

# ***Shaking-and-Baking on a Grid***

**Russ Miller & Mark Green**

**Center for Computational Research, SUNY-Buffalo  
Hauptman-Woodward Medical Inst**



**NSF ITR  
ACI-02-04918**

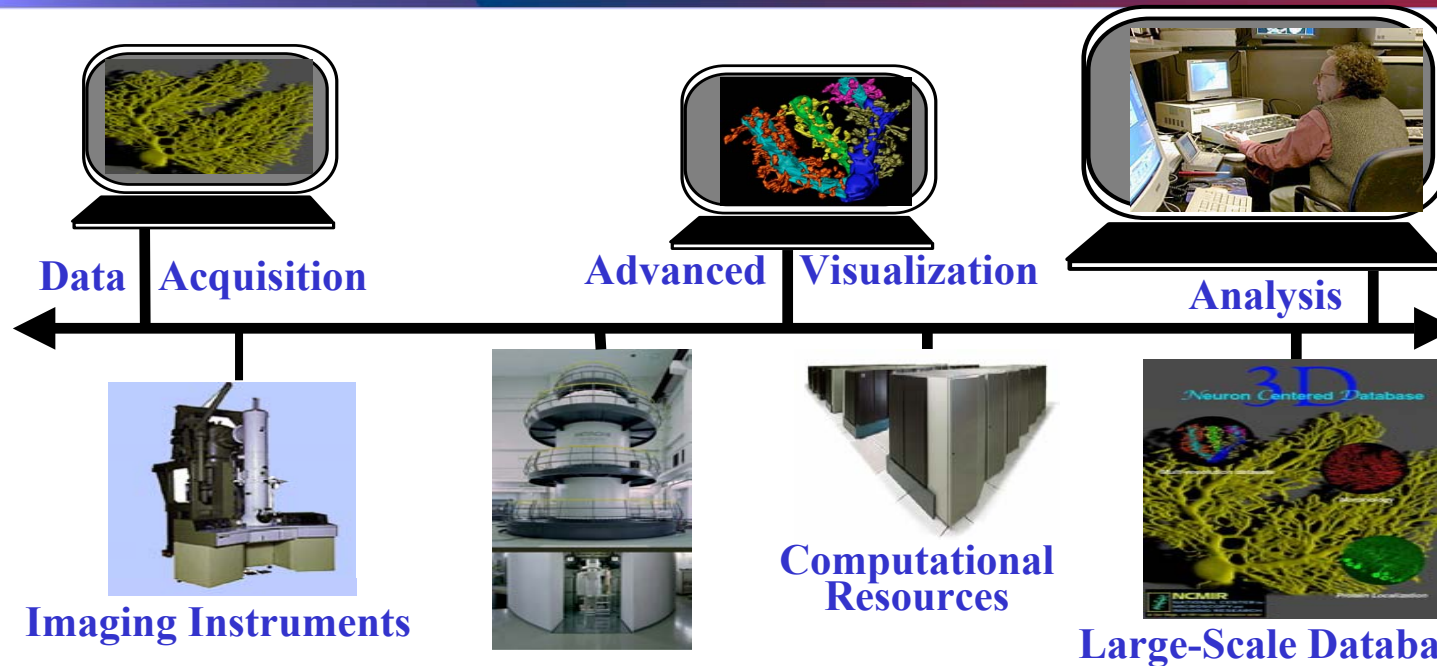


**University at Buffalo**  
*The State University of New York*

# Grid Computing



# 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



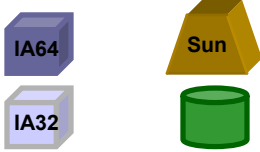
# Computational Grids & Electric Power Grids

- **Similarities/Goals of CG and EPG**
  - Ubiquitous
  - Consumer is comfortable with lack of knowledge of details
- **Differences Between CG and EPG**
  - Wider spectrum of performance & services
  - Access governed by more complicated issues
    - Security
    - Performance
    - Socio-political factors

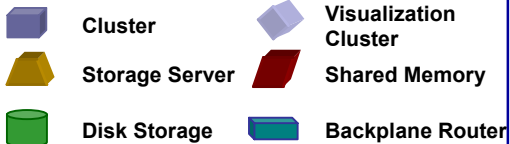
# NSF Extensible TeraGrid Facility

## Caltech: Data collection analysis

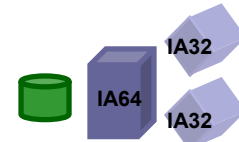
0.4 TF IA-64  
IA32 Datawulf  
80 TB Storage



## LEGEND

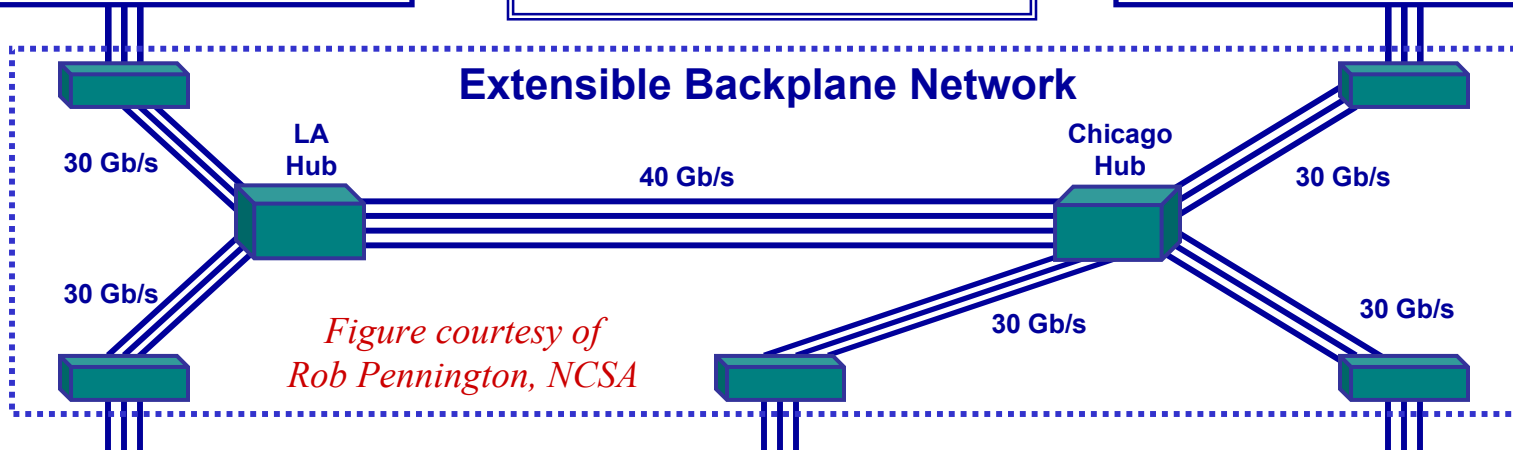


## ANL: Visualization

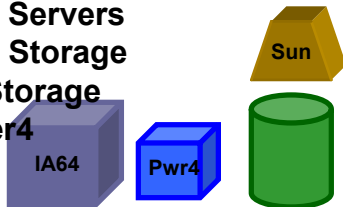


1.25 TF IA-64  
96 Viz nodes  
20 TB Storage

## Extensible Backplane Network

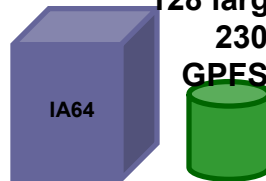


4 TF IA-64  
DB2, Oracle Servers  
500 TB Disk Storage  
6 PB Tape Storage  
1.1 TF Power4



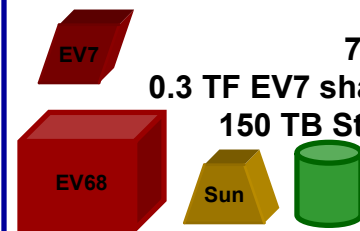
**SDSC: Data Intensive**

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**

# Critical Resources: WNY Computational & Data Grids

- Computational & Data Resources (CCR)

