

# *UGP and the UC Grid Portals*

**OGF 2007**

Documentation at:  
<http://www.ucgrid.org>

**Prakashan Korambath & Joan Slottow  
Research Computing Technologies  
UCLA Academic Technology Services**

## Academic Technology Services

The image displays three overlapping screenshots of university grid portals, each showing a web browser window with a navigation bar, a main content area, and a sidebar with login options.

- UC Grid Portal:** The top-left window shows the UC Grid Portal with a "Welcome" button and links for "Home" and "Apply".
- UCLA Grid Portal:** The top-right window shows the UCLA Grid Portal with a "Welcome" button and links for "Home" and "Apply". It also features a language dropdown menu set to "English".
- UCI Grid Portal:** The middle window shows the UCI Grid Portal with a "Welcome" button and links for "Home" and "Apply".
- UCSB Grid Portal:** The bottom window shows the UCSB Grid Portal with a "Welcome" button and links for "Home" and "Apply". It also features a language dropdown menu set to "English".

The UCI Grid Portal provides a single web interface to those computational clusters outside of the UCI Grid, including the TeraGrid.

**Compute Power Available in the UCI Grid:**

- Number of Clusters: 1
- Number of Nodes: 340
- Aggregate Peak Performance: 136

**Outside Clusters Accessible from the UCI Grid:**

The UCSB Grid Portal also features a "Help" link and a language dropdown menu set to "English".

The UCSB Grid Portal also features the UCSB and CNSi logos.

## **UGP (UCLA Grid Portal)**

- Joins computational clusters into a Grid
- Under development at UCLA since 2002.
- Built on top of:
  - Globus Toolkit 4.n
  - GridSphere Portlet Framework 2.n (production systems) and 3.n (under development)
  - Zimbra, YUI, and GWT toolkits
  - Shibboleth
  - Tomcat
  - MySQL

## Motivation

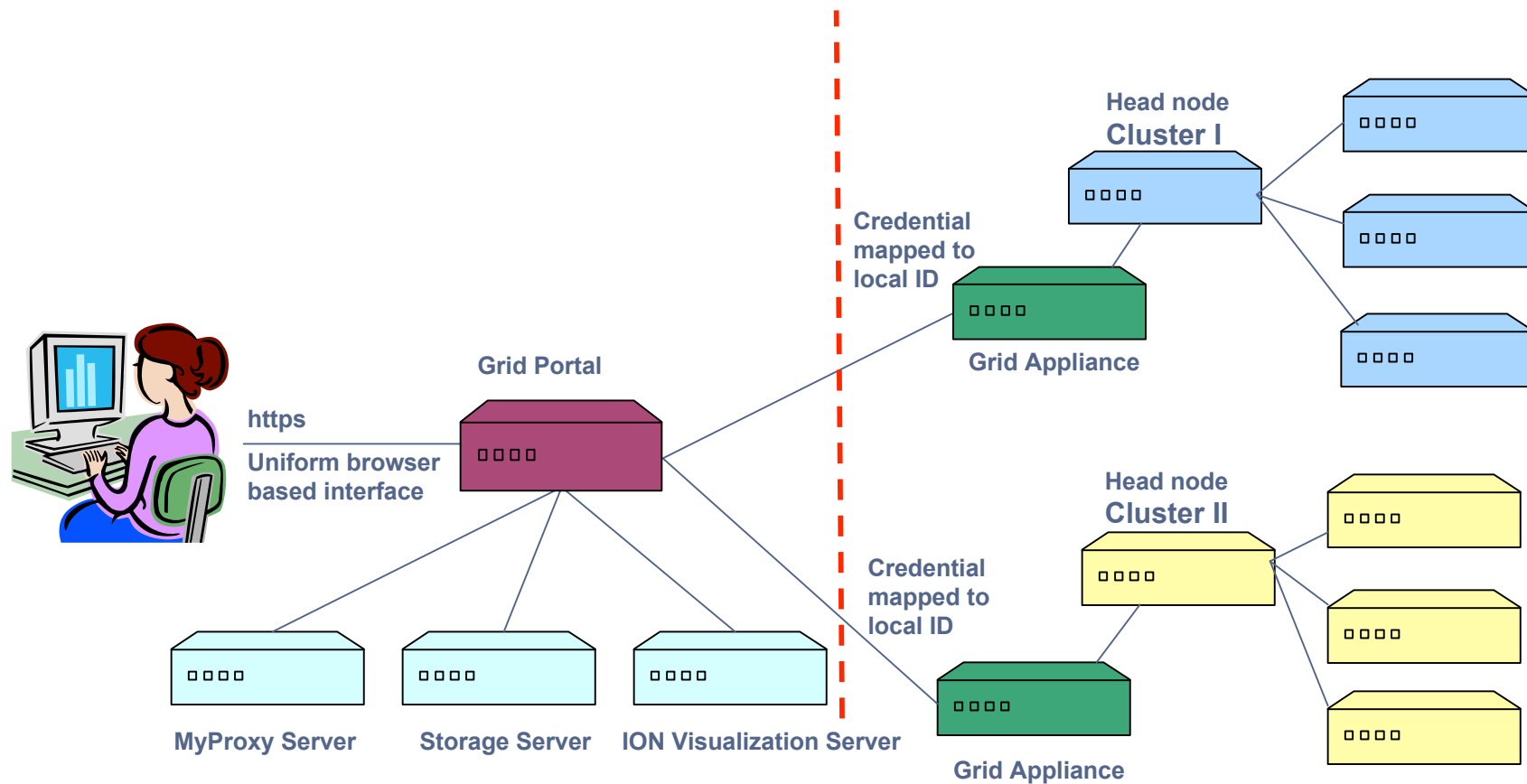
- Need to share resources among the campus clusters:
  - Better equipment utilization
  - Conserve energy
- Diverse cluster ownership and operation
  - Owners reluctant to give login ids to any but their users
- A number of users have login ids on multiple clusters
  - Need to get to them from one interface/location

## Design Goals:

- Do not change the way individually-managed clusters do business. --> Grid Appliance.
- Web Portal:
  - Hide certificates and the complexity of the Grid from users.
    - No software for users to install.
  - Single sign on.
  - Common web interface to all clusters.

## I. Architecture

# Single Campus Architecture



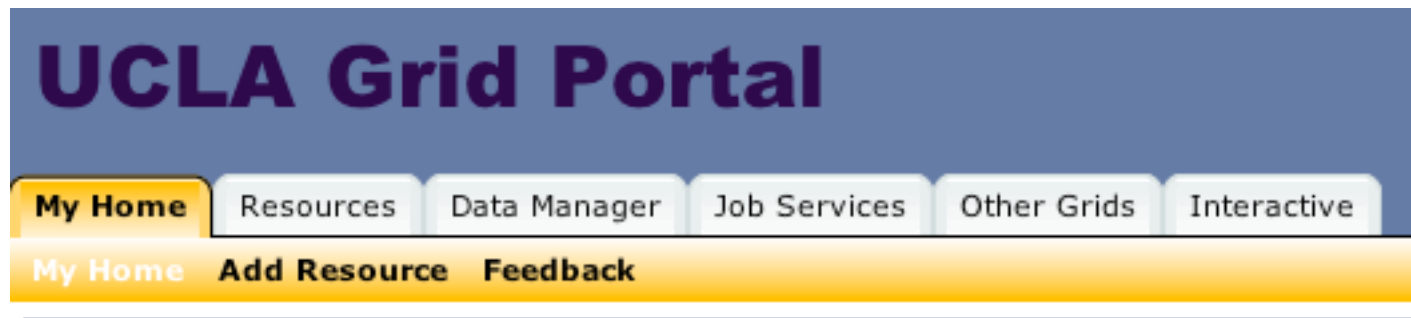
## Grid Appliance

- Enables a cluster to participate in the Grid.
- Provided by the Grid administrators.
- In no way modifies policy decisions at the cluster level.
- Cluster head node can always also be used directly.
- Globus Toolkit is installed on the Appliance.
- Port forwarding for VNC done here.



