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

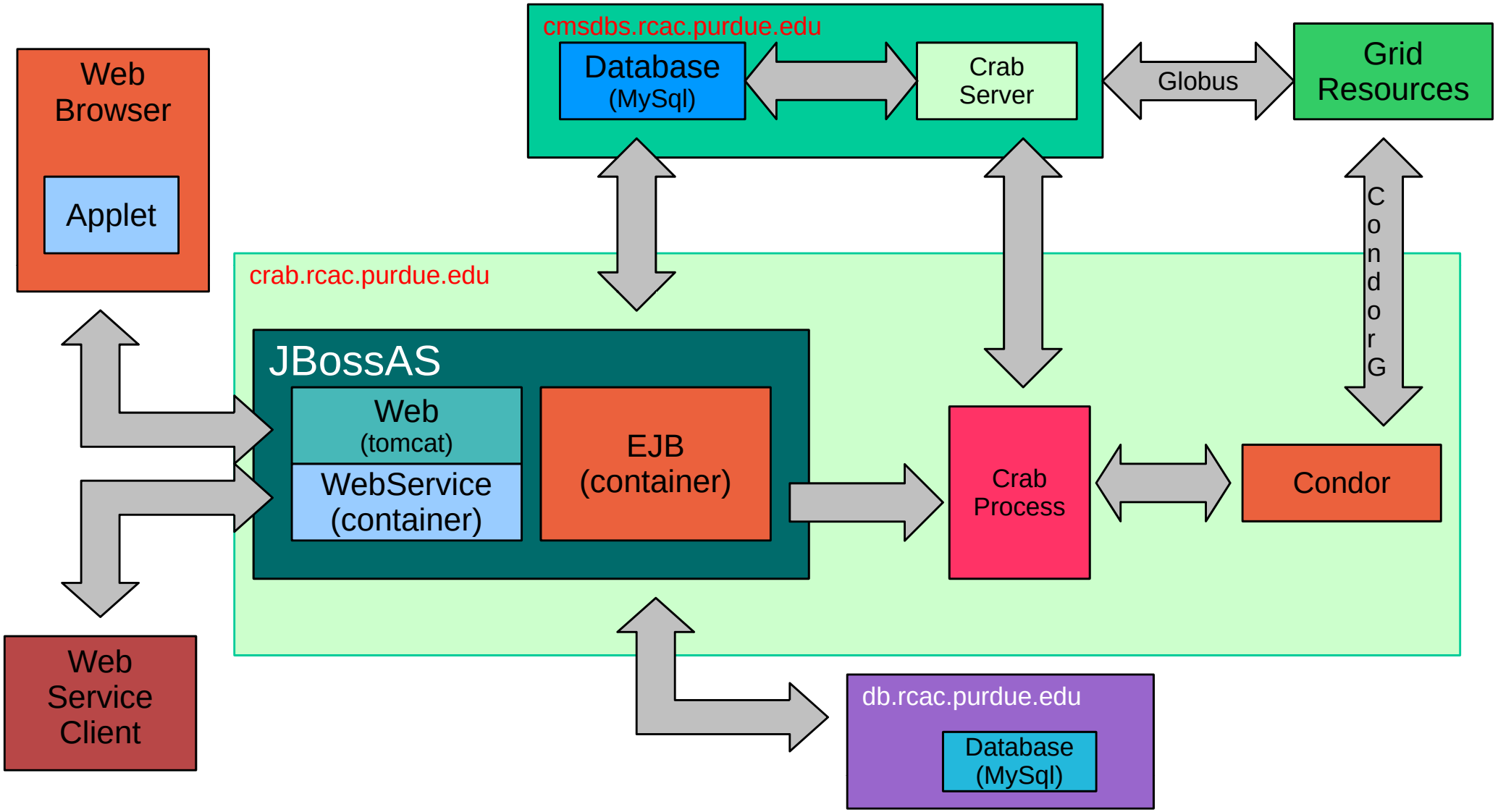
Why Are Portals Important?

