

---

CS 696

Introduction to Grid Computing:  
Lecture #17: Data Mangement on  
the TeraGrid

---

Mary Thomas

San Diego State

Thursday, 17-Apr-07

---

# Comment

- These slides are mostly adapted and updated from an SC04 tutorial.

---

# GSI-Authenticated Data Transfer

- TeraGrid File Management
- Data Transfer Performance
- Tools:
  - GridFTP, uberFTP (new)
  - tgcp (new)
  - gsissh (new)
- Hands-on Exercises
  - Use of GridFTP clients & servers to transfer files

---

# TeraGrid File Placement

- No common cross-site filesystems (currently)
- User controls where their data resides
  - Appropriate sites(s)
  - Appropriate storage
    - Online Filesystem(s)
      - Speed, visibility, quotas, backup policy
      - Each filesystem directly accessible from single site
    - Mass Storage Systems
      - Long-term storage, slower access
      - Accessible from all sites

---

# TeraGrid File Movement

- File movement responsibility of user
  - Between Online Filesystems
    - Intra-site
    - Cross-site\*
  - Between Mass Storage and Online Filesystems
    - Intra-site\*
    - Cross-site\*

\* Session focuses on these types of transfers

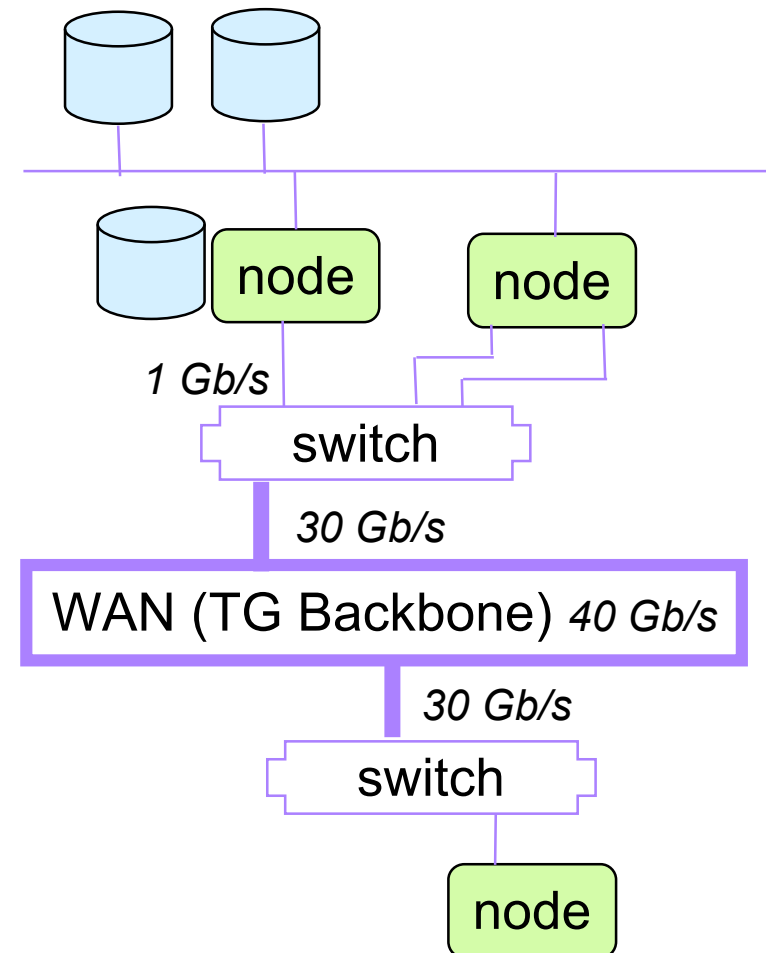
---

# TeraGrid Transfer Environment

- Many sites have nodes dedicated to transferring files
- TeraGrid backbone bandwidth (40 Gb/sec) means Wide Area Network is rarely a bottleneck
- GSI authentication and proxy certificates provide security for transfers
- Transfer requests can be integrated into job execution scripts
  - Moving input data to site(s) of job execution
  - Moving results to another filesystem, site, or archive

# Data Transfer Performance

- What impacts transfer rates?
  - Disk speed
  - Connectivity of disk to node
  - Node characteristics & load
  - Connectivity of node to WAN
  - For all networks
    - Bandwidth
    - Latency
    - Buffer Size
    - Protocol
    - Load
    - Encryption ...
- Don't expect 40 Gb/sec!



