Explaining Wide Area Data Transfer Performance

Rick Wagner
rick@globus.org

University of Chicago
Argonne National Laboratory

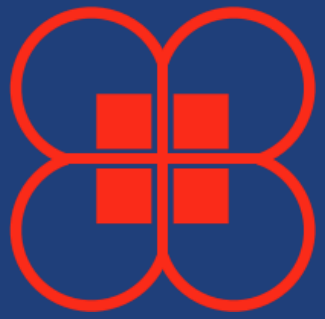
Data Over Distance – July 19, 2018

Publication

Explaining Wide Area Data Transfer Performance

Zhengchun Liu
Rajkumar Kettimuthu

Prasanna Balaprakash
Ian Foster



The 26th International ACM Symposium on High-Performance Parallel and Distributed Computing (HPDC)
HPDC '17, June 26-30, 2017, Washington, DC, USA

26th HPDC

https://doi.org/10.1145/3078597.3078605

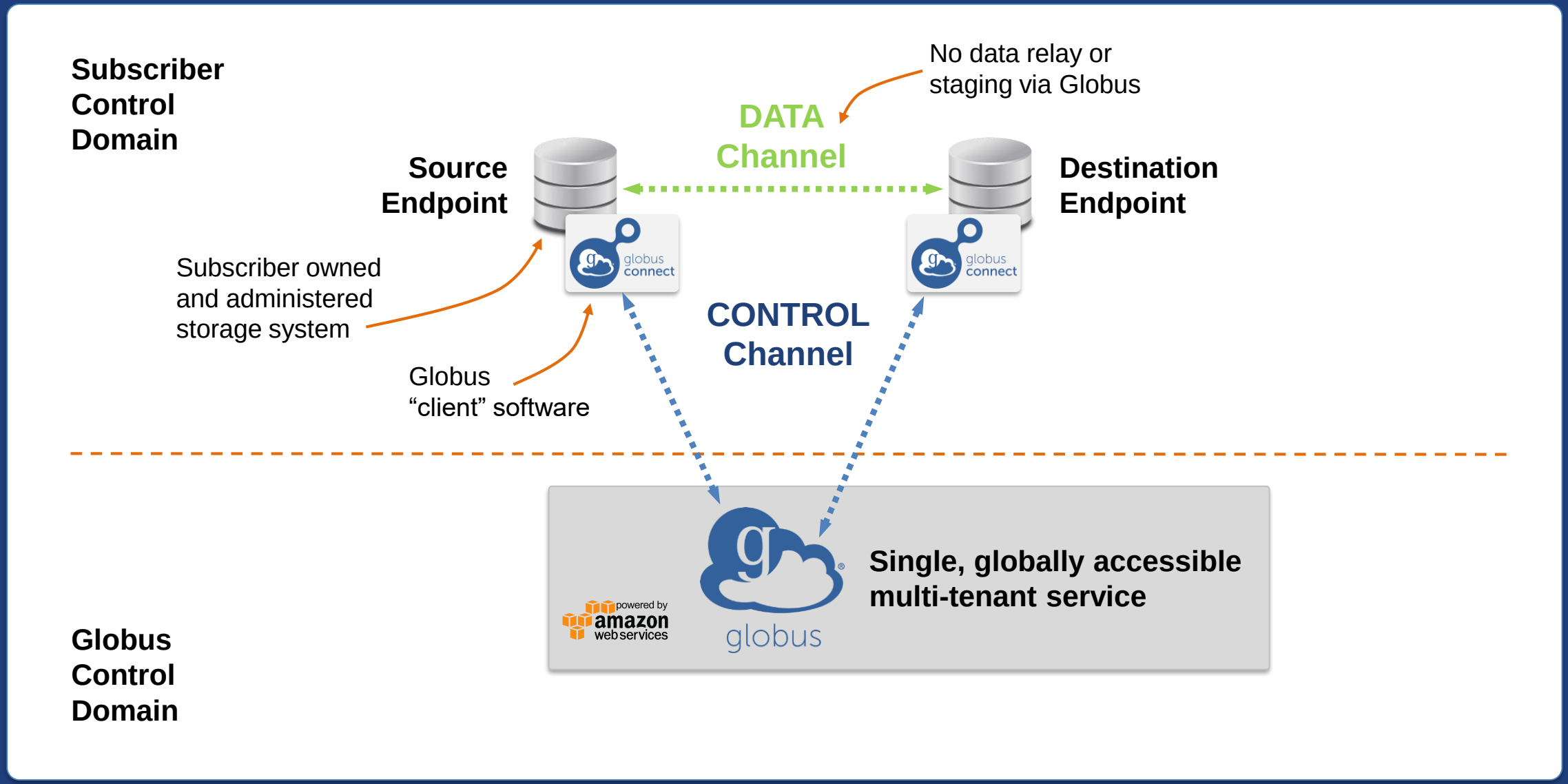
Argonne NATIONAL LABORATORY

Motivation

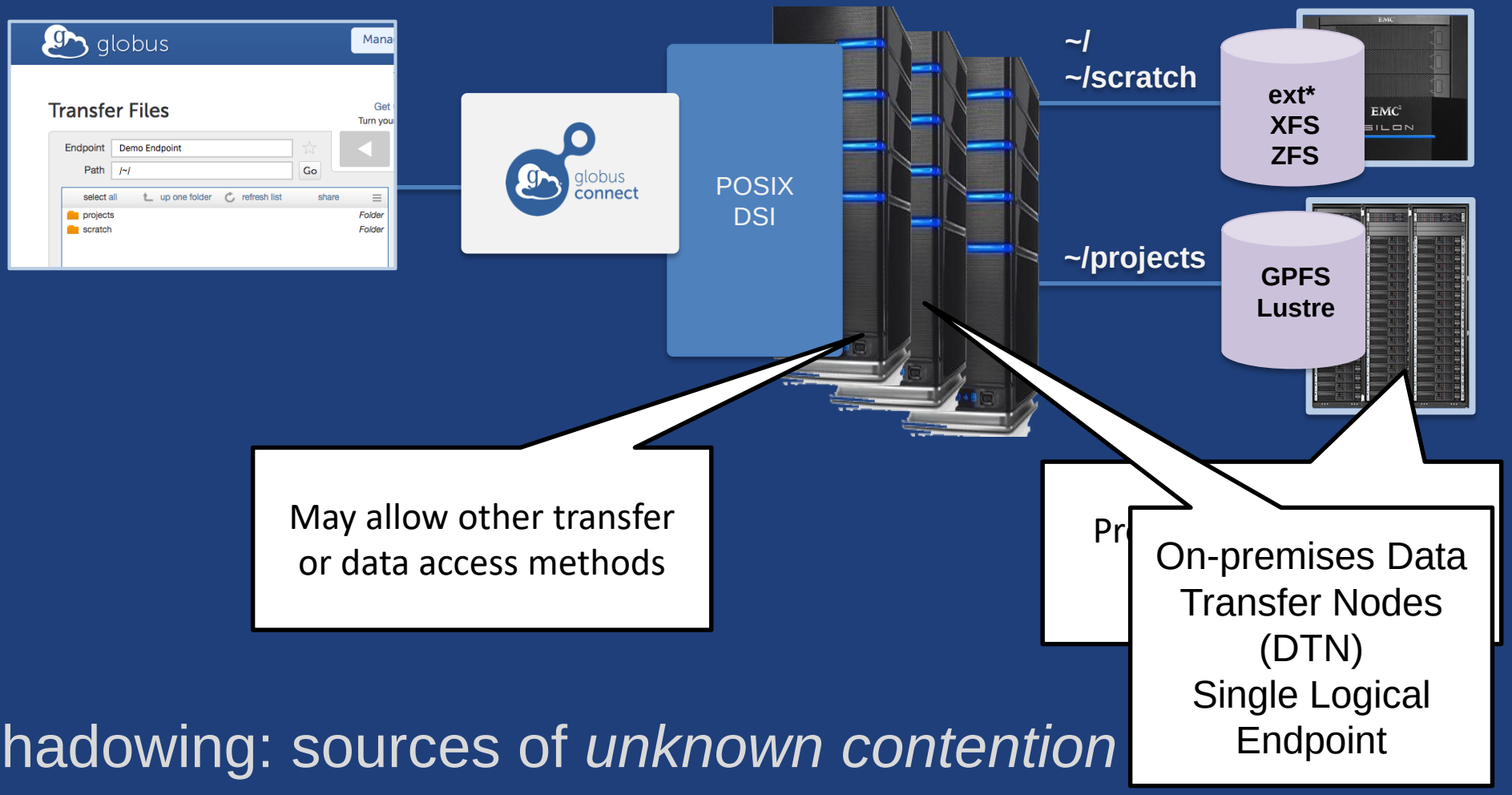
Armed with a large collection of Globus wide-area file transfer records, and experiments performed in the ESnet testbed environment, we want to:

- Extract factors that affect the transfer performance based on domain knowledge, and study their importance **(explanation)**;
- Build models to predict transfer performance **(prediction)**;
- Model based performance optimization **(optimization, future work)**.

