

A Web Portal for CMS Grid Job Submission and Management

David Braun, Norbert Neumeister
Purdue University

CHEP09
21-27 Mar 2009, Prague, Czech Republic

Introduction

- Investigated possibilities to provide a Grid Portal for OSG and TeraGrid
 - Reduce the complexity of the user Grid experience
 - Reuse experience from others portals like *nanoHub.org*
 - Evaluated different technologies
- Developed a prototype
 - Makes use of CMS as a use case
 - It is not part of the official CMS tool set
 - Currently this is a missing functionality of the official CMS tool set
 - We provide this service to a local CMS Tier-2 community at Purdue

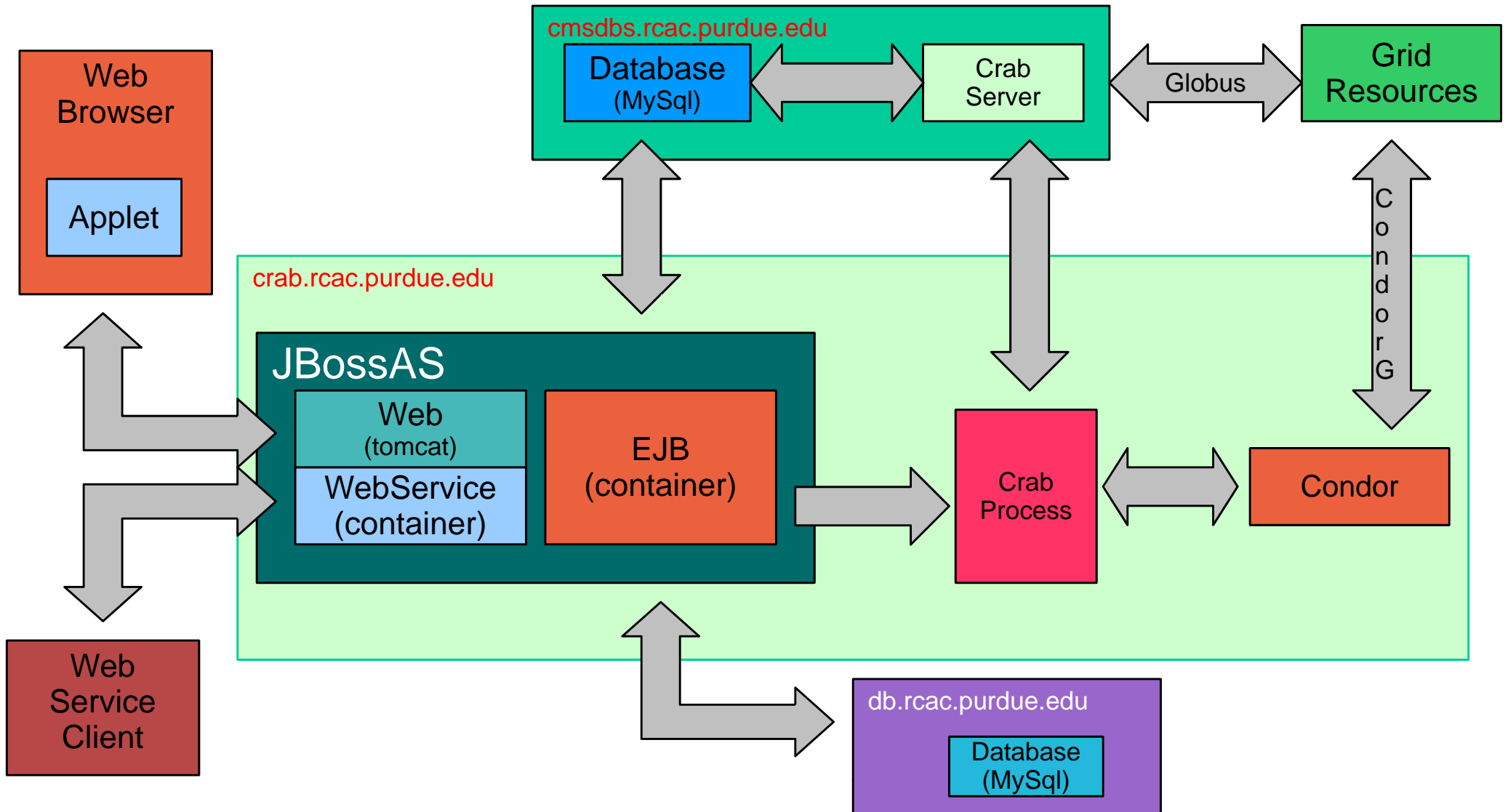
Why Are Portals Important?

- Users
 - Installation of software and maintaining versions is not easy
 - Currently exposed to different flavours of Grid middleware
 - GUI assisted interaction for the non-hardcore UNIX user
- Grid
 - Not so user friendly at times
 - Requires even more software and services to be installed
 - No mobility of public/private keys
 - Few platforms supported
- **How can we reduce the complexity of the user Grid experience?**