# Performance – Choices Matter

- Transfer large files for best performance
- Use fast filesystems, dedicated transfer nodes, optimized transfer parameters
- Transfer 1 GByte file from NCSA to SDSC (10/6/2004)

Choices	Transfer Time	Transfer Rate
Home filesystems Login nodes Default parameters	20 min 18 sec	.845 MBytes/sec (.0066 Gbits/sec)
Parallel filesystems Transfer nodes Optimized parameters	11 sec	93.091 MBytes/sec (.727 Gbits/sec)

---

# Transfer times...

- See:
  - <http://gridinfo.psc.edu/gridftp/speedpage.php>

---

# GridFTP Tools

---

---

# GridFTP Terminology - Protocol

- “GridFTP is a high-performance, secure, reliable data transfer ***protocol*** optimized for high-bandwidth, wide-area networks. GridFTP is based on FTP, the highly popular Internet file transfer protocol.”

- Quoted from Globus Alliance website

---

# Terminology - Server

- A GridFTP **server** process understands requests that adhere to the GridFTP protocol, and performs authentication and data transfer operations based on those requests
  - A system that is configured to automatically start GridFTP server processes is sometimes referred to as a *GridFTP server*
  - Not all systems (nodes) on TeraGrid machines are GridFTP servers
  - Some mass storage front-ends are GridFTP servers

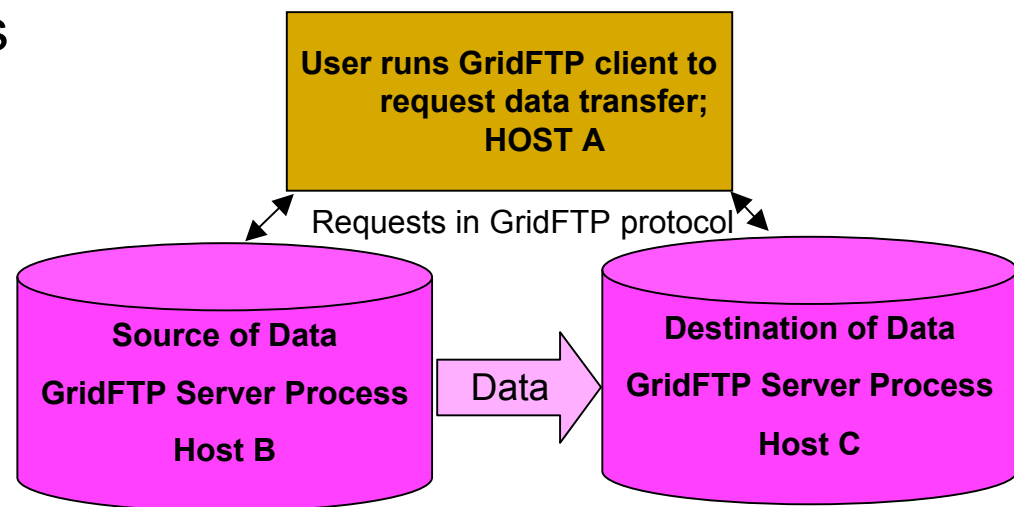
---

# Terminology - Client

- GridFTP *client* programs issue requests that adhere to the GridFTP protocol
  - Users run GridFTP client programs to transfer files
  - *globus-url-copy* and *uberftp* are two GridFTP client programs that are part of the Common TeraGrid Software Stack (CTSS)
  - There is no client program named *gridFTP*, which can be confusing because users are told “use gridFTP to transfer your files”

# Terminology – 3<sup>rd</sup> Party Transfer

- A GridFTP transfer between two GridFTP servers, rather than between a server and a client, is called a ***third-party transfer***
  - A third-party transfer occurs when the GridFTP client initiating the transfer is run on a system that is neither the source nor the destination of the transfer operation
  - Allows use of dedicated transfer nodes



