



# Globus GridFTP and RFT: An Overview and New Features

Raj Kettimuthu  
Argonne National Laboratory and  
The University of Chicago



the globus alliance

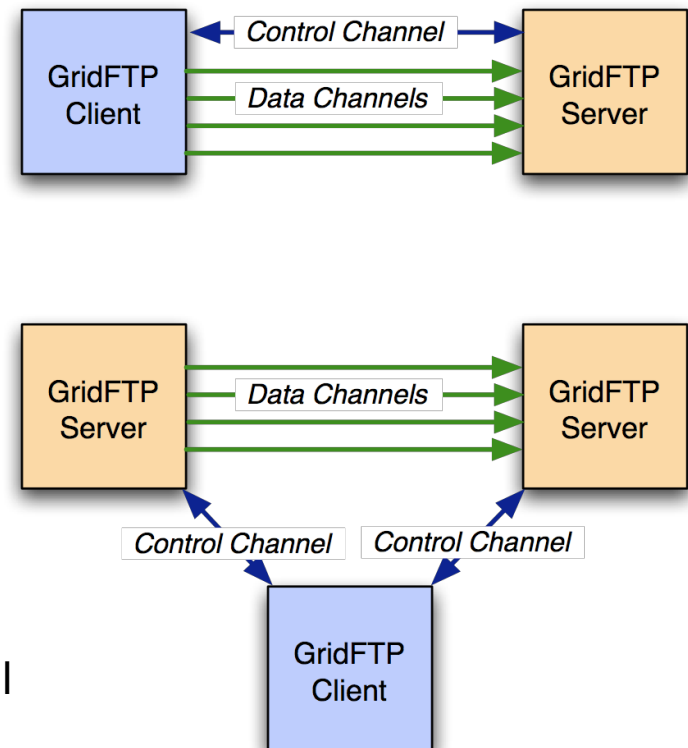
[www.globus.org](http://www.globus.org)

## What is GridFTP?

- High-performance, reliable data transfer protocol optimized for high-bandwidth wide-area networks
- Based on FTP protocol - defines extensions for high-performance operation and security
- We supply a reference implementation:
  - ◆ Server
  - ◆ Client tools (globus-url-copy)
  - ◆ Development Libraries
- Multiple independent implementations can interoperate
  - ◆ Fermi Lab and U. Virginia have home grown servers that work with ours.

# GridFTP

- Two channel protocol like FTP
- Control Channel
  - ◆ Communication link (TCP) over which commands and responses flow
  - ◆ Low bandwidth; encrypted and integrity protected by default
- Data Channel
  - ◆ Communication link(s) over which the actual data of interest flows
  - ◆ High Bandwidth; authenticated by default; encryption and integrity protection optional



