

High-Performance Computing, Computational and Data Grids

Russ Miller

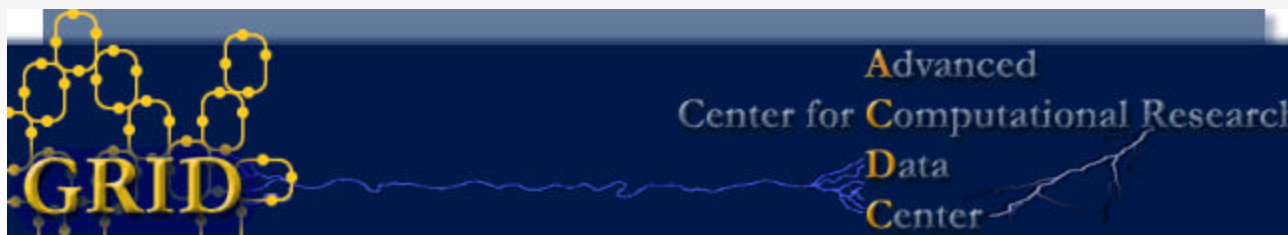
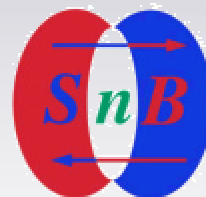
Center for Computational Research

Computer Science & Engineering

SUNY-Buffalo

Hauptman-Woodward Medical Inst

NSF, NIH, DOE
NIMA, NYS, HP



University at Buffalo

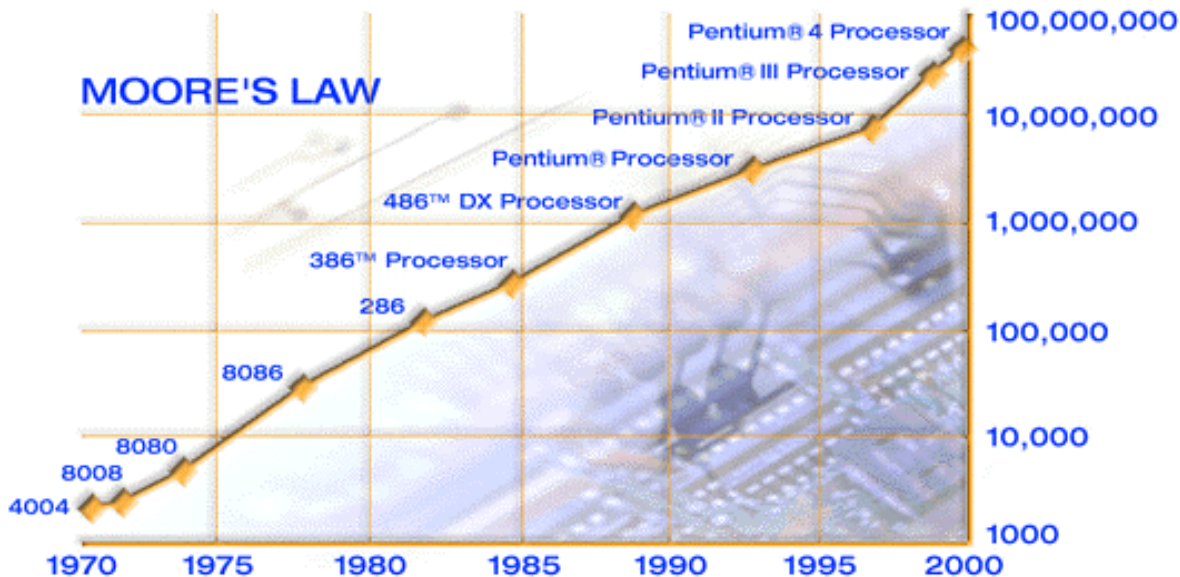
The State University of New York

Gordon E. Moore

- Co-Founder of Intel
- Predicted (1965/75) that transistor density would double every 12/18 months
- Processing speed doubling every 18 mos.
- Disk storage doubling every 12 mos.
- Aggregate bandwidth doubling every 9 mos.



Gordon E. Moore

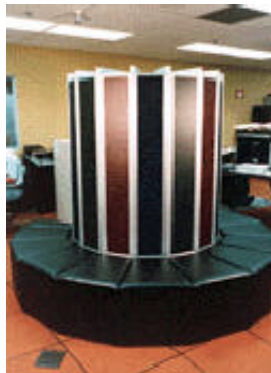


- A computation that took 1 year to run on a PC in 1985 would only take 5 mins to run on a PC today!
- A computation that runs in 2 hours on a PC today would have taken 24 years to run on a PC in 1985!

Supercomputers

- Fastest computers at any point in time
- Used to solve large and complex problems
- Machines 1000 times faster than a PC
- Machines 10 times slower than what you need to solve the most challenging problems

“Seymour Cray is the Thomas Edison of the supercomputing industry”
- Larry L. Smarr



Cray1 - 1976



Seymour Cray
1925-1996

Beowulf Clusters

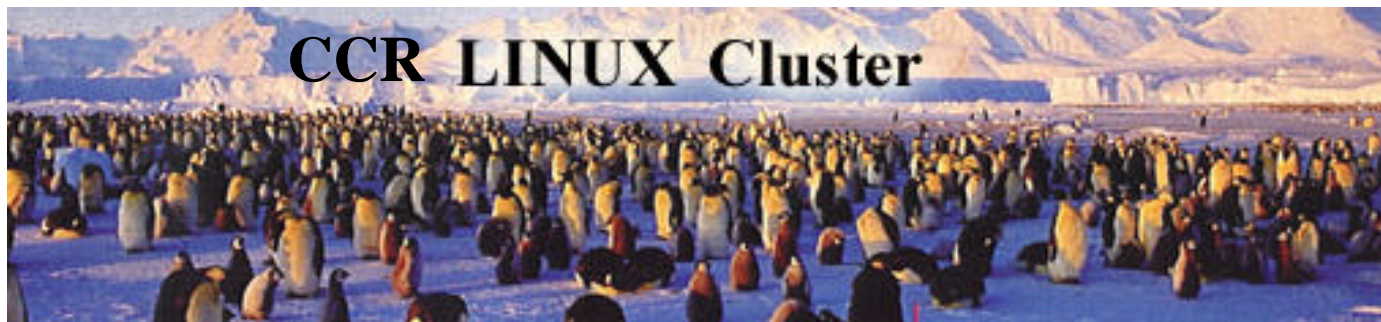
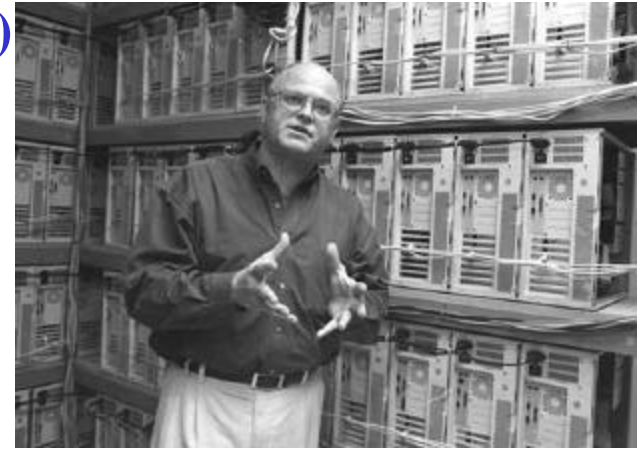
■ Industry Standard Hardware and Software

