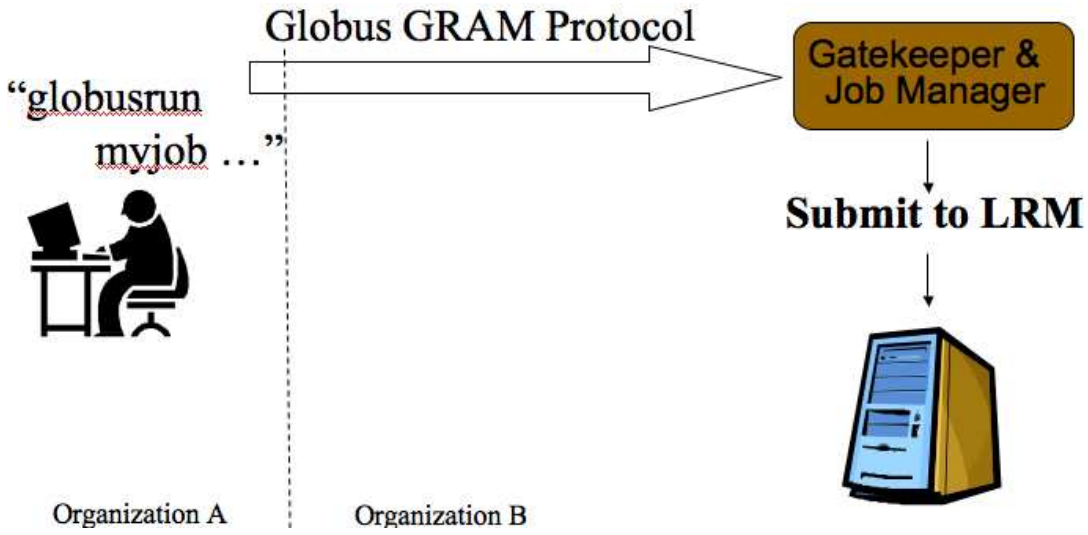
Job and Resource Management[1]

Local Resource Manager (LRM)

- A batch scheduler to run jobs on a computing cluster
- Popular LRMs
  - PBS – Portable Batch System
  - LSF – Load Sharing Facility
  - SGE – Sun Grid Engine
  - Condor
    - Originally designed as a cycle scavenger
    - Evolved into a comprehensive system to manage computing
- LRM executes on a cluster's head node
- Simple LRM
  - Allows a user to fork jobs quickly
  - Runs on the head node (gatekeeper) for fast utility functions
  - No queuing
- In GRAM, each LRM is handled with a *job manager*

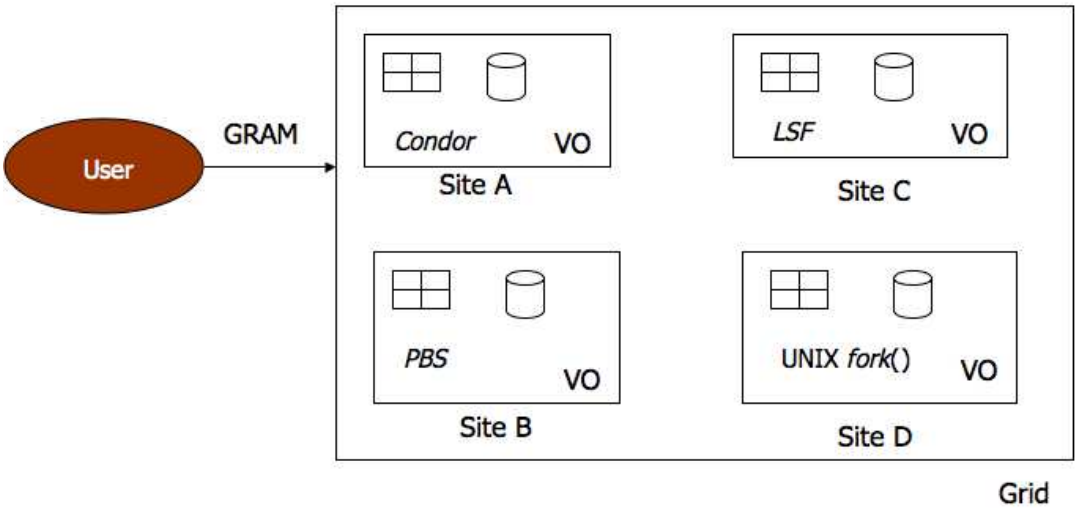
GRAM

- Globus Resource Allocation Manager

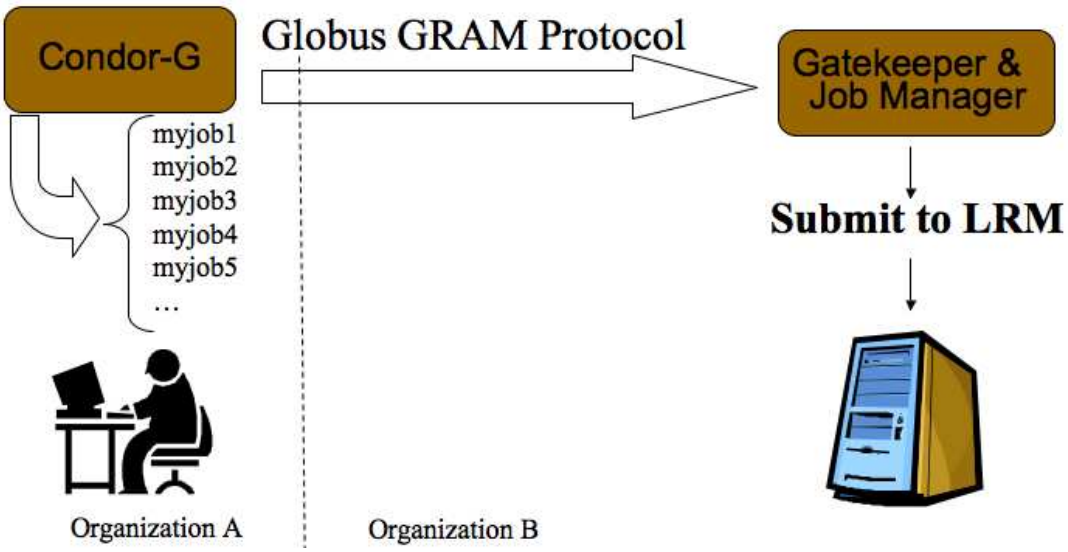


- Provides a uniform interface to diverse cluster schedulers

[1] Most of the material in this set of notes is from the Educational division of Open Science Grid.



