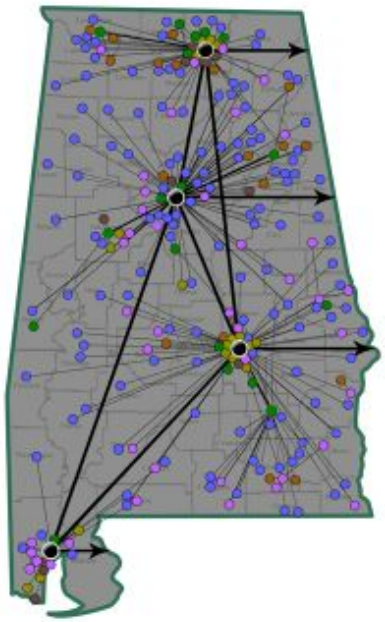
Alabama Supercomputer Center
Alabama Research and Education Network



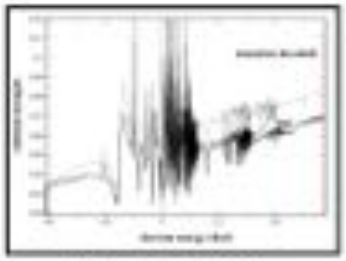
New and Cool HPC Developments at the Alabama Supercomputer Center

presented at

UAB Research Computing Day 2013

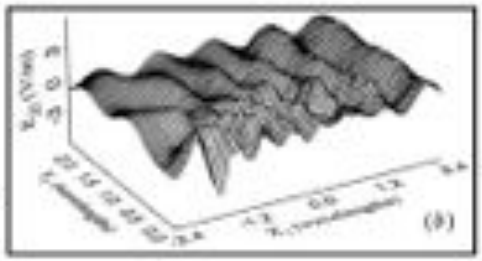
- Ultraviolet 2000 & processor models
- Dense Memory Cluster (DMC)
- GPUs & programming them
- Documentation updates
- non-HPC activities at ASA

How Supercomputers Are Used



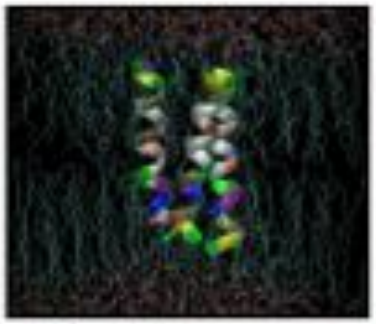
Auburn University
Dr. Loch

University of Alabama
Dr. Tanner



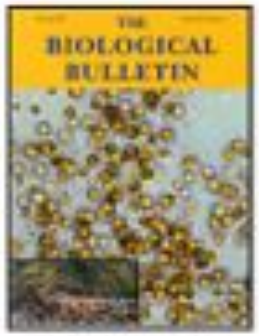
UAH Dr. Jarem

Computer Science



Alabama A&M University
Dr. Kim

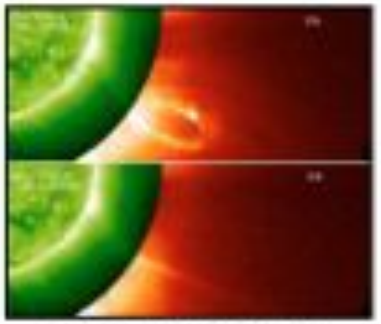
Mathematics



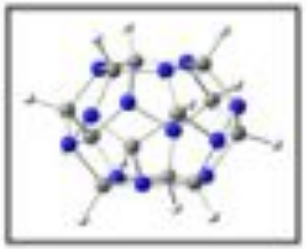
Auburn University
Dr. Santos

UAB Dr. Shin

Music

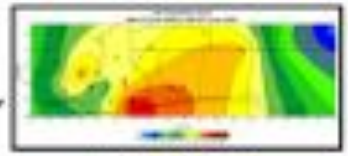


Auburn University Dr. Lin



ASU Dr. Streut

Business



USA Dr. Kimball

BSCC
Dr. Freeman

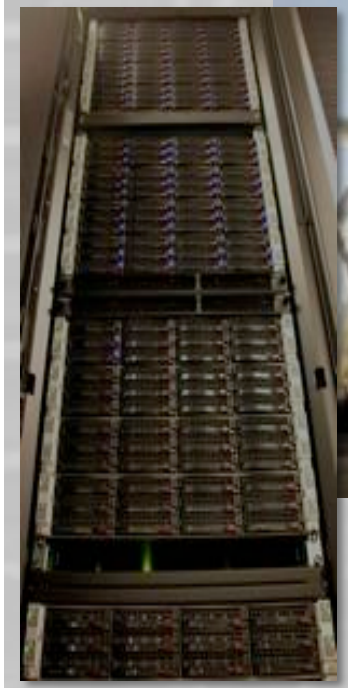
Nuclear Physics
Quantum Chemistry
Semiempirical
Molecular Dynamics
Bioinformatics
Materials Science
Neurology
Electromagnetics
Medicine
Design Analysis, CFD
Agriculture
Weather Modeling
Social Science
Earth Science
Solar System
Galaxy