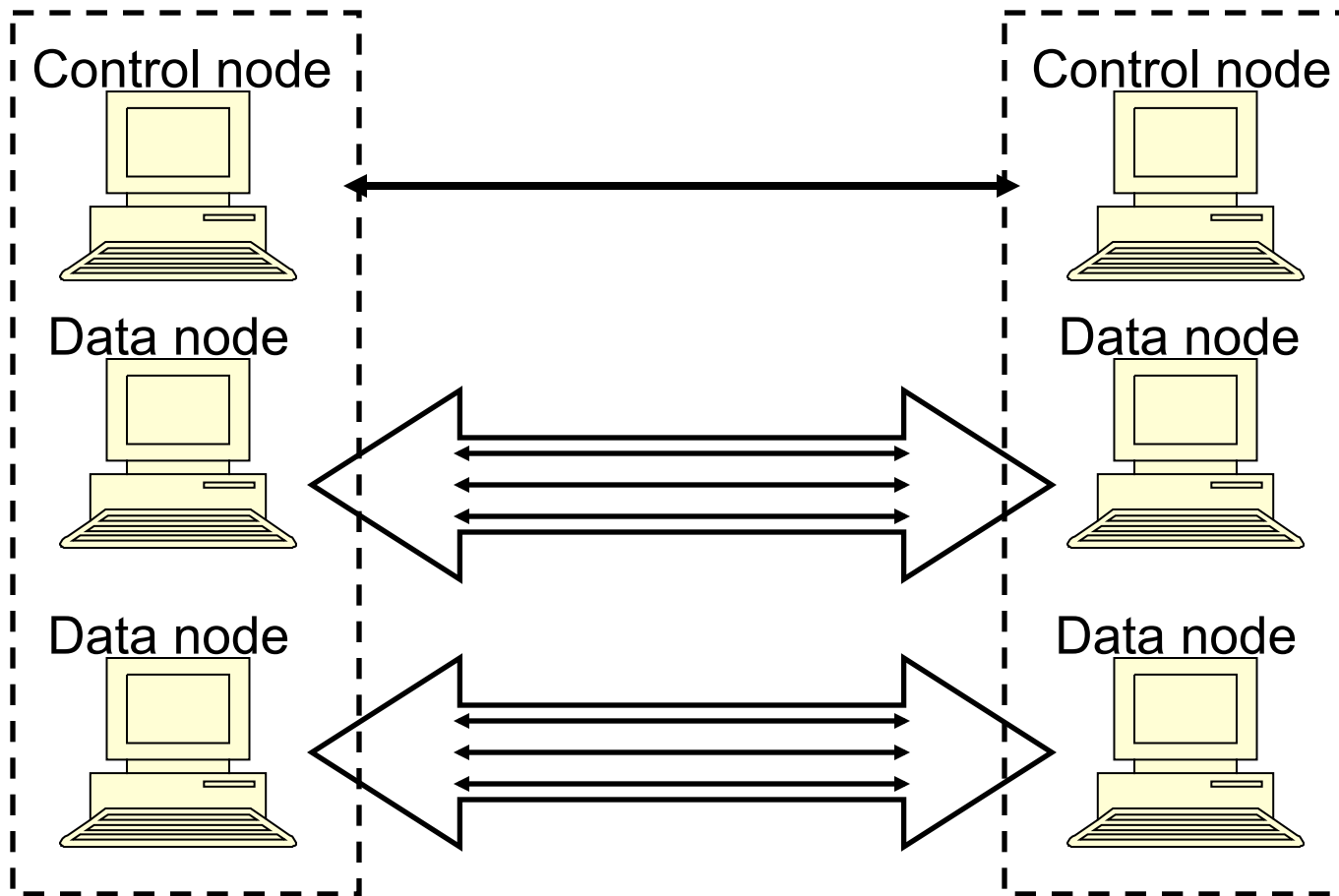


# Globus GridFTP

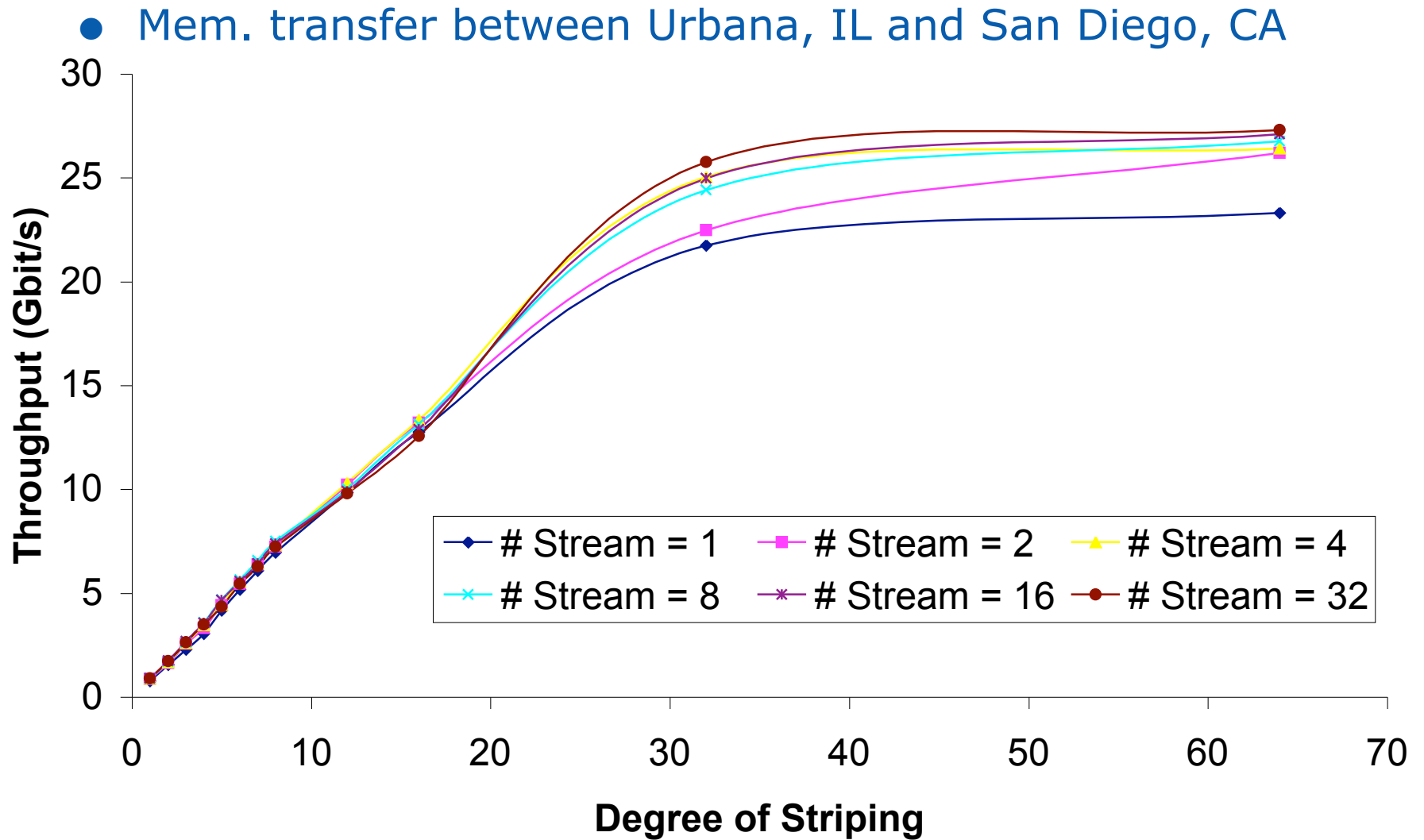
- Performance
  - ◆ Parallel TCP streams
  - ◆ Non TCP protocol such as UDT
  - ◆ Order of magnitude greater
- Cluster-to-cluster data movement
  - ◆ Another order of magnitude
- Support for reliable and restartable transfers
- Multiple security options
  - ◆ Anonymous, password, SSH, GSI
- Modular and easy to optimize for various storage
  - ◆ HPSS, SRB



