

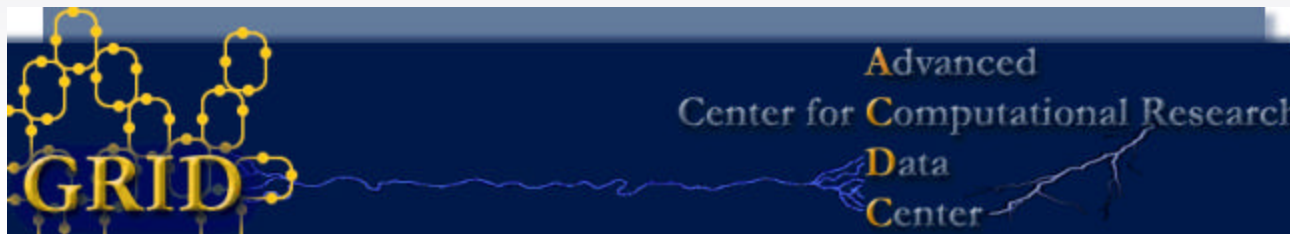
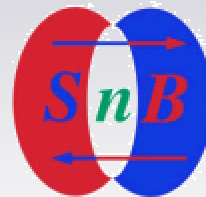
# The Center for Computational Research: High-End Computing & Visualization and Grid Computing

**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*

# Center for Computational Research 1998-2005 Snapshot

## ■ High-End Computing, Storage, Networking, and Visualization

### □ ~140 Research Groups in 37 Depts

○ Physical Sciences

○ Life Sciences

○ Engineering

○ Scientific Visualization, Medical Imaging, Virtual Reality

### □ 13 Local Companies

### □ 10 Local Institutions

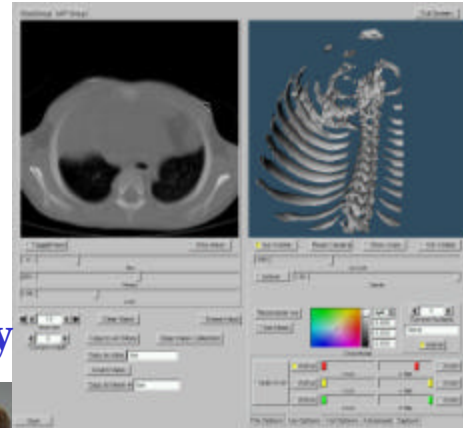
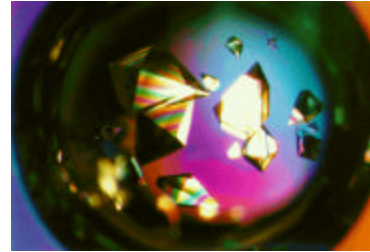
## ■ External Funding: \$300M+

## ■ Total Leveraged WNY: \$500M+

## ■ Deliverables

### □ 1100+ Publications

### □ Software, Media, Algorithms, Consulting, Training, CPU Cycles...



# Major Compute/Storage Resources

## (22 TF Peak Compute & 130 TB of User Storage)

- **Dell Linux Cluster (10TF peak)**
  - ❑ 1600 Xeon EM64T Processors (3.2 GHz)
  - ❑ 2 TB RAM; 65 TB Disk
  - ❑ Myrinet / Force10
  - ❑ 30 TB EMC SAN
- **Dell Linux Cluster (2.9TF peak)**
  - ❑ 600 P4 Processors (2.4 GHz)
  - ❑ 600 GB RAM; 40 TB Disk; Myrinet
- **Dell Linux Cluster (6TF peak)**
  - ❑ 4036 Processors (PIII 1.2 GHz)
  - ❑ 2TB RAM; 160TB Disk; 16TB SAN
- **IBM BladeCenter Cluster (3TF peak)**
  - ❑ 532 P4 Processors (2.8 GHz)
  - ❑ 5TB SAN
- **SGI Intel Linux Cluster (0.1TF peak)**
  - ❑ 150 PIII Processors (1 GHz)
  - ❑ Myrinet
- **SGI Altix3700 (0.4TF peak)**
  - ❑ 64 Processors (1.3GHz ITF2)
  - ❑ 256 GB RAM
  - ❑ 2.5 TB 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



# CCR Visualization Resources

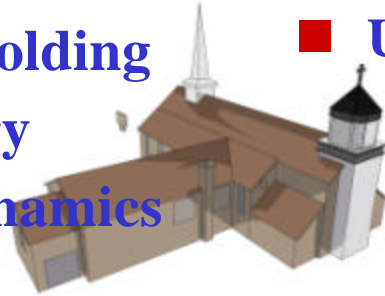
- **Fakespace ImmersaDesk R2**
  - ❑ Portable 3D Device
  - ❑ Onyx2: 6 R10000 @ 250MHz
  - ❑ 2 IR2 Pipes; 3 64MB texture memory mgrs
- **Tiled-Display Wall**
  - ❑ 20 NEC projectors: 15.7M pixels
  - ❑ Screen is 11' x 7'
  - ❑ Dell PCs with Myrinet2000
- **Access Grid Nodes (2)**
  - ❑ Group-to-Group Communication
  - ❑ Commodity components
- **SGI Reality Center 3300W**
  - ❑ Dual Barco's on 8' x 4' screen
  - ❑ Onyx300: 10 R14000 @ 500MHz
  - ❑ 2 IR4 Pipes; 1 GB texture mem per pipe





# CCR Research & Projects

- Archaeology
- Bioinformatics/Protein Folding
- Computational Chemistry
- Computational Fluid Dynamics
- Data Mining/Database
- Earthquake Engineering
- Environ Modeling & Simulation
- Grid Computing
- Molecular Structure Determination



- Videos: MTV
- Urban Simulation and Viz

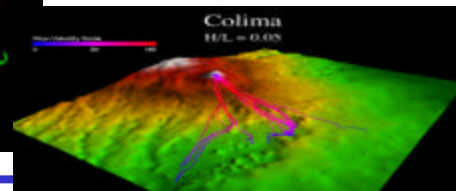
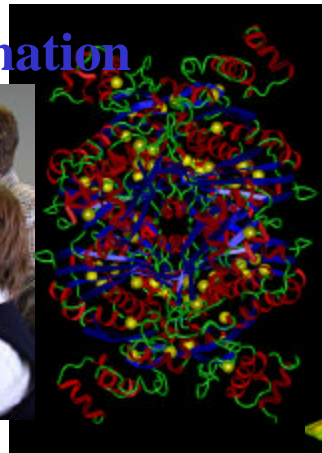
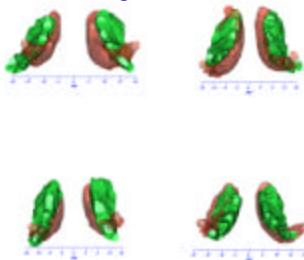
- StreetScenes
- I-90 Toll Barrier
- Medical Campus
- Peace Bridge

- Accident Reconstruction
- Scientific Viz

- Dental
- Surgery
- MRI/CT Scan
- Confocal Microscopy
- Crystallization Wells
- Collaboratories