- 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?**

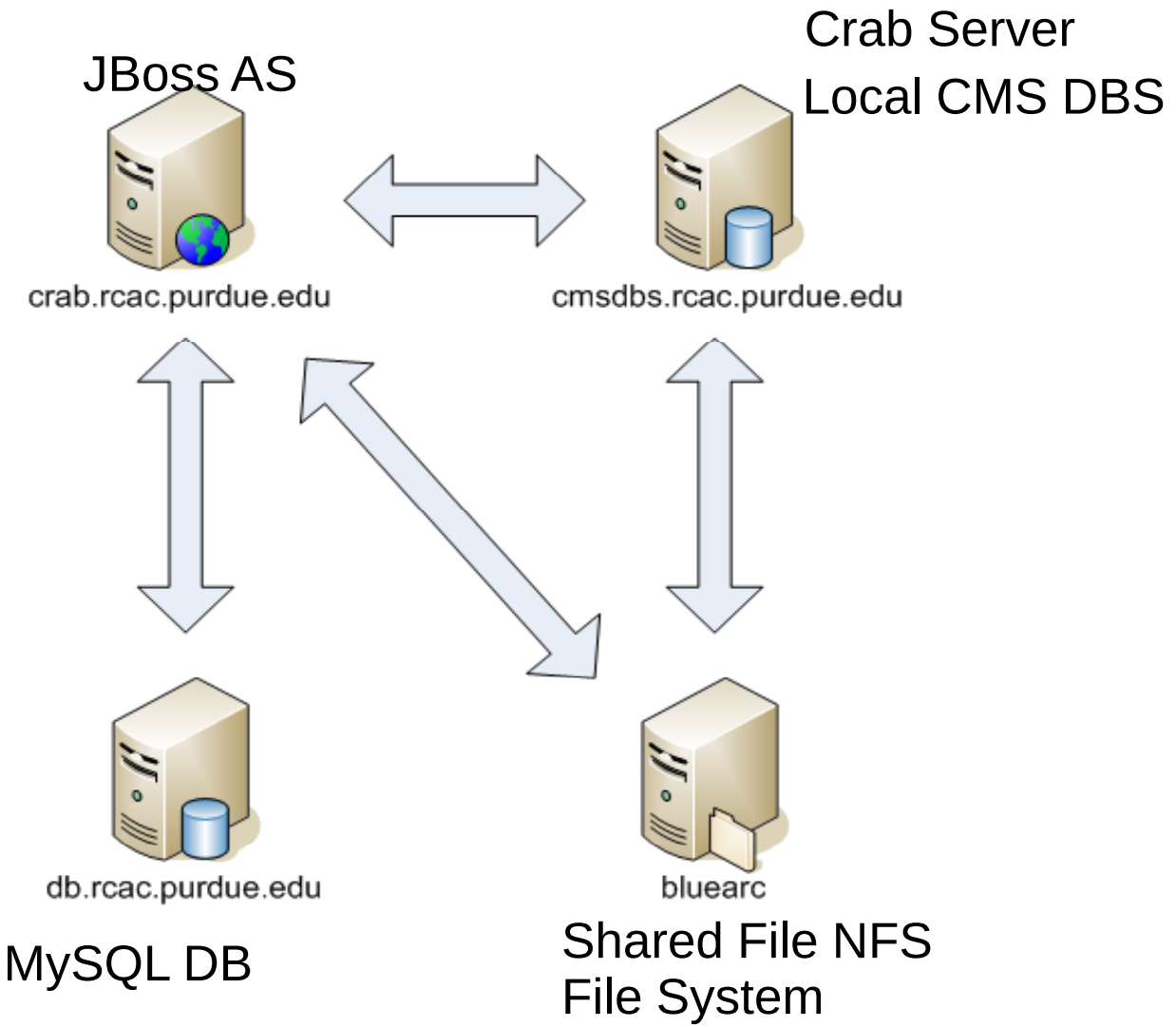
Design Goals

- 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



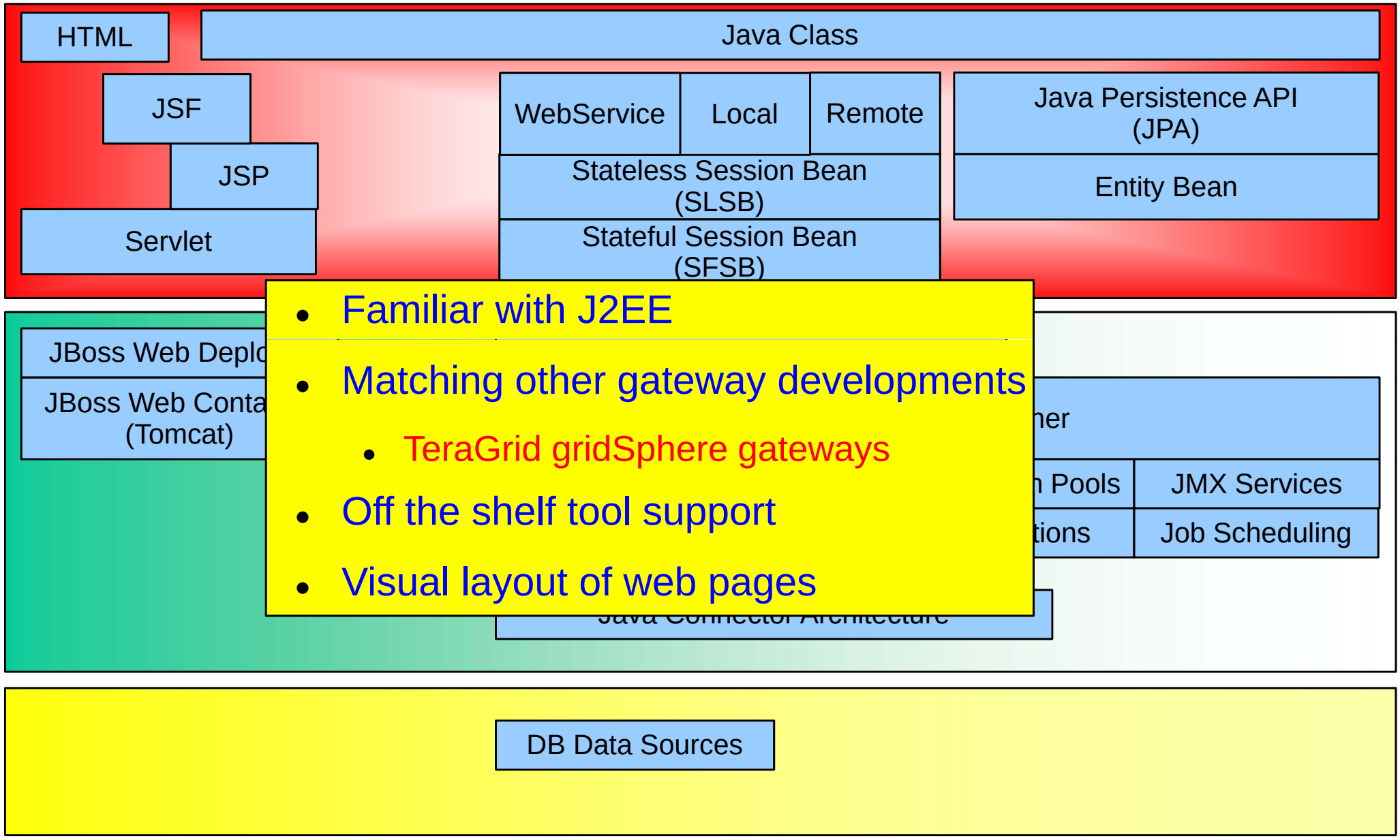
System Diagram



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

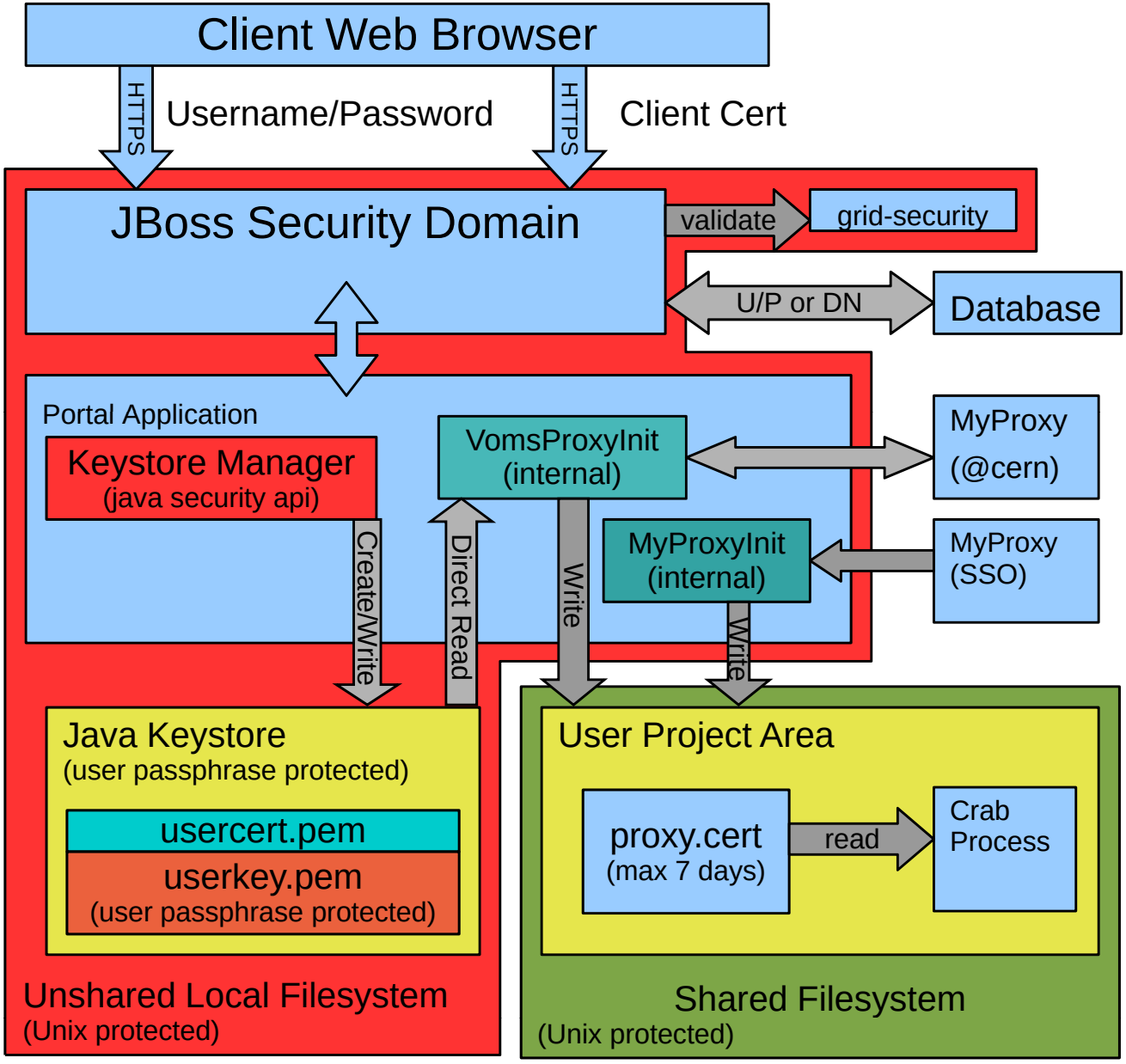


- Familiar with J2EE
- Matching other gateway developments
  - TeraGrid gridSphere gateways
- Off the shelf tool support
- Visual layout of web pages

