

A Data Transfer Framework for Large-Scale Science Experiments

Raj Kettimuthu
Argonne National Laboratory and
The University of Chicago

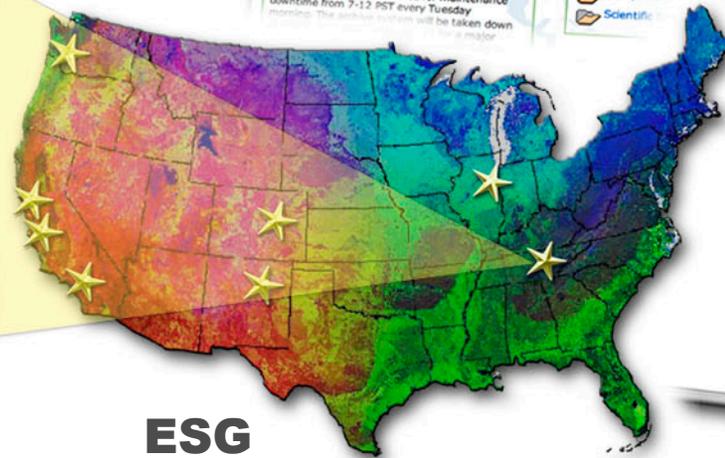


the globus alliance
www.globus.org

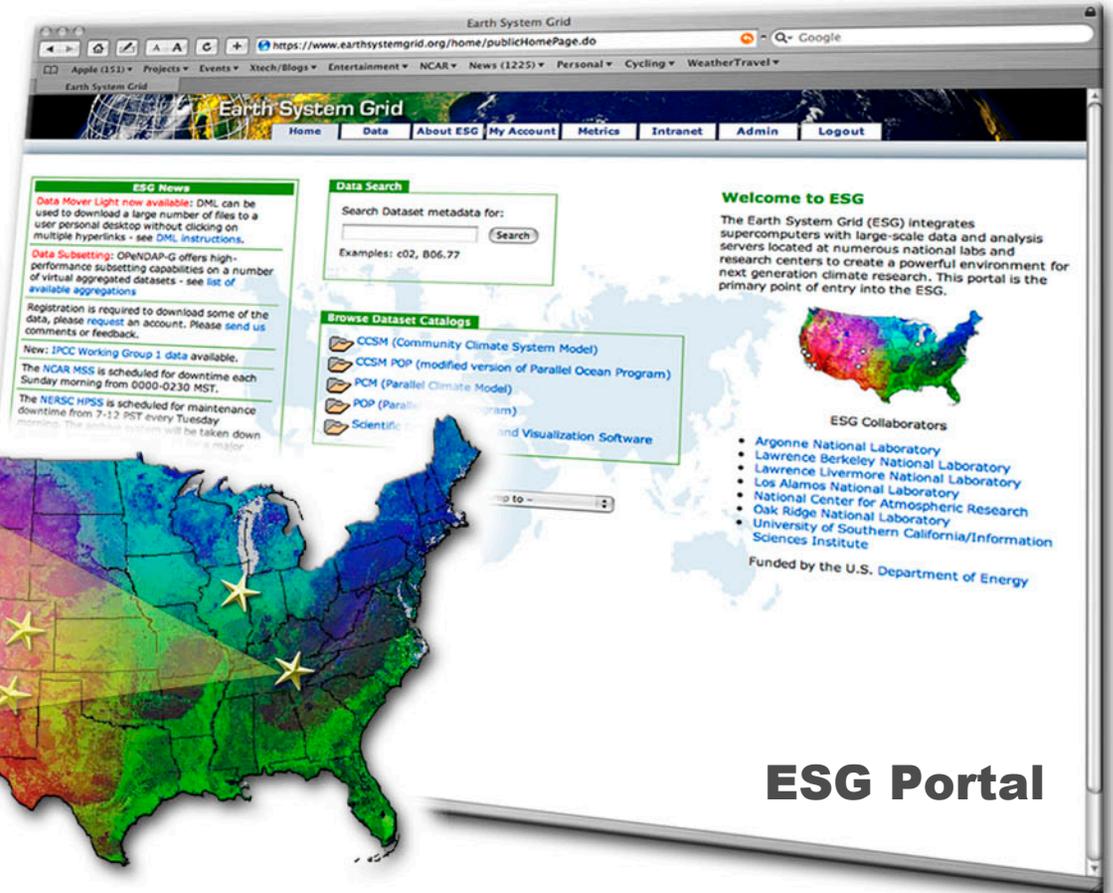
Earth System Grid

To support the infrastructural needs of the national and international climate community, ESG is providing crucial technology to securely access, monitor, catalog, transport, and distribute data in today's grid computing environment.

