Developing a Powerful yet Inexpensive Computational Infrastructure for the UT Dept. of Nuclear Engineering

David D. Dixon

April 8, 2009

Overview

- Status of Existing Computational Infrastructure
- General Design Approach
- Description of System Hardware and Software
- Auxiliary Experiments
- Future Opportunities

Status of Existing Infrastructure

- Single 20-CPU cluster
- Xeon 2.4 GHz processors (130nm, 65W TDP)
- ~73GB disk space
- Red Hat Linux, 2.4 series kernel
- 40GB hard drive in each node

General Design Approach

- Use commodity desktop hardware
- Utilize diskless nodes
- Freely available, open source software where possible
- Focus on energy efficiency, low construction cost, speed
- Consideration given to application software of interest at time of construction

Design Considerations

- Limited power and cooling available
  - Test/identify energy efficient processors
- Limited space
  - Look to high-density configurations (e.g. multi-core CPUs, alternative cooling)
- Less than $20k to spend
  - Desktop hardware instead of server hardware

CPU Evaluation

- Kill-A-Watt used to measure power consumption, power factor

- Testing at idle and at full load

- Where possible, the same motherboard was used for multiple CPUs

- 2 MCNP test cases used – one ABTR full-core k-code model, one ORNL critical benchmark experiment

Intel Xeon 2.4 GHz

- 2.4 GHz
- Single Core
- 130nm
- 65W TDP

- $235 each when purchased
- 2/$10 on eBay today

AMD Athlon 64 X2 5200+

- 2.7 GHz
- Dual-Core
- 65nm
- 90W TDP

- $130 each when purchased (1/08)
- ~$55 today

AMD Phenom X4 9850

- 2.5 GHz
- Quad-core
- 65nm
- 125W TDP

- Released 3/08 at $235
- ~$160 today

Intel Core 2 Quad Q6700

- 2.66 GHz
- Quad-core
- 65nm
- 95W TDP
- Released 4/07 at $533
- ~$275 in May 2008
- ~$215 today

Intel Core 2 Quad Q9300

- 2.5 GHz
- Quad-core
- 45nm
- 95W TDP
- Released 3/08 at $266
- ~$285 in May 2008
- ~$240 today

Intel Core 2 Quad Q8200

- 2.33 GHz
- Quad-core
- 45nm
- 95W TDP

- Released 8/2008 at $224
- ~$165 today

Performance Summary

Processor	Run Time	Idle Power (W)	Full Load Power (W)	Energy Cons. (Wh)	CPU Cost
AMD Athlon X2 5200+	18:09	74	145	43.86	$54.99
AMD Phenom X4 9850	07:12		279	33.46	$154.99
Intel Core 2 Quad Q6700	06:28	63	127	13.69	$214.99
Intel Core 2 Quad Q9300	05:24	45	97	8.73	$239.99
Intel Core 2 Quad Q8200	05:48	45	91	8.8	$165.99

Overclocking Summary

Processor	Run Time	Idle Power (W)	Full Load Power (W)	Energy Cons. (Wh)
Intel Core 2 Quad Q6700 @ 2.66GHz	06:28	63	127	13.69
@ 3.06GHz	06:05	64	138	13.99
@ 3.15GHz	05:48	63	140	13.53
@ 3.25GHz	05:35	64	143	13.31
Intel Core 2 Quad Q9300 @2.5GHz	05:24	45	97	8.73
@3.00GHz	04:30	47	101	7.58
@3.15GHz	04:17	47	102	7.28
@3.00GHz w/Larger Heat Sink	04:30	47	93	6.98

Dual-Core Cooling

- Stock aluminum heat sink (top) unable to manage heat generated in 2U rack space