Thomas Sterling
Caltech

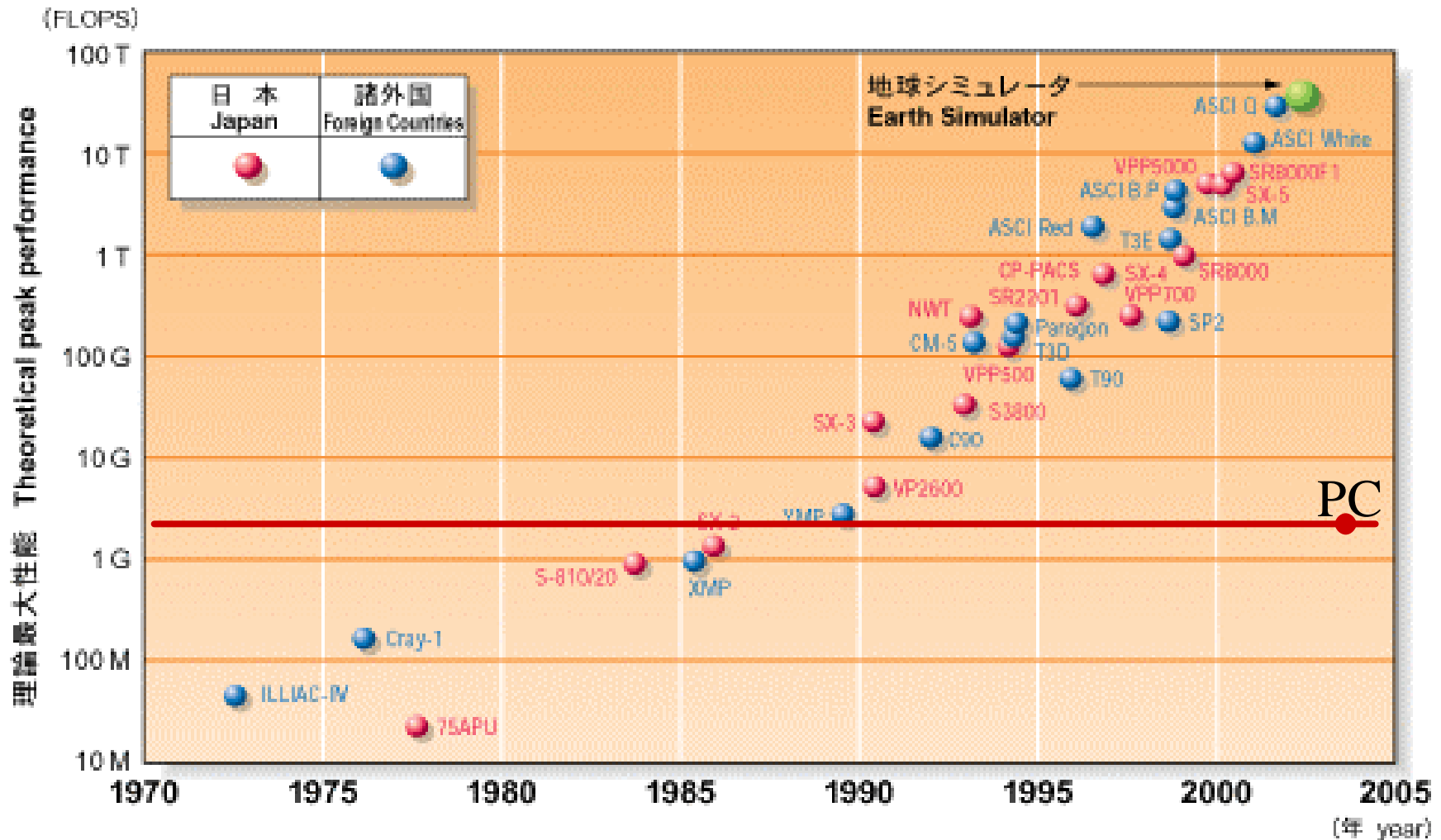
- ❑ PC-Based Components (Intel or AMD)
- ❑ Ethernet, Myrinet, InfiniBand
- ❑ Linux, PBS, MPI
- ❑ “Commodity Off-The-Shelf” (COTS)

■ Operates as a Single System

■ Rivals Performance of Traditional Supercomputer at a Fraction of the Price



Growth of Peak Performance



TOP500 List for June 2004

R_{max} and R_{peak} values are in GFlops. For more details about other fields, please click on the button "Explanation of the Fields"


[EXPLANATION OF THE FIELDS](#)

Rank	Site Country/Year	Computer / Processors Manufacturer	Computer Family Model	Inst. type Installation Area	R_{max} R_{peak}	N_{max} N_{half}
1	Earth Simulator Center Japan/2002	Earth-Simulator / 5120 NEC	NEC Vector SX6	Research	35860 40960	1.0752e+06 266240
2	Lawrence Livermore National Laboratory United States/2004	<i>Thunder</i> Intel Itanium2 Tiger4 1.4GHz - Quadrics / 4096 California Digital Corporation	NOW - Intel Itanium Itanium2 Tiger4 Cluster - Quadrics	Research	19940 22938	975000 110000
3	Los Alamos National Laboratory United States/2002	ASCI Q - AlphaServer SC45, 1.25 GHz / 8192 HP	HP AlphaServer Alpha-Server-Cluster	Research	13880 20480	633000 225000
4	IBM - Thomas Watson Research Center United States/2004	BlueGene/L DD1 Prototype (0.5GHz PowerPC 440 w/Custom) / 8192 IBM/ LLNL	IBM BlueGene/L BlueGene/L	Research	11680 16384	331775
5	NCSA United States/2003	<i>Tungsten</i> PowerEdge 1750, P4 Xeon 3.06 GHz, Myrinet / 2500 Dell	Dell Cluster PowerEdge 1750, Myrinet	Academic	9819 15300	630000
6	ECMWF United Kingdom/2004	eServer pSeries 690 (1.9 GHz Power4+) / 2112 IBM	IBM SP SP Power4+, Federation	Research Weather and Climate Research	8955 16051	350000
7	Institute of Physical and Chemical Res. (RIKEN) Japan/2004	RIKEN Super Combined Cluster / 2048 Fujitsu	Fujitsu Cluster Fujitsu Cluster	Research	8728 12534	474200 120000
8	IBM - Thomas Watson Research Center United States/2004	BlueGene/L DD2 Prototype (0.7 GHz PowerPC 440) / 4096 IBM/ LLNL	IBM BlueGene/L BlueGene/L	Research	8655 11469	294911
9	Pacific Northwest National Laboratory United States/2003	<i>Mpp2</i> Integrity rx2600 Itanium2 1.5 GHz, Quadrics / 1936 HP	HP Cluster Integrity rx2600 Itanium2 Cluster	Research	8633 11616	835000 140000
10	Shanghai Supercomputer Center China/2004	Dawning 4000A, Opteron 2.2 GHz, Myrinet / 2560 Dawning	NOW - AMD NOW Cluster - AMD - Myrinet	Research	8061 11264	728400 180000

Peak

40 TF

23 TF

20 TF

16 TF

15 TF

16 TF

13 TF

11 TF

12 TF

11 TF

Linpack

36 TF

20 TF

14 TF

12 TF

10 TF

9 TF

9 TF

9 TF

9 TF

8 TF

Capacity vs Capability Computing

■ Capacity Computing

- ❑ Commodity Processors
- ❑ Commodity Networks
- ❑ Leverage Millions of Processors Designed for Home Use
- ❑ Attractive Price Point

Commodity (Clusters, NOWs)

Commodity Processors

Commodity Interconnects

IA32, IA64, AMD, Alpha, PowerPC

GigE, IB, Myrinet, Quadrics

Clusters, HP Alpha, NEC TX7

■ Capability Computing

- ❑ Special Purpose Processors
- ❑ Custom Networks
- ❑ Designed for Science
- ❑ Few Machines Sold
- ❑ High Price