HPC
hardware running
climate models



ESG
Sites



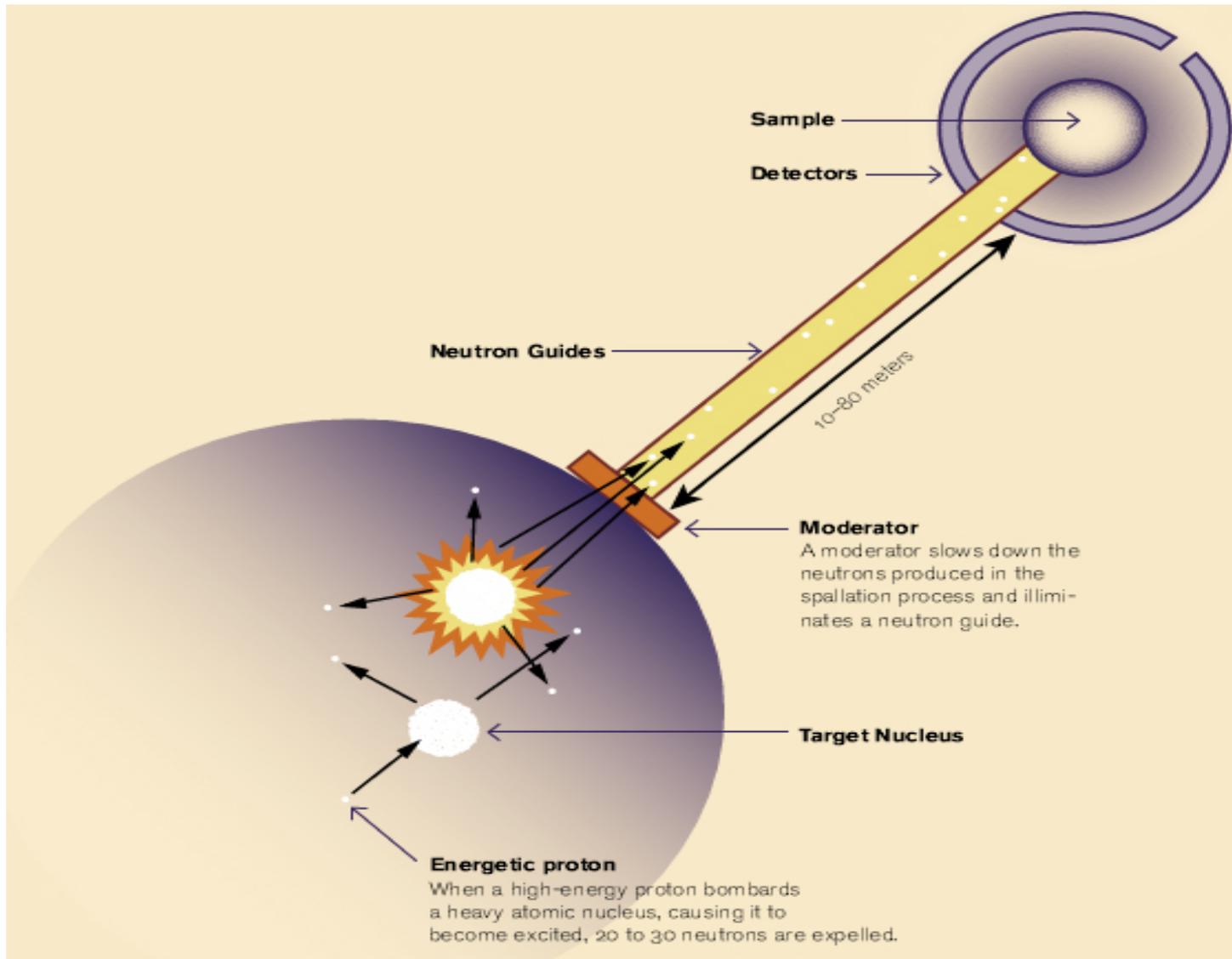
ESG Portal



the globus alliance

www.globus.org

Spallation Neutron Source





the globus alliance

www.globus.org

Data Movement

- Characteristics of scientific data
 - Typically stored in file
 - Modest file size
 - Huge number of files
 - A dataset is organized into hierarchical directories
- Requirements for data movement
 - Reliability
 - Performance
 - Scalability
 - Ease of use