# Cluster-to-Cluster transfers

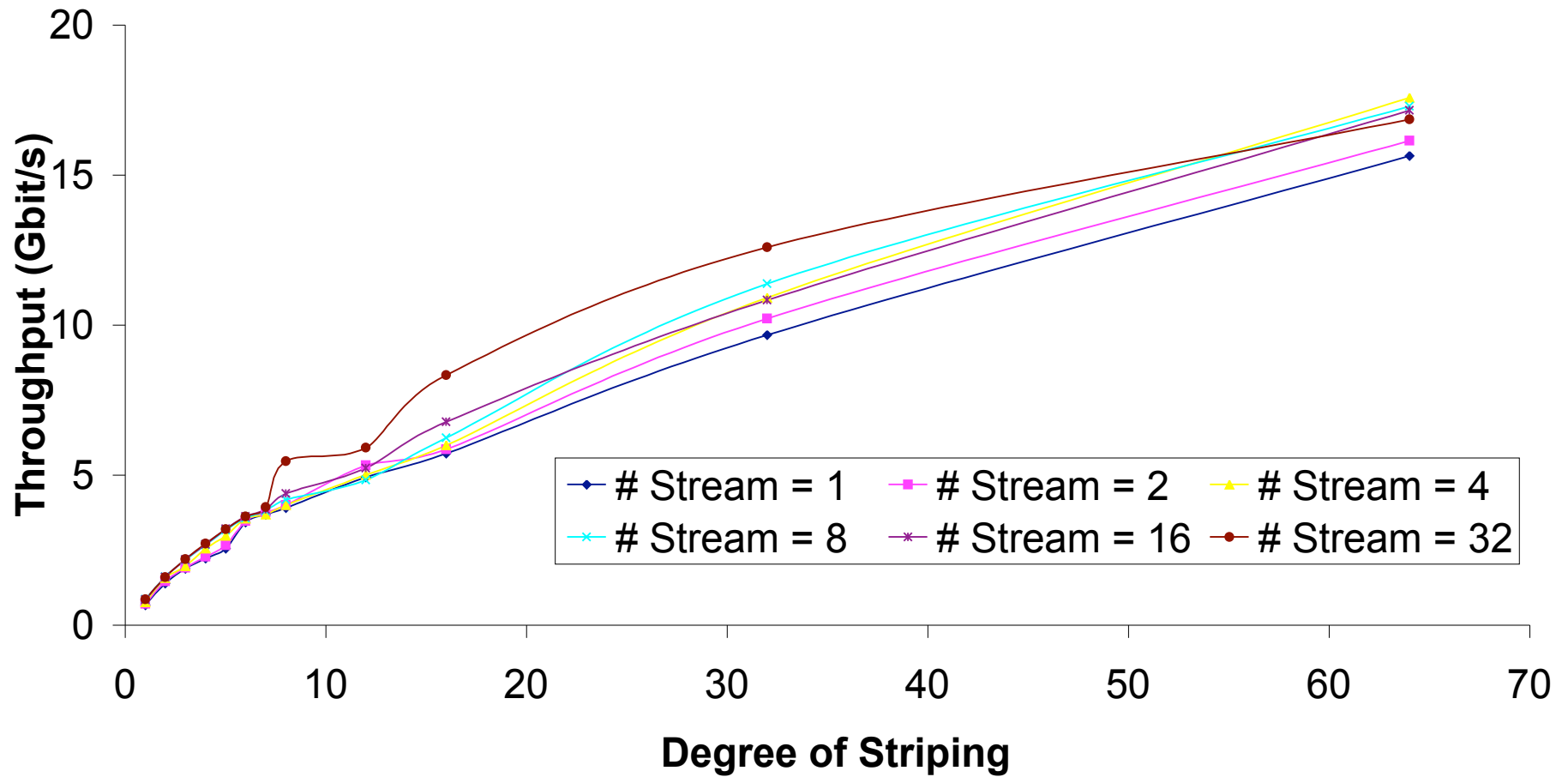


# Performance



# Performance

- Disk transfer between Urbana, IL and San Diego, CA





the globus alliance

[www.globus.org](http://www.globus.org)

## Users

- HEP community is basing its entire tiered data movement infrastructure for the LHC computing Grid on GridFTP
- Southern California Earthquake Center (SCEC), Laser Interferometer Gravitational Wave Observatory (LIGO), Earth Systems Grid (ESG) use GridFTP for data movement
- European Space Agency, Disaster Recovery Center in Japan move large volumes of data using GridFTP
- An average of more than 2 million data transfers happen with GridFTP every day





## New Features

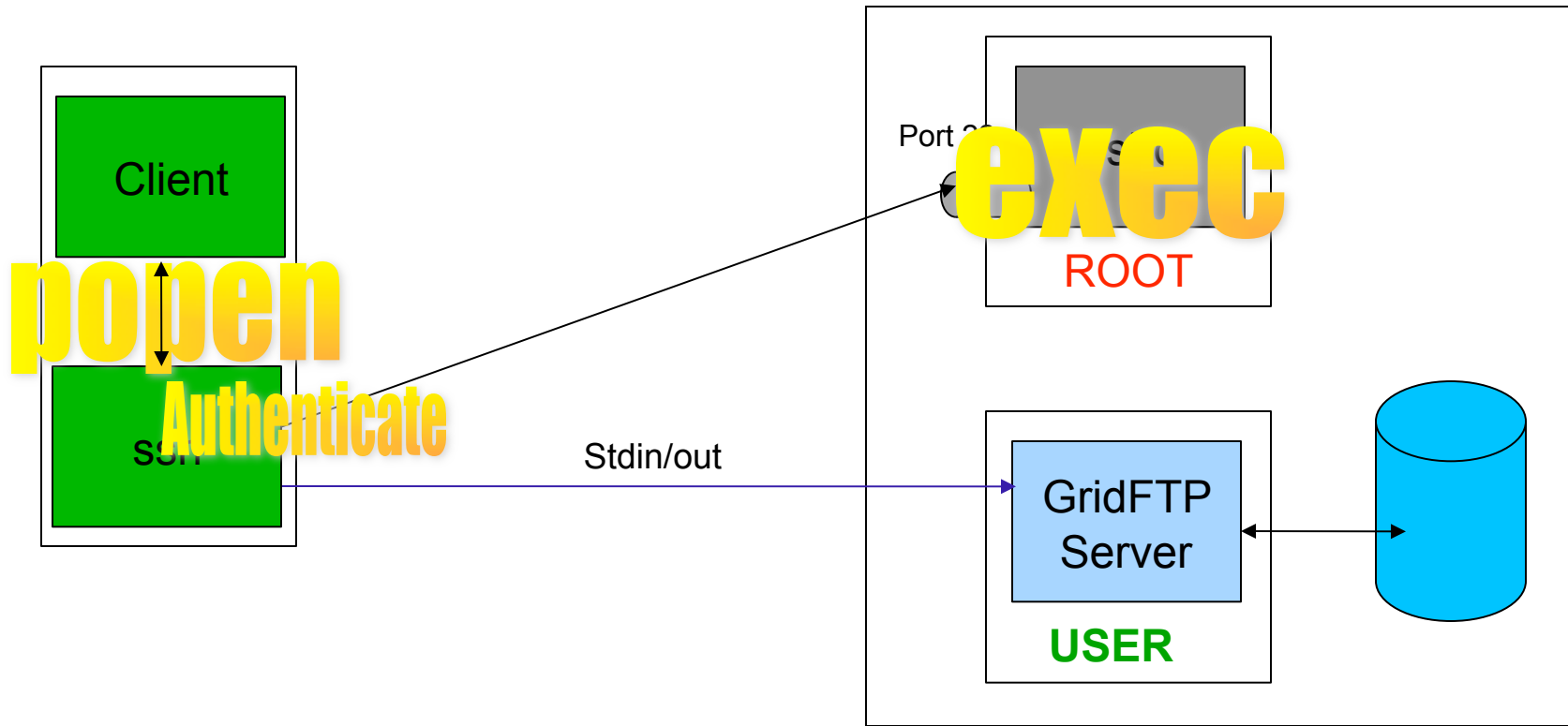
- GUI client
- SSH security for GridFTP
- GridFTP over UDT
- Pipelining
- Multicasting / Overlay Routing
- Scalability
- Lotman Storage plugin
- Anomaly and bottleneck detection using Netlogger