Globus Conceptual architecture: Hybrid SaaS



Endpoints & Storage

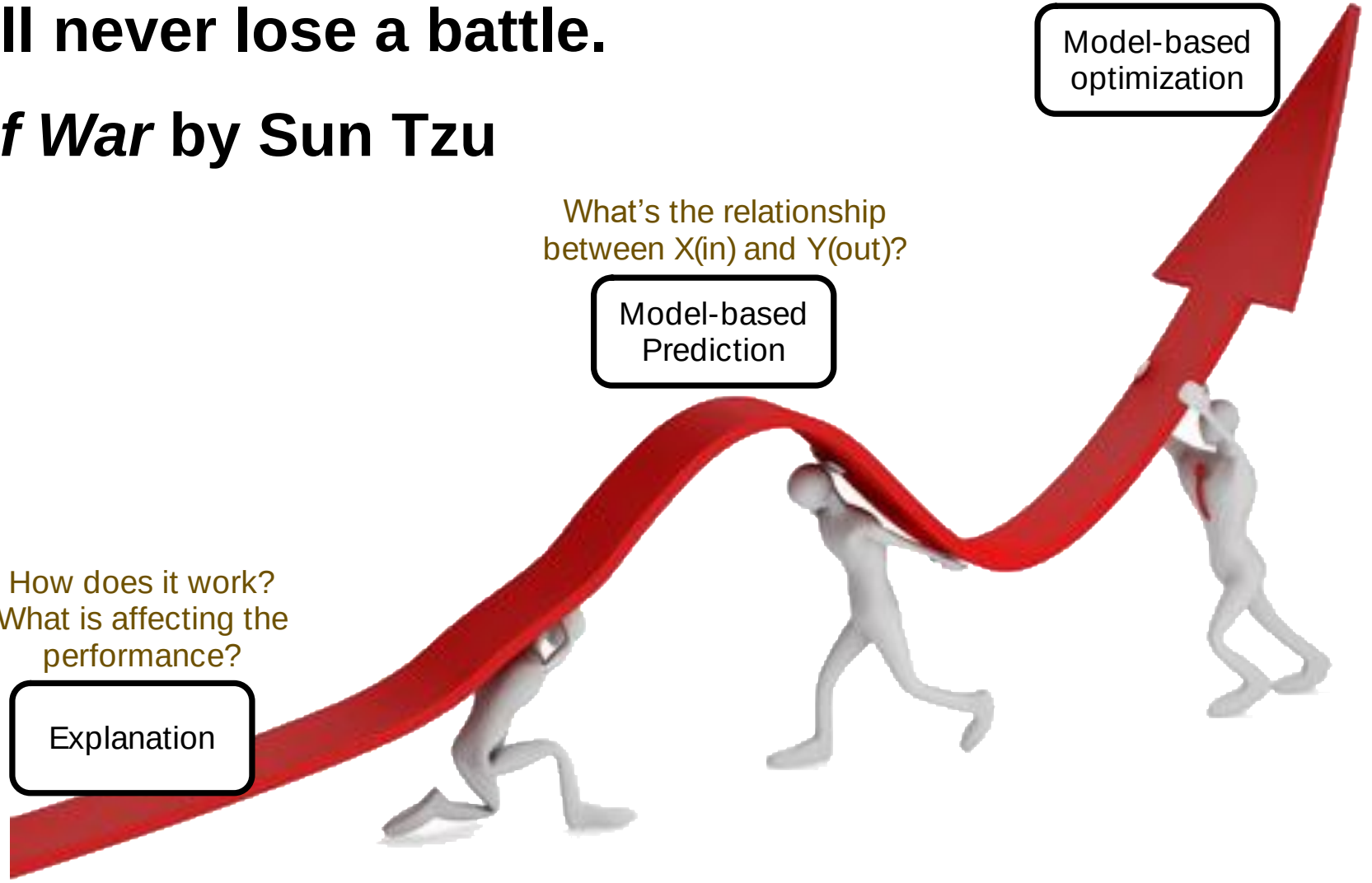


Foreshadowing: sources of *unknown contention*

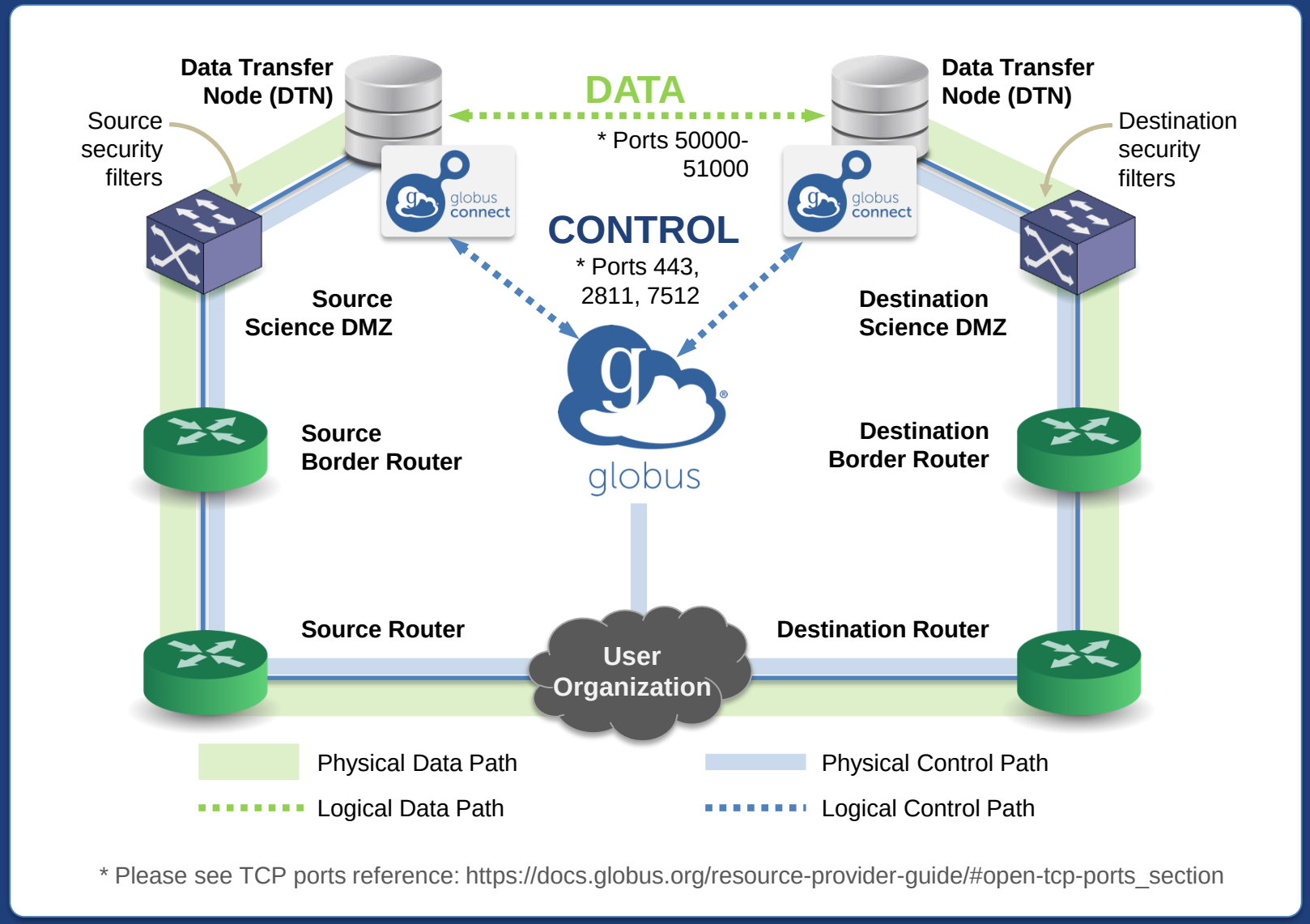
Motivation

If you know yourself and your enemy, you'll never lose a battle.

— *The Art of War* by Sun Tzu

