the globus alliance
www.globus.org

Globus GridFTP: What's New in 2007

Raj Kettimuthu
Argonne National Laboratory and
The University of Chicago

GridFTP

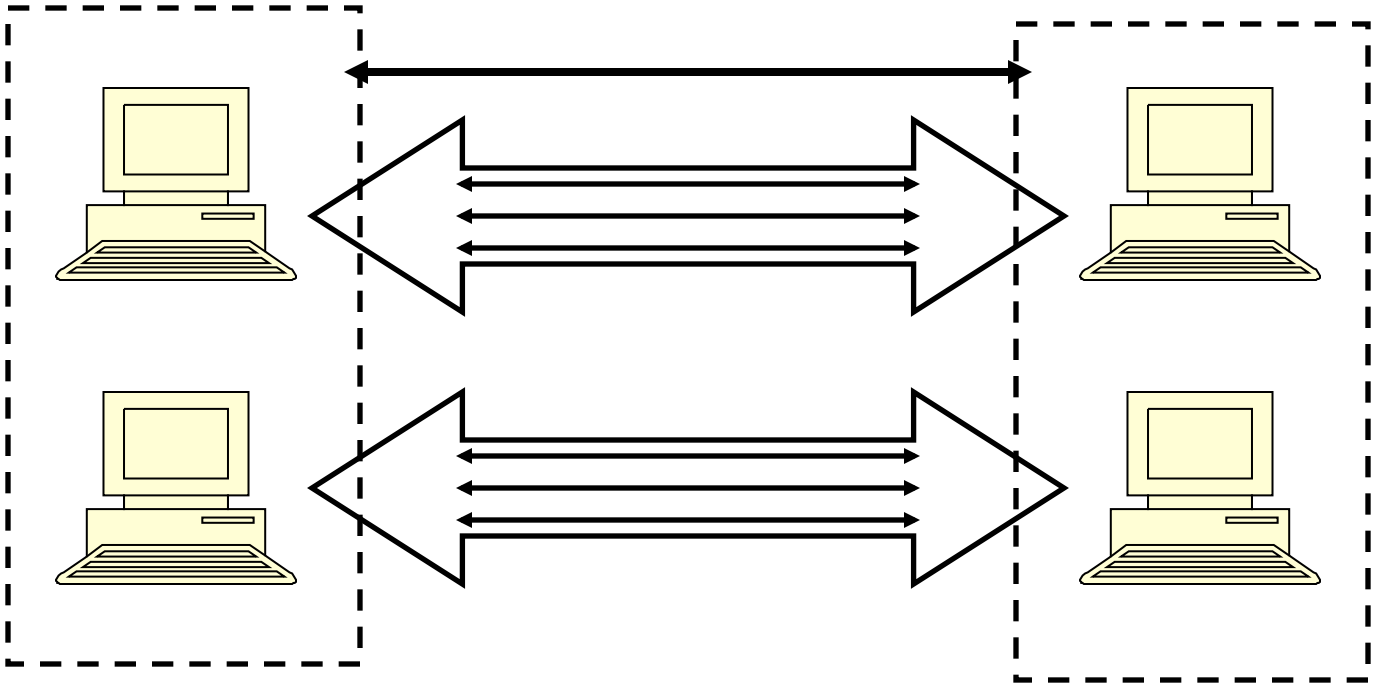
- A secure, robust, fast, efficient, standards based, widely accepted data transfer protocol
- Multiple independent implementations can interoperate
  - Both the Condor Project at Uwis and Fermi Lab have home grown servers that work with ours.
- Lots of people have developed clients independent of the Globus Project.
- We also supply a reference implementation:
  - Server
  - Client tools (globus-url-copy)
  - Development Libraries

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

Striping

- GridFTP offers a powerful feature called striped transfers (cluster-to-cluster transfers)



