



A GridFTP Transport Driver for Globus XIO

Raj Kettimuthu, Liu Wantao, Joe Link and
John Bresnahan

Argonne National Laboratory and
The University of Chicago

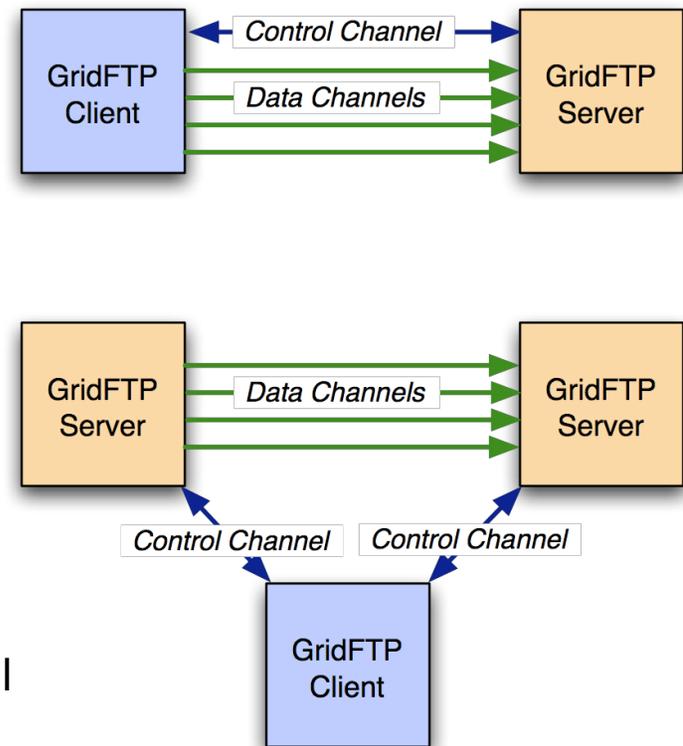


What is GridFTP?

- 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





GridFTP

- GridFTP is the de facto standard in projects requiring secure, robust, high-speed bulk data transport
- HEP community is basing its entire tiered data movement infrastructure for the LHC computing Grid on GridFTP
- LIGO routinely uses GridFTP to move 1 TB a day during production runs
- GridFTP is the recommended data transfer mechanism to maximize data transfer rates on the TeraGrid.



the globus alliance

www.globus.org

Common GridFTP clients

- Globus-url-copy - command line client
- UberFTP - interactive client
- CoG transfer client - graphical user interface client
- Globus Reliable File Transfer Service (RFT) - web service client
- Topaz enabled Firefox - a web browser client



