

Pigeon Tools

Storage Discovery & Testing

OSG All-Hands Meeting

3/7/2010

The Problem

- The Grid offers a tremendous opportunity!
- Many storage elements with large capacity.

However...

- Diverse resources with different limitations
- Not all are configured correctly for all VO's.
- Not all sites actually do what they advertise.
 - BDII information inaccurate
- Difficult to find, discover, and use storage.

Solutions

Discovery Tools

- Answers: What SE's support my VO?
 - Contacts BDII Information Service
 - Reports storage elements per VO
 - Command-line interface

Pigeon Tools

- Answers: What SE's support my VO and actually work?
 - Gets storage elements from Discovery Tools
 - Tests and verifies information
 - Web interface

Discovery Tools

Developer: Ted Hesselroth

- OSG provides BDII information which aggregates LDAP information about site configuration
- Discovery Tools contacts BDII to receive storage element information
 - GridFTP and SRM storage elements
- Caches information for quick access
- Command-line scripts show storage element information
- Available with OSG client:
 - `pacman -get http://software.grid.iu.edu/osg-1.2:client`

Discovery Tool Scripts

SRM Search Scripts

- **get_surl**: Retrieve SRM for each SE
- **get_srmcp_cmd / get_srm_copy_cmd**: Show SRM copy command (either FNAL or LBNL clients) for SRM storage elements
- **get_srm_reserve_space_cmd / get_srm_sp_reserve_cmd**: Show reserve space command (FNAL/LBNL clients) for each SRM storage element
- **get_storage_element_id**: Show list of ids of each SE
- **get_srm_storage_element_id**: Show list of ids of each SE (SRM endpoints only)
- **get_site_name**: Get the name of the site for each SE
- **get_site_name_for_seid**: Takes id and returns the site name

Discovery Tool Scripts

Gridftp Search Scripts

- **get_gridftp_url:** Shows the gsiftp URL for each GridFTP server (classic Storage Element)
- **get_gridftp_storage_element_id:** Gets the storage id for each GridFTP server

Discovery Tool Scripts

Compute Scripts

- **get_mount_path:** Shows local mount path for storage element on worker nodes that have a POXIS mount.
- **get_os_versions:** Shows the operating system and version for each site
- **get_runtime_versions:** Finds runtime software versions for clusters and sites (e.g. OSG VERSION)

Discovery Tools - Example

```
[dstrain@gw014kx ~]$ get_surl --vo Engage
```

```
srm://bsrm-1.t2.ucsd.edu:8443/srm/v2/server?SFN=/hadoop/engage/TESTFILE
srm://charm.ucr.edu:10443/srm/v2/server?SFN=/data/bottom/cms/TESTFILE
srm://cit-se.ultraviolet.org:8443/srm/v2/server?SFN=/mnt/hadoop/osg/engage/TESTFILE
srm://cmssrm.hep.wisc.edu:8443/srm/managerv2?SFN=/pnfs/hep.wisc.edu/data5/engage/TESTFILE
srm://dcsrm.usatlas.bnl.gov:8443/srm/managerv2?SFN=/pnfs/usatlas.bnl.gov/osg/engage/TESTFILE
srm://ff-se.unl.edu:8443/srm/v2/server?SFN=/panfs/panasas/CMS/data/engage/TESTFILE
srm://fndca1.fnal.gov:8443/srm/managerv2?SFN=/pnfs/fnal.gov/usr/fermigrid/volatile/engage/TESTFI
...
```

```
[dstrain@gw014kx ~]$ get_mount_path --vo Engage
```

COMPUTE ELEMENT ID	MOUNT POINT
ff-grid2.unl.edu:2119/jobmanager-pbs-default	/panfs/panasas/CMS/data/engage
ff-grid3.unl.edu:2119/jobmanager-pbs-default	/panfs/panasas/CMS/data/engage
ce.grid.unesp.br:2119/jobmanager-pbs-default	/store/engage
ce.grid.unesp.br:2119/jobmanager-pbs-long	/store/engage
ce.grid.unesp.br:2119/jobmanager-pbs-short	/store/engage
antaeus.hpcc.ttu.edu:2119/jobmanager-sge-cms	/lustre/hep/osg/engage
antaeus.hpcc.ttu.edu:2119/jobmanager-sge-tigre	/lustre/hep/osg/engage
...	

Pigeon Tools

Problems:

- Information reported to BDII inaccurate
- Storage Elements are not always configured to support all VOs.
- Pigeon **queries** BDII via Discovery Tools to find SRM and GridFTP storage elements that advertise support for a VO.
- Pigeon will **verify** this information by testing via SRM clients or globus-url-copy (for GridFTP) to test access
- Pigeon can catch mount point / path problems, authentication issues, etc by using the VO's actual credentials.