Topics for discussion

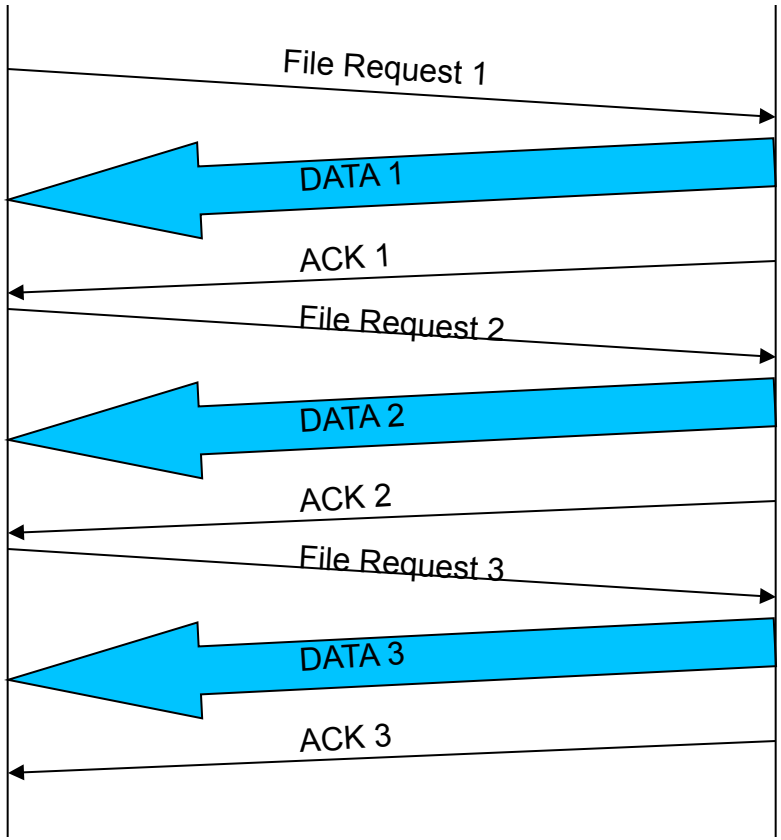
- Performance enhancements
  - LOSF problem and solution
  - GridFTP over UDT
- Ease of Use enhancements
  - Alternate security mechanism
  - GridFTP Where there's FTP
- Resource Management in GridFTP
- Future directions

Lots of Small Files (LOSF) Problem

- GridFTP and FTP - command response protocols
- A client can sends one command and then wait for a "Finished response" before sending another
- Overhead added on a per file basis
- Performance is best on large files
  - Overhead has less impact
- Performance suffers for a large data set partitioned into many small files

