Annual Users' Meeting

TAMU Supercomputing Facility <u>http://sc.tamu.edu</u>

May 5, 2005

Outline

Formal part

- Overview of Facility
 - Steering Committee (members, functions)
 - Budget summary (sources, expenses, worries)
- Overview of operations
 - Help desk
 - Short courses
 - Machines (Age, Strengths/Weaknesses, etc.)
 - Issues for discussion
- Round-Table of Cluster/Grid Computing
- Break: Pick up Pizza
- Informal part
 - General Discussion
 - "Exit survey"

Texas A&M Supercomputing Facility

"The Texas A&M Supercomputing Facility is committed to providing advanced computing resources and support to Texas A&M faculty and students engaged in all aspects of large-scale computation and to advancing the expertise and use of high-end computation at TAMU."

- Foci: Computational Support AND Computational Science
- Models:
 - □ Scale: Intermediate between PI's means and National Center capacity
 - □ Access: "University Library" rather than. "University Book Store"

Issues for Discussion

Budgeting \$\$\$ and Staff Time

- Maintenance vs.. Acquisition (Hardware and Software)
- Retiring old equipment"
- Balancing User Needs (there are conflicts!)
 - □ Turn-around time vs idle cycles
 - Availability of dedicated time
- Cycle Shop" vs. ``HPC Center"
- Next Step(s)
 - Acquisitions: (Shared memory vs Cluster)
 - Campus Grid efforts (TIGRE)
- (Other Issues Raised by Users)

High Performance Computing Steering Committee

- Faculty Steering Committee (meets biweekly)
 - Sets policy on resource use
 - □ Reviews computer time requests for "large research" grants
 - Plans hardware acquisition
 - Advises Provost regarding the University's plans and guidelines for High Performance Computing
- Membership
 - Lee Panetta, Chair, Atmospheric Sciences
 - □ Mike Hall, Chemistry
 - Dave Adelson, Animal Science
 - Paul Roschke, Civil Engineering
 - Lawrence Rauchwerger, Computer Science
 - Pierce Cantrell, Associate Provost for IT
 - □ Spiros Vellas, Assoc Director CIS for Supercomputing
 - Tom Putnam, Director CIS

Supercomputing Staff

- Associate Director
 - Spiros Vellas
- Analyst Staff
 - Keith Jackson
 - Zdenko Tomasic
 - Michael Thomadakis
 - Francis Dang
 - Faisal Chaudhry
- Administration
 - Greta Thompson

- Helpdesk Staff (GANTS)
 - Nageswar Rao
 - Santosh Srinivasan
 - Jayesh Krishna

Total Funds



Operating Funds



Capital Funds



Operating Expenses



Top 500 Comparison



Supercomputing Services

- Helpdesk
 - □ M-F, 9:00 am to 5:00 pm
 - Staffed by three rotating GANT's
 - □ All full-time staff also participate
 - Services for users include help with:
 - code parallelization (MPI and OpenMP)
 - debugging
 - math library selection and call interfacing
 - script writing (especially for batch)
 - I/O and scalar code optimization
 - code porting
 - software installation
 - account activation/management



Supercomputing Services

- Short-courses
 - Each semester 6 to 8 courses are taught
 - □ Spring 2005 courses
 - Intro to Code Parallelization Using MPI (4 x 2 hrs)
 - Intro to Shared Memory Parallelization Using OpenMP (3 x 2 hrs)
 - Intro to Scalar Code Optimization (3 x 2 hrs)
 - Computer Organization for Scientific Programmers (3 x 2 hrs)
 - Intro and Intermediate Unix (4 x 2 hrs)
 - Intro to SGI Altix (2 x 2 hrs)
 - Intro to IBM Regatta (2 x 2 hrs)
 - Other courses taught in the past
 - Matlab Graphics
 - Scientific Visualization Using Ensight
 - Code Debugging and Common Errors

Supercomputing Services

- Website Maintenance (<u>http://sc.tamu.edu</u>)
 - □ Technical Documentation (user guides, etc.)
 - Online Account Applications
 - Policies
 - Pertinent Announcements for Users
 - Other