A new client

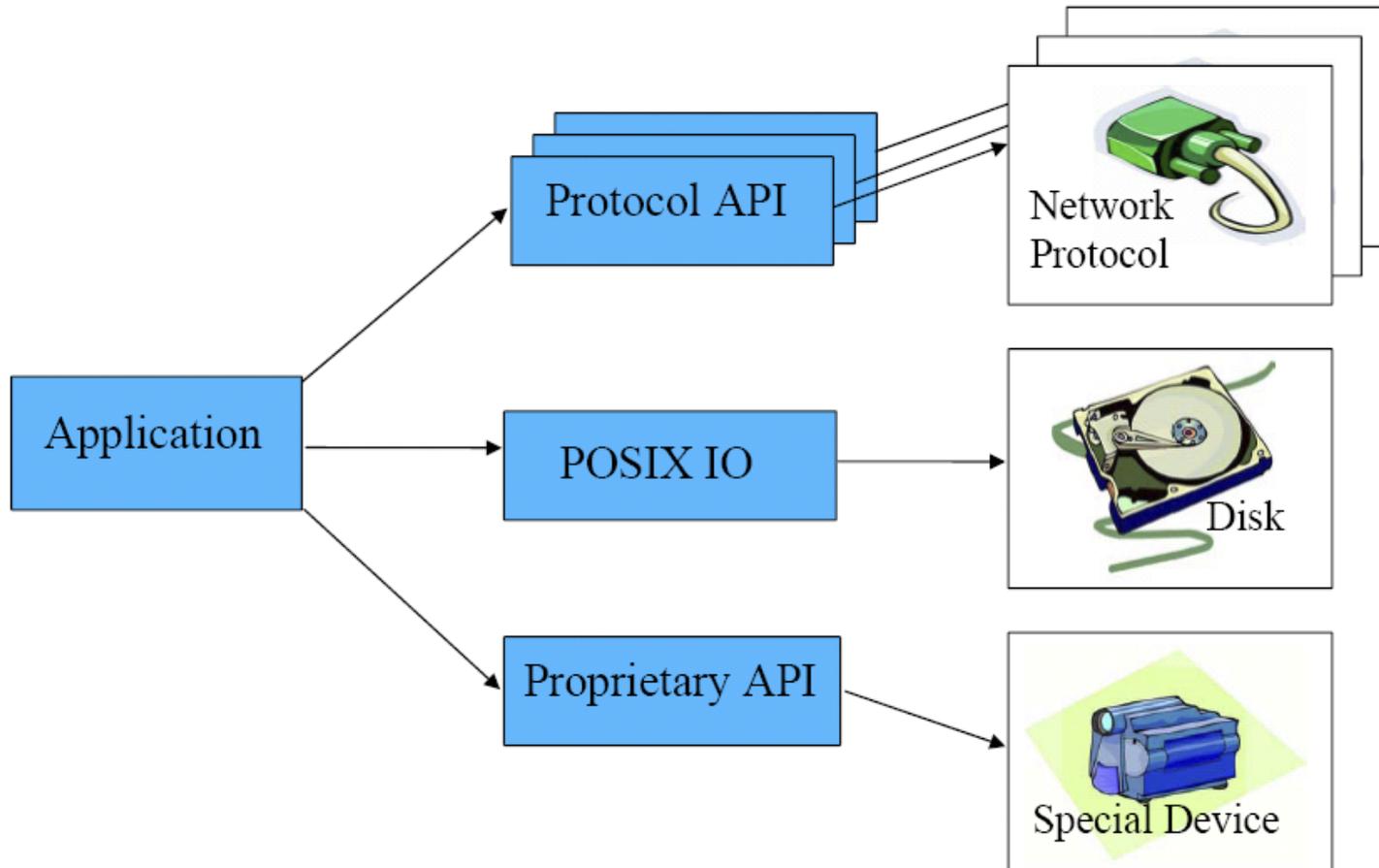
- Open/close/read/write (OCRW) interface
- Globus FTP client library provides an extensive interface
- OCRW - Simpler interface
- Eases the addition of GridFTP support to third-party programs, such as SRB and MPICH-G2
- OCRW is a familiar interface for applications to efficiently access remote files
- New client is based on Globus XIO