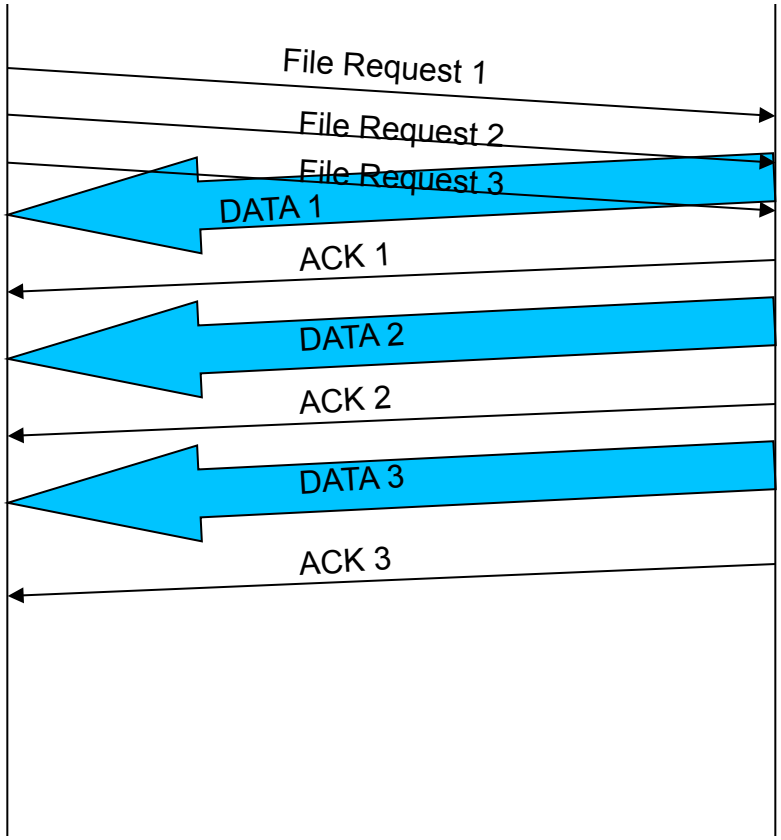
LOSF

- Traditional data transfer pattern



Pipelining

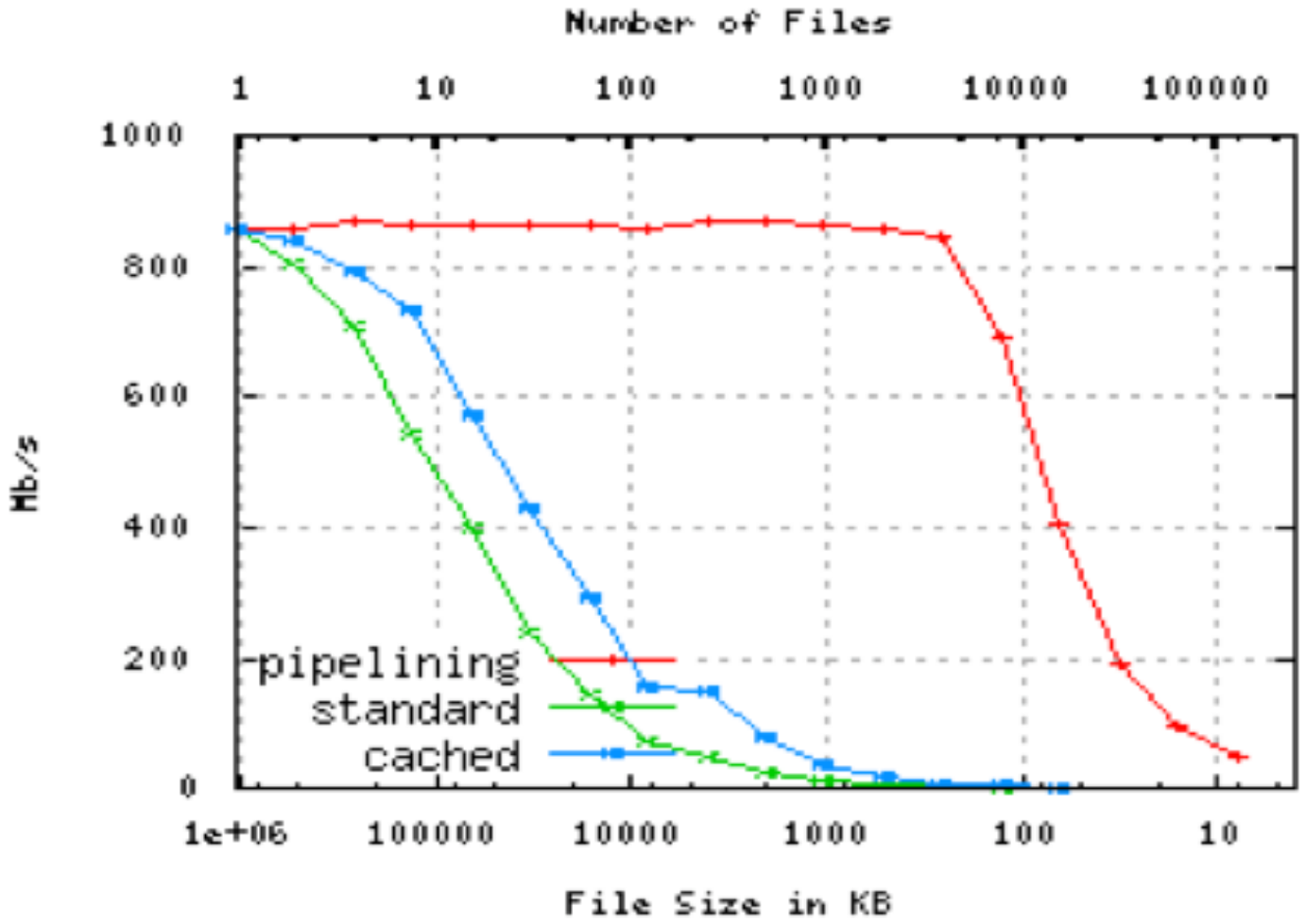
- Data transfer pattern with pipelining



the globus alliance
www.globus.org

Pipelining

WAN - SDSC and ANL with GSI security



GridFTP over UDT

- UDT is an application-level data transport protocol that uses UDP to transfer data
- Implement its own reliability and congestion control mechanisms
- Achieves good performance on high-bandwidth, high-delay networks where TCP has significant limitations
- GridFTP uses Globus XIO interface to invoke network I/O operations