## A GUI client for GridFTP

- An alpha version is available at <http://www.globus.org/cog/demo/>
- Java web start application
- Integrated with myproxy-logon
  - ◆ Certificates can be completely hidden from the user
- If certificates are in place, proxy can be generated through the GUI
- Provides support for RFT as well

# SSH Security for GridFTP



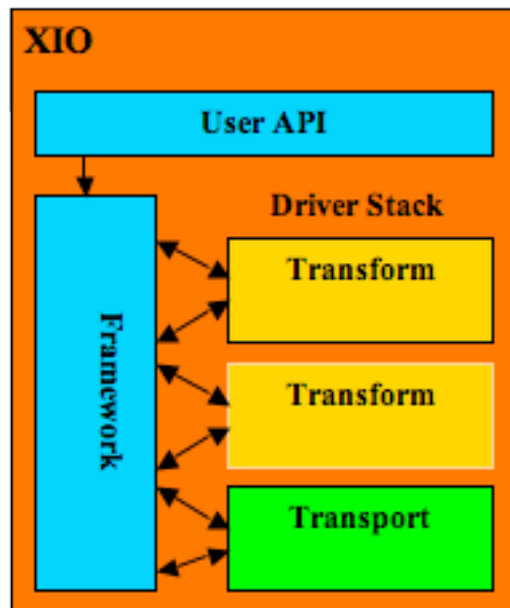


## SSH Security for GridFTP

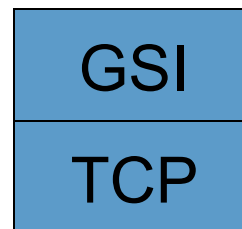
- Client support for using SSH is automatically enabled
- On the server side (where you intend the client to remotely execute a server)
  - ◆ `setup-globus-gridftp-sshftp -server`
- In order to use SSH as a security mechanism, the user must provide urls that begin with `sshftp://` as arguments.
  - ◆ `globus-url-copy sshftp://<host>:<port>/<filepath> file:/<filepath>`
  - ◆ `<port>` is the port in which `sshd` listens on the host referred to by `<host>` (the default value is 22).

## GridFTP over UDT

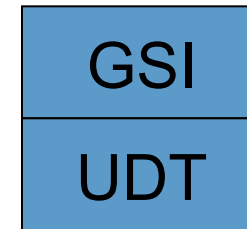
- GridFTP uses XIO for network I/O operations
- XIO presents a POSIX-like interface to many different protocol implementations