Technology List

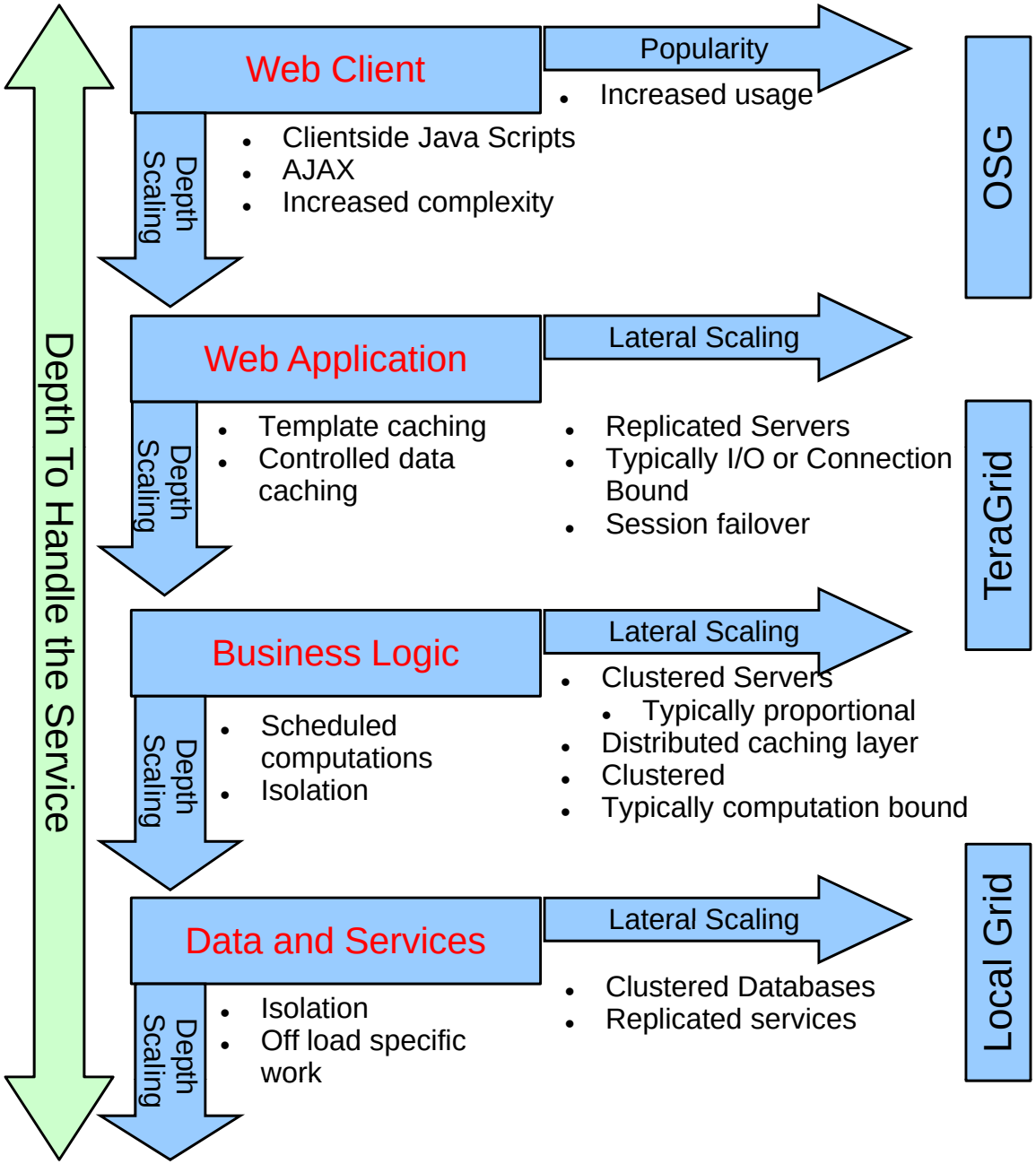
- Java SDK1.6
- Java Server Faces (JSF) JSR-314
  - Woodstock Components Implementation
  - Facelets, Dojo, DWR, Json
- Java Server Pages (JSP)
  - JSP Standard Tag Library (JSTL)
- Java Architecture for XML Binding (JAXB 2.1)
  - XML object serialization
- Java Management Extensions (JMX)
  - Service 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 excelent science that the CMS project will provide.