Supercomputing Systems

Component	SGI Altix 3700 (cosmos)	IBM Regatta p690 (agave)	SGI Origin 3800 (k2)	SGI Origin 2000 (titan)
Processors	128 Intel Itanium @ 1300 MHz	32 Power4 @ 1300 MHz	64 R14000 @ 500 MHz	32 R10000 @ 250 MHz
Peak Performance	665.6 GFLOPS	166 GFLOPS	64 GFLOPS	16 GFLOPS
Memory	256 GB	64 GB	64 GB	8 GB
Disk Space	4 TB	1,500 GB	1,150 GB	300 GB
Networking	Gigabit Ethernet	Gigabit Ethernet	Gigabit Ethernet	Gigabit Ethernet
Installation Date	February 2004	July 2002	June 2001	July 1998
Operating System	Linux	AIX	IRIX	IRIX

Supercomputing Systems: Issues

- Cost of maintenance on new machines
- Staff time requirements with multiple OS's
- How to treat aging systems

Supercomputing Systems (continued)

Archive Tape Library

- 28 TB STK L180
- Summer 2004
- Replaces 1996 vintage EMASS
- Issues, policies...

Supercomputing Accounts

Type of Applicant	Need < 100 CPU Hrs?	Need up to 500 CPU Hrs?	Need > 500 CPU Hrs?
Students	Basic		
Post-Docs, Research Associates, Research Scientists	Basic	Small Research	
Faculty	Basic	Small Research	Large Research

- Students can obtain "Basic" account for research or class
 - **30** CPU hours on the SGI Altix or the IBM Regatta
 - □ 50 CPU hours on the SGI Origin 3800 or 2000
- Currently, the facility maintains:
 - □ 337 basic accounts
 - □ 79 small research accounts
 - □ 28 large research accounts

Peak Load Per System

(Feb 1 - Apr 30 2005)



Use of the High CPU Queues (Cosmos - FY05)



Usage of Queues on Cosmos (FY05)



Average Queue Wait Times (Cosmos - FY05)



Xblast Bioinformatics Cluster

David Adelson Associate Professor, Department of Animal Science

Xblast.tamu.edu

First of its kind

- 20 dual processor xserves + one head node
- Mac OS X (UNIX) command line
- Queue system (Sun Grid Engine)
- User friendly web interface for "command line phobics"
- Many useful applications
- Custom BLAST dbs available on request

Xblast Accounts by Department



Jobs Submitted By Xblast Users



Consortia & Associations

Lee Panetta Chair, HPC Steering Committee

HiPCAT – High Performance Computing

Across Texas

- Consortium of Texas institutions that used advanced computational techniques to enhance research, development, and education.
- Technologies include
 - High Performance Computing
 - Computational Grids
 - Clusters
 - □ Scientific Visualization
 - Large Storage
- Founding Members
 - □ Texas A&M Lee Panetta is TAMU representative
 - UT-Austin
 - Rice
 - University of Houston
 - Texas Tech
- TIGRE Texas Internet Grid for Research and Education
 - Build a computational grid that integrates computing systems, storage, visualization labs, displays, and sensor across Texas
 - **2** \$2.5 million over two years from Texas Enterprise Fund
 - Current plan is \$500K to each of the five founding members to hire two grid experts for two years

High Performance Computing Across Texas (HiPCAT) TIGRE Project

- Texas Internet Grid for Research & Education (TIGRE)
- Vision: develop Higher Education grid that benefits all of Texas
 - Leading-edge R&D
 - Collaboration and resource sharing
 - Competing for federal cyberinfrastructure dollars
 - □ Training for students
 - Platform for R&D partnership with industry and government

Requires:

- High bandwidth network
- Grid computing software deployment and configuration

HiPCAT's TIGRE Project

- Build-out conducted by the five HiPCAT universities with grid computing expertise, projects, and staff:
 - UT Austin: TACC
 - Texas Tech: HPC Center
 - Texas A&M: HPC Facility
 - □ Rice: Computing & Information Technologies Institute
 - Houston: Texas Learning & Computation Center
- Goal is to build a computational grid that integrates
 - Computing systems
 - □ Storage systems and databases
 - Visualization laboratories and displays
 - Instruments and sensors
- TIGRE funding will support
 - Salary for two senior grid software developers at each of the five schools for two years
 - Deliverables to all HiPCAT and LEARN members