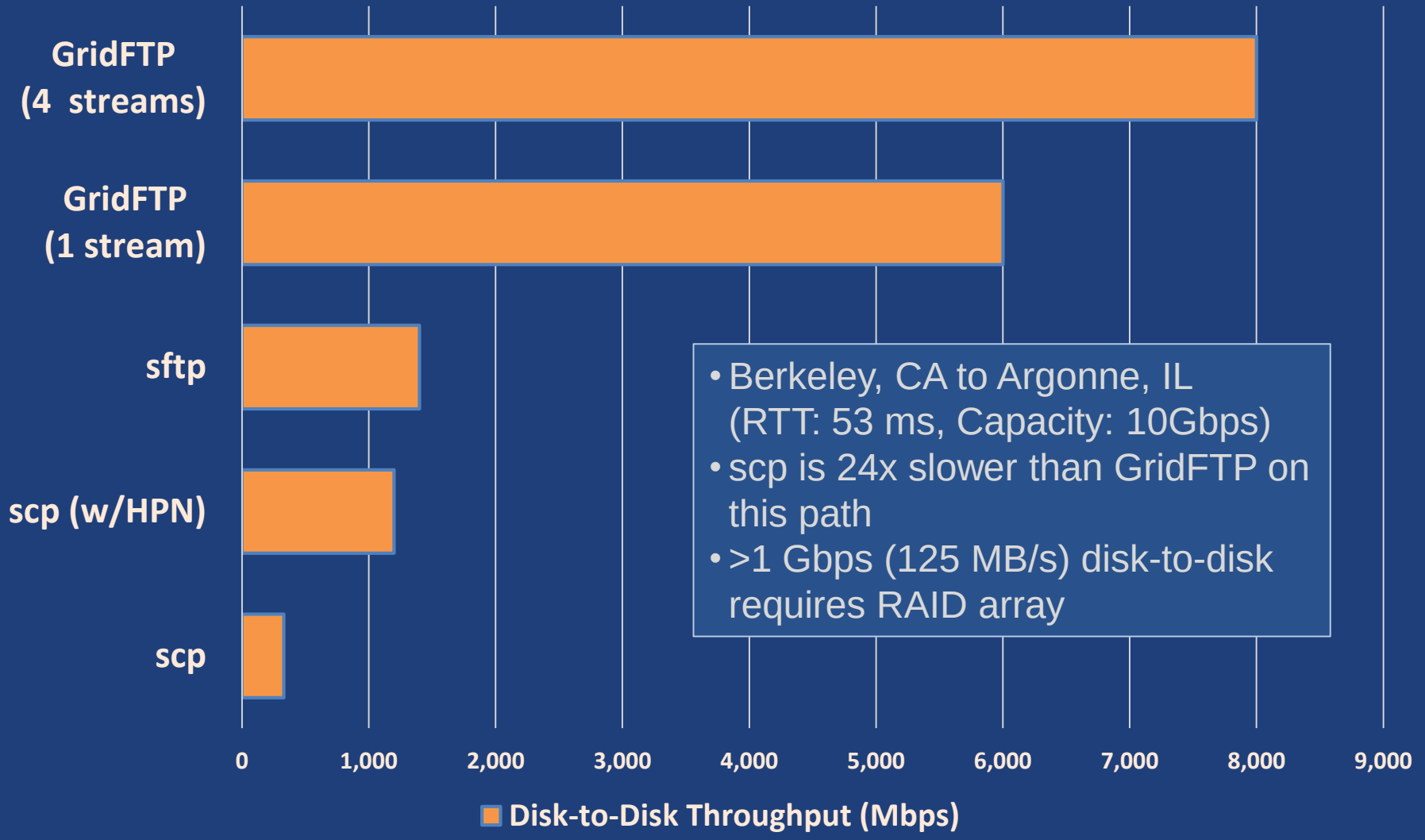


Best practice network configuration



* Please see TCP ports reference: https://docs.globus.org/resource-provider-guide/#open-tcp-ports_section

Disk-to-Disk Throughput: ESnet Testing



Source: ESnet (2016)

What affects transfer performance?