- 10TF Computing & 78TB Storage

- Instruments (HWI, RPCI)

- High-Throughput Crystallization Laboratory

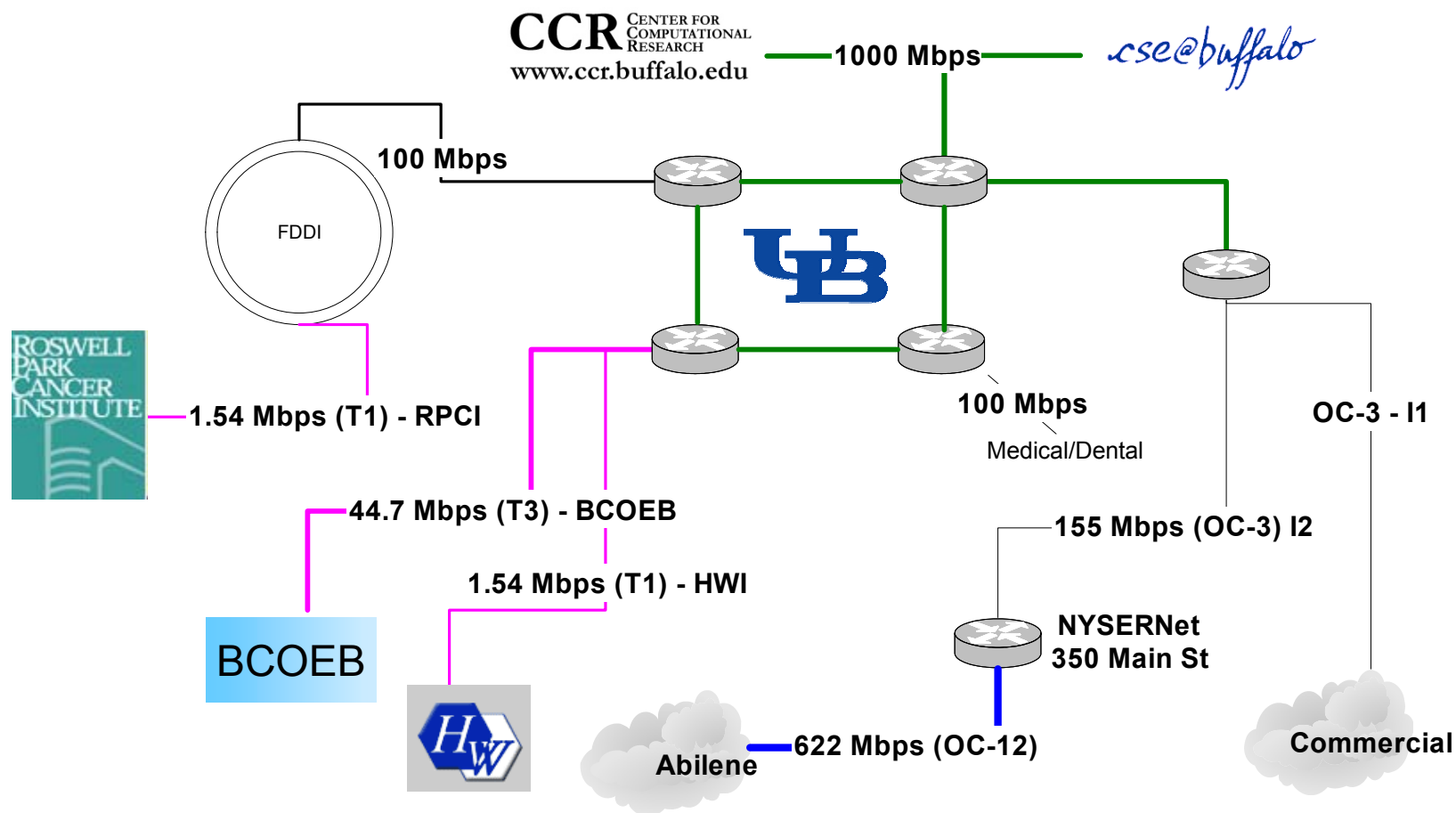
- Data Generation (HWI)

- 7TB per year

- Databases (UB-N, UB-S, BGH, CoE)