- Replaced with copper heat-pipe heat sink (bottom)
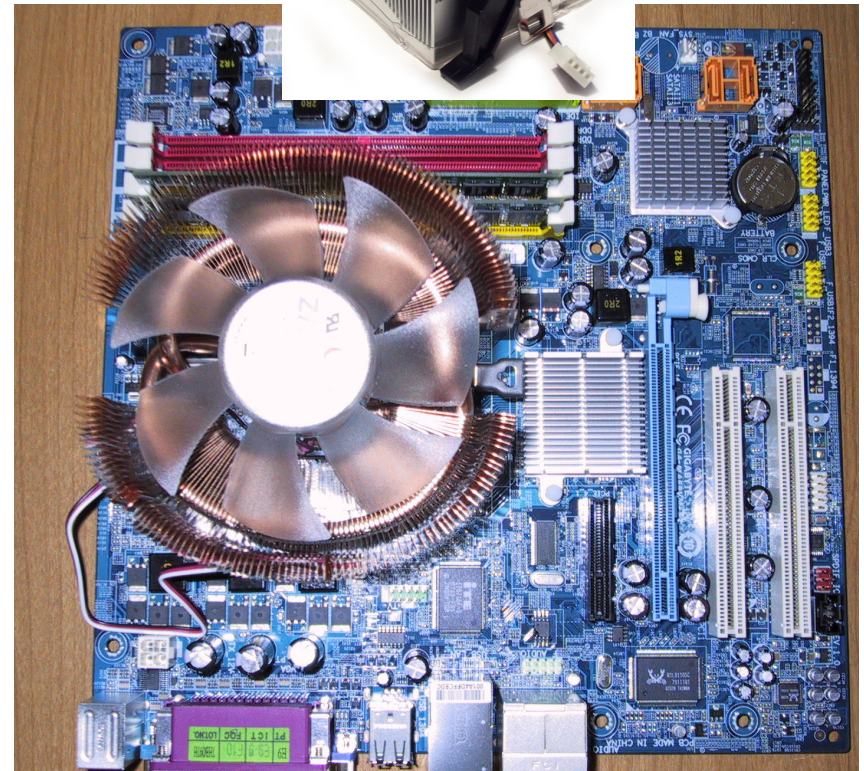
Phenom Cooling

- Stock heat sink too tall for 2U rack space with neighboring nodes

- No suitable substitute air-cooled substitute for high density nodes
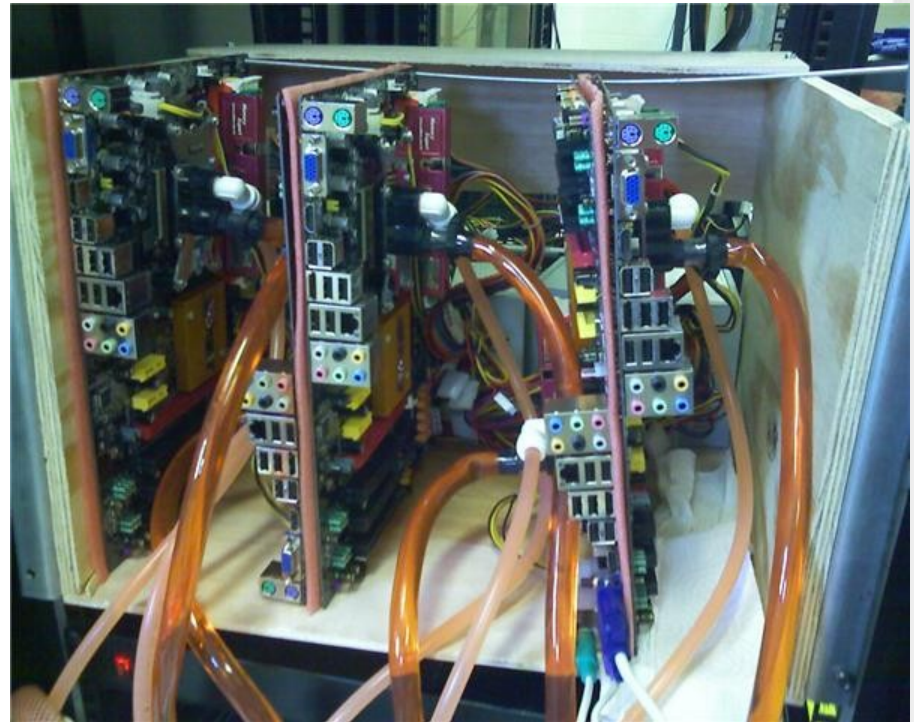
- Water cooling as alternative
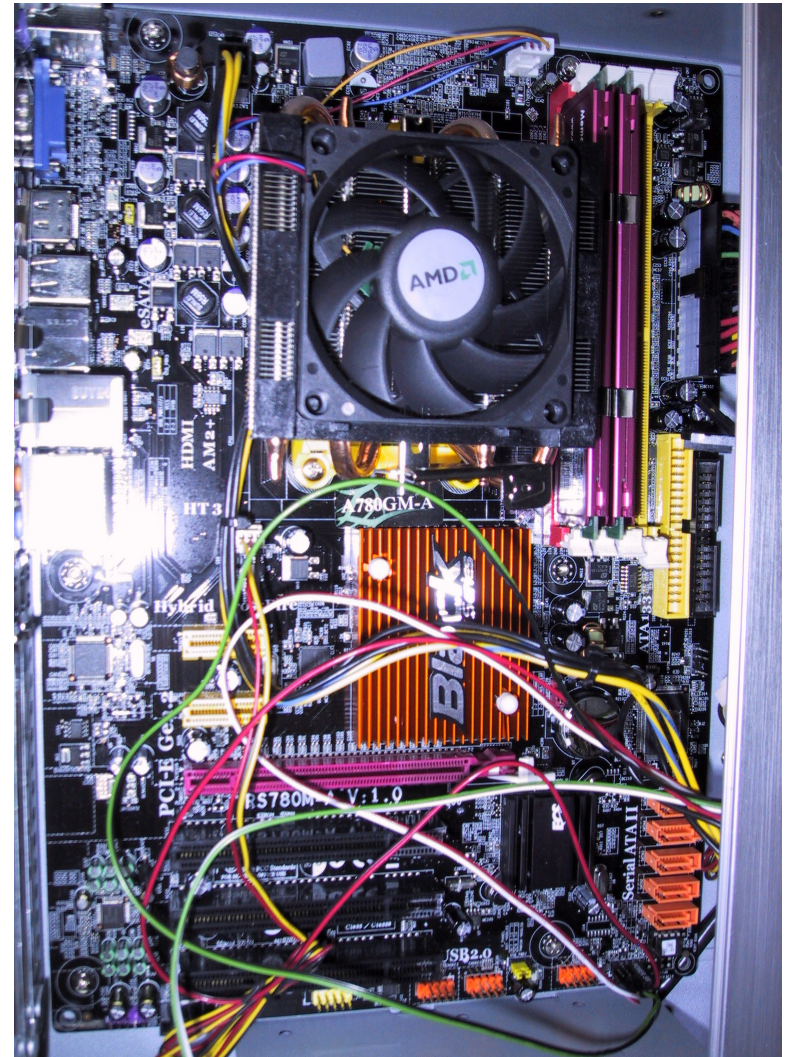
Phenom Cooling