Data Transfer Framework

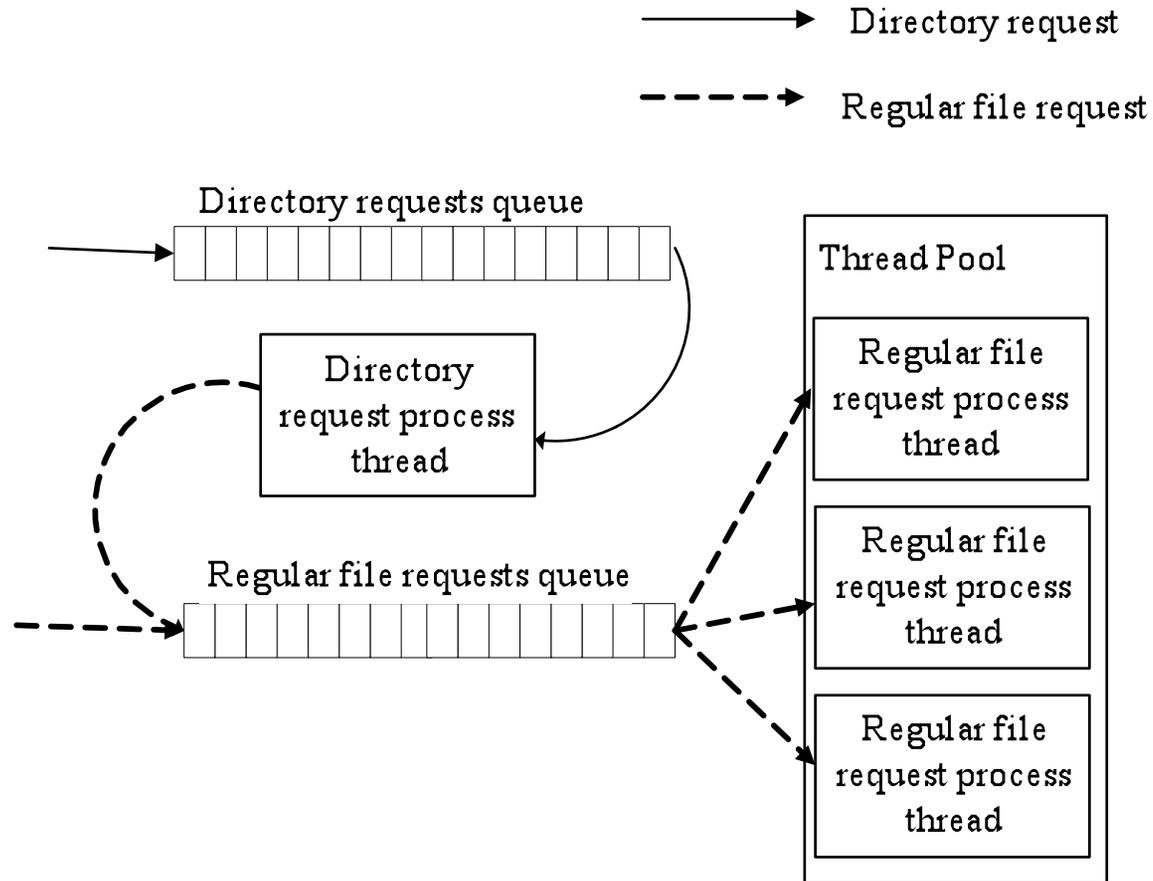
GridFTP GUI

Data Scheduler

Data Transfer Library

CoG jglobus (including GridFTP client)

