



Earth Science Grid (ESG) prototype using Globus GSI-FTP

Reported by
Arie Shoshani

PPDG meeting
July 13-14, 2000

PPDG meeting, July 2000

GSI-FTP



- GSI-FTP is a generic term that refers to security enhances FTP clients and servers
- The GSI-FTP that are considered reliable
 - **gsi-nc-ftp** client
 - **gsi-wu-ftpd** server (developed at WU)
 - for access from disk
 - **gsi-pftpd** (developed at SDSC)
 - for access from HPSS

PPDG meeting, July 2000

Other Services Used



- **Globus Replica Catalog**
 - developed using LDAP
 - has concept of a logical file collection
 - registers logical file name by collection
 - uses URL format for location of replica
 - this includes host machine, (port), path, file_name
 - May contain other parameters, e.g. file size
 - provides hierarchical partitioning of a collection in the catalog (does not have to reflect physical organization at any site)
- **Network weather service (NWS)**
 - developed and installed by U of Tennessee
 - require installation at each participating host
 - provides pair-wise bandwidth estimates

PPDG meeting, July 2000

Setup of prototype



- **Goal: access files from multiple sites, selecting the "best" replica, in a secure manner**
- **participating institutions for file replicas**
 - SDSC: all the files for the demo on HPSS (about 60 files)
 - about 15 disjoint files on disk in each of 4 locations: ISI, ANL, NCAR, LBNL
 - some files are only on tape
 - size of files 100-500 MBs
- **The entire dataset stored at NERSC (LBNL)**
 - use HRM (using CORBA) to request staging of files to HRM's disk
 - use GSI-wuftp to transfer file after it is staged

PPDG meeting, July 2000

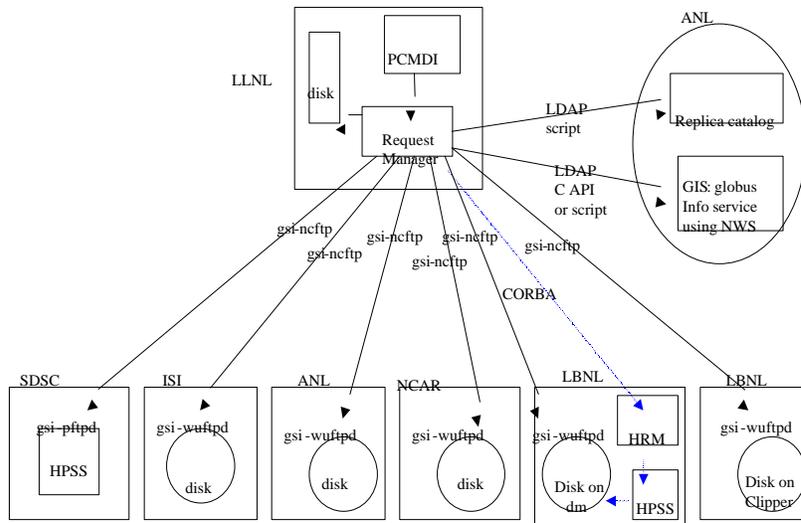
Setup of prototype



- **Client is PCMDI software**
 - it has its own “metadata” catalog
 - lookup in the catalog generates a set of files that are needed to satisfy a user’s request
- **Request Manager (ReqM) was newly developed software at LBNL**
 - accepts a request to cache a set of logical file names
 - checks for each replica location
 - gets for each replica location NWS bandwidth
 - selects “lowest” cost location
 - initiates transfer using GSI-FTP
 - monitors progress, responds to status command

PPDG meeting, July 2000

Inter-communication Diagram



PPDG meeting, July 2000

Status



- **Definitely a prototype**
- **Some of the software used first time (Request Manager, Replica Catalog)**
- **Test ran successfully**
- **Limited experience so far**
- **Only one client tested**
- **A lot of work to coordinate communication between modules**
- **Need a really fast network for demo**
- **Maybe will be demo'ed at SC2000**
- **Amazing that it all worked !**