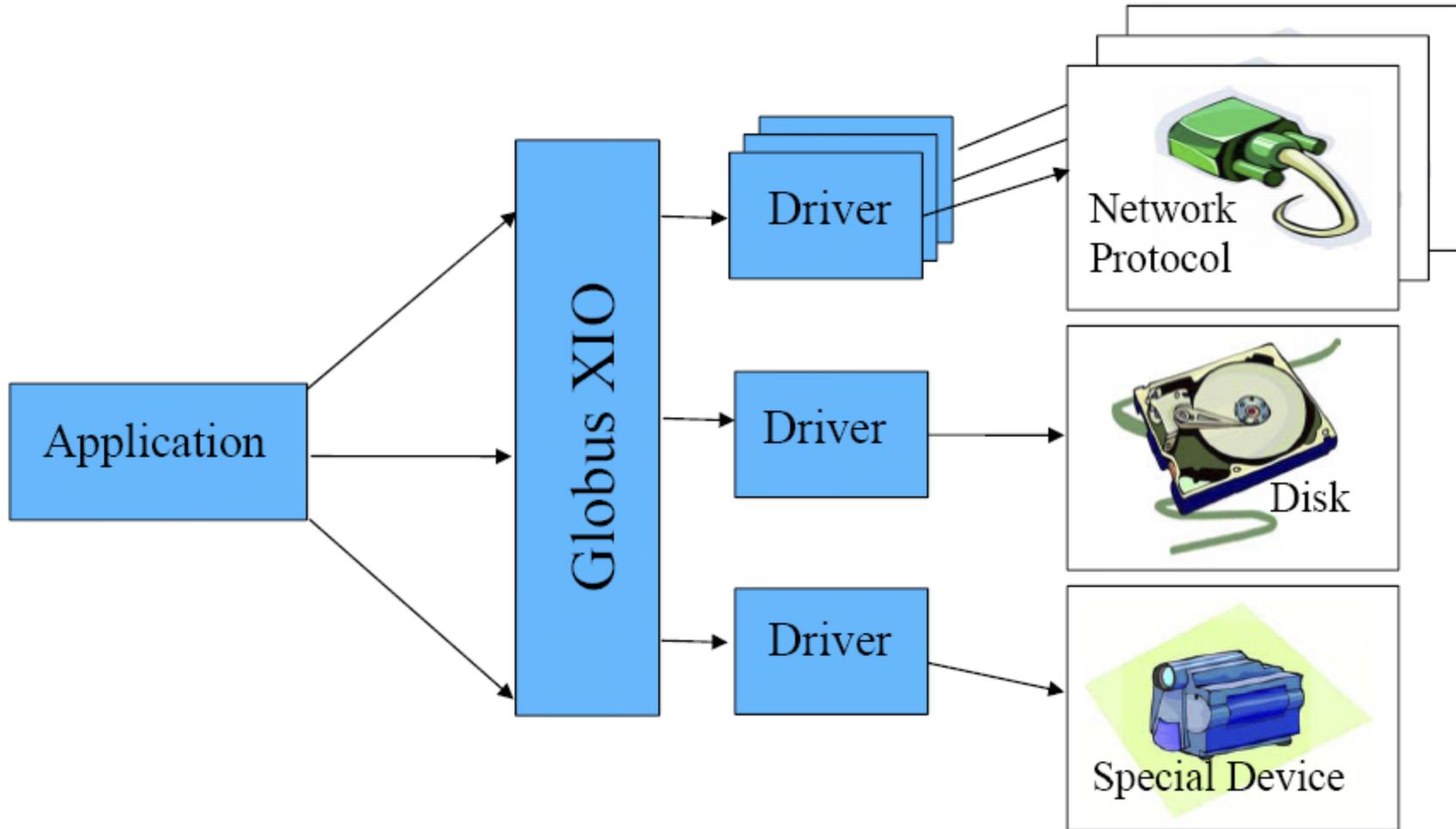
Globus XIO

- XIO framework presents a standard open/close/read/write interface to many different protocol implementations
 - ◆ Including TCP, UDP, HTTP, UDT and now GridFTP
- The protocol implementations are called drivers.
 - ◆ A driver can be dynamically loaded and stacked by any Globus XIO application

Typical Approach



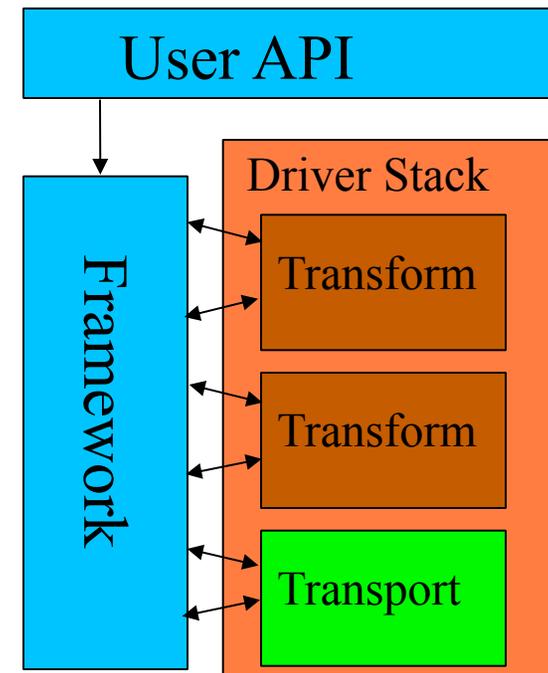
Globus XIO Approach





Globus XIO Framework

- Moves the data from user to driver stack
- Manages the interactions between drivers
- Transport
 - ◆ Moves data out of process space
 - ◆ TCP, UDP, File etc
- Transform
 - ◆ Manipulate or examine data
 - ◆ Do not move data outside of process space
 - ◆ Compression, security etc