Data Transfer Library





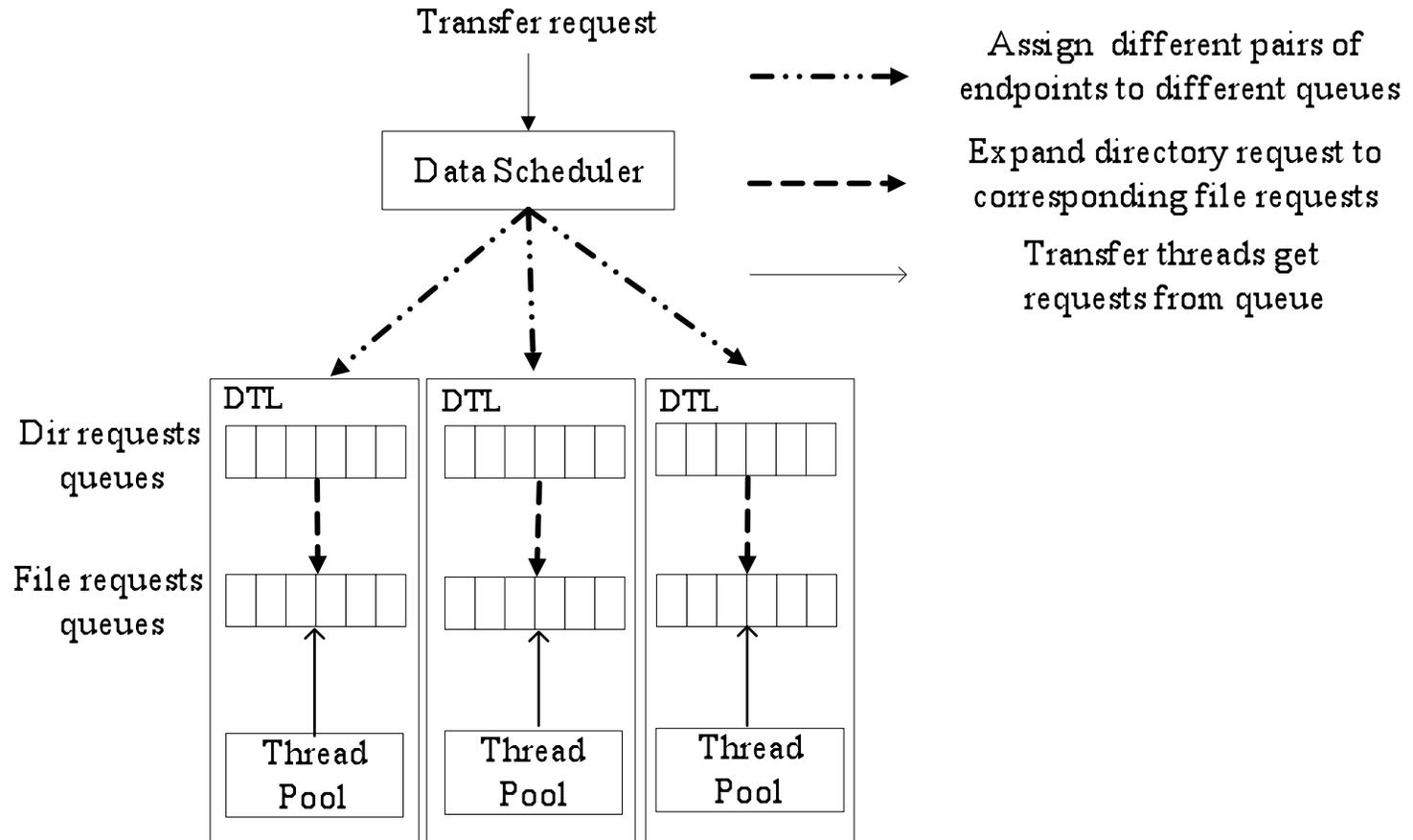
the globus alliance

www.globus.org

Data Transfer Library

- **Fault detection and recovery**
 - Network outage, server crash
 - > Client (DTL) still active
 - > Retries
 - Client failures
 - > Maintain transfer request checkpoint
 - > Recover when program starts again
- **Support multiple data transfer protocols**
 - Provides interface for integrating new protocol
 - The prototype is based on GridFTP

Data Scheduler





Data Scheduler

- FCFS
- Dependency-aware scheduling
 - User specify dependency between requests
 - Scheduler parse the dependency relationship and enqueue requests accordingly
- Multiple-Pair Transfer Scheduling
 - Different transfer requests may have different source and destination
 - Process them concurrently



GridFTP GUI

- **Java web start application**
 - No installation
 - Always up-to-date version
- **Drag and drop**
 - Zero configuration
 - Integrated with myproxy
 - Automatically trusts the CAs part of IGTF distribution
 - Fault tolerant
 - Transfer status monitoring
 - Optimized for performance



the globus alliance

www.globus.org

GridFTP GUI

The screenshot displays the GridFTP GUI interface. At the top, there is a menu bar with 'File', 'Connect', 'Security', 'Log', and 'Help'. Below the menu bar are five buttons: 'Credential' (with a padlock icon), 'GridFTP' (with a grid icon), 'Local' (with a folder icon), 'Advanced' (with a gear icon), and 'Remote Dir' (with a padlock icon).