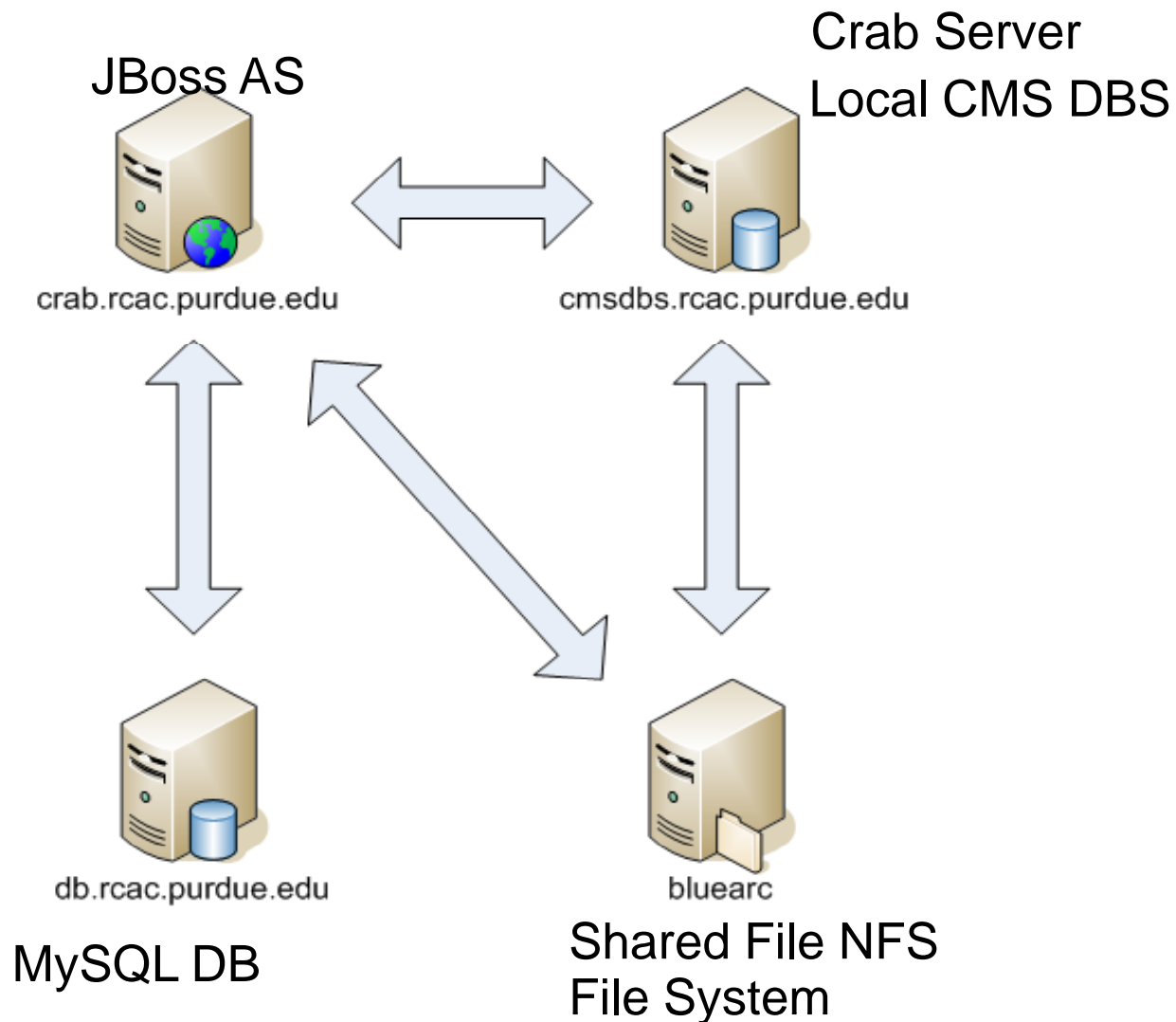
Design Goals

- Develop Web Portal for CMS Grid Job *Submission, Management and Monitoring*
- Industry standards based development
 - J2EE, IDE, Web Frameworks, etc.
- Rapid development of GUI interfaces
 - IDE assisted development
 - Java Server Faces component model
- Multiple service entry points, Web Service, RMI, etc.
- Scalability
- Both portal and Grid security management
- Reusable services and components
- Multiple Grid domains

Architecture



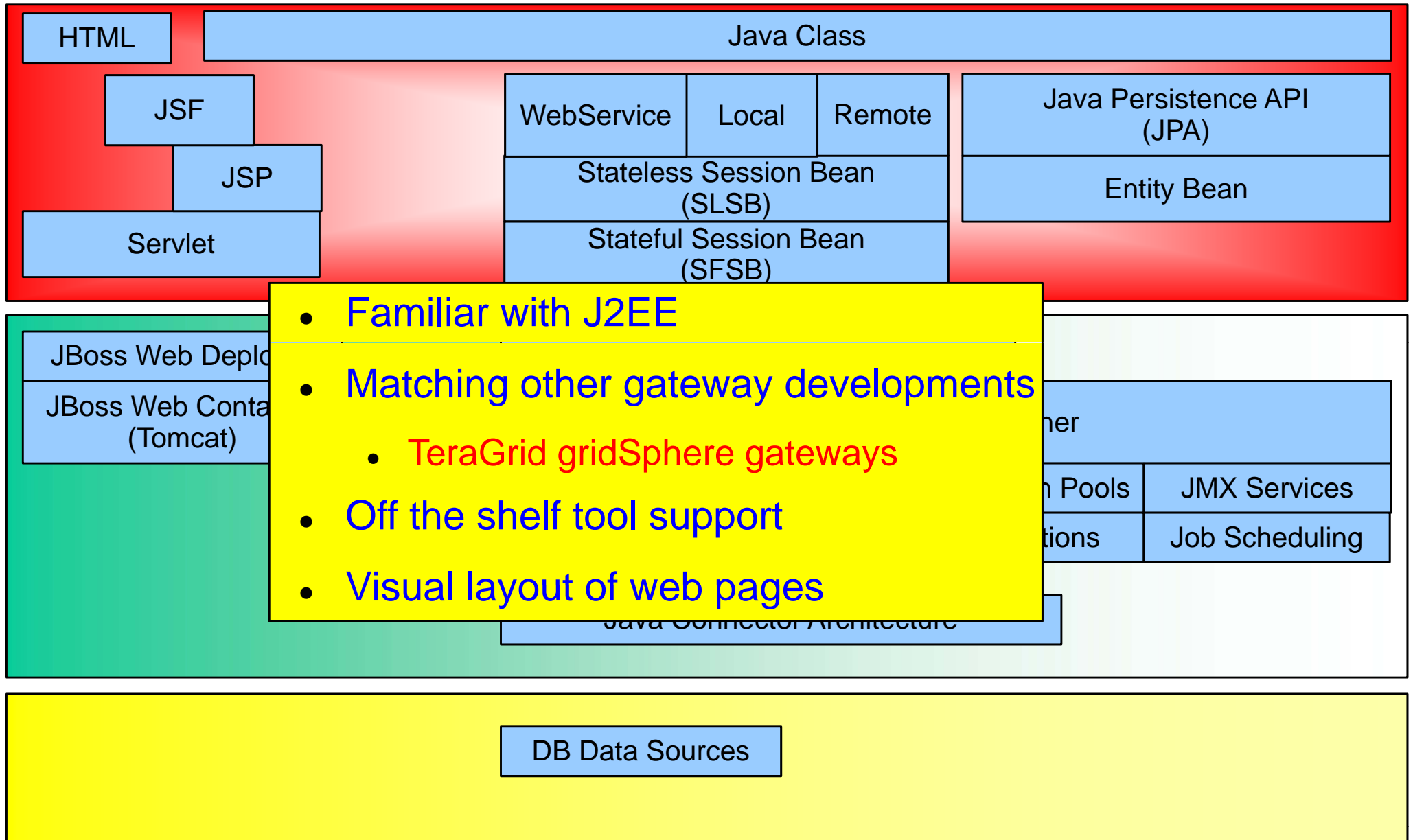
System Diagram



Key Functionality

- Reduce the complexity of Grid certificates
- Interactive GUI within a web browser or deployable in a seamless browser centric manner
- The ability to scale to both in breadth and depth to handle a diverse population of users and services
- Upload, Templates, Defaults
 - Allow for full range of complexity, default, modify default, custom user configurations, upload private modules and libraries
- Job Submission, Status, Termination
 - Using both Grid or non-Grid resources
 - Flexibility: Not bound to a given Grid submission tool (CRAB, etc.)
- Project History, Feedback, Archiving Project/Files, Bookkeeping