GridFTP driver

- Transport driver
- For open/close/read/write calls, invoke one or more Globus FTP client library calls
- Support partial file transfers - seeking is allowed
- Allows users to set optimization parameters
 - ◆ TCP buffer size, Parallel streams etc

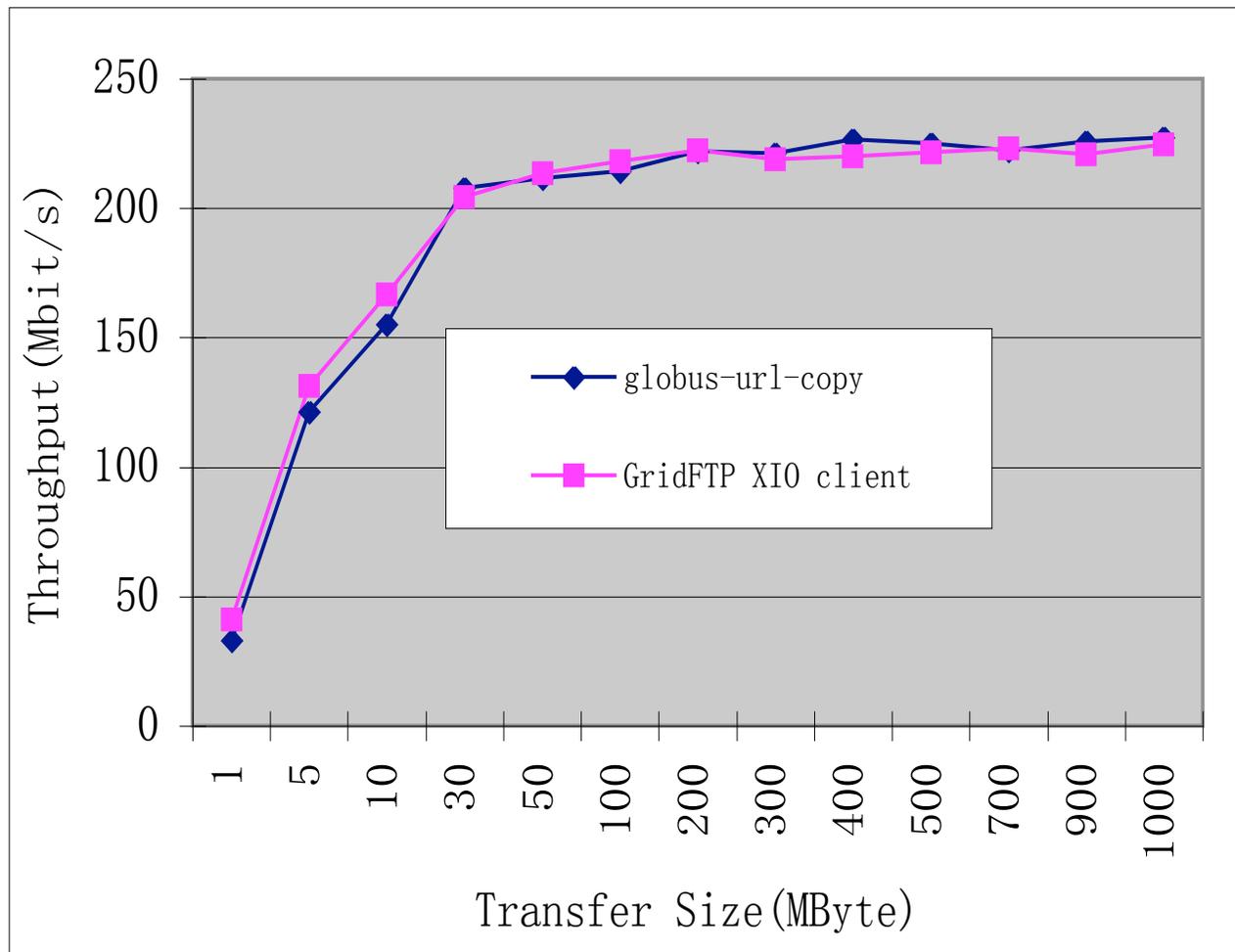


Experimental results

- Compare the performance of XIO client with that of globus-url-copy
 - ◆ Transferred files of various sizes on multiple pairs of sites on Teragrid
 - ◆ Network links with 4ms and 72ms RTTs
- Compare the performance of our new client with a wide area networked file system
 - ◆ GPFS WAN physically available at SDSC and is mounted at other sites on TeraGrid