The main interface is divided into three panes:

- Monitor Frame:** Contains a 'Transfer Queue' table with columns: Job, From, To, Sta, Cur, %, Err, RFT. The table shows 12 rows of transfer progress.
- Remote System ->clutch.aps.anl.gov:2811/:** Shows a directory tree with folders: /data/test/, raw/, reconstructed/, test1/, test_liuwt/, tomography/, and a file: sam01_exp.hdf.
- Local Sy...:** Shows the local system directory tree with folders: C:\, D:\, E:\, F:\, and G:\.

At the bottom of the window, the following text is displayed:

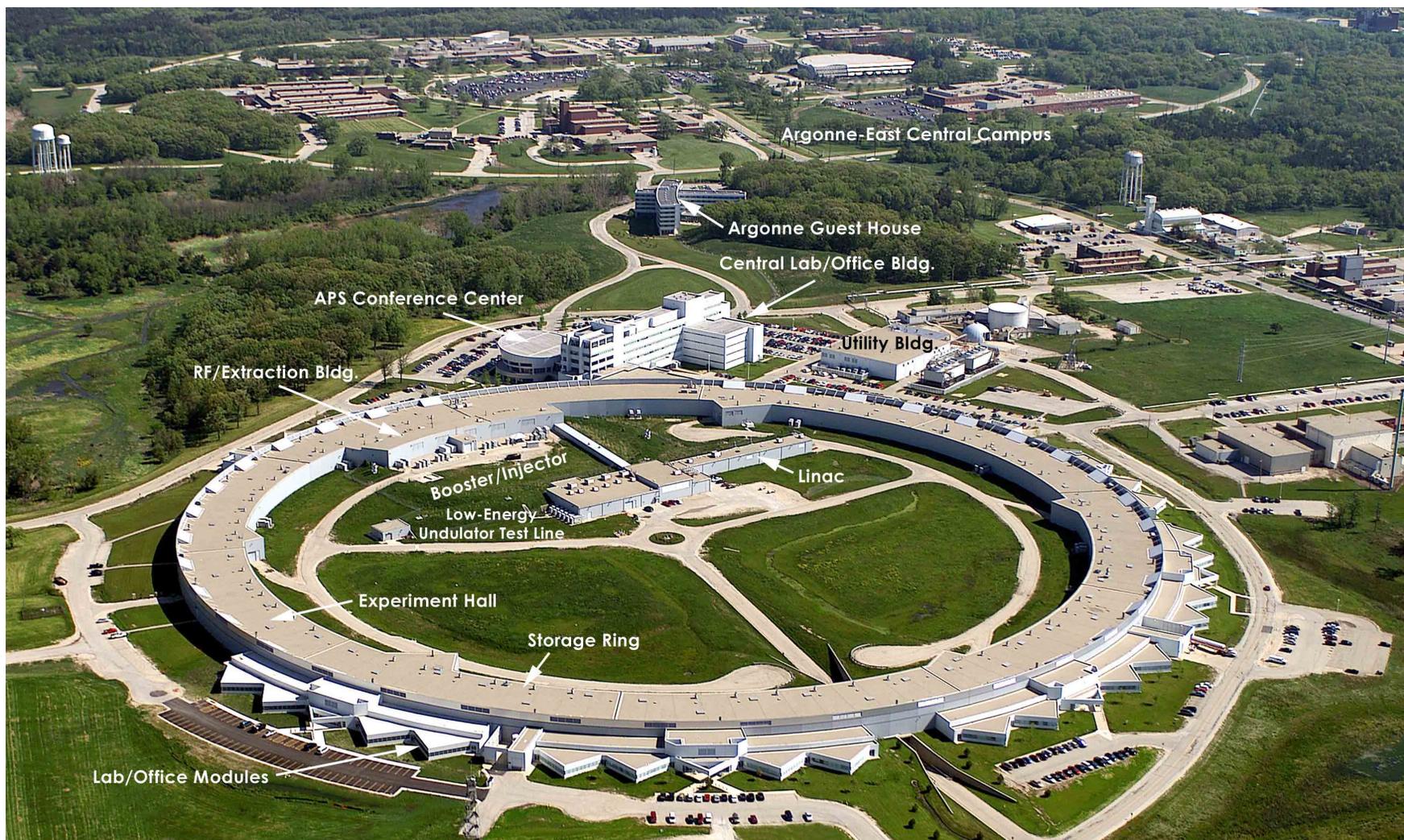
Proxy Subject: /DC=org/DC=doegrids/OU=People/CN=Wantao Liu 896661/CN=764692481
Time Left: 21 days, 16 h, 55 min, 51 sec