- Heat-pipe heat sink replaced with water block (right)
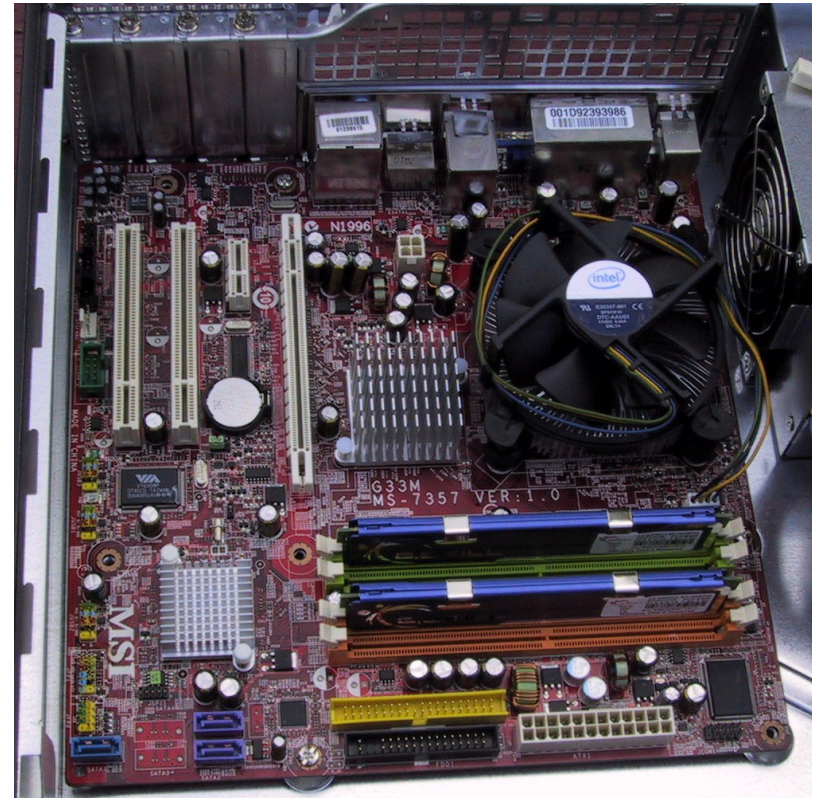- Connected to pumped loop rejecting heat via large radiator