GridFTP over UDT

- XIO framework presents a standard open/close/read/write interface to many different protocol implementations
  - including TCP, UDP, HTTP -- and now UDT
- The protocol implementations are called drivers.
  - A driver can be dynamically loaded and stacked by any Globus XIO application.
- Created an XIO driver for UDT reference implementation
- Enabled GridFTP to use it as an alternate transport protocol

the globus alliance
www.globus.org

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

Alternate security mechanism

- GridFTP traditionally uses GSI for establishing secure connections
- In some situations, preferable to use SSH security mechanism
- Leverages the fact that an SSH client can remotely execute programs by forming a secure connection with SSHD
  - The client (globus-url-copy) acts as an SSH client and remotely executes a Globus GridFTP server
  - All of the standard IO from the remote program is routed back to the client.

SSH security mechanism

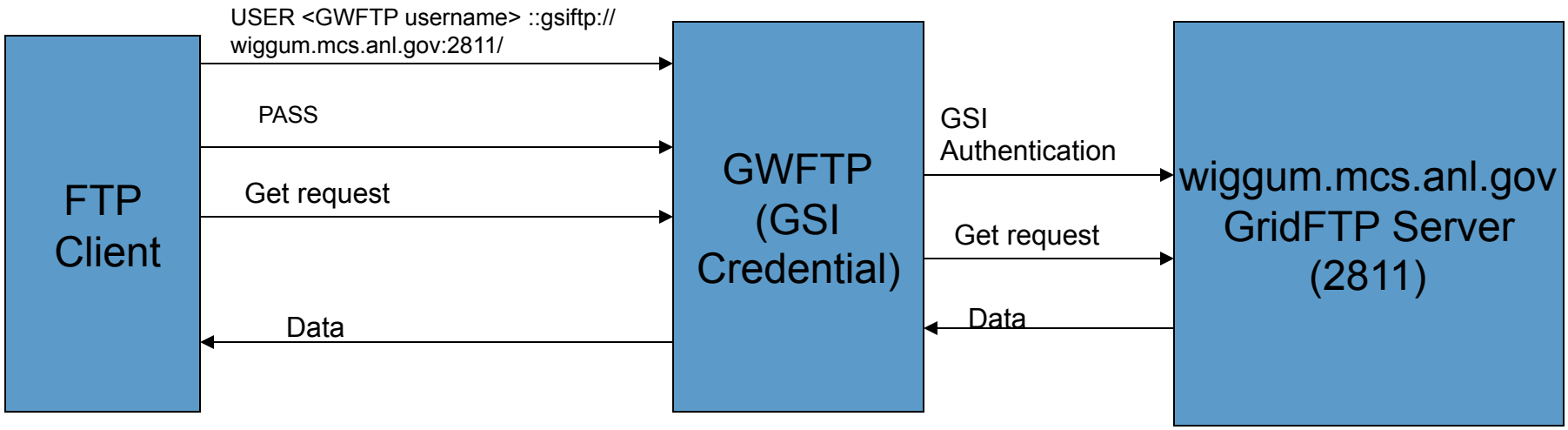
- Client support for using SSH is automatically enabled
- On the server side (where you intend the client to remotely execute a server)
  - setup-globus-gridftp-sshftp -server
- In order to use SSH as a security mechanism, the user must provide urls that begin with sshftp:// as arguments.
  - globus-url-copy 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 Where there's FTP (GWFTP)

- GridFTP has been in existence for some time and has proven to be quite robust and useful
- Only few GridFTP clients available
- FTP has innumerable clients
- GWFTP - created to leverage the FTP clients
- A proxy between FTP clients and GridFTP servers