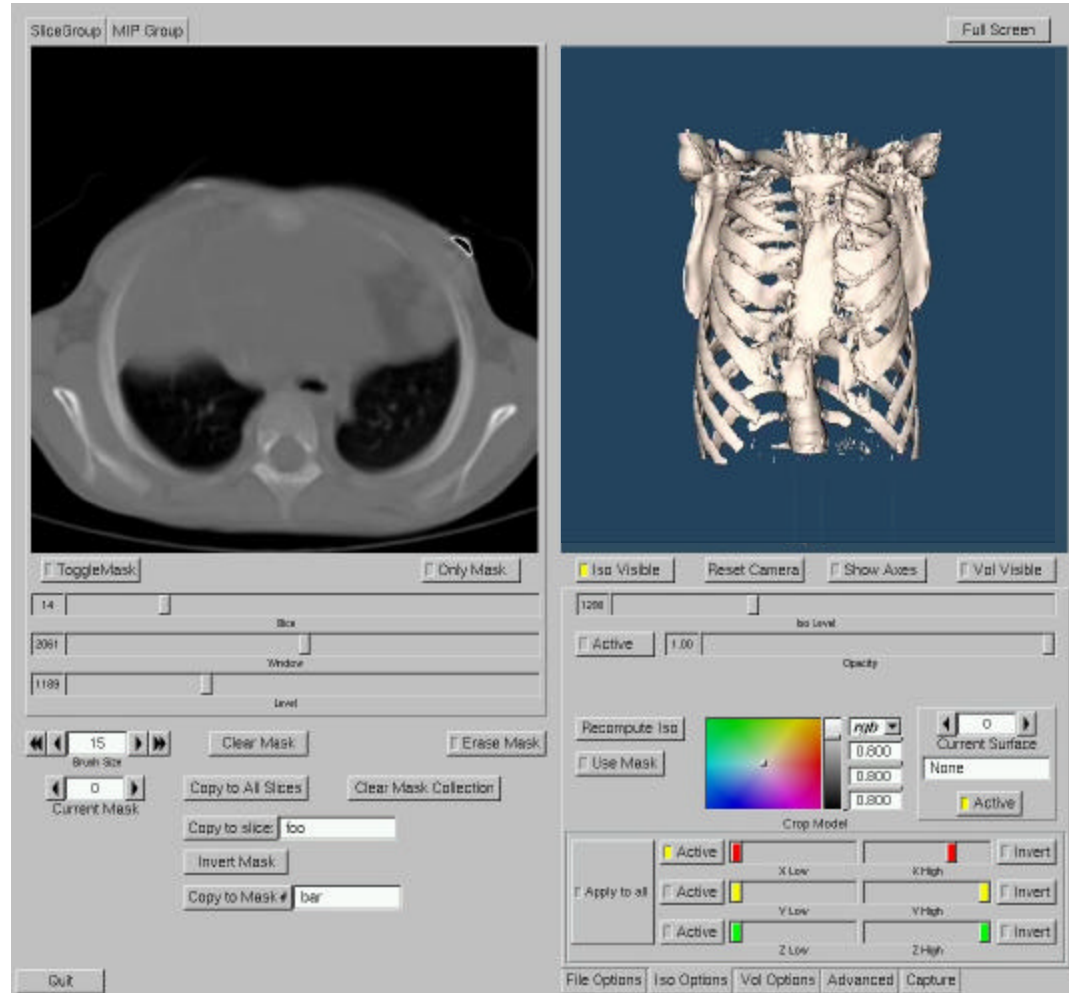
## Physics



# Real-Time Visualization

# 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: Real-Time 3D Traffic Simulation

- Accurate local landmarks: Bridges, Street Signs, Business, Homes
- Can be viewed from driver's perspective
- Real-Time Navigation
- Works with
  - Corsim
  - Synchro
- Generate AVI & MOV
- Simultaneously multiple
  - Traffic Loads
  - Simulation
  - Varying POV



# Animation & Simulation

## **Rendered Scenes**

# Thruway HOT Lanes Animation





# Peace Bridge Visualization: Animation & Simulation

## ■ Proposed Options

- ❑ Relocate US plaza
- ❑ Build a 3-lane companion span & rehab existing bridge

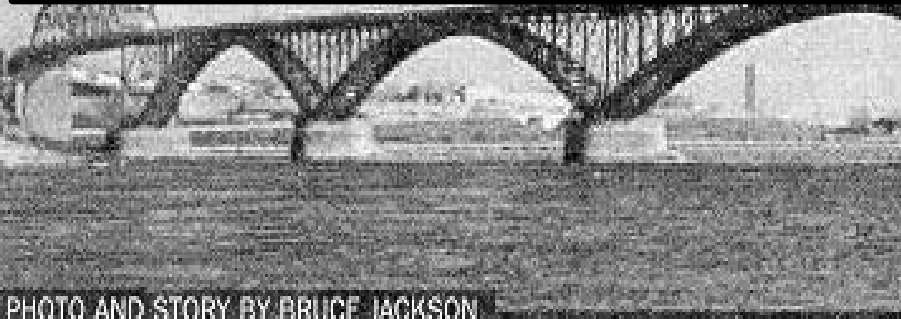
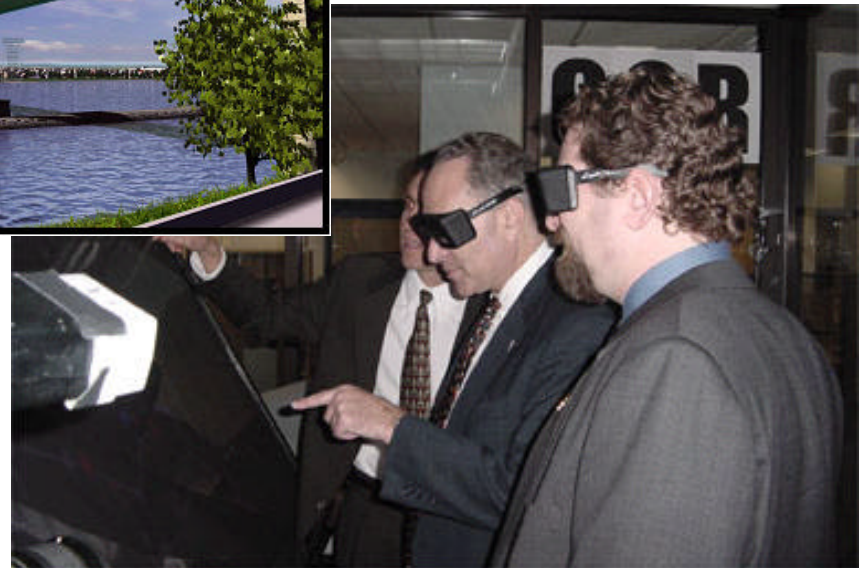


PHOTO AND STORY BY BRUCE JACKSON



# Public Forum



# Grid Computing



DISCOM

SinRG

APGrid

IPG ...



Asia-Pacific Advanced Network



University at Buffalo

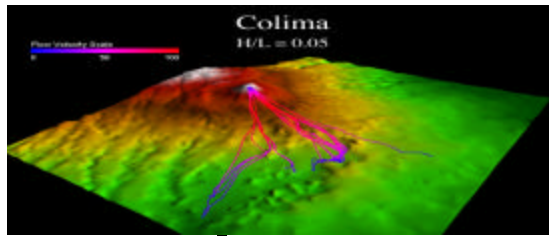
The State University of New York

Center for Computational Research

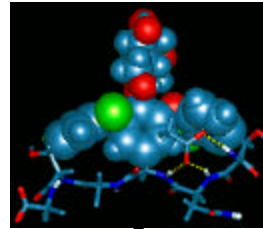
CCR



# Grid Computing Overview



Data Acquisition



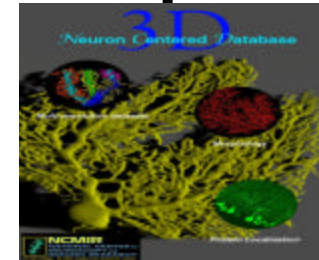
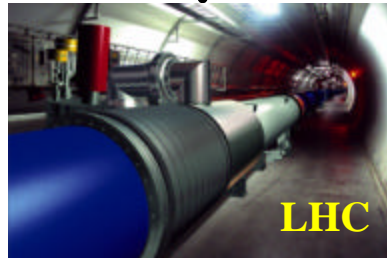
Advanced Visualization



Analysis



Imaging Instruments



Large-Scale Databases

- 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

# ACDC-Grid Cyber-Infrastructure

## ■ Integrated Data Grid

- ❑ Automated Data File Migration based on profiling users.

## ■ Lightweight Grid Monitor (Dashboard)

## ■ Predictive Scheduler

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

## ■ Dynamic Resource Allocation

- ❑ Develop automated procedures for dynamic computational resource allocation.

## ■ High-Performance Grid-Enabled Data Repositories

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

# ACDC-Grid Collaborations

- **High-Performance Networking Infrastructure**
- **WNY Grid Initiative**
- **Grid3+ Collaboration**
- **iVDGL Member**
  - ❑ Only External Member
- **Open Science Grid Member**
  - ❑ Organizational Committee
  - ❑ Blueprint Committee
  - ❑ Security Working Group
  - ❑ Data Working Group
  - ❑ GRASE VO
- **Grid-Lite: Campus Grid**
  - ❑ HP Labs Collaboration
- **Innovative Laboratory Prototype**
  - ❑ Dell Collaboration