Phenom Cooling

- 5 full-size ATX motherboards operated with water cooling in 10U rack space

Water Cooling Decommissioned

- Availability of 45-nm Intel processors made system unnecessary

- Power consumption of Intel 45-nm quad-core less than AMD 65-nm dual-core

Smaller Heat Sink with Smaller Process

- Reduced TDP in moving from 65-nm to 45-nm process led to smaller included heat sink

- Larger heat sink on 45-nm CPU saves 8W

- Intel's 8088 used a 3,000 nm process, original Pentium was 800 nm
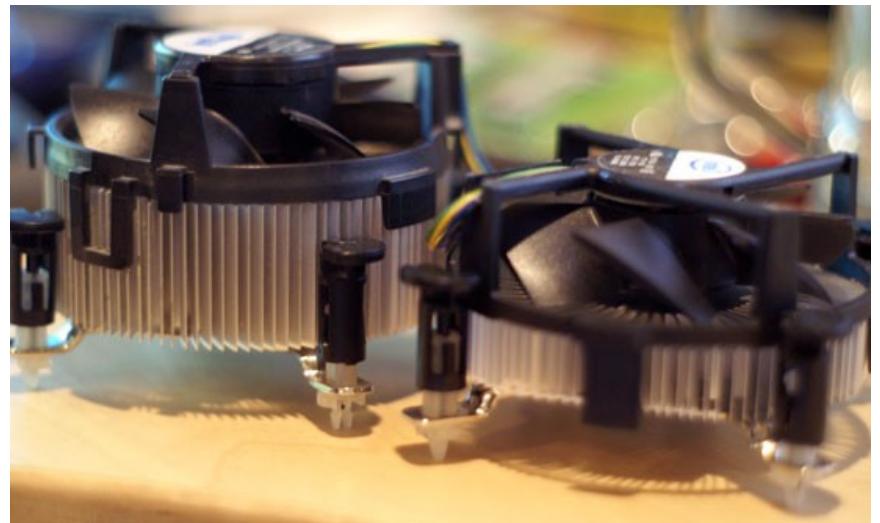
New Cluster

- 10 dual-core AMD CPUs @ 2.7 GHz

- 5 quad-core AMD CPUs @ 2.5 GHz

- 4 quad-core Intel CPUs @2.7 GHz

- 2TB disk space

- Ubuntu 8.04 LTS Server Edition, 2.6 series kernel

Software Overview

- Ubuntu 8.04 LTS Server Edition
- Diskless Remote Boot Linux (DRBL)
- Ganglia
- Webmin
- Application Software
  - CoreMaster 2, SCALE, MCNP, MCNPX, MonteBurns, POLCA, PHOENIX, NESTLE, Space07 Pointwise-Continuous Cross sections

Ubuntu/DRBL

- Ubuntu is a Debian-based Linux distribution maintained by Canonical
  - Focus is on out-of-the-box hardware support, package manager
- Diskless Remote Boot Linux (DRBL) provides boot images and management of cluster compute nodes

Modifications to DRBL Script

- Addition of modules for hardware sensor data and CPU frequency scaling on compute nodes

- Implementation of on-demand CPU frequency scaling governor
 - This reduces clock speed at idle to as low as 1 GHz, reducing idle power consumption