10^{-15} 10^{-10} 10^{-5} 10^{0} 10^{5} 10^{10} 10^{15} 10^{20}

Length Scale in Meters

A Tale of Two Supercomputers

HPC



The DMC has many nodes with 8 or 16 cores, similar to having many small rockets.

The UV has one big node with 256 cores and much more memory per node, similar to having a Saturn V rocket.

Utilizing Processors

HPC

- **Serial Processing** – Traditionally, most software has used a single computer processor core.
  - Both computers can run serial software, but the UV has more memory.
- **Shared Memory Parallelism** – Software that runs on multiple processor cores that can access the same memory using programming tools like OpenMP.
  - Example: Running World of Warcraft on a dual core laptop.
  - The DMC can run shared memory programs on the 8-16 cores in a given node.
  - The SGI UV2000 can run shared memory programs on the 256 cores in a compute node.
  - GPU math co-processors provide a type of shared memory parallelism
- **Distributed Memory Parallelism** – Software that utilizes multiple computers on a network using programming tools like MPI.
  - Example: SETI@home
  - Both the DMC and the SGI UV can run distributed memory programs.

SGI UV Supercomputer

UV

- 268 Xeon “Sandy Bridge” Processors
  - 5,194 GFLOPS Peak
- Shared Memory Architecture
  - NUMAlink shared memory network
- Memory (4TB per node)
  - 4,160 GB Total
- Disk Storage
  - 15 TB shared

UV came online Jan 2013, Altix offline July 2013

Vector/SIMD extensions

CPU

- 4 operations to add two single precision (32 bit), four-component vectors

```
vector_result.x = vector_1.x + vector_2.x;
vector_result.y = vector_1.y + vector_2.y;
vector_result.z = vector_1.z + vector_2.z;
vector_result.w = vector_1.w + vector_2.w;
```

- Using 128-bit SSE registers, four-component vectors are added in a single operation

- Intel's Sandy Bridge architecture (used in UV) introduced AVX instructions for 256 bit vector operations, potentially resulting in up to a 2x performance improvement for some applications

AVX capable chips will be added to DMC Q4 2013

Dense Memory Cluster (DMC)

DMC

- 1,800 x86-64 Processors (AMD/Intel)
  - 16,462 GFLOPS Peak
- Shared/Distributed Memory Architecture
  - InifiniBand high speed/low latency network
- Memory (24-128GB per node)
  - 10,136 GB
- Disk Storage
  - 225 TB internal, 20 TB shared

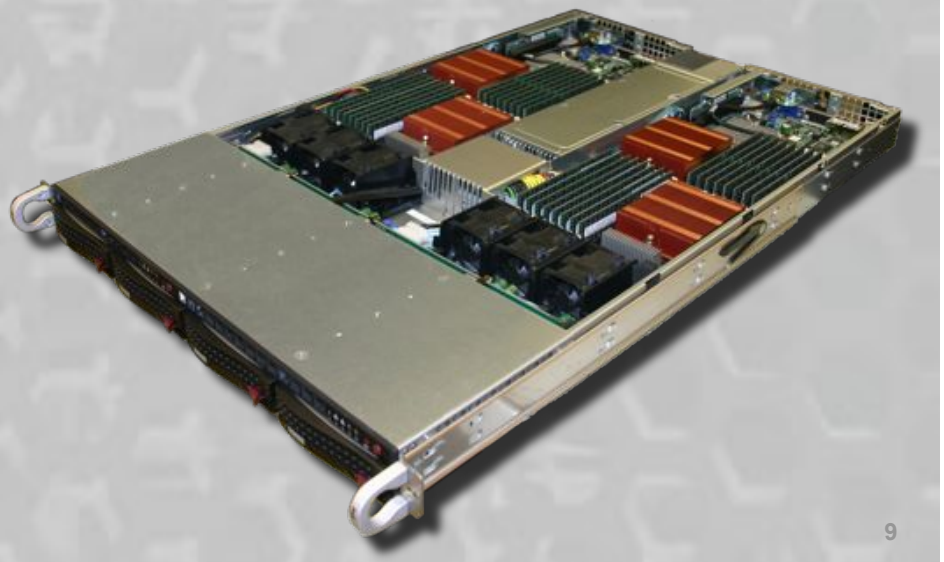
40 nodes (20 cores, 128 GB each) will be added to DMC Q4 2013