Mixed

Commodity Processors

Custom Interconnects

SGI Altix, Cray Red Storm, Blue Gene/L

Custom

Custom Processors

Custom Interconnects

Cray X1, NEC SX-7, IBM Regatta

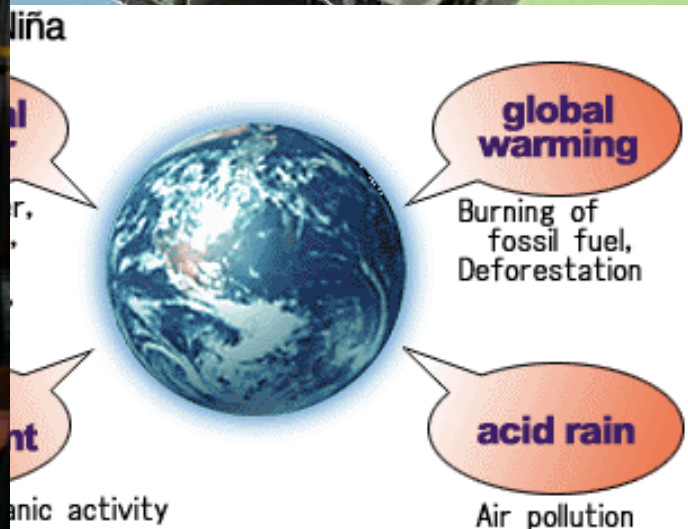
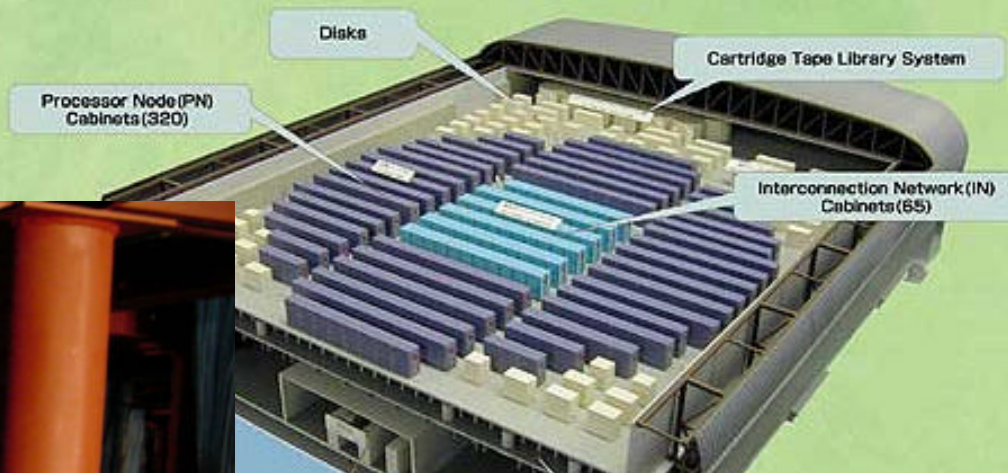
← **Loosely Coupled**

Tightly Coupled →



Earth Simulator

- 40 TFlops Peak
- Homogeneous, Centralized,



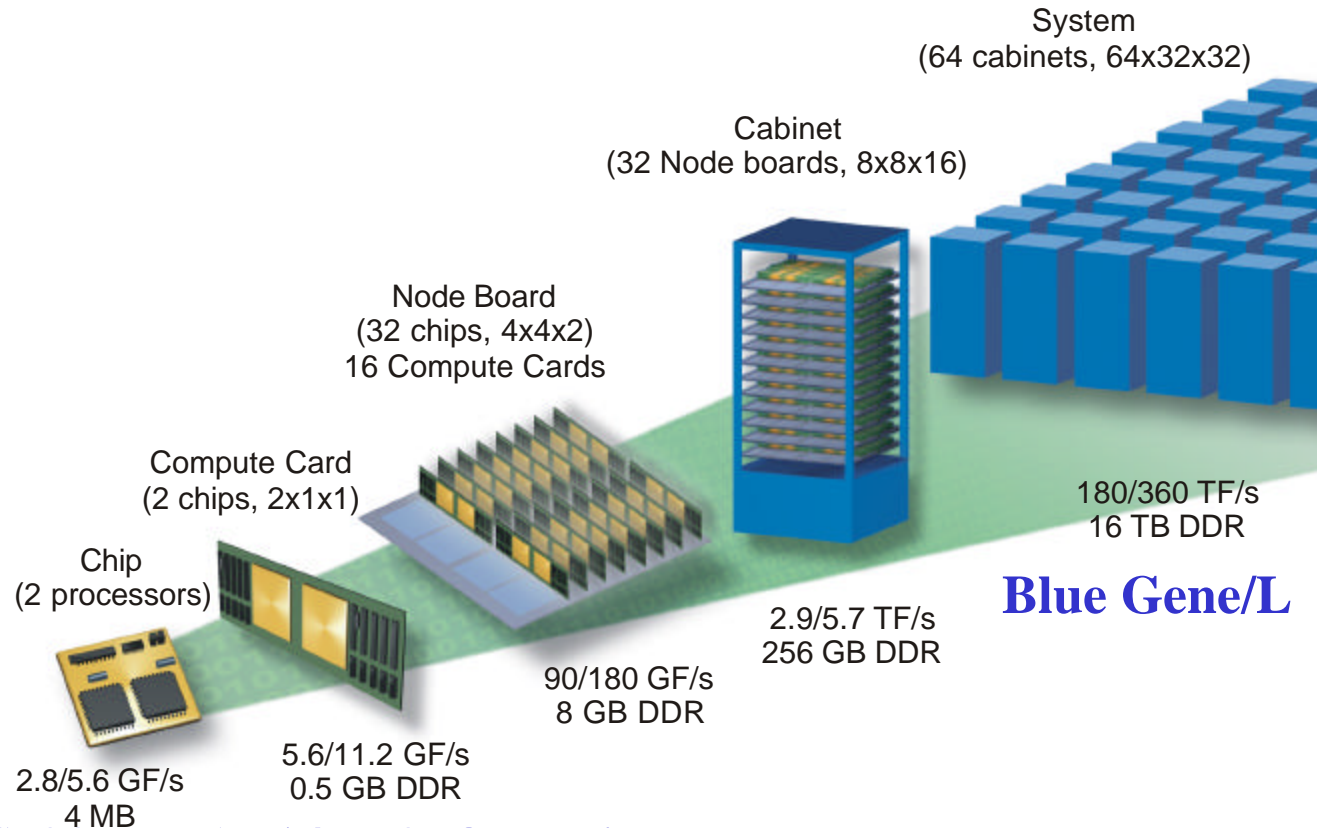
ASCI Purple and Blue Gene/L

■ **ASCI Purple (2004)**

- ❑ LLNL
- ❑ IBM Power5
- ❑ 12K Processors
- ❑ 100 TFlops

■ **Blue Gene/L (2005)**

- ❑ LLNL
- ❑ PowerPC
- ❑ 128K Processors
- ❑ Linux
- ❑ 360 Tflops



■ **LLNL Thunder (2004) 23 TF; 1K 4 IT2, Quadrics**

■ **Sandia/PSC Cray Red Storm (2004) 40 TF, Opteron, Hypertransport**

■ **LANL Lightning (2004) 11 TF; Linux Networx; 2816 Opteron; Myrinet**

■ **ORNL Cray X1 (2005) 50 TF; Vector PEs; 2006: 100 TF**

■ **Dawning (China, 2004) 11 TF; 2560 Operon; Myrinet**

Center for Computational Research 1999-2004 Snapshot

■ High-Performance Computing and High-End Visualization

- ❑ 110 Research Groups in 27 Depts
- ❑ 13 Local Companies
- ❑ 10 Local Institutions

■ External Funding

- ❑ \$116M External Funding
 - \$16M as lead
 - \$100M in support
- ❑ \$43M Vendor Donations
- ❑ Total Leveraged: \$0.5B

■ Deliverables

- ❑ 400+ Publications
- ❑ Software, Media, Algorithms, Consulting, Training, CPU Cycles...

