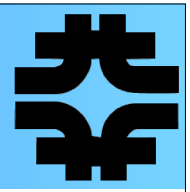


User Facility Projects

June 22, 2004

User Facility Project

US-CMS User Facility (UF) Project one of two level two tasks of the US-CMS Software and Computing Project (the other being Core Application Software, CAS)

- User Facility Project is responsible for building computing centers in the US
  - US-CMS Tier-1 at FNAL
  - Approximately 5 Tier-2 centers at US universities
- Also responsible for building computing infrastructure to facilitate analysis, production of simulated data, and eventually reconstruction of real data
  - Developing local services
    - User Analysis Facility at FNAL
  - Developing the distributed processing environment (DPE) user to produce simulated events in the US
  - Developing distributed analysis prototypes and higher level services

User Facility Projects

Today we will give a high level overview of User Facility Projects

- Interoperability with LCG
- VO Management
- Privilege and Authorization work
- File Cataloging
- US-CMS Distributed Production
- CMS Dataset preparations and discovery
- Monte Carlo Generation Services
- RunJob

Today is an introduction we can go into more depth at future meetings

LCG Interoperability

UF is responsible for ensuring that US-CMS computing resources are used to enable physics analysis in the US and meet our obligations to international CMS efficiently

- We are working on enabling the LHC Computing Grid to submit requests for processing and data to US-CMS grid enabled resources
  - Tier-1 and Tier-2 centers
- We are working to allow US developed grid infrastructure like Grid3 and eventually OSG to use resources deployed with LCG interfaces

Interoperating and Federating grid resources can be approached in a number of ways

- Essentially one trades efficiency for development effort
- CMS is working on a number of short and medium term solutions

LCG Interoperability

There are several methods available to for the Tier-1 facility to interoperate with multiple grid installations

- Partition the cluster between the projects
  - Currently used at FNAL due to time pressure, but an inflexible and inefficient method
- Deploy multiple sets of interfaces to the same physical resources
  - Requires decomposing the interfaces to determine how to integrate into the existing structure most efficiently.
    - Working on enabling this for processing through Condor-G
    - Often requires development on the interfaces to make them compliant with the site security policies and configuration techniques.
- Define and deploy compatible interfaces between multiple grid projects
  - A flexible and efficient method from the facility standpoint, but requires cooperation and coordination between grid installations

US-CMS is working on the final two methods. The middle is seen as a viable short term solution. The final is seen as the most desirable.

VO Management

US-CMS successfully finished the VOX project which developed the VOMRS (Virtual Organization Management Registration Service)

- Developed a database to store VO information and could be used to populate VOMS servers with user and group information
- Developed authorization module to check the status of an incoming user against registered FNAL users
- Developed workflow to enable a transparent registration process for users
  - Register with CMS VO
  - Automatically work through FNAL registration process

Working with LCG

- CERN registration issues have all the same issues as FNAL

VO Management

US-CMS planned to implement the VOMRS to control the membership of Grid2003

- During the project itself the number of users was small enough to manage manually.
- We wanted to enable the concept of groups and roles
  - For official CMS production there are only production coordinators, but as one adds unofficial production and analysis it becomes important to specify role priority

Also becoming important as we try to allow LCG interoperability.

- LCG has a VO, if we want to enable those users to utilize resources in the US we need to get them registered as FNAL users in a transparent way
  - We are working to enable the workflows proposed

Privilege and Authorization Project

The goal of the Privilege and Authorization Project is to enable finer grained authorization to grid enabled services

- Areas of authentication management have had longer to develop
  - Verification of user identity using a grid certificate and a trusted certificate authority
- How to establish and enforce policy for what that user can do and how many resources can be utilized is not well implemented
  - In Grid2003 users were mapped to group accounts
    - Mapping was many to one and static
  - Access was controlled on a very course level using the UNIX access writes of the group account the user is mapped to

We would like to enable the site to define and enforce finer grained auth.

- Enfore quotas, provide access to specific services based on activity

Allow users to define their roles and activities dynamically

- Obtain different authorization for different activities

Currently Activities

Adapting prototype authorization module to work with current gatekeeper

- Working on enabling authorization module to work with storage gatekeeper
  - Also investigating other techniques for storage plugin

Adapting prototype authorization module to parse extended proxies

- Enabling VOMRS infrastructure for end to end testing
  - Register Users
  - Assign group memberships
  - Enable users to create extended proxies that indicate membership and role
  - Allow processing and storage gatekeepers to change authorization and policy based on extended proxy

Developing prototype policy decision point

- Policy communicated through developing standard XACML
- Working with our BNL colleagues to have GUMS interface with this

Cataloging and data Replica Research

During DC04 CMS had experience with a single instance of the EDG RLS which was used as both a replica service and a file meta data catalog.