Pigeon Versions

- Stand-alone Pigeon
 - Independent tarball
 - Good for quickly starting up
 - Can be used without root privileges
 - Has ability to link to GOC tickets
- Pigeon for RSV
 - Part of VDT
 - Good for sites that already use RSV
 - More supported and tested (ITB cycle)

Pigeon Features

- Can find SRM servers and tests:
 - srm-ping, srm-ls, srm-copy, srm-mkdir, srm-rm, srm-rmdir
- Can find classicSE's (gridFTP on head node):
 - Tests globus-url-copy
- Archives previous test results
- Understands Dependencies
 - Won't run further tests if earlier tests fail

Requirements

To use Pigeon, you need:

- Python 2.4+: provided in SL4+
- Valid x509 cert w/VO proxy: needed to run grid commands
- Apache: comes with VDT
- Discovery Tools: comes with VDT.
- Pigeon cannot be on a read-only partition.

Installation Overview - RSV

- Download and setup VDT
 - `pacman -get http://software.grid.iu.edu/osg-1.2:pigeon-tools`
 - `source setup.sh`
 - `vdt-ca-manage setupca --location local --url osg`
 - `vi osg/etc/config.ini`
- Configure Pigeon
 - `configure_pigeon --rsvuser rsv_user --vo voname --server y`
- Run Pigeon
 - Run “pigeon” (manual execution)
 - `vdt-control --on pigeon-tools` (enables nightly cron execution)


HTML Results - RSV

RSV Status - 2011-01-24 09:42:24 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

https://gw014k1.fnal.gov:9443/pigeon/

RSV Status - 2011-01-24 09:42:24



Open Science Grid

FAQ

Site: --srm-destination-dir (srm-dcache.rcac.purdue.edu:8443)

Metric	Last Executed	Status
org.osg.globus.gridftp-simple	2011-01-24 06:05:55 CST	OK

Site: BNL-ATLAS (gridgk01.racf.bnl.gov:2811)

Metric	Last Executed	Status
org.osg.globus.gridftp-simple	2011-01-24 05:47:44 CST	UNKNOWN

Site: BNL-ATLAS-ITB (gridtest01.racf.bnl.gov:2811)

Metric	Last Executed	Status
org.osg.globus.gridftp-simple	2011-01-24 05:47:44 CST	UNKNOWN

Site: BNL-ATLAS_2 (dcsrm.usatlas.bnl.gov:8443)

Metric	Last Executed	Status
org.osg.srm.srmclient-ping	2011-01-24 05:30:20 CST	OK
org.osg.srm.srmclient-put	2011-01-24 05:30:24 CST	OK
org.osg.srm.srmclient-rm	2011-01-24 05:30:43 CST	OK
org.osg.srm.srmclient-mkdir	2011-01-24 05:29:53 CST	OK
org.osg.srm.srmclient-rmdir	2011-01-24 05:30:05 CST	OK

Site: CIT_CMS_T2 (cit-se.ultralight.org:8443)

Metric	Last Executed	Status
org.osg.srm.srmclient-ping	2011-01-24 05:42:44 CST	OK
org.osg.srm.srmclient-put	2011-01-24 05:42:50 CST	OK
org.osg.srm.srmclient-rm	2011-01-24 05:43:06 CST	OK
org.osg.srm.srmclient-mkdir	2011-01-24 05:43:27 CST	OK
org.osg.srm.srmclient-rmdir	2011-01-24 05:43:39 CST	OK

Site: Firefly (ff-se.unl.edu:8443)

Metric	Last Executed	Status
org.osg.srm.srmclient-ping	2011-01-24 05:24:47 CST	OK
org.osg.srm.srmclient-put	2011-01-24 05:24:52 CST	OK
org.osg.srm.srmclient-rm	2011-01-24 05:25:06 CST	OK

Done

Results

Pigeon has found many problems that were later fixed via GOC tickets:

- SE defaults to support all VOs
 - Even if intention is to support only 1 VO
- Wrong storage path
- Wrong port
 - Not updated after upgrade

Results

- Authentication problems
 - DN not mapped
 - DN mapped to wrong VO
 - Found Bug with gridmap file and users with multiple Vos
- Test installations reporting to BDII
- Better documentation of config.ini options

Summary

Procedure:

- Install Pigeon / Discovery Tools
- Find Storage Elements (Discovery)
- Test Storage Elements (Pigeon)
- Fix any broken SE's (Submit GOC ticket)

Result:

- More Storage Elements Available
- Better Information Reporting in BDII
- In Use with Engage, LIGO, GlueX VOs

More Information

- Pigeon Tools
 - <https://twiki.grid.iu.edu/bin/viewauth/ReleaseDocumentation/RSVPigeonTools>
- Discovery Tools
 - <https://twiki.grid.iu.edu/bin/viewauth/ReleaseDocumentation/OSGStorageDiscoveryTool>