the globus alliance

www.globus.org

Advanced Photon Source





the globus alliance

www.globus.org

Advanced Photon Source

- APS beamline experiments generate TBs data
- For example, in tomography beamline users get 3 days of beam time per cycle (4 cycles/yr)
 - Generate 3-4 TBs of data in that 3 days
 - About 15-20 users per cycle
 - 45-80 TBs of data generated per cycle
 - 180-320 TBs of data per year
- Move data back to their home institute after their experiments



the globus alliance

www.globus.org

TomoScript

TomoScript--V2.4.0

File Utilities Security Help

Data Acquisition Acquisition Movement Data Reconstruction

Move Samples

Sample ...	User	Row	Column	Experim...	Experim...
Sampl...	Brian Ti...	1	A	Sam01...	D:/data/...

Add Directory...
Add Sample...
Remove Sample
Edit...
Move Up
Move Down

Transfer Sample

Current Sample: CurrentSample
 Automate Transfers

Source Location

Server: bsi2.xor.aps.anl.gov
Root Path: SampleRecord.RootPath
URL:
gsiftp://<server>/<root path>

Destination Location

Server: blacklab.aps.anl.gov
Root Path: /data/test/tomography
URL:
gsiftp://<server>/<root path>

Progress

Pending: 0
Completed: 0
Failed: 0

Subject: /DC=org/DC=doegrids/OU=People/CN=Brian Tieman 17315/CN=proxy
Time Left: 0 sec



Code Snippet

```
Listener l = new DefaultListener();  
DataTransferExecutor executor  
    = new GridFTPTransferExecutor(  
        Constants.DEFAULT_THREADS_NUM,  
        l, "logfile");  
executor.setSchedulePolicy(Constants.FCFS);  
Transfer t1 = new DirTransfer(  
    "gsiftp://clutch.aps.anl.gov:2811/data/tomo/",  
    "gsiftp://qb1.loni.org:51000/work/tomo/ ");  
executor.addTransfer(t1);
```



the globus alliance
www.globus.org

Experiments Setup

- Move data from APS to LSU and PSC
 - 32 ms RTT
- TCP tuning is enabled in Linux OS
- Dataset is about 696GB
- Compared with globus-url-copy