## Demo

1.



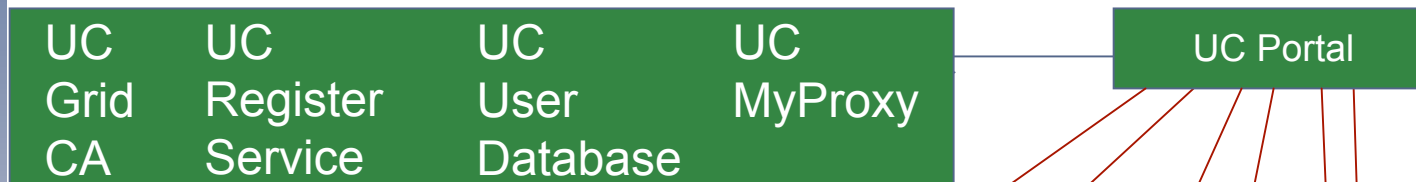
2.

The screenshot shows the "Resources" page of the UCLA Grid Portal. It features a navigation bar with tabs: "My Home", "Resources" (highlighted in yellow), "Data Manager", "Job Services", "Other Grids", and "Interactive". Below the navigation bar is a yellow banner with the text "Resources". The main content area is titled "Resources" and contains a "Resource Discovery" section. This section includes an "Update" button and a table with the following data:

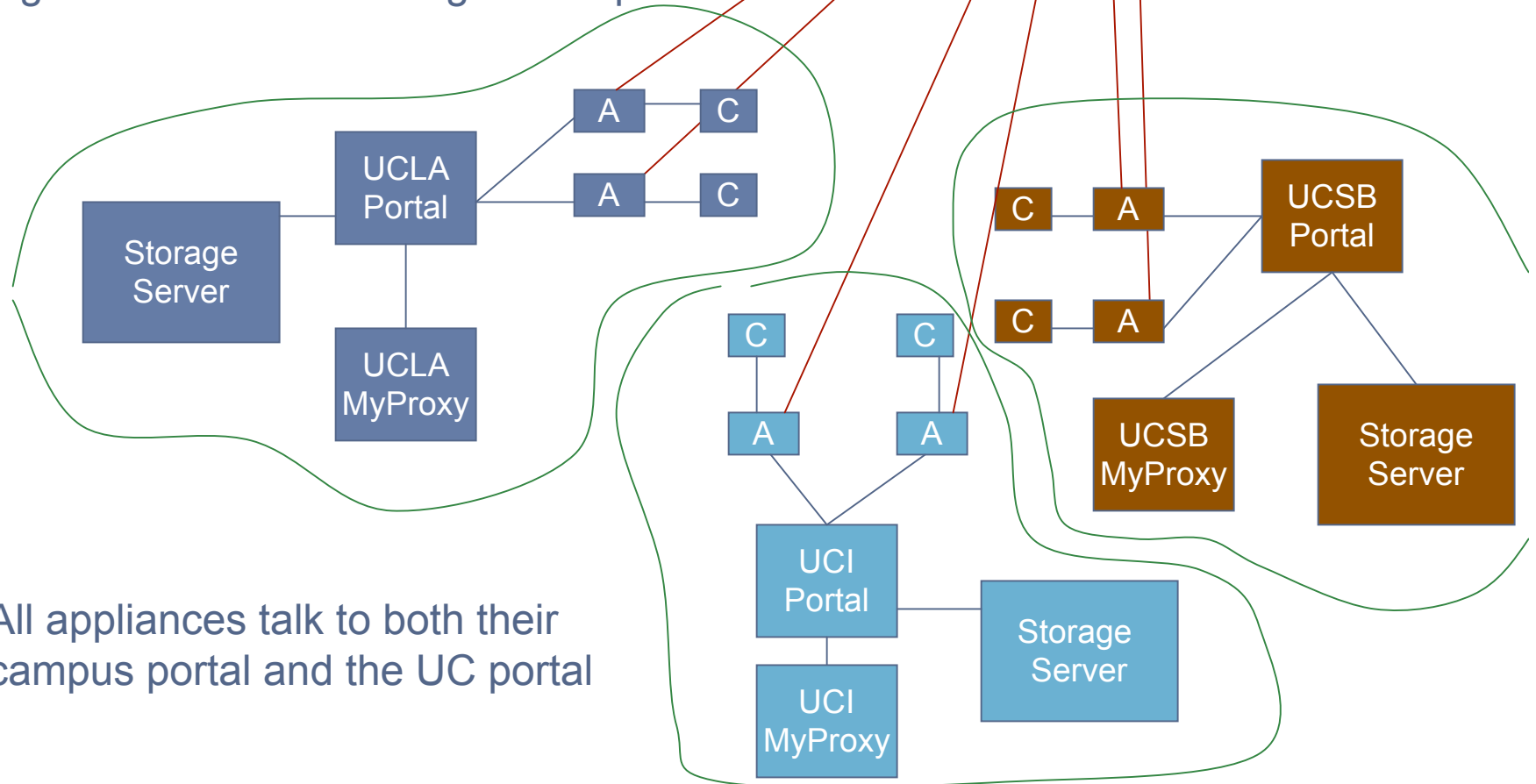
Cluster Name	Status	Load%	Total Nodes	Free Nodes	Down Nodes	Running	Queued	Peak Performance (GFlops)	Queues	Jobs
Hoffman Cluster	↑	61.2	58	19	9	33	37	440	Queues	Jobs
Dawson Cluster	↑	65.2	273	95	0	3	0	4470	Queues	Jobs
Miles Cluster	↑	0.0	10	9	1	0	0	160	Queues	Jobs

## Academic Technology Services

### UC Grid Multi-Campus Architecture



Single CA is shared among all campuses



All appliances talk to both their campus portal and the UC portal

## The UC Grid

- The umbrella Grid and aggregation point for all the campuses.
  - Users can work with clusters from all the campuses
  - Every appliance is open to the Campus Grid Portal and the UC Grid Portal
  - Each user of a Campus Grid can also use the UC Grid Portal
  - Single sign on to transfer files across the campuses.
  - Single Sign on to make use of the cluster and pooled resources across the campuses.

## UC Grid -- Status

- Received UC approval in November 2006.
- Currently have three campuses connected
  - UCLA, UCSB, and UCI
  - 16 clusters (UCLA 10, UCSB 5, UCI 1)
  - 1214 nodes, > 2428 CPUs
  - 12.92 TFlops
- UC Davis and UC Berkeley are in the process of bringing their portals up.

