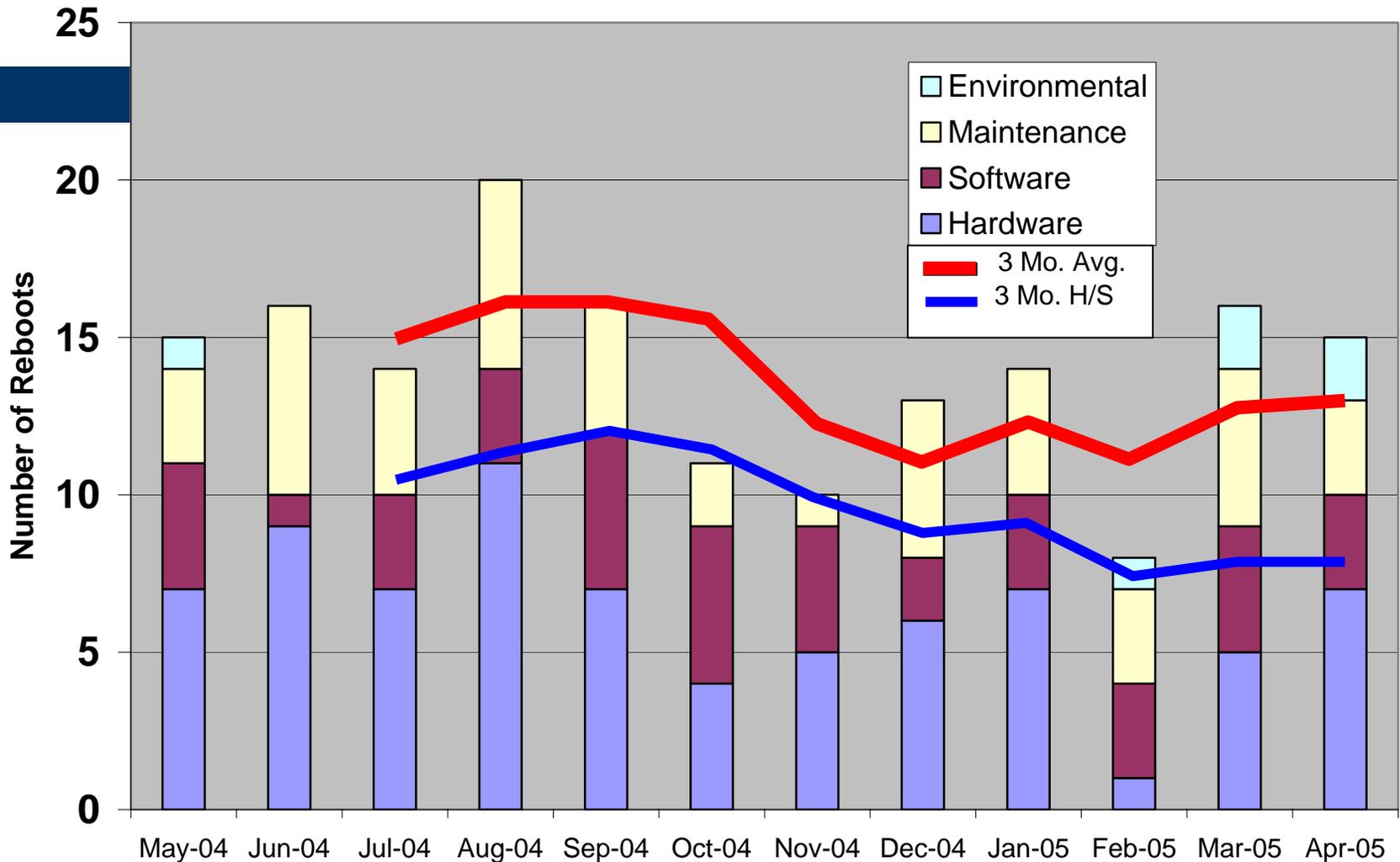


Phoenix – Usage by job size



Cray X1 - Reboots

Cray X1 Reboots by Month and Cause With Three Month Running Averages



Phoenix – Last three months

- 55.4 hour MTBF from all causes
- 93.4 hour MTBF from Hardware and Software (excludes maintenance and environmental problems)
- Target is to be greater than 168 hours (less than one reboot per week)
- Currently averaging 1.7 reboots per month due to water quality problems at ORNL

Phoenix – What are the problems?

