Monitoring and Accelerating GridFTP

Martin Swany, Indiana U.

Dan Gunter, LBL

Jason Zurawski, Internet2
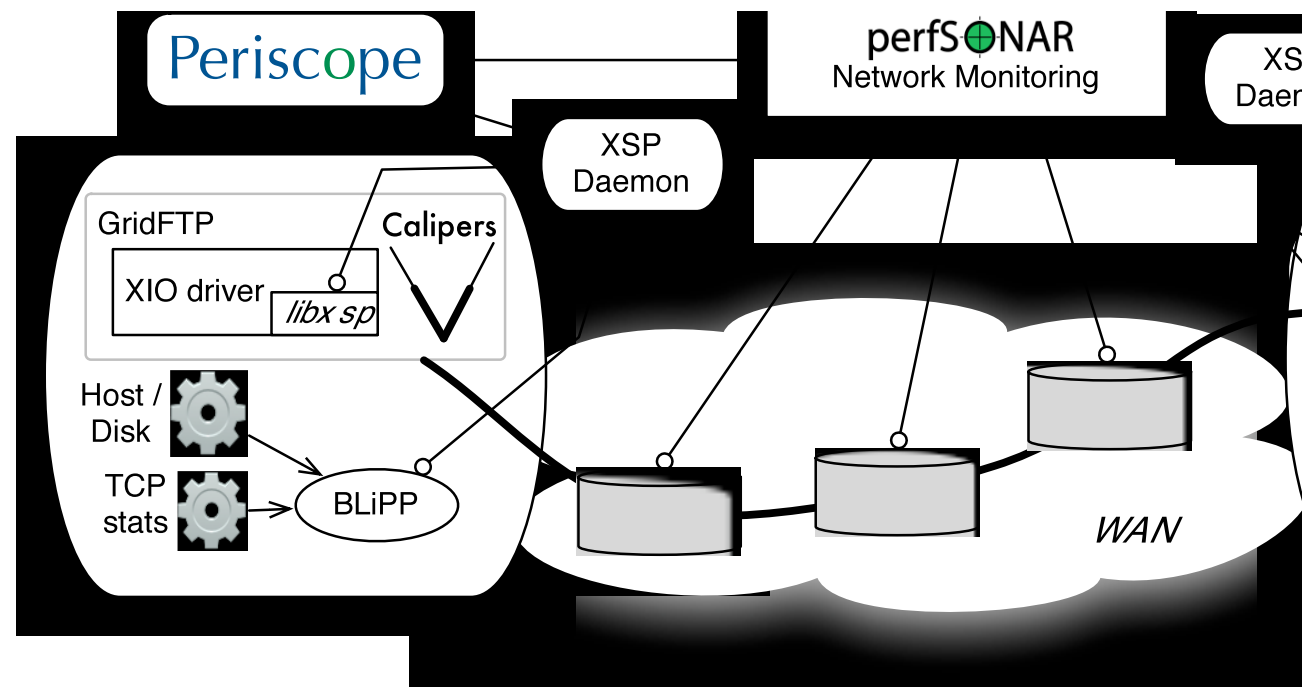
XSP - eXtensible Session Protocol
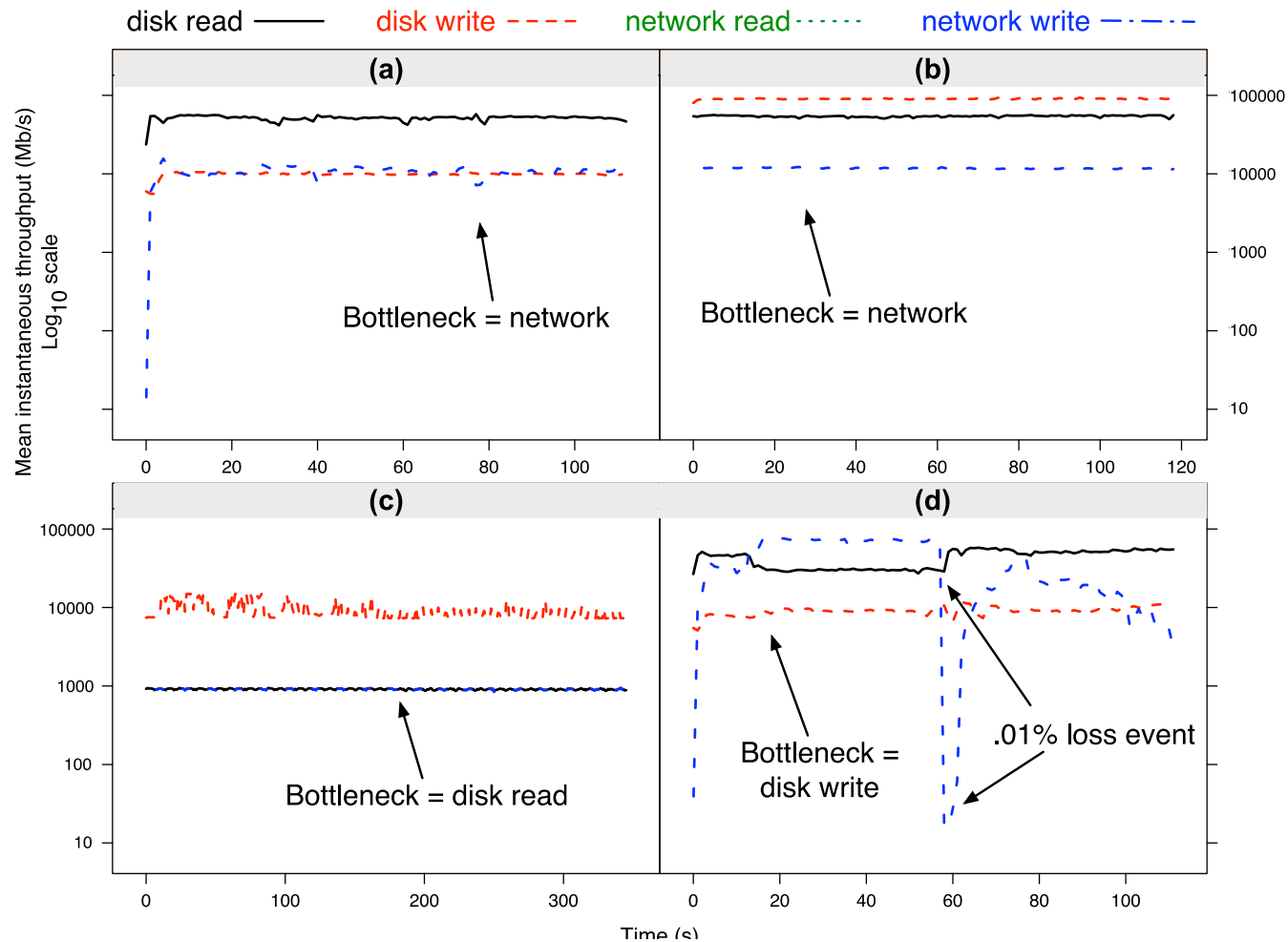
- Session layer (Layer 5 in the OSI model)
- Implemented as a Globus XIO driver
  - Loadable with –dcstack and –fsstack
- Use cases:
  - Monitoring with NetLogger, Calipers and Periscope
  - Dynamic network provisioning
  - WAN Acceleration with Phoebus

