



A Canada-Wide Computational and Storage Grid

Copy of this talk:	http://grid.phys.uvic.ca/index.php?content=presentations
Victoria Grid Home Page:	http://grid.phys.uvic.ca/
Installation instructions	http://grid.phys.uvic.ca/gcprod/requirement.php
Grid Canada:	http://www.gridcanada.ca/

Grid Canada



-
- Grid Canada originated with CANARIE, NRC and C3
 - C3 not really active (HPC group)
 - Victoria, Alberta, Pacific Forestry are other interested groups
 - WestGrid uses GC Certificates
 - CANARIE wants GC to succeed
 - Darcy Quesnel is the “Grid Master”
 - Helps fund high speed network tests (Carleton, TRIUMF, Alberta)
 - Bill St Arnaud looking to firm up Grid Canada but for many fields
 - Value to ATLAS?
 - Bill is a strong supporter of ATLAS and has good connections to HEP
 - HEP is the only user of the computational grid (Babar and ATLAS)
-

Outline



- GC Certificate Authority
 - GC Computational Grid
 - Site requirements
 - Grid Information Service
 - User access
 - Job Submission
 - Grid/Job Monitoring
 - Grid Manager and Resource Broker
-

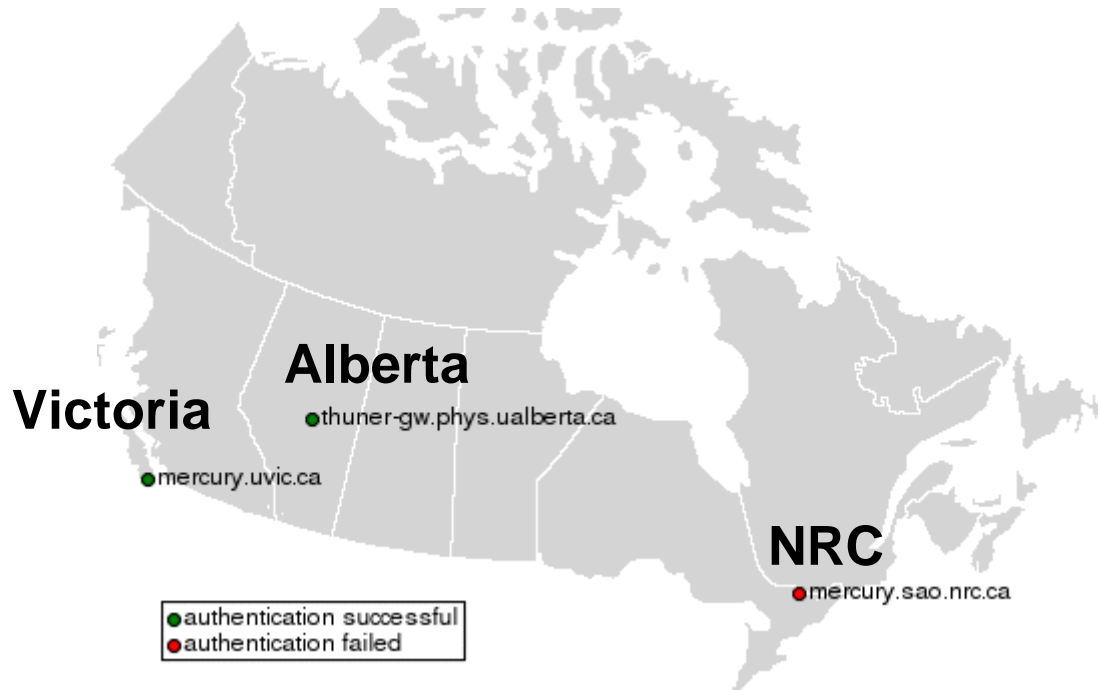
GC Certificate Authority



Darcy Quesnel, CANARIE (Grid Master)

- GC must abide by international CA agreements (EU DataGrid, GGF)
 - GC Certificates are internationally recognized
 - Each resource has a Host Certificate
 - Each users has a User Certificate
 - An organization/person must vouch for the user
 - Users create a time-limited User Proxy which is passed to a resource
 - A Certificate does not allow users to access any resource
 - User must have permission to use the resource
 - GC Computational Grid has a Gridmap File containing the User Certificates of authorized users
-