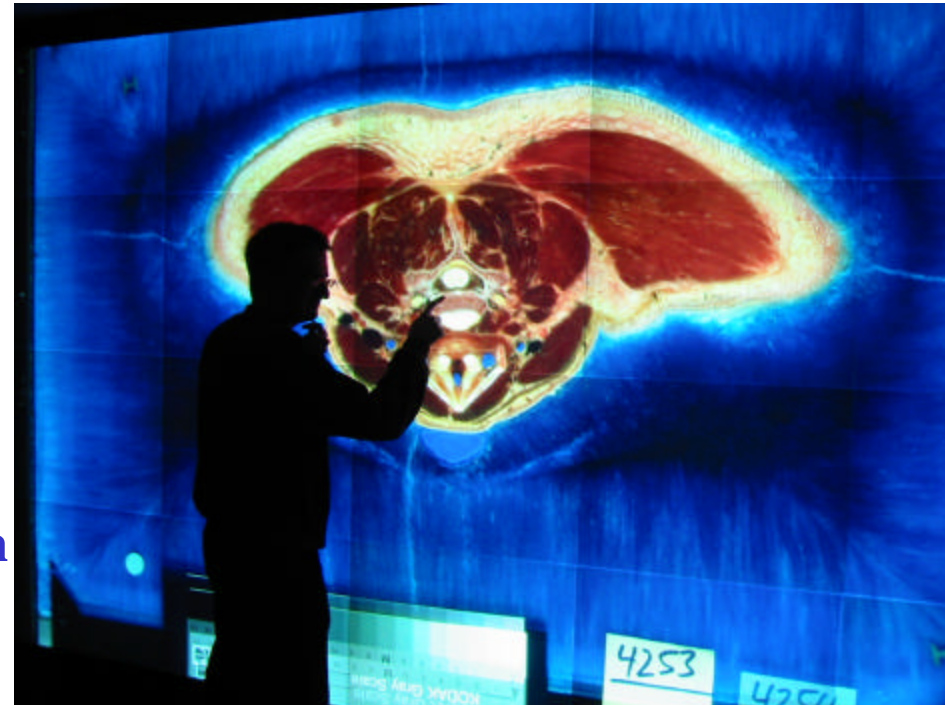


Major CCR Resources (12TF & 290TB)

- **Dell Linux Cluster: #22® #25® #38® #95**
 - ❑ 600 P4 Processors (2.4 GHz)
 - ❑ 600 GB RAM; 40 TB Disk; Myrinet
 - **Dell Linux Cluster: #187® #368® off**
 - ❑ 4036 Processors (PIII 1.2 GHz)
 - ❑ 2TB RAM; 160TB Disk; 16TB SAN
 - **IBM BladeCenter Cluster: #106**
 - ❑ 532 P4 Processors (2.8 GHz)
 - ❑ 5TB SAN
 - **SGI Origin3700 (Altix)**
 - ❑ 64 Processors (1.3GHz ITF2)
 - ❑ 256 GB RAM
 - ❑ 2.5 TB Disk
 - **SGI Origin3800**
 - ❑ 64 Processors (400 MHz)
 - ❑ 32 GB RAM; 400 GB Disk
 - **Apex Bioinformatics System**
 - ❑ Sun V880 (3), Sun 6800
 - ❑ Sun 280R (2)
 - ❑ Intel PIIIs
 - ❑ Sun 3960: 7 TB Disk Storage
 - **HP/Compaq SAN**
 - ❑ 75 TB Disk
 - ❑ 190 TB Tape
 - ❑ 64 Alpha Processors (400 MHz)
 - ❑ 32 GB RAM; 400 GB Disk
-
- **IBM RS/6000 SP: 78 Processors**
 - **Sun Cluster: 80 Processors**
 - **SGI Intel Linux Cluster**
 - ❑ 150 PIII Processors (1 GHz)
 - ❑ Myrinet

CCR Visualization Resources

- **Fakespace ImmersaDesk R2**
 - Portable 3D Device
- **Tiled-Display Wall**
 - 20 NEC projectors: 15.7M pixels
 - Screen is 11' ´ 7'
 - Dell PCs with Myrinet2000
- **Access Grid Node**
 - Group-to-Group Communication
 - Commodity components
- **SGI Reality Center 3300W**
 - Dual Barco's on 8' ´ 4' screen



Peace Bridge Visualization

■ Proposed Options

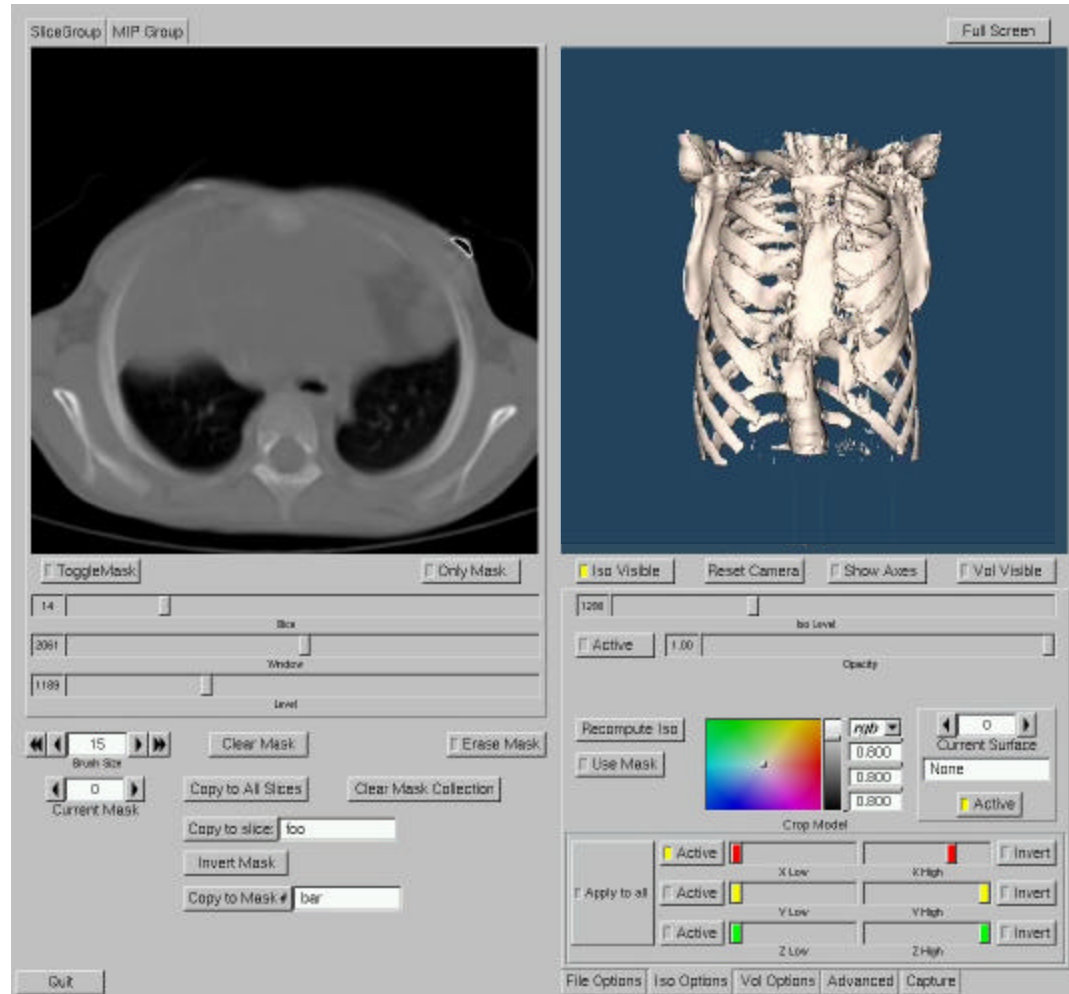
- Relocate US plaza
- Build a 3-lane companion span, rehab existing bridge
- Build a six lane signature span