DMC Nodes

DMC

Year	Processor	Cores/node	Memory/node	SPECFP/node
2008	2x 2.3GHz Opteron	8	64GB	89
2009	2x 2.26GHz Xeon	8	24GB	155
2010	2x 2.3GHz Opteron	16	128GB	252
2011	2x 2.3GHz Opteron	16	128GB	252

Each DMC node has 8-16 CPU cores, 24-128GB of memory and 1-2TB of local disk space. Annual upgrades take advantage of higher density /performance and react to user needs.



GPUs

GPU

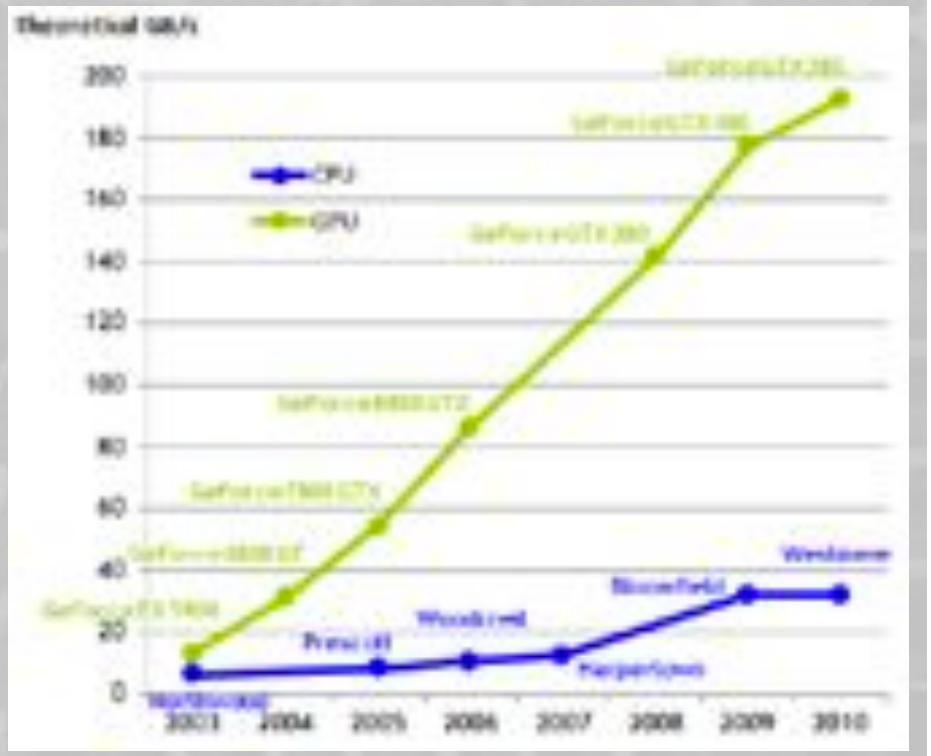
Graphic Processing Units (GPUs) are graphics chips typically found in video cards. There has been an experimental movement in supercomputing to utilize these chips as math coprocessors.

Chips have evolved from specialized graphics hardware into more conventional massive multithreaded, manycore SIMD processors.

Previously programmed using standard graphics APIs (DirectX/OpenGL), but new software development kits enable more direct/straightforward programming in C/C++ and other high-level languages.

Why GPUs?

GPU



Comparison of peak theoretical GFLOPs and memory bandwidth for NVIDIA GPUs and Intel CPUs over the past few years.

Graphs from the NVIDIA CUDA C Programming Guide 4.0.

NVIDIA Tesla GPUs (DMC)