## Demo

**UCLA Grid Portal** Logout  
Joan Slottow

My Home **Resources** Data Manager Job Services Other Grids Interactive

Resources

? Resources

Resource Discovery

Update

Cluster Name	Status	Load%	Total Nodes	Free Nodes	Down Nodes	Running	Queued	Peak Performance (GFlops)	Queues	Jobs
Hoffman Cluster	↑	61.2	58	19	9	33	37	440	Queues	Jobs
Dawson Cluster	↑	65.2	273	95	0	3	0	4470	Queues	Jobs
Miles Cluster	↑	0.0	10	9	1	0	0	160	Queues	Jobs
Neutrino Cluster	↑	50.0	40	18	4	33	0	400	Queues	Jobs

**UC Grid Portal** Logout  
Joan Slottow

My Home **Resources** Data Manager Job Services Other Grids Interactive

Resources

? Resources

Resource Discovery

Update

Cluster Name	Status	Load%	Total Nodes	Free Nodes	Down Nodes	Running	Queued	Peak Performance (GFlops)	Queues	Jobs
Hoffman Cluster	↑	67.3	58	16	9	34	37	440	Queues	Jobs
Dawson Cluster	↑	52.7	273	129	0	4	0	4470	Queues	Jobs
Miles Cluster	↑	0.0	10	9	1	0	0	160	Queues	Jobs
Neutrino Cluster	↑	50.0	40	18	4	33	0	400	Queues	Jobs

## Web Services that make the Hierarchy Possible

- Register Service -- Synchronizes User Creation
- Sync Service -- Synchronizes changes to the Grid
  - When an administrator of a Campus Grid Portal adds/removes a cluster, application or pool
    - The UC Grid Portal is updated in real time accordingly.

## Pools

- Resource Pools
  - Applications and cycles
- Pool-Only Users
  - Students and faculty members who:
    - Do not have accounts on any campus cluster
    - Have low-level or sporadic usage requirements
    - Need for a specific application, compiler, visualization tool
- Cluster users are also pool users
- Target cluster selected by the UGP
- Currently runs applications only

## User Types vs. Portals

### Portal View

		Campus Portal	UC Portal
User Type	Pool-only User	Can submit jobs to the campus pool	Can submit jobs to campus and UC pools
	Cluster User	Can use those clusters on campus he/she can access  Can submit jobs to the campus pool	Best choice for those with access to clusters on different campuses -- can use all clusters, UC-wide, he/she can Access  Can submit jobs to campus and UC pools



## II. Workflow for Handling Usernames/Passwords

## To Login to a Portal a User Needs:

- A certificate
- A gridsphere account

Additionally:

A Cluster User must be added to the GridMap file on the clusters on which he/she has login ids.

A Pool-Only user needs to be assigned disk space on the Grid Portal's file server.

## Apply for Grid Access

Login

Grid Username

Grid Password

Login

Forgot your Grid password?

[Apply for Grid Access](#)

# UCLA Grid Portal

[Welcome](#)[Home](#) [Apply](#)**Grid Account Service**

Application to Access the Grid

To apply for a Grid Username:

- You must be able to prove that you are a member of the campus community authenticating using your local campus credential.
- Your application will then be sent to both the administrator of the resource and the administrator of the Grid as a whole for approval.
- You will be notified as to the status of your application after these administrators have reviewed it.
- After you have received their approval, you will be able to login to both your local and Grid accounts.

**Please Authenticate:**

University of California-Los Angeles ▼



## UCLA Federated Authentication Service

### Sign In

LOGON ID:

PASSWORD:

LOGON TYPE:

[Sign In Now](#)

[forget your ID or password?](#)

### Useful Links:

- [More about UCLA Single Sign-on](#)
- [Sign up for a UCLA Logon ID](#)
- [Frequently Asked Sign-In Questions](#)

### Help for UCLA Federated Authentication Service

- web: [helpdesk@ais.ucla.edu](mailto:helpdesk@ais.ucla.edu)
- email: [helpdesk@ais.ucla.edu](mailto:helpdesk@ais.ucla.edu)
- phone: 310.206.6951

## UCLA Grid Portal

[Welcome](#)[Home](#) [Apply](#)

Grid Account

### Application to Access the Grid

- Please save the password in a safe place. You will need the password to access the Grid.
- Your password should contain only: a-z, A-Z, 0-9, @ # \$ \_
- Password length should not be less than 6 characters.

### Please provide the following information:

Grid Username: (Proposed Grid Username)

Desired Grid Password:

Confirm Grid Password:

First Name

Last Name

Email Address:

Organization or Department:

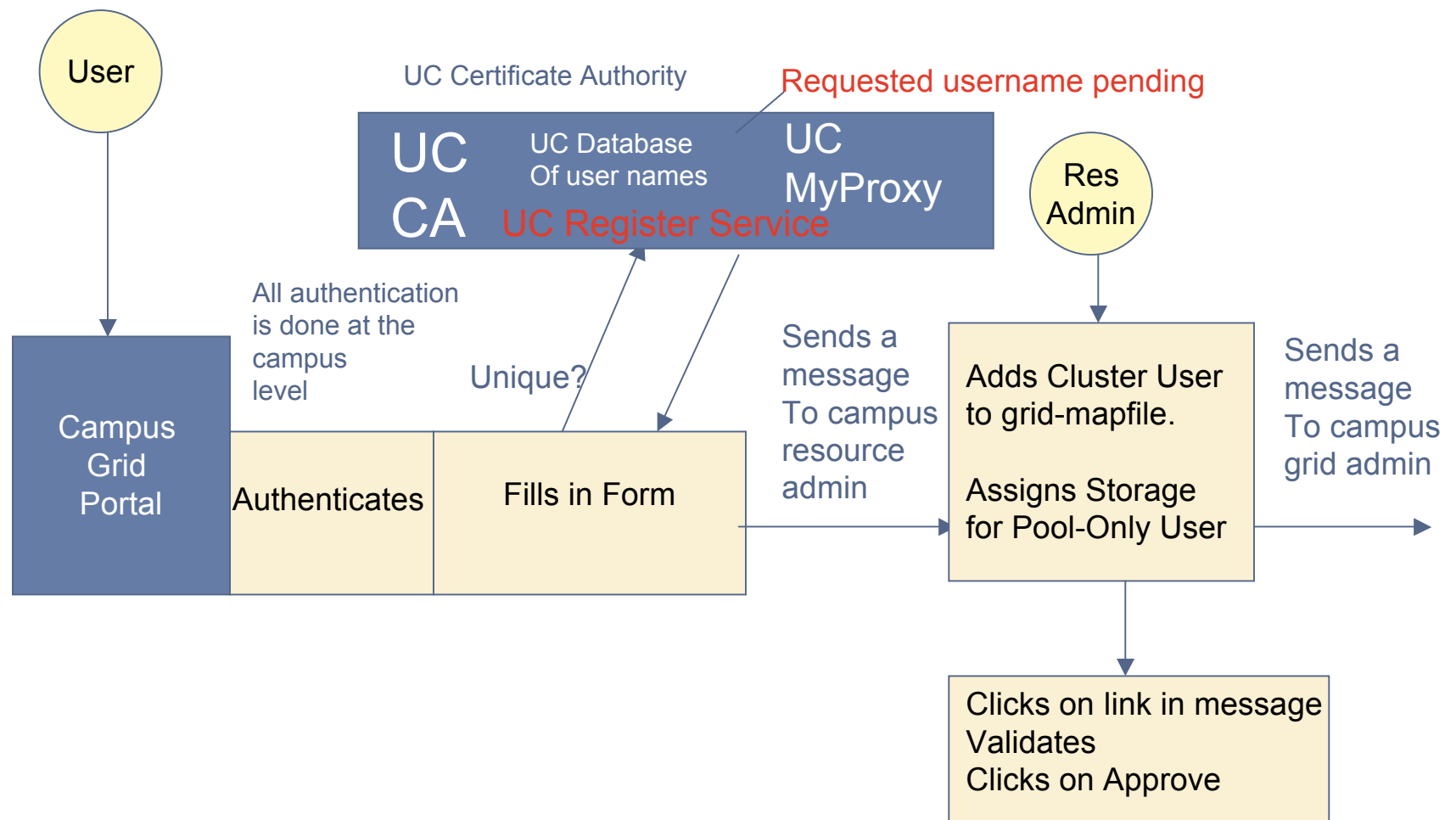
Phone Number:

Project Name:

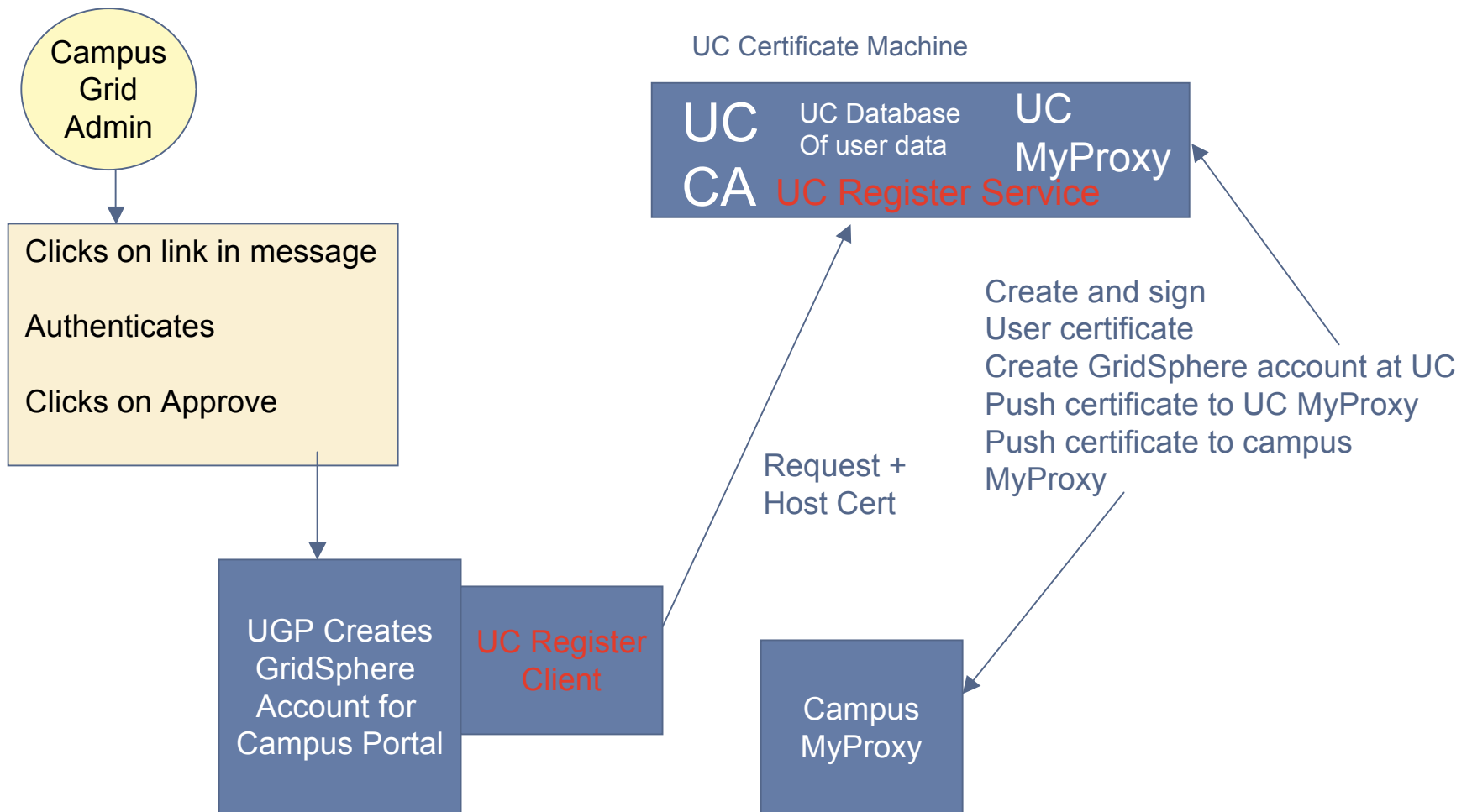
Statement of Work:

Select a Resource:

## New User Workflow - 1



## New User Workflow - 2





### **III. Grid Portal Services and User Interface**

## Demo

File Manager View of /u/home/oacuser:

Name	Mode	Size	Modified
.QVis	0755	472	20071009173219
.ajax	0755	136	20080807234103
.ajax.tar.gz	0644	497	20080821194001
.alkane8.output	0644	158992	20080815191143
.amber	0755	178	20071008155915
.boe401a2.pdf	0644	589532	20080821202335
.cea.pdf	0644	4518931	20080822173802
.debug.out	0644	16	20071008190019
.fadorial	0755	108	20071008155408
.gaussian_s			
hello.sh	#!/bin/sh		
hello.sh.c2i	echo "Hello World"		
hello.sh.o2i	sleep 100		
.mathematic			
.mathematic			
.mathematic			
.matlabtest			
.mpitest			
.MxEnlar			

Edit File: /u/home/oacuser/hello.sh

Application Submission Dialog:

Predefined Applications:

- Hoffman Cluster
  - Gaussian03-Parallel
  - Gaussian03-Serial
  - XMD-Serial
  - Mathematica
  - Q-Chem
  - Stata
  - Matlab
  - Amber8
  - Amber8-Parallel
  - CPMD-Parallel
  - Mathematica 64 bit
- CNSI Cluster
  - Gaussian03-Serial
  - Q-Chem
  - Amber8-Parallel

Mathematica

Submit to: Hoffman Cluster

**Job To Submit**  
Required entries have bold labels.

Application Description: A software package to solve mathematical problems

Job Name: Mathematica-Hoffman

Arguments:

Directory:

Stdin: < /dev/null

**JobType:** Serial

**Job Requirements**

- For serial jobs, the number of processors must be one.
- Some schedulers terminate jobs that have reached their maximum CPU time, others use the maximum elapsed time.

Number of Processors: 1

Memory Per: 400

Job Status

Job Status shows the status of jobs submitted through the Grid Portal Job Services. Jobs submitted from Advanced and Jobs submitted from a cluster head node do not appear here.