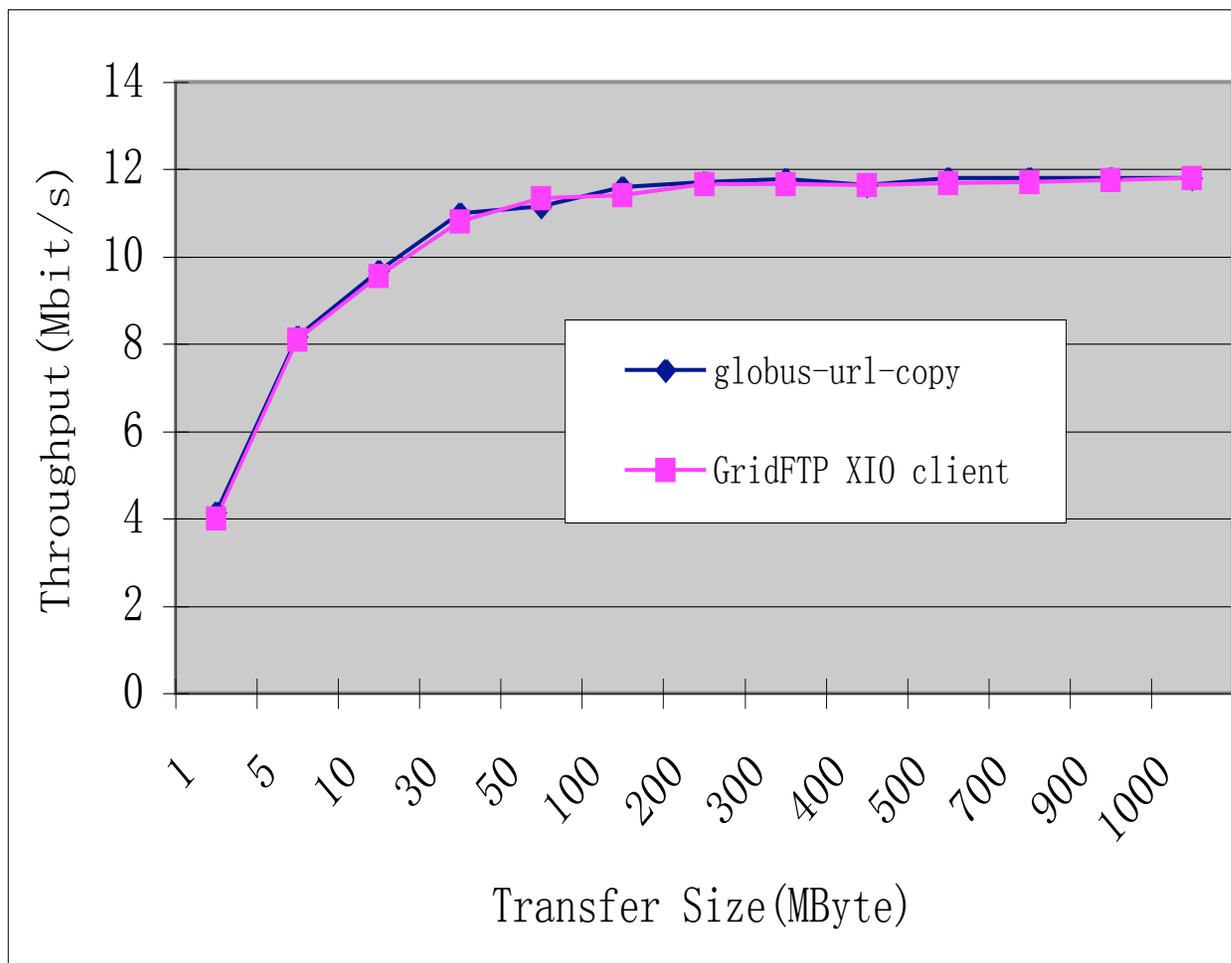


XIO client vs GUC - ANL and NCSA



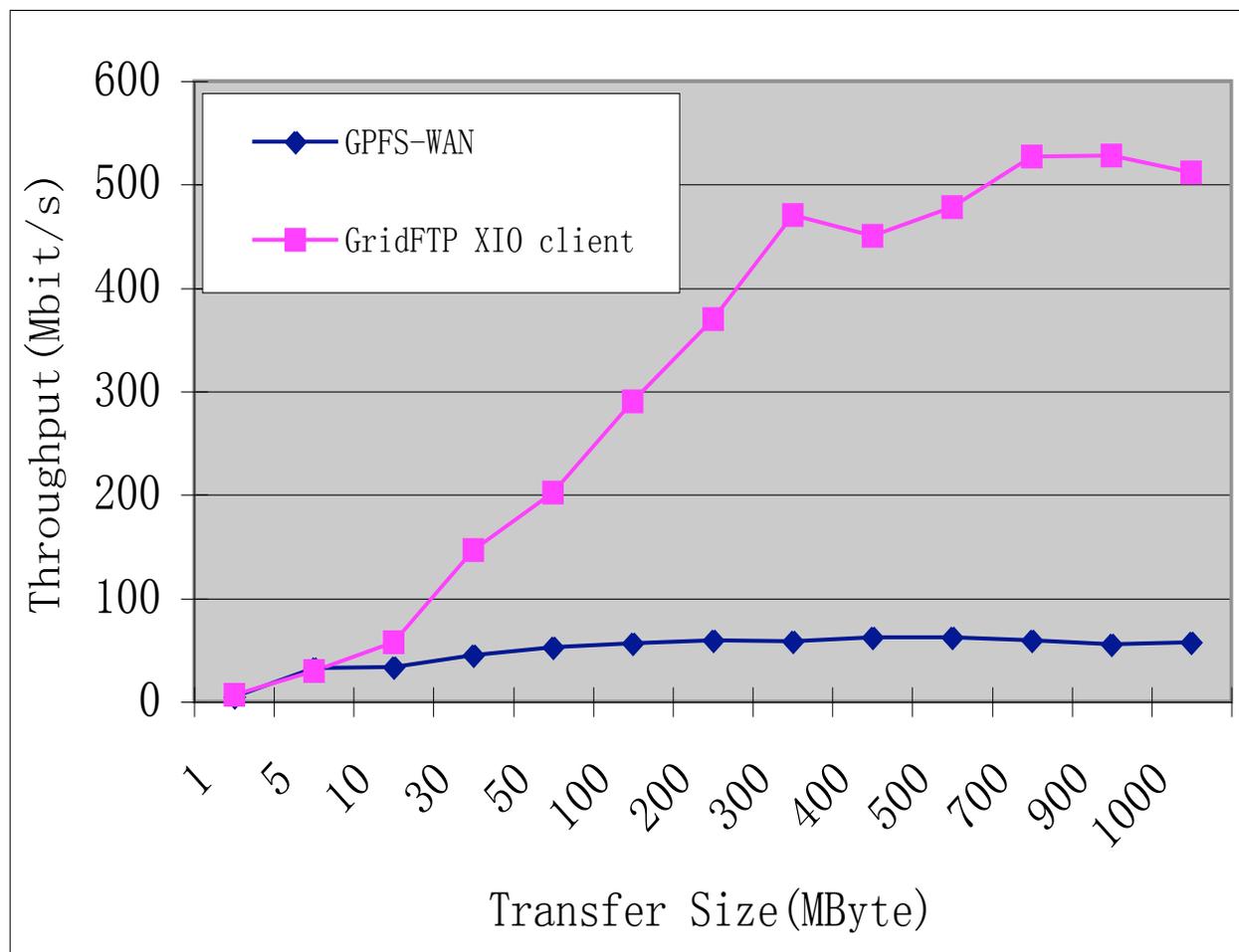


XIO client vs GUC - ORNL & SDSC





XIO client vs GPFS WAN at ANL