PHOTO AND STORY BY BRUCE JACKSON

3D Medical Visualization App

- Collaboration with Children's Hospital
 - Leading miniature access surgery center
- Application reads data output from a CT Scan
- Visualize multiple surfaces and volumes
- Export images, movies or CAD representation of model

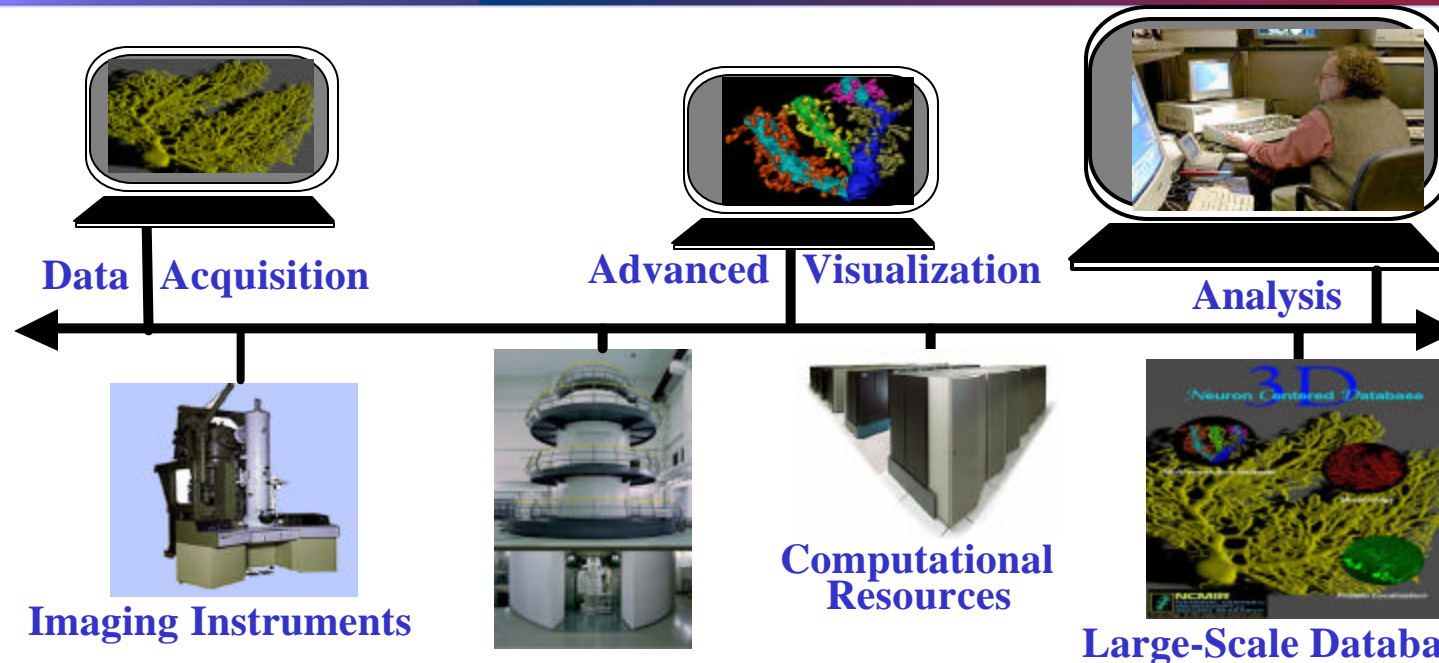


StreetScenes[®] Demo

- *StreetScenes*[®] is a Virtual Reality (VR) software solution for 3D visualization of surface traffic
- 3D model of proposed soccer stadium in Rochester
- Used *StreetScenes*[®] to import output file from Synchro traffic simulation



Grid Computing Overview



Thanks to
Mark Ellisman

- **Coordinate Computing Resources, People, Instruments in Dynamic Geographically-Distributed Multi-Institutional Environment**
- **Treat Computing Resources like Commodities**
 - ❑ Compute cycles, data storage, instruments
 - ❑ Human communication environments
- **No Central Control; No Trust**

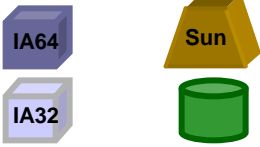
Factors Enabling the Grid

- **Internet is Infrastructure**
 - ❑ Increased network bandwidth and advanced services
- **Advances in Storage Capacity**
 - ❑ Terabyte costs less than \$5,000
- **Internet-Aware Instruments**
- **Increased Availability of Compute Resources**
 - ❑ Clusters, supercomputers, storage, visualization devices
- **Advances in Application Concepts**
 - ❑ Computational science: simulation and modeling
 - ❑ Collaborative environments ® large and varied teams
- **Grids Today**
 - ❑ Moving towards production; Focus on middleware

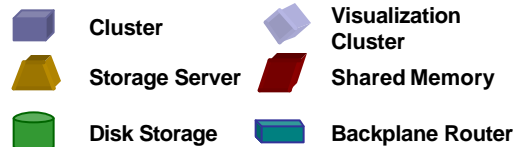
NSF Extensible TeraGrid Facility

Caltech: Data collection analysis

0.4 TF IA-64
IA32 Datawulf
80 TB Storage

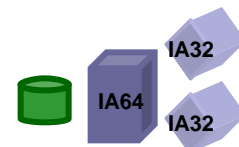


LEGEND

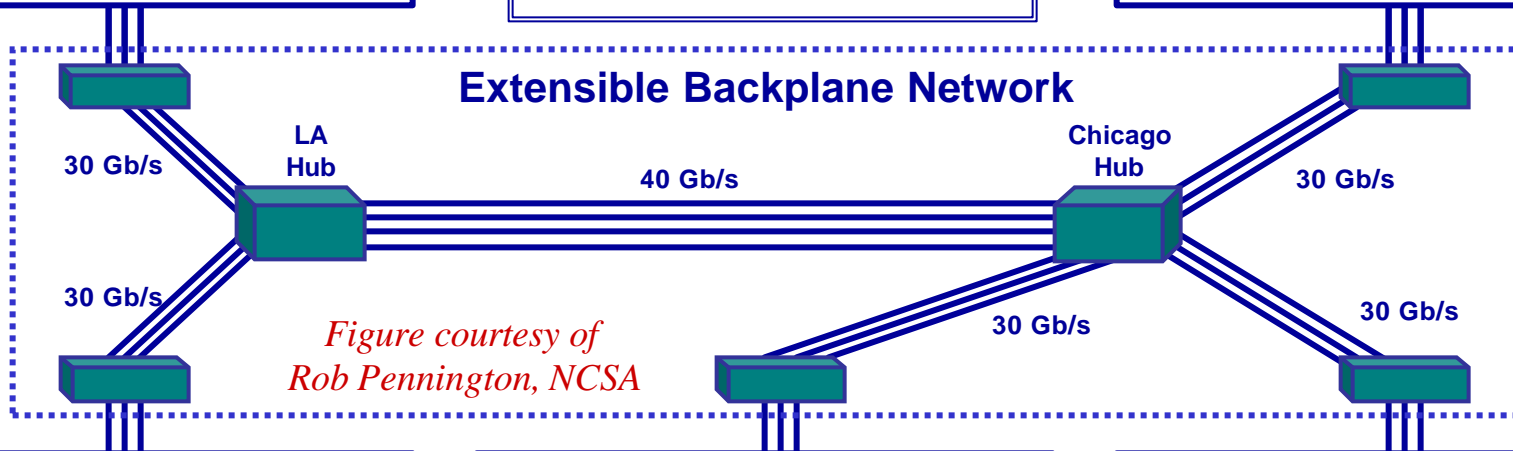


ANL: Visualization

1.25 TF IA-64
96 Viz nodes
20 TB Storage

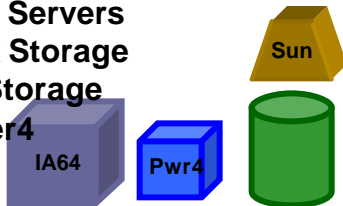


Extensible Backplane Network



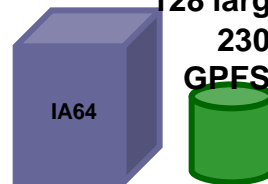
*Figure courtesy of
Rob Pennington, NCSA*