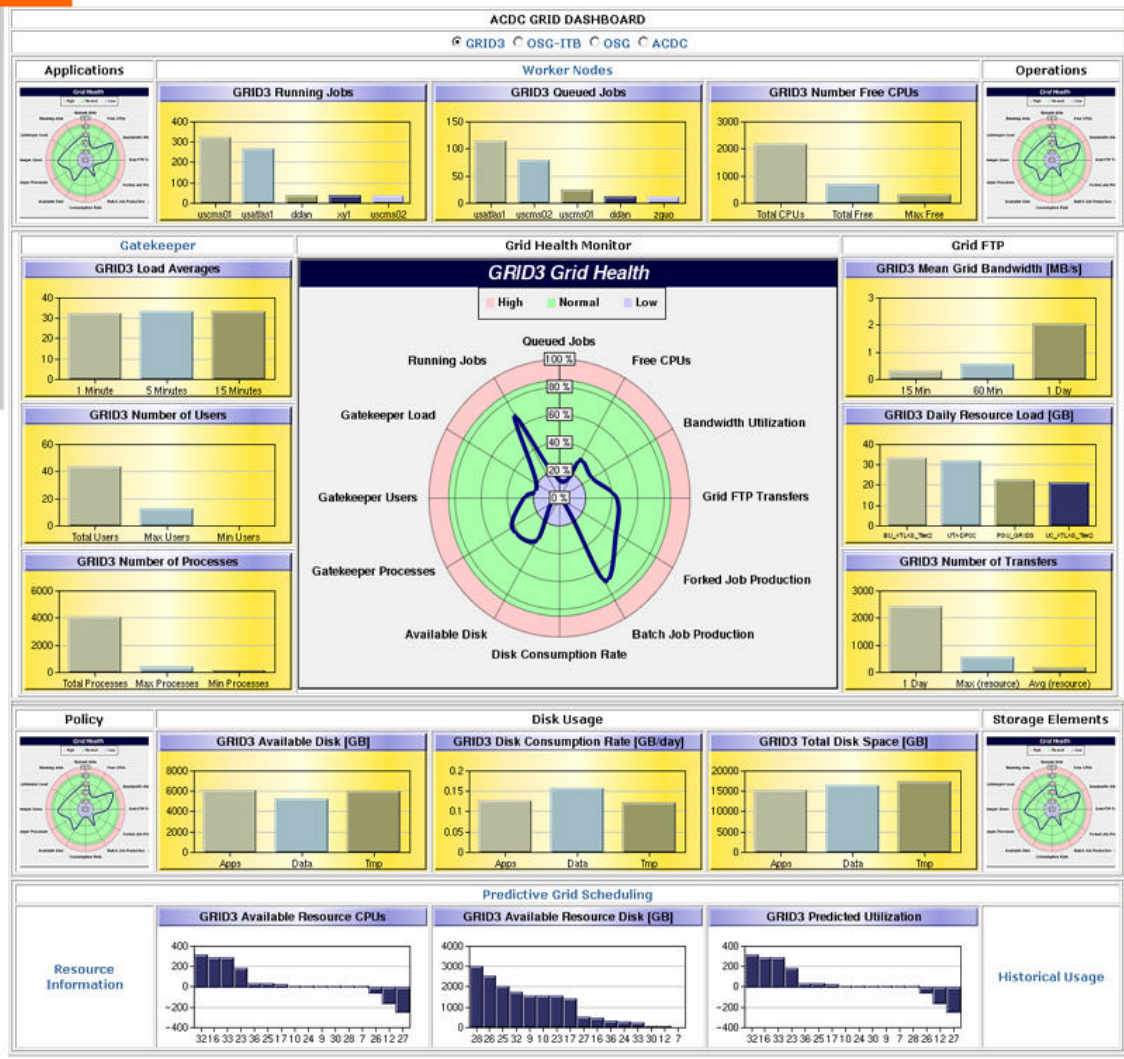






# ACDC-Grid Monitoring: The ACDC-Grid DASHBOARD

- Grid Resources
- ACDC Monitoring
- ACDC Grid Dashboard
- Running/Queued Jobs
- Job History
- Detailed Job History
- Detailed GridFTP History
- Resource Queue Visualization
- Resource User Visualization
- Self Application Demonstrator
- Presentations
- ACDC Site Status
- Contact Us
- Staff Only



# Grid-Enabling Application Templates (GATs)

## ■ Structural Biology

- ❑ *SnB* and *BnP* for Molecular Structure Determination/Phasing

## ■ Groundwater Modeling

- ❑ *Ostrich*: Optimization and Parameter Estimation Tool
- ❑ *POMGL*: Princeton Ocean Model Great Lakes for Hydrodynamic Circulation
- ❑ *Split*: Modeling Groundwater Flow with Analytic Element Method

## ■ Earthquake Engineering

- ❑ *EADR*: Evolutionary Aseismic Design and Retrofit; Passive Energy Dissipation System for Designing Earthquake Resilient Structures

## ■ Computational Chemistry

- ❑ *Q-Chem*: Quantum Chemistry Package

## ■ Geographic Information Systems & BioHazards

- ❑ *Titan*: Computational Modeling of Hazardous Geophysical Mass Flows

# 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

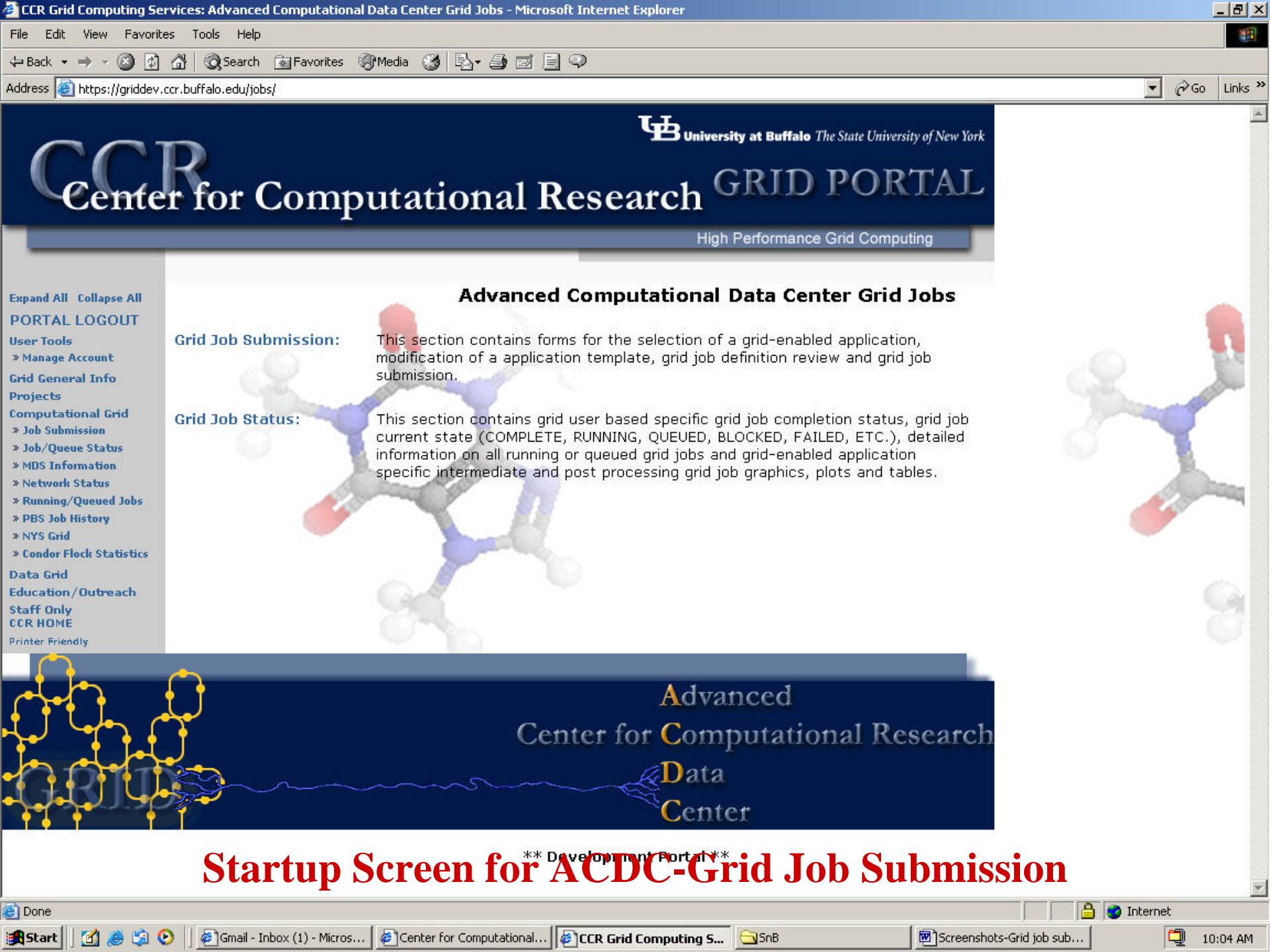
**ACDC-Grid  
Data  
Resources**



Adapted from Ian Foster and Carl Kesselman







# CCR Center for Computational Research GRID PORTAL

High Performance Grid Computing

## Advanced Computational Data Center Grid Jobs

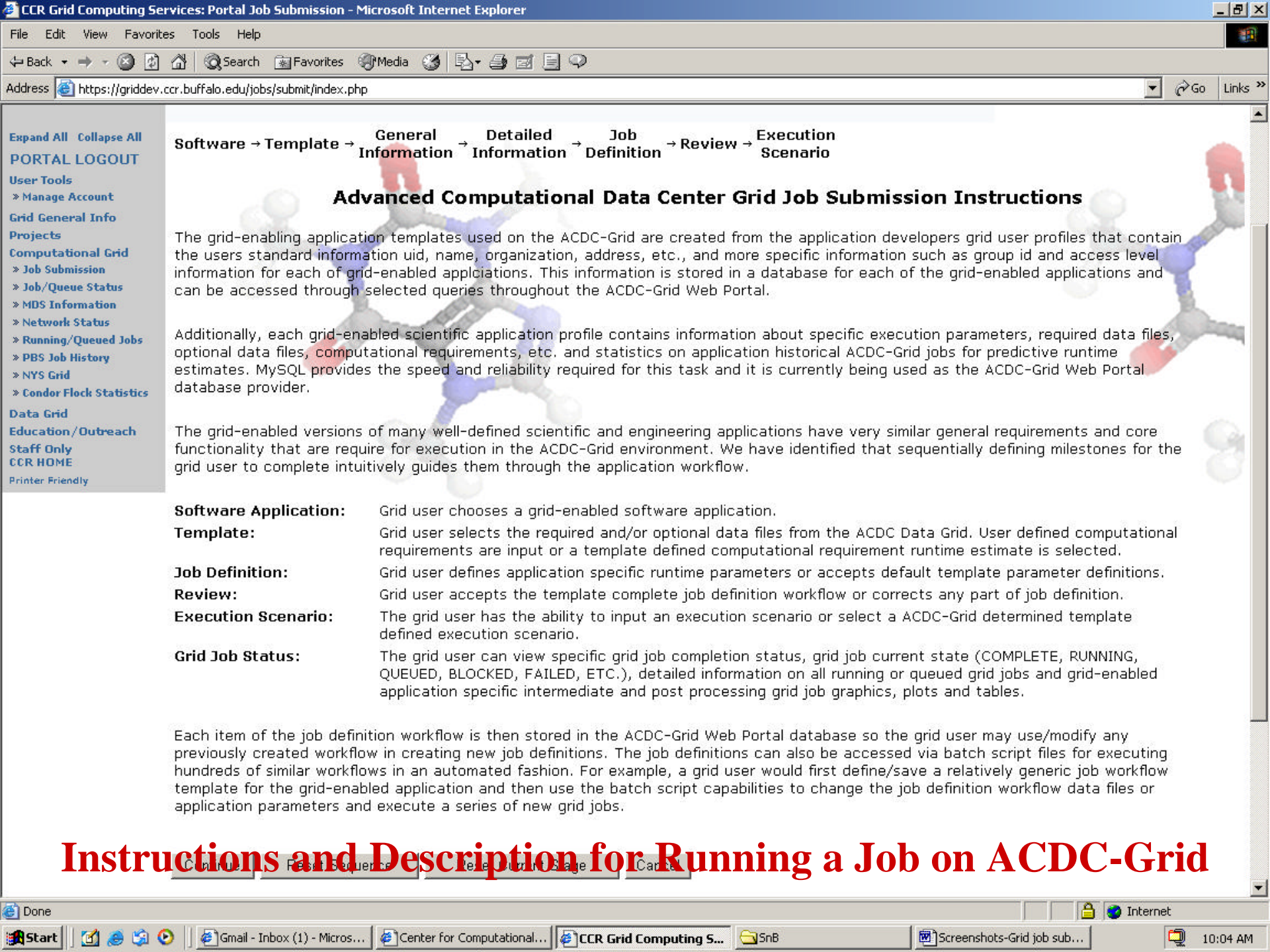
**Grid Job Submission:** This section contains forms for the selection of a grid-enabled application, modification of an application template, grid job definition review and grid job submission.

