GT'04 Panel: Storage Considerations for Grid Computing Environments

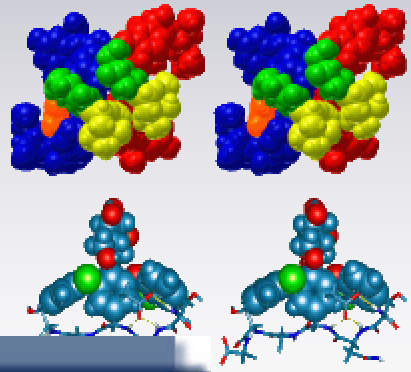
Russ Miller & Mark Green

Center for Computational Research &
Computer Science & Engineering
SUNY-Buffalo

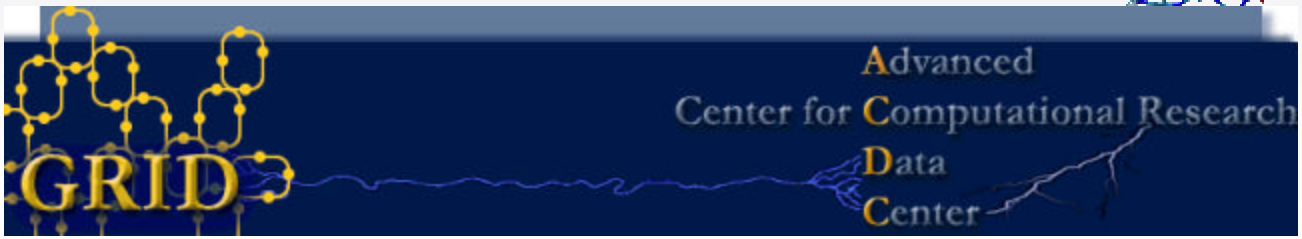
Hauptman-Woodward Medical Inst

NSF, NIH, DOE, NYS

Grid3 2003



SnB



University at Buffalo
The State University of New York

Major CCR Resources (12TF & 290TB)

- Dell Linux Cluster: #22 ® #25 ® #38
  - 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
  - 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