Sincerely, CMS Tier-2 Facility

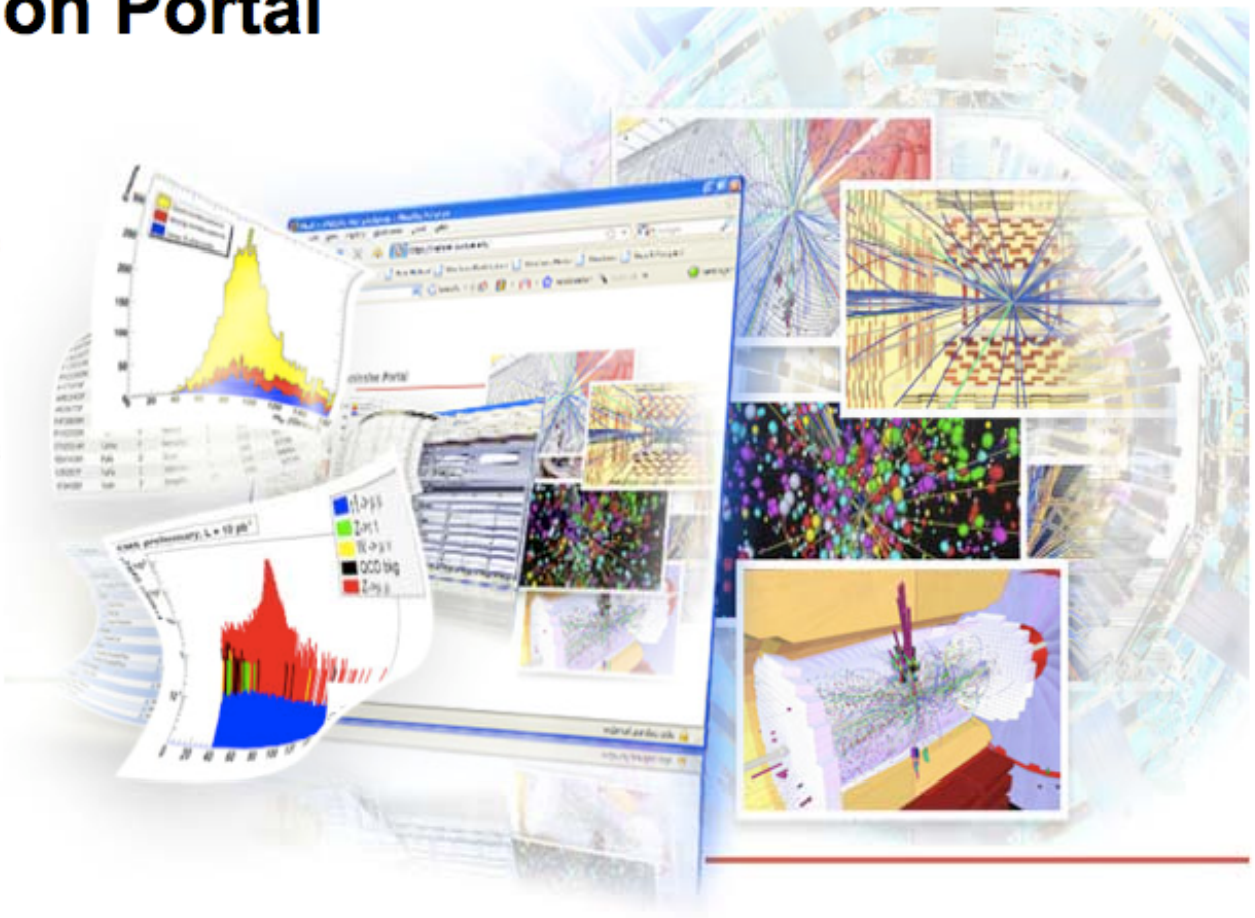
CMS General | OSG News | RCAC News

iCMS interruption 2009-03-19
by Gilles Raymond - Mar 18, 2009
Dear all, due to this ...