**Grid Job Status:** This section contains grid user based specific grid job completion status, grid job current state (COMPLETE, RUNNING, QUEUED, BLOCKED, FAILED, ETC.), detailed information on all running or queued grid jobs and grid-enabled application specific intermediate and post processing grid job graphics, plots and tables.

- Expand All Collapse All
- PORTAL LOGOUT
- User Tools
  - » Manage Account
- Grid General Info
- Projects
- Computational Grid
  - » Job Submission
  - » Job/Queue Status
  - » MDS Information
  - » Network Status
  - » Running/Queued Jobs
  - » PBS Job History
  - » NYS Grid
  - » Condor Flock Statistics
- Data Grid
- Education/Outreach
- Staff Only
- CCR HOME
- Printer Friendly



**Startup Screen for ACDC-Grid Job Submission**



Software → Template → General Information → Detailed Information → Job Definition → Review → Execution Scenario

## Advanced Computational Data Center Grid Job Submission Instructions

The grid-enabling application templates used on the ACDC-Grid are created from the application developers grid user profiles that contain the users standard information uid, name, organization, address, etc., and more specific information such as group id and access level information for each of grid-enabled applications. This information is stored in a database for each of the grid-enabled applications and can be accessed through selected queries throughout the ACDC-Grid Web Portal.

Additionally, each grid-enabled scientific application profile contains information about specific execution parameters, required data files, optional data files, computational requirements, etc. and statistics on application historical ACDC-Grid jobs for predictive runtime estimates. MySQL provides the speed and reliability required for this task and it is currently being used as the ACDC-Grid Web Portal database provider.

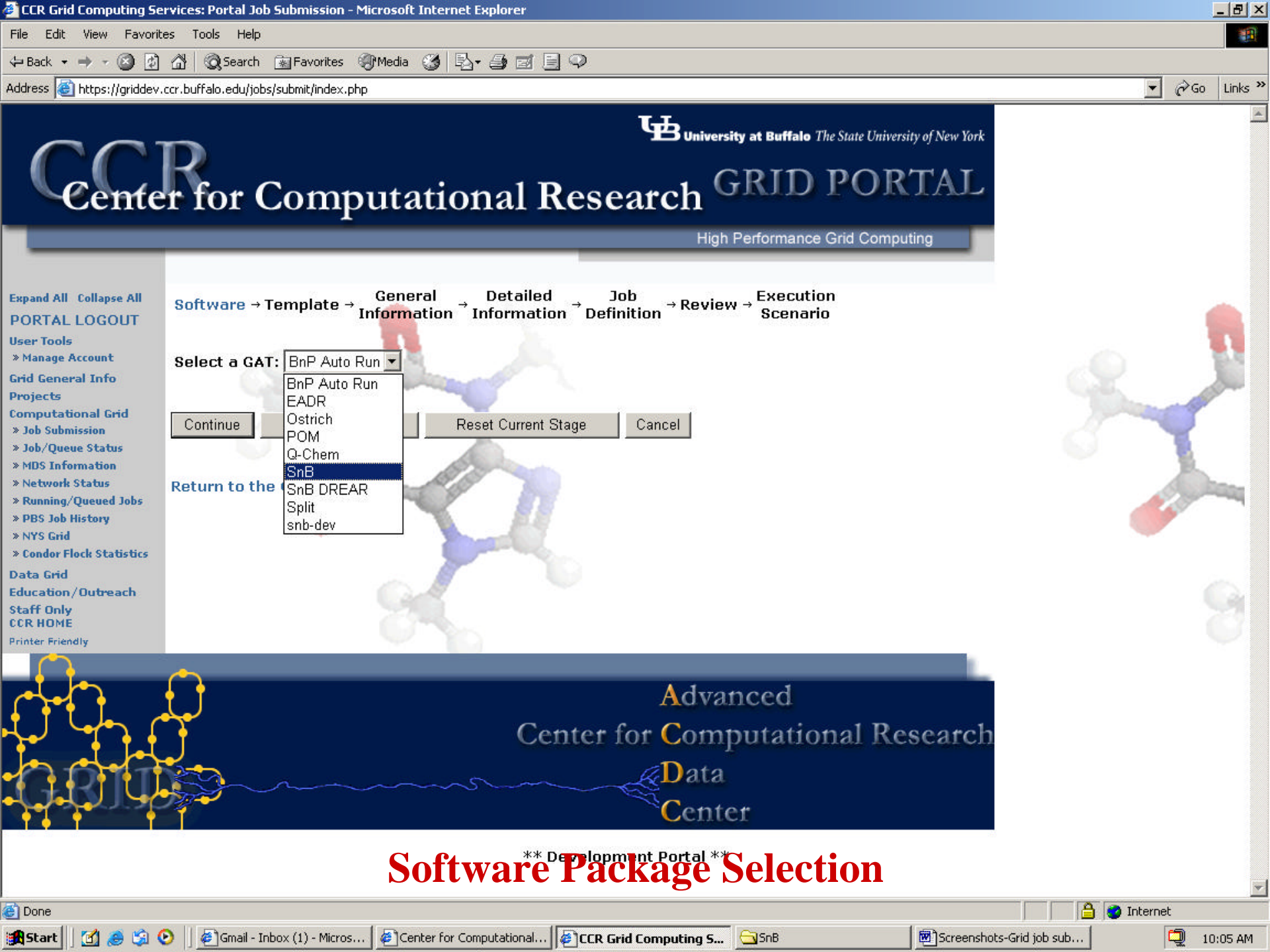
The grid-enabled versions of many well-defined scientific and engineering applications have very similar general requirements and core functionality that are require for execution in the ACDC-Grid environment. We have identified that sequentially defining milestones for the grid user to complete intuitively guides them through the application workflow.

- Software Application:** Grid user chooses a grid-enabled software application.
- Template:** Grid user selects the required and/or optional data files from the ACDC Data Grid. User defined computational requirements are input or a template defined computational requirement runtime estimate is selected.
- Job Definition:** Grid user defines application specific runtime parameters or accepts default template parameter definitions.
- Review:** Grid user accepts the template complete job definition workflow or corrects any part of job definition.
- Execution Scenario:** The grid user has the ability to input an execution scenario or select a ACDC-Grid determined template defined execution scenario.
- Grid Job Status:** The grid user can view specific grid job completion status, grid job current state (COMPLETE, RUNNING, QUEUED, BLOCKED, FAILED, ETC.), detailed information on all running or queued grid jobs and grid-enabled application specific intermediate and post processing grid job graphics, plots and tables.

Each item of the job definition workflow is then stored in the ACDC-Grid Web Portal database so the grid user may use/modify any previously created workflow in creating new job definitions. The job definitions can also be accessed via batch script files for executing hundreds of similar workflows in an automated fashion. For example, a grid user would first define/save a relatively generic job workflow template for the grid-enabled application and then use the batch script capabilities to change the job definition workflow data files or application parameters and execute a series of new grid jobs.