Start with 4 parameters (3 known, 1 unknown)

For a given endpoint pair:

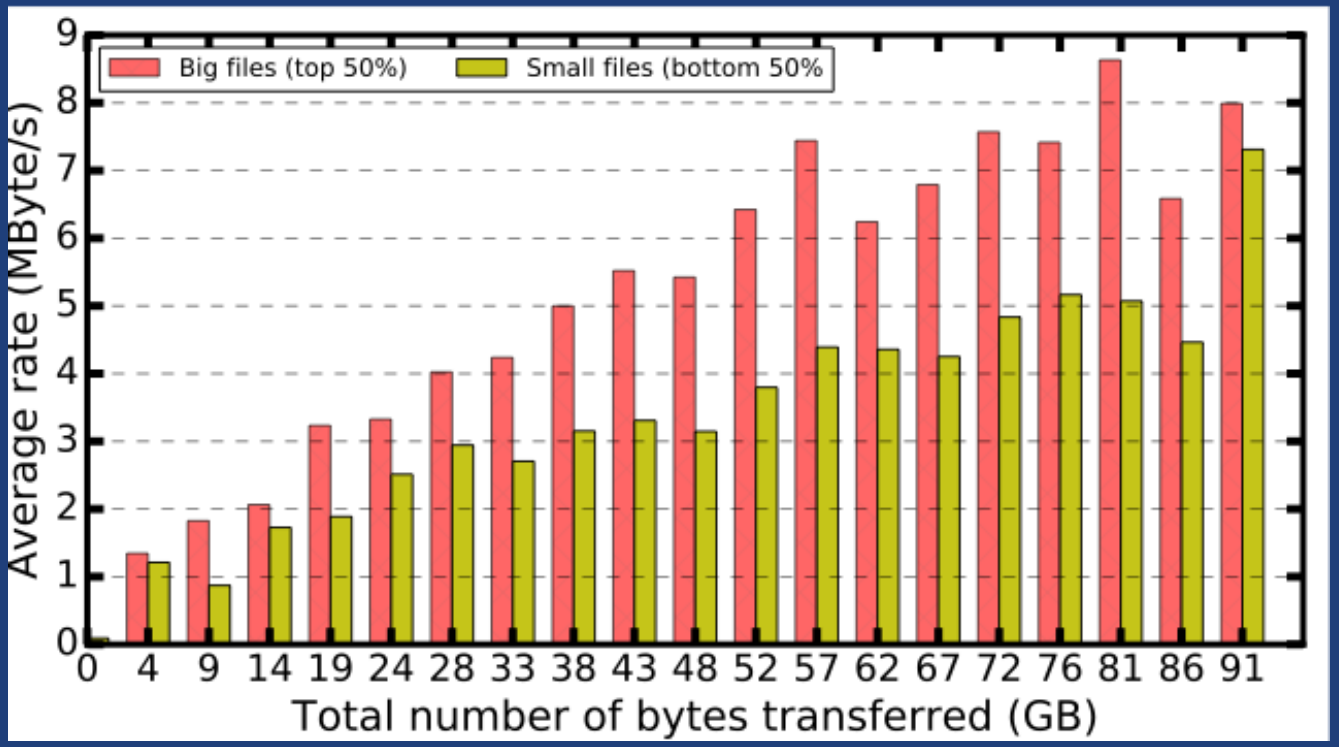
KNOWN
- Transfer file characteristic, e.g., file size