GPU

Tesla S1070
- 8 T10 GPUs
- 4GB memory/GPU
- 240 cores

Tesla M2070
- 8 Fermi GPUs
- 6GB memory/GPU with ECC support
- 448 cores

16 Kepler K20 (2496 core) GPUs scheduled for Q4 '13

CUDA GPU programming example

GPU

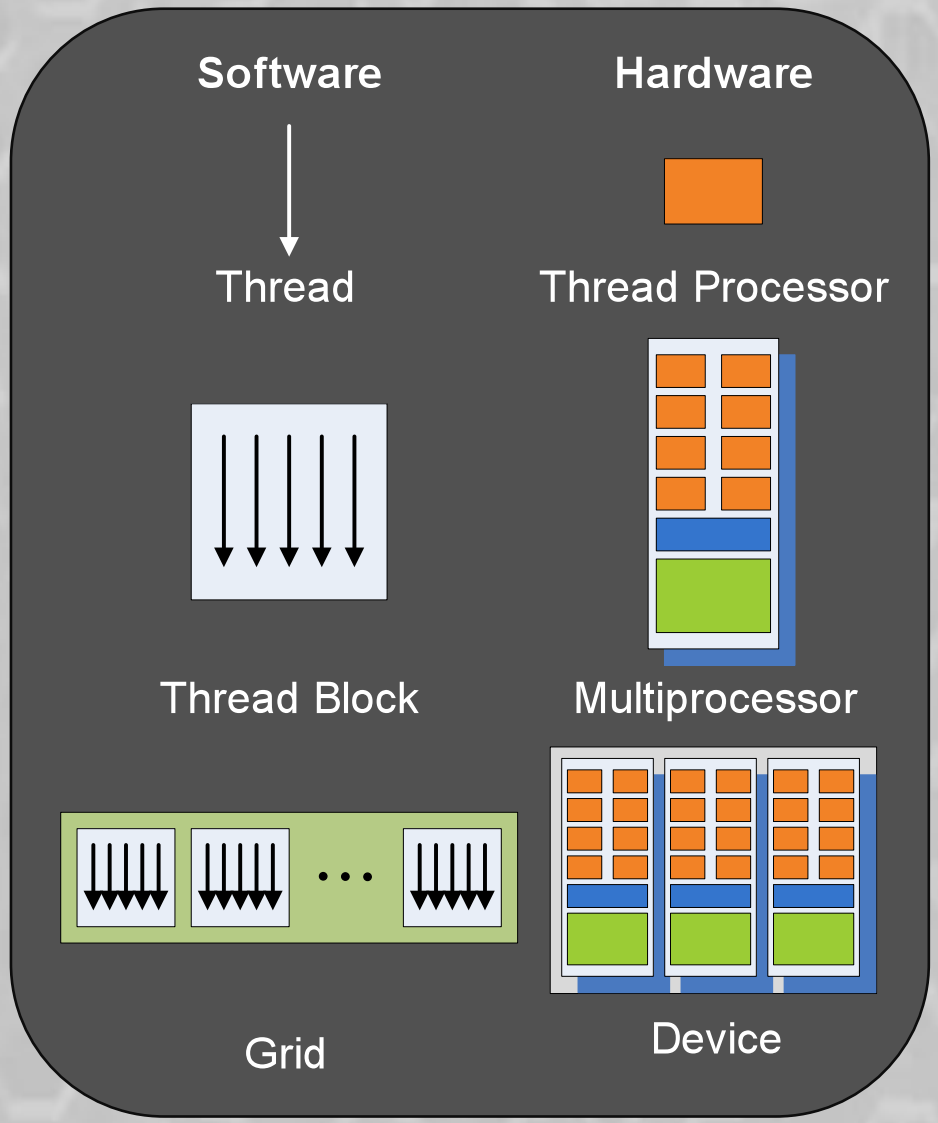
```
// CPU only matrix addition
int main() {
 int i, j;
 for (i=0;i<N;i++) {
  for (j=0;j<N;j++) {
   C[i][j]=A[i][j]+B[i][j];
  }
 }
}
```

```
// GPU kernel
__global__ gpu(A[N][N], B[N
    ][N], C[N][N]) {
 int i = threadIdx.x;
 int j = threadIdx.y;
 C[i][j]=A[i][j]+B[i][j];
}

int main() {
 dim3 dimBlk(N,N);
 gpu<<1,dimBlk>>(A,B,C);
}
```

GPU Execution Model

GPU



Thread is a single execution of a kernel, and all execute the same code

Threads within a block have access to shared memory for local cooperation

Kernel launched as a grid of independent thread blocks, and only a single kernel executes at a time

Number of Software Packages

Apps

Bioinformatics	62
Programming	43
Mathematics	28
Quantum Chemistry	24
Visualization	21
Fluid Dynamics	13
Molecular Dynamics	12
Weather Modeling	10
Materials Science	9

- These numbers include libraries and utilities, as well as the core packages.