- About 500k entries were stored in the catalog
  - Multiple physical file names
- A number of performance issues were noted

US-CMS is looking at the performance of the Globus RLS under the same conditions

- Populating the Globus implementation of RLS and trying the same variety of queries

Also examining the new cataloging services from the EGEE/ARDA distributed analysis prototype middleware (g-lite)

CMS Event Production

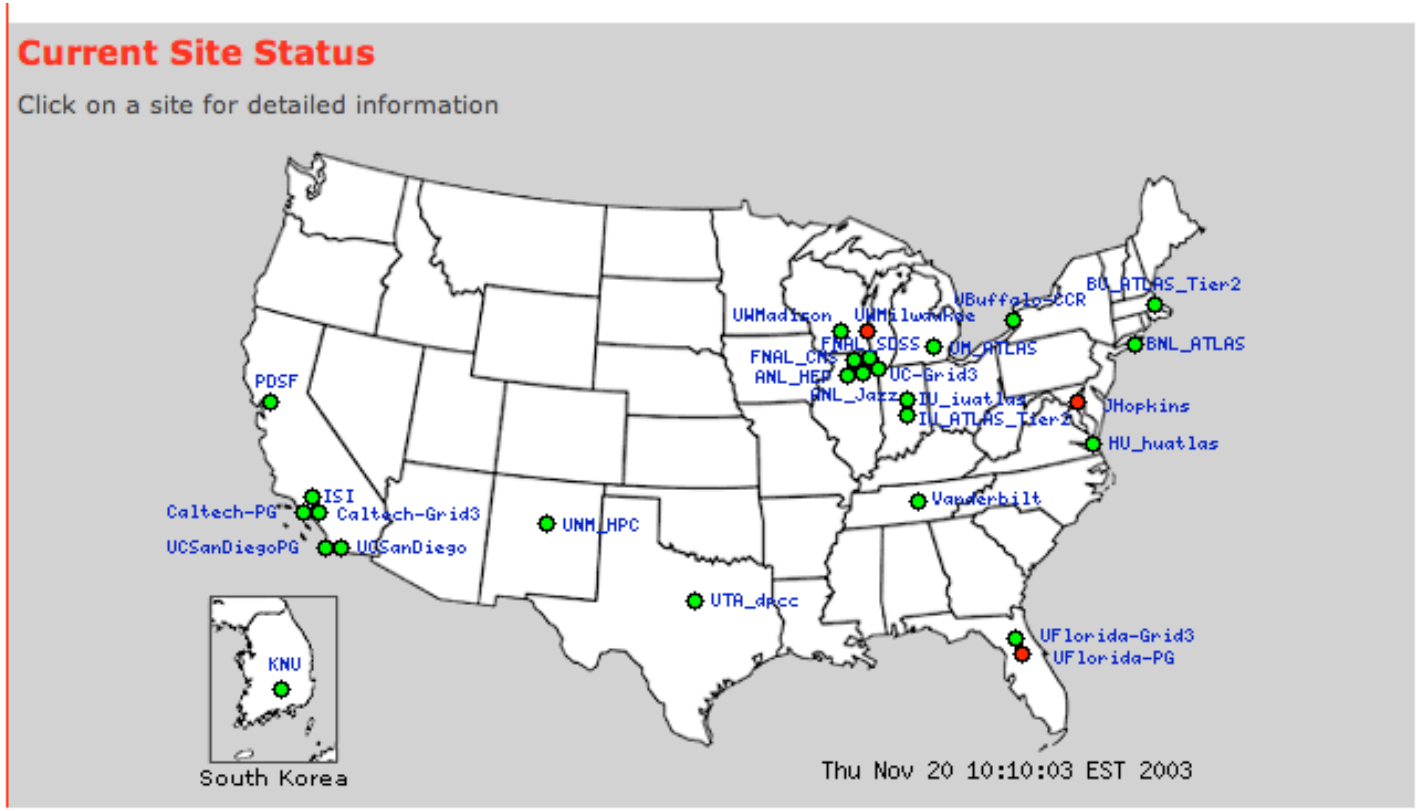
CMS is in a continuous running mode for production

- The USCMS effort to created the Distributed Production Environment for stable persistent running of production from late 2002.
  - VDT Based
  - Plus CMS Production tools and runtime environment
  - Compatible grid prototype components
- Rolling Prototypes deployed across three test environments
  - DGT Development Grid Testbed
    - R&D Efforts, Evaluation of new components, shake-out testing
  - IGT Integration Grid Testbed
    - Scalability and stability testing and development
  - PG Production Grid
    - Stable production running
- Infrastructure allowed us to make use of Grid2003 resources from US-ATLAS and other VOs

CMS Event Production

US-CMS continues to use both dedicated and opportunistic computing resources, though at a lower level than before the challenge

- Production is expected to ramp up over the summer to a level of 10M events per month worldwide



Recent Plots From Production