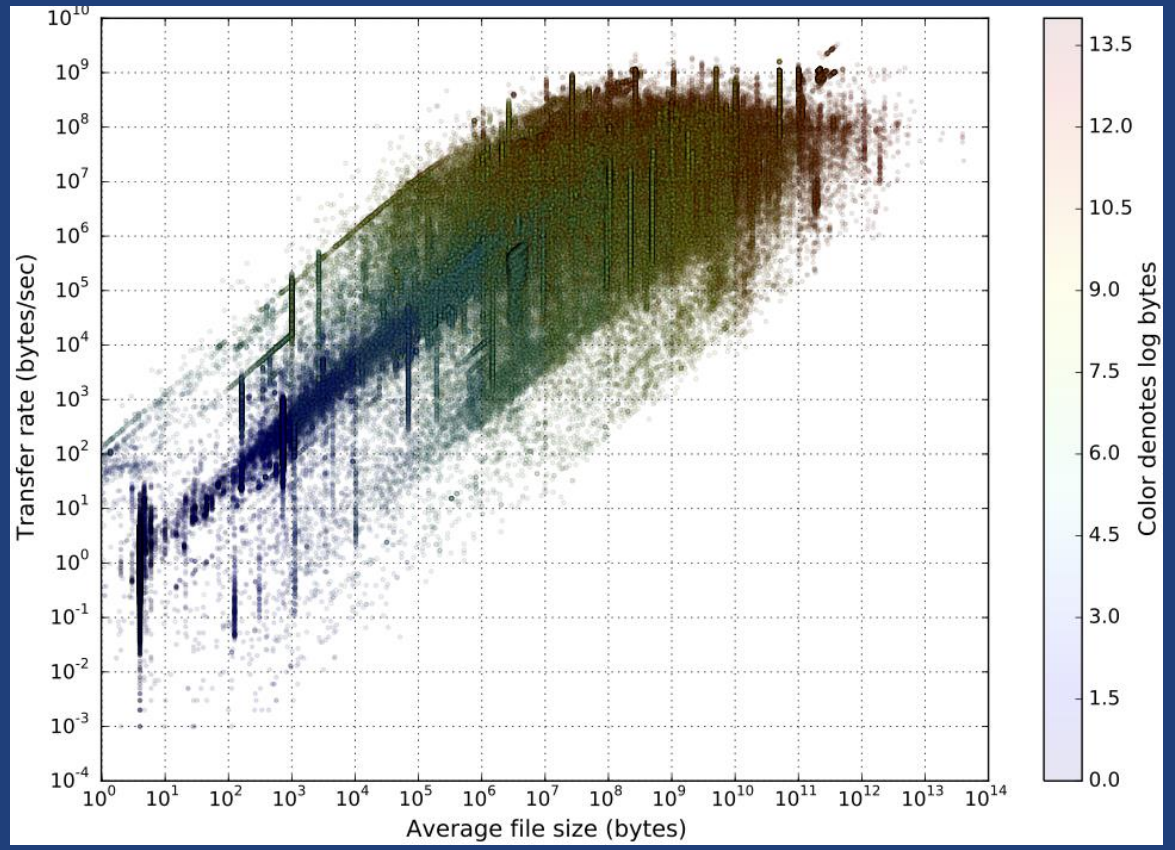
KNOWN
- Tunable transfer parameters, e.g., concurrency (flying files), parallelism

KNOWN
- Contentions from other simultaneous Globus transfers

????
- Contentions from other programs, e.g., sharing file system, network

E.g., larger transfer: better performance

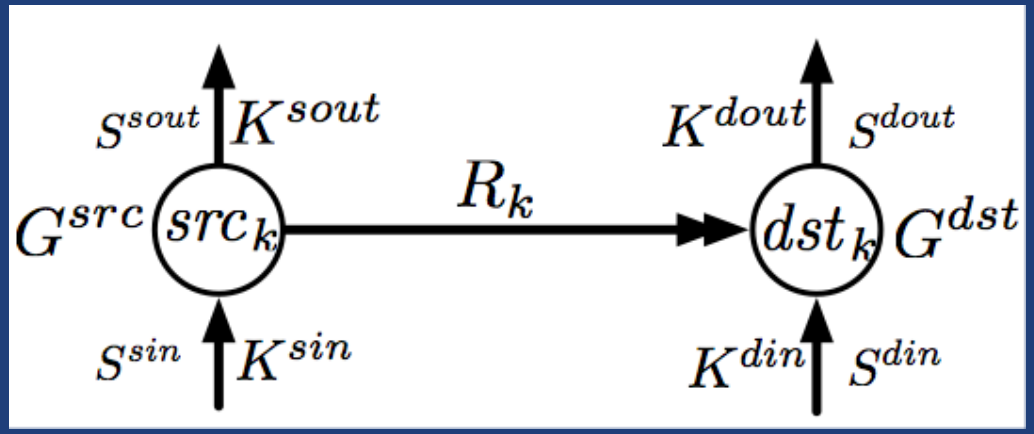




Large transfers with big average file size are more likely to have better performance.

I.e., the startup cost is high.

Modeling Simultaneous Globus Transfers



Load experienced by a Globus transfer *k* from src_k to dst_k with rate R_k

Relative external load:

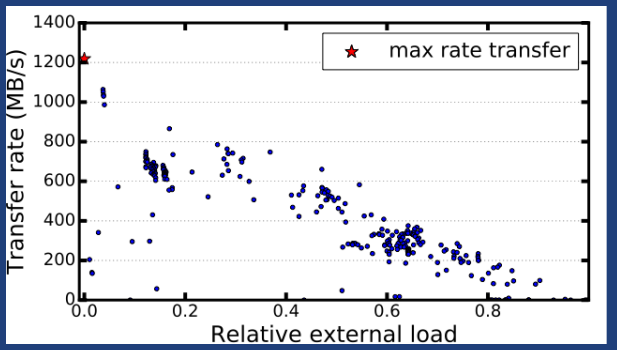
$$ReL = max\left(\frac{K^{sout}}{R_k + K^{sout}}, \frac{K^{din}}{R_k + K^{din}}\right)$$