GC Computational Grid



Victoria ~168 CPUs
Alberta ~120 CPUs
NRC ~50 CPUs

Intel CPUs with Linux OS
Globus Toolkit V2.4
OpenPBS/Maui scheduler

No root access at UVic and NRC

Victoria
Grid Manager/Broker

CANARIE
Grid Information Server

Site Requirements I



Goal is to make minimal hardware and software requirements

- Globus
 - Batch scheduler (PBS, Condor)
 - Ganglia
 - “gcprod00, gcprod01, ...” accounts (single group)
 - No user login privileges
 - Worker nodes access a common file system (NFS)
 - “gcprod” batch queue
 - Root access is not required
-

Site Requirements II



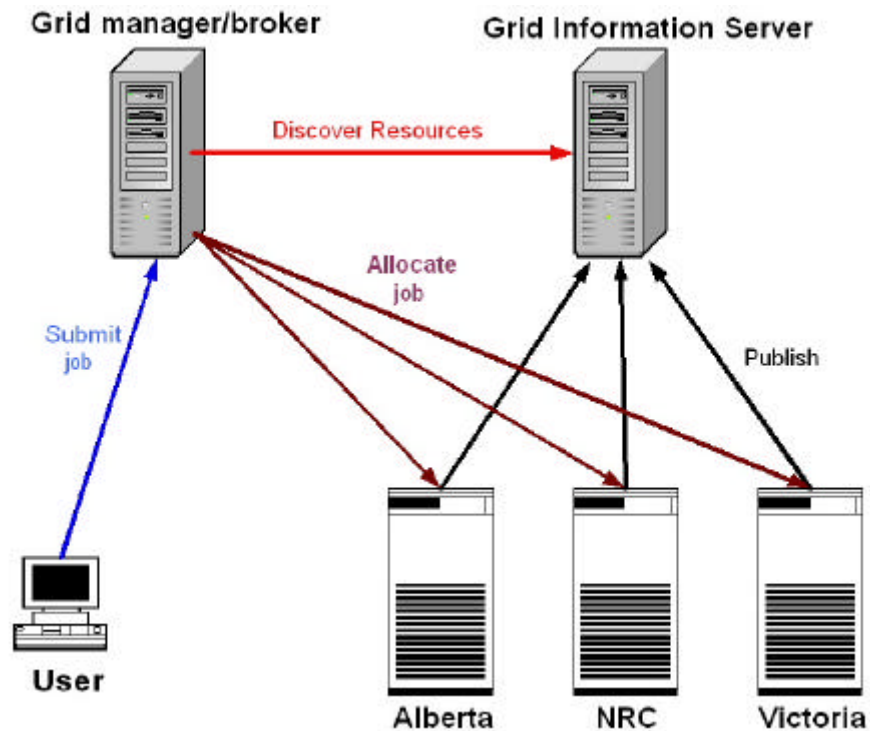
The site must:

- Have a GC Host Certificate
- Publish to the GC GIIS
- Install the GC grid map file
- Provide ~100 GB disk space

ATLAS Software requirements:

- Need to install MySQL client
 - We have a grid-enabled installation procedure based on the ATLAS rpm installation used in DC1
 - This is OS dependent
-

Grid Overview

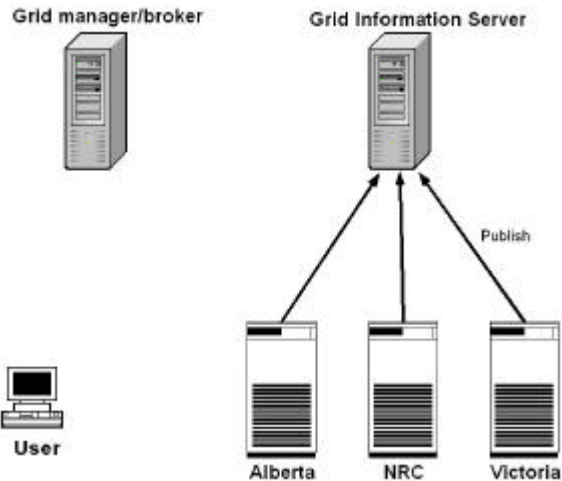


Standard Model of Grids

Services

- information services
- user access
- resource broker
- monitoring

Grid Information Service



Resources publish information on system information and load to the GC GIIS
<http://giis.gridcanada.ca>

A resource broker can query the GIIS to find available resources

The screenshot shows the LDAP Browser/Editor v2.8.2 interface. The title bar indicates the connection to `ldap://giis.gridcanada.ca:2135/Mds-Vo-name=gc-production, o=grid`. The left pane shows a tree view of LDAP entries, with the selected entry being `Mds-Host-hn=mercury.sao.nrc.ca`. The right pane displays a table of attributes and their values for this entry.