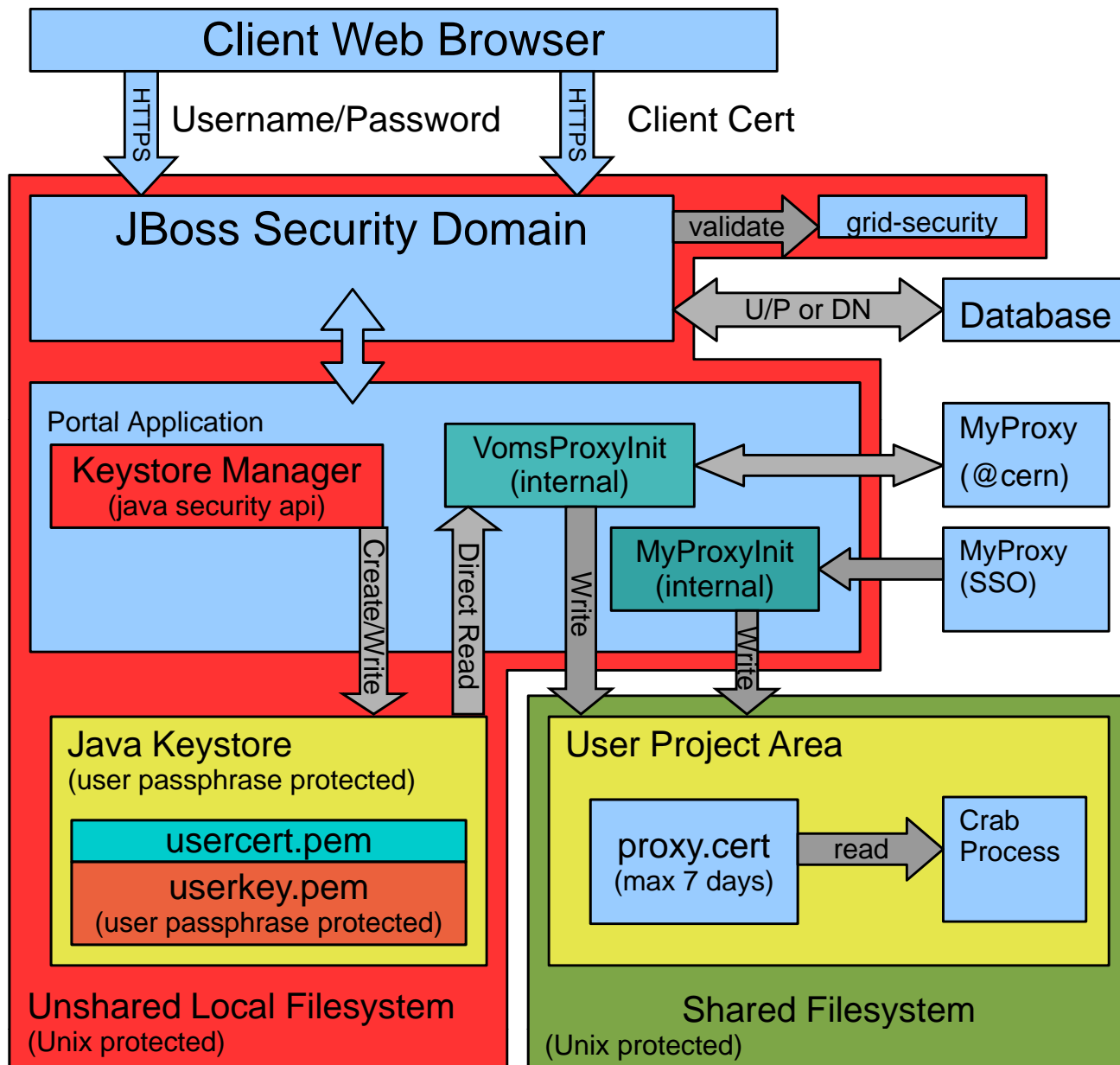
Technology



Technology List

- Java SDK1.6
- Java Server Faces (JSF) JSR-314
 - Woodstock Components Implementation
 - Facelets, Dojo, DWR, Json
- Java Server Pages (JSP)
 - JSP Standard Tag Library (JSTL)
- Java Architecture for XML Binding (JAXB 2.1)
 - XML object serialization
- Java Management Extensions (JMX)
 - Service management
- Java Persistence API (JPA)
- J2EE
 - Servlets and Enterprise Java Beans (EJB)
- Netbeans Integrated Development Environment (IDE)
 - Ant – build management
 - Visual Web Page Tools – graphical layout of web pages
- JBoss
 - Tomcat – web server
 - Quartz – job scheduler
- Thirdparty
 - Apache Commons FileUpload,IO
 - JUpload Applet – file upload
 - Bean Shell – java scripting
 - Jglobus, glite – grid interfaces
 - Ini4j – ini file parser

Security



- **Local Users**

- Portal runs as Apache user not as root

- **Keystore**

- Unique user passphrases required
- Key passphrase required
- Passphrases will never be cached
- Can impose stronger passphrase restrictions

