OSG 1.2 and GUMS Installation at FLTECH

Patrick Ford

Following the OSG Administrator workshop in August 2009, we have upgraded the site to the latest version of the virtual data toolkit (2.0) and therefore the OSG 1.2 middleware. We also implemented the Grid User Management System (GUMS) as a replacement for the grid-mapfile.
Previously, we had the Bestman storage element installed on our compute element (rocks frontend). In order to accomplish this we had to manually change the ports that the VDT Apache process ran on. Unfortunately, GUMS refused to run under this hacked configuration, so we are forced to run a separate Storage Element server. This will be implemented in the near future.

Listed are some important acronyms or terms related to the OSG Middleware.

- VDT - Virtual Data Toolkit: The package that contains and installs all OSG software.
- DN - Distinguished Name: A personal identity linked to a grid certificate.
- VO - Virtual Organization: An organization that is responsible for a set of DN
- GUMS - Grid User Management System: A piece of software that maps grid users (DN) to local cluster accounts.
- CE - Compute Element: System that runs most of the critical software (e.g. globus, apache, RSV).
- SE - Storage Element: System that runs a storage manager (such as BeStMan or dCache), sometimes is the same as the CE.

1. Pacman

To install Pacman we created a directory: /usr/local/pacman – and then ran the following commands:

```
# wget http://physics.bu.edu/pacman/sample_cache/tarballs/pacman-latest.tar.gz
```

This should be the latest version of Pacman to get the latest VDT.

```
# tar –no-same-owner –xzvf pacman-latest.tar.gz
# cd pacman-<version>
# ./setup.sh
```

2. Condor

It is possible to install Condor through the VDT, but I don't recommend this. This guide will assume that you have an existing Condor infrastructure.

Set the following variables (use your install directory) either manually or in /etc/profile

```
# export VDTSETUP_CONDOR_LOCATION=/opt/condor/
# export VDTSETUP_CONDOR_CONFIG=$VDTSETUP_CONDOR_LOCATION/etc/
condor_config
```

3. OSG:CE

Set up the versioning.
Note: You can follow along in the OSG documentation, this guide just simplifies it.
https://twiki.grid.iu.edu/bin/view/ReleaseDocumentation/ComputeElementInstall

```
# mkdir /opt/osg-1.2
# ln -s /opt/osg-1.2 /opt/osg
# cd /opt/osg
```

Now install OSG services for the compute element. At the time of install, we had to use a direct link to the VDT software cache rather than the OSG:<package> link.

```
# pacman -get http://software.grid.iu.edu/osg-1.2:ce
# source setup.sh
```

Also, you may want to add the script to /etc/profile as you will need to run it a lot during the setup of OSG.
Now make sure that the condor environmental variables have set correctly and run the following command:

```
# pacman -get http://software.grid.iu.edu/osg-1.2:Globus-Condor-Setup
```

Before moving on to the next steps, your /etc/profile should look like this:

```
export VDT_LOCATION=/opt/osg
export VDTSETUP_CONDOR_LOCATION=/opt/condor
export VDTSETUP_CONDOR_CONFIG=$VDTSETUP_CONDOR_LOCATION/etc/condor_config
export GLOBUS_TCP_PORT_RANGE=40000,40200
export GLOBUS_TCP_SOURCE_RANGE=40000,40200
source $VDT_LOCATION/setup.sh
source $VDT_LOCATION/globus/etc/globus-user-env.sh
```

Installing Managed Fork is a good idea to avoid “fork-bombs” where a user can run too many jobs on your cluster.

```
# pacman -get  http://software.grid.iu.edu/osg-1.2:ManagedFork
# $VDT_LOCATION/vdt/setup/configure_globus_gatekeeper --managed-fork y --server y
```

Install the CA Certificates in the local configuration.

```
# $VDT_LOCATION/vdt/bin/vdt-ca-manage setupca --location local --url osg
```

```
# ln -s $VDT_LOCATION/globus/TRUSTED_CA /etc/grid-security/certificates

# vdt-post-install
```