---

# TG GridFTP Server Deployment

- `tg-login1.<site>.teragrid.org` is a GridFTP server
  - Shared resource; Many tasks
  - See [http://www.teragrid.org/userinfo/data/transfer\\_location.php#deployment](http://www.teragrid.org/userinfo/data/transfer_location.php#deployment) for list of servers
- `tg-gridftp.<site>.teragrid.org` resolves to one or more machines that are GridFTP servers
  - Dedicated file transfer resources at many sites
  - Fewer tasks, possibly better connectivity
  - “GridFTP Server”

---

# TG GridFTP Client Deployment

## **globus-url-copy** <source\_url> <destination\_url>

- ❑ command line interface
- ❑ -tcp-bs <size> | -tcp-buffer-size <size>
  - specify the size (in bytes) of the buffer to be used by the underlying ftp data channels
- ❑ -p <parallelism> | -parallel <parallelism>
  - specify the number of streams to be used in the ftp transfer

## **uberftp**

- ❑ interactive GridFTP transfer client
- ❑ configurable tcp buffersize and number of parallel streams

# Using globus-url-copy

```
tg-login3 thomasm/data.tests> globus-url-copy file:`pwd`/python.tar.gz gsiftp://gridftp-
co.ncsa.teragrid.org:2811:python_trans.tar.gz
tg-login3 thomasm/data.tests> ls
total 10860
drwxr-x---  2 thomasm mpk          26 2007-04-17 19:45 .
drwx--x--x 10 thomasm ac          4096 2007-04-17 19:45 ..
-rw-r--r--  1 thomasm mpk    11108613 2007-04-17 19:14 python.tar.gz
globus-url-copy file:`pwd`/python.tar.gz gsiftp://tg-
gridftp.sdsc.teragrid.org:2811:python_trans.tar.gz
tg-login3 thomasm/data.tests> globus-url-copy file:`pwd`/python.tar.gz gsiftp://tg-
gridftp.sdsc.teragrid.org:2811/~/python_trans.tar.gz
tg-login3 thomasm/data.tests> gsissh tg-login.sdsc.teragrid.org "/bin/ls"
python_trans.tar.gz
...
```

---

## Hands-on: Exercise 2

*Copy a 1 MByte file from the current directory at NCSA to your home directory at SDSC. Use a third-party transfer and the GridFTP server nodes at both NCSA and SDSC. Use optimized transfer parameters.*

**Look at the transfer script:**

```
tg-login1> cat ./ex2
/usr/bin/time -f %E
  globus-url-copy -tcp-bs 8388608
  gsiftp://tg-gridftp.ncsa.teragrid.org/`pwd`/OneMBfile
  gsiftp://tg-gridftp.sdsc.teragrid.org/~OneMBfile-Ex2
```

**Run the transfer script:**

```
tg-login1> ./ex2
0:02.72
```

---

## Hands-on: Exercise 3

*Copy a 1 MByte file from your home directory at SDSC to your home directory at ANL/UC. Use a third-party transfer. Use optimized transfer parameters.*

**Look at the transfer script:**

```
tg-login1> cat ./ex3
/usr/bin/time -f %E
  globus-url-copy -tcp-bs 8388608
  gsiftp://tg-gridftp.sdsc.teragrid.org/~/OneMBfile-Ex2
  gsiftp://tg-gridftp.uc.teragrid.org/~/OneMBfile-Ex3
```

**Run the transfer script:**

```
tg-login1> ./ex3
0:02.77
```

---

# Hands-on: Exercise 4

*Copy a 1 MByte file from the current directory at NCSA to Mass Storage at NCSA. Use optimized transfer parameters.*

**Look at the transfer script:**

```
tg-login1> cat ./ex4
/usr/bin/time -f %E
  globus-url-copy -tcp-bs 8388608
  file:`pwd`/OneMBfile
  gsiftp://mss.ncsa.teragrid.org/~OneMBfile-Ex4
```

**Run the transfer script:**

```
tg-login1> ./ex4
0:00.80
```

---

## Hands-on: Exercise 5

*Copy a 1 MByte file from your home directory at SDSC to Mass Storage at NCSA. Disable data channel authorization, use 3<sup>rd</sup> party transfer, and use optimized transfer parameters.*