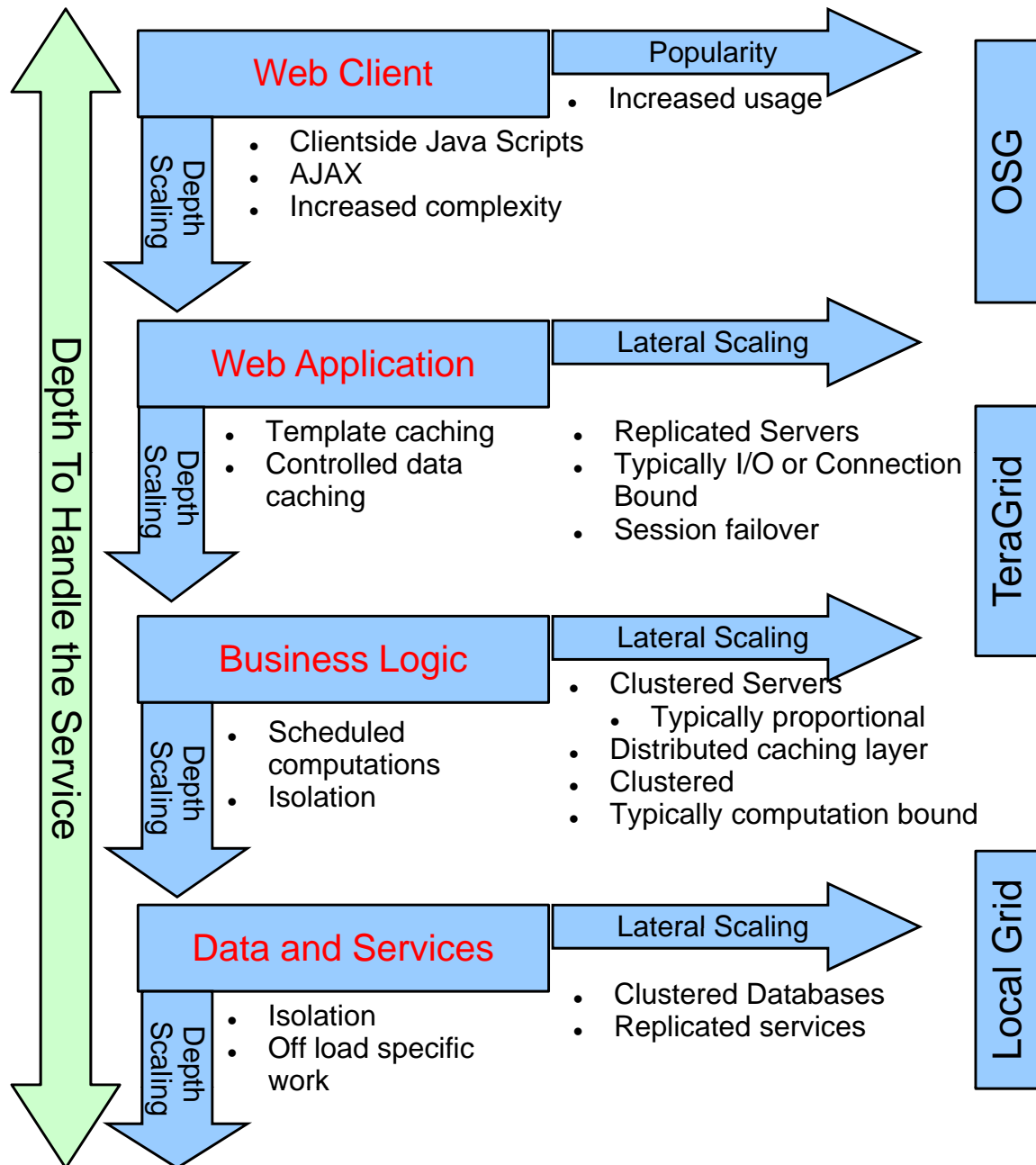
- **Proxy**

- Max lifespan
- Generation handled internally
- Can impose voms rules

- **Disaster**

- Delete keystores and proxies
- Email users

Scaling Where/Why



Bottlenecks

- CPU Bound
- Network Bound
- Connection Bound

Distribute the Load?

- More hardware
 - Coupled/Uncoupled
- Smarter software
 - Complexity increases
 - Polling vs Pushing
 - Looking to enterprise tools

Current Functionality

- Grid Security
- Projects: create, share, clone
- CRAB functions (Cms Remote Analysis Builder)
- Project monitoring
- CRAB Server integration
- CMS users only
 - Users validated against CMS VOMS server for job submission
 - Grid Certificate required to submit jobs

Grid Security

- Username/password authentication
- Client certificate authentication
- Dynamic fallback to username/password if no certificate is detected
- Key management based upon Java keystore
- Valid VOMS proxy can be created from either local keys or fetched from remote proxy

Project Management

- Project creation
 - CMSSW and crab_0_* directories
- Crab wizard tool for easy configuration
- Template with macro substitution
- Project sharing and cloning
- Project file browsing and download
- CMSSW file upload based upon an applet
- Edit CRAB and CMSSW config files with auto saving

CRAB Functions

- CRAB is a Python program intended to simplify the process of creation and submission of CMS analysis jobs into a grid environment
- Subset of CRAB functions currently supported
 - Submit, kill, clean, getoutput, status, postmortem, resubmit

Project Monitoring

- Projects can be selected for monitoring
- Selected projects are added to a round robin pool to be periodically updated
 - Caching results allows for periodic polling from clients
- Ajax based table
- Sorting and colorization of table fields
- Underlying data model matches CRAB server implementation
- Single click access to job log files

CRAB Server Integration

- Clone user CRAB jobs from server into the portal
- Updates are based upon user requests
- Do have the ability to poll CRAB server database to get current job status and merge this status with jobs running on portal

Typical Workflow

[Home](#)

[Register](#)

[Start](#)

[About](#)

CMS Grid Submission Portal

Welcome to the CMS Grid Submission Portal

Now that the LHC is on its way, start doing analysis of the new data that is provided. This portal is provided by Purdue CMS Tier-2 as a service to the CMS user community to enable the excellent science that the CMS project will provide.

Sincerely, [CMS Tier-2 Facility](#)

[CMS General](#) [OSG News](#) [RCAC News](#)

[iCMS interruption 2009-03-19](#)

by Gilles Raymond - Mar 18, 2009

Dear all, due to this ...

[iCMS reboot](#)

by Gilles Raymond - Jan 23, 2009

Dear all, IT requested a reboot (after a kernel upgrade) on the iCMS web servers. The web servers will therefore be rebooted ...

[iCMS stop](#)