Compilers and Programming

HPC

- Compilers
- GNU C/C++ Fortran 77/90/95
- Intel C/C++ Fortran 77/90/95
- Portland Group C/C++ Fortran 77/90/HP

- Parallel Programming
- Shared memory: OpenMP, pthreads, Java threads
- Distributed memory: MPI
- Math libraries: ACML, SLATEC, MKL, SCSL, IMSL
- GPU: CUDA, OpenCL **OpenACC coming soon**

User Support

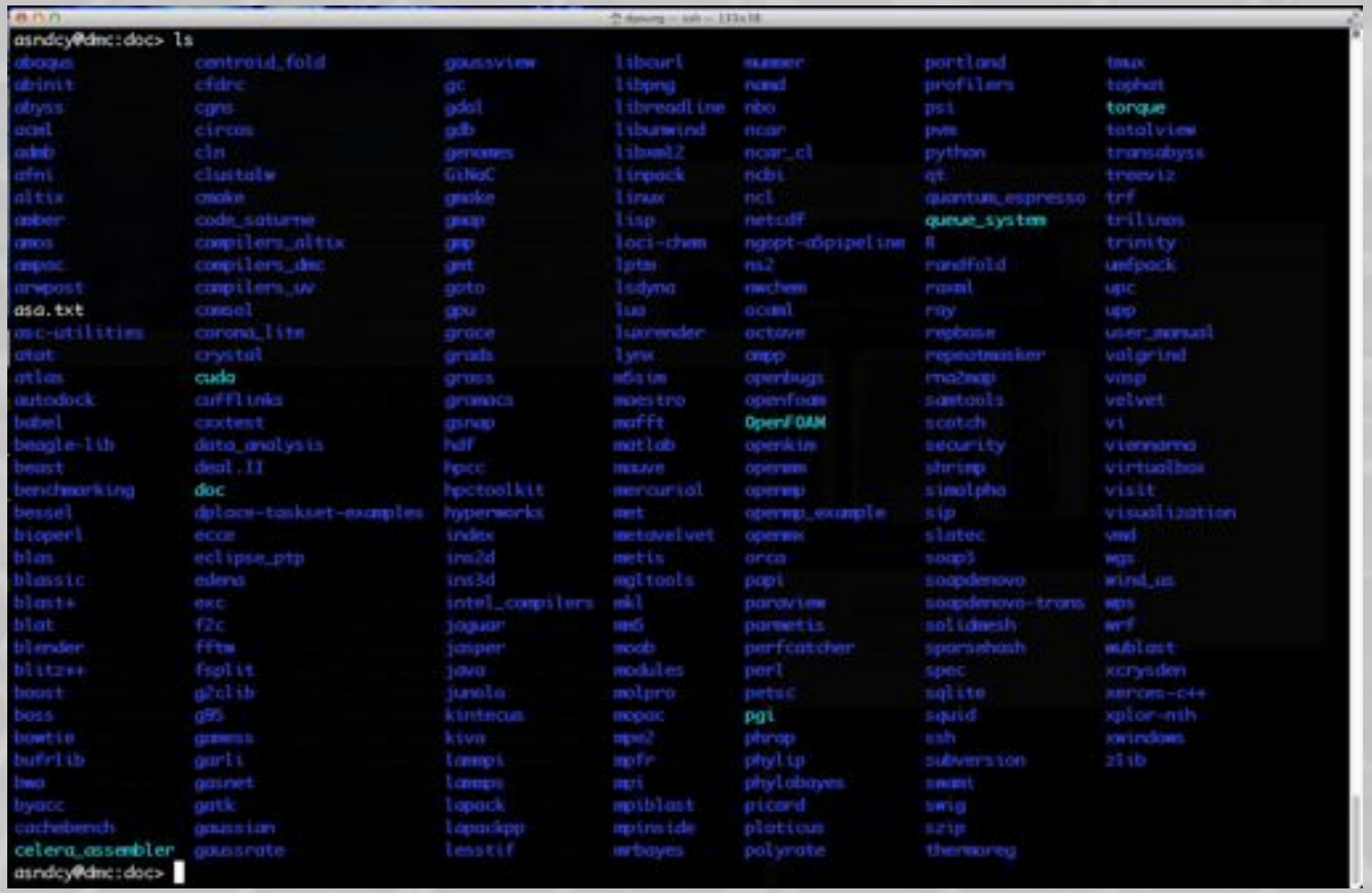
Doc

- **Documentation**
  - HPC User Manual
  - Application specific README files
  - Man pages
  - Programming examples
  - Best practices white papers
- **Queue scripts**
  - A uniform front end for submitting all jobs to the queue that hides the details of queue command syntax.
- **Technical support staff**
  - Our HPC staff have degrees in chemistry, mathematics, business, MIS, computer science, and electrical engineering.
- **Software installation**