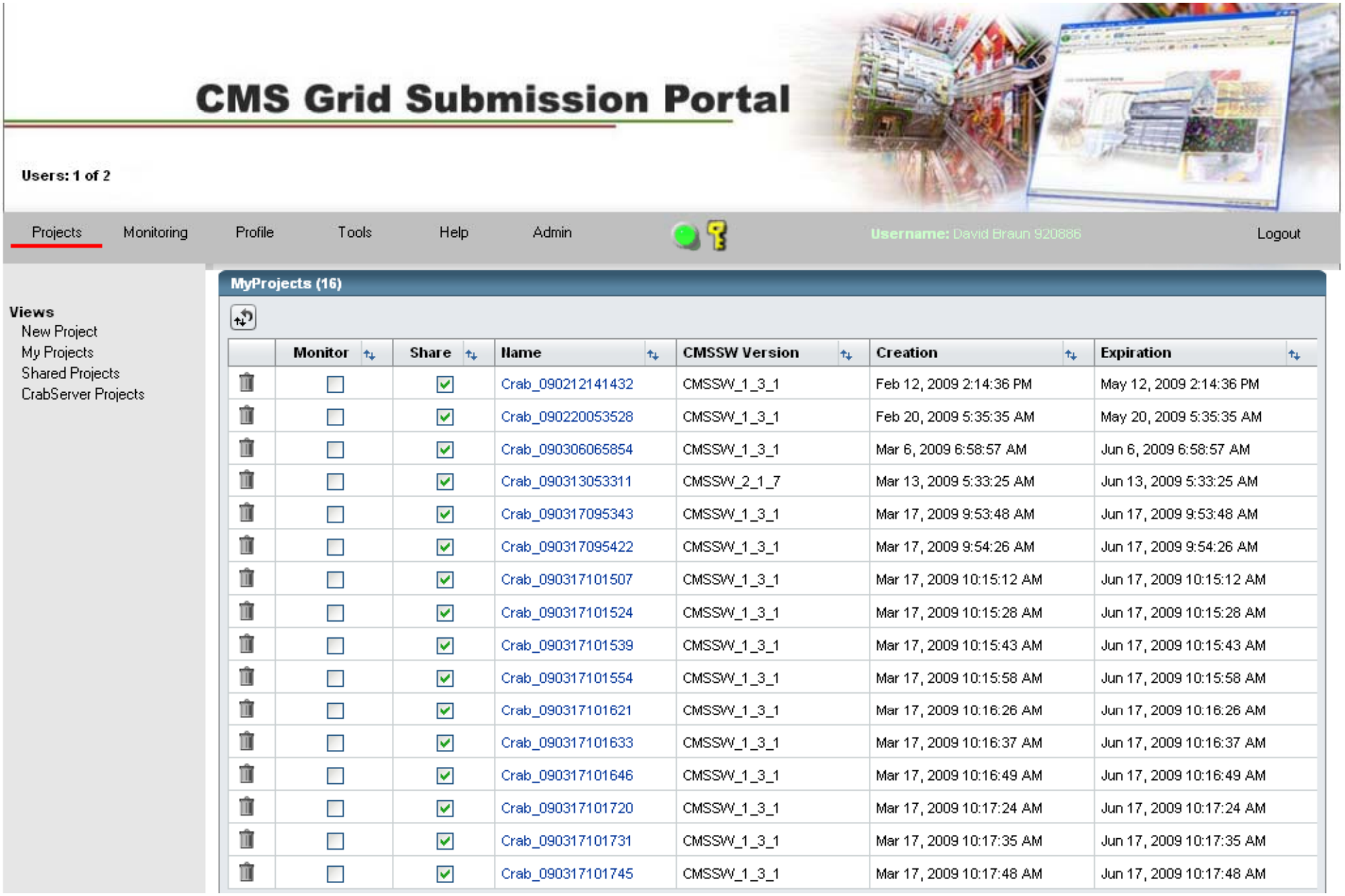
iCMS reboot
by Gilles Raymond - Jan 23, 2009
Dear all, IT requested a reboot (after a kernel upgrade) on the iCMS web servers. The web servers will therefore be rebooted ...

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 authentification certificate no ...