# Instructions and Description for Running a Job on ACDC-Grid





# CCR Center for Computational Research GRID PORTAL

High Performance Grid Computing

- Expand All Collapse All
- PORTAL LOGOUT
- User Tools
  - » Manage Account
- Grid General Info
- Projects
- Computational Grid
  - » Job Submission
  - » Job/Queue Status
  - » MDS Information
  - » Network Status
  - » Running/Queued Jobs
  - » PBS Job History
  - » NYS Grid
  - » Condor Flock Statistics
- Data Grid
- Education/Outreach
- Staff Only
- CCR HOME
- Printer Friendly

Software -> Template -> General Information -> Detailed Information -> Job Definition -> Review -> Execution Scenario

Select a GAT:

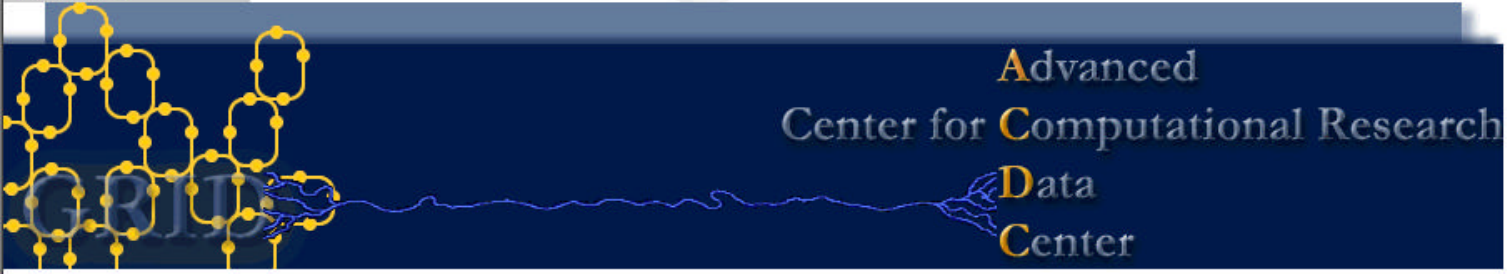
- BnP Auto Run
- BnP Auto Run
- EADR
- Ostrich
- POM
- Q-Chem
- SnB
- SnB DREAM
- Split
- snb-dev

Continue

Reset Current Stage

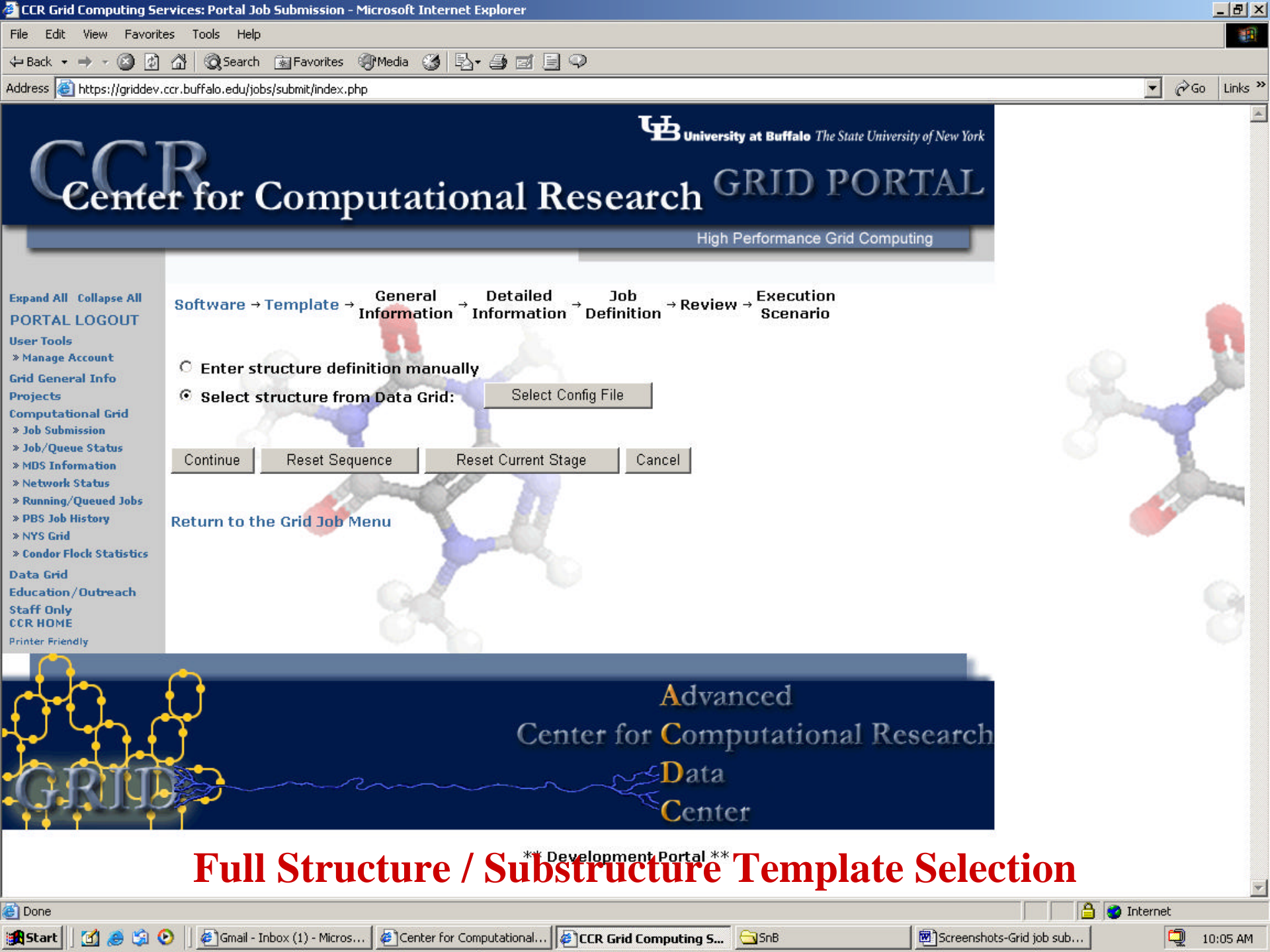
Cancel

Return to the



**\*\* Development Portal \*\***  
**Software Package Selection**



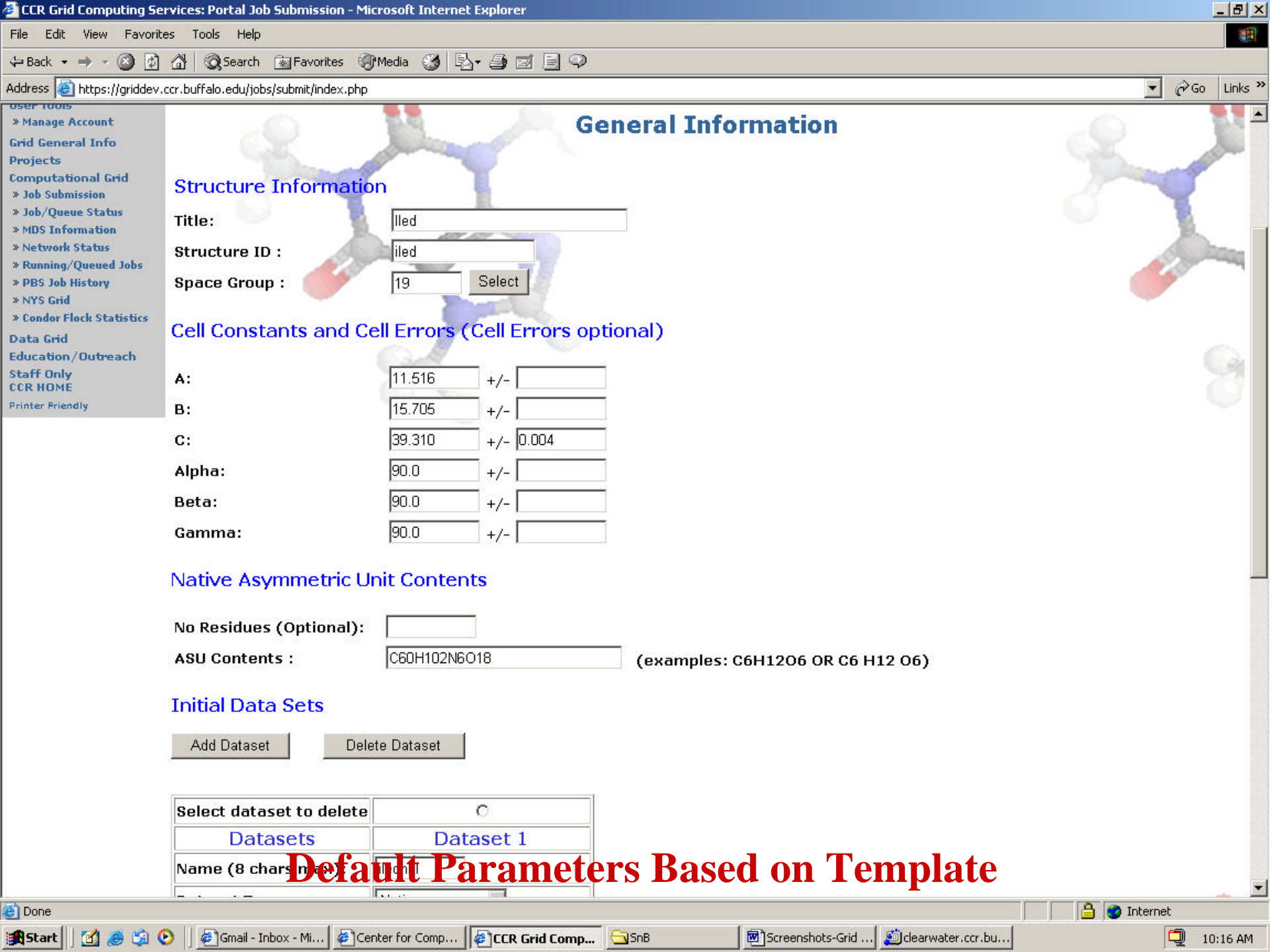


Software → **Template** → General Information → Detailed Information → Job Definition → Review → Execution Scenario

