OSG PUBLIC STORAGE USING IRODS

PROGRESS UPDATE

7-JAN-2013

Tanya Levshina

Motivations

- Enable VO whose computation requires “large” data to use OSG sites more easily
  - LHC VOs have solved this problem (FTS, Phedex, LFC)
  - Smaller VOs are still struggling with large data in a distributed environment
- Ease the task of VO data management:
  - Providing quota management
  - Moving data and software to the sites
  - Retrieving the output data from the sites
  - Providing metadata catalog

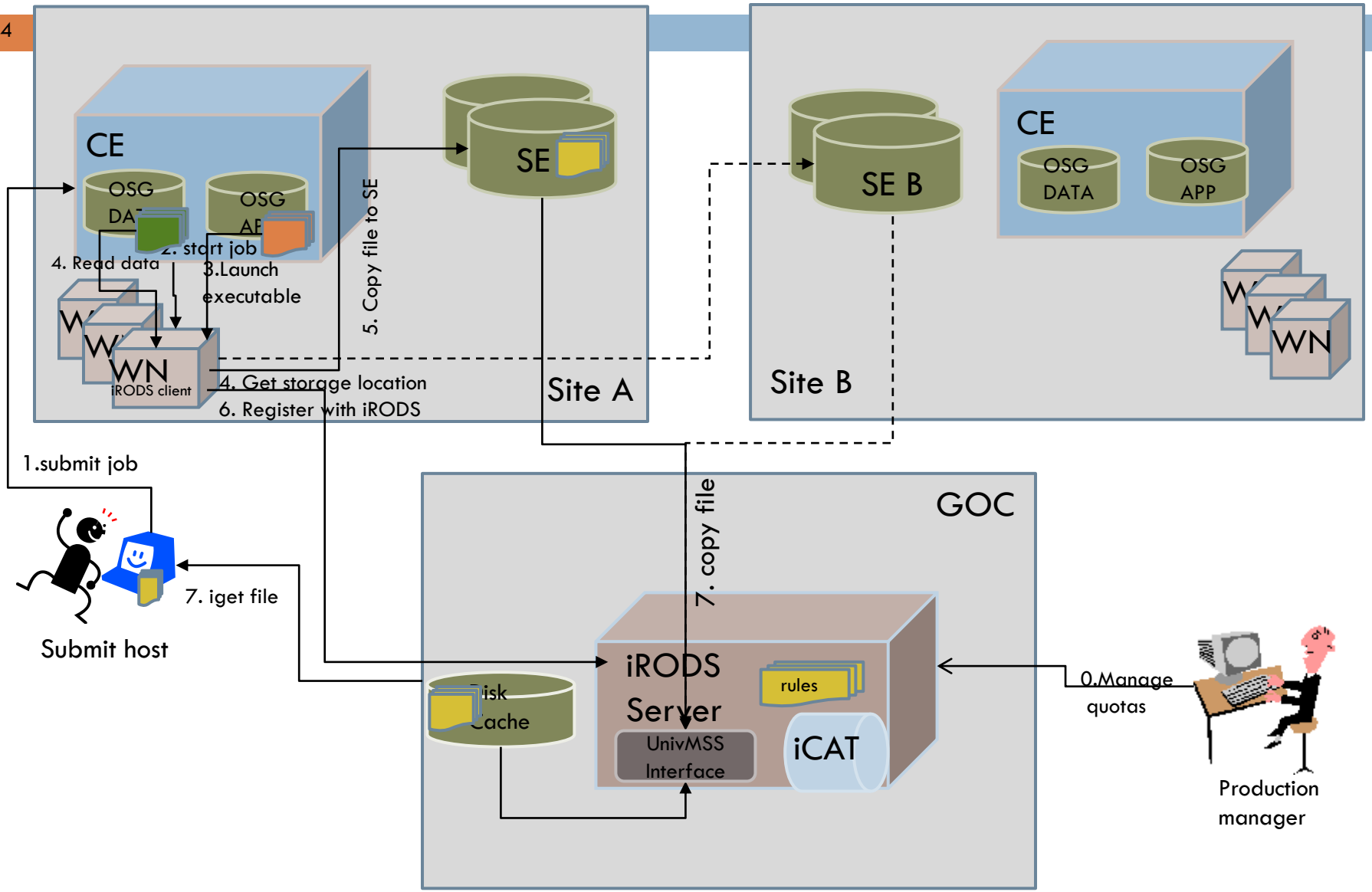
Timeline

- Concept level meeting with iRODS team in July 2011
- Phase I (proof of concept) completed in June 2012
  - Presented at the OSG Tech. Area meeting