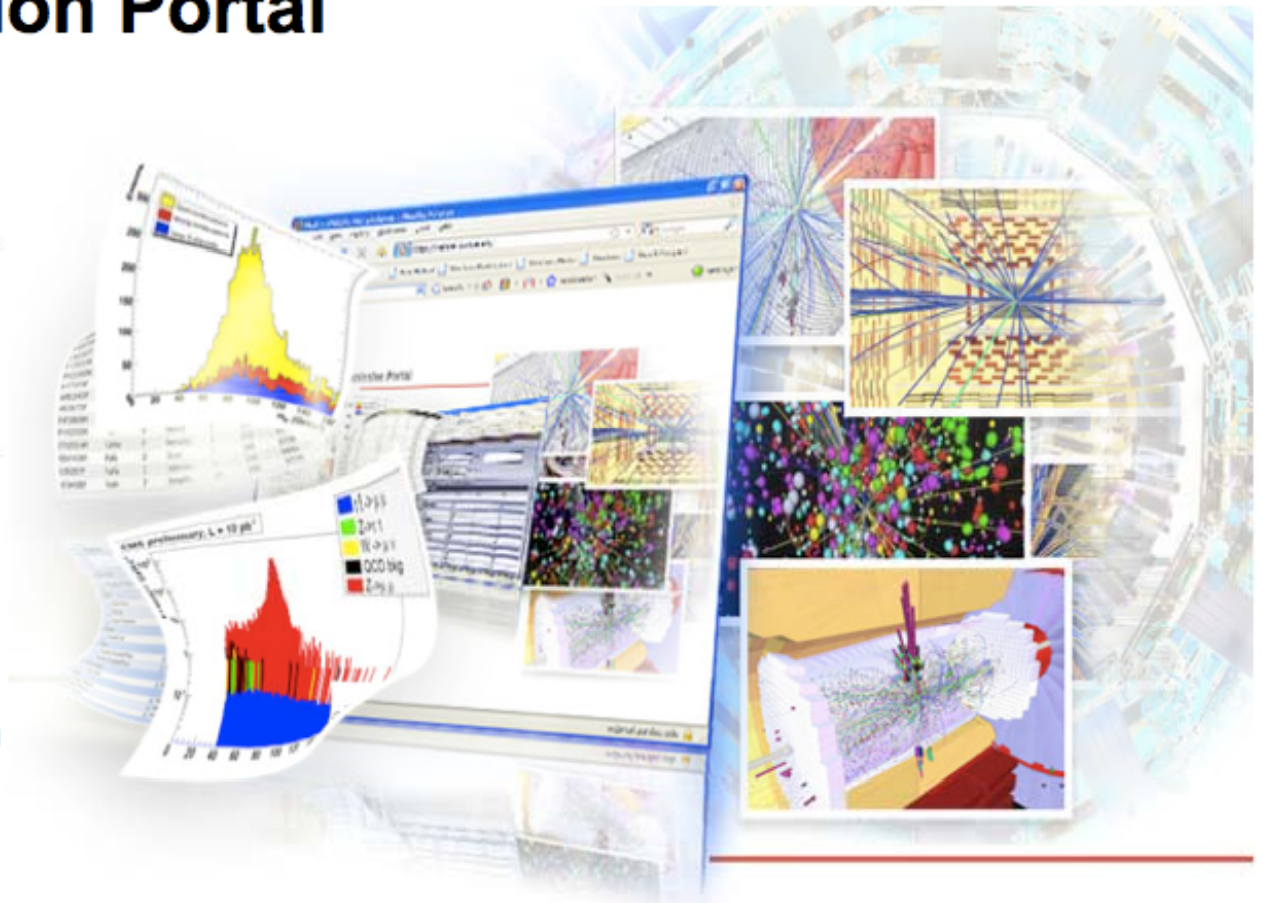
by Gilles Raymond - Jan 19, 2009

Dear all, iCMS will be stopped tomorrow Tue 20/01 between 7.30 and 8.00 AM
Reason: update IT authentication certificate no ...

[Annual Power cut 4/01/2009](#)

by Gilles Raymond - Dec 19, 2008

Dear all, In accordance to the general power cut on Sunday 4th 2009 from 7.00 AM to 9.00 ...



Typical Workflow

CMS Grid Submission Portal

Users: 1 of 2

Projects Monitoring Profile Tools Help Admin Username: David Braun 920886 Logout

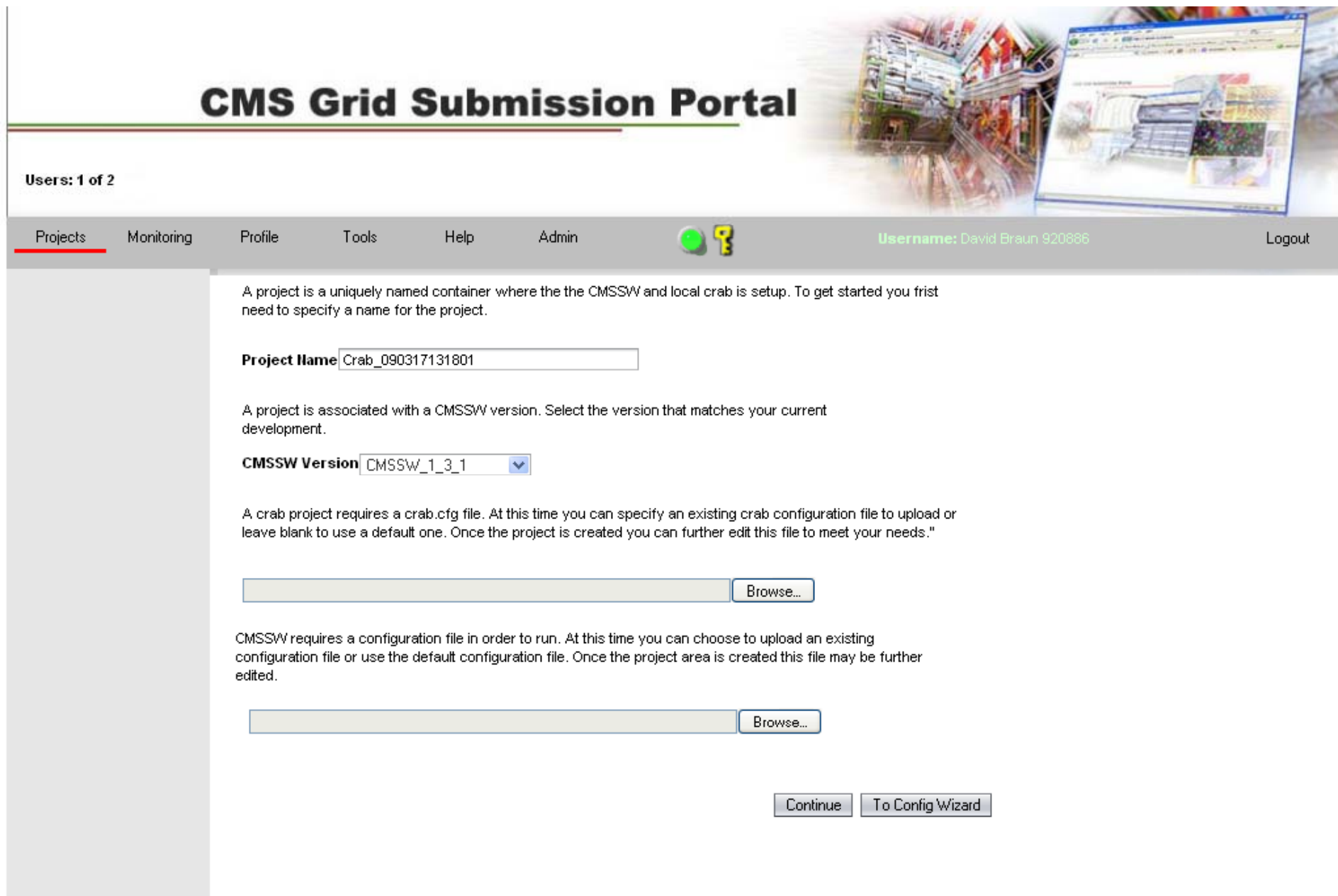
MyProjects (16)