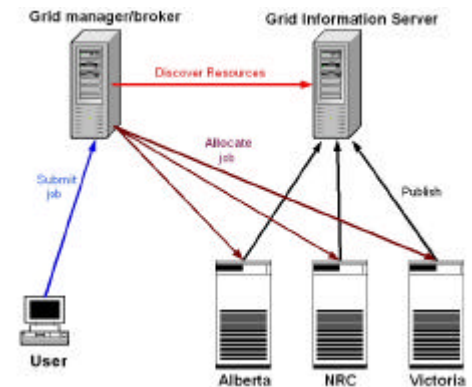
Attribute	Value
Mds-Cpu-speedMHz	1995
Mds-Memory-Ram-Total-freeMB	1570
Mds-Fs-freeMB	1009
Mds-Fs-freeMB	10207
Mds-Fs-freeMB	18383
Mds-Fs-freeMB	7158
Mds-Fs-freeMB	78
Mds-Fs-freeMB	9275
Mds-Fs-freeMB	9286
Mds-Cpu-Free-5minX100	198
Mds-Net-Total-count	3
Mds-validfrom	20031014234930Z
Mds-Cpu-Total-count	2
Mds-Memory-Vm-sizeMB	4095
Mds-Cpu-vendor	GenuineIntel
Mds-Net-name	eth0
Mds-Net-name	eth1
Mds-Net-name	lo
Mds-validto	20031014234930Z

User access



- Generic accounts have been set up at each site
 - gcprod01, gcprod02, gcprod03, ... gcprod20
- Users create a Grid Proxy based on their User Certificate
- Users are mapped on a 1:1 basis into one of these accounts
 - Gridmap File contains a copy of the User Certificate
 - Identical at all 3 sites

- Users have no log in privileges
 - Software and data transfer via grid copy
 - Software compilation/linking done in batch
 - Job execution via batch scheduler



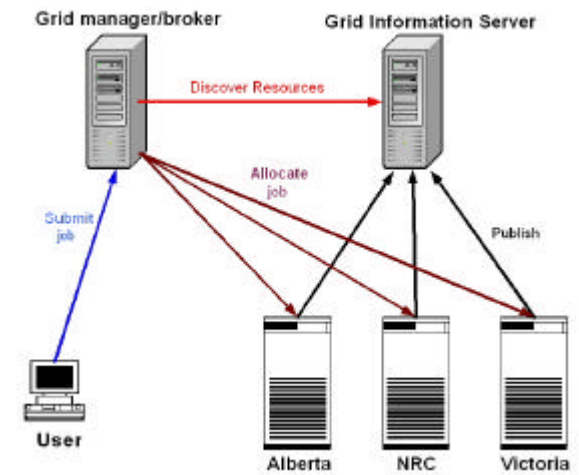
Job Submission



Globus Toolkit provides the means to discover resources and run applications on those resources.

It does not provide utilities to determine where the job will run

- ***grid-proxy-init***
 - create Grid Proxy
- ***gcsub <parameters>***
 - job is sent to Grid Manager/Broker
- **We have 3 methods for job submission**
 - gcsub (like NorduGrid's *ngsub*)
 - gcsub2 (sends request to a RB)
 - CondorG