  https://twiki.grid.iu.edu/twiki/pub/VirtualOrganizations/IRODSOSG/OSGStorage_iRODS_phaseI_final_report.pdf
- Phase II (beta trials) is wrapping up now:
  - Presentation for Operation and Production Managers

  https://twiki.grid.iu.edu/twiki/pub/VirtualOrganizations/IRODSOSG/OSGStorage_iRODS_deployment.pdf
  - Various "beta" users (EIC, SAGA, Pheno)
- What's next for Public Storage?

Architecture



Capabilities

- Storage resource management
- User management
- Resource per VO quota management
- Three resource types:
  - srmGroup consists of the OSG SRM SEs. Used for input/output data by a grid job
  - gridFTPGroup consists of the OSG Classical SEs (Allow access to OSG_DATA via gridftp server). Used for common input data pre-staging.
  - gridFTPAPPGroup consists of the gridftp servers that allow access to OSG_APP area. Used for software installation (should be phased out as soon as OASIS becomes a production service).
- Data access via unified namespace
- The intricacy of resource access (server, port, specific path for each site) is hidden from a user
- User can pre-stage/remove data to/from resource group with a single command
- User can manage data on a worker node by either using condor transfer plugin mechanism or by specifying an appropriate command in a grid job
- iRODS client commands is installed on worker nodes via glide-in pilot.

Use Cases Demonstrated (I)

- EIC tested pre-staging files to a site with OSG_DATA area and accessing these files from the worker nodes
- Pheno tested iRODS while running MPI jobs on the OSG:
  - Start jobs on any site that allows to use a "whole" node
  - Upload files to a SE from worker nodes
    - Check if a local SE is available and has enough quota or find other available SE
    - Upload file
    - Register file in iRODS
  - Download all the files to user's laptop using irods client commands
  - Ran several jobs

Use Cases Demonstrated (II)

- SAGA
  - SAGA-based BigJob applications:
    - A reference genome dataset (2-100 GB) that needs to be shared among jobs. They want to use OSG_DATA area to pre-stage this dataset.
    - Need to install application (short-read aligner called 'BFAST', specific version of python, etc) in OSG_APP area
  - SAGA-Python group uses proprietary software to deal with storage management. They are trying to adopt iRODS for their needs.
  - Several users are running tests for couple of weeks.

Issues

- We are using iRODS in non-orthodox way. We wrote plug-ins and modified commands to fit with OSG.
- iRODS developers are not interested in our use cases. Support (bugs and capability) has been minimal – consulting has been fine.
- We have put iRODS "front" so our users are shielded from the complexity and unpredictability of SEs behavior. OSG staff still needs to deal with all the SEs issues; a cost of providing this service.

Relationships & Interactions

- OASIS – for distributing applications
- AAA – for data access
  - When for smaller OSG VO?

What's next?

- Stop now:
  - Advantage:
    - 0.20 (?) FTE could be spent on something else
  - Drawbacks
    - We don't have any solution for managed public storage, resource discovery, unified namespace (catalog)
- Continue to General Deployment in OSG (Phase III)
  - Work to be done:
    - Packaging
    - Improve fault tolerances
    - Implement "watchdog" service
    - Improve monitoring
    - Deploy as OSG Service
    - Negotiate with sites quotas and space partition
  - Advantages:
    - Manageable and easy to use public storage
  - Drawbacks:
    - Maintenance and support of yet another central service
    - Involvement with SE malfunctions and misconfigurations
- Others ?