	Monitor	Share	Name	CMSSW Version	Creation	Expiration
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090212141432	CMSSW_1_3_1	Feb 12, 2009 2:14:36 PM	May 12, 2009 2:14:36 PM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090220053528	CMSSW_1_3_1	Feb 20, 2009 5:35:35 AM	May 20, 2009 5:35:35 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090306065854	CMSSW_1_3_1	Mar 6, 2009 6:58:57 AM	Jun 6, 2009 6:58:57 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090313053311	CMSSW_2_1_7	Mar 13, 2009 5:33:25 AM	Jun 13, 2009 5:33:25 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090317095343	CMSSW_1_3_1	Mar 17, 2009 9:53:48 AM	Jun 17, 2009 9:53:48 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090317095422	CMSSW_1_3_1	Mar 17, 2009 9:54:26 AM	Jun 17, 2009 9:54:26 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090317101507	CMSSW_1_3_1	Mar 17, 2009 10:15:12 AM	Jun 17, 2009 10:15:12 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090317101524	CMSSW_1_3_1	Mar 17, 2009 10:15:28 AM	Jun 17, 2009 10:15:28 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090317101539	CMSSW_1_3_1	Mar 17, 2009 10:15:43 AM	Jun 17, 2009 10:15:43 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090317101554	CMSSW_1_3_1	Mar 17, 2009 10:15:58 AM	Jun 17, 2009 10:15:58 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090317101621	CMSSW_1_3_1	Mar 17, 2009 10:16:26 AM	Jun 17, 2009 10:16:26 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090317101633	CMSSW_1_3_1	Mar 17, 2009 10:16:37 AM	Jun 17, 2009 10:16:37 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090317101646	CMSSW_1_3_1	Mar 17, 2009 10:16:49 AM	Jun 17, 2009 10:16:49 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090317101720	CMSSW_1_3_1	Mar 17, 2009 10:17:24 AM	Jun 17, 2009 10:17:24 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090317101731	CMSSW_1_3_1	Mar 17, 2009 10:17:35 AM	Jun 17, 2009 10:17:35 AM
	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Crab_090317101745	CMSSW_1_3_1	Mar 17, 2009 10:17:48 AM	Jun 17, 2009 10:17:48 AM

Views


- New Project
- My Projects
- Shared Projects
- CrabServer Projects

Typical Workflow



CMS Grid Submission Portal


Users: 1 of 2

Projects Monitoring Profile Tools Help Admin  Username: David Braun 920886 Logout

A project is a uniquely named container where the the CMSSW and local crab is setup. To get started you frist need to specify a name for the project.

Project Name

A project is associated with a CMSSW version. Select the version that matches your current development.

CMSSW Version 

A crab project requires a crab.cfg file. At this time you can specify an existing crab configuration file to upload or leave blank to use a default one. Once the project is created you can further edit this file to meet your needs."

CMSSW requires a configuration file in order to run. At this time you can choose to upload an existing configuration file or use the default configuration file. Once the project area is created this file may be further edited.