New User Manual in 2013

Old Documentation System
A directory for every software package

Doc



New Documentation System
A browser to navigate the doc directory

Doc

dyoung — ssh — 80×24

```
                ascdocs - Documentation Browser

/opt/asn/doc/index

 1.  all                          2.  bioinformatics
 3.  crystallography              4.  fluid_dynamics
 5.  general_information          6.  materials_science
 7.  mathematics                  8.  molecular_dynamics
 9.  other                       10.  programming
11.  quantum_chemistry           12.  semiempirical
13.  structural_engineering      14.  utilities
15.  visualization               16.  weather_modeling
17.  Go up one directory.
18.  Go back to the main index.
19.  Exit.

Enter the number of your selection:
```

Examples of White Papers

Doc

- Introduction to Big Data Analysis for Scientists and Engineers
- Software Development Methodologies
- Choosing a Version Control System
- Getting Started with Visualization
- Switching from CUDA 4.x to CUDA 5.0
- Introduction to GPU Programming with CUDA

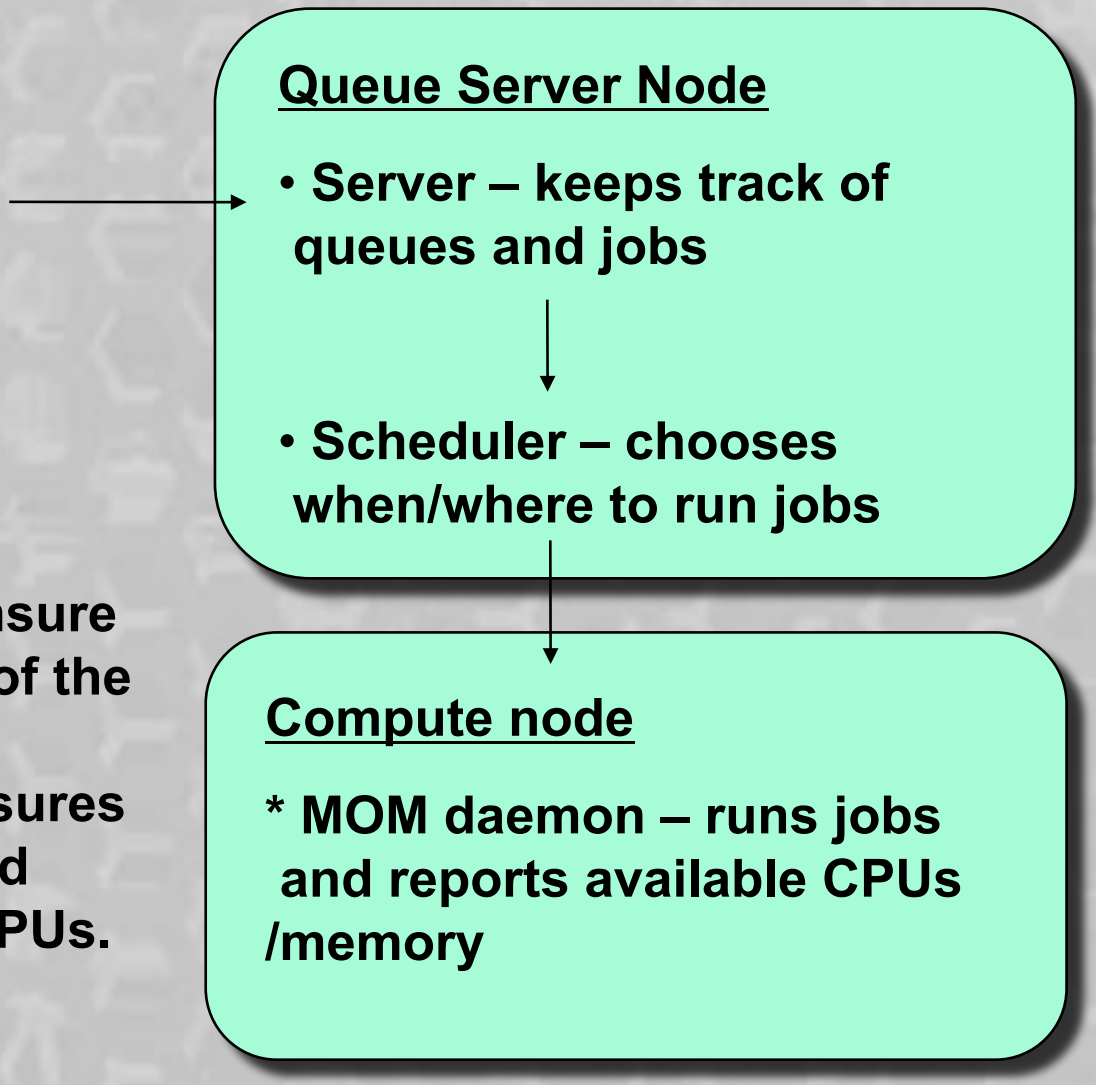
Torque/MOAB queue system

HPC

User commands

- qsub – run a job
- qstat – see status
- checkjob – see job information

MOAB runs jobs to ensure maximum utilization of the system, without over -subscription, and ensures jobs get the requested amount of memory/CPUs.