Advanced CCR Data Center (ACDC) Computational Grid Overview

Nash: Compute Cluster

75 Dual Processor
1 GHz Pentium III
RedHat Linux 7.3
1.8 TB Scratch Space

Joplin: Compute Cluster

300 Dual Processor
2.4 GHz Intel Xeon
RedHat Linux 7.3
38.7 TB Scratch Space

Mama: Compute Cluster

9 Dual Processor
1 GHz Pentium III
RedHat Linux 7.3
315 GB Scratch Space

ACDC: Grid Portal

4 Processor Dell 6650
1.6 GHz Intel Xeon
RedHat Linux 9.0
66 GB Scratch Space

Young: Compute Cluster

16 Dual Sun Blades
47 Sun Ultra5
Solaris 8
770 GB Scratch Space

Crosby: Compute Cluster

SGI Origin 3800
64 - 400 MHz IP35
IRIX 6.5.14m
360 GB Scratch Space

Fogerty: Condor Flock Master

1 Dual Processor
250 MHz IP30
IRIX 6.5

Expanding

RedHat, IRIX, Solaris, WINNT, etc

CCR

19 IRIX, RedHat, & WINNT Processors

T1 Connection

Computer Science & Engineering

25 Single Processor Sun Ultra5s

School of Dental Medicine

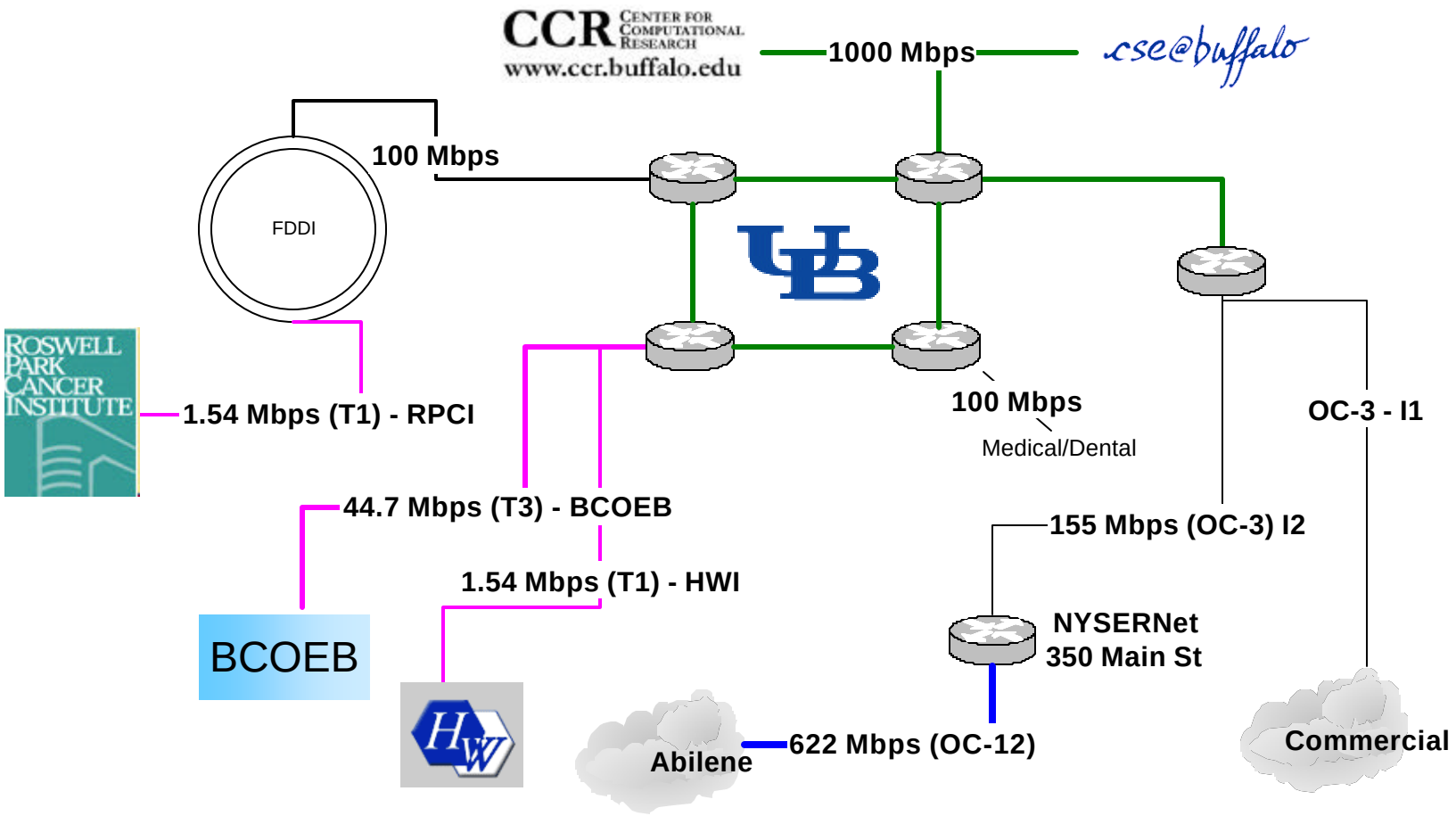
9 Single Processor Dell P4 Desktops

Hauptman-Woodward Institute

13 Various SGI IRIX Processors

Note: Network connections are 100 Mbps unless otherwise noted.

Network Connections



ACDC Data Grid Overview
(Grid-Available Data Repositories)

