Introduction to OSG Fundamentals

Suchandra Thapa
Computation Institute
University of Chicago

Overview

- Introduction to OSG terms and operations
- Will discuss OSG fundamentals
- Have about ?? minutes for a ?? timeslot
- Questions encouraged
- Q&A time afterwards

Introduction to OSG

- OSG stands for Open Science Grid
- Provides high-throughput computing across US
  - Currently more than 70 sites
  - Recent stats:
    - 282,912 jobs for 433,051 hours
    - Used 75 sites
    - Jobs by ~20 different virtual organizations
    - 92% of jobs succeeded
    - Underestimate: 4 sites didn't report anything
  - Provides opportunistic computing for VOs
- Focus on high-throughput computing rather than high performance computing

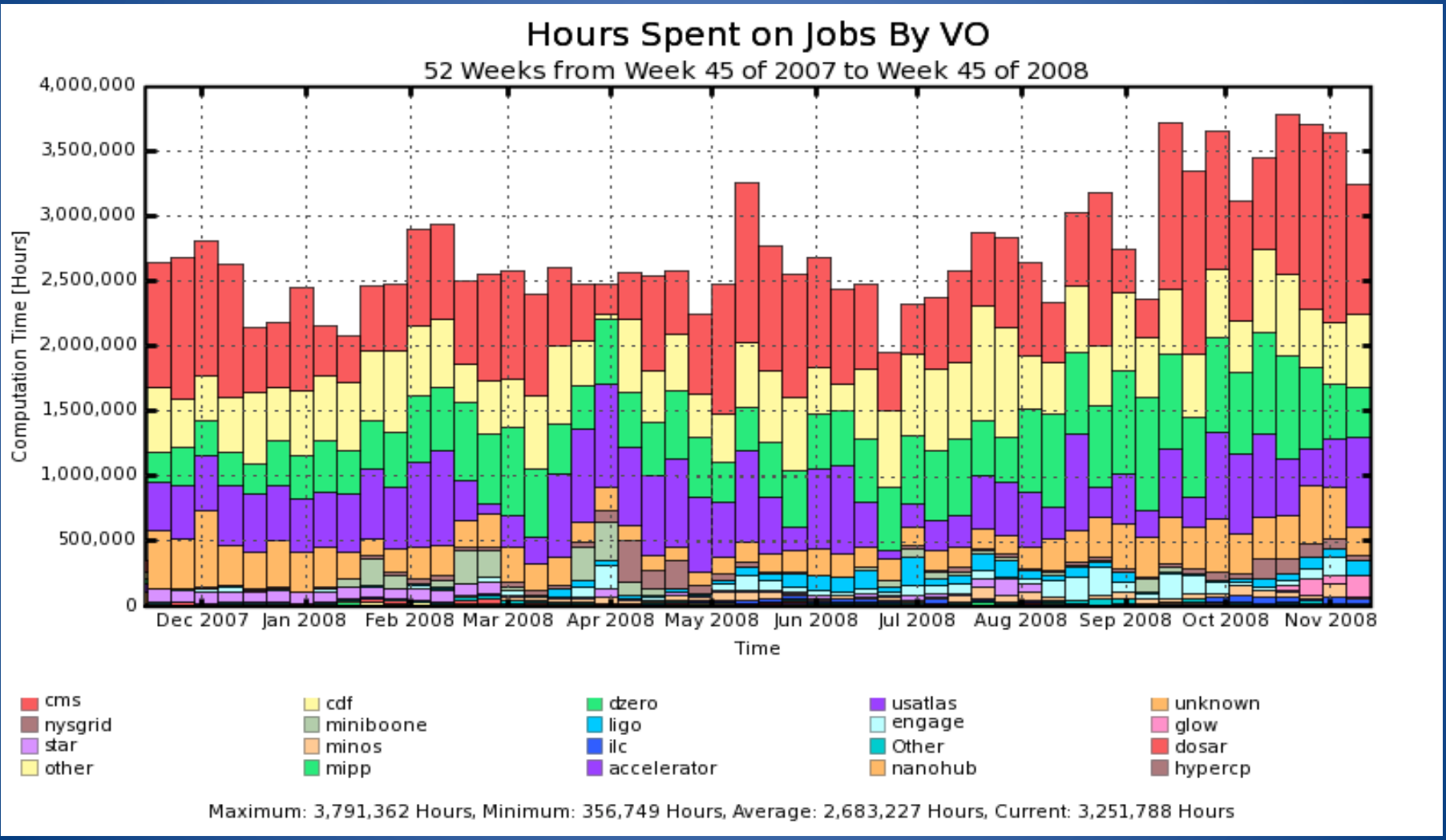
Basic Terms

- CE – Compute Element
- SE – Storage Element
- VO – Virtual Organization
- WN – Worker Node
- GOC – Grid Operations Center
- VDT – Virtual Data Toolkit
- DN – Distinguished name

Who uses OSG?

- About 20 virtual organizations
  - High-energy physics uses a large chunk of OSG
  - But several other sciences are actively using OSG.
    - nanoHUB: nanotechnology simulations
    - LIGO: detecting gravitational waves
    - CHARMM: molecular dynamics
- More at http://www.opensciencegrid.org/About/What_We're_Doing/Research_Highlights

Typical usage breakdown



Overriding principle: Autonomy

- Sites and VOs are autonomous
  - Admins are free to make decisions about site
  - OSG provides software and recommendations about configuration