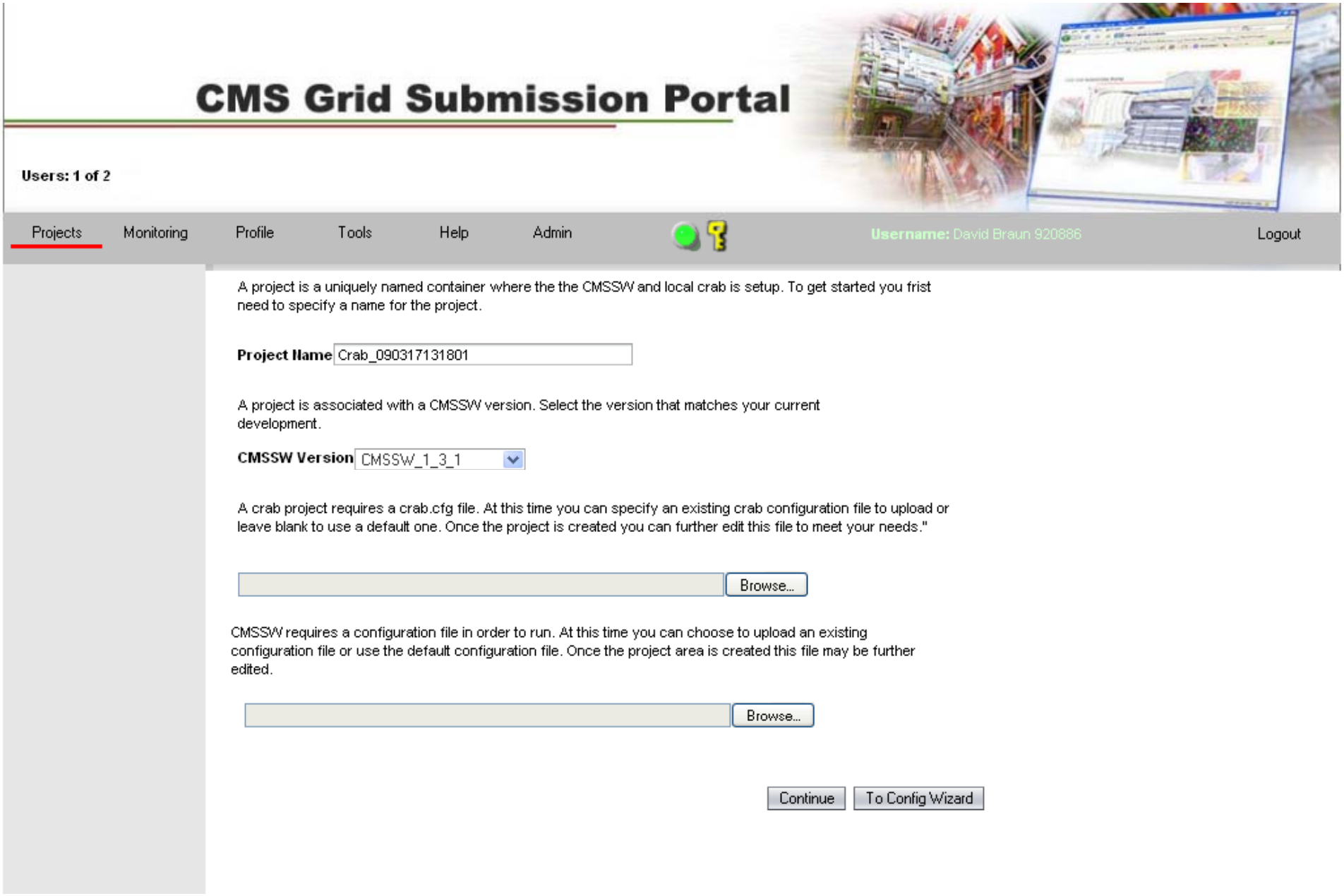
Annual Power cut 4/01/2009
by Gilles Raymond - Dec 19, 2008
Dear all, In accordance to the general power cut on Sunday 4th 2009 from 7.00 AM to 9.00 ...