182 GB Storage

Joplin: Compute Cluster

300 Dual Processor
2.4 GHz Intel Xeon
RedHat Linux 7.3
38.7 TB Scratch Space

Nash: Compute Cluster

75 Dual Processor
1 GHz Pentium III
RedHat Linux 7.3
1.8 TB Scratch Space

100 GB Storage

70 GB Storage

Mama: Compute Cluster

9 Dual Processor
1 GHz Pentium III
RedHat Linux 7.3
315 GB Scratch Space

ACDC: Grid Portal

4 Processor Dell 6650
1.6 GHz Intel Xeon
RedHat Linux 9.0
66 GB Scratch Space

56 GB Storage

100 GB Storage

Young: Compute Cluster

16 Dual Sun Blades
47 Sun Ultra5
Solaris 8
770 GB Scratch Space

136 GB Storage

Crosby: Compute Cluster

SGI Origin 3800
64 - 400 MHz IP35
IRIX 6.5.14m
360 GB Scratch Space

CSE Multi-Store
2 TB

Storage Area Network
75 TB

Network Attached Storage
1.2 TB

Fogerty: Condor Flock Master

1 Dual Processor
250 MHz IP30
IRIX 6.5

CCR
19 IRIX, RedHat, & WINNT Processors

T1 Connection

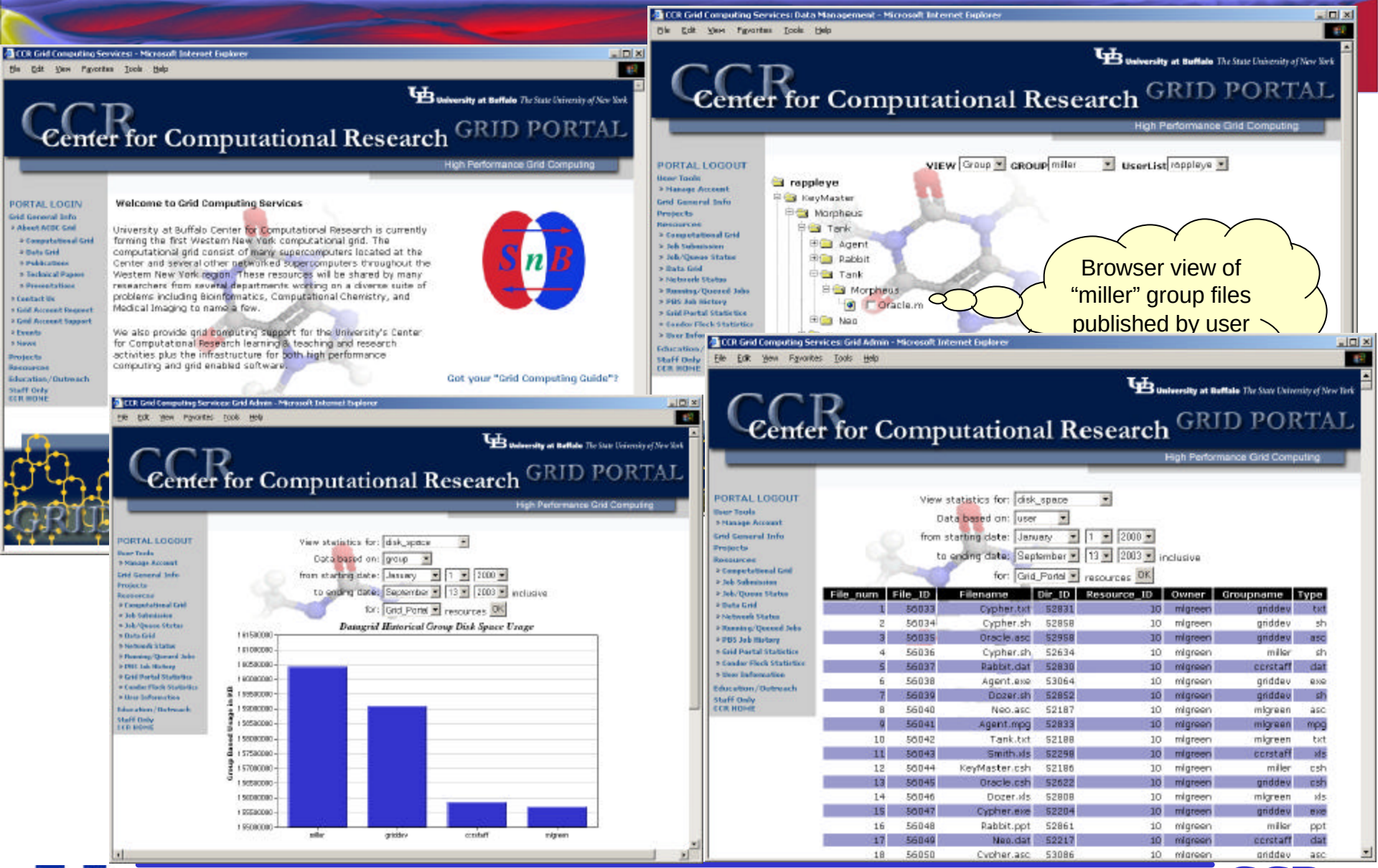
Computer Science & Engineering
25 Single Processor Sun Ultra5s