Enter structure definition manually  
 Select structure from Data Grid:

[Return to the Grid Job Menu](#)

**Full Structure / Substructure Template Selection**  
\*\* Development Portal \*\*



- USER TOOLS
- » Manage Account
- Grid General Info
- Projects
- Computational Grid
- » Job Submission
- » Job/Queue Status
- » MDS Information
- » Network Status
- » Running/Queued Jobs
- » PBS Job History
- » NYS Grid
- » Condor Flock Statistics
- Data Grid
- Education/Outreach
- Staff Only
- CCR HOME
- Printer Friendly

## General Information

### Structure Information

Title:

Structure ID :

Space Group :

### Cell Constants and Cell Errors (Cell Errors optional)

A:  +/-

B:  +/-

C:  +/-

Alpha:  +/-

Beta:  +/-

Gamma:  +/-

### Native Asymmetric Unit Contents

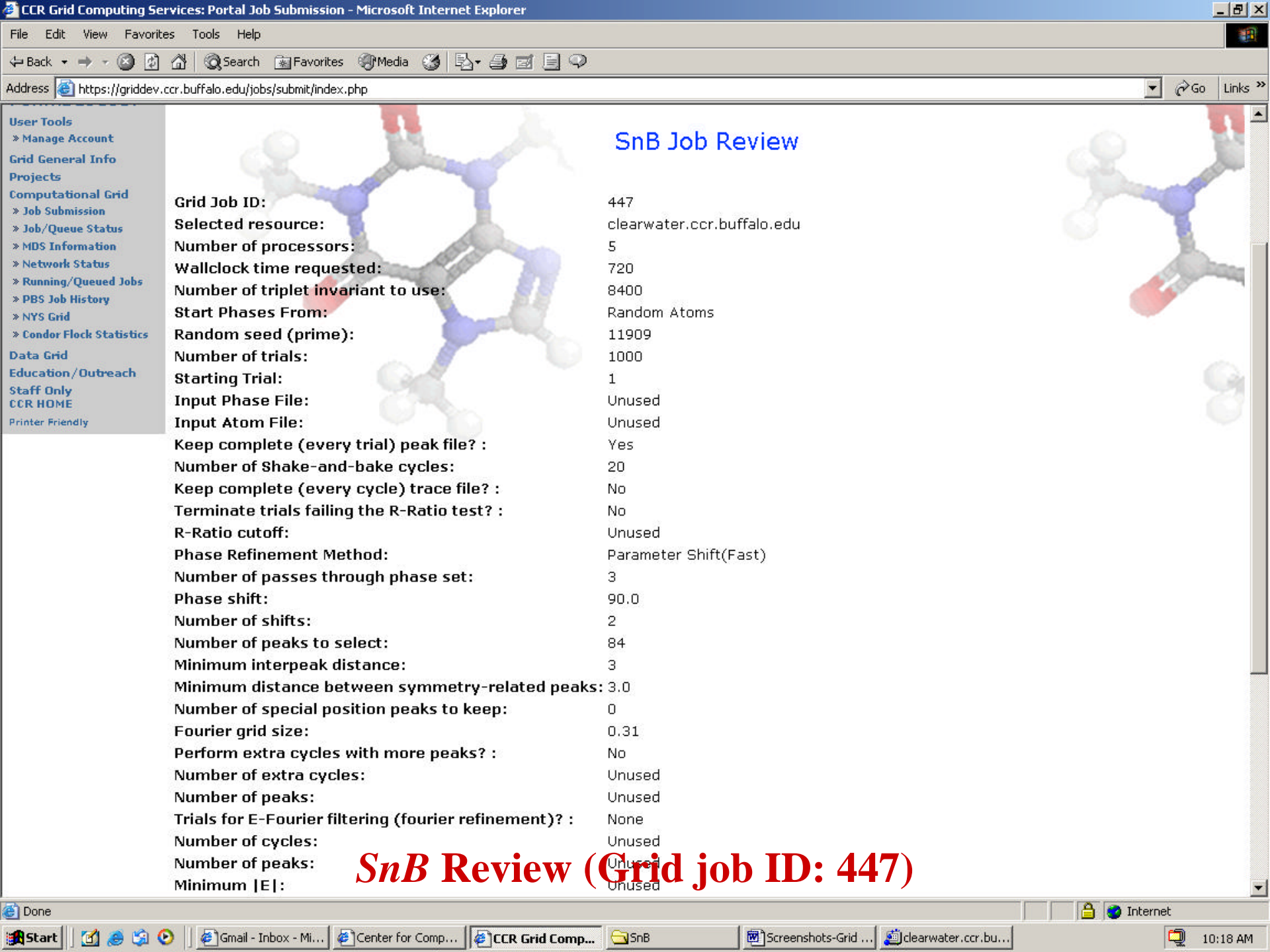
No Residues (Optional):

ASU Contents :  (examples: C6H12O6 OR C6 H12 O6)

### Initial Data Sets

Select dataset to delete	
<a href="#">Datasets</a>	<a href="#">Dataset 1</a>
Name (8 chars max)	

**Default Parameters Based on Template**



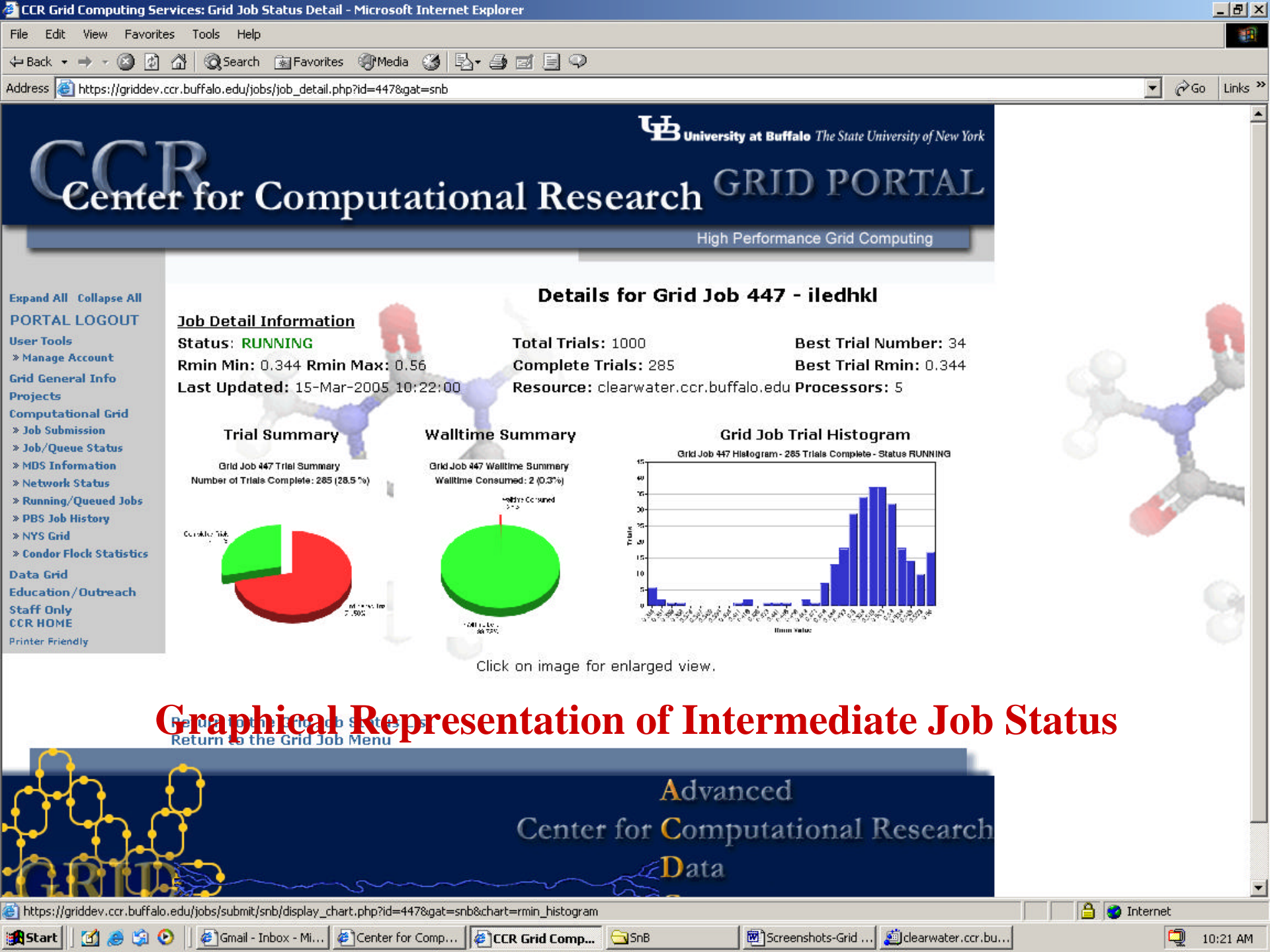
- User Tools
  - » Manage Account
- Grid General Info
- Projects
- Computational Grid
  - » Job Submission
  - » Job/Queue Status
  - » MDS Information
  - » Network Status
  - » Running/Queued Jobs
  - » PBS Job History
  - » NYS Grid
  - » Condor Flock Statistics
- Data Grid
- Education/Outreach
-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