Default  
GridFTP



GridFTP  
over UDT





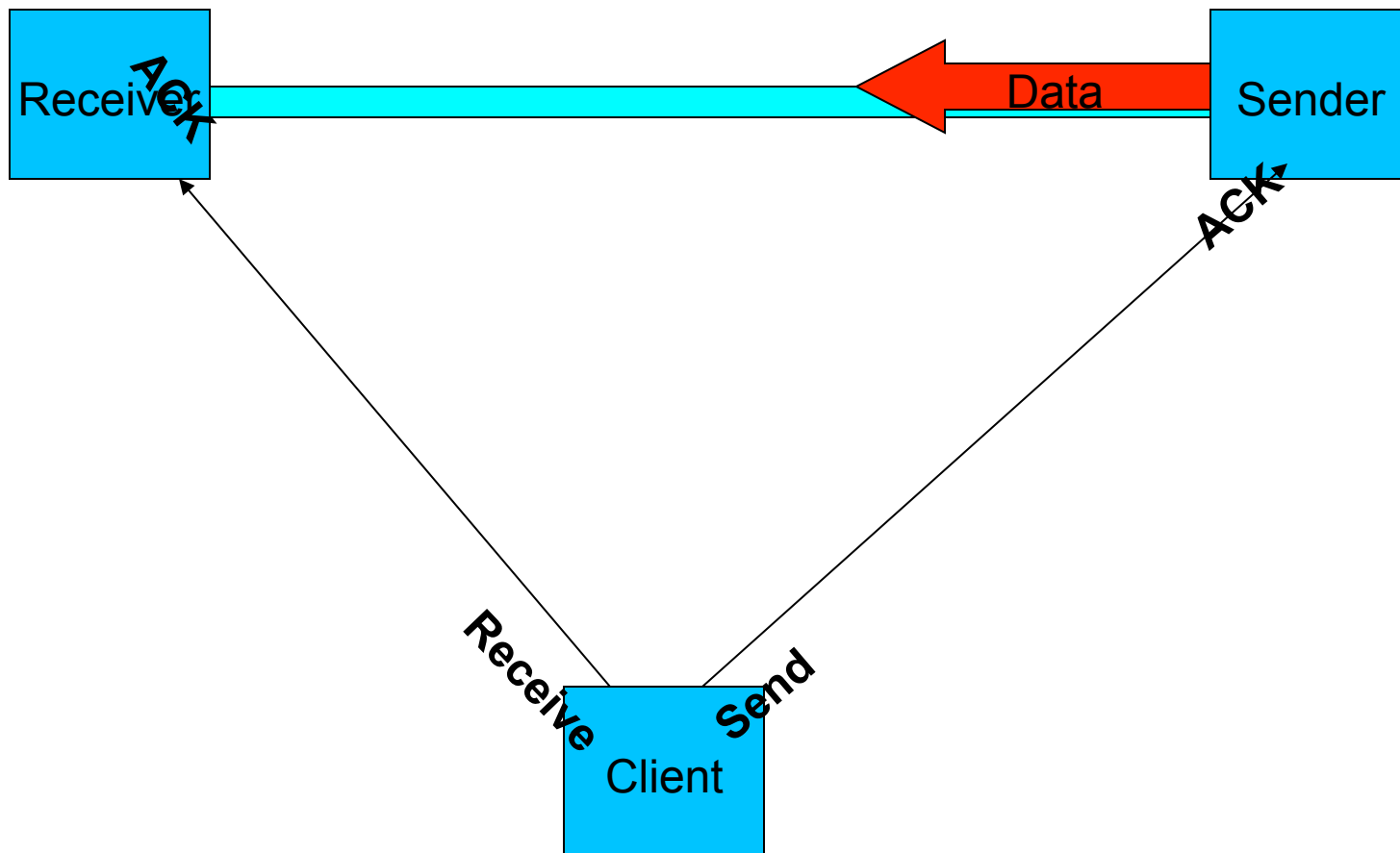
## GridFTP over UDT

	Argonne to NZ Throughput in Mbit/s	Argonne to LA Throughput in Mbit/s
Iperf – 1 stream	19.7	74.5
Iperf – 8 streams	40.3	117.0
GridFTP mem TCP – 1 stream	16.4	63.8
GridFTP mem TCP – 8 streams	40.2	112.6
GridFTP disk TCP – 1 stream	16.3	59.6
GridFTP disk TCP – 8 streams	37.4	102.4
GridFTP mem UDT	179.3	396.6
GridFTP disk UDT	178.6	428.3
UDT mem	201.6	432.5
UDT disk	162.5	230.0



# Lots of Small Files (LOSF) Problem

- Traditional transfer pattern





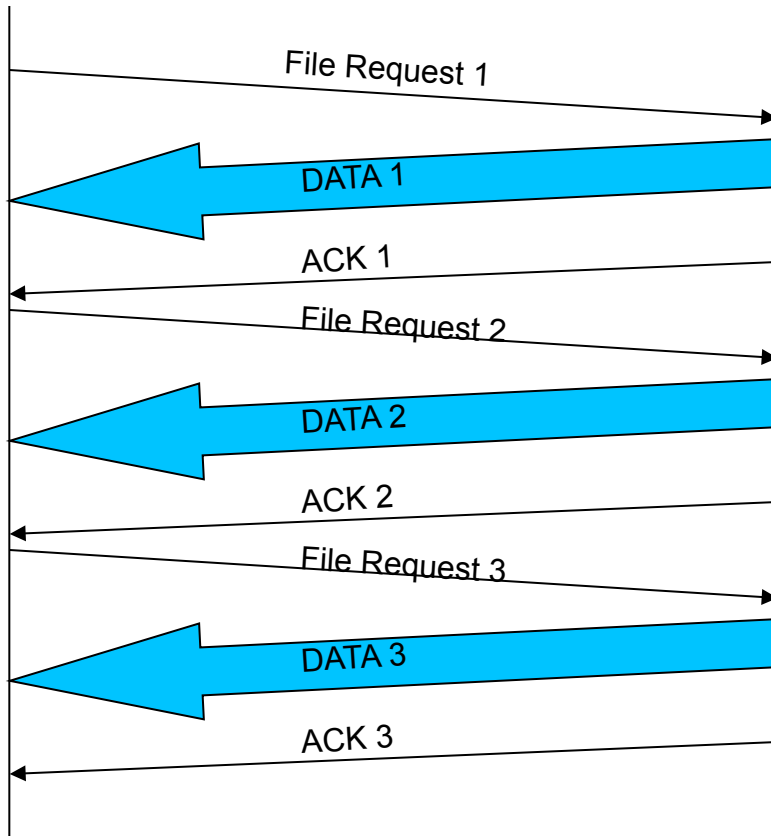
## Pipelining

- Allow many outstanding transfer requests
- Send next request before previous completes
  - ◆ Latency is overlapped with the data transfer
- Backward compatible
  - ◆ Wire protocol doesn't change
  - ◆ Client side sends commands sooner