the globus alliance
www.globus.org

GWFTP



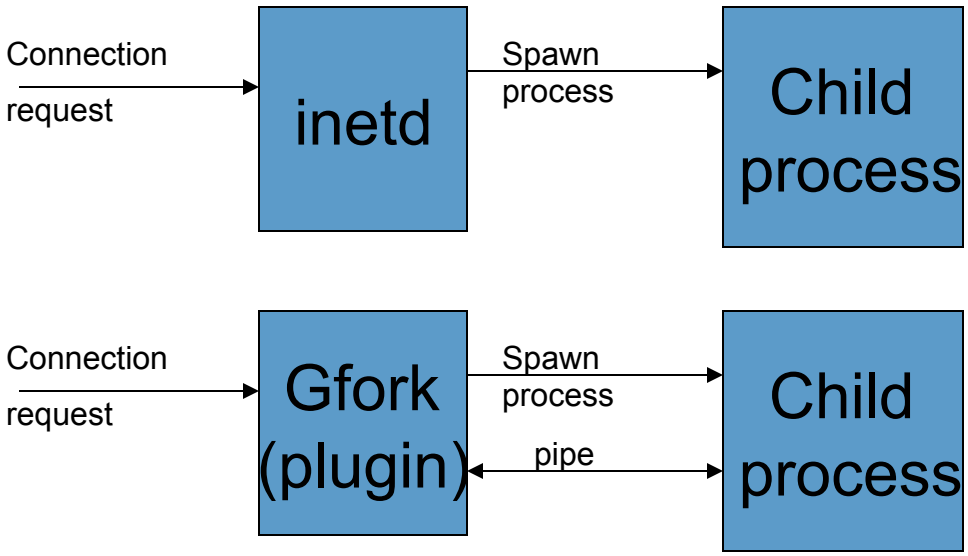
- Two security options provided with GWFTP to authenticate its client
  - Password based and host based

Resource management

- Under extreme loads it is possible that GridFTP servers require more memory than the system has and cause the system to fall over
- Developed a service called gfork to help avoid this situation
- Gfork - a service like inetd that listens on a TCP port and runs a configurable executable in a child process whenever a connection is made
- Associated with Gfork is a user defined master process
- Master process runs for the lifetime of the gfork daemon

Gfork

- GFork creates bi-directional pipes between the child processes and the master service.
- These pipes are used for communication between the child process and a master process plugin.



Resource Management

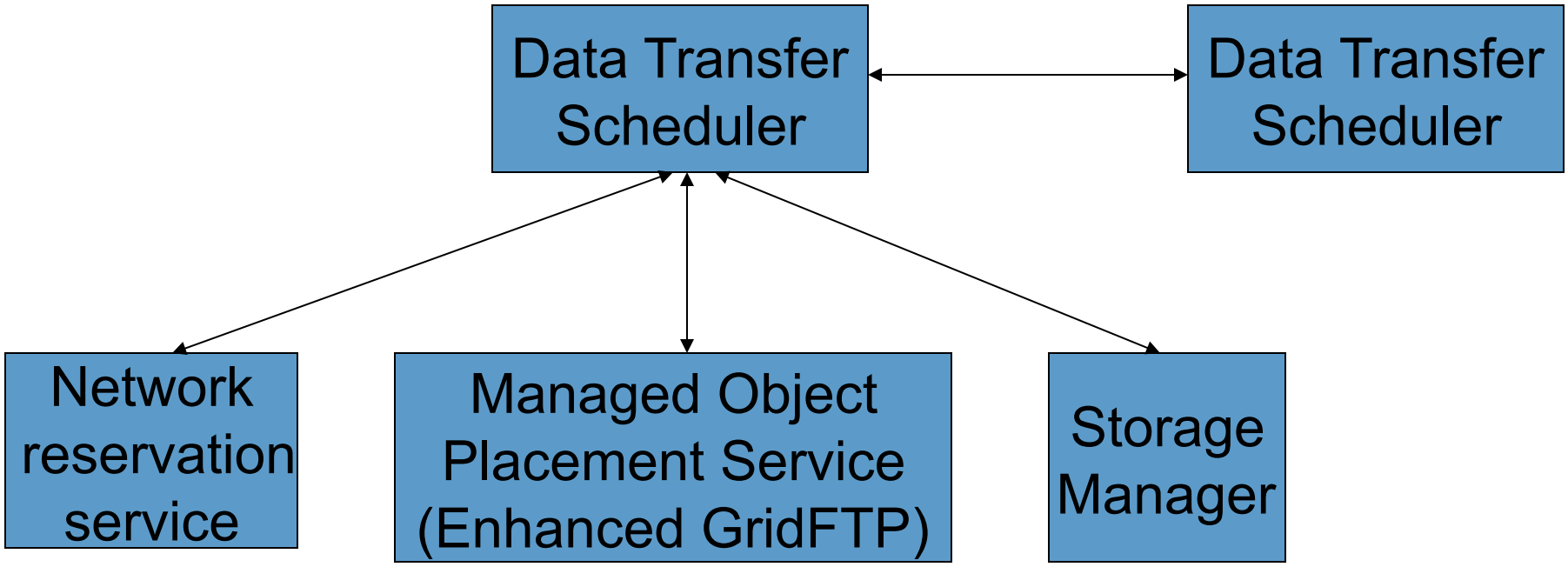
- Created a GridFTP Gfork plugin that has a memory limiting option
- Limit memory usage to a given value or to the maximum amount of RAM in the system.
- Most of the memory is given to the first few connections
- When the plugin detects that it is overloaded, each session is limited to half the available memory.