Refresh

Program	Submitted	Started	Ended	Target Cluster	Output	Error	Status	Purge
a.out	Fri Feb 23 16:05:54 PST 2007			Hoffman Cluster	Stdout	Stderr	Done	
hello_world	Wed Feb 28 13:32:38 PST 2007	Wed Feb 28 13:32:56 PST 2007	Wed Feb 28 13:33:42 PST 2007	CNSI Cluster	Stdout	Stderr	Done	
Q-Chem-CNSI-Parallel	Wed Feb 28 13:52:13 PST 2007	Wed Feb 28 14:40:32 PST 2007	Wed Feb 28 15:02:07 PST 2007	CNSI Cluster	Stdout	Stderr	Done	

Page 1 out of 1 | 1 | Show all

Pool Job Status

Pool Job Status shows the status of jobs submitted to pooled resources. You MUST Download your output if you want to keep it. Pool job output is periodically DELETED from the cluster on which the job has run after a suitable interval.

Refresh

Program	Submitted	Target Cluster	Pool User	Started	Ended	Stdout	Stderr	Download	Status	Purge
math	Tue Mar 20 14:39:11 PST 2007	neutrino.grid.hosted.ats.ucla.edu	gpool01		Tue Mar 20 14:40:42 PST 2007	Stdout	Stderr	Download	Done	

Generic Job Submission

Submit to: Hoffman Cluster

**Job To Submit**  
Required entries have bold labels.

Job Name:

Executable:

Arguments:

Directory:

Stdin: < /dev/null

**JobType:** Serial

Environment Variables:

**Job Requirements**

- For serial jobs, the number of processors must be one.
- Some schedulers terminate jobs that have reached their maximum CPU time, others use the maximum elapsed time.

Number of Processors: 1

Maximum Memory (MB): 400

Maximum Time (in hours): 1

## Pool Job Setup

- Currently applications only
- Each cluster must have:
  - Guest login id
    - With a certificate
    - In the grid-mapfile
  - Scheduling mechanism for pool jobs
- UGP has:
  - Table of pool applications by cluster
  - Passwords for guest user certificates

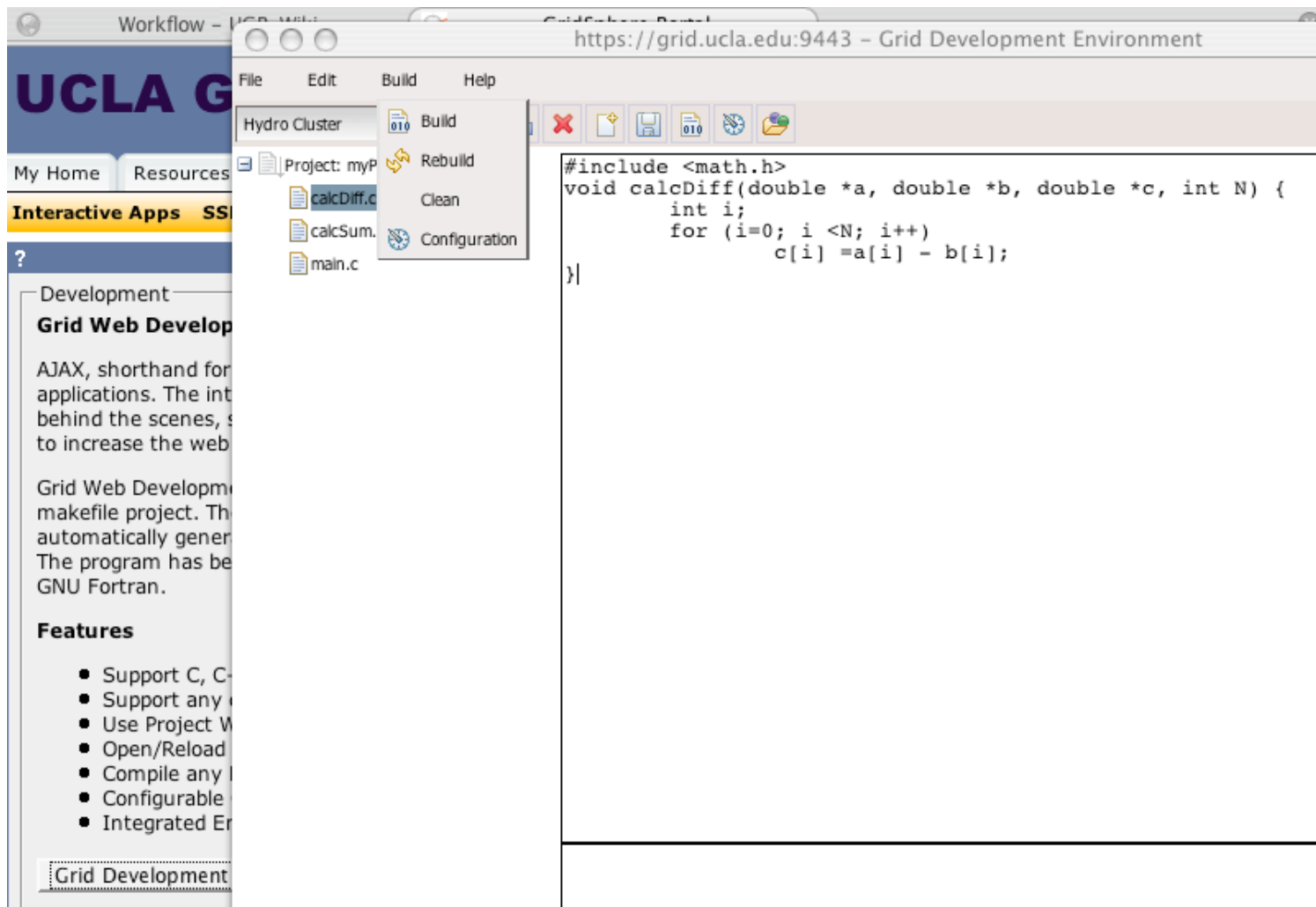
## Pool Job Submission

- UGP:
  - Chooses the best cluster for the job
  - Maps the portal user to the guest user
  - Generates the proxy certificate for the guest user
  - Stages the input files to the target cluster
  - Submits the job to the target cluster
  - Retrieves the output for downloading

## Grid Development Environment

- Like a desktop program development environment.
- Grid Appliances and clusters normally of different architecture
  - Uses an instantaneous queue
- ssh or xterm under Interactive tab is an alternative.

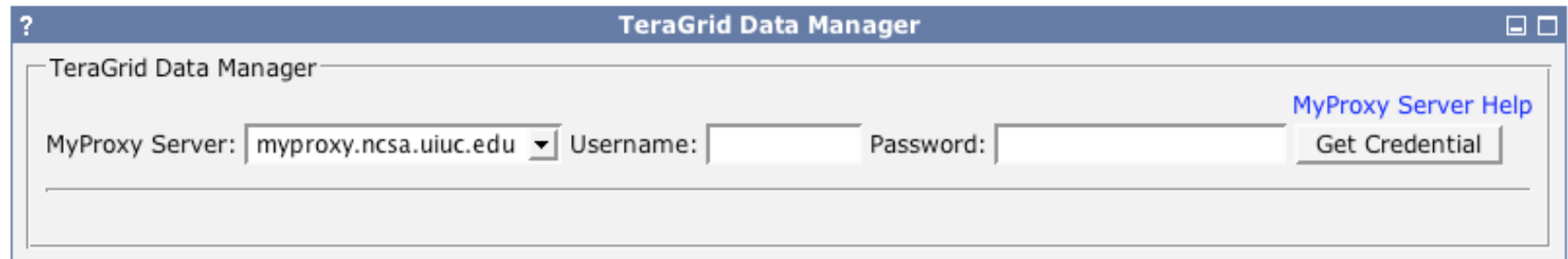
## Demo



## Federation with Other Grids

- UGP can act as a client to some other Grids provided:
  - Their Globus Toolkit nodes are open to all of their users and any computer.
  - They allowed us to get the CA Signing Policy and CA Certificate Public Key for each of their clusters.
  - They have an accessible MyProxy Server.
    - Their users push their proxy certificates to that MyProxy.
- Current policy prevents Globus access from outside:
  - We don't give out the CA Signing Policy and CA public keys for our Grid Appliances
  - We don't give the user's their user certificates
  - Our Grid Appliances are behind firewalls.