  - Admins are allowed to decided when and if to upgrade
  - Admins are responsible for site but OSG provides operational support

Your role as an admin

- As a site admin, you should:
  - Keep in touch with OSG (downtime, security, etc.)
  - Respond to trouble tickets or inquiries from GOC
  - Plan your site's layout
  - Update software as needed (within limits)
  - Participate and be a good community member

Support provided for admins

- OSG provides:
  - Software and ancillary information (configuration tools, documentation, recommendations)
  - Assistance in keeping site running smoothly
  - Help in troubleshooting and installing software
  - Users for your site
  - An exciting, cutting-edge, 21st-century collaborative distributed computing grid cloud buzzword-compliant environment

VDT

- Stands for Virtual Data Toolkit
- Team based in Madison at UW-Madison
- Integrates and provides a large collection of software
- Provides the software distribution that many US grids use (including OSG)
- http://vdt.cs.wisc.edu

VDT Example

- GUMS
  - Authorizes users at a site
  - Maps global user name to local UID
- VDT includes dependencies. For example, GUMS needs:

Apache	CA Certificates
Tomcat	Configuration scripts
Mysql	Infrastructure

OSG Software Stack

- Consists of:
  - VDT Software
    PLUS
  - Additional OSG Specific bits
- E.g. CE
  - VDT Subset
    - Globus
    - RSV
    - PRIMA
    - … and another dozen
  - OSG bits:
    - Information about OSG VOs
    - OSG configuration script (configure_osg.py)