Monitoring

- NetLogger monitors read and write system calls, and Calipers summarizes these in situ
- The XSP collector daemon collates and forwards to Periscope
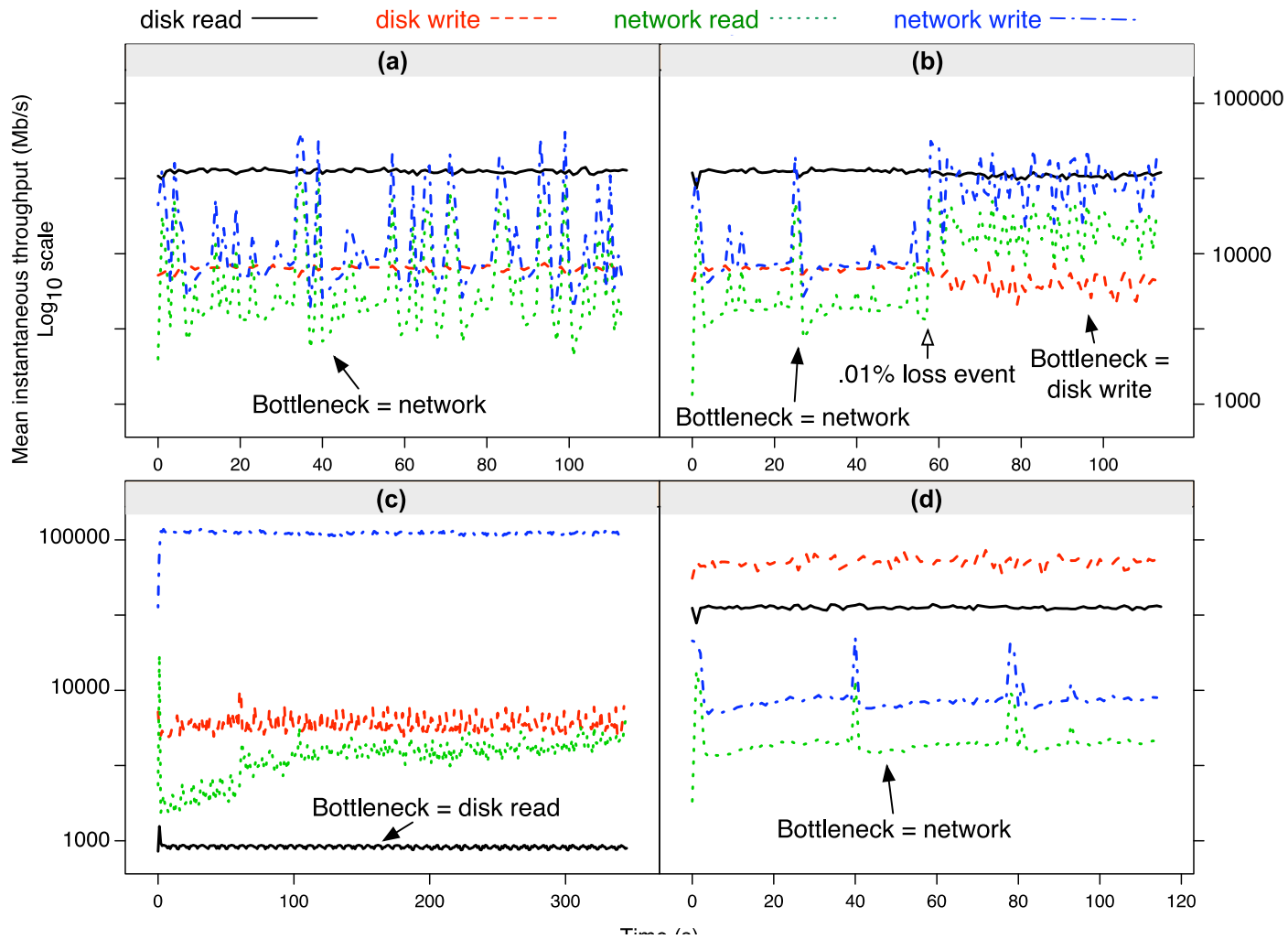 - Periscope is a perfSONAR front end and cache
- BLiPP:

Basic
Lightweight
Periscope
Probes

disk read —— **disk write** – – – **network read** ······ **network write** —·—·—

- (a) Bottleneck = network
- (b) Bottleneck = network
- (c) Bottleneck = disk read
- (d) Bottleneck = disk write, .01% loss event