School of Dental Medicine
9 Single Processor Dell P4 Desktops

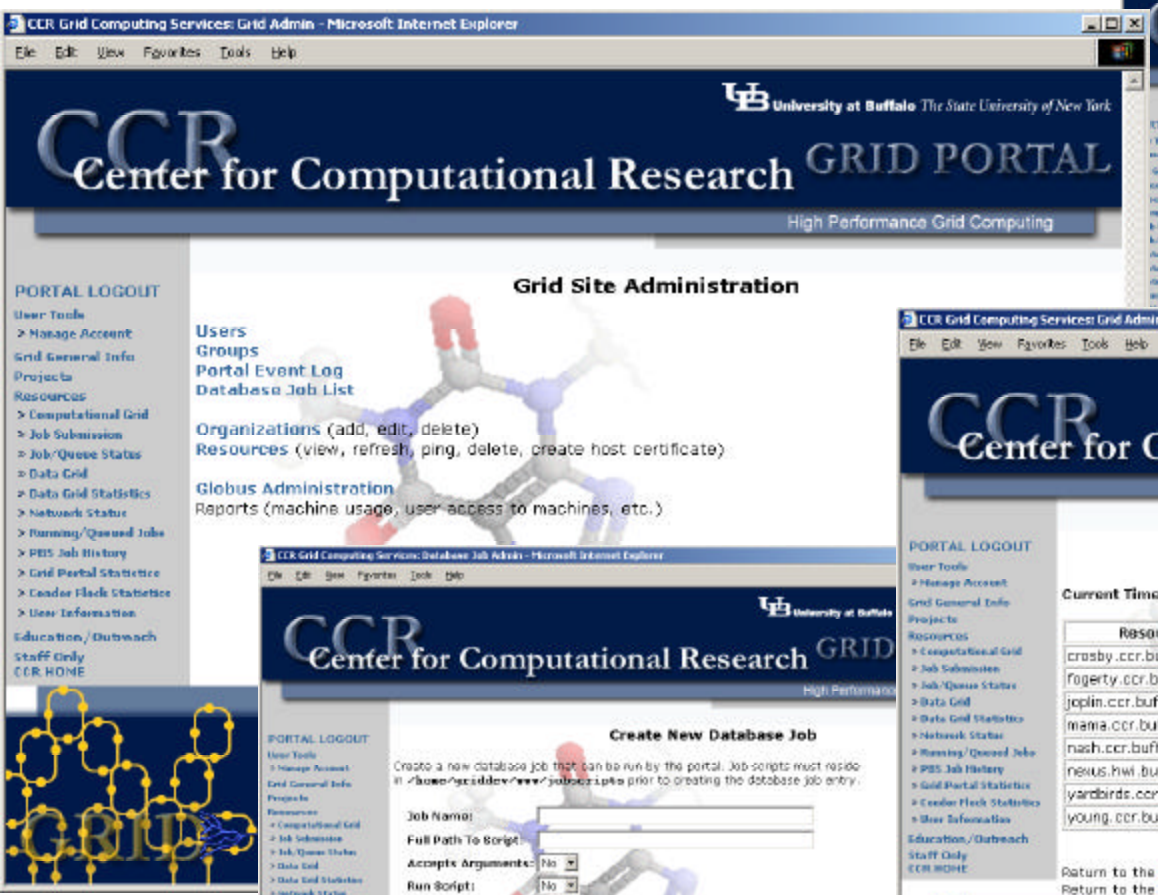
Hauptman-Woodward Institute
13 Various SGI IRIX Processors

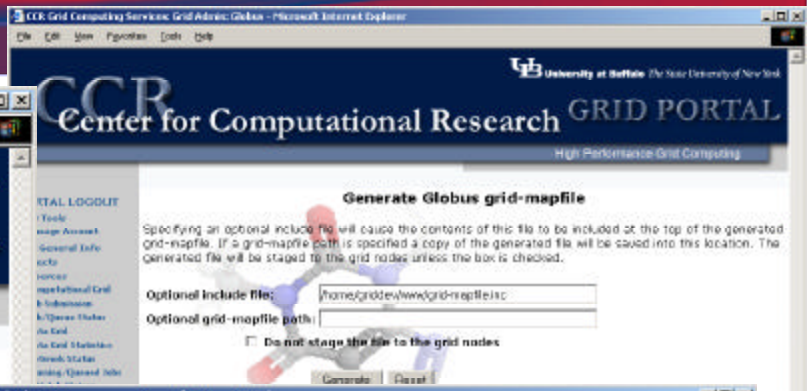
Note: Network connections are 100 Mbps unless otherwise noted.

