Multi-VO Support in IHEP’s Distributed Computing Environment

YAN, Tian
On behalf of distributed computing group
Institute of High Energy Physics (IHEP), CAS, China
CHEP-2015, Apr. 13-17th, OIST, Okinawa
Outline

- Distributed computing for BESIII
- Other experiments wish to use DIRAC
- Multi-VO support
  - architecture refinement
  - configurations in DIRAC
  - configurations in StoRM SE
  - frontend system
  - monitoring and accounting system
Distributed Computing for BESIII

- Start running at 2012
- Based on DIRAC
- Figures:
  - 10 sites
  - ~ 3000 CPU cores
  - ~ 400 TB storage
  - ~ 50k jobs/month
Architecture of BES-DIRAC

- DIRAC as middleware
- ganga as frontend, with extension for BESIII
- dCache with 126TB disk array as central SE
- A high level data transfer system for data transferring between SEs
- VMDIRAC plugin for cloud sites
- Use CERN’s cvmfs server for software deployment
They Wish to Use DIRAC

Circular Electron Positron Collider (CEPC)

Large High Altitude Air Shower Observatory (LHAASO)

Jiangmen Underground Neutrino Observatory (JUNO)

Hard X-ray Modulation Telescope (HXMT)
Making DIRAC as a Service

- One DIRAC setup for one VO is expensive:
  - need dedicated hardware
  - need expert manpower to maintain
  - small VOs are not willing to afford that

- Universities in China joined several experiments above
  - one site belongs to several VOs
  - they may have quota and priority policies
  - a single DIRAC setup will be easy to manage these resources

- BES-DIRAC is already running and can be extended without too much effort

- Easily to support new experiments in future
Restructure for Multi-VO Support

- multiple DIRAC servers for load balances
- dedicated DB server for DFC, accounting and monitoring
- lightweight frontend JSUB
- local CVMFS server for software deployment
- StoRM+Lustre for integration of grid and local data
Multi-VO Registry in DIRAC

- **VO and VOMS settings**
  - disable the global VO
  - add VO items in the configuration
  - set VOMS server and URL for each VO

- **User Group settings**
  - add new user group for each VO
  - add a generic pilot group for each VO
  - specify VO in user group and generic pilot group
VO Permissions Control in Site Configurations

- **Cloud site**
  - cloud site has different setting scheme
  - we plan to add direct VO control for cloud sites

- **Temporary solutions for cloud site**
  - specify the groups belong to the VO in requirements
  - only cloud from that group can launch VMs and accept jobs

- **non-Cloud site**
  - VO can be set for each site in /CE/queue
  - multi-VO in one site can be specified
Workload Management

- **SiteDirector agents**
  - one agent for one VO, with it’s own configuration file
  - `/opt/dirac/pro/etc/WorkloadManagement_SiteDirector<VO>.cfg`
  - can be distributed in slaver DIRAC server for load balance

- **Multi-VO’s job control**
  - SiteDirector will send generic pilot to job queue under the same VO as the job
  - The generic pilot of the VO is set in the configuration
  - Generic pilot can pull all jobs in the same VO

```python
1 Systems
2 {
3 WorkloadManagement
4 {
5     Production
6     {
7         Agents
8         {
9             #@@-zhangxm@dirac_admin - 2014-07-21 10:57:29
10             SiteDirector
11             {
12                 PollingTime = 120
13                 CETypes = CREAM
14                 CETypes += SSHTorque
15                 CETypes += SSMGE
16                 CETypes += SSHCondor
17                 MaxJobsInFillMode = 1
18                 PilotLogLevel = VERBOSE
19                 ExtraPilotOptions =
20                 GetPilotOutput = False
21                 UpdatePilotStatus = True
22                 SendPilotAccounting = True
23                 FailedQueueCycleFactor = 10
24                 PilotStatusUpdateCycleFactor = 10
25             #@@-zhangxm@dirac_admin - 2014-07-21 10:52:16
26             VO = bes
27             Community = bes
28         }
29     }
30 }
31 }
32 ~
```
Credential management in StoRM
- rely on user credential for what concern user authentication and authorization
- use VOMS extension to define access policy

Multi-VO configurations in StoRM
- list of supported VOs:
  - VOS variable in storm.def
- storage areas for each VO:
  - STORM_STORAGEAREA_LIST
  - STORM_{SA}_VONAME
  - STORM_{SA}_ONLINE_SIZE
  - STORM_{SA}_DEFAULT_ACL_LIST
- VO specific users and groups in user.conf, group.conf files
- VOMS information for each VO in siteinfo/vo.d/<vo>
A lightweight frontend system (JSUB) is designed
- ganga is too complicated for us
- it takes a long time for adding plugin for new VOs

Features of JSUB
- easy for adding a plugin when new VO is joined
- support massive job splitting and submission
- support workflow control
- support task-based job management
- support Condor backend

A Prototype is completed
- support bes/cepc/juno VO
- has been used by juno and cepc users
Monitoring and Accounting

- **Monitoring system**
  - system admin will benefit from site and server monitoring
  - help find problems quickly and locating exactly
  - save manpower during the management of several VO’s resources

- **VO-related Accounting**
  - can give statistics of resource usage for different users, and different VOs
  - more functionality can be added beyond DIRAC’s accounting

- **Progress**
  - a dedicated DB server is prepared
  - prototype is under design
Summary

- Based on BES-DIRAC platform, we are making DIRAC as a service for multi-VOs (bes, cepc, juno, etc.).
- DIRAC and StoRM SE are configured to support multi-VO workload management and data management.
- Frontend and monitoring system are under developing with multi-VO considerations.