NLR – Southern Route Preferred Alternative Wiltel Fiber in the West -- AT&T Fiber in the Southeast





Cluster Computing

Steven L. Johnson Sr. Systems Analyst, Department of Mathematics & Institute for Scientific Computation

Distributed Memory Computing

- Memory is local to each node.
- Each node runs a copy of the OS.
- Shared and local filesystems. Shared authentication (flat files, NIS, LDAP, Krb5).
- Use message passing library (MPI) to exchange data for parallel applications.
- Run many single-node jobs concurrently.
- Can use OpenMP or Pthreads to parallelize on individual MP nodes.

Typical Cluster Configuration

- x86 processor(s) 32 or 64 bit, Intel or AMD; PowerPC, Apple or IBM.
- 512MB to 4GB+ RAM per node
- diskful or diskless compute nodes
- head nodes for login/edit/compile/debug/submit
- dedicated I/O nodes
- Interconnect: 100Mbps Ethernet to 10Gbps Infiniband. Usually on private net.

Configuration Factors

- Type of application: CPU bound? Memory hog? I/O bound? Communication bound?
- Price/performance ratio for cpu, memory, I/O, interconnect options.
- Density of nodes. Space, power, cooling.
- Local disk or diskless? Cluster FS or NFS?
- Backup policy.
- Static or dynamically linked applications?

Cluster Setup: OS and Networking

- Non-trivial for first time users. Strong Linux/Unix knowledge required.
- Use prepackaged cluster distro or build your own? Diskboot or netboot (pxelinux, Etherboot)?
 - NPACI Rocks
 - Build skeletal SuSE or RedHat (< 50MB)</p>
 - □ Build busybox env (< 8MB)
- OS driver support issues: network, filesystem.
- MPI implementation: MPICH or LAM? Both?

Cluster Setup: Management

- Monitoring/Management: off the shelf or build your own?
 - CPU load, memory use, I/O and network rates
 - □ Temperature, voltage.
 - Remote reboot, reset, power cycle.
 - Disk imaging.
- Clock synchronization.
- Authentication.

Cluster Setup: Scheduling

- Scheduling choices
 - OpenPBS/Torque
 - LSF
 - SGE
 - Condor
- Scheduling parameters
 - Number of concurrent jobs per queue
 - □ CPU time, wall time
 - Node limits. Usually one job per node.
 - Dedicated time

Discussion

Issues for Discussion

Budgeting \$\$\$ and Staff Time

- Maintenance vs. Acquisition (Hardware and Software)
- Retiring old equipment"
- Balancing User Needs (there are conflicts!)
 - □ Turn-around time vs. idle cycles
 - Availability of dedicated time
- "Cycle Shop" vs. "HPC Center"
- Next Step(s)
 - Acquisitions: (Shared memory vs Cluster)
 - Campus Grid efforts (TIGRE)
- (Other Issues Raised by Users)

Major Application Software

Name	Vendor	Cosmos	Agave	K2
Abaqus	Abaqus Inc.	✓	✓	✓
Ansys	ANSYS Inc.	\checkmark	\checkmark	\checkmark
Blast	NCBI	\checkmark	\checkmark	\checkmark
CFX	ANSYS Inc.	\checkmark	\checkmark	\checkmark
Ensight	CEI	\checkmark		\checkmark
Fluent	Fluent Inc.	\checkmark	\checkmark	\checkmark
Gaussian	Gaussian Inc.	\checkmark	\checkmark	\checkmark
GCG	Accelrys			\checkmark
LS-DYNA	LSTC	\checkmark		\checkmark
Matlab	The Mathworks		\checkmark	\checkmark
Migrate	Florida State U	✓	✓	\checkmark
Patran	MSC Software			\checkmark
PHASE	Manchester U	\checkmark		
StarCD	CD-adapco	\checkmark	\checkmark	\checkmark

Normalized CPU Hours Used in FY04



Normalized CPU Hours Used in FY05