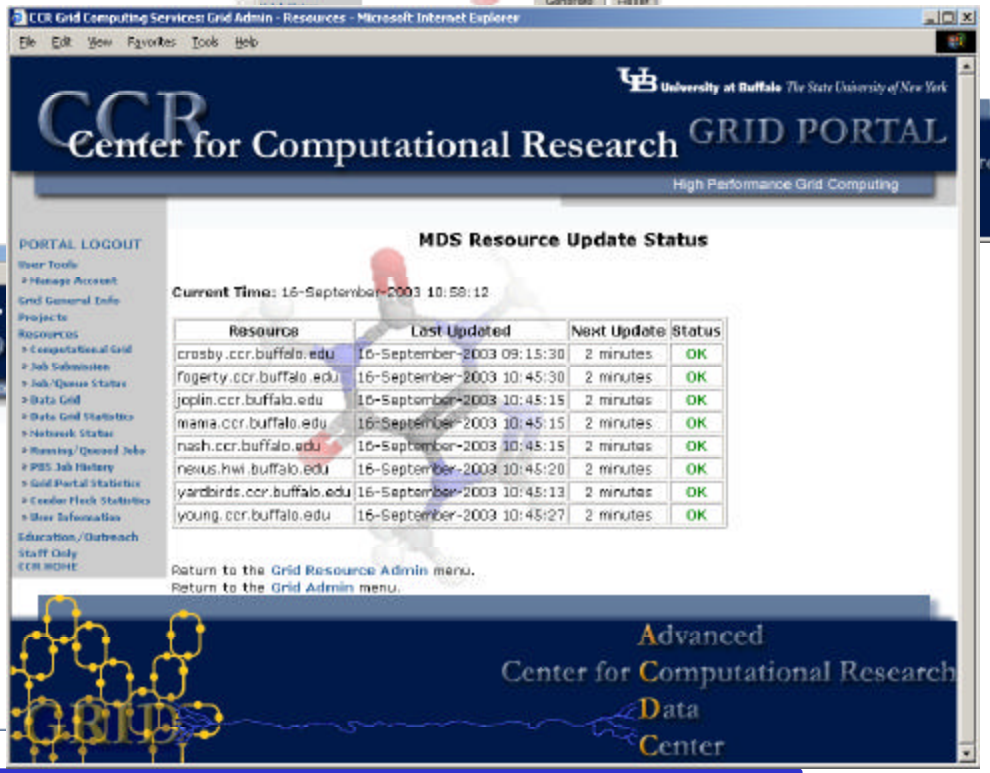
ACDC-Grid


CCR Center for Computational Research GRID PORTAL
High Performance Grid Computing
Welcome to Grid Computing Services
PORTAL LOGIN
PORTAL LOGOUT
Browser view of "miller" group files published by user
Daugrid Historical Group Disk Space Usage
File_num File_ID Filename Dir_ID Resource_ID Owner Groupname Type