**Look at the transfer script:**

```
tg-login1> cat ./ex5
/usr/bin/time -f %E
  globus-url-copy -nodcau -tcp-bs 8388608
  gsiftp://tg-gridftp.sdsc.teragrid.org/~/OneMBfile-Ex1
  gsiftp://mss.ncsa.teragrid.org/~/OneMBfile-Ex5
```

**Run the transfer script:**

```
tg-login1> ./ex5
0:03.01
```

---

# Hands-on: Exercise 6 – pg 1

*Copy a 1 MByte file from your current directory to Mass Storage System at NCSA. Use optimized transfer parameters. Interactive session.*

**Start uberftp and set transfer parameters:**

```
tg-login1> uberftp
```

```
uberftp> parallel 2
```

```
uberftp> tcpbuf 4194304
```

```
TCP buffer set to 4194304 bytes
```

**Open connection to Mass Storage System:**

```
uberftp> open mss.ncsa.teragrid.org
```

```
%%% BANNER %%%
```

```
220 UNIX Archive FTP server ready.
```

```
230 User xxx logged in.
```

# Hands-on: Exercise 6 – pg 2

**Copy the file:**

```
uberftp> put OneMBfile OneMBfile-Ex6
```


```
150 Opening BINARY connection(s) for OneMBfile-Ex6.
```

```
226 Transfer complete.
```

**Get a listing of the Mass Storage System directory:**

```
uberftp> ls
```

```
-rw---- user group DK common 10485760 date OneMBfile-Ex4  
-rw---- user group DK common 10485760 date OneMBfile-Ex5  
-rw---- user group DK common 10485760 date OneMBfile-Ex6
```



File is on disk. AR used to indicate file on tape. *stage* and *mstage* commands move files from tape to disk. See TeraGrid UniTree online documentation for details.

---

# Hands-on: Exercise 7 – pg 1

*Continuing previous interactive uberftp session, transfer three 1 MByte files from Mass Storage System at NCSA to home directory at ANL/UC. This will be a 3<sup>rd</sup> party transfer.*

**Establish “local” connection to UC:**

```
uberftp> lopen tg-gridftp.uc.teragrid.org
220 tg-grid1.uc.teragrid.org GridFTP Server ... ready.
230 User xxx logged in.
```

---

# Hands-on: Exercise 7 – pg 2

**Get multiple files from MSS to the “local” (UC) site:**

```
uberftp> mget OneMBfile*
```

```
dst> 500 `SBUF 4194304`: command not understood
```

```
dst> 500 `WIND 4194304`: command not understood
```

```
src> 150 Opening BINARY connection(s) for OneMBfile-Ex4 (1048576 bytes).
```

```
dst> 150 Opening BINARY mode data connection.
```

```
src> 226 Transfer complete.
```

```
dst> 226 Transfer complete.
```

```
. . .
```

```
src> 150 Opening BINARY connection(s) for OneMBfile-Ex5 (1048576 bytes).
```

```
. . .
```

```
src> 150 Opening BINARY connection(s) for OneMBfile-Ex6 (1048576 bytes).
```

```
dst> 150 Opening BINARY mode data connection.
```

```
src> 226 Transfer complete.
```

```
dst> 226 Transfer complete.
```

---

# Hands-on: Exercise 7 – pg 3

**List OneMB files at “local” (UC) site:**

```
uberftp> lls OneMBfile*
```

```
150 Opening BINARY mode data connection  
-rw-r--r- user 1048576 date OneMBfile-Ex3  
-rw-r--r- user 1048576 date OneMBfile-Ex4  
-rw-r--r- user 1048576 date OneMBfile-Ex5  
-rw-r--r- user 1048576 date OneMBfile-Ex6
```

**Quit uberftp:**

```
uberftp> quit
```

```
221-You have transferred 3145728 bytes in 3 files.  
221- Total traffic for this session was 3163276 bytes in 4 transfers.  
221-Thank you for using the FTP service on tg-grid1.uc.teragrid.org.  
221 Goodbye.  
221 Goodbye.
```