Mean instantaneous throughput (Mb/s) Log$_{10}$ scale
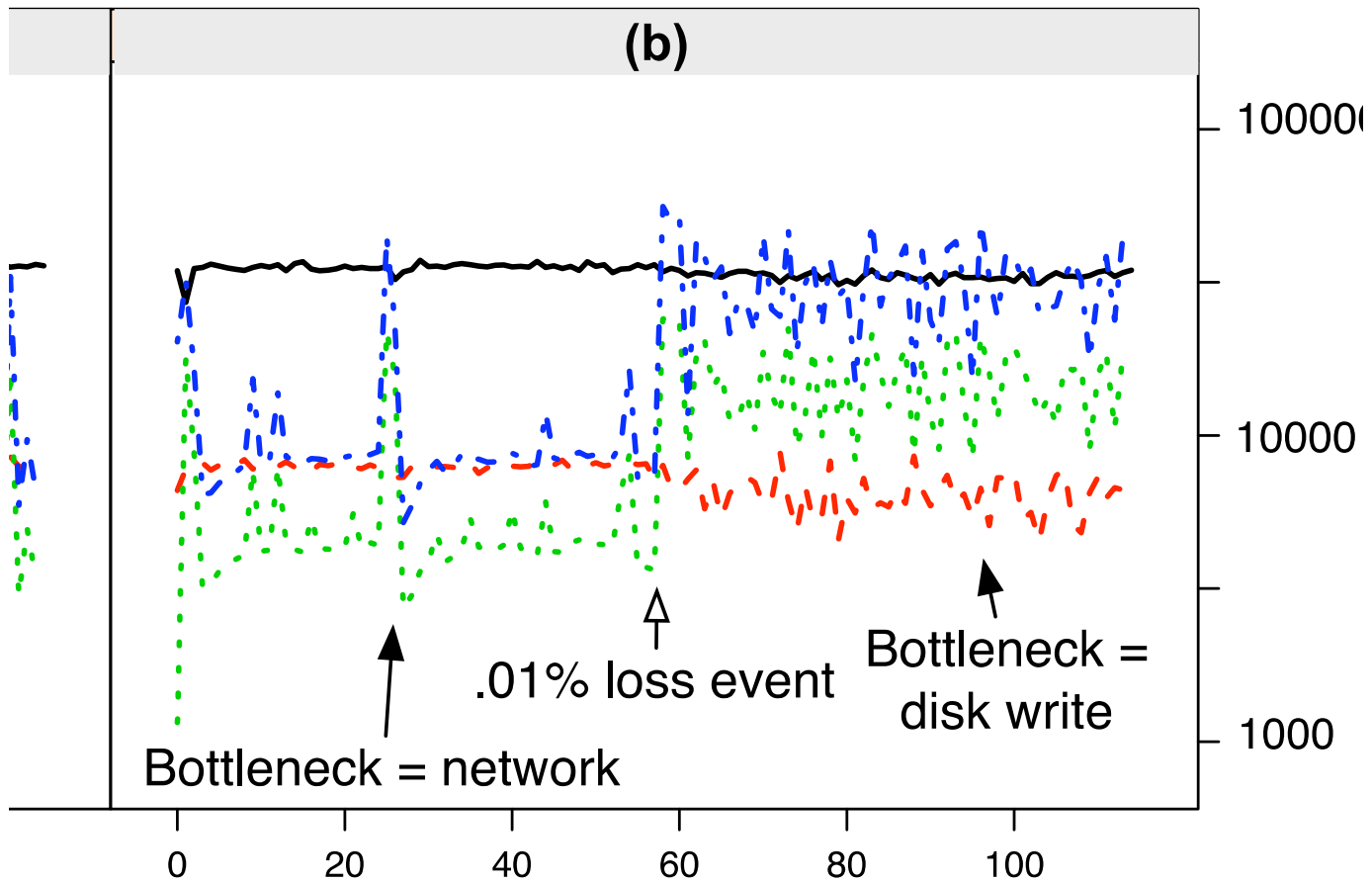
Time (s)

TCP throughput

Time series of throughput for representative TCP experiments: (a) 1 stream memory-to-disk with 100ms latency, (b) 1 stream memory-to-memory with no latency, (c) 1 stream disk-to-disk with no latency, (d) 4 streams memory-to-disk with 100ms latency and 1% loss added at 60 seconds.