## Demo



The screenshot shows a web browser window titled "TeraGrid Data Manager". Inside the window, there is a form with the following elements:

- A title bar with a question mark icon, the text "TeraGrid Data Manager", and window control icons.
- A sub-header "TeraGrid Data Manager" inside the form area.
- A "MyProxy Server:" label followed by a dropdown menu showing "myproxy.ncsa.uiuc.edu".
- A "Username:" label followed by a text input field.
- A "Password:" label followed by a text input field.
- A "Get Credential" button.
- A blue link "MyProxy Server Help" in the top right corner.
- A horizontal line below the input fields.



## Conventional Web Interfaces vs. AJAX User interfaces

- AJAX = asynchronous XMLHttpRequests and JavaScript
- JavaScript runs in the Browser
- Java code on the server processes the requests
- Interactivity gains and efficiency benefits

## Demo

The screenshot displays the 'Data Manager' web interface. At the top, there is a header bar with a question mark icon and the title 'Data Manager'. Below this, a navigation bar includes a dropdown menu set to 'Hoffman Cluster' and a 'Go' button. To the right of the navigation bar is a '[Home]' link and a trash can icon. The main content area is divided into two sections. The top section displays the 'Target: /u/home2/joan/snake/runit' in red text, followed by a list of actions: View | Edit | Visualize | Rename | Download | Change Permissions | Compress File | Remove | Move to Another Directory | Copy to Another Directory | Copy To Another Cluster. Below this, the 'Current Directory: /u/home2/joan/snake' is shown in red text, followed by another list of actions: Create New File | Create New Directory | Upload File | Copy From Another Directory | Copy From Another Cluster. The bottom section features a file list on the left and a code editor on the right. The file list includes icons for various file types and names: data\_cube\_gray\_back.bin, good\_data\_cube.bin, Makefile, MasterSlave.cc, MasterSlave.hh, params3D4P, PI16304, PI16957, PI17765, PI18781, PI30024, PI31069, PI4471, and runit (highlighted in red). The code editor on the right contains the text: 'Enter the text for this file in the box below and click the Save button.' followed by a save icon. The code being edited is: 

```
source /u/local/compilers/intel9/cc/9.0/bin/iccvars.sh
./snake3D4PMainParallel
```

an Cluster

File:

Grid FTP Client...

**Hoffman Cluster** ☐ Overwrite

Hydro Cluster

..  
Command  
gaussiantest  
MPIcodes  
mytestData  
PPLIBNEW  
qchemtest  
ucla\_workspace  
xmd-test  
04019878.pdf  
ajaxdivtag.jar  
alkane8.output  
apache-worm.c  
apiCallsDemo.zip  
AssessorMap.pdf  
babyNeeds.txt  
CPMDLIB.tar.gz

..  
Factor  
Fortran  
gaussiantest  
MPIcodes  
myTEST  
test  
ucla\_workspace  
alkane8.output  
apache\_pb.gif  
debug.out  
DesignDoc.doc.pdf  
gaussian.output  
GridDesign\_July2006.ppt  
macosxlogo.gif  
MPIcodes.zip  
mytest2.sh

Cancel

ome  
ajaxdivtag.jar  
alkane8.output  
apache-worm.c  
apiCallsDemo.zip  
AssessorMap.pdf  
babyNeeds.txt  
Command  
CPMDLIB.tar.gz  
Credit\_cash.pdf  
creditform.doc  
gaussian.output  
gaussiantest  
gdb.pdf  
GT4BuildAServiceV19.pdf  
IMG\_0010.jpg  
IMG\_0086.JPG  
IMG\_0131.JPG  
IMG\_8484.JPG  
info.pl  
installer.sh  
java.log.30776  
java.log.30835  
matlab.m  
meshviewer4.bin  
moodle-latest-17.tgz

```
#!/usr/bin/perl

# thistab is working
my @loadarray = ();

@loadarray = `ghost | awk '{print \$4}'`;

my $count = 0;
my $totalload = 0.0;
my $totalnode = 0;
my
my
fo:
{
#
}

$totalload = 1 - ($freenode / ($totalnode - $downnode));

print("\nTotal Load: $totalload \n");
print("\nTotal Node: $totalnode \n");
print("\nFree Node: $freenode \n");
print("\nDown Node: $downnode \n");

# "gw") gram_status="Pending";
# "s") gram_status="Suspended";
# "h") gram_status="Suspended";
```

## Upload

Browse...

Remove

Add

Browse...

Remove

Add

Browse...

Remove

Add

## Upload Status:

OK

Cancel

## IV. Interactive Applications

## Interactive Applications via VNC

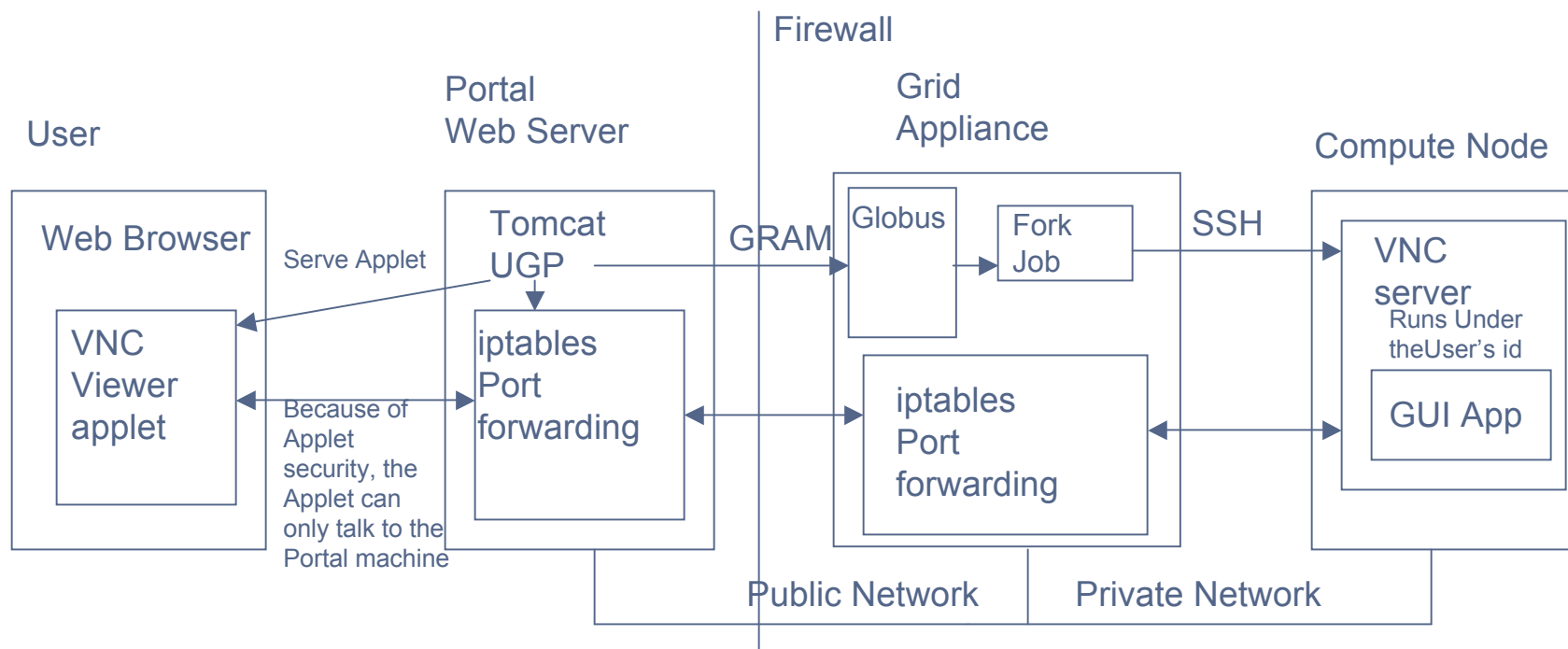
- Pioneers in this area are:
  - Purdue University's nanoHUB
  - University of Florida's In-VIGO
  - University of Texas Advanced Computer Center (TACC)

