Helping Scientists Concentrate on Science: Providing a Transparent View of Data on the Grid

Scientific Data Management Group, Lawrence Berkeley National Laboratory

Terascale computing (performing trillions of calculations per second) and large scientific experiments produce enormous quantities of data that require effective and efficient management, a task that can distract scientists from focusing on their core research. As part of the U.S. Department of Energy's Scientific Discovery through Advanced Computing program (SciDAC), researchers at Berkeley Lab and other institutions are establishing the Scientific Data Management Integrated Software Infrastructure Center. The goal of this project is to provide a coordinated framework for the unification, development, deployment, and reuse of scientific data management software. This demo demonstrates an example of middleware that was developed to aid applications that generate a large number of objects (such as High Energy Physics experiments) to manage their data transparently on the grid. Requests expressed by scientists in their own terms are supported transparently using tools that interpret the requests and gather the data from the data grid in the most efficient way possible.

PROBLEM STATEMENT

- Large number of objects reside in files on a distributed Data Grid
  - 10^8 – 10^9 objects
  - 0.5 – 5 million files
  - 15,000 – 150,000 tapes
- Distributed system can be across continents
  - 100's of sites
- Some of the data is replicated based on demand or pre-assigned replication
- Request expressed as logical request by user
- Systems and network may fail
- Problem:
  given a logical request, get relevant data to local system without human intervention

DEMO SETUP AND GOALS

- Storage servers are distributed in different physical locations
- 4 types of servers:
  - Server with HRM, connected to HPSS
  - Server with DRM
  - Server with GridFTP only
  - Server with FTP only
- All Files are on the HPSS system (30 files)
- Files are partially replicated on other servers (10 each)
- BitMap index on 2 million objects X 100 attributes
- Goals:
  - Demonstrate use of BitMap index
  - Demonstrate access from all 4 types of servers
  - Demonstrate file transfer monitoring

APPROACH

- Use BitMap index
  - For efficient processing of logical request
- Use Request Executer
  - For selecting replica and issuing multiple file request to Disk Resource Manager (DRM)
- Use local DRM
  - To communicate to other remote Storage Resource Managers
  - Initiate file transfers in parallel
- Use Hierarchical Resource Managers (HRM)
  - To schedule file staging from HPSS tape
- Use GridFTP to transfer files
- Use monitoring tool to track file transfer progress

MIDDLEWARE SOFTWARE SHOWN IN DEMO

1) BitMap index
   - in: logical request
     - `(73.39 < zdc2Energy < 94.94 AND -24.99 < qxb < -7.25)`
   - out: a set of logical files
     - `{uvcloud_csm_1982_1.nc,.., B04.16.atm.0079.nc}`
   - Size of data to be indexed:
     10^8 objects x 500 attributes x 4 bytes = 200 GB

2) Request Executer
   - in: a set of files
     - `{uvcloud_csm_1982_1.nc,.., B04.16.atm.0079.nc}`
   - out: selected URLs
     - `gsiftp://dg0n1.mcs.anl.gov/homes/asim/gsiftp/uvcloud_csm_1982_1.nc`
     - `hrm://DRMServerAlone@srm.lbl.gov:4000/home/dm/srm/data1/gsiftp/sc2001/B04.23.atm.0079.nc`
   - Uses Replica Catalog
   - Monitors transfer progress

3) Disk Resource Manager (DRM)
   - Manages space
   - Invokes HRM, DRM, GridFTP, FTP

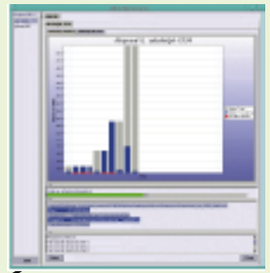
4) Hierarchical Resource Manager (HRM)
   - Manages space
   - Gets files from HPSS

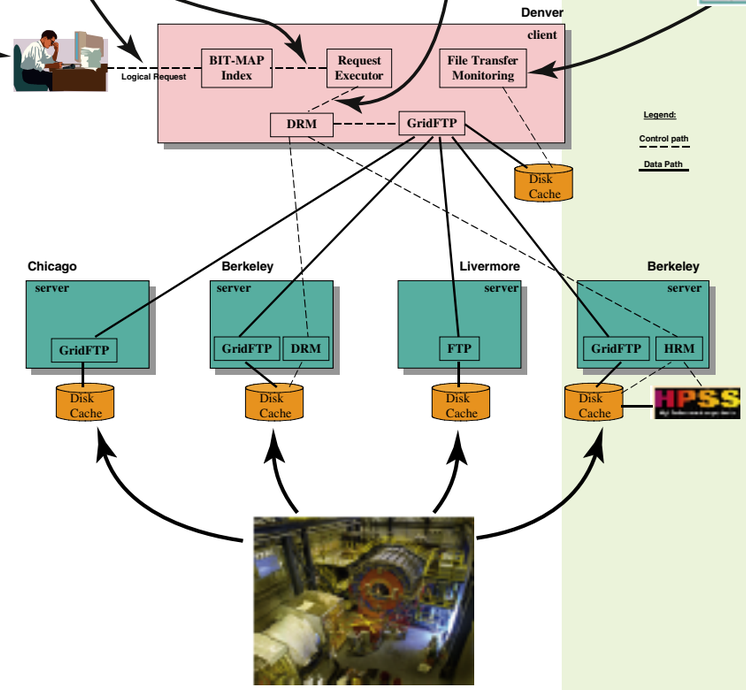
5) GridFTP
   - Globus product for secure & efficient transfer

6) File Transfer Visualization tool (FTV)
   - View by file size and fraction of file transferred
   - View by % of files transferred

Query Specification Tool

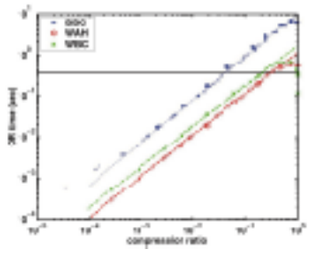
Monitoring File Transfer



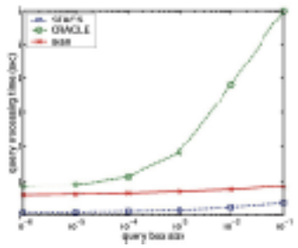


Particle Detection System Generates Terabytes of Data

Compressed BitMap Index performance compared to known methods



Compressed BitMap index performance compared to ORACLE



Arie Shoshani, Vijaya Natarajan, Junmin Gu, Alex Sim, John Wu
E-mail: AShoshani@lbl.gov

We acknowledge the help of the Globus team at ANL in providing and supporting the GridFTP and security infrastructure used in this demo.