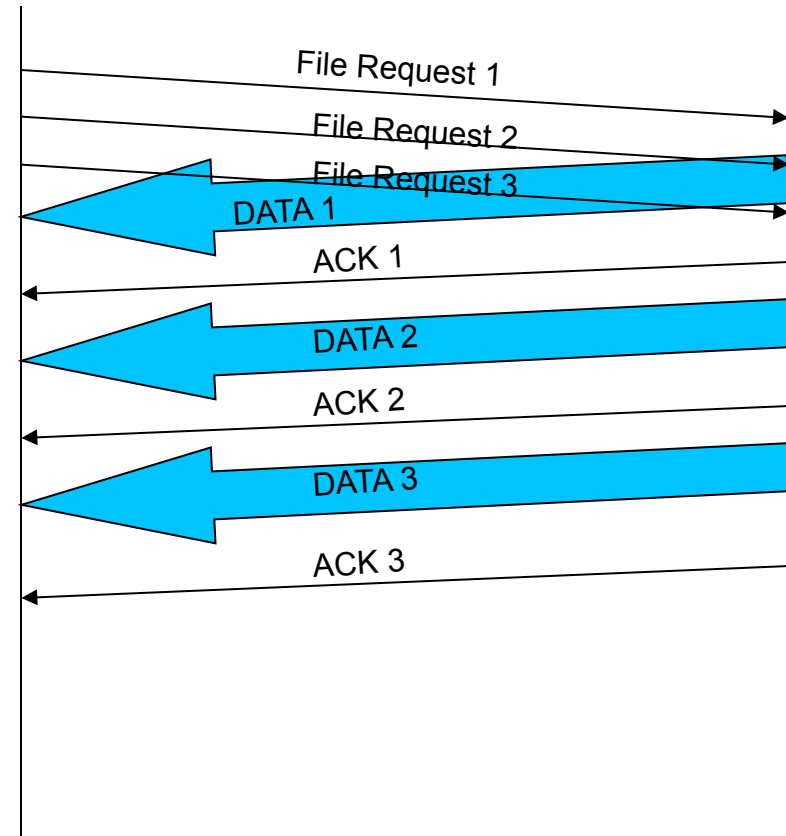


# Pipelining

## Traditional



## Pipelining

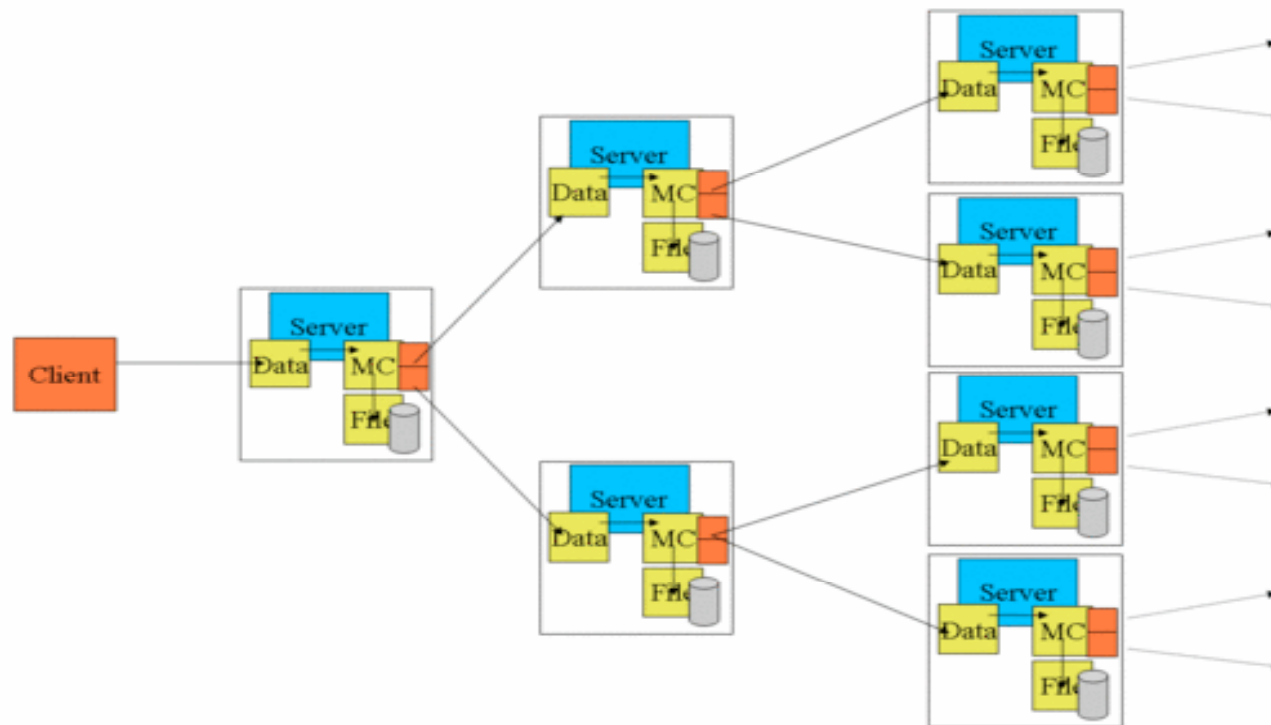


- Significant performance improvement for LOSF



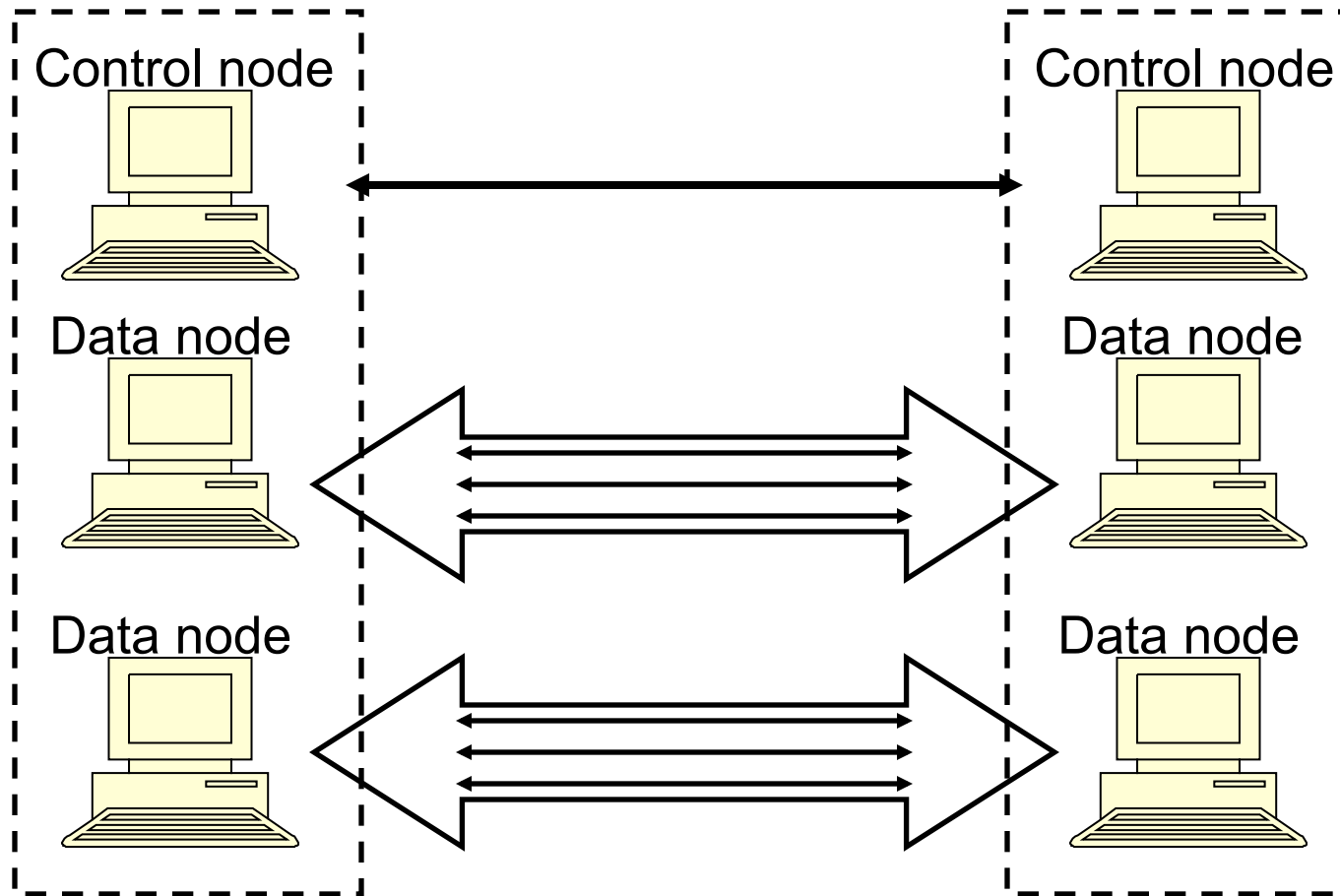
# Multicast / Overlay Routing

- Enable GridFTP to transfer single data set to many locations or act as an intermediate routing node





# Scalability



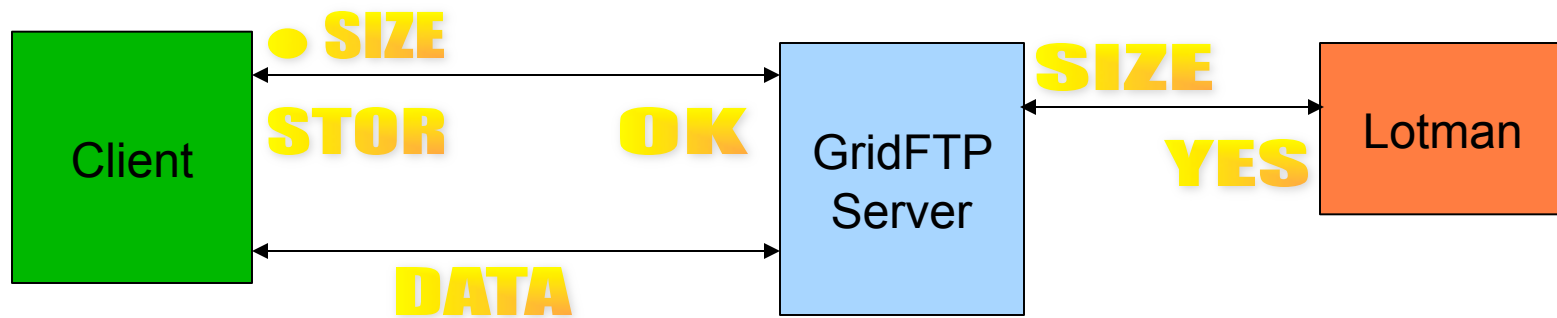
- Data nodes can be added dynamically - need more throughput, add more data nodes



## Storage Plugin

- Destination storage might run out of space in the middle of a GridFTP transfer
- Lotman - tool from univ. of wisconsin that manages storage
- Developed plugin for GridFTP to interact with Lotman
- Space availability (for individual file transfers) determined ahead of transfers to Lotman enabled storage

# GridFTP with Lotman





the globus alliance

[www.globus.org](http://www.globus.org)