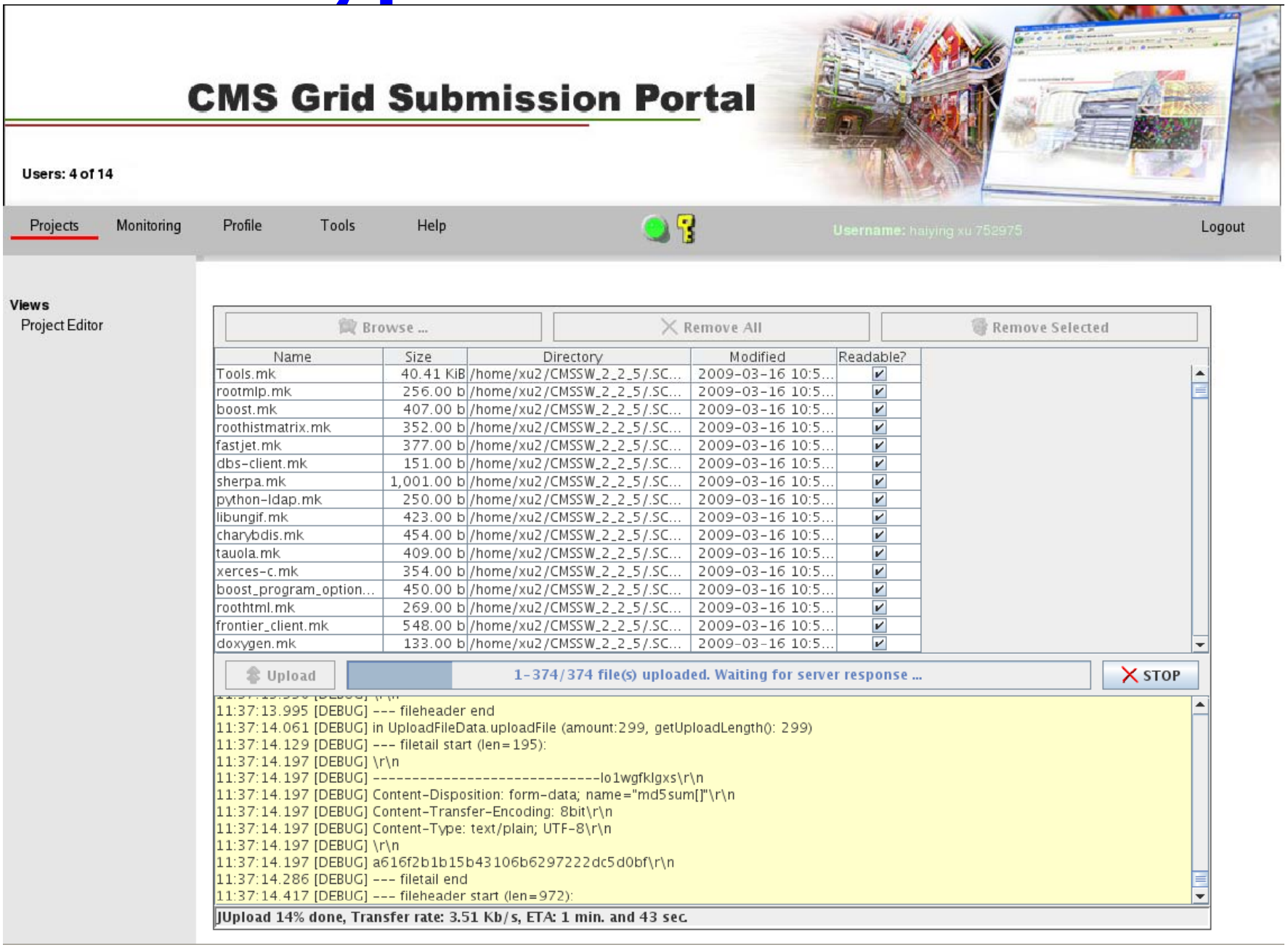
Typical Workflow



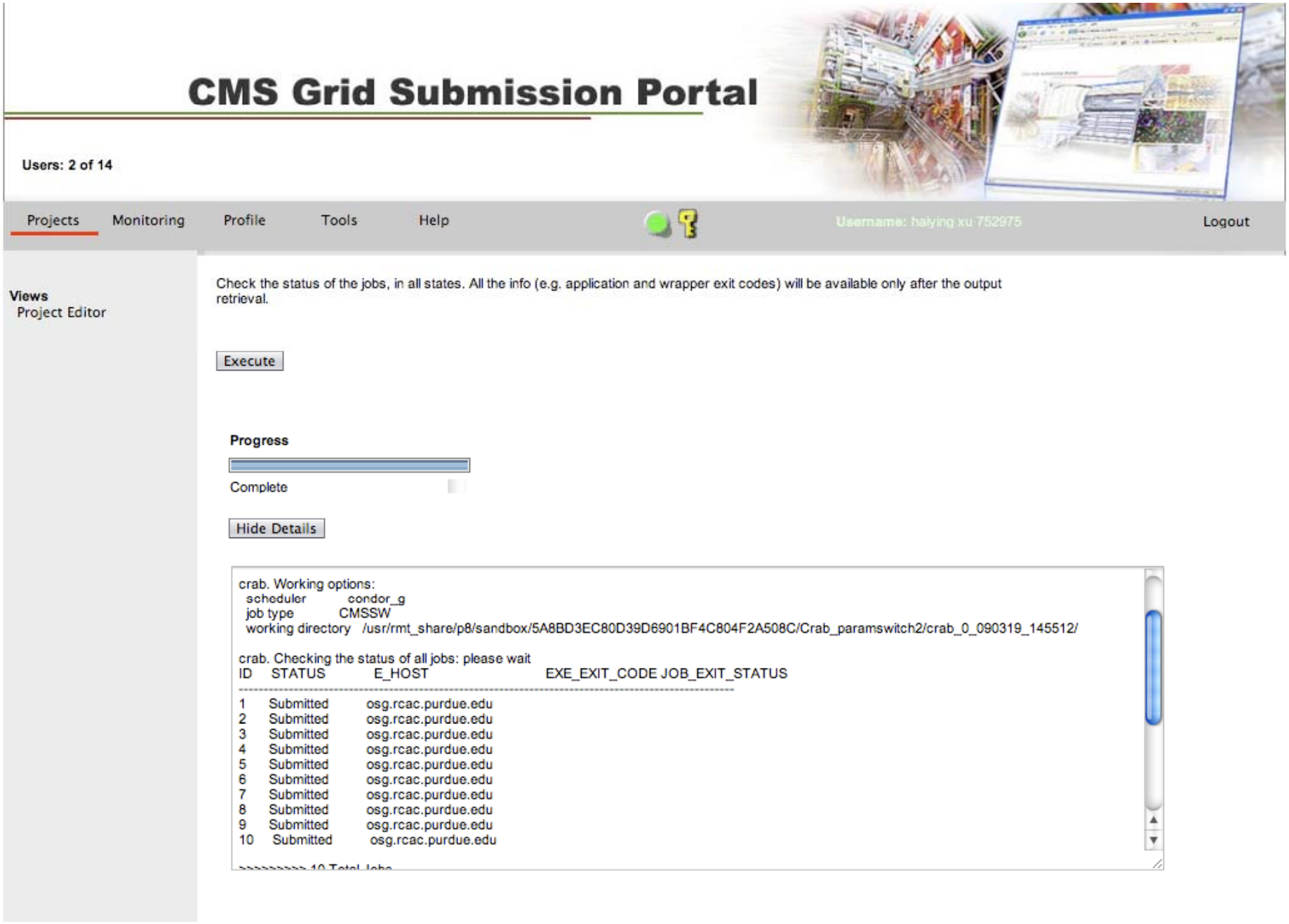
Typical Workflow



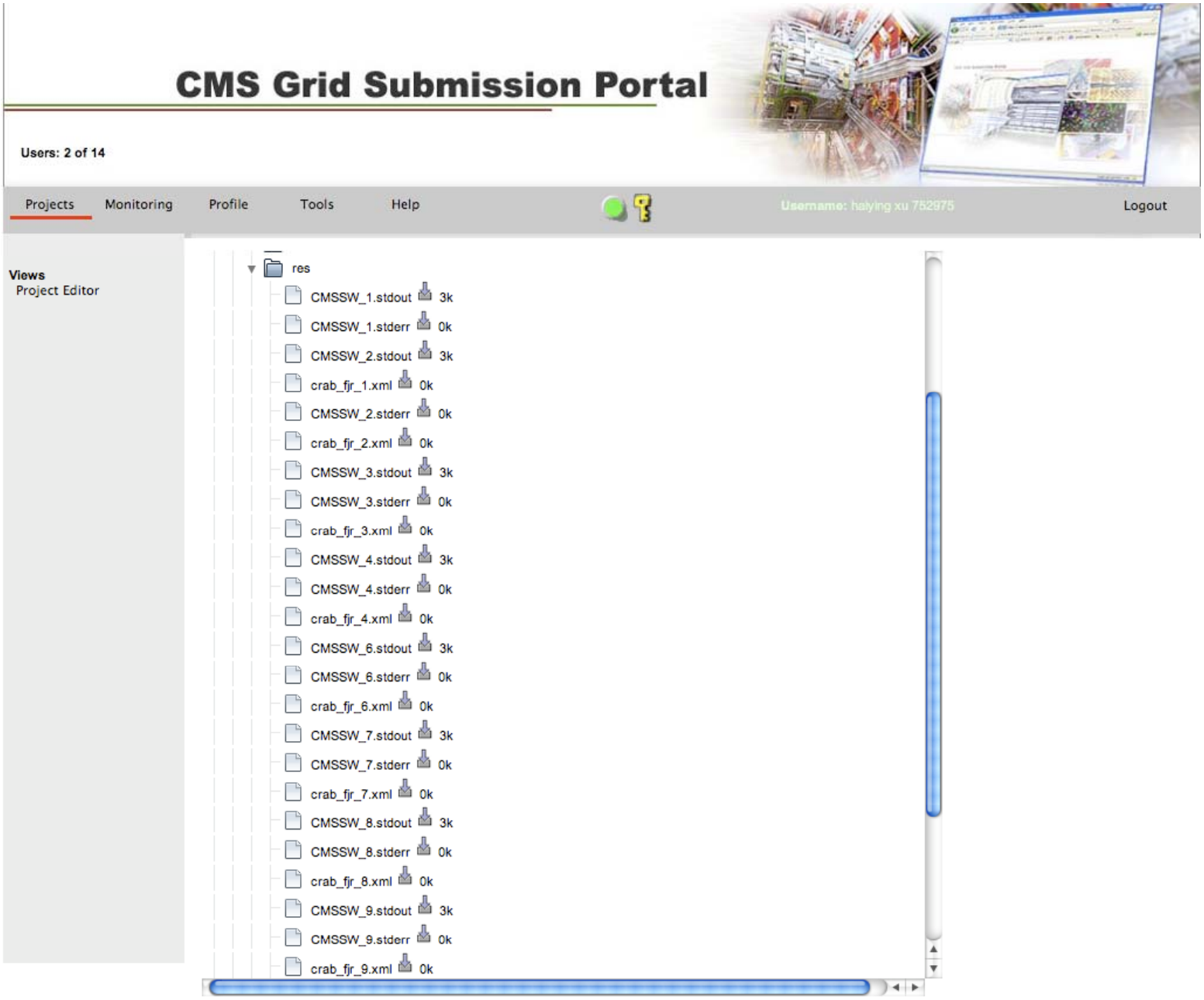
