



# NERSC Online CA: Enabling Grid Data Transfers

**Shreyas Cholia**  
**Open Software and Programming Group**

NERSC Division, LBNL  
October 2008





## Goals

- **Introduce new easy-to-use certificate service**
- **Introduce NERSC Grid Client package**
- **Demonstrate grid data clients**
- **Discuss performance and firewall issues**



# Grids and Certificates

- **NERSC supports the OSG/Globus software toolkits to provide a common interface across our systems**
  - Remote Job Submission Services (GRAM)
  - Remote File Transfer Services (GridFTP)
- **We use grid certificates to access these grid resources**



## Certificates - Old Way

<http://www.doe grids.org/pages/cert-request.html>

- Fill out online form in browser
- Click on email link to retrieve cert
- Import Cert into Browser
- Export Cert into File
- Convert File into PEM format
- Transfer certificate information into NIM
- Initialize proxy to generate a short lived certificate for use with grid tools

**Too Cumbersome for many users!**



## Certificates - NERSC CA Way

```
myproxy-logon -T -s nerscca.nersc.gov
```

(Supply NIM password)

**That's it!! You have a ready to use short lived grid certificate (12 hours). NERSC already has your certificate information set up in NIM.**



## Additional Flags

- l <username> (specify your username)
- t <hours> (specify a lifetime - max 277 hrs)



# Clients

On desktop/remote machine:

- Pacman is a simple package installer from BU
- NERSC provides a pacman installer for a simple grid client
  - Out of the box lightweight grid client
  - No need to deal with complicated Globus installs
- To install
  - Download pacman
  - Run:

```
pacman -get \
```

```
http://www.nersc.gov/nusers/services/Grid/pacman:nersc-grid-client
```

On NERSC systems:

- module load osg



# Using GridFTP

## Get Certificate

- `myproxy-logon -T -s nerscca.nersc.gov`

## Set up Environment

- `source /nersc-grid-client-dir/setup.sh`

## Invoke Client

- `globus-url-copy`  
`file:///home/shreyas/testfile`  
`gsiftp://davinci.nersc.gov//scratch/scr`  
`atchdirs/shreyas/testfile`

or

- `uberftp davinci.nersc.gov`  
– interactive client

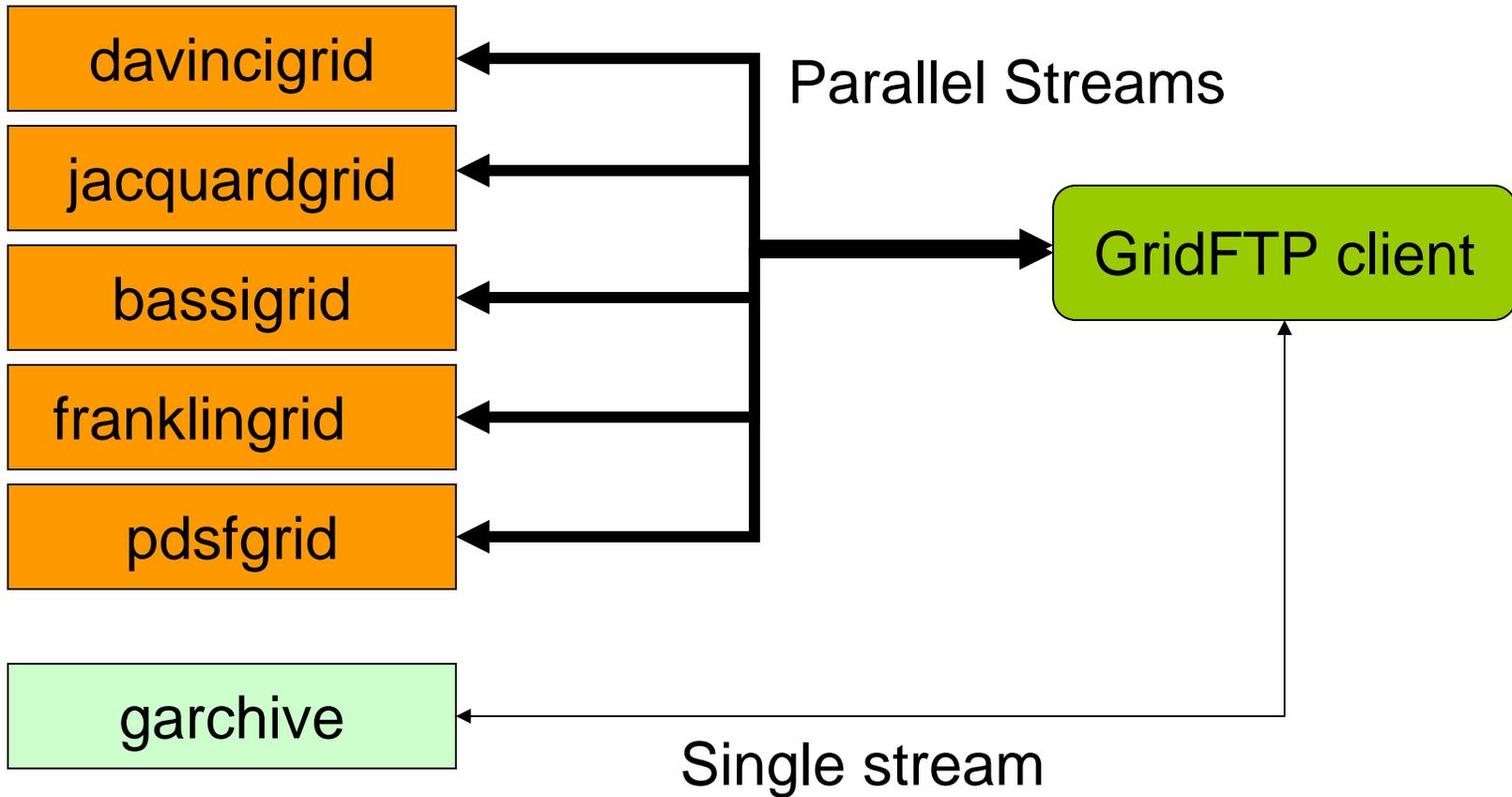


# Endpoints

- **gsiftp://hostgrid.nersc.gov//full/path/file**
  - davincigrd, pdsfgrid, franklingrid, jacquardgrid, bassgrid
  - Also: garchive
- **file://local/path/to/file**
- **data.nersc.gov is the recommended host for NGF**
- **You can use two gsiftp endpoints for third party transfer**



# NERSC GridFTP Hosts





# Performance Tips

- **Increase number of parallel streams (2-8)**
  - globus-url-copy: -p flag
  - Uferftp: parallel command
- **Increase TCP buffer size**
  - globus-url-copy: -tcp-bs flag
  - Uferftp: tcpbuf command



# Firewalls

- High performance data transfer tools do not always play well with firewalls, because they need an open range of ports for parallel transfers
- Port range 60000-65000 open on NERSC hosts to allow incoming data connections
  - Set automatically in server
- If your client or remote gridftp server is blocking connections, you may need to create a hole in the firewall and set:

```
export GLOBUS_TCP_PORT_RANGE=60000,65000
```



# Documentation

<http://www.nersc.gov/nusers/services/Grid/grid.php>