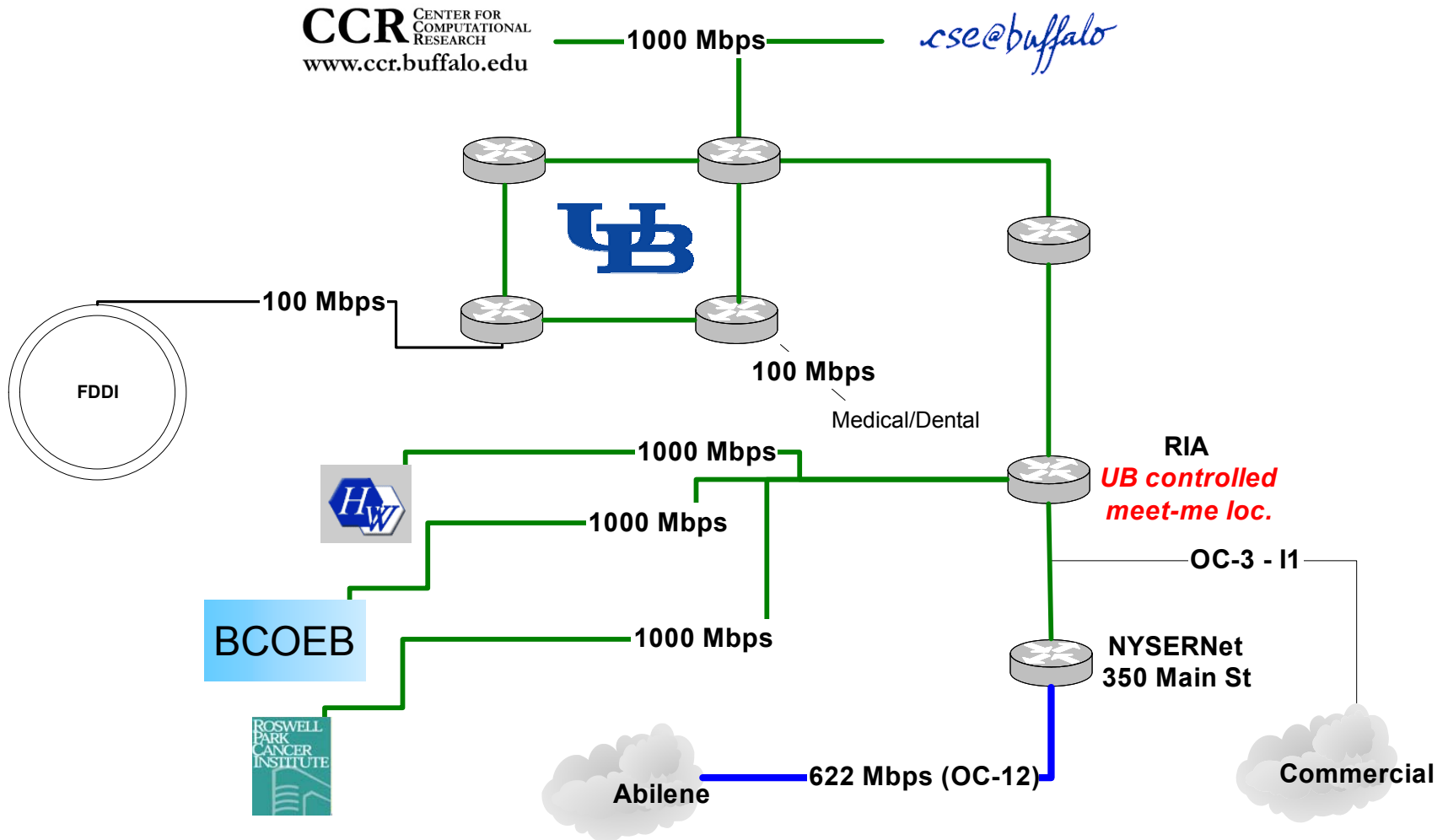
- *SnB*

# Network Connections





# Network Connections (New)



# Advanced CCR Data Center (ACDC) Computational Grid Overview

## 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

## 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



## Expanding

RedHat, IRIX, Solaris,  
WINNT, etc

## Fogerty: Condor Flock Master

1 Dual Processor  
250 MHz IP30  
IRIX 6.5



## CCR

19 IRIX, RedHat, &  
WINNT Processors

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

T1 Connection

Note: Network connections are 100 Mbps unless otherwise noted.



University at Buffalo

The State University of New York

Center for Computational Research

CCR

# ACDC Data Grid Overview

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

70 GB Storage

## Mama: Compute Cluster

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



100 GB Storage

## ACDC: Grid Portal

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



100 GB Storage

## Crosby: Compute Cluster

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



136 GB Storage

## Young: Compute Cluster

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



56 GB Storage

CSE Multi-Store  
2 TB

Storage Area Network  
75 TB

Network Attached  
Storage  
480 GB

## Fogarty: 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.



University at Buffalo

The State University of New York

Center for Computational Research

CCR

# WNY Grid Highlights

- Heterogeneous Computational & Data Grid
- Currently in Beta with *Shake-and-Bake*
- WNY Release in March
- Bottom-Up General Purpose Implementation
  - Ease-of-Use User Tools
  - Administrative Tools
- Back-End Intelligence
  - Backfill Operations
  - Prediction and Analysis of Resources to Run Jobs (Compute Nodes + Requisite Data)

# Grid-Based *SnB* Objectives

- Install Grid-Enabled Version of *SnB*
- Job Submission and Monitoring over Internet
- *SnB* Output Stored in Database
- *SnB* Output Mined through Internet-Based Integrated Querying Tool
- Serve as Template for Chem-Grid & Bio-Grid
- Experience with Globus and Related Tools



# Grid Enabled *SnB* Execution

## □ User

- defines Grid-enabled *SnB* job using Grid Portal or *SnB*
- supplies location of data files from Data Grid
- supplies *SnB* mode of operation

## □ Grid Portal

- assembles required *SnB* data and supporting files, execution scripts, database tables.
- determines available ACDC-Grid resources.

## □ ACDC-Grid job management includes:

- automatic determination of appropriate execution times, number of trials, and number/location of processors,
- logging/status of concurrently executing resource jobs, &
- automatic incorporation of *SnB* trial results into the molecular structure database.

# ACDC-Grid Portal

CCR Grid Computing Services - Microsoft Internet Explorer

File Edit View Favorites Tools Help

University at Buffalo The State University of New York

## CCR Center for Computational Research GRID PORTAL

High Performance Grid Computing

**PORTAL LOGIN**

- Grid General Info
  - » About ACDC Grid
    - » Computational Grid
    - » Data Grid
    - » Publications
    - » Technical Papers
    - » Presentations
  - » Contact Us
  - » Grid Account Request
  - » Grid Account Support
  - » Events
  - » News
- Projects
- Resources
- Education/Outreach
- Staff Only
- CCR HOME

### 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 Bioinformatics, 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.

Got your "Grid Computing Guide"?

Do you want to learn about 'Grid Computing'?

Advanced Center for Computational Research Data Center

# ACDC-Grid Portal Login

CCR Grid Computing Services: - Microsoft Internet Explorer

File Edit View Favorites Tools Help

UB University at Buffalo The State University of New York

## CCR Center for Computational Research GRID PORTAL

High Performance Grid Computing

**PORTAL LOGIN**

Grid General Info

- » About ACDC Grid
  - » Computational Grid
  - » Data Grid
  - » Publications
  - » Technical Papers
  - » Presentations
- » Contact Us
- » Grid Account Request
- » Grid Account Support
- » Events
- » News

Projects

Resources

Education/Outreach

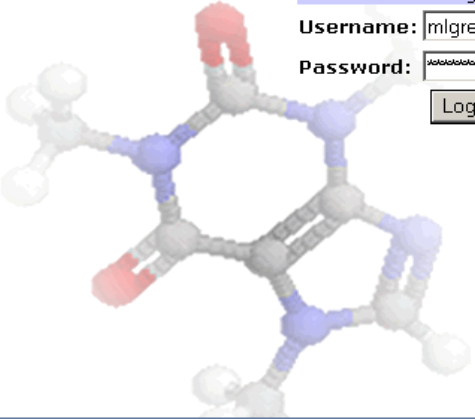
Staff Only

CCR HOME

**Login**

Username:

Password:



Grid Portal login screen

Advanced  
Center for Computational Research  
Data  
Center

# Data Grid Capabilities

CCR Grid Computing Services: Data Management - Microsoft Internet Explorer

File Edit View Favorites Tools Help

University at Buffalo The State University of New York

CCR Center for Computational Research GRID PORTAL

High Performance Grid Computing

PORTAL LOGOUT

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

VIEW User

USER

- mlgreen
  - Dozer
  - Morpheus
    - Agent
      - Oracle
        - Morpheus.sh
    - Cypher
    - KeyMaster
    - Oracle
  - Neo
  - Oracle
  - Smith
  - Tank

Browser view of "mlgreen" user files stored in the Data Grid

Advanced Center for Computational Research Data Center



# Data Grid Capabilities

CCR Grid Computing Services: Data Management - Microsoft Internet Explorer

File Edit View Favorites Tools Help

University at Buffalo The State University of New York

## CCR GRID PORTAL

Center for Computational Research

High Performance Grid Computing

PORTAL LOGOUT

User Tools

- » Manage Account

Grid General Info

Projects

Resources

- » Computational Grid
- » Job Submission
- » Job/Queue Status
- » Data Grid
- » Network Status
- » Running/Queued Jobs
- » PBS Job History
- » Grid Portal Statistics
- » Condor Flock Statistics
- » User Information

Education/Outreach

Staff Only

CCR HOME

VIEW  GROUP  UserList

Folder structure:

- Folder: rappleye
  - Folder: KeyMaster
  - Folder: Morpheus
    - Folder: Tank
      - Folder: Agent
      - Folder: Rabbit
      - Folder: Tank
        - Folder: Morpheus
          - File: Oracle.m
  - Folder: Neo
  - Folder: Cypher
  - Folder: Neo
  - Folder: Morpheus
  - Folder: Oracle

Browser view of "miller" group files published by user "rappleye"

Advanced Center for Computational Research Data





# Grid Portal Job Status

■ Grid-enabled jobs can be monitored using the Grid Portal web interface dynamically.