Queue scheduler algorithm HPC

- Jobs run only if the requested memory and CPUs are available.
- One person can use a larger percentage of the system, if it isn't being used.
- Multiple people requesting many resources will get equal number of CPU cores.
- Jobs from small users jump ahead of jobs from big users.
- Reservations ensure resources for class work and types of jobs that take a long time to queue.
- The jobs that have been waiting the longest are labelled as "starving" and other jobs can no longer jump ahead of them.

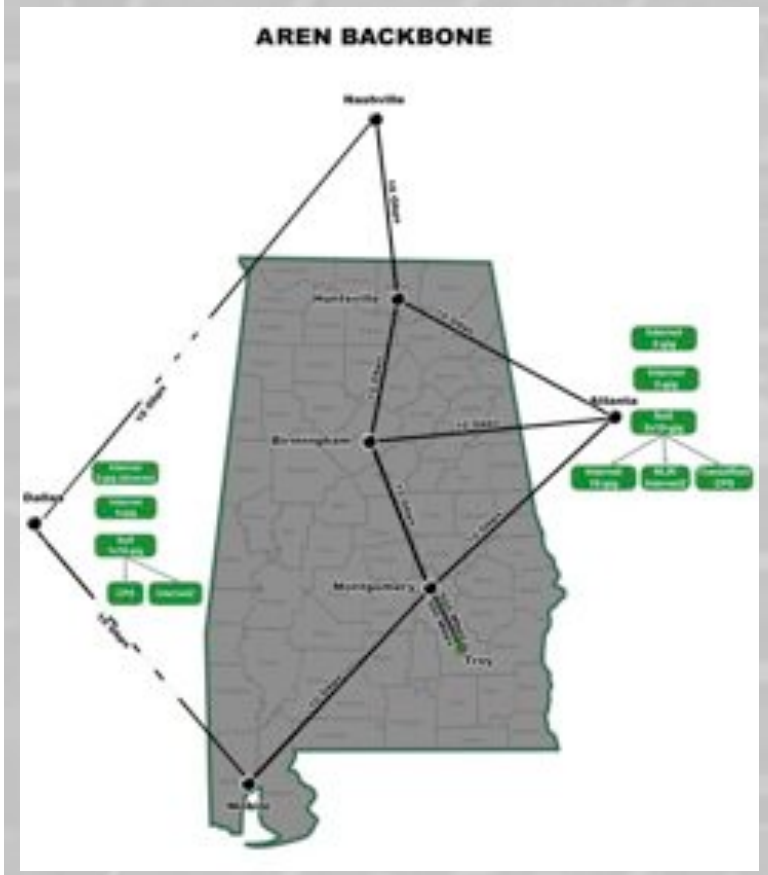
ASC queue list

HPC

Queue	CPU	Mem	# CPUs
small-serial	40:00:00	4gb	1
medium-serial	90:00:00	16gb	1
large-serial	240:00:00	120gb	1
small-parallel	48:00:00	8gb	2-8
medium-parallel	100:00:00	32gb	2-16
large-parallel	240:00:00	120gb	2-64
class	2:00:00	64gb	1-64
commercial	1008:00:00	360gb	1-128
daytime	4:00:00	16gb	1-4
express	01:00:00	500mb	1
special	1008:00:00	700gb	1-100

Statewide Network

Net



- Alabama Research and Education Network (AREN)
  - Universities & Colleges
  - K-12 School Systems
  - Public Libraries
- State Internet2 Network Operations
- 24x7 Operations Center
- Excellent Statewide Network Infrastructure

Additional ASA Services

ASA

- Web & Email Hosting
- Distance Learning
- Disaster Recovery
- Software Development
- Application Server Hosting
  - Alabama Virtual Library (AVL)
- On-Demand Computing
- Economic Development

Summary

HPC

- The Alabama Supercomputer Authority provides two high performance computing systems. These are free of charge for use by state funded educational institutions in Alabama.
- The SGI UV is a large NUMA shared-memory system using Intel Xeon CPUs.
- The DMC is a distributed memory system using an Infiniband interconnect and AMD Opteron and Intel Xeon CPUs. The DMC has GPU math coprocessors.
- A variety of software is available.
- ASA also provides network, hosting, and software development services to the academic community.