4 TF IA-64
DB2, Oracle Servers
500 TB Disk Storage
6 PB Tape Storage
1.1 TF Power4



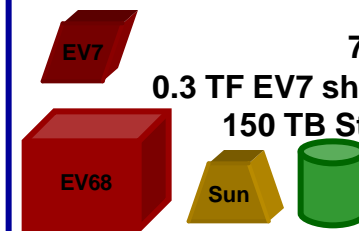
SDSC: Data Intensive

10 TF IA-64
128 large memory nodes
230 TB Disk Storage
GPFS and data mining



NCSA: Compute Intensive

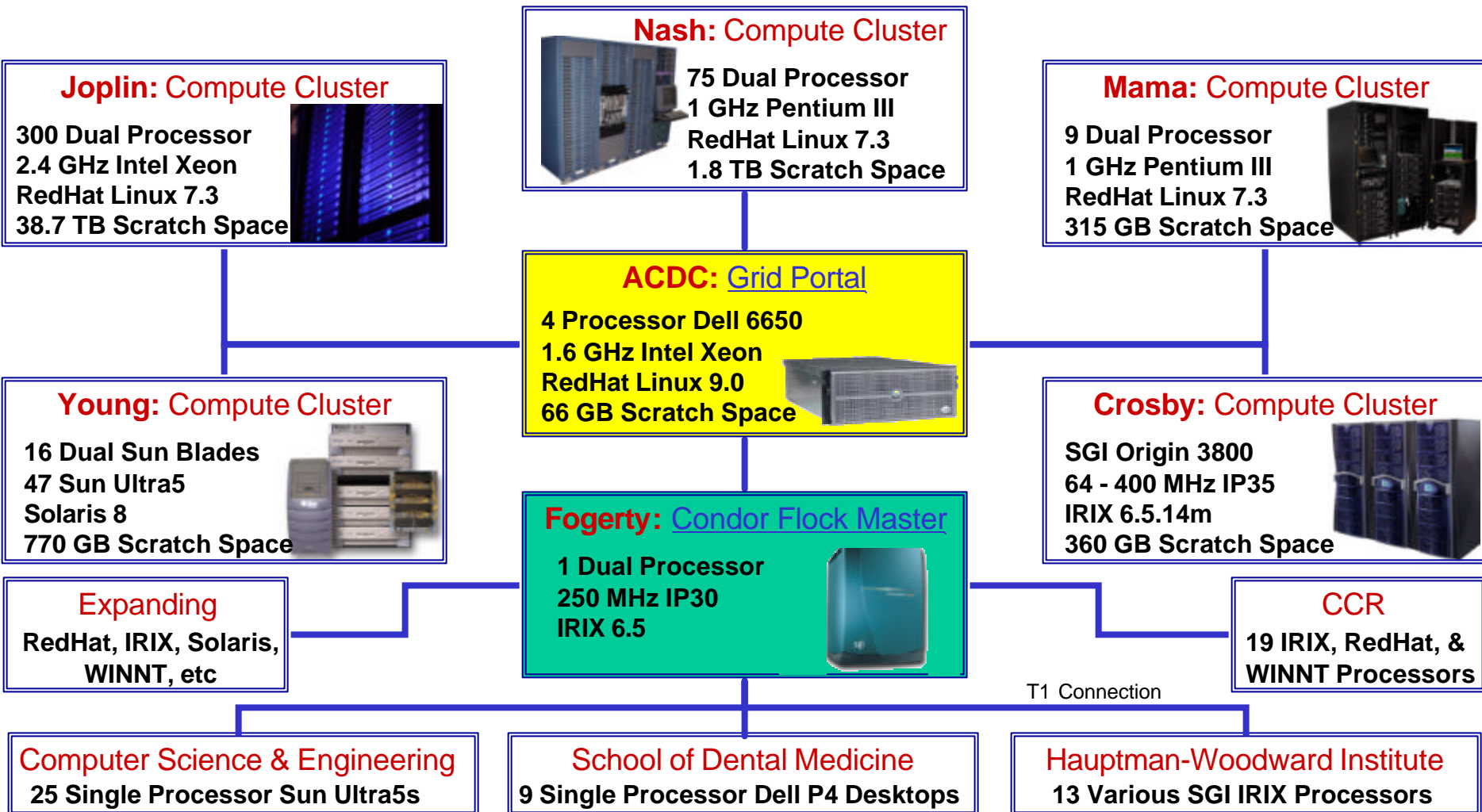
6 TF EV68
71 TB Storage
0.3 TF EV7 shared-memory
150 TB Storage Server