□ Charts are based on:

- total CPU hours, or
- total jobs, or
- total runtime.

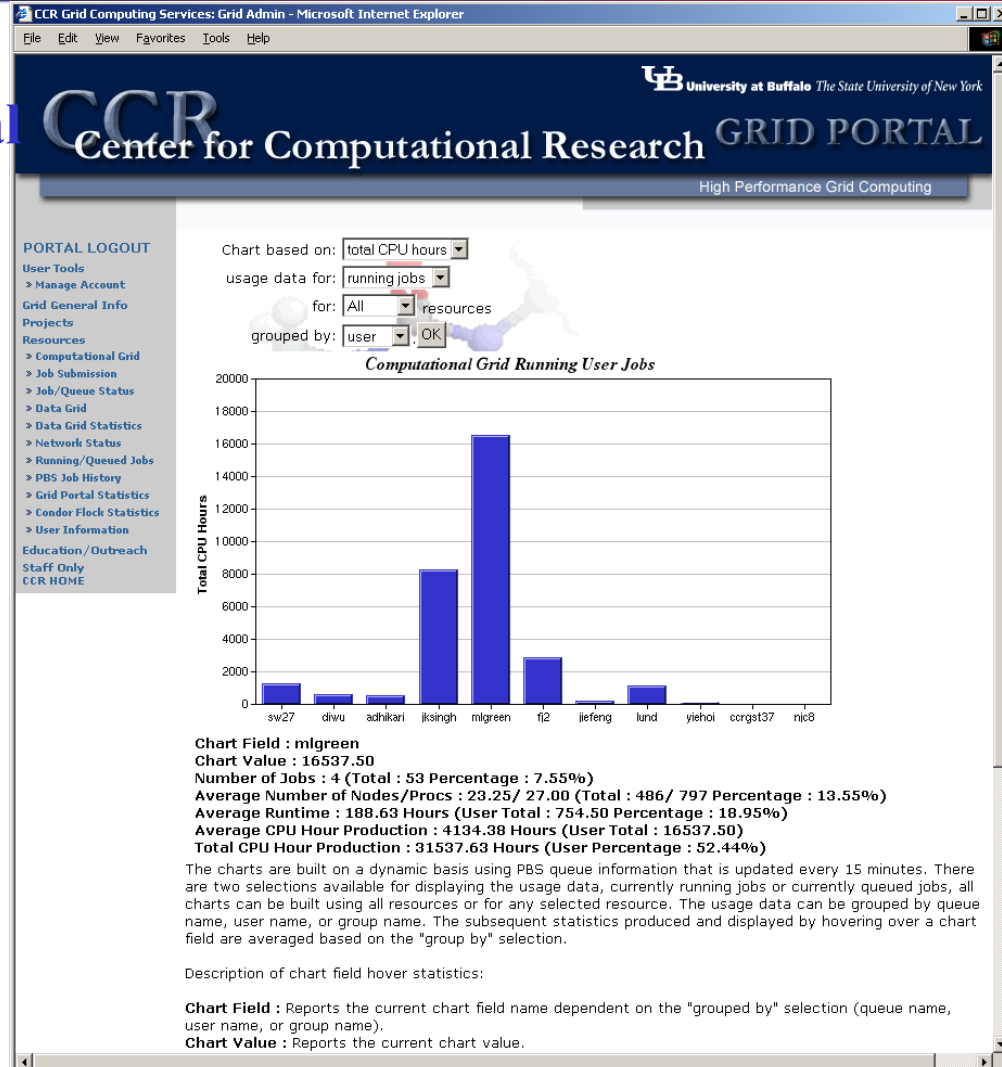
□ Usage data for:

- running jobs, or
- queued jobs.

□ Individual or all resources.

□ Grouped by:

- group, or
- user, or
- queue.



# Grid Portal Job Status

CCR Grid Computing Services: Grid Admin - Microsoft Internet Explorer

File Edit View Favorites Tools Help

CCR University at Buffalo The State University of New York  
Center for Computational Research GRID PORTAL

High Performance Grid Computing

## 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  
> Condo Flock Statistics  
> User Information  
Education/Outreach  
Staff Only  
CCR HOME

Chart based on:   
usage data for:   
for:  resources  
grouped by:  OK

Computational Grid Queued User Jobs

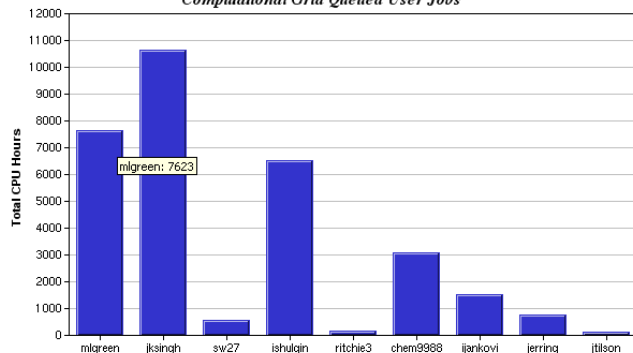


Chart Field : mlgreen

Chart Value : 7623.17

Number of Jobs : 27 (Total : 60 Percentage : 45.00%)

Average Number of Nodes/Procs : 1.96/ 3.93 (Total : 1641/ 3206 Percentage : 3.31%)

Average Runtime : 71.92 Hours (User Total : 1941.75 Percentage : 71.90%)

Average CPU Hour Production : 282.34 Hours (User Total : 7623.17)

Total CPU Hour Production : 31015.17 Hours (User Percentage : 24.58%)

The charts are built on a dynamic basis using PBS queue information that is updated every 15 minut are two selections available for displaying the usage data, currently running jobs or currently queue charts can be built using all resources or for any selected resource. The usage data can be grouped name, user name, or group name. The subsequent statistics produced and displayed by hovering ov field are averaged based on the "group by" selection.

Description of chart field hover statistics:

**Chart Field :** Reports the current chart field name dependent on the "grouped by" selection (queue user name, or group name).