- Jobs submitted by using Condor-G
  - Grid job submission manager



Data Management

- Services to provide mechanisms to find, move, and share data
  - GridFTP
    - Fast, flexible, secure, ubiquitous data transport
    - Often embedded in higher level services
  - RFT
    - Reliable file transfer service
    - Uses GridFTP
  - Replica location service
    - Tracks multiple copies of data for speed and reliability

  - Storage resource manager
    - Manages storage space allocation, aggregation, and transfer
  - Metadata management services

- GridFTP
  - Optimized for high-bandwidth wide-area networks
  - Based on standard FTP
    - Supports reliable and restartable data transfers
    - Provides extensions for high performance operation and security
    - Not software but specification, with standards defined within the Open Grid Forum
    - Reference implementation provided by Globus Alliance
  - Performance improvement
    - Supports parallel TCP streams over high-speed WAN links
    - Allows users to set optimal TCP buffer size for a transfer
  - Coordinated data transfer
    - Supported over multiple nodes, both source and destination
  - Security through GSI (Grid Security Infrastructure)
    - Also supports `ssh` as an alternative to GSI
    - Authentication and authorization
    - Can also provide encryption
    - Pipelining feature to allow clients to send requests while previous requests are still completing
      - Reduces overhead
      - Speeds up transfer of many small files
  - Multicasting/Striping
    - Provides for one-to-many transfers
  - Overlay routing
    - Servers can be used as intermediate nodes to route data
  - Client tools
    - `globus-url-copy`
    - `uberftp`
    - Custom clients

- Replicating data files for faster access
  - Effective use of grid resources – more parallelism
  - Each *logical* file can have multiple *physical* copies
  - Avoids single point of failure
  - Manual or automatic replication
    - Automatic replication considers factors such as demand for a file and transfer bandwidth

- File catalogs
  - Provides the means to locate data
  - File catalog services
    - Replica location service (RLS)
    - Phedex

    - RefDBPupDB
  - Requirements
    - Abstract the logical file name (LFN) for a physical file
    - Maintain the mappings between LFN and PFN (physical file name)
    - Maintain the location information of a file

Grid Security

- Crucial component
  - Problems being solved might be sensitive
  - Resources are typically valuable
  - Resources located in distinct administrative zones
    - Different policies, procedures, security mechanisms
  - Implementation must be broadly available and applicable
    - Standard, well-tested, and well-understood protocols
    - Protocols integrated with a wide variety of tools
- Grid security infrastructure (GSI)
  - Provides secure communications for all the higher level grid services
  - Secure *authentication* and *authorization*
    - Authentication
      - Ensures you *are* who you claim to be
      - ID card, fingerprint, passport, username/password
    - Authorization
      - Controls what you are permitted to *do*
      - Run a job, read or write a file
  - GSI provides Uniform Credentials
  - Single Sign-On
    - User authenticates once; then can perform many tasks

National Grid Infrastructure

- Open Science Grid
  - Provides shared compyting resources, benefiting a broad set of disciplines
  - Consortium of universities and national laboratories, building a sustainable grid infrastructure for science
  - Incorporates advanced networking and focuses on general services, operations, and end-to-end performance
  - Composed of a large number (> 50) of shared computing facilities, or “sites”, with more than 15.000 CPUs
  - Can handle from 400 to more than 1000 concurrent jobs
  - Supports many aplications and CS experiments, including long-running production operations
  - Up since 2003; few FTEs at central ops
- TeraGrid
  - Provides vast resources via a number of huge computing facilities

- Locating and monitoring resources
  - Required for efficient use of the grid
  - Check the availability of different grid sites
  - Discover different grid services
  - Check the status of "jobs"
  - Make better scheduling decisions with information maintained on the "health" of sites
  - Resource selection service in OSG provided by VORS – Virtual Organization Resource Selector

Grid Workflow

- A typical workflow pattern in image analysis runs many filtering applications
- Image classification operations include steps like raster pixel processing, segmentation, raster to object conversion, raster object operations, raster object vector conversion, and vector cleanup
- The processes can be divided among different sites for large images but need to be synchronized for different steps

Why grids?

- New approaches to inquiry based on
  - Deep analysis of huge quantities of data
  - Interdisciplinary collaboration
  - Large-scale simulation and analysis
  - Can explore a large parameter space at a finer granularity
  - Smart instrumentation
  - Dynamically assembling resources to tackle a new scale of problem
- Enabled by access to resources and services without regard for location and other barriers
- Grid as a platform
- Applications
  - Bandwidth provision and management
  - Service provision
  - Medicine and health services
  - Engineering
  - Oracle Grid Interfaces
- Science needs community
  - Teams organized around common goals
    - People, resources, software, data, instruments
  - With diverse membership and capabilities
    - Expertise in multiple areas required for many applications
  - And geographic and political distribution
    - No location/organization posesses all required skills and resources
  - Must adapt as a function of the situation
    - Adjust membership, reallocate responsibilities, renegotiate resources