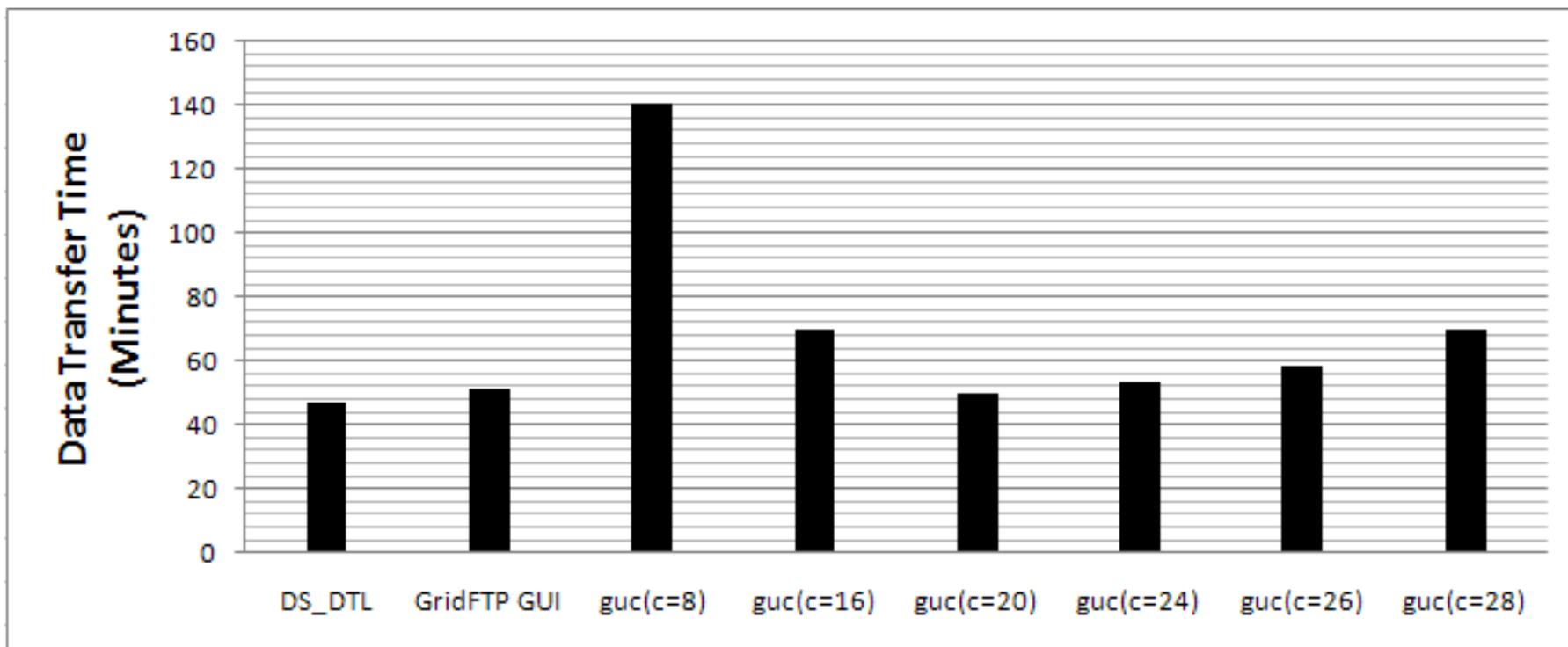




the globus alliance

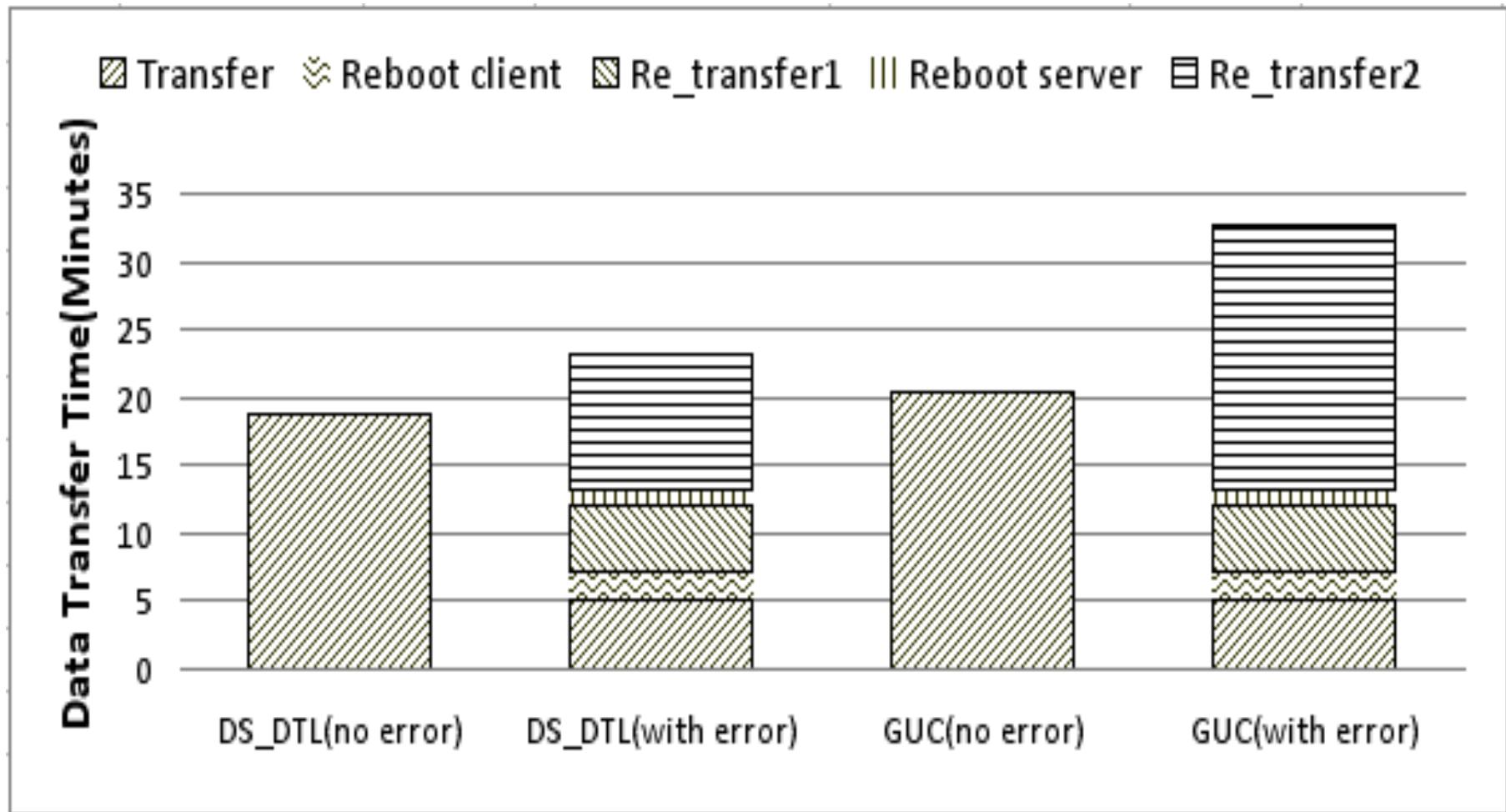
www.globus.org

Experiments Results