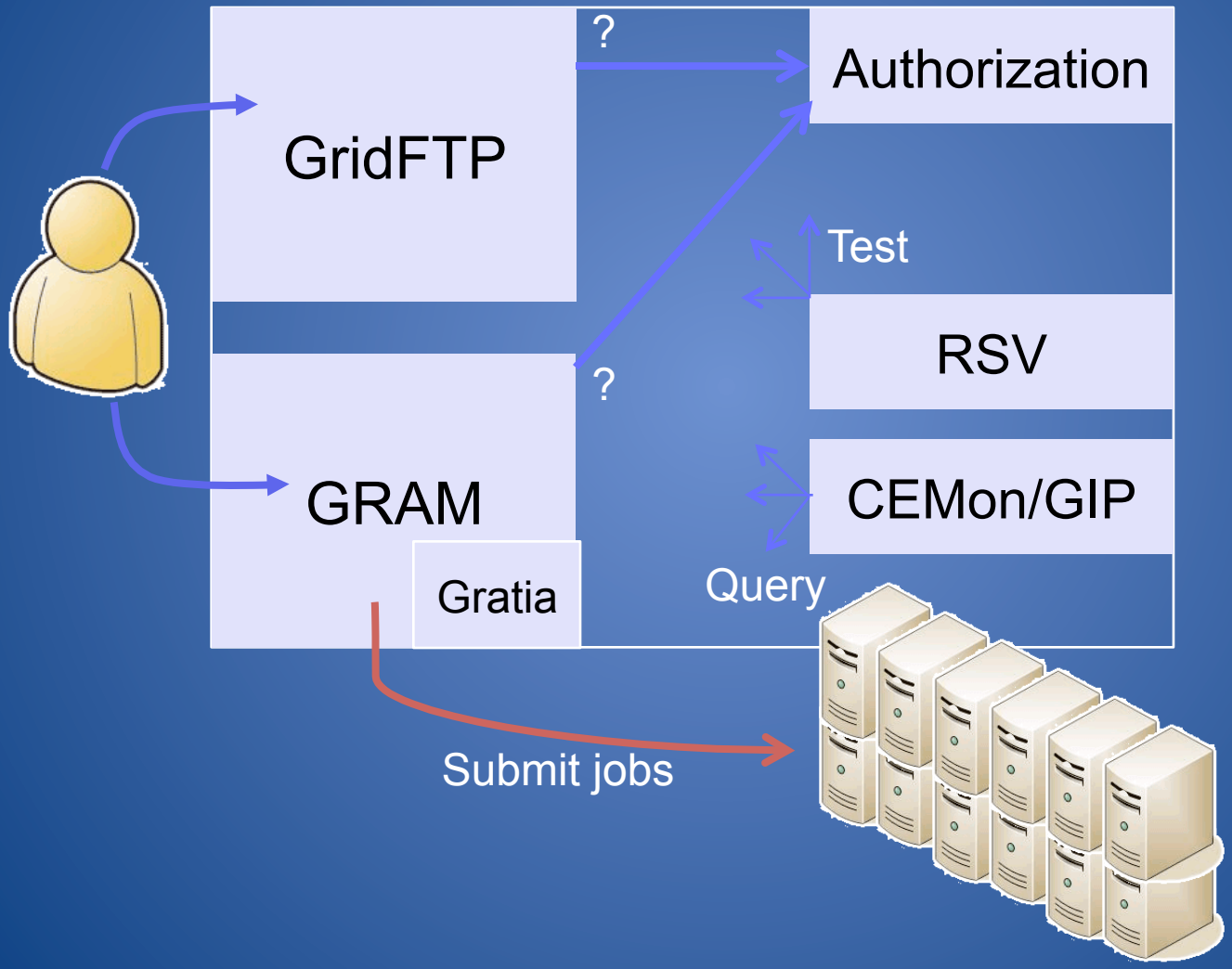
Overview of OSG components

- CE – Compute Element
  - Provides point of interface for tools attempting to run jobs or work on a cluster
  - Users submit jobs to this system
  - OSG provides a package that installs all software needed for this component
- SE – Storage Element
  - Several implementations
    - dCache
    - Bestman
  - Manages data and storage services on cluster
- WN – Worker Node
  - Software found on each compute node on grid
  - Provides software that incoming jobs may depend on (e.g. curl, srmcp, gsiftp, etc.)
- Client – Client Software
  - Provides software that users can use to submit and manage jobs and data on OSG
  - May be superseded by VO specific software
- Other tools (more specific and not necessarily used by many people)

5000 meter overview of CE

- GRAM : Allows job submissions and passes them on to local batch manager
- Gridftp : Provides data transfer services into and out of cluster
- CEMon / GIP : Provides information to central services
- Gratia : Sends accounting information on jobs run to central server
- RSV : Provides probes to monitor health of the CE
- User authorization : Needed to connect certificates to user accounts

Basic CE


?
Authorization
GridFTP
Test
RSV
?
GRAM
CEMon/GIP
Gratia
Query
Submit jobs

GRAM

- Two different flavors
  - OSG provides and supports both
  - Very different implementations
- GT2
  - What most users and VOs use
  - Very stable and well understood
  - On the other hand, fairly old
- GT4 (aka ws-gram)
  - Web services enabled job submission
  - Currently in transition
  - Used primarily by LIGO

Gratia

- Collects information about what jobs have run on your site and by whom
- Hooks into GRAM and/or job manager
- Cron job also present
- Sends information to a central server
- Can connect and query central service to get reports and graphs
- Option exists for a local server

CEMon / GIP

- These work together
  - Essential for accurate information about your site
  - End-users see this information
- Generic Information Provider (GIP)
  - Scripts to scrape information about your site
  - Some information is dynamic (queue length)
  - Some is static (site name)
- CEMon
  - Reports information to OSG GOC's BDII
  - Reports to OSG Resource Selector (ReSS)