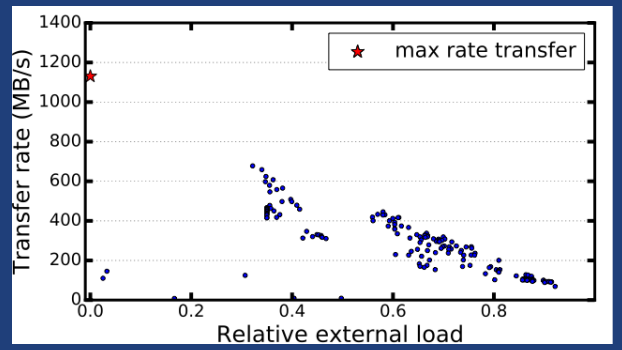
Features to explain a transfer

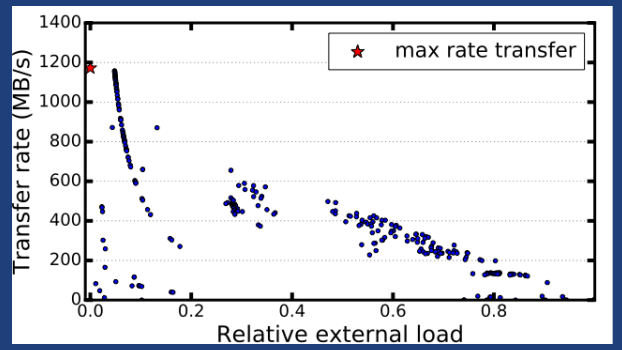
K^{sin}	Contending incoming transfer rate on src_k.
K^{sout}	Contending outgoing transfer rate on src_k.
K^{din}	Contending incoming transfer rate on dst_k.
K^{dout}	Contending outgoing transfer rate on dst_k.
C	Concurrency: Number of GridFTP processes.
P	Parallelism: Number of TCP channels per process.
S^{sin}	Number of incoming TCP streams on src_k.
S^{sout}	Number of outgoing TCP streams on src_k.
S^{din}	Number of incoming TCP streams on dst_k.
S^{dout}	Number of outgoing TCP streams on dst_k.
G^{src}	GridFTP instance count on src_k.
G^{dst}	GridFTP instance count on dst_k.
Nf	Number of files transferred.
Nd	Number of directories transferred.
Nb	Total number of bytes transferred.

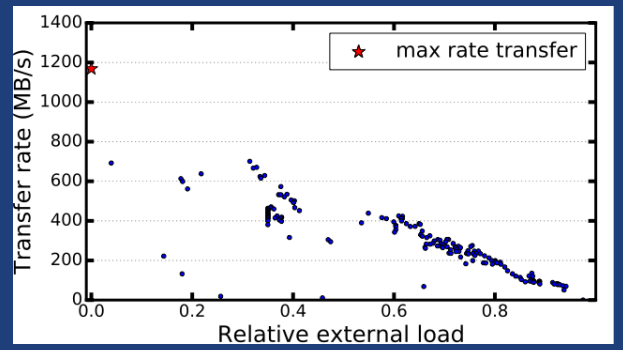
Unknown Contention Matters

Hard to Quantify

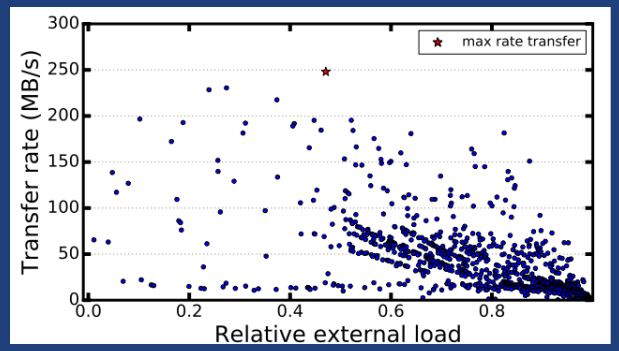


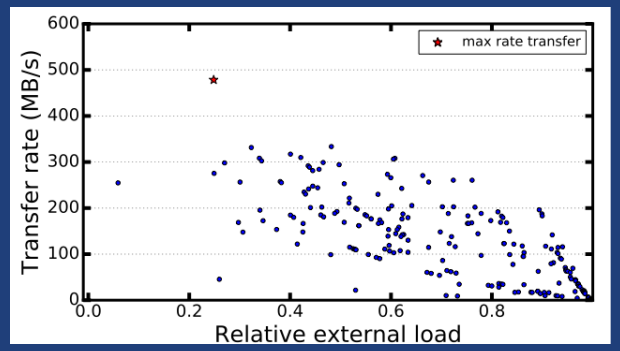


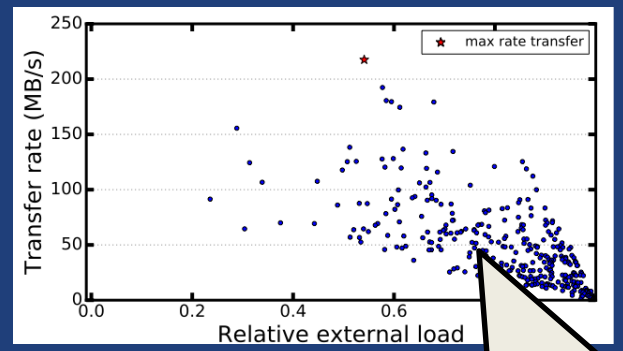


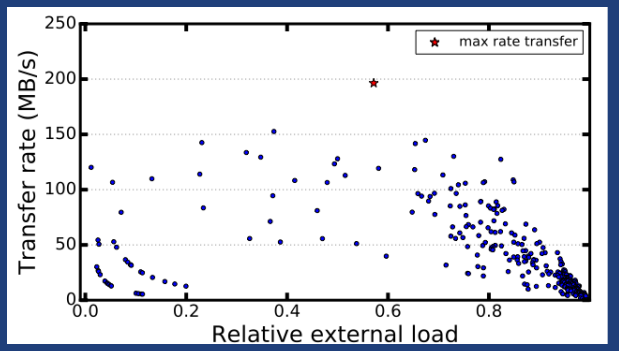


Transfers over ESnet testbed
(less likely to have non-Globus load on endpoints)