Failure Recovery

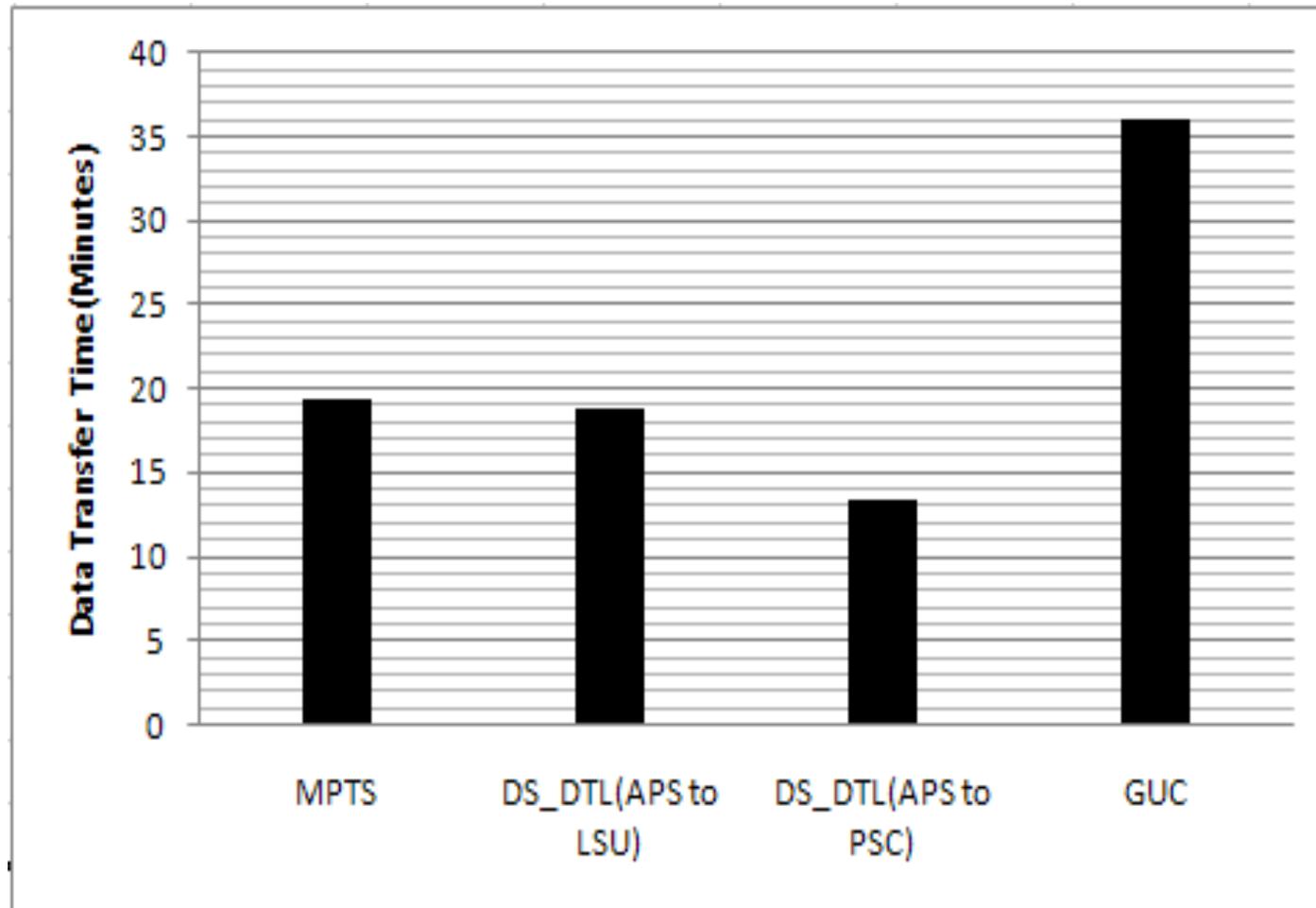




the globus alliance

www.globus.org

Multiple Pair Transfer Scheduling



Questions