PSC: Compute Intensive

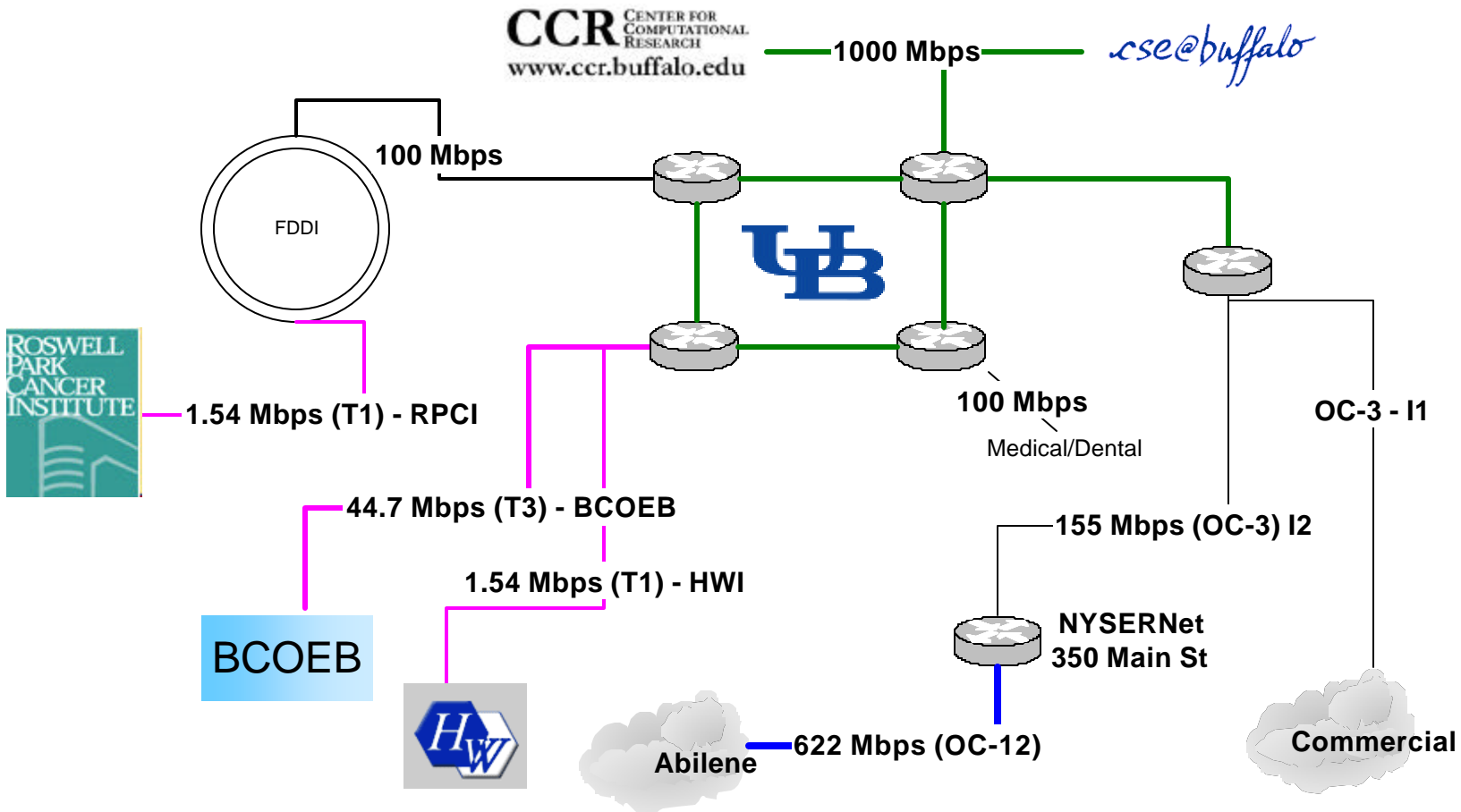
Advanced Computational Data Center

ACDC: Grid Overview



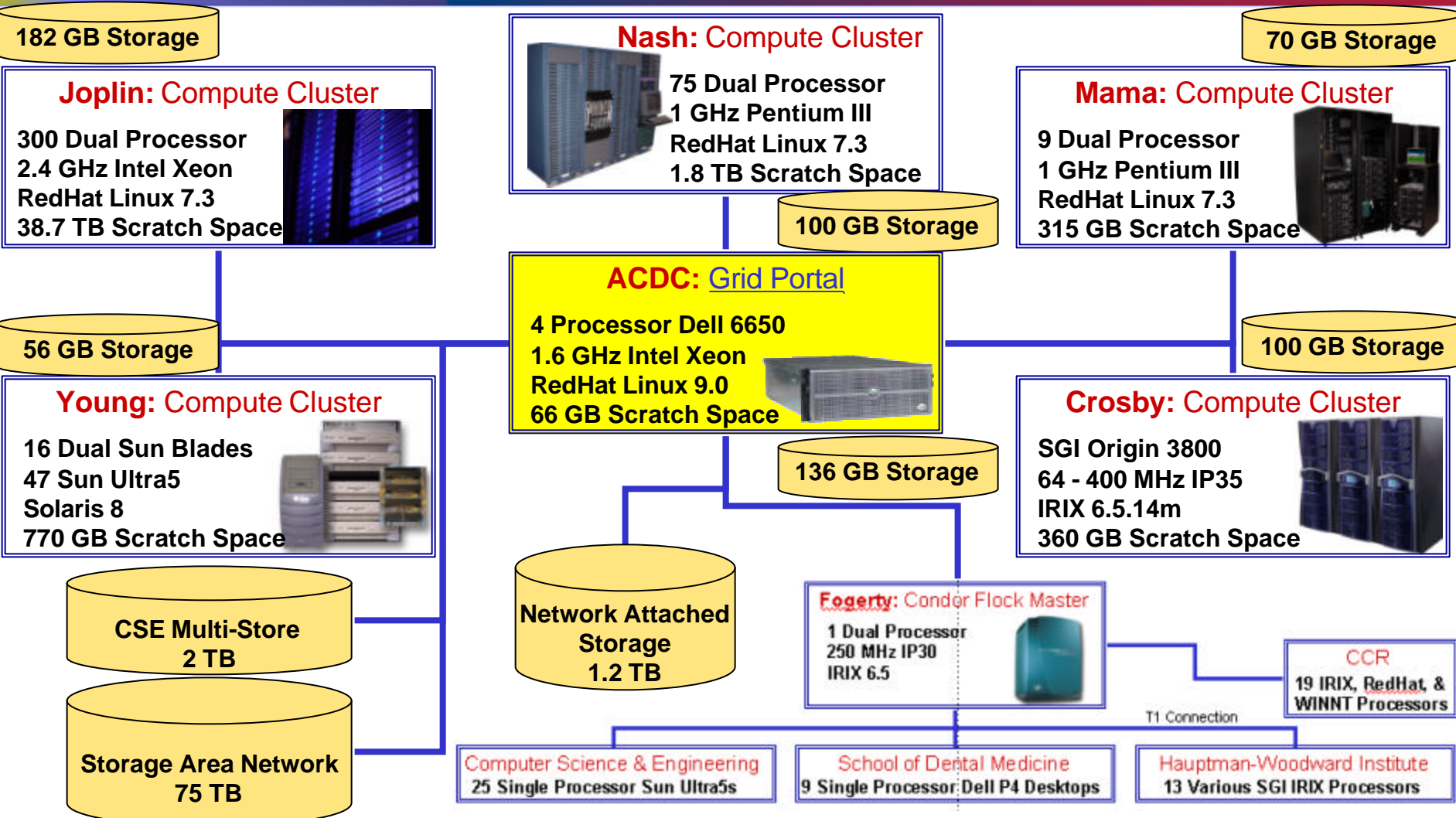
Note: Network connections are 100 Mbps unless otherwise noted.

Network Connections



ACDC Data Grid Overview

(Grid-Available Data Repositories)



Note: Network connections are 100 Mbps unless otherwise noted.



ACDC-Grid

CCR Grid Computing Services - Microsoft Internet Explorer

CCR University at Buffalo The State University of New York

Center for Computational Research GRID PORTAL

High Performance Grid Computing

WELCOME TO GRID COMPUTING SERVICES

University at Buffalo Center for Computational Research is currently forming the first Western New York computational grid. The computational grid consist of many supercomputers located at the Center and several other networked supercomputers throughout the Western New York region. These resources will be shared by many researchers from several departments working on a diverse suite of problems including Biomimetics, Computational Chemistry, and Medical Imaging to name a few.



We also provide grid computing support for the University's Center for Computational Research learning, teaching and research activities plus the infrastructure for both high performance computing and grid enabled software.

Get your "Grid Computing Guide"?

CCR Grid Computing Services Data Management - Microsoft Internet Explorer

CCR University at Buffalo The State University of New York

Center for Computational Research GRID PORTAL

High Performance Grid Computing

VIEW: Group UserList:

replaye

- KeyMaster
- Morpheus
 - Tank
 - Agent
 - Rabbit
 - Tank
 - Morpheus
 - Oracle.m
 - Neo

Browser view of "miller" group files published by user

CCR Grid Computing Services Grid Admin - Microsoft Internet Explorer

CCR University at Buffalo The State University of New York

Center for Computational Research GRID PORTAL

High Performance Grid Computing

View statistics for:

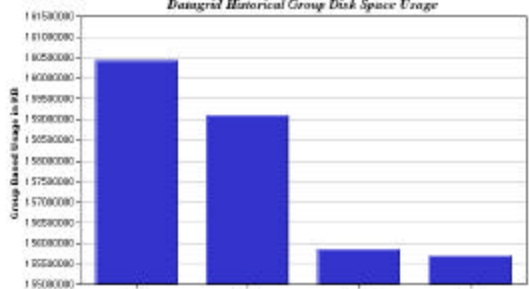
Data based on:

from starting date:

to ending date: inclusive

for: resources:

Baagrid Historical Group Disk Space Usage



Group	Disk Space Usage (KB)
miller	~1,950,000
griddev	~1,850,000
ccrstaff	~1,550,000
mlgreen	~1,500,000

CCR Grid Computing Services: Grid Admin - Microsoft Internet Explorer

CCR University at Buffalo The State University of New York

Center for Computational Research GRID PORTAL

High Performance Grid Computing

View statistics for:

Data based on:

from starting date:

to ending date: inclusive

for: resources:

File_num	File_ID	Filename	Dir_ID	Resource_ID	Owner	Groupname	Type
1	56033	Cypher.txt	52831	10	mlgreen	griddev	txt
2	56034	Cypher.sh	52858	10	mlgreen	griddev	sh
3	56035	Oracle.asc	52958	10	mlgreen	griddev	asc
4	56036	Cypher.sh	52634	10	mlgreen	miller	sh
5	56037	Rabbit.dat	52830	10	mlgreen	ccrstaff	dat
6	56038	Agent.exe	53064	10	mlgreen	griddev	exe
7	56039	Dozer.sh	52852	10	mlgreen	griddev	sh
8	56040	Neo.asc	52187	10	mlgreen	mlgreen	asc
9	56041	Agent.mpg	52833	10	mlgreen	mlgreen	mpg
10	56042	Tank.txt	52188	10	mlgreen	mlgreen	txt
11	56043	Smith.xls	52258	10	mlgreen	ccrstaff	xls
12	56044	KeyMaster.csh	52186	10	mlgreen	miller	csh
13	56045	Oracle.csh	52622	10	mlgreen	griddev	csh
14	56046	Dozer.xls	52808	10	mlgreen	mlgreen	xls
15	56047	Cypher.exe	52204	10	mlgreen	griddev	exe
16	56048	Rabbit.ppt	52861	10	mlgreen	miller	ppt
17	56049	Neo.dat	52217	10	mlgreen	ccrstaff	dat
18	56050	Cypher.asc	53086	10	mlgreen	griddev	asc

ACDC-Grid Administration

CCR Grid Computing Services: Grid Admin - Microsoft: Internet Explorer

Center for Computational Research **GRID PORTAL**
High Performance Grid Computing

Grid Site Administration

PORTAL LOGOUT
User Tools
Manage Account
Grid General Info
Projects
Resources
Computational Grid
Job Submission
Job/Queue Status
Data Grid
Data Grid Statistics
Network Status
Running/Queued Jobs
PBS Job History
Grid Portal Statistics
Center Fleck Statistics
User Information
Education/Outreach
Staff Only
CCR HOME

Users
Groups
Portal Event Log
Database Job List

Organizations (add, edit, delete)
Resources (view, refresh, ping, delete, create host certificate)

Globus Administration
Reports (machine usage, user access to machines, etc.)

Generate Globus grid-mapfile

Specifying an optional include file will cause the contents of this file to be included at the top of the generated grid-mapfile. If a grid-mapfile path is specified a copy of the generated file will be saved into this location. The generated file will be staged to the grid nodes unless the box is checked.

Optional include file:
Optional grid-mapfile path:
 Do not stage this file to the grid nodes

CCR Grid Computing Services: Database Job Admin - Microsoft: Internet Explorer

Center for Computational Research **GRID PORTAL**
High Performance Grid Computing

Create New Database Job

Create a new database job that can be run by the portal. Job scripts must reside in `home/griddev/www/jobscripts` prior to creating the database job entry.

Job Name:
Full Path To Script:
Accepts Arguments:
Run Script:
Run As User:

CCR Grid Computing Services: Grid Admin - Resources - Microsoft: Internet Explorer

Center for Computational Research **GRID PORTAL**
High Performance Grid Computing

MDS Resource Update Status

Current Time: 16-September-2003 10:59:12

Resource	Last Updated	Next Update	Status
crasby.ccr.buffalo.edu	16-September-2003 09:15:30	2 minutes	OK
rogerty.ccr.buffalo.edu	16-September-2003 10:45:30	2 minutes	OK
joplin.ccr.buffalo.edu	16-September-2003 10:45:15	2 minutes	OK
mama.ccr.buffalo.edu	16-September-2003 10:45:15	2 minutes	OK
nash.ccr.buffalo.edu	16-September-2003 10:45:15	2 minutes	OK
newus.hwi.buffalo.edu	16-September-2003 10:45:20	2 minutes	OK
yardbirds.ccr.buffalo.edu	16-September-2003 10:45:13	2 minutes	OK
young.ccr.buffalo.edu	16-September-2003 10:45:27	2 minutes	OK

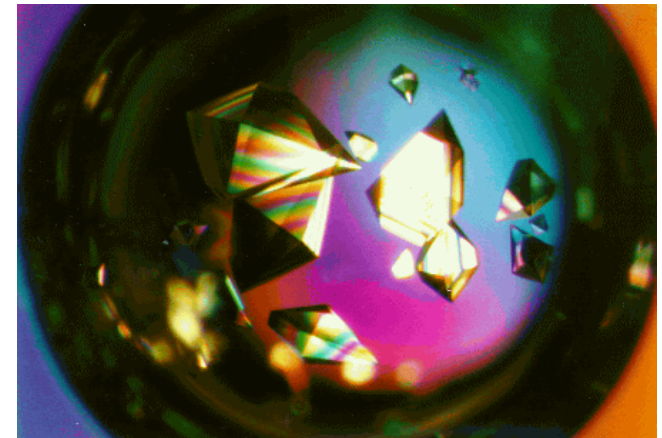
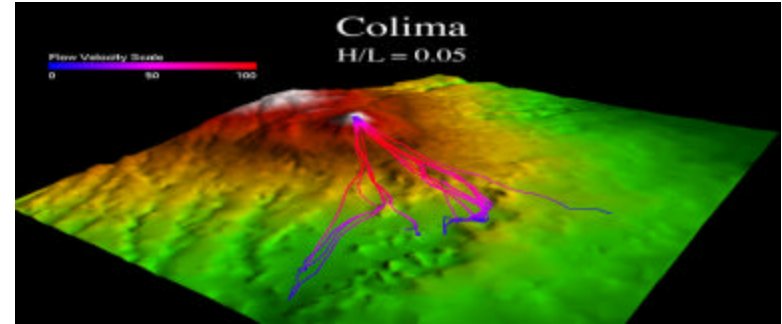
Return to the Grid Resource Admin menu.
Return to the Grid Admin menu.

Advanced
Center for Computational Research
Data
Center



Grid-Enabling Application Templates

- Structural Biology
- Earthquake Engineering
- Pollution Abatement
- Geographic Information Systems & BioHazards



ACDC-Grid Cyber-Infrastructure

■ Predictive Scheduler

- Define quality of service estimates of job completion, by better estimating job runtimes by profiling users.

■ Data Grid

- Automated Data File Migration based on profiling users.

■ High-performance Grid-enabled Data Repositories

- Develop automated procedures for dynamic data repository creation and deletion.

■ Dynamic Resource Allocation

- Develop automated procedures for dynamic computational resource allocation.

ACDC-Grid Collaborations

- High-Performance Networking Infrastructure
- WNY Grid Initiative
- Grid3+ Collaboration
- iVDGL Member
- Open Science Grid Member
- Grid-Based Visualization
 - SGI Collaboration
- Grid-Lite
 - HP Labs Collaboration
- Innovative Laboratory Prototype
 - Dell Collaboration



Grid Services and Applications

**ACDC-Grid
Computational
Resources**

Applications

Shake-and-Bake

Apache

MySQL

Oracle

High-level Services and Tools

Globus
Toolkit

NWS

MPI

MPI-IO

C, C++, Fortran, PHP

globusrun

Core Services

Metacomputing
Directory
Service

Globus
Security
Interface

GRAM

GASS

Local Services

Condor

Stork

MPI

RedHat Linux

WINNT

LSF

PBS

Maui Scheduler

TCP

UDP

Irix

Solaris

**ACDC-Grid
Data
Resources**

Adapted from Ian Foster and Carl Kesselman



Middleware

- **Globus Toolkit 2.2.4 ® direct upgrade WSRF**
- **Condor 6.6.0**
- **Network Weather Service 2.6**
- **Apache2 HTTP Server**
- **PHP 4.3.0**
- **MySQL 3.23**
- **phpMyAdmin 2.5.1**



www.ccr.buffalo.edu