# Anomaly and Bottleneck Detection using Netlogger

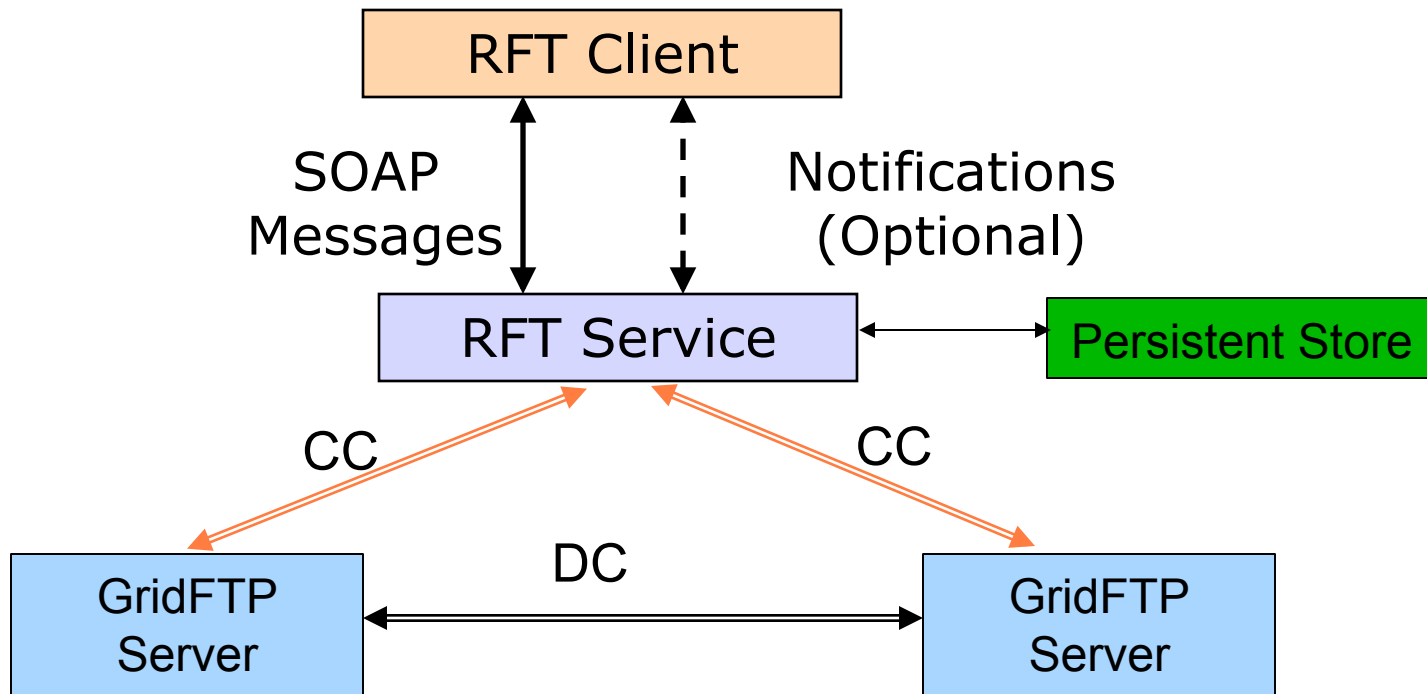
- GridFTP server can be instrumented with Netlogger
- Log messages which can be post processed using Netlogger tools
- Fine grained disk and net I/O characteristics can then be visualized and analyzed



# Reliable File Transfer Service (RFT)

- GridFTP - on demand transfer service
  - ◆ Not a queuing service
- RFT - GridFTP client
  - ◆ Queues requests
  - ◆ Orchestrates transfers on client's behalf
  - ◆ Third party transfers
  - ◆ Interacts with many GridFTP servers
  - ◆ Retry requests on failure
  - ◆ Recovers from GridFTP and RFT service failures

# RFT

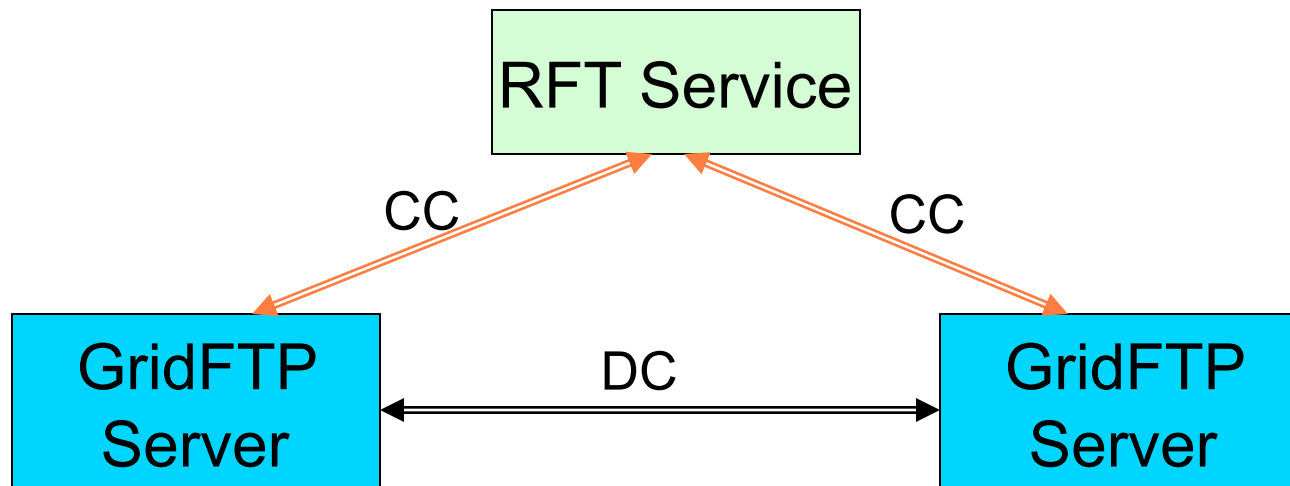






## RFT - Connection Caching

- Control channel connections (and thus the data channels associated with it) are cached to reuse later (by the same user)





## RFT - Connection Caching

- Reusing connections eliminate authentication overhead on the control and data channels
- Measured performance improvement for jobs submitted using Condor-G
- For 500 jobs - each job requiring file stageIn, stageOut and cleanup (RFT tasks)
  - ◆ 30% improvement in overall performance
  - ◆ No timeout due to overwhelming connection requests to GridFTP servers