Alabama Supercomputer Authority

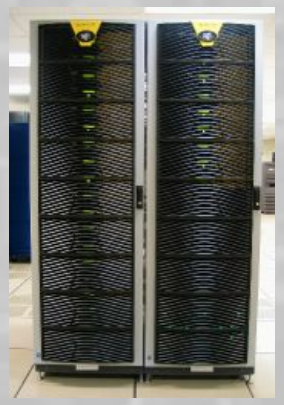
State of Alabama Leader and Trusted Partner for Technology

Alabama Supercomputer Authority
Historical Perspective

HPC

Cray X-MP
1987

Cray C90
1994



SGI Altix 350
2004

Altix 450
2006

SGI UV 2000
2012

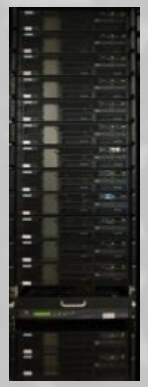
An eternity in computer years

9 node
network

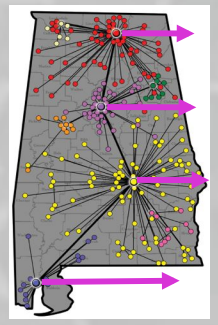
nCube
1991

Cray SV1
1999

Cray XD1
2004



DMC
2008



640 node
network

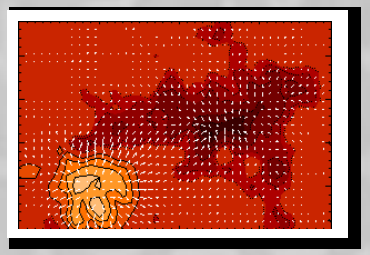
Performance Comparison

HPC

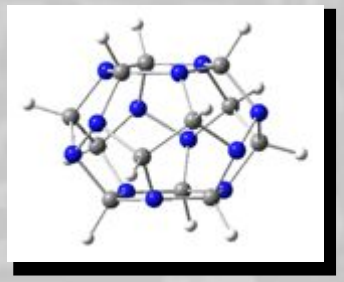
	DMC	SGI UV
Number of CPUS	1800	268
CPU Type	Xeon / Opteron	Xeon
SPECFP / core	11.2 – 19.4	24.375 – 34.75
Relative processing capacity	3.6	1.0
Memory (GB)	10,136	4,160
Internal Disk (TB)	225	2.2
Shared Disk (TB)	20	15
GFLOP	16,462	5,194
Clock (GHz)	2.26, 2.3, 2.4	2.4, 2.9

Who uses HPC?

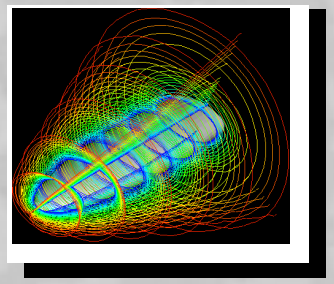
HPC



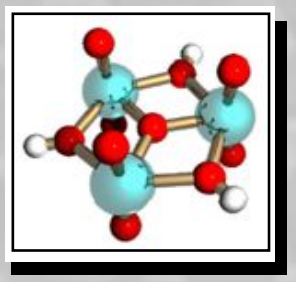
University of Alabama
in Huntsville



Alabama State University

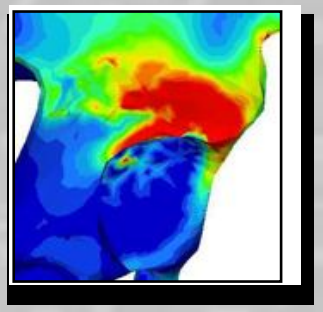


Alabama A&M University



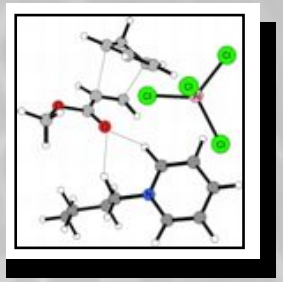
University of Alabama

Athens State University
Auburn University -Montgomery
Bevill State College
Jacksonville State University
Troy University
Tuskegee University
University of West Alabama
University of Montevallo
U.S. Air Force
U.S. Army
NASA
Intel Corporation
Operon Biotechnologies
Time Domain



University of Alabama
at Birmingham

University of South
Alabama



Auburn University