**Chart Value :** Reports the current chart value.

CCR Grid Computing Services: Grid Admin - Microsoft Internet Explorer

File Edit View Favorites Tools Help

CCR University at Buffalo The State University of New York  
Center for Computational Research GRID PORTAL

High Performance Grid Computing

## 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  
> Condo Flock Statistics  
> User Information  
Education/Outreach  
Staff Only  
CCR HOME

Chart based on: total CPU hours

usage data for: queued jobs

for: All resources

grouped by: user mlgreen

| Job_Num | Job_Id | Username | Groupname | Queue | Node_Cnt | Proc_Cnt | Rank | Request_Time(hr) | Start_Time(hr) |
|---------|--------|----------|-----------|-------|----------|----------|------|------------------|----------------|
| 1       | 23364  | mlgreen  | miller    | grid  | 1        | 2        | 1    | 71.916666666667  | queued         |
| 2       | 23365  | mlgreen  | miller    | grid  | 1        | 2        | 2    | 71.916666666667  | queued         |
| 3       | 23366  | mlgreen  | miller    | grid  | 2        | 4        | 3    | 71.916666666667  | queued         |
| 4       | 23367  | mlgreen  | miller    | grid  | 2        | 4        | 4    | 71.916666666667  | queued         |
| 5       | 23368  | mlgreen  | miller    | grid  | 2        | 4        | 5    | 71.916666666667  | queued         |
| 6       | 23369  | mlgreen  | miller    | grid  | 3        | 6        | 6    | 71.916666666667  | queued         |
| 7       | 23370  | mlgreen  | miller    | grid  | 2        | 4        | 7    | 71.916666666667  | queued         |
| 8       | 23371  | mlgreen  | miller    | grid  | 2        | 4        | 8    | 71.916666666667  | queued         |
| 9       | 23372  | mlgreen  | miller    | grid  | 2        | 4        | 9    | 71.916666666667  | queued         |
| 10      | 23373  | mlgreen  | miller    | grid  | 2        | 4        | 10   | 71.916666666667  | queued         |
| 11      | 23374  | mlgreen  | miller    | grid  | 2        | 4        | 11   | 71.916666666667  | queued         |
| 12      | 23375  | mlgreen  | miller    | grid  | 2        | 4        | 12   | 71.916666666667  | queued         |
| 13      | 23376  | mlgreen  | miller    | grid  | 2        | 4        | 13   | 71.916666666667  | queued         |
| 14      | 23377  | mlgreen  | miller    | grid  | 2        | 4        | 14   | 71.916666666667  | queued         |
| 15      | 23378  | mlgreen  | miller    | grid  | 2        | 4        | 15   | 71.916666666667  | queued         |
| 16      | 23379  | mlgreen  | miller    | grid  | 2        | 4        | 16   | 71.916666666667  | queued         |
| 17      | 23380  | mlgreen  | miller    | grid  | 2        | 4        | 17   | 71.916666666667  | queued         |
| 18      | 23381  | mlgreen  | miller    | grid  | 2        | 4        | 18   | 71.916666666667  | queued         |
| 19      | 23382  | mlgreen  | miller    | grid  | 2        | 4        | 19   | 71.916666666667  | queued         |
| 20      | 23383  | mlgreen  | miller    | grid  | 2        | 4        | 20   | 71.916666666667  | queued         |
| 21      | 23384  | mlgreen  | miller    | grid  | 2        | 4        | 21   | 71.916666666667  | queued         |
| 22      | 23385  | mlgreen  | miller    | grid  | 2        | 4        | 22   | 71.916666666667  | queued         |
| 23      | 23386  | mlgreen  | miller    | grid  | 2        | 4        | 23   | 71.916666666667  | queued         |
| 24      | 23387  | mlgreen  | miller    | grid  | 2        | 4        | 24   | 71.916666666667  | queued         |
| 25      | 23388  | mlgreen  | miller    | grid  | 2        | 4        | 25   | 71.916666666667  | queued         |
| 26      | 23389  | mlgreen  | miller    | grid  | 2        | 4        | 26   | 71.916666666667  | queued         |
| 27      | 23393  | mlgreen  | miller    | grid  | 2        | 4        | 27   | 71.916666666667  | queued         |

Description of table fields:

**Job\_Num :** Sequential counter for row number.

**Job\_Id :** PBS job identification number corresponding to local PBS queue job number.

**Username :** PBS job owner local username.

**Groupname :** PBS job owner local primary group name.



University at Buffalo The State University of New York

Center for Computational Research

CCR

# ACDC-Grid Portal User Management

CCR Grid Computing Services: User Admin: Manage Users - Microsoft Internet Explorer

File Edit View Favorites Tools Help

**CCR** University at Buffalo The State University of New York  
Center for Computational Research GRID PORTAL  
High Performance Grid Computing

**Manage User Accounts**

In order to select which user accounts to manage, you can select one or more usernames from the list below or search for users based on specified criteria. The "Last Name" and "Organization" fields are case sensitive. Selecting "Fuzzy Search" will search on fields containing the text entered. When entering search dates, if both start and end dates are entered then values falling within that range (inclusive) will be returned. Entering only a start date will search for all entries starting with that date while entering only an end date will search for all entries up to and including that date.

Username:

Account State:

Last Name:  Fuzzy Search: ☐

Organization:  Fuzzy Search: ☐

Date Added:  through

Last Login:  through

Sort by:

Return to the User Admin menu.

Advanced Center for Computational Research Data Center

Administrator based

CCR Grid Computing Services: User Admin: Edit User Information - Microsoft Internet Explorer

File Edit View Favorites Tools Help

**CCR** University at Buffalo The State University of New York  
Center for Computational Research GRID PORTAL  
High Performance Grid Computing

**Edit information for user: mlgreen**

Username: mlgreen  
State: Active  
Last Login: 2003-09-16 09:36:23  
Last Logout: 2003-09-14 18:16:47  
Date Added: 2003-09-22

Password:

First Name:

Last Name:

Organization:

Address 1:

Address 2:

City:

State:

Country:

Postal Code:

Phone:

Fax:

Email:

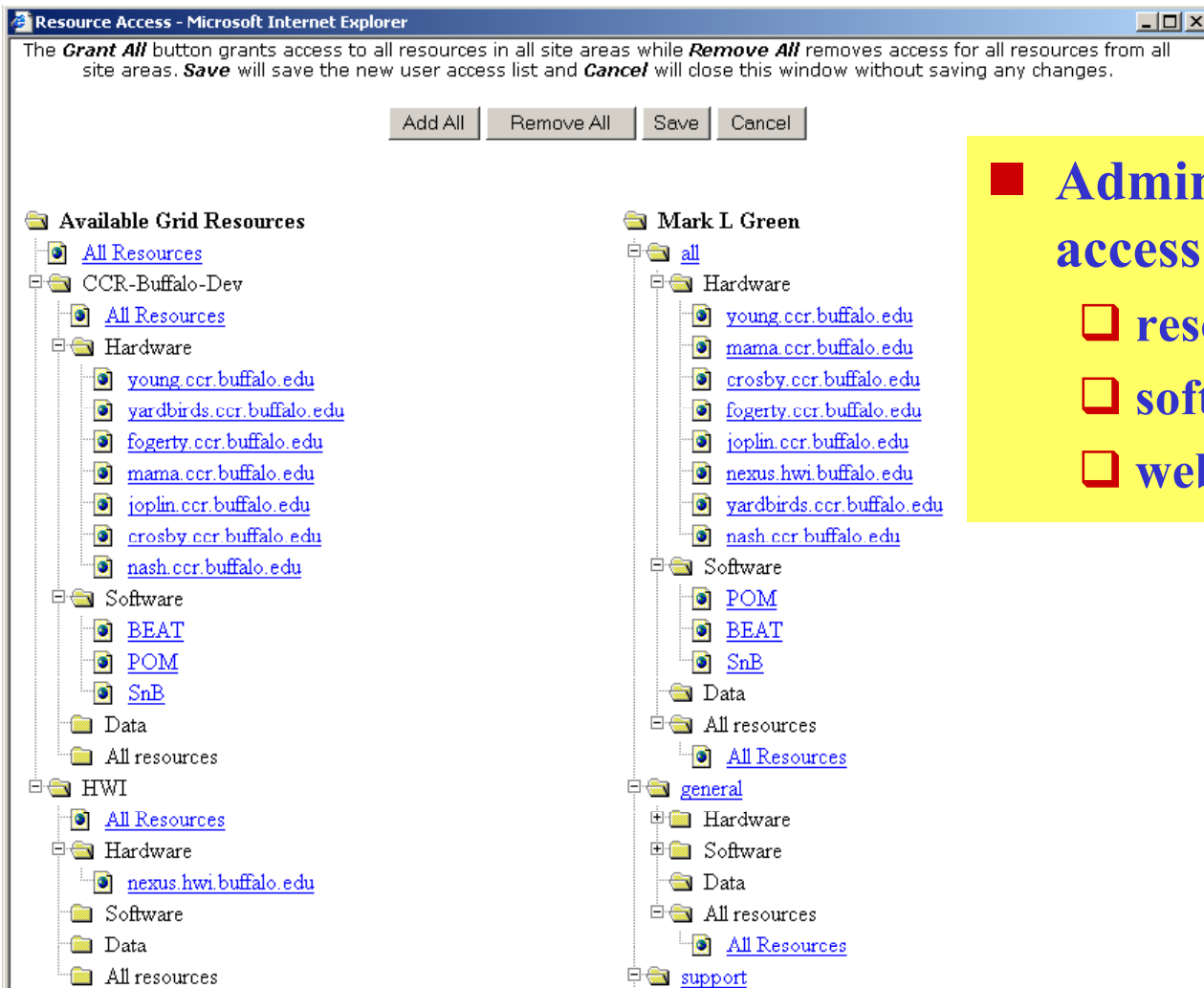
Url:

Return to the User Admin menu.

Advanced Center for Computational Research Data Center

user based

# ACDC-Grid Portal Resource Management



- Administrator grants a user access to ACDC-Grid
  - resources,
  - software, and
  - web pages.



# 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**

**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 the 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:

Return to the Database Job Admin menu.  
Return to the Grid Admin menu.

**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 10:58:12

| Resource                  | Last Updated               | Next Update | Status |
|---------------------------|----------------------------|-------------|--------|
| crosby.ccr.buffalo.edu    | 16-September-2003 09:15:30 | 2 minutes   | OK     |
| fogerty.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     |
| nexus.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 Enabled Data Mining

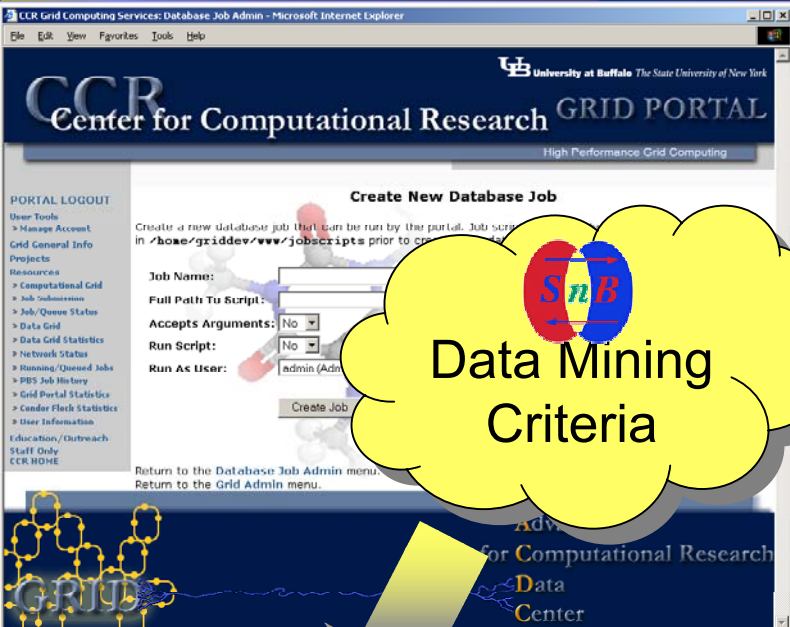
## ■ Problem Statement

- Use all available resources in the ACDC-Grid for executing a data mining genetic algorithm optimization of *SnB* parameters for molecular structures having the same space group.

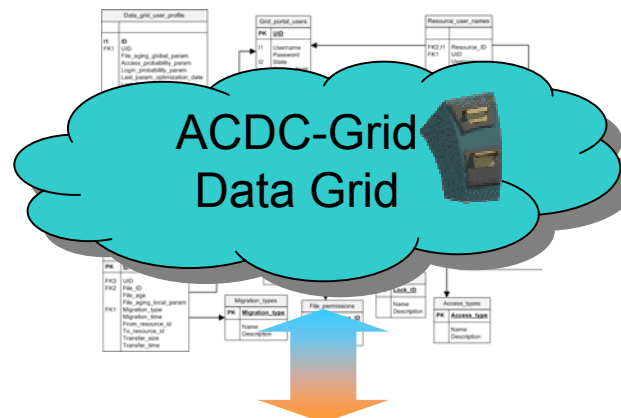
## ■ Grid Enabling Criteria

- All heterogeneous resources in the ACDC-Grid are capable of executing the *SnB* application.
- All job results obtained from the ACDC-Grid resources are stored in a corresponding molecular structure databases.

# Grid Enabled Data Mining



  
Data Mining  
Criteria



ACDC-Grid Computational  
Resources

  
Grid Portal  
Workflow Job  
Manager

Molecular  
Structure  
Database

# SnB Molecular Structure Database

domain\_snb.evo\_results running on Grid Portal - phpMyAdmin 2.5.1 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Structure Browse SQL Select Insert Export Operations Options Empty Drop

Home

domain\_snb (5)