## Prerequisites

- Password-less SSH access from the Grid Appliance to the cluster node on which the VNC Server will run.
- The VNC Server and TWM Window Manager installed on the cluster node. (They are part of the standard linux installation.)
- OpenSSH on the cluster node must be version 3.6.1 or higher.

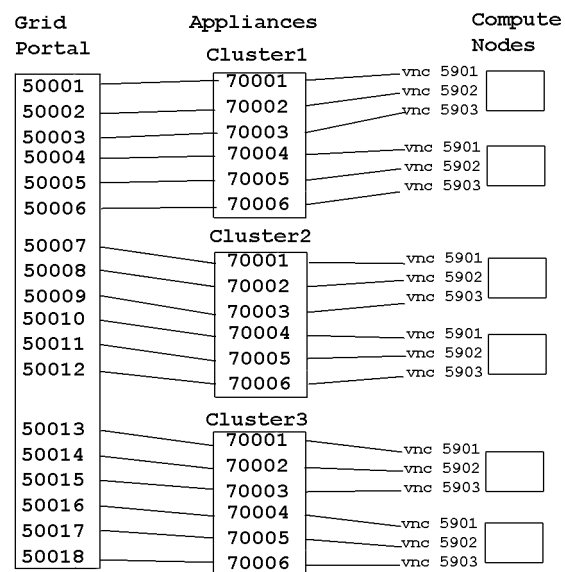
## Academic Technology Services

There is 1 Grid Portal  
One Grid Appliance per cluster  
Multiple VNC servers can be  
Running on an Appliance simultaneously

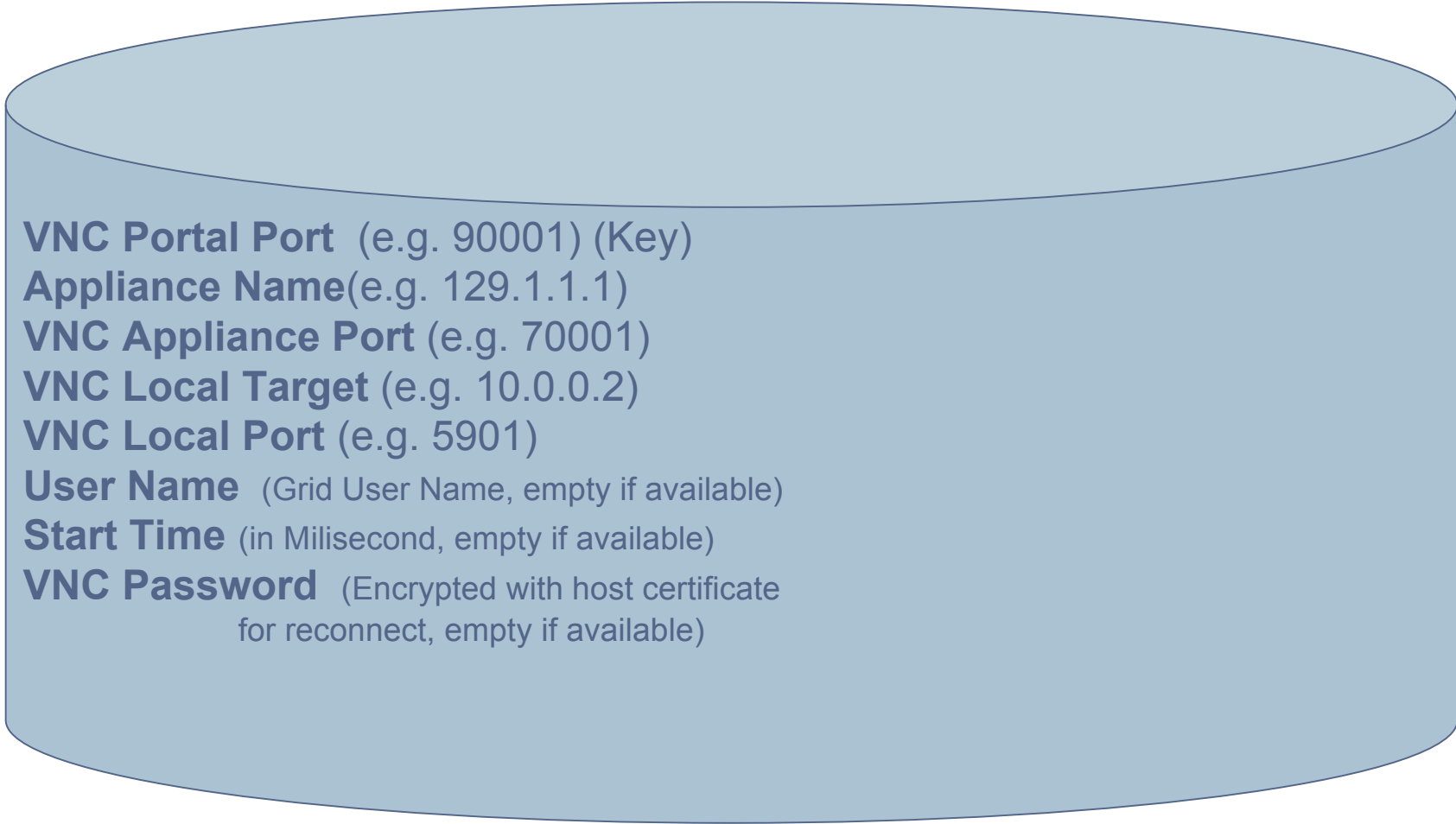




## Iptables set up



## VncPortMap Table



**VNC Portal Port** (e.g. 90001) (Key)  
**Appliance Name**(e.g. 129.1.1.1)  
**VNC Appliance Port** (e.g. 70001)  
**VNC Local Target** (e.g. 10.0.0.2)  
**VNC Local Port** (e.g. 5901)  
**User Name** (Grid User Name, empty if available)  
**Start Time** (in Milisecond, empty if available)  
**VNC Password** (Encrypted with host certificate  
for reconnect, empty if available)

## Killing VNC Server Sessions

- Users can have up to n concurrent sessions
- They can logout of the grid portal and later log back in and reconnect to the same VNC session.
- They can purge the VNC session.
- The VncPortMap table is updated when a VNC Server session times out.