4. Certificates

A host machine needs a certificate on the root account to authenticate the gatekeeper, and the admin needs a user certificate in order to run jobs via globus. If you intend to run RSV you will also need the rsv and http service certificates.
To get the host certificate, simply run the following command and follow instructions:

```
# cert-request –ou s –dir . –label uscms1
```

To get the service certificates, run the following commands.

```
# cert-request -ou s -service rsv -host uscms1.fltech-grid3.fit.edu -label rsv-uscms1
# cert-request -ou s -service http -host uscms1.fltech-grid3.fit.edu -label http-uscms1
```

When you receive your host certificate and produced a private key, you must login as root and put them both into ~/.globus/ naming them usercert.pem and userkey.pem
Do the same for your own user cert in your own user account. Keep a backup copy of all certificates and keys. Also put a copy of the host certificate and key in /etc/grid-security/ named hostcert.pem and hostkey.pem (with root owner), and also named containercert.pem and containerkey.pem (with globus owner).
When you do the same for the service certificates, place the rsvcert and key into /etc/grid-security (owned by rsvuser, create it) and the httpcert and key into /etc/grid-security/http.

To renew grid certificates, you must request new ones in the same fashion and overwrite the old ones. For personal certificates you can use cert-renew.

5. Configuring OSG

You must now set up the configuration python script.

```
# cd $VDT_LOCATION
# ./setup.sh (if you have not added it to profile)
# cd osg/etc (not a typo)
# vim config.ini
```

Set up this file with the information related to your site. Most of it is self-explanatory and documentation is included in comments, but there are detailed instructions on the OSG TWiki. For FLTECH we have stored this configuration file in /opt/osg/Scripts so it can be copied over.
Run the following command:

```
# configure-osg -c
```

6. Start VDT Control

Ensure that all standard services are enabled (all default except gums):

```
$ vdt-control --list
```

Run the following command as root:

```
# vdt-control --on
```

This will set up everything standard for OSG, but now we will want to set up the GUMS Server. You may get some errors since our config.ini is set up for GUMS authentication, but running this once should serve its purpose.

```
# vdt-control --off
```

Now make absolutely sure there are no processes running from the VDT or bad things may happen (such as cryptic errors that will make no sense). This can usually be accomplished by checking if any processes are running under osg-1.2 and killing them.

```
# ps aux | grep osg-1.2
# ps aux | grep globus
# kill -9 <offending process>
```

7. GUMS

GUMS can be a little tricky. Since you started up the VDT, a link should have been made from /var/lib/mysql/mysql.sock to /opt/osg-1.2/vdt-app-data/mysql5/var/mysql.sock
If not, the mysql5 server that GUMS depends on will not work properly. We are running GUMS on the CE, which is apparently frowned upon for security/failsafe reasons, but I figure that if someone gets control of the CE, you will have bigger problems. Many other sites also do the same for simple installations, although it is possible to set up high availability redundant GUMS servers.

First, the following is the mode we run GUMS in. We want XACML2 GUMS.
https://twiki.grid.iu.edu/bin/view/ReleaseDocumentation/FullPrivPreConfig

```
# cd $VDT_LOCATION
# pacman -get http://software.grid.iu.edu/osg-1.2:gums
# source setup.sh
# vdt-post-install
# configure-osg -c
```

This should have set up the services correctly, you can check if edg-mkgridmap has been disabled and gums-host-cron enabled:

```
# vdt-control -list
# vdt-control -on
```

You should check if the web server works. Load your personal certificate into your browser and go to:
http://uscms1.fltech-grid3.fit.edu:8080
That should link to the private area (port 8443) and you can check the RSV test status and GUMS interface.