## SnB Job Review

<b>Grid Job ID:</b>	447
<b>Selected resource:</b>	clearwater.ccr.buffalo.edu
<b>Number of processors:</b>	5
<b>Wallclock time requested:</b>	720
<b>Number of triplet invariant to use:</b>	8400
<b>Start Phases From:</b>	Random Atoms
<b>Random seed (prime):</b>	11909
<b>Number of trials:</b>	1000
<b>Starting Trial:</b>	1
<b>Input Phase File:</b>	Unused
<b>Input Atom File:</b>	Unused
<b>Keep complete (every trial) peak file? :</b>	Yes
<b>Number of Shake-and-bake cycles:</b>	20
<b>Keep complete (every cycle) trace file? :</b>	No
<b>Terminate trials failing the R-Ratio test? :</b>	No
<b>R-Ratio cutoff:</b>	Unused
<b>Phase Refinement Method:</b>	Parameter Shift(Fast)
<b>Number of passes through phase set:</b>	3
<b>Phase shift:</b>	90.0
<b>Number of shifts:</b>	2
<b>Number of peaks to select:</b>	84
<b>Minimum interpeak distance:</b>	3
<b>Minimum distance between symmetry-related peaks:</b>	3.0
<b>Number of special position peaks to keep:</b>	0
<b>Fourier grid size:</b>	0.31
<b>Perform extra cycles with more peaks? :</b>	No
<b>Number of extra cycles:</b>	Unused
<b>Number of peaks:</b>	Unused
<b>Trials for E-Fourier filtering (fourier refinement)? :</b>	None
<b>Number of cycles:</b>	Unused
<b>Number of peaks:</b>	Unused
<b>Minimum  E :</b>	Unused

**SnB Review (Grid job ID: 447)**





# CCR Center for Computational Research

## GRID PORTAL

High Performance Grid Computing

- Expand All Collapse All
- PORTAL LOGOUT
- User Tools
  - Manage Account
- Grid General Info
- Projects
- Computational Grid
  - Job Submission
  - Job/Queue Status
  - MDS Information
  - Network Status
  - Running/Queued Jobs
  - PBS Job History
  - NYS Grid
  - Condor Flock Statistics
- Data Grid
- Education/Outreach
- Staff Only
- CCR HOME
- Printer Friendly

### Details for Grid Job 447 - iledhkl

#### Job Detail Information

Status: **RUNNING**  
Rmin Min: 0.344 Rmin Max: 0.56  
Last Updated: 15-Mar-2005 10:22:00

Total Trials: 1000  
Complete Trials: 285  
Resource: clearwater.ccr.buffalo.edu

Best Trial Number: 34  
Best Trial Rmin: 0.344  
Processors: 5

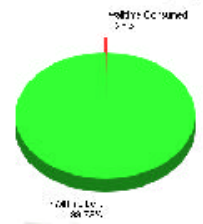
#### Trial Summary

Grid Job 447 Trial Summary  
Number of Trials Complete: 285 (28.5%)



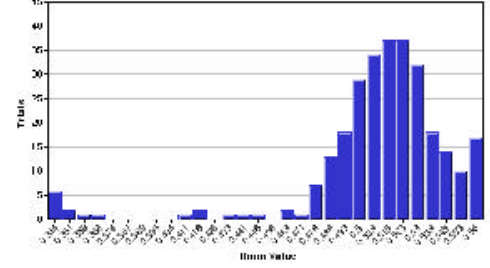
#### Walltime Summary

Grid Job 447 Walltime Summary  
Walltime Consumed: 2 (0.3%)



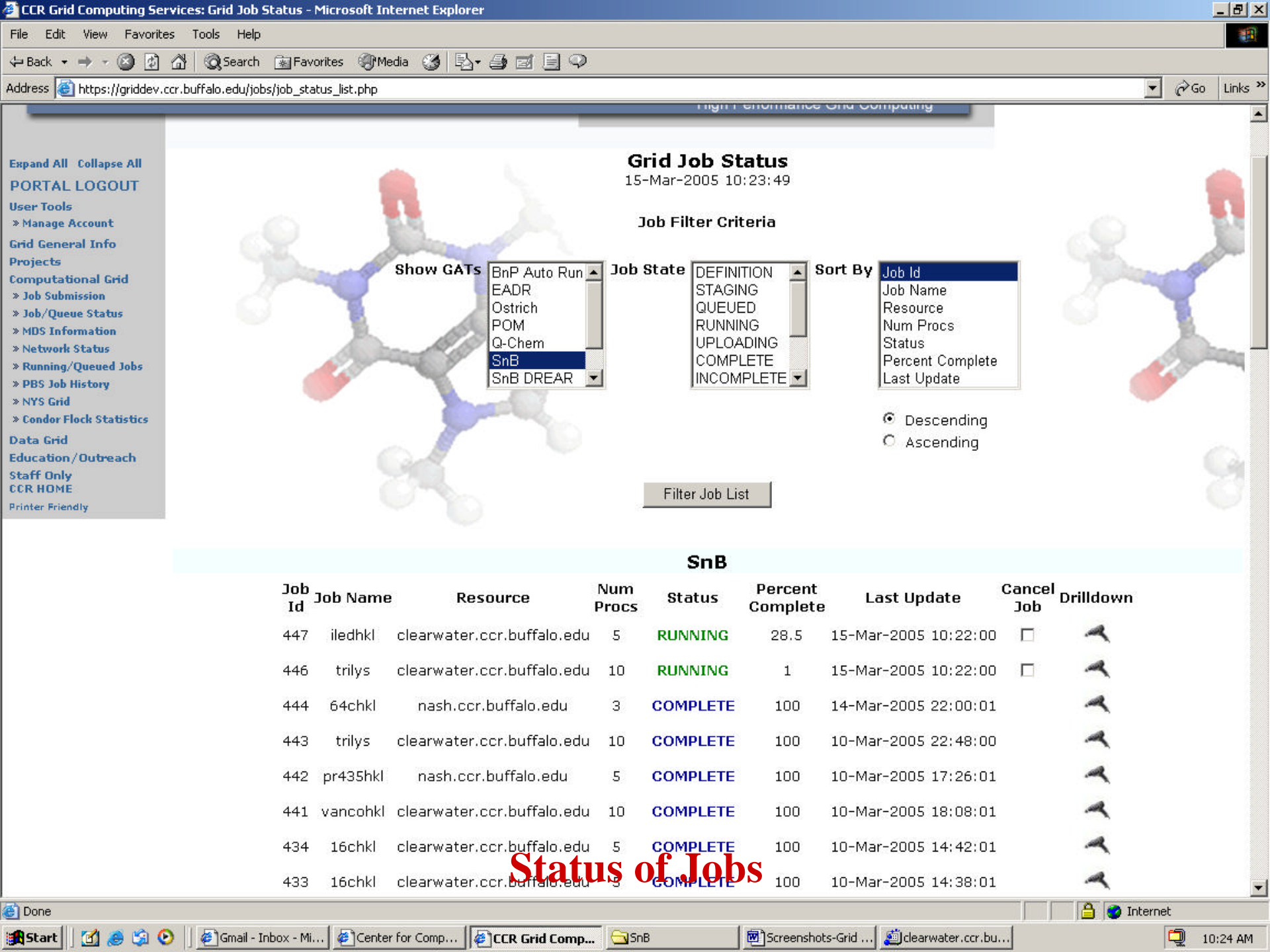
#### Grid Job Trial Histogram

Grid Job 447 Histogram - 285 Trials Complete - Status RUNNING



Click on image for enlarged view.

## Graphical Representation of Intermediate Job Status



- Expand All Collapse All
- PORTAL LOGOUT
- User Tools
  - » Manage Account
- Grid General Info
- Projects
- Computational Grid
  - » Job Submission
  - » Job/Queue Status
  - » MDS Information
  - » Network Status
  - » Running/Queued Jobs
  - » PBS Job History
  - » NYS Grid
  - » Condor Flock Statistics
- Data Grid
- Education/Outreach
- Staff Only
- CCR HOME
- Printer Friendly

## Grid Job Status

15-Mar-2005 10:23:49

### Job Filter Criteria

Show GATS

- BnP Auto Run
- EADR
- Ostrich
- POM
- Q-Chem
- SnB**
- SnB DREAR

Job State

- DEFINITION
- STAGING
- QUEUED
- RUNNING
- UPLOADING
- COMPLETE
- INCOMPLETE

Sort By

- Job Id**
- Job Name
- Resource
- Num Procs
- Status
- Percent Complete
- Last Update