Typical Workflow

CMS Grid Submission Portal

Users: 4 of 14

Projects Monitoring Profile Tools Help Username: haiying xu 752975 Logout

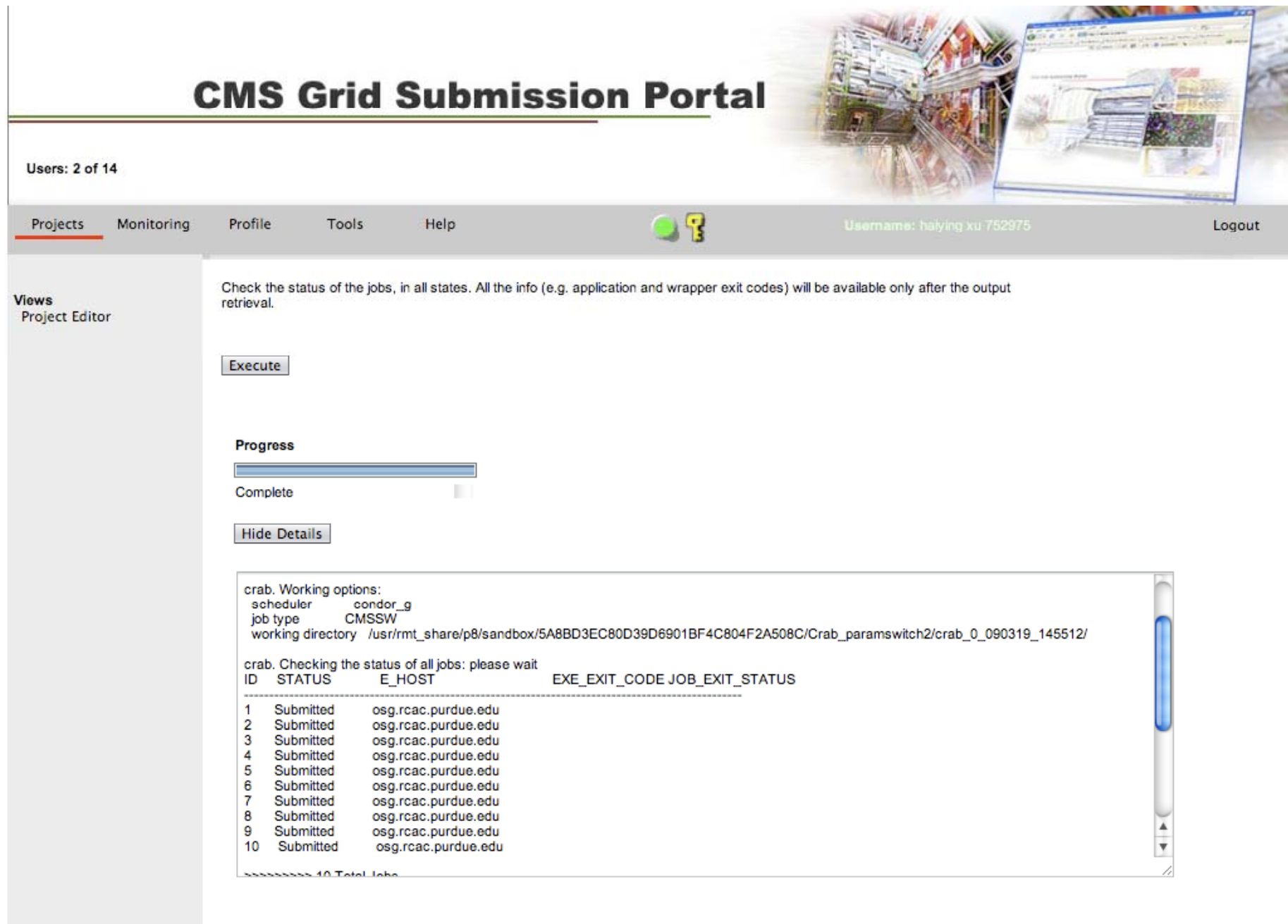
Views Project Editor

Name	Size	Directory	Modified	Readable?
Tools.mk	40.41 KiB	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
rootmlp.mk	256.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
boost.mk	407.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
roothistmatrix.mk	352.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
fastjet.mk	377.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
dbs-client.mk	151.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
sherpa.mk	1,001.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
python-ldap.mk	250.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
libungif.mk	423.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
charybdis.mk	454.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
tauola.mk	409.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
xerces-c.mk	354.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
boost_program_option...	450.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
roothtml.mk	269.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
frontier_client.mk	548.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>
doxygen.mk	133.00 b	/home/xu2/CMSSW_2_2_5/SC...	2009-03-16 10:5...	<input checked="" type="checkbox"/>

Upload 1-374/374 file(s) uploaded. Waiting for server response ... STOP

```
11:37:13.995 [DEBUG] --- fileheader end
11:37:14.061 [DEBUG] in UploadFileData.uploadFile (amount:299, getUploadLength(): 299)
11:37:14.129 [DEBUG] --- filetail start (len=195):
11:37:14.197 [DEBUG] \r\n
11:37:14.197 [DEBUG] -----lo1wgfkIgxS\r\n
11:37:14.197 [DEBUG] Content-Disposition: form-data; name="md5sum[]"\r\n
11:37:14.197 [DEBUG] Content-Transfer-Encoding: 8bit\r\n
11:37:14.197 [DEBUG] Content-Type: text/plain; UTF-8\r\n
11:37:14.197 [DEBUG] \r\n
11:37:14.197 [DEBUG] a616f2b1b15b43106b6297222dc5d0bf\r\n
11:37:14.286 [DEBUG] --- filetail end
11:37:14.417 [DEBUG] --- fileheader start (len=972):
Upload 14% done, Transfer rate: 3.51 Kb/s, ETA: 1 min. and 43 sec.
```

Typical Workflow



CMS Grid Submission Portal

Users: 2 of 14

Projects Monitoring Profile Tools Help Username: haiying xu 752975 Logout

Views
Project Editor

Check the status of the jobs, in all states. All the info (e.g. application and wrapper exit codes) will be available only after the output retrieval.

Execute

Progress

Complete

Hide Details

```
crab. Working options:
scheduler      condor_g
job type       CMSSW
working directory /usr/rmt_share/p8/sandbox/5A8BD3EC80D39D6901BF4C804F2A508C/Crab_paramswitch2/crab_0_090319_145512/

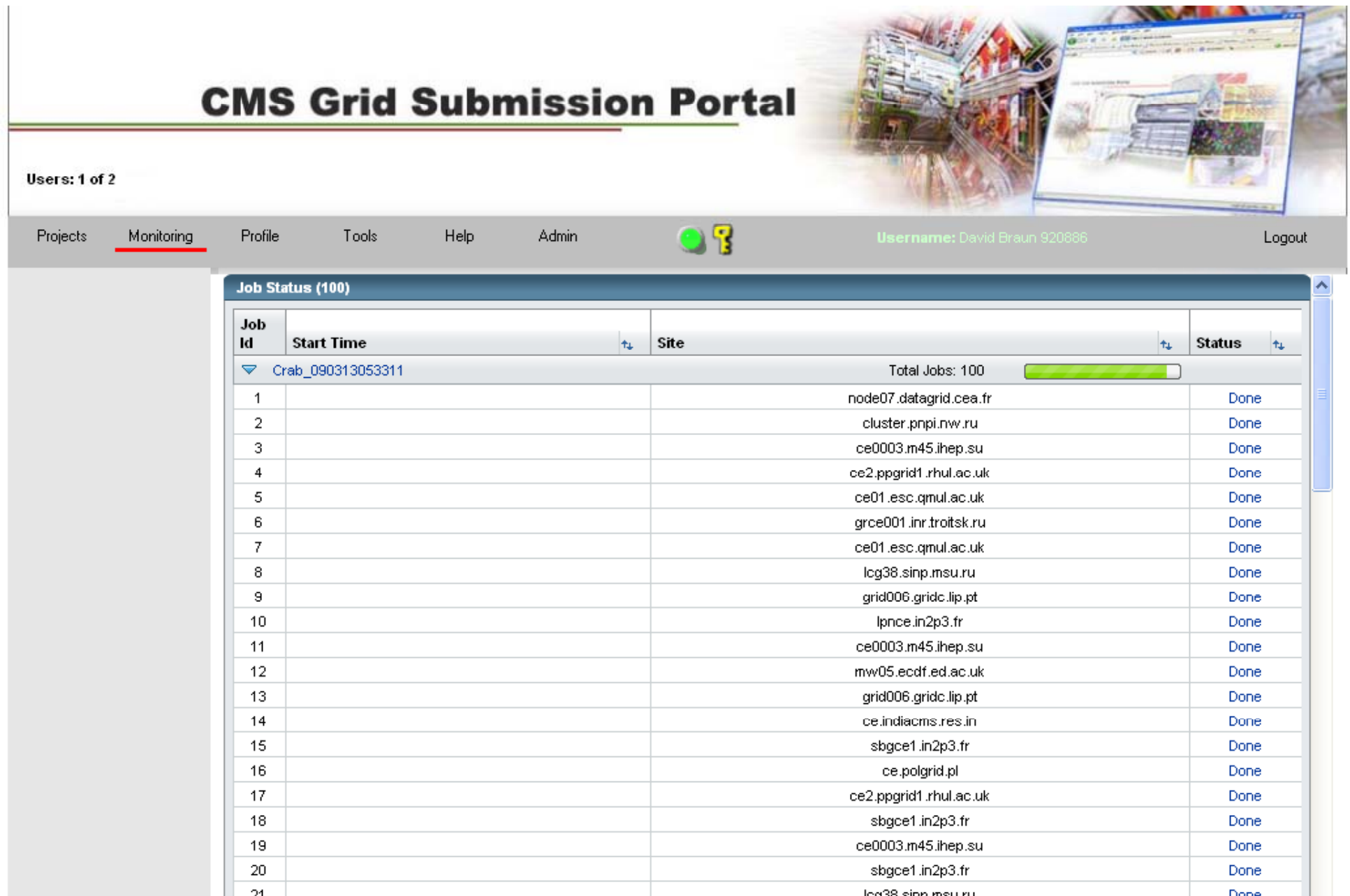
crab. Checking the status of all jobs: please wait
ID  STATUS      E_HOST      EXE_EXIT_CODE JOB_EXIT_STATUS
-----
1   Submitted   osg.rcac.purdue.edu
2   Submitted   osg.rcac.purdue.edu
3   Submitted   osg.rcac.purdue.edu
4   Submitted   osg.rcac.purdue.edu
5   Submitted   osg.rcac.purdue.edu
6   Submitted   osg.rcac.purdue.edu
7   Submitted   osg.rcac.purdue.edu
8   Submitted   osg.rcac.purdue.edu
9   Submitted   osg.rcac.purdue.edu
10  Submitted   osg.rcac.purdue.edu
..... 10 Total Jobs
```