You can force the RSV probes to run immediately (to check if everything is working). NOTE: RSV will probably NOT be working right now, as GUMS does not map the rsvuser account by default. We will cover this later. To run this command you will have to su to root, and then su rsvuser - it will only work if you are rsvuser.

```
$ ${VDT_LOCATION}/osg-rsv/bin/misc/run-rsv-probe-by-hand.sh uscms1.fltech-grid3.fit.edu ${VDT_LOCATION}
```

The GUMS web interface is not ready yet. First we have to set up the administrator and config template (using my DN as example)

```
# cd tomcat/v55/webapps/gums/WEB-INF/scripts
# ./gums-add-mysql-admin "/DC=org/DC=doegrids/OU=People/CN=Patrick Ford 571150"
# ./gums-create-config --osg-template
```

You can now navigate to the GUMS web interface as before, and click on Update VO Members to get the user lists from the VOs.

In order to make RSV work, we have to create a group for the rsvuser (which RSV probes run under) and set up an account mapper for it.
Go to User Groups and scroll to the bottom and click add to create the group. I called it rsv, and it should be of type "manual" and persistence factory "mysql" and GUMS access "read all".
Now go to Account Mappers and add a mapper named rsv, of type "group" and account "rsvuser".
Now go to Group to Account Mappings and add a mapping from User Group "rsv" to Account Mapper "rsv".
Lastly, go to Host to Group Mappings and edit the existing Host to include the rsv group.

Phew! It's an extensive process. It's almost easier to register your DN with a VO and be updated automatically. My DN is part of the CMS VO so it is added already, but if you have a basic personal cert you will have to repeat the steps above to be authenticated.

Since I have not yet set up more than one local account for CMS users, I mapped all CMS subgroups to the uscms01 account. To do this you change the cms groups in Account Mappers to map to uscms01 user. Then you can go to Group to Account Mappers and check that the groups are linked to the mappers correctly.

This should be enough for a basic configuration.

8. IPTables

Several OSG services need certain ports opened and available. Ensure that your iptables file has the following ports open:

```
## INCOMING ##
tcp 2119, tcp 2811, tcp 2135, tcp 9443, udp 9000, tcp 40000 to 40200, tcp 9000 to 9010,
tcp 39281, tcp 7512, tcp 2136, tcp 8443, tcp 8080

## OUTGOING ##
tcp 40000 to 40200, tcp 9443
```

9. Running Tests

Up to this point, you should have the environmental variables I listed above set correctly, and an updated grid-mapfile, CRL, and ports opened in IPTables. You should be logged into your own user account and have your usercert and userkey in your ~/.globus/ directory. Create a proxy by running the following commands.

```
$ grid-proxy-init
$ grid-proxy-info –identity
```

Now test the fork and condor queue

```
$ globus-job-run $(hostname –f):2119/jobmanager-fork /usr/bin/id
$ globus-job-run $(hostname –f):2119/jobmanager-condor /usr/bin/id
```

These commands may take a while and should return your VO as the ID (in my case it was uscms01)

Now test the Grid FTP: GSIFTP – this includes creating a test file, copying it to the location on your cluster reserved for grid data (/mnt/nas0/OSG/DATA) through globus, and then verifying that the file copied.

```
$ echo “My test gsiftp file” > /tmp/gsiftp.test
$ source $VDT_LOCATION/monitoring/osg-attributes.conf (for data variable)
```

```
$ globus-url-copy file:/tmp/gsiftp.test gsiftp://$(hostname)${OSG_DATA}/gsiftp.test
$ ls –l $OSG_DATA/gsiftp.test
```

This file should exist in the directory /mnt/nas0/OSG/DATA/

10. Site Verify

NOTE: Site Verify (or gridscan) is becoming outdated, RSV is the best way to see if your site is operating normally. The new service myOSG keeps track of this. You should also check that you are broadcasting information to BDII. At this point you should force the RSV probes to run again (see earlier in the guide).

http://myosg.grid.iu.edu
http://is.grid.iu.edu/cgi-bin/status.cgi

Run the site verification script as your user with an open proxy.

```
$ $VDT_LOCATION/verify/site_verify.pl --host=uscms1.fltech-grid3.fit.edu
```

You should pass all tests if you followed the instructions completely. Now you need to register your site with the Grid Operations Center in order to get external verification.