Data Transfer Scheduling

- Collaborating with Prof. Saday's group at Ohio State to develop a data transfer scheduling framework
- Enhance GridFTP servers to expose status as WS resource properties
- Take advantage of the network provisioning services
  - Collaborating with Terapath and LambdaStation projects

the globus alliance
www.globus.org

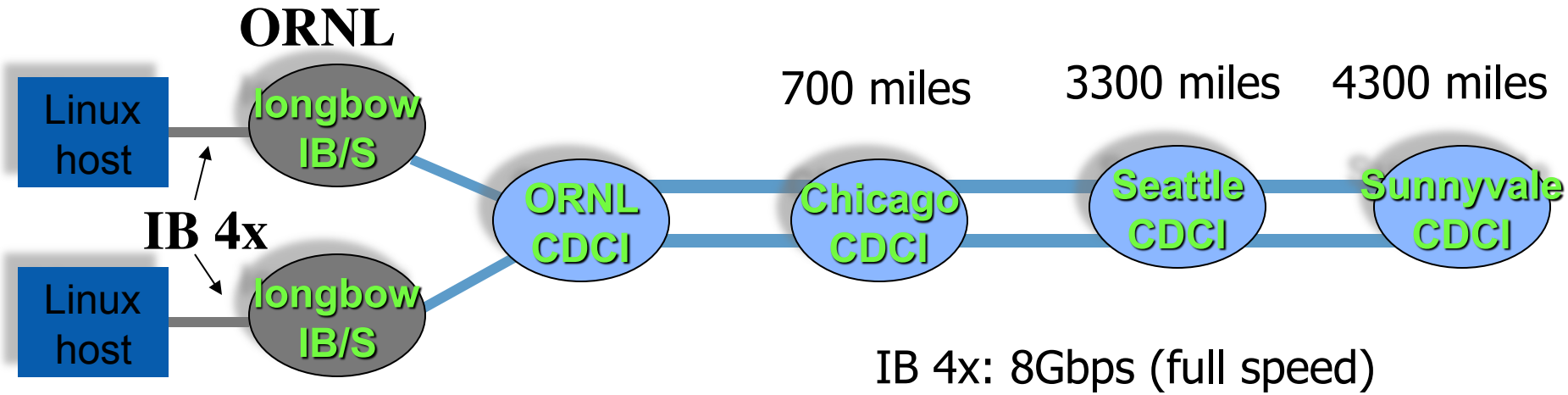
Data Transfer Scheduling



Infiniband Over SONET

Need specialized hardware: Obsidian longbow

1. IB over SONET/Ethernet – frame conversion
2. Buffer-based termination of IB flow control



IB 4x: 8Gbps (full speed)
Host-to-host local switch:7.5Gbps

ORNL loop -0.2 mile: **7.5**Gbps

ORNL-Chicago loop – 1400 miles: **7.46**Gbps

ORNL- Chicago - Seattle loop – 6600 miles: **7.23**Gbps

ORNL – Chicago – Seattle - Sunnyvale loop – 8600 miles: **7.20**Gbps

GridFTP over Infiniband

- Can use infiniband through Sockets Direct Protocol (SDP)
- SDP provides a socket interface for Open Fabrics software stack (a standard implemented by infiniband and iwarp)
  - No kernel bypass
- User level verbs to interface directly with infiniband hardware
  - Develop a XIO driver for verbs interface

the globus alliance
www.globus.org

GridFTP Tutorial at SC 2007

- When: Nov 12, 2007 8:30 AM - 12:00 PM
- Where: SC07 conference at Reno-Sparks Convention Center, Reno, NV
- The tutorial consists on hands-on exercises and demonstrations
  - Installation of the Globus GridFTP server and configuring the environment
  - Running the server in various security modes
  - Setting up the proper GSI security environment
  - Running client/server and 3rd party transfer
  - Performance optimizations
  - Setting up striped server & running striped transfers
  - GridFTP over UDT