Grid Monitoring

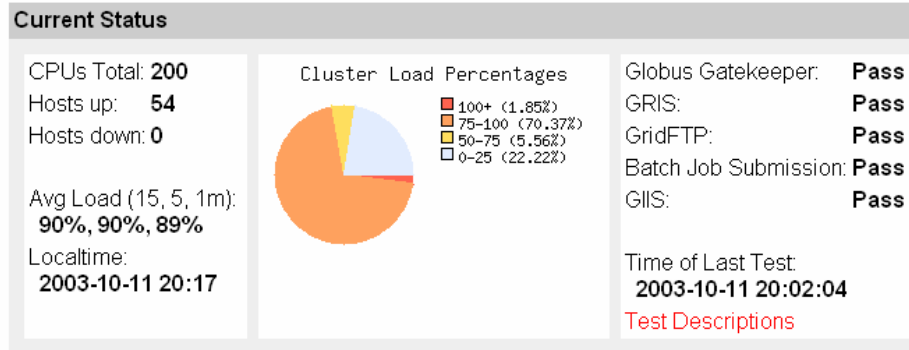


Ganglia is used to monitor the status and load of the resource

mercury.uvic.ca

ALBERTA • CANARIE • NRC • VICTORIA • PFC

Located at the University of Victoria in Victoria, BC



System properties

Status of site

Current load on the system

Job Monitoring



Web-based monitor of job status

Grid Status

ALBERTA • CANARIE • NRC • VICTORIA • PFC

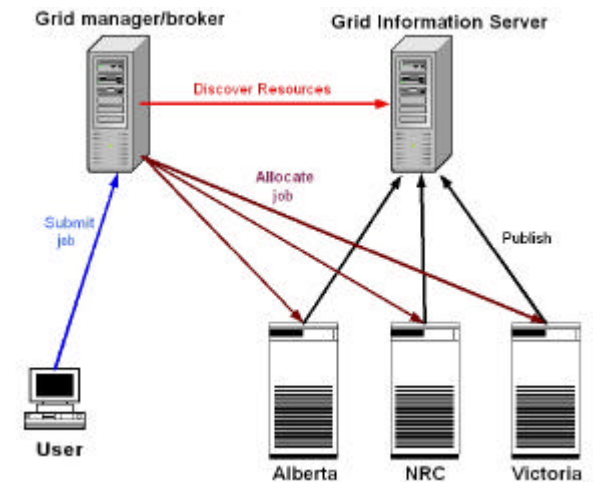
Status of Globus Job Queue

User	Filename	Arguments	Staged	Time Submitted	Status	Time Started	Resource	Globus Job ID
lmk	/usr/bin/whoami		no	2003-10-09 10:41:36	pending	2003-10-09 10:43:11	mercury.uvic.ca	https://mercury.uvic.ca:40001/9210/1065721366/
lmk	/usr/bin/whoami		no	2003-10-09 10:19:42	done	2003-10-09 10:20:22	mercury.uvic.ca	https://mercury.uvic.ca:40001/3066/1065719996/
dvanders	/homes/dvanders/test.sh		yes	2003-10-09 10:19:15	done	2003-10-09 10:20:13	mercury.sao.nrc.ca	https://mercury.sao.nrc.ca:17279/12893/1065720011/
dvanders	/homes/dvanders/test.sh		yes	2003-10-09 10:14:20	done	2003-10-09 10:16:17	mercury.sao.nrc.ca	https://mercury.sao.nrc.ca:17247/11864/1065719776/

Grid Manager/Broker



- Reviewing our current model for the Grid Manager/Broker
 - Some grids avoid the use of a central Grid Manager and put the intelligence into the client application (such as *gcs*)
 - We see advantages in a Grid Manager for job monitoring, error recovery and job management
- Next step is to utilize the GIIIS information
 - Workqueue (WQ) scheduler
 - Send jobs to free resources that meets specs
 - CPU, memory, disk space
- Looking at more sophisticated brokers
 - Optimize jobs on faster resources
 - Jobs on slow resources are replicated



Work in Progress



- GC Grid is a work in progress
 - It will work today for ATLAS DC with expert control (ie the same way we run BaBar MC Simulation at UVic)
 - JWhite: only 20% of ATLAS DC jobs are success on NorduGrid
- Lots of room for improvements
 - Monitoring
 - data management
 - resource broker
 - interface to TRIUMF/LCG
- GC is primarily a deployment rather than development project but we have targeted areas of development where we can contribute
 - We are doing work on Resource Brokers (collaborating with Engineering)
 - We want to develop the tools so that the Grid can utilize CANARIE UDLP

Plans



-
- Bring on new sites
 - Start with TRIUMF
 - Prefer to have at least 5 nodes (head node and worker nodes)
 - Continue running ATLAS DC1 application
 - ATLAS DC Interface
 - Grid sphere, NorduGrid interface
 - Monitoring
 - Data management (many choices)
 - How to access data on worker nodes
-