---

# Hands-on: Wrapup

*Log into SDSC and UC sites and verify files were copied.*

```
tg-login> gsissh tg-login.sdsc.teragrid.org
```

```
> ls -l
```

```
-rw-r--r--    user group      1048576 date OneMBfile-Ex1  
-rw-r--r--    user group      1048576 date OneMBfile-Ex2
```

```
> exit
```

```
tg-login> gsissh tg-login.uc.teragrid.org
```

```
> ls -l
```

```
-rw-r--r--    user group      1048576 date OneMBfile-Ex3  
-rw-r--r--    user group      1048576 date OneMBfile-Ex4  
-rw-r--r--    user group      1048576 date OneMBfile-Ex5  
-rw-r--r--    user group      1048576 date OneMBfile-Ex6
```

```
> exit
```

---

# GridFTP: Data Transfer Summary

- GridFTP clients *globus-url-copy* and *uberftp* can be used to perform transfers between many TeraGrid online filesystems and mass storage systems accessible via GridFTP servers.
- Users responsible for managing data transfers, including job-related data movement which can be incorporated into job scripts.
- Choose servers, filesystems, and transfer parameters wisely to optimize performance.
- Performance (usually) limited by end node connectivity, not WAN bandwidth.
- Ongoing efforts to improve rates, usability, add servers.

---

# UberFTP

---

<http://dims.ncsa.uiuc.edu/set/uberftp/>

---

# Interactive GridFTP client

- Supports:
  - standard FTP client mechanics
  - GSI authentication
  - parallel data channels
  - striping.
- session within which files may be transferred and directories and files may be manipulated.
- Some users prefer UberFTP over globus-url-copy because the syntax is less verbose and less prone to typographical errors.
  - globus-url-copy is more versatile for imbedding data transfers into job scripts.

# UberFTP

## Usage

```
uberftp [-?] [-d]
          [ host [-P port] [-a auth] [-c N] [-t type] ]
uberftp [-?] [-d] host [-P port] [-a auth]
          [-c N] [-t type] "cmd"
uberftp -v
```

## Options

host     Connect to host

-P port   Connect to port (default 2811 for GSI)

-a auth   Use auth authentication (default is GSI)

-c N     Use N parallel data connections. Defaults to 1.

-t type   Use ASCII or BINARY transfer type. Valid type options are 'a' (ASCII) or 'b' (BINARY). Default is BINARY.

cmd     This specifies the FTP commands to run once the control connection is establish. This list must be enclosed in quotes. Multiple commands are semicolon delimited  
uberftp will execute these commands and then exit

-v     Print version information and exit

-?     Print this help message

-d     Turn on debug (verbose) output

# Interactive session

```
tg-login3 thomasm/data.tests> uberftp
uberftp> help
Usage "help [topic]" where topic is one of:
active      ascii      bget      bput      binary    blksize
bugs        bye        cat        cd         chmod     close
debug       dir        get        hash      help      lcat
lcd         lchmod    lclose    lls        lmkdir    lopen
lpwd        lrename   lrm        lrmdir    lquote    ls
mget        mkdir     mode      mput      open      order
passive     parallel  pget      pput      pwd       put
rename      rm        rmdir     quit      quote     tcpbuf
```

Examples at:

<http://www.teragrid.org/userinfo/data/examples.php>

---

tgcp

---

---

`tgcp [options] <sourceURL> <destURL>`

- command-line user tool intended to provide high-performance file transfer
  - efficient copying of files and directories between and within gridFTP enabled clusters.
- tgcp is a wrapper for globus-url-copy and RFT (Reliable Transfer Service),
- options include
  - '-big', for a striped globus-url-copy
  - '-rft', to manage transfers using RFT.
- invokes third-party transfers between GridFTP servers at the TeraGrid source and the TeraGrid destination sites.
- Tg administrators maintain configuration data such as TCP buffer sizes for TeraGrid host/domain source-destination pairs, users do not need to enter these values.

---

gsiscp

---

---

# Gsiscp

- Works like standard scp, but automatically uses the GSI proxy credential generated from myproxy-logon
- See previous slides for example
- Some indication that performance better than tgcp.
  - CTSS4 will drop tgcp and support gsissh