RSV

- System to run tests on various components of your site
- Presents a web page with red/green overview and links to more specific information on test results
- Optional interface to nagios
- Can be run on a server other than CE

Site planning

- Bureaucratic details
- Cluster layout
- Disk layout / sharing
- Authorization

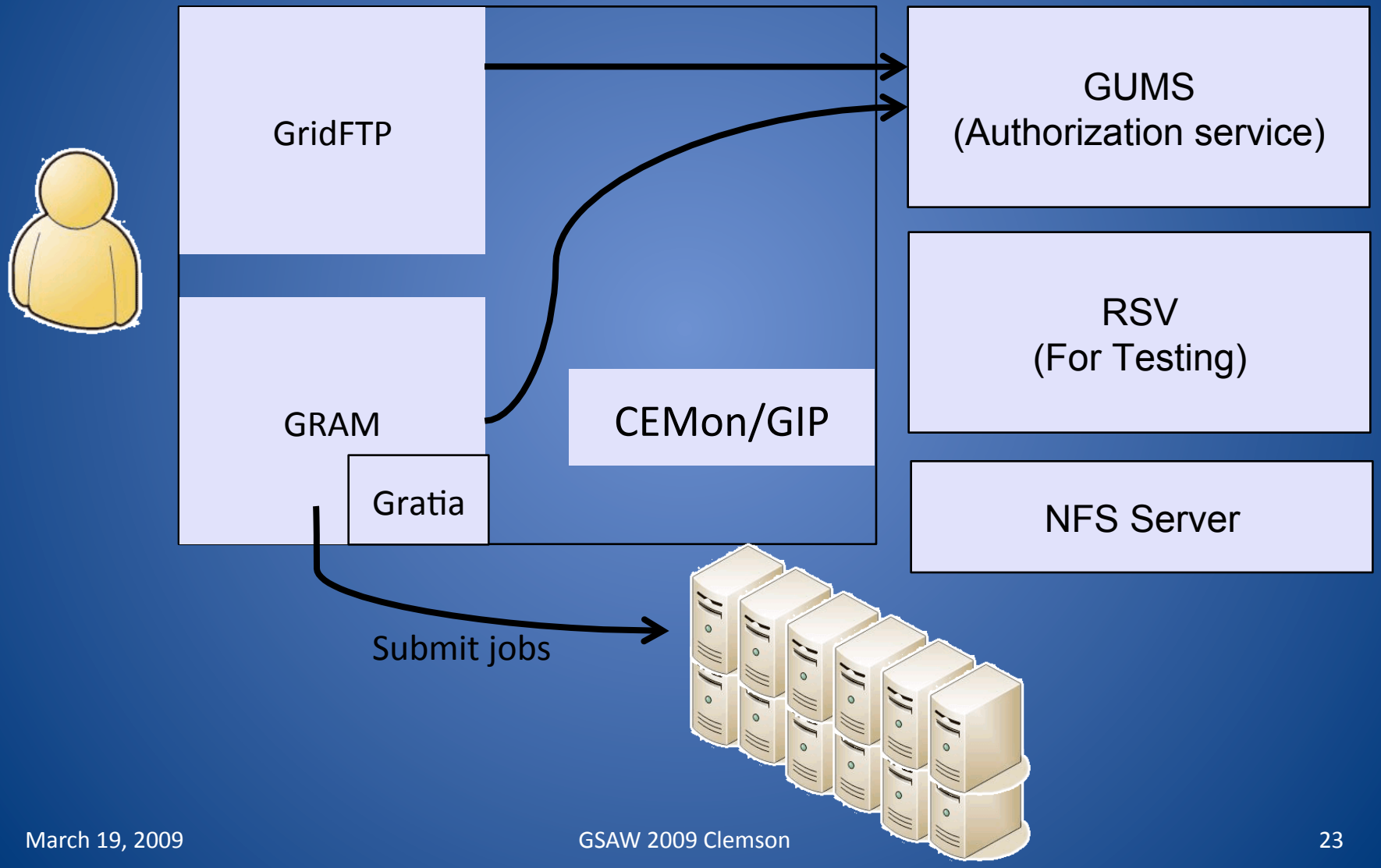
Bureaucracy

- Certificates (personal/host)
- VO registrations
- Registration with OSG
  - Need a site name (e.g. UC_ITB)
  - Need contacts (security, admin, etc.)
- Site policy on web

Site planning

- How is software / data being shared
  - NFS can work but gets bogged down with larger workloads
  - Where do services run?
    - Single server vs. dedicated servers
  - Worker node software?
    - Locally present on worker nodes vs. served over nfs
  - Certificates shared?