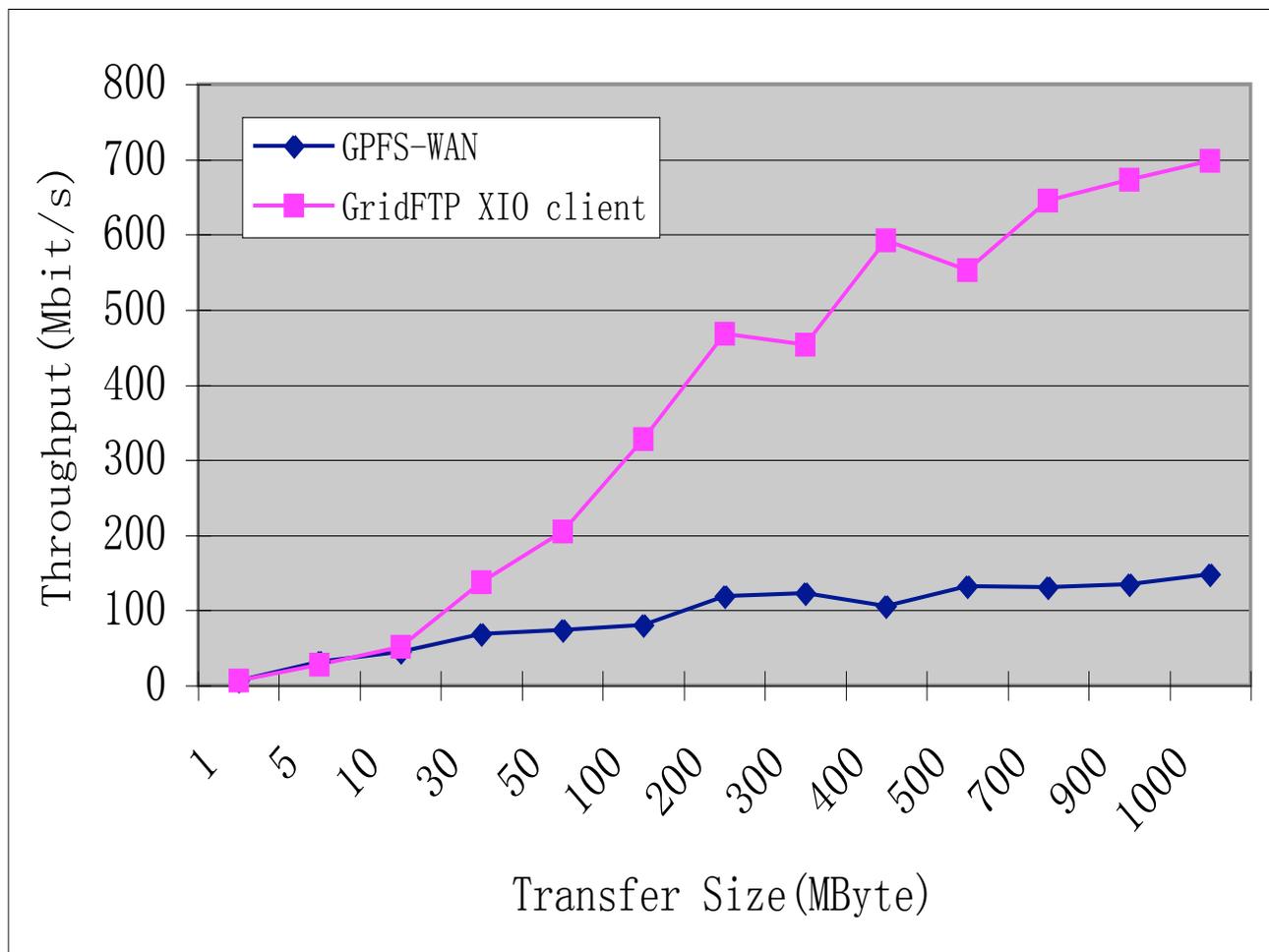




the globus alliance

www.globus.org

XIO client vs GPFS WAN at IU



Links and contacts

- Client is available in the Globus toolkit
 - ◆ www.globus.org
 - ◆ Source available in the source installer at `source-trees/gridftp/gridftp_driver`
 - ◆ Xio-user@globus.org, xio-dev@globus.org
 - ◆ Gridftp-user@glou.org,
gridftp-dev@globus.org

Questions