Ganglia

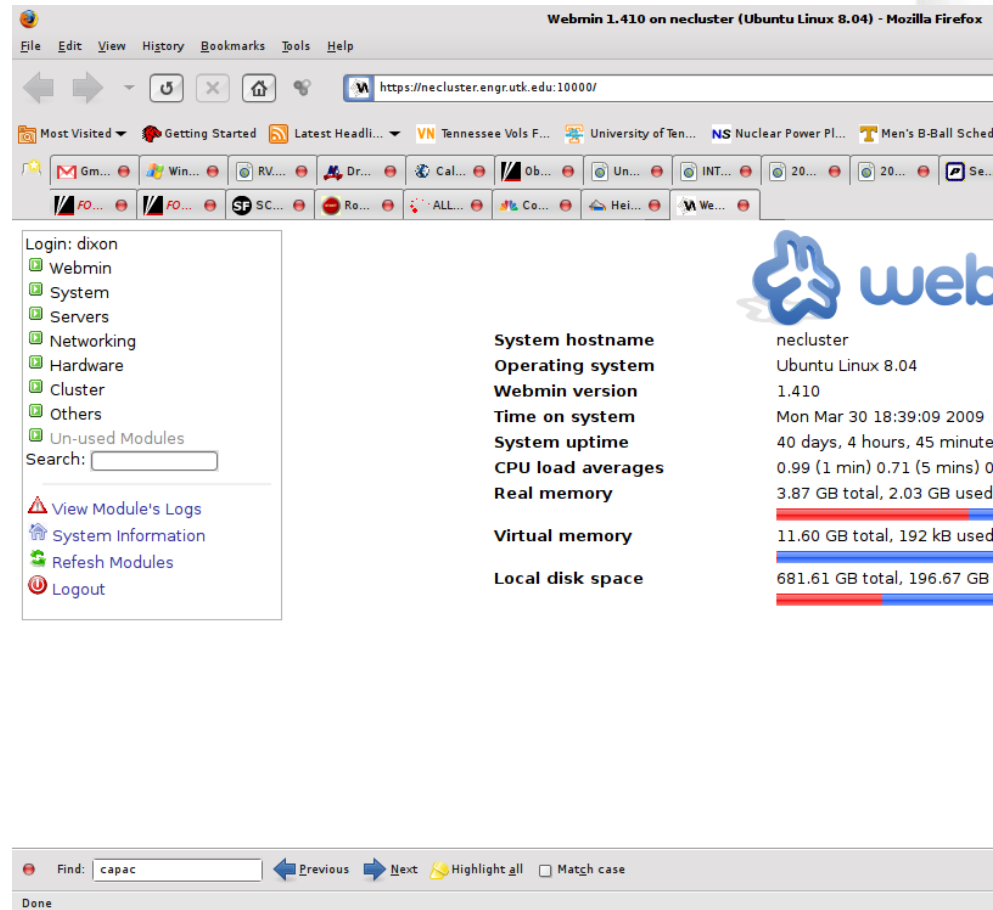- Ganglia provides a web-based view of cluster status/health
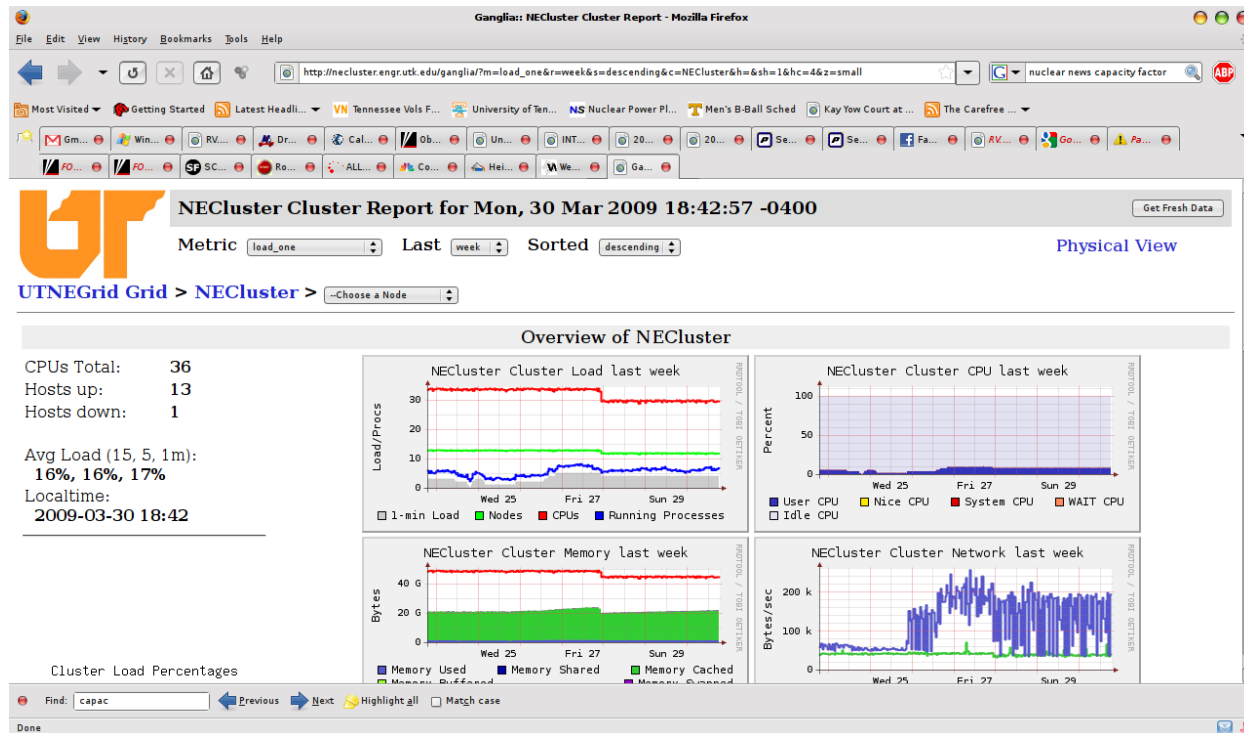
Ganglia Mods