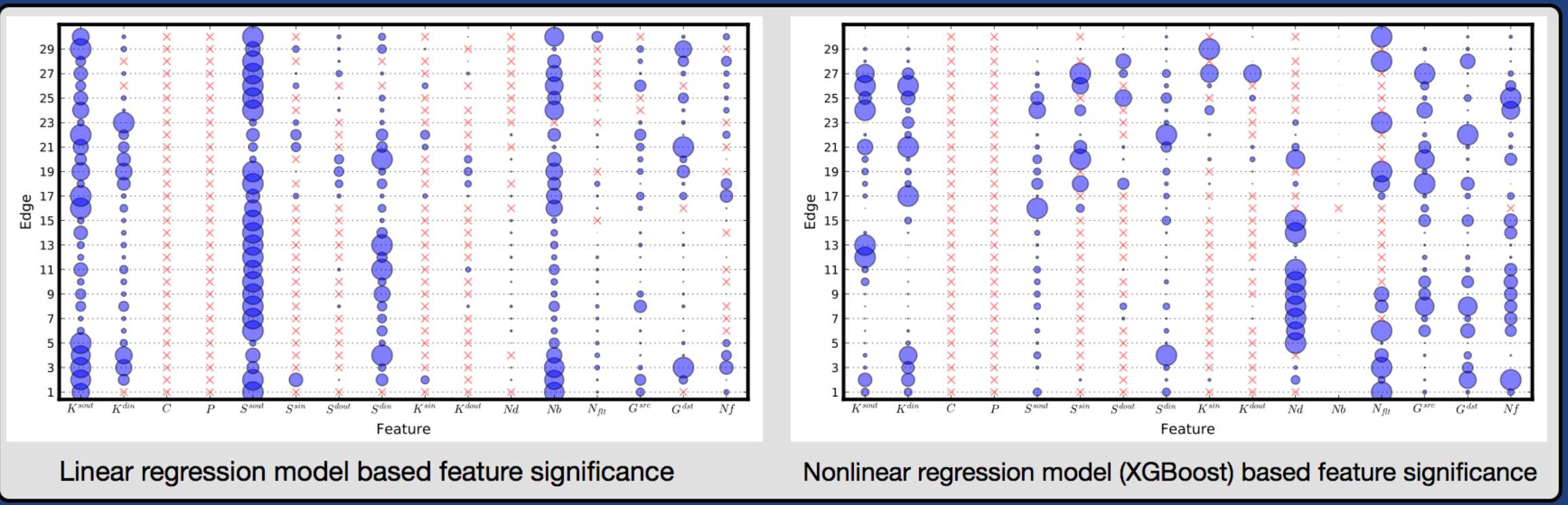


Transfer over production DTN
(more likely to have non-Globus load on endpoints)

Noisy: Small files?
File system load?

Machine Learning Results

Resource contention at endpoint is clear: K^{sout}, K^{din}, S^{sout} and S^{din} are significant in the models. Total transfer bytes also matters, means that the startup cost is high.



Circle size indicates the relative significance of features in the linear model, for each of 30 edges. A red cross means that the corresponding feature is eliminated because of low variance.

Applicability to other tools

These methods apply to non-Globus WAN transfers, too

- The data (e.g. # of TCP connections, # of concurrent files, transfer size, # of files) related to all WAN transfers, regardless of the tool
- The data can be obtained from other transfer tools. E.g: FTP; rsync; scp; BBCP; FDT; and XDD.

Thank you

The authors would also like to thank:

- U.S. Department of Energy, Office of Science, ASCR, and the program manager *Richard Carlson*;
- *Nagi Rao* for useful discussions
- *Brigitte Raumann* for help with Globus log analysis;
- *Glenn Lockwood* for help with experiments at NERSC.
- *A. A. Otto* for the espresso

Thank you to our sponsors…

U.S. DEPARTMENT OF
ENERGY

THE UNIVERSITY OF
CHICAGO

NIST
National Institute of
Standards and Technology
U.S. Department of Commerce

Argonne
NATIONAL LABORATORY

Thank you to our users…

48
most server endpoints at a single organization

384 PB
transferred

64 billion
tasks processed

76,000
registered users

500
100TB+ users

14,000
active users

3 months
longest running managed transfer

10,000
active endpoints

350+
federated identities

1 PB
largest single transfer to date

5,000
active shared endpoints

99.5%
uptime

Our supporters

Join the Globus community

- Access the service: **globus.org/login**
- Create a personal endpoint: **globus.org/app/endpoints/create-gcp**
- Documentation: **docs.globus.org**
- Engage: **globus.org/mailing-lists**
- Subscribe: **globus.org/subscriptions**
- Need help? **support@globus.org**
- Follow us: **@globusonline**