More complicated setup



Required Directories for CE / Cluster

- **OSG_APP**: Store VO applications
  - Must be shared (usually NFS)
  - Must be writeable from CE, readable from WN
  - Must be usable by whole cluster
- **OSG_GRID**: Stores WN client software
  - May be shared or installed on each WN
  - May be read-only (no need for users to write)
  - Has a copy of CA Certs & CRLs, which must be up to date
- **OSG_WN_TMP:** temporary directory on worker node
  - May be static or dynamic
  - Must exist at start of job
  - Not guaranteed to be cleaned by batch system

Optional directories for CE

- **OSG_DATA**: Data shared between jobs
  - Must be writable from the worker nodes
  - Potentially massive performance requirements
  - Cluster file system can mitigate limitations with this file system
  - Performance & support varies widely among sites
  - 0177 permission on OSG_DATA (like /tmp)
- **Squid server**: HTTP proxy can assist many VOs and sites in reducing load
  - Reduces VO web server load
  - Efficient and reliable for site
  - Fairly low maintenance
  - Can help with CRL maintenance on worker nodes

Space Requirements

- Varies between VOs
  - Some VOs download all data & code per job (may be Squid assisted), and return data to VO per job.
  - Other VOs use hybrids of OSG_APP and/or OSG_DATA
- OSG_APP used by several VOs, not all.
  - 1 TB storage is reasonable
  - Serve from separate computer so heavy use won't affect other site services.
- OSG_DATA sees moderate usage.
  - 1 TB storage is reasonable
  - Serve it from separate computer so heavy use of OSG_DATA doesn't affect other site services.
- OSG_WN_TMP is not well managed by VOs and you should be aware of it.
  - ~100GB total local WN space
  - ~10GB per job slot.

Worker Node Storage

- Provide about 12GB per job slot
- Therefore 100GB for quad core 2 socket machine
- Not data critical, so can use RAID 0 or similar for good performance

Authorization

- Two major setups:
  - Gridmap setup
    - File with list of mappings between DN and local account
    - Can be generated by edg-mkgridmap script
    - Doesn't handle users in mulitple VOs or with VOMS roles
  - Service with list of mappings (GUMS)
    - A little more complicated to setup
    - Centralizes mappings for entire site in single location
    - Handles complex cases better (e.g. blacklisting, roles, multiple VO membership)
    - Preferred for sites with more complex requirements
    - Ideally on dedicated system (can be VM)
    - Can add SAZ service for authorization

CE Installation Overview

- Prerequistes
  - Certificates
  - Users
- Installation
  - Pacman
- Configuration
- Getting things started

PKI Certificates

- Used for all authentication
- Your site needs PKI certificates
  - I assume you understand basics
    - You need a public cert
    - You need a private key
    - Often referred to informally, incorrectly as “certificate”
- Your site needs two certificates
  - Host certificate
  - HTTP certificate
  - Best to get these in advance
  - Optionally you may need RSV certificate
- Online documentation on getting them:
- https://twiki.grid.iu.edu/bin/view/ReleaseDocumentation/GetGridCertificates

Local accounts

- You need following local accounts
- User for RSV
- Daemon account used by most of vdt
- Globus user is optional but will be used if found

Pacman

- The OSG Software stack is installed with Pacman
  - Yes, custom installation software
- Why?
  - Mostly historical reasons
  - Makes multiple installations and non-root installations easy
- Why not?
  - It's different from what you're used to
  - It sometimes breaks in strange ways
  - Updates can be difficult
- Will we always use Pacman?
  - Maybe
  - Investigating alternatives but changing existing infrastructure is hard
  - Work ongoing to support RPM/deb in the future

Pacman (part deux)