domain\_snb

- evo\_pearson\_prod
- evo\_results
- evo\_runtime
- evo\_scores
- evo\_stats

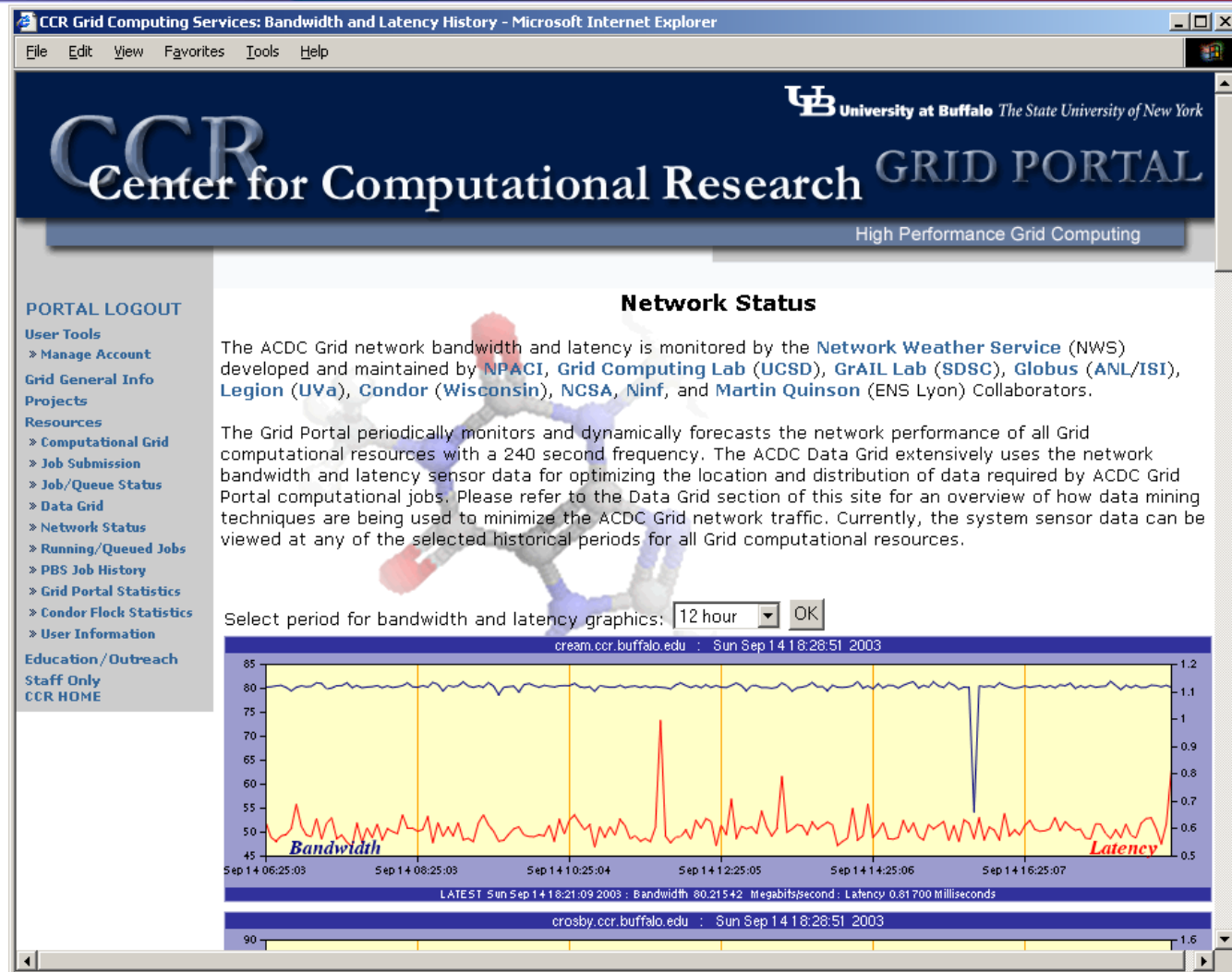
| Field                                      | Type         | Attributes | Null | Default | Extra          | Action                                    |
|--|--------------|------------|------|---------|----------------|---|
| <input type="checkbox"/> DIR_LOC           | varchar(255) |            | No   |         |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> PREFIX_OUT        | varchar(255) |            | No   |         |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> ATOMSIZE          | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> NUM_REF           | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> RESO_MAX          | float        |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> E_SIG_CUT         | float        |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> NUM_INV           | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> NUM_CYCLE         | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> PH_REFINE_METHOD  | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> PS_INIT_SHIFT     | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> PS_NUM_SHIFT      | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> PS_NUM_ITER       | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> TAN_NUM_ITER      | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> MIN_MAP_RESO      | float        |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> NUM_PEAKS_TO_OMIT | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> INTERPOLATE       | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> C1                | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> C2                | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> P1                | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> P2                | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> NUM_TRIAL         | int(11)      |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> FUNC_VALUE        | float        |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> AVG_RMIN          | float        |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> RMIN_CUTOFF       | float        |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> RUNTIME           | float        |            | No   | 0       |                | Change Drop Primary Index Unique Fulltext |
| <input type="checkbox"/> ID                | bigint(20)   | UNSIGNED   | No   |         | auto_increment | Change Drop Primary Index Unique Fulltext |

Query window

Molecular  
Structure  
Database

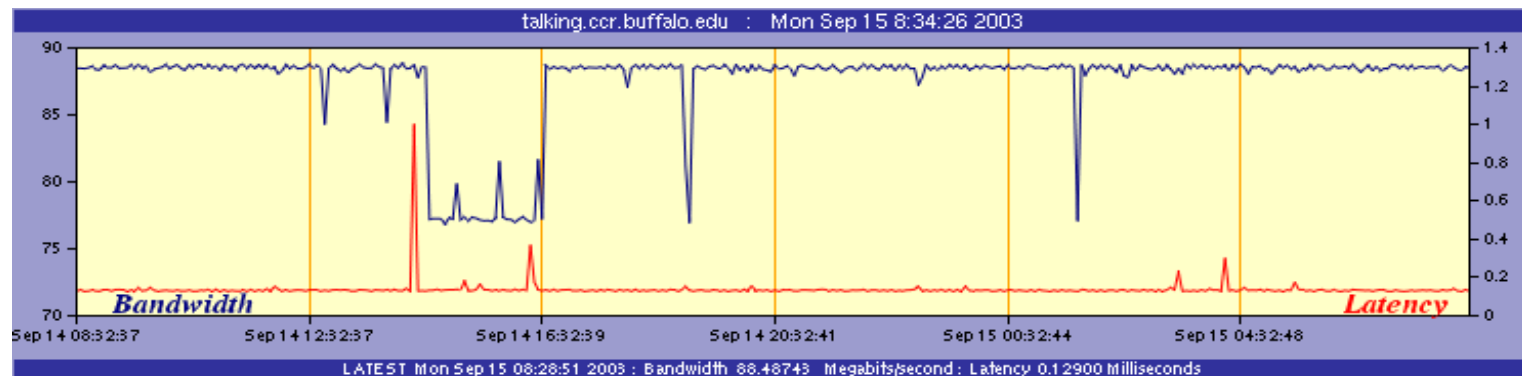
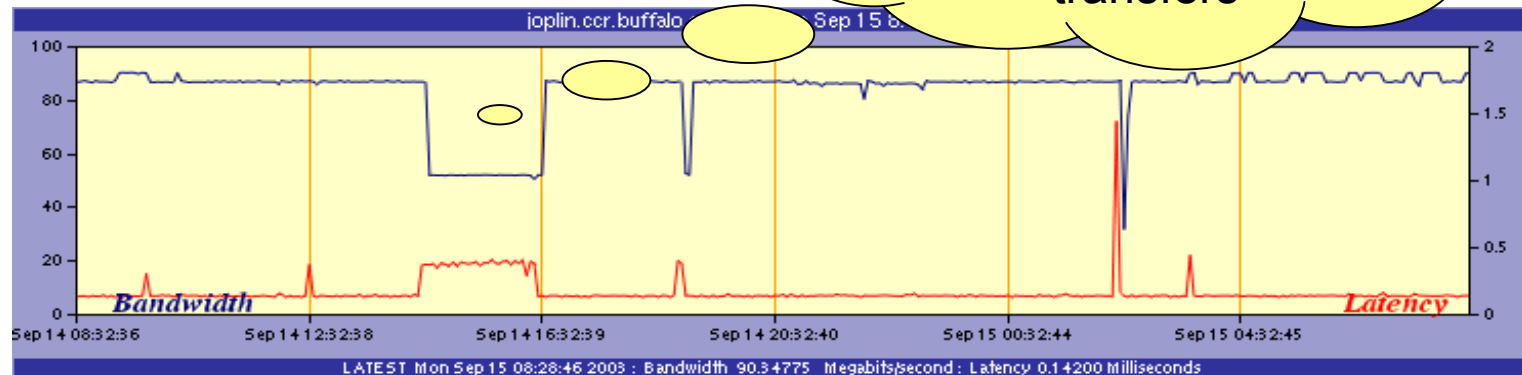