# Demo

The screenshot displays the UCLA Grid Portal interface on the left and the MATLAB 7.4.0 (R2007a) desktop environment on the right.

**UCLA Grid Portal (Left):**

- Header: My Home, Resources
- Section: Interactive Apps SSH
- Section: Interactive Application
- Section: Interactive Appli
- Text: To run the interactive
- Section: Hoffman Cluster
- Text: Xterm
- Section: CNSI Cluster
- Text: Xterm
- Section: Hydro Cluster
- Text: Xterm
- Section: Dawson Cluster
- Text: Xterm
- Table with columns: Application Name, h
- Page 1 out of 1 | 1 |
- Refresh

**MATLAB 7.4.0 (R2007a) (Right):**

- File Edit Debug Desktop Window Help
- Current Directory: /u/home2/joan
- Shortcuts How to Add What's New
- Current Directory Workspace
- Table of files and folders:

All Files	Type	Size
.fontconfig	Folder	
.globus	Folder	
.matlab	Folder	
.mozilla	Folder	
.netscape	Folder	
.ssh	Folder	
.trash	Folder	
.vnc	Folder	
lglobus	Folder	
ajax	Folder	

Command Window

To get started, select [MATLAB Help](#) or [Demos](#) from the

Command History

```
%-- 8/17/07 7:10 AM --%  
%-- 8/17/07 7:23 AM --%
```

## appForm

- Previous Work
  - Purdue/nanoHub Rappture
  - San Diego Supercomputer Center (SDSC) and National Biomedical Computation Resource (NBCR) Project Gemstone

Welcome

Administration

G

Job Status

Generic Jobs

Appli

?

Application Submission

Predefined Applications

**Hoffman Cluster**Gaussain03-  
ParallelGaussain03-  
Serial

XMD-Serial

Mathematica

Q-Chem

Stata

Matlab

Ameber8

Amber8-  
Parallel**CNSI Cluster**

Gaussain03-

Submit to: **Hoffman Cluster**

Application Description: Ab initio quantum chemistry package Instructions to use arguments, namely the inputfile name and scratch directory added by default. Users need to add only the input file

Job Name: 

Executable: /u/local/apps/qchem-3.0/exe/qcprog.exe

Arguments: **Directory**

- If you specify a directory, your job will be run in that directory
- If you don't specify a directory, your job will run in your home directory
- Unless an absolute path is specified for any file used in the job, the file will be searched in the directory specified or your home directory, if omitted.

Directory: 

Stdin: /dev/null

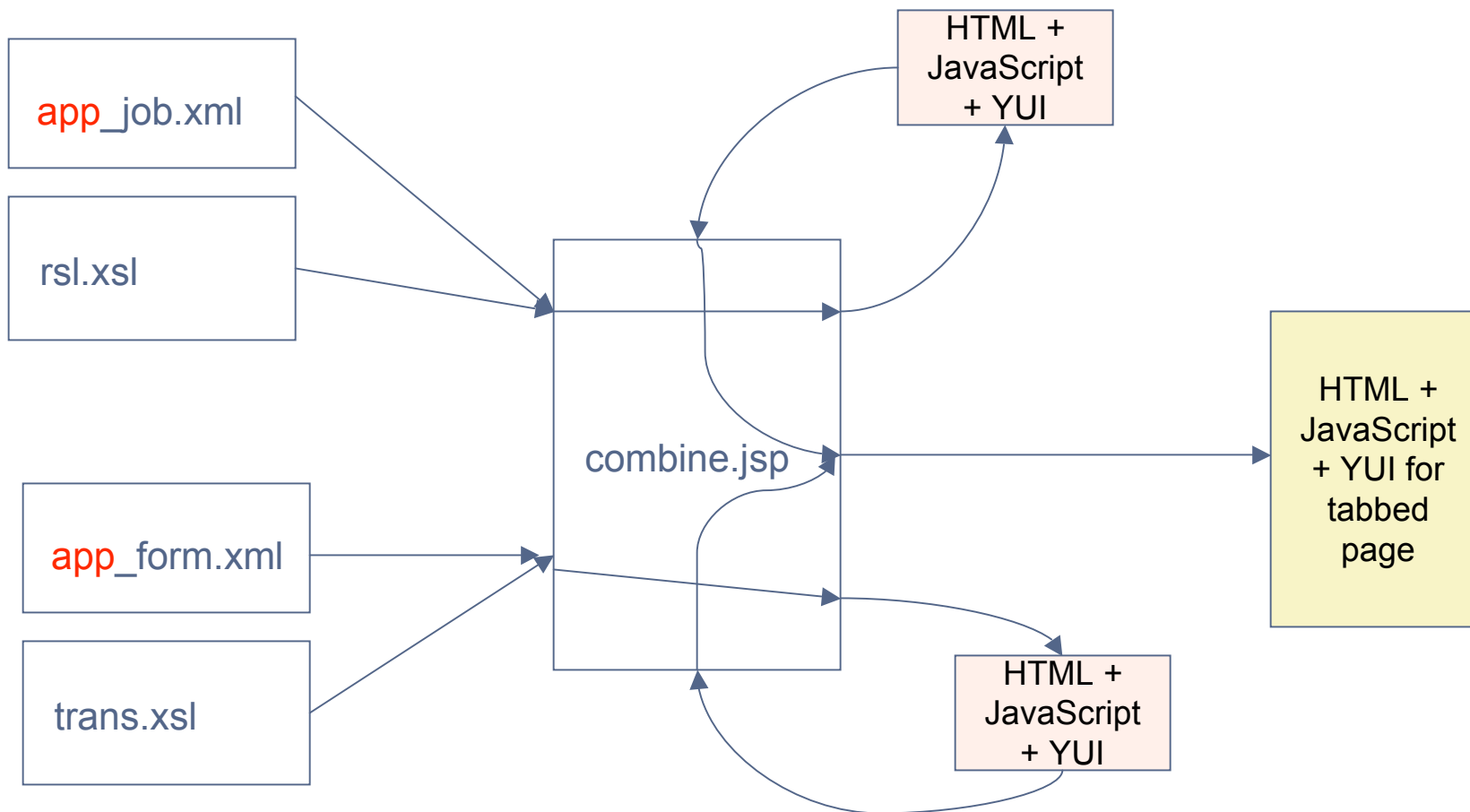
Job Type: mpi

**Job Requirements**

- For serial jobs, the number of processors must be one.
- Some schedulers terminate jobs that have reached their maximum elapsed time.

Number of  
Processors:Max Time  
(in hours):

## How appForm Works



`app_job.xml`, `app_form.xml` and `app_form_input.xsl` must be written for each app.

## Demo

Save

Load

[Main Page](#)**REM**[Optional REM Parameters](#)

### General Q-Chem Processing Options for This Job

**Basis:**

STO-2G

**Current Basis:**

test

**2nd Basis:**

STO-3G



Must be smaller than 1st Basis

**Exchange Level:**

(HF) Hartree-Fock

**Effective Core Potential (ECP):**

No Pseudopotential



Useful for large molecules

**Correlation:**

None

**Job Type:**

(SP) Single Point Energy



Specifies the type of calculation.



## How appForm Works on Save