ACDC-Grid Administration

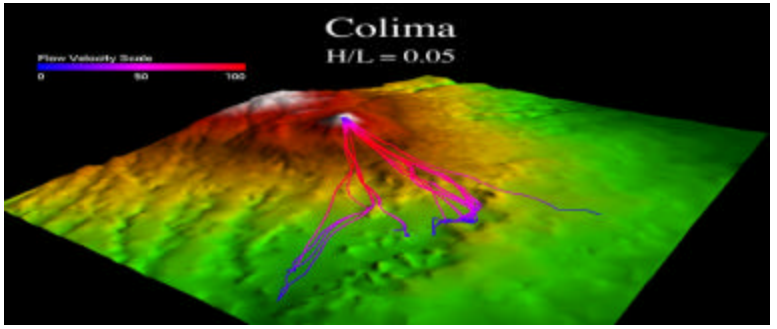


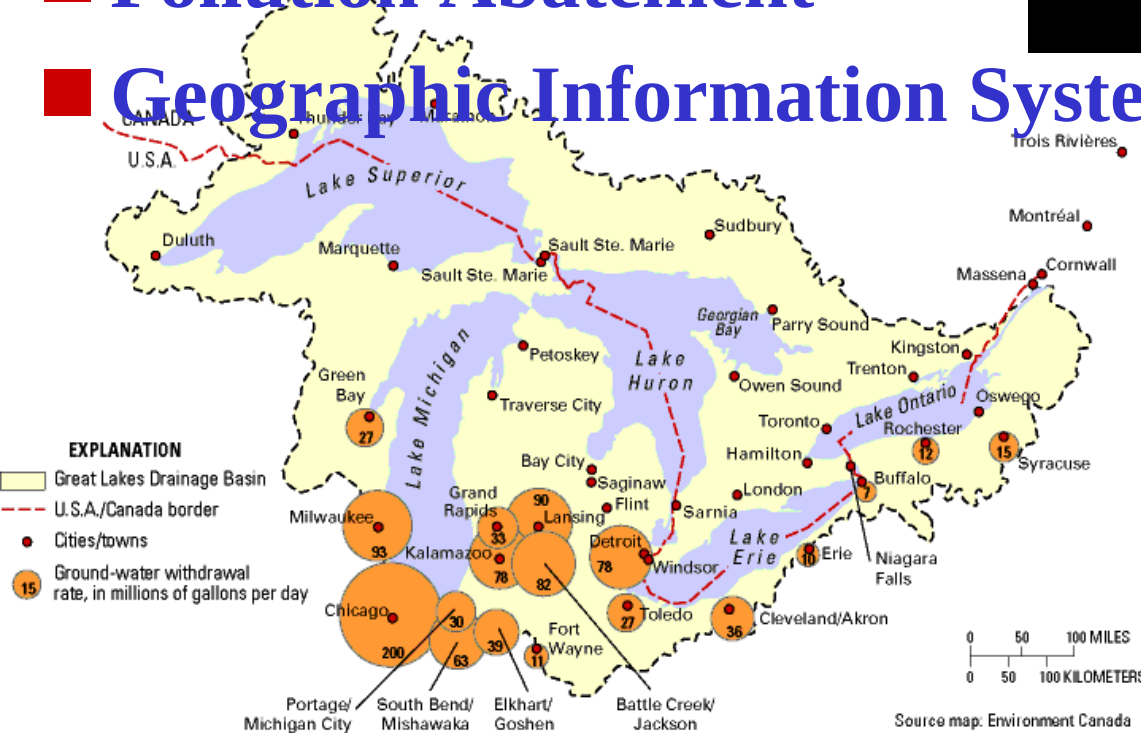


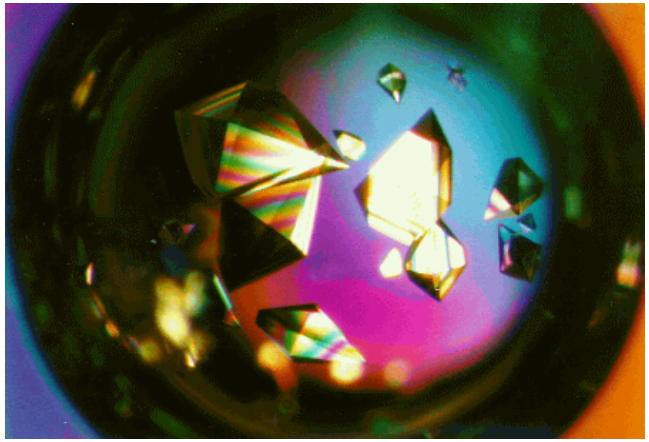


Grid-enabling Application Templates

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







ACDC-Grid Collaborations

- **Advanced Computational Data Center – Grid (ACDC-Grid)**
- **Innovative Laboratory Prototype**
- **Grid3+ Collaboration**
- **High-Performance Networking Infrastructure**
- **HP Labs Collaboration**
- **IBM (under discussion)**
- **Open Science Grid (Future)**

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.**

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