# Data Grid Resource Info



# Data Grid Resource Info

Both platforms have reduced bandwidth available for additional transfers





# Data Grid File Age

Database *data\_grid* - Table *file\_management* running on *Grid Portal*

Structure Browse SQL Select Insert Export Operations Options Empty Drop

Showing rows 0 - 29 (13000 total, Query took 0.0022 sec)

SQL-query : [Edit] [Explain SQL] [Create PHP Code]  
`SELECT 'File_ID', resource_id, filename, dir_id, access_time, file_age  
FROM 'file_management'  
WHERE 1 LIMIT 0, 30`

Show : 30 row(s) starting from record # 30  
in horizontal mode and repeat headers after 100 cells Page number: 1

|      |        | File_ID<br>Unique File ID Value | resource_id | filename      | dir_id | access_time         | file_age |
|------|--------|---------------------------------|-------------|---------------|--------|---------------------|----------|
| Edit | Delete | 62844                           | 10          | Dozer.csh     | 54030  | 2003-08-16 11:24:52 | 925832   |
| Edit | Delete | 57120                           | 10          | Tank.mpg      | 53304  | 2003-04-01 01:45:23 | 12837001 |
| Edit | Delete | 57121                           | 10          | Neo.ksh       | 53505  | 2003-06-08 06:41:05 | 6904459  |
| Edit | Delete | 57122                           | 10          | Trinity.m     | 53499  | 2003-07-30 11:34:54 | 2394030  |
| Edit | Delete | 57123                           | 10          | Rabbit.ksh    | 53541  | 2003-07-13 01:43:34 | 3898310  |
| Edit | Delete | 57124                           | 10          | Neo.ksh       | 53407  | 2003-06-22 06:19:03 | 5696181  |
| Edit | Delete | 57049                           | 10          | Agent.ppt     | 53928  | 2003-02-24 12:15:39 | 15909585 |
| Edit | Delete | 61710                           | 10          | Neo.txt       | 52724  | 2003-07-26 09:44:48 | 2746236  |
| Edit | Delete | 61711                           | 10          | Morpheus.sh   | 52710  | 2003-07-31 07:03:43 | 2367101  |
| Edit | Delete | 61712                           | 10          | Morpheus.ppt  | 52761  | 2003-08-26 08:04:38 | 117046   |
| Edit | Delete | 61713                           | 10          | Tank.jpg      | 52929  | 2003-06-26 09:59:37 | 5337347  |
| Edit | Delete | 61714                           | 10          | Rabbit.dat    | 52624  | 2003-08-26 05:57:43 | 124661   |
| Edit | Delete | 61715                           | 10          | KeyMaster.mpg | 52770  | 2003-06-17 04:16:44 | 6178720  |

■ File age, access time, and resource id denote:

- the amount of time since a file was accessed,
- when the file was accessed, and
- where the file currently resides respectively.

# ACDC-Grid

## Development/Maintenance

### ■ Development Requirements

#### □ 7 – Person months for Grid Services Coordinator

- Including Grid and Database conceptual design and implementation

#### □ 5 – Person months for Grid Services Programmer

- Web portal programming

#### □ 5 – Person months for System Administrator

- Globus, NWS, MDS, etc. installations

#### □ 3 – Person months for Database Administrator

- Grid Portal Database implementation

### ■ Minimum Maintenance Requirements

#### □ 1 – Grid Services Coordinator

- 100% level of effort

#### □ 1 – Grid Services Programmer

- 100% level of effort

#### □ 1 – System Administrator

- 50% level of effort

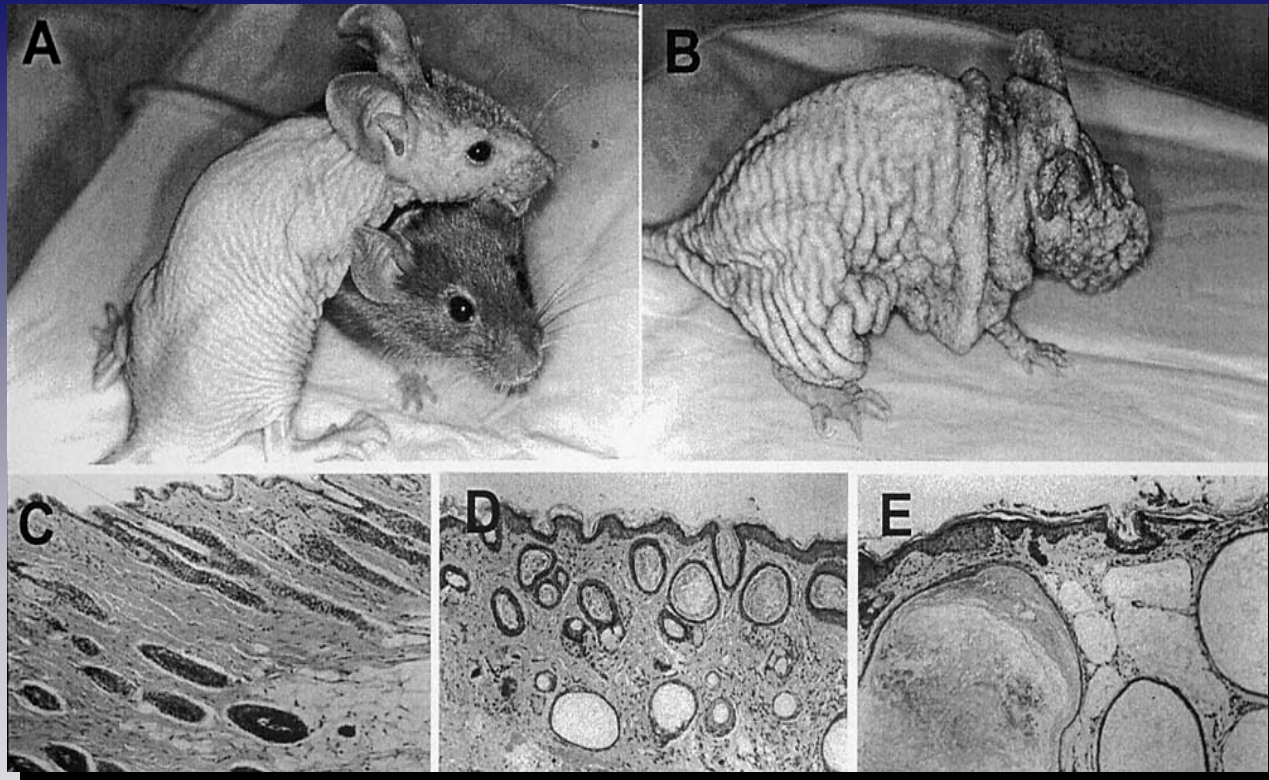
#### □ 1 – Database Administrator

- 10% level of effort

# Acknowledgments

- Mark Green
- Steve Gallo
- Jason Rappleye
- Jeff Tilson
- Martins Innus
- Cynthia Cornelius
- George DeTitta
- Herb Hauptman
- Charles Weeks
- Steve Potter

# Contact Information



[miller@buffalo.edu](mailto:miller@buffalo.edu)  
[www.ccr.buffalo.edu](http://www.ccr.buffalo.edu)