EDG Configuration and Authentication

October 29, 2001

CERN

Anders Wäänänen, Niels Bohr Institute, NorduGrid

waananen@nbi.dk

Summary

- Overview of the EDG Globus Configuration
  - Gatekeeper and Jobmanager
  - GridFTP
  - Information system
- Authentication and Security Issues
  - Certificate Authorities
  - Certificates for Testbed 1
  - Open ports

Running a Testbed site 101

- To run a testbed site or "being part of the Grid" means providing services to Grid users
- The users need to be able to contact these services (jobmanagers) through a single point (gatekeeper).
- Information about the resources should also be provided (MDS) to let the user and brokers get detailed information about the site.
- The site needs to now the users (authentication) and whether to allow them access or not (authorization).

Globus Configuration

- Has historically been quite complex
- Many different configuration files
- Same information occur in several places

- Very easy to make errors
- Not easy to update information
- Steep learning curve for busy site administrators

EDG Globus Configuration

- Centralized configuration
- Information should only be entered once
- Should be easy to update/change information
- Sensible default values
- System administrators should only need to enter site specific configuration
- Allow "Expert" configuration
- Should work "out of the box" (almost)
- No hard-wired site specific values – Not even EDG specific

File locations

- Globus root: /opt/globus
  - Relocation not (yet) supported – use soft links
- Red Hat / Linux file locations instead of Globus centric locations:
  - /opt/globus/sbin -> /etc/rc.d/init.d (SysV scripts)
  - /opt/globus/etc -> /etc/profile.d (user profiles)
  - /opt/globus/var -> /var/log (log files)
  - /opt/globus/var -> /var/run (pid files)
  - …
- In the future /opt/globus could be made read-only and on a shared filesystem

/etc/globus.conf

- Common configuration file /etc/globus.conf
- Works identically to Red Hat's /etc/sysconfig/ files
- Simple format:
  - # at start of line means a comment
  - MACRO=VALUE
- Example:
  - # This is a comment line
  - GLOBUS_LOCATION=/opt/globus
- It is not a shell script, but can be used in shell scripts by sourcing

Notes on MACRO names

- Try to be consistent with MACRO names across sub-packages
  - Same MACROS for same functionality
    - GLOBUS_LOCATION is used by all packages to specify Globus install path
  - Consistent MACRO naming while trying to maintain the original Globus namespace as much as possible:
    - X509_*GATEKEEPER*_KEY
    - X509_*GIIS*_KEY
    - *GATEKEEPER*_LOGFILE
    - *GRID_INFO*_LOGFILE

EDG Config packages

- Should replace the Globus setup packages
- No manual setup using these packages are needed
- EDG config RPMs are named as: globus_<package>-edgconfig
- Where "package" as a first approximation is the name of the corresponding setup package which it replaces
- Current EDG configuration packages:
  - core, common, profile, gatekeeper, gsi_wuftpd, mds, doc
- These packages (together with the CA packages) fully replaces the standard GLOBUS setup procedure
- No RPM post-install scripts (requirement)
- Local (internal) configuration files are (re)created by each restart of service with parameters from /etc/globus.conf

Core, Common and Profile

- Core
  - Replaces globus_core_setup
  - Basic scripts (Globus file system layout, Unix tools locations)
- Common
  - Replaces globus_common_setup
  - Common scripts (hostname, domainname)
- Profile
  - Set user environment including path

Common options

- GLOBUS_LOCATION
  - Root of Globus installation (/opt/globus)
- X509_CERT_DIR
  - CA certificate directory (/etc/grid-security/certificates)
- GRIDMAP
  - Grid Mapfile (/etc/grid-security/grid-mapfile)

Gatekeeper & Jobmanager

- The gatekeeper machine act as an interface to the local scheduler. It runs the gatekeeper daemon and the jobmanagers which calls the low level schedulers – all on the same host.
- Started through the SysV script called globus-gatekeeper:

  `/etc/rc.d/init.d/globus-gatekeeper`

- Maintained on Red Hat with service and chkconfig(8):
  - `chkconfig globus-gatekeeper on`
  - `service globus-gatekeeper start`
- Not (yet) support for start-up through (x)inetd

Gatekeeper Options

- GATEKEEPER_PORT
  - Port number of Gatekeeper (2119)
- GATEKEEPER_LOG
  - Gatekeeper logfile (/var/log/globus-gatekeeper.log)
- GATEKEEPER_USER
  - Username of the gatekeeper (not specfied (root))
- X509_GATEKEEPER_CERT
  - Gatekeeper certificate (/etc/grid-security/hostcert.pem)
- X509_GATEKEEPER_KEY
  - Gatekeeper key (/etc/grid-security/hostkey.pem)

Jobmanager Options

- To setup the local job-managers simply define GLOBUS_JOBMANAGERS:
  - Eg.: GLOBUS_JOBMANAGERS="fork pbs"
- An empty GLOBUS_JOBMANAGERS makes a very dull gatekeeper
- The contact string is then given by:
  - hostname/jobmanager-<jobmanager_type>
  - Eg: grid.cern.ch/jobmanager-lsf or grid.cern.ch:12345/jobmanager-fork
- The default job-manager (simply called "jobmanager") will be the first jobmanager listed in GLOBUS_JOBMANAGERS
- At the moment the fork manager is needed by some client programs (globus-job-get-output and globus-job-clean)
  - You should provide a fork job-manager if you want to support these clients
  - If fork is not the default job-manager this must be specfied by the users when running these programs

Jobmanager Options (continued)