- Easy installation
  - Download pacman
  - Untar and source shell script
  - Start using
  - Look ma! No root!
- Gotcha:
  - Installs into current directory

Documentation

- Twiki
  - OSG collaborative documentation
  - Used throughout OSG

  `https://twiki.grid.iu.edu/twiki/bin/view/`
- Installation documentation

  `https://twiki.grid.iu.edu/twiki/bin/view/ReleaseDocumentation/`

Basic installation and configuration

- Install Pacman
  - Download
    `http://physics.bu.edu/pacman/sample_cache/tarballs/pacman-3.26.tar.gz`
  - Untar (keep in own directory)
  - Source setup
- Make OSG directory
  - Example: /opt/osg symlink to /opt/osg-1.0
- Run pacman commands
  - Get CE
  - Get job manager interface
- Configure
  - Edit config.ini
  - Run configure_osg.py
- Start services

CA Certificates

- What are they?
  - Public certificate for certificate authorities
  - Used to verify authenticity of user certificates
- Why do you care?
  - If you don't have them, users can't access your site

Installing CA Certificates

- The OSG installation will **not** install CA certificates by default
  - Users will not be able to access your site!
- To install CA certificates
  - Edit a configuration file to select what CA distribution you want

```
vdt-update-certs.conf
```

  - Run a script

```
vdt-setup-ca-certificates
```

Choices for CA certificates

- You have two choices:
  - Recommended: OSG CA distribution
    - IGTF + TeraGrid-only
  - Optional: VDT CA distribution
    - IGTF only (Eventually)
    - Same as OSG CA (Today)
- IGTF: Policy organization that makes sure that CAs are trustworthy
- You can make your own CA distribution
- You can add or remove CAs

Why all this effort for CAs?

- Certificate authentication is the first hurdle for a user to jump through
- Do you trust all CAs to certify users?
  - Does your site have a policy about user access?
  - Do you only trust US CAs? European CAs?
  - Do you trust the IGTF-accredited Iranian CA?
    - Does the head of your institution?

Updating CAs

- CAs are regularly updated
  - New CAs added
  - Old CAs removed
  - Tweaks to existing CAs
- If you don't keep up to date:
  - May be unable to authenticate some user
  - May incorrectly accept some users
- Easy to keep up to date
  - vdt-update-certs
    - Runs once a day, gets latest CA certs

CA Certificate RPM

- There is an alternative for CA Certificate installation: RPM
  - We have an RPM for each CA cert distribution
  - No deb package yet
  - Install and keep up to date with yum
  - Some details not discussed here: read the docs

Certificate Revocation Lists (CRLs)

- It's not enough to have the CAs
- CAs publish CRLs: lists of certificates that have been revoked
  - Sometimes revoked for administrative reasons
  - Sometimes revoked for security reasons
- You really want up to date CRLs
- CE provides periodic update of CRLs
  - Program called fetch-cr
  - Runs once a day (today)
  - Will run four times a day (soon)

Updates

- We periodically release updates to OSG software stack
- Announced by VDT team on vdt-discuss mailing list
  - Not OSG-specific announcement or update procedure
- Announced by GOC
  - OSG-specific instructions

Two kinds of updates

- Incremental updates
- Major updates

Incremental Updates

- Frequent (Every 1-4 weeks)
- Can be done within a single installation
- Process:
  - Turn off services
  - Backup installation directory
  - Perform update
  - Re-enable services

Major Updates

- Irregular (Every 6-12 months)
- Must be a new installation
- Can copy configuration from old installation
- Process:
  - Point to old install
  - Perform new install
  - Turn off old services
  - Turn on new services

Incremental updates are a mess

- If you apply each update as it's available, it's not *too* bad.
- VDT supplies instructions for each update
  - Sometimes there are picky details
  - It's unclear how to do multiple updates at once
    - What order should steps be done in?
    - What are all the appropriate picky details?
- New updater script being released

The grand future of updates

- VDT team is working on improved mechanism for doing updates
  - Fully scripted to get picky details right
  - Hopefully easy to keep right, even with multiple updates
  - Still in progress
  - Release date – end of this month

Discussion, Questions

- Questions? Thoughts? Comments?

Acknowledgements

- Alain Roy
- Terrence Martin