- Hardware – almost all recent problems have been memory DIMM failures
- Software – Variety of problems. Trend is flat, not improving
- Maintenance – Weekly preventive Maint.
- Environmental – Poor water quality. Installing filters to clean this up.

Outline

- Overview of National Center for Computational Sciences (NCCS)
- Cray X1 experiences
- Cray XT3 experiences
- Cray XD1 experiences

Cray XT3 - Jaguar

	Cabinets	Processors	Memory	TFLOPS
Today	40	3,748	7.5 TB	18
June '05	56	5,212	10.5 TB	25.1
2006	120	22,000+	44 TB	100+



Jaguar

- Power – Using between 12 and 15 KW per cabinet today (2.4 GHz processors + 2 GB per processor)
- Cooling – These cabinets take a LOT of air. You need to have both cooling capacity and airflow for 3,000 CFM per cabinet
- Hardware is remarkably stable (14 days without a hardware problem)

Jaguar Software

- ORNL is rebooting 2-10 times per day, depending on what bugs the users are hitting
- Performance of any codes that are latency sensitive is terrible.
- Bandwidth is great and codes that can tolerate the large latency are getting excellent performance.
- CAM overflow problem causes programs to hang if a node has too many outstanding messages. This limits the scalability of some important codes.

Jaguar - Software

- Portals must be completed with full functionality and performance before the XT3 will be a production ready system. This impacts both stability and performance of the system.
- Diagnostics are essentially non-existent. The on-site engineers have worked out trial-an-error methods for isolating problems, but it takes too long and doesn't really work well. It appears that the RAS system is collecting lots of useful information, but there is no way to get it out of the system, yet.
(Documentation?)

Jaguar - Lustre

- We have 16 service nodes connected to DDN RAID disks today. Lustre installation is happening as we speak.
- We have tested Lustre on a 1-cabinet test XT3 and are getting promising results. Performance is scaling well with the number of OSTs. We are told that as the number increases, that stability will suffer. Again, a portals problem.
- We are seeing a problem with Fibre Channel cards running at 1 Gb/s instead of 2 Gb/s. We are told this is a configuration problem with the cards coming from the supplier. We can put the cards in a PC and switch them, or wait until Cray provides a tool to do it under Linux.
- CFS is focusing all of its efforts on a version of Lustre that requires the Linux 2.6 Kernel (for extents). Cray's current plan is to start supporting 2.6 in September. That makes getting CFS to focus on Lustre on 2.4 problematic.

Jaguar - Summary

- The hardware performance and stability are very good, even excellent.
- Software is still Alpha quality. Cray has a lot of work to make the software stable.
- Portals is THE key technology for the XT3 and it is just not ready, yet.
- Lustre looks to be a good choice for a file system, but is yet another new technology in this product.
- I expect that the XT3 will be an exceptional system for capability applications when these problems are resolved.

Outline

- Overview of National Center for Computational Sciences (NCCS)
- Cray X1 experiences
- Cray XT3 experiences
- Cray XD1 experiences

Cray XD1 - Tiger

- 12 Chassis
- 144 processors (2.2 GHz)
- 6 FPGA Application accelerators
- 4 GB per processor



Tiger - Summary

- The system hardware has been very stable.
- Software is in pretty good shape. We have just installed the software that allows us to combine all 12 chassis into a single system.
- Physical installation is easy. Essentially no problems
- There were early problems with ActiveManager, but these are largely resolved.
- Excellent performance from the system. Bandwidth, latency, and stability are all exceptional.
- Compilers are the weak point. We find both PGI and Pathscale compilers to be better at some problems, and worse at others.

Questions?