Typical Workflow

The screenshot displays the CMS Grid Submission Portal interface. At the top, the title "CMS Grid Submission Portal" is centered. Below the title, it indicates "Users: 2 of 14". A navigation bar contains links for "Projects", "Monitoring", "Profile", "Tools", and "Help". On the right side of the navigation bar, the user's "Username: halying xu 752975" is shown, along with a "Logout" button. A sidebar on the left is labeled "Views" and contains the "Project Editor" option. The main content area shows a file explorer view of a directory named "res". The file list includes:

- CMSSW_1.stdout (3k)
- CMSSW_1.stderr (0k)
- CMSSW_2.stdout (3k)
- crab_fjr_1.xml (0k)
- CMSSW_2.stderr (0k)
- crab_fjr_2.xml (0k)
- CMSSW_3.stdout (3k)
- CMSSW_3.stderr (0k)
- crab_fjr_3.xml (0k)
- CMSSW_4.stdout (3k)
- CMSSW_4.stderr (0k)
- crab_fjr_4.xml (0k)
- CMSSW_6.stdout (3k)
- CMSSW_6.stderr (0k)
- crab_fjr_6.xml (0k)
- CMSSW_7.stdout (3k)
- CMSSW_7.stderr (0k)
- crab_fjr_7.xml (0k)
- CMSSW_8.stdout (3k)
- CMSSW_8.stderr (0k)
- crab_fjr_8.xml (0k)
- CMSSW_9.stdout (3k)
- CMSSW_9.stderr (0k)
- crab_fjr_9.xml (0k)

Typical Workflow



The screenshot displays the CMS Grid Submission Portal interface. At the top, the title "CMS Grid Submission Portal" is prominently displayed. Below the title, the user count "Users: 1 of 2" is shown. A navigation menu includes "Projects", "Monitoring" (which is underlined), "Profile", "Tools", "Help", and "Admin". On the right side of the menu, the user's name "Username: David Braun 920886" and a "Logout" link are visible. The main content area is titled "Job Status (100)" and features a table with columns for "Job Id", "Start Time", "Site", and "Status". A progress bar indicates "Total Jobs: 100" with a green bar at 100%. The table lists 21 job entries, all of which are in a "Done" status. The sites listed include various international locations such as node07.datagrid.cea.fr, cluster.pnpi.nw.ru, ce0003.m45.ihep.su, ce2.ppgrid1.rhul.ac.uk, ce01.esc.qmul.ac.uk, grce001.inr.troitsk.ru, lcg38.sinp.msu.ru, grid006.gridc.lip.pt, lpnce.in2p3.fr, mw05.ecdf.ed.ac.uk, ce.indiacms.res.in, sbgce1.in2p3.fr, ce.polgrid.pl, and lcg38.sinp.msu.ru.

Job Id	Start Time	Site	Status
Crab_090313053311		Total Jobs: 100	
1		node07.datagrid.cea.fr	Done
2		cluster.pnpi.nw.ru	Done
3		ce0003.m45.ihep.su	Done
4		ce2.ppgrid1.rhul.ac.uk	Done
5		ce01.esc.qmul.ac.uk	Done
6		grce001.inr.troitsk.ru	Done
7		ce01.esc.qmul.ac.uk	Done
8		lcg38.sinp.msu.ru	Done
9		grid006.gridc.lip.pt	Done
10		lpnce.in2p3.fr	Done
11		ce0003.m45.ihep.su	Done
12		mw05.ecdf.ed.ac.uk	Done
13		grid006.gridc.lip.pt	Done
14		ce.indiacms.res.in	Done
15		sbgce1.in2p3.fr	Done
16		ce.polgrid.pl	Done
17		ce2.ppgrid1.rhul.ac.uk	Done
18		sbgce1.in2p3.fr	Done
19		ce0003.m45.ihep.su	Done
20		sbgce1.in2p3.fr	Done
21		lcg38.sinp.msu.ru	Done

Integration with CMS Tools

- Frame includes or open window Java script calls
 - **DBS**
- File system for local crab server
 - **CRABSERVER**
- Crab through process calls
 - **SRAM, CRAB**
- Investigated but not implemented
 - **JSON RPC**
 - requires an additional trust manager
 - **Web Service**
 - XML payload needed parser
 - Globus proxy required
 - **HTML hijacking**
 - Complicated but do able

Future Developments

- Most of the key functionality implemented
- JMS bridge to CRAB server – provide async job status
- Expose/consume JSON RPC services
- Google Gadget Job Status
- Integrate SRM Web based file browser
- CRAB configuration wizard
 - Direct interface into DBS/DLS
- Statistics, user jobs
- Service interfaces to provide additional user interface

Portability

- JBoss 4.2.3-GA
- Jdk 1.6
- Two Enterprise Application Archives
 - `user.ear`
 - `crab.ear`
- Default Database or MySQL
- Crab and CMSSW installations
- OSG and Glite stacks
- Bash
- Reuse
 - Security
 - VOMS Tools
 - Basic tools
 - File upload
 - Browsing
 - Process control
 - File templates
 - General structure so long as work flow matches.

Outlook

- Prototype portal for CMS users to tryout
- Gain insight into how to reduce the complexity of the users experience
- Evaluation of industry standard technologies to produce a reusable and flexible deployment platform
- Give it a try: <http://crab.rcac.purdue.edu:8080/Crab>