- Supported Testbed 1 job-managers: fork, pbs and lscf
- The location of the backend jobmanager programs must be specified in /etc/globus.conf (if they are not in root's path):
- Example for PBS:
  - GLOBUS_GRAM_JOB_MANAGER_QDEL=/usr/local/pbs/bin/qdel
  - GLOBUS_GRAM_JOB_MANAGER_QSTAT=/usr/local/pbs/bin/qstat
  - GLOBUS_GRAM_JOB_MANAGER_QSUB=/usr/local/pbs/bin/qsub
- The GRAM reporter (reporting system for the backend scheduler) setup is handled by the MDS system

Information system

- At the moment 2 different information system exists
  - MDS 2.1 (integrated with Globus)
  - OpenLDAP Ftree (does not require Globus)
- Both have advantages and disadvantages
- Can be run in parallel on the same machine (uses different ports)
- Will hopefully be merged soon

Globus MDS 2.1

- Very new – Globus just delivered a couple of weeks ago
- The new MDS 2.1 provides among other things support for virtual organizations (VO)
- Runs off a single port (2135) with both GRIS and GIIS
- Started through SysV with name: globus-mds
- Support non-anonymous queries to the GIIS
- Configured completely through /etc/globus.conf
- Configuration not completely finalized
- Run as non-root using GRID_INFO_USER

Globus GRIS Configuration

- GRIS – Grid Resource Information Service
- GRIS startup:
  - GRID_INFO_GRIS_ACTIVE
    - Start the GRIS (y)
- Registration to GIIS:
  - GRID_INFO_GRIS_REG_HN=
  - GRID_INFO_GRIS_REG_PORT=
  - GRID_INFO_GRIS_REG_PERIOD=
- Void values means a local GIIS
- GRAM reporter automatically setup using GLOBUS_JOBMANAGERS

Globus GIIS Configuration

- GIIS - Grid Index Information Server
- Certificate and key for non-anonymous queries
  - X509_GIIS_CERT
  - X509_GIIS_KEY
- Multiple VirtualOrganizations (VO's) on different/same node.
- Many more details in the afternoon demo

OpenLDAP Ftree

- New OpenLDAP backend "ftree"
- GRIS and GIIS on separate ports
- Interface to various backend programs from
  - WP1, WP4, WP5, WP7
- Port setup:
  - Site GIIS:
    - cached: 2173
    - Using referrals: 2169
  - Compute Elements (ce):
    - PBS: 2171
    - LSF: 2172
- Afternoon demo?

Testbed 1 Certification Authorities

- To use Globus and the Grid one needs certificates
- Certificate authorities can issue certificates
- Testbed 1 approved certificate authorities:
  - http://marianne.in2p3.fr/datagrid/ca/ca-table-ca.html
- From the list (currently 13) one should identify which CA one belongs to.
- This CA will provide user, host and server certificates
- A host might need several certificates (gatekeeper, mds, ...)
- Instructions for interacting with each CA can be found from the local CA's web pages

Authenticating on Testbed 1

- Each host running a Grid service needs to be able to authenticate users and other hosts
- Therefore the CA certificate for each CA needs to be installed
  - (This certificate can be thought of the public key of the CA).
- RPMs can be found from:
  - http://datagrid.in2p3.fr/distribution/config/security.html
- RPM naming: ca_<CA_ALIAS> Eg: ca_CERN
- Installing these RPMs does not allow access (*authorization*) to local grid services. It just provides *authentication*.
- All CA certificates should and can safely be installed without compromising local site security.

Authenticating on Testbed 1 (continued)

- CA certificate location: /etc/grid-security/certificates/
  - Controlled by X509_CERT_DIR in /etc/globus.conf
- CA filenames are made of the hash of the CA certificate with .0 appended: eg. 1f0e8352.0
- To specify which certificates a CA can sign a policy file for each CA is needed.
  - The policy file is located in the same directory as the CA certificate with a .signing-policy appended to the hash: eg. 1f0e8352.signing-policy
  - This replaces ca-signing-policy.conf from Globus 1.
- Other CA specific files used by various tools are located in the CA certificate directory
- To get the hash of a certificate: openssl x509 -in certificate -noout -hash

grid-cert-request

- Some CA support certificate requests through the Globus grid-cert-request program.
- For this to work one should install the "local" CA RPMs:

  http://datagrid.in2p3.fr/distribution/config/security.html

- These packages replaces the globus-sslutils-setup scripts.
- Current CA with "local" support:
  - CERN, GridPP, INFN, NorduGrid, Russia

Certificate Revocation

- Certificates can be revoked for several reasons:
  - Comprised keys, forgotten passwords, no longer used,..
- To communicate this to site maintainers each CA provides online *Certificate Revocation Lists* (CRL's).
- Each CA makes this available in different ways.
- The tool edg-fetch-crl provides a common interface to all the CRLs and should be run on a regular basis (eg. each day) from cron(8) on all machines providing Grid services
- It is provided by the RPM package edg-utils
- Basic usage:
  - edg-fetch-crl -o /opt/grid-security/certificates
- CRL filenames is the hash of the CA certificate with .r0 appended.

Default Port numbers / firewalls

- Port numbers can be specified in /etc/services eg.:
  - globus-gatekeeper 2119/tcp # Globus Gatekeeper
- Useful for netstat(8) and lsof(8), but not required
- The default port numbers used by the Testbed software:
  - 2119 (Gatekeeper)
  - 2811 (GSI WU-FTPD)
  - 2135 (Globus MDS)
  - 2169, 2173 (Ftree site GIIS)
  - 2171, 2172 (Ftree compute elements)

Documentation

- Under Construction...
- Contents:
  - Administrators Guide (exists)
  - User's Guide
  - Developer's Guide
  - globus.conf template (exists)
    - Includes all valid configuration options