- Descending
- Ascending

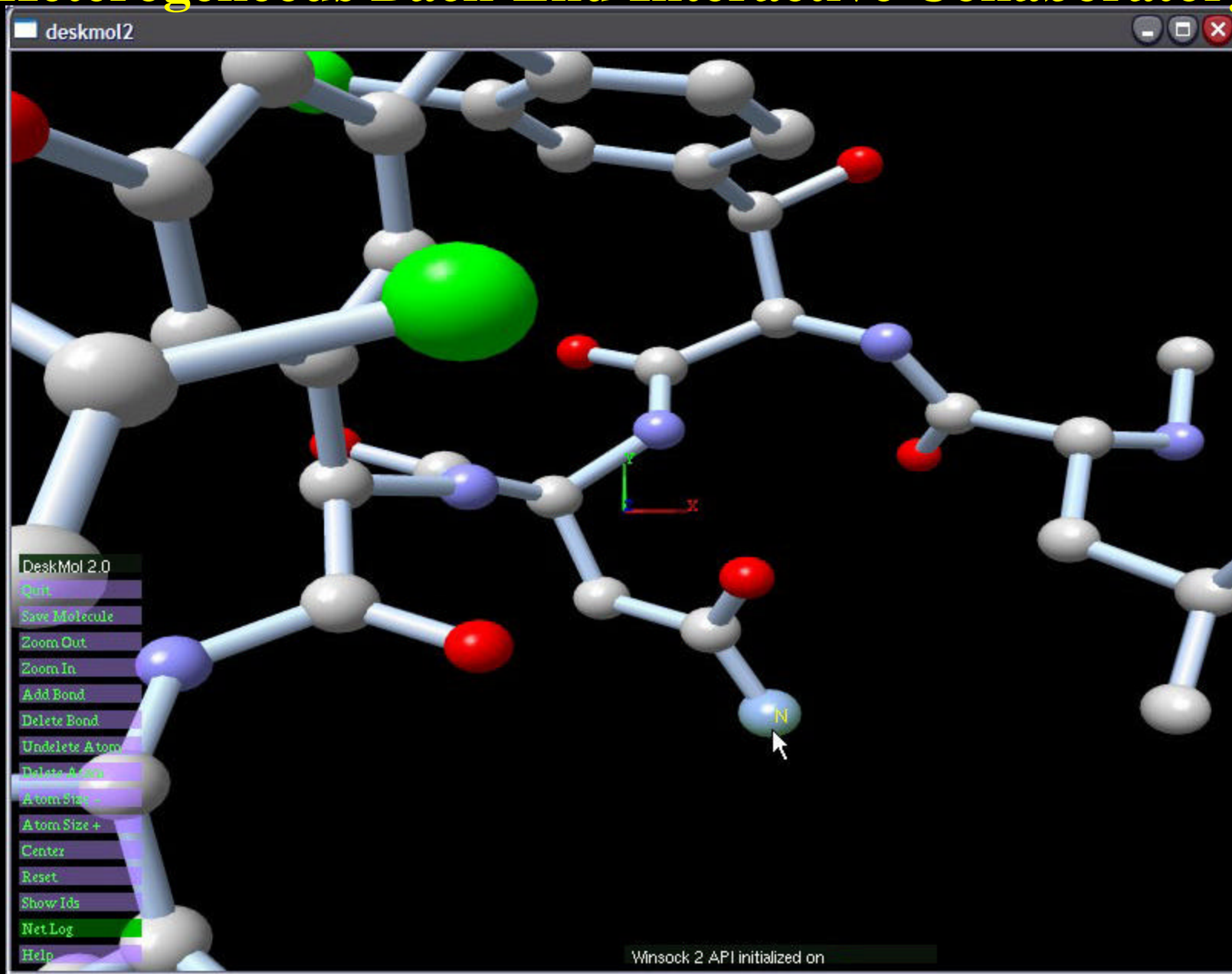
Filter Job List

### SnB

Job Id	Job Name	Resource	Num Procs	Status	Percent Complete	Last Update	Cancel Job	Drilldown
447	iledhkl	clearwater.ccr.buffalo.edu	5	<b>RUNNING</b>	28.5	15-Mar-2005 10:22:00	<input type="checkbox"/>	
446	trilys	clearwater.ccr.buffalo.edu	10	<b>RUNNING</b>	1	15-Mar-2005 10:22:00	<input type="checkbox"/>	
444	64chkl	nash.ccr.buffalo.edu	3	<b>COMPLETE</b>	100	14-Mar-2005 22:00:01		
443	trilys	clearwater.ccr.buffalo.edu	10	<b>COMPLETE</b>	100	10-Mar-2005 22:48:00		
442	pr435hkl	nash.ccr.buffalo.edu	5	<b>COMPLETE</b>	100	10-Mar-2005 17:26:01		
441	vancohkl	clearwater.ccr.buffalo.edu	10	<b>COMPLETE</b>	100	10-Mar-2005 18:08:01		
434	16chkl	clearwater.ccr.buffalo.edu	5	<b>COMPLETE</b>	100	10-Mar-2005 14:42:01		
433	16chkl	clearwater.ccr.buffalo.edu	3	<b>COMPLETE</b>	100	10-Mar-2005 14:38:01		

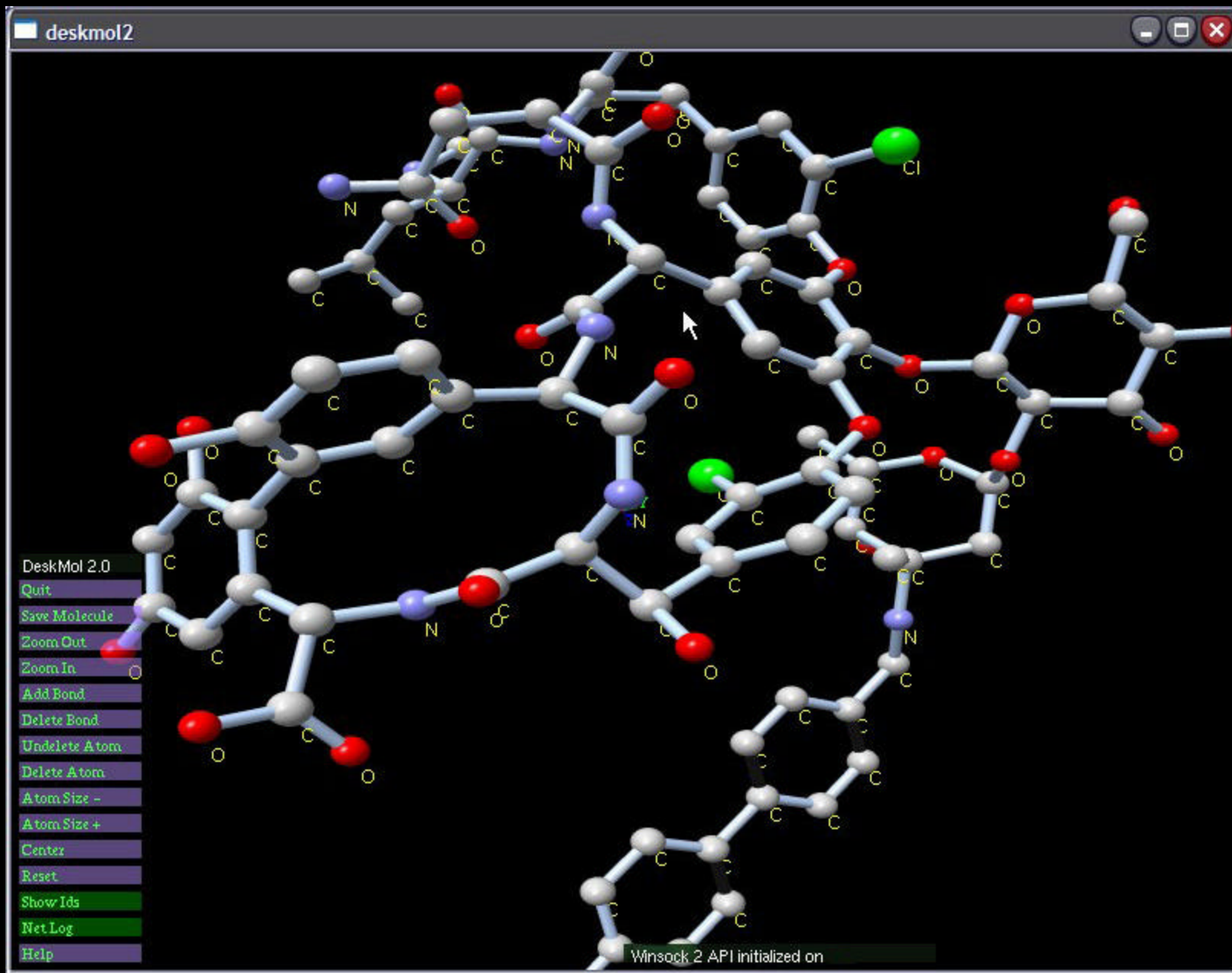
Status of Jobs

# Heterogeneous Back-End Interactive Collaboratory



User starts up – default image of structure.





Molecule scaled, rotated, and labeled.

# Current Efforts

## ■ Grass Roots NYS Grid

- SUNY-Buffalo
- Niagara University
- Canisius College
- SUNY-Geneseo
- SUNY-Binghamton
- Columbia
- Hauptman-Woodward Inst.

## ■ Expand to Ontario??

## ■ Harden

- Dashboard
- Predictive Scheduler

## ■ GRASE VO: Grid Resources for Advanced Science and Engineering Virtual Organization

- (Non-Physics Research)
- Structural Biology
- Groundwater Modeling
- Earthquake Engineering
- Computational Chemistry
- GIS/BioHazards

# Acknowledgments

- **Mark Green**
- **Cathy Ruby**
- **Amin Ghadersohi**
- **Naimesh Shah**
- **Steve Gallo**
- **Jason Rappleye**
- **Jon Bednasz**
- **Sam Guercio**
- **Martins Innus**
- **Cynthia Cornelius**
- **NSF, NIH, NYS,  
NIMA, NTA, Oishei,  
Wendt, DOE**