Typical Workflow



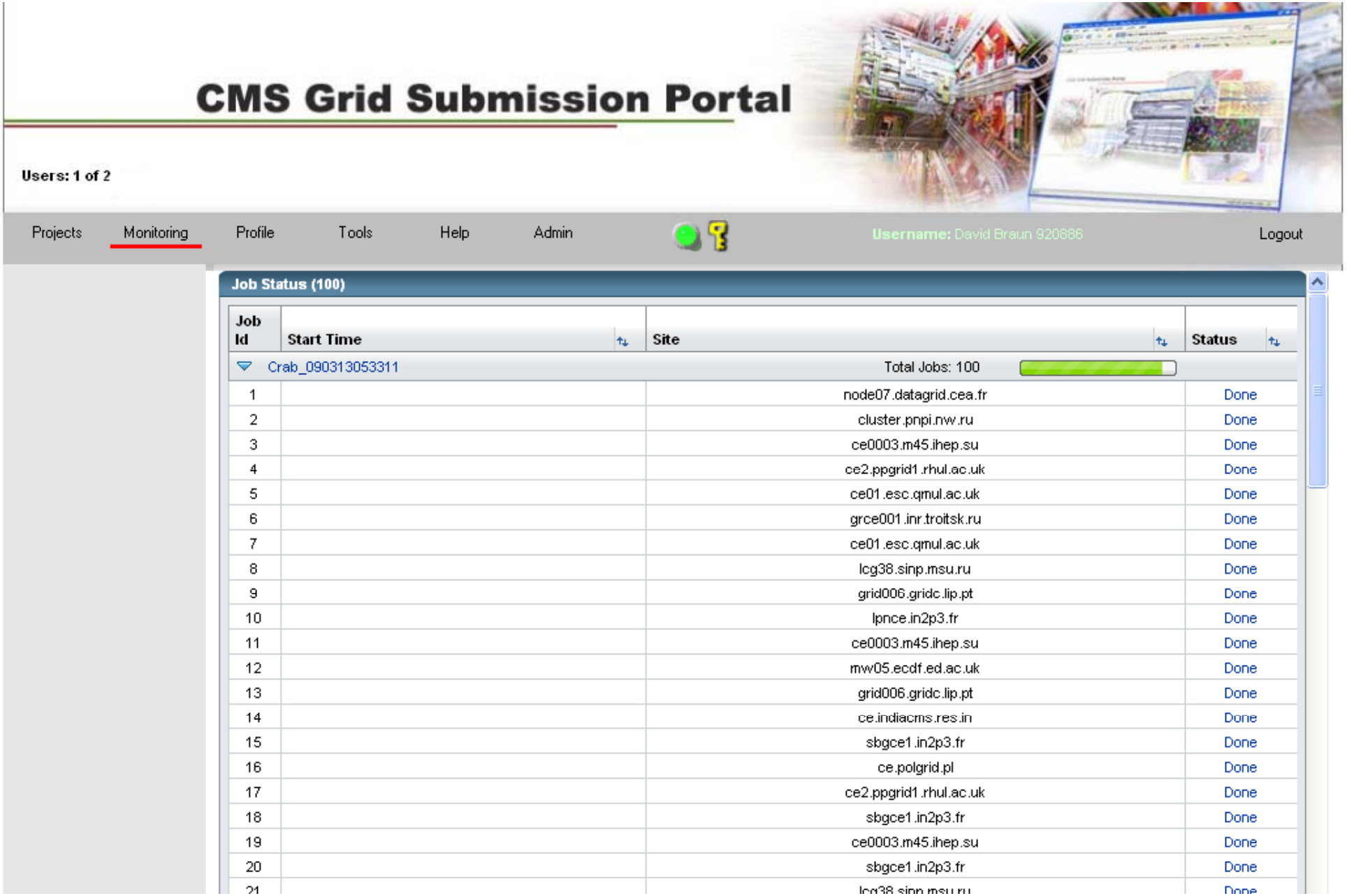
Typical Workflow



Typical Workflow



Typical Workflow



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