- Cron job to access motherboard's sensor data (fan speeds, temperatures, voltages) sent to Ganglia

  - Randomization needed to avoid network collisions

- System-average plots generated

- User metrics added

NECluster Cluster Temperatures last week
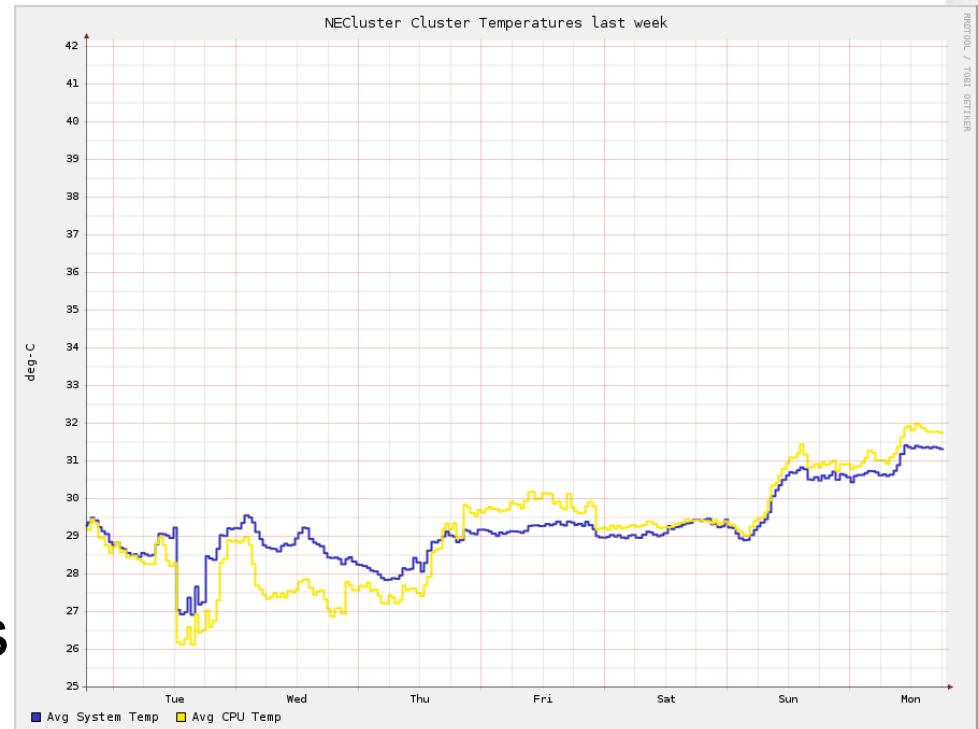
Webmin

- Web-based system administration tool

- Allows non-Linux savvy admin to perform user management tasks

Application Software

- MCNP, MCNPX, MonteBurns
- VirtualBox
- Matlab
- CoreMaster 2, PHOENIX, POLCA
- NESTLE, SCALE
- ANSYS

Development Tools

- Intel, Portland Group, and GNU Fortran Compilers

- C++, Perl, PHP, MySQL, Java, Apache2, Tomcat

- MPICH, LAM MPI

Future Development Opportunities

- MPI-compatible job scheduler

  – Include logic to power nodes on/off as needed

- Larger heat sinks to reduce energy use

- Higher density, mini-ITX (6" square) motherboards

  – May be able to mount as many as 6 quad-core motherboards in a single case

- Implement independent RAID network attached storage (NAS) device