CMS Connect

Planning

Goal

Explore creating a distributed job service for USCMS, similar to ATLAS Connect and OSG Connect

ATLAS CONNECT

http://connect.usatlas.org/

osg connect

http://osgconnect.net/

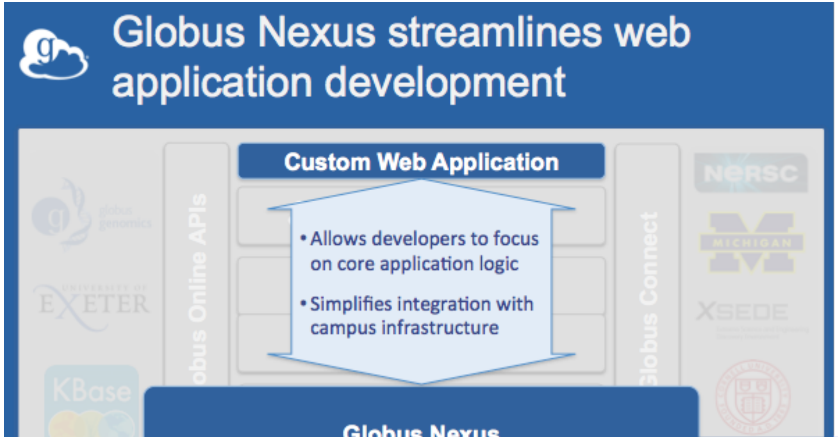
- “Virtual cluster” to augment local resources
- Ease burden of site admins (resource targets)

Service Elements

- Identity & group management service
- Auto account provisioning to login host
- Flocking service (BOSCO-based)
- Data service (for quasi-transient job data)
- Job monitoring service
- Accounting service
- Web page for navigation
- Service documentation page

Identity and Group Management

- Leverage the Nexus service provided by Globus
- Use campus identity
- Use CI-Logon, InCommon federation
- Users authorized in standard way (acceptable use agreement, personal verification) but no X509 or VOMS needed



Globus Nexus streamlines web application development

Custom Web Application

- Allows developers to focus on core application logic
- Simplifies integration with campus infrastructure

Globus Nexus
(Identity, group, & profile management)

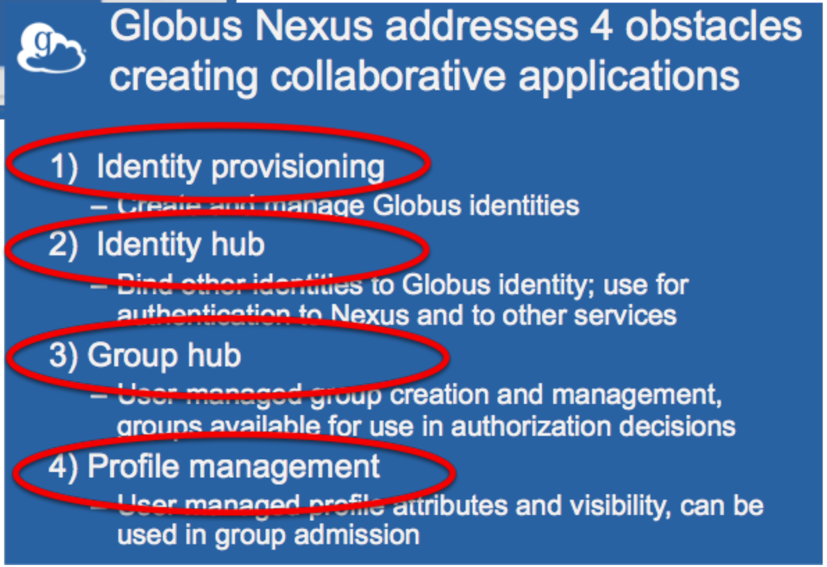
Globus Nexus makes it easy for individuals, teams, and institutions to create web applications for the science community

It provides a flexible, powerful Platform-as-a-Service to which developers can outsource their identity, group, and profile management needs.

CMS Connect would leverage all these components.

CI-Logon plus home university ID provider used to login.

CMS Operations staff control authorization.



Globus Nexus addresses 4 obstacles creating collaborative applications

1) Identity provisioning
   - Create and manage Globus identities
2) Identity hub
   - Bind other identities to Globus identity; use for authentication to Nexus and to other services
3) Group hub
   - User managed group creation and management, groups available for use in authorization decisions
4) Profile management
   - User managed profile attributes and visibility, can be used in group admission

Auto provisioning of user accounts

- A cron runs on the login hosts and checks for newly authorized members in the CMS Connect group.
- Creates Unix account, and users can quickly login and start work (no admin needed) (~10 mins)
- Has `/cvmfs/cms.cern.ch/`, OASIS mounted plus other CMS compatibility libs as needed.

Flocking service

- `cms-connect` user account created on target clusters (runs Glideins)
- Only ssh-based protocol to reach clusters
- Resource targets can advertise attributes back to the schedd as usual
- Usual ClassAds and Condor submit scripts

Resource targets

- Any cluster running Condor, PBS, SLURM, or SGE
- User account created
- Best if local squid service is available
- Can also submit to any GlideinWMS VO front-end (also OSG VO front-end for opportunistic cycles)

CVMFS

There are a number of options is the target cluster does not have CVMFS. We are exploring these in ATLAS + CCTools group at Notre Dame

Storage service

- Provide a quasi-transient storage service for job input/output data
- Implemented with Ceph object store
- POSIX access provided to login host
- Globus Server for managed xfers
- Xrootd, and http service endpoints
- Would provision ~ 20 TB

Other services

- Web skin for navigation to services
- Cycle Server for job monitoring
- Gratia for accounting
- Confluence wiki for service documentation (CMS user documentation found elsewhere)

How to do it

- Meeting to discuss services; capabilities and limitations
- Identify one or more resource targets, setup technical phone call
- Meeting to discuss web branding, information, domain name, email support lists, etc.

Possible timeline

1. Meeting next week or soon thereafter
2. Two-three weeks for UC team to deploy services, including Globus-based backend
3. Week of testing with first couple of resource targets. New targets can be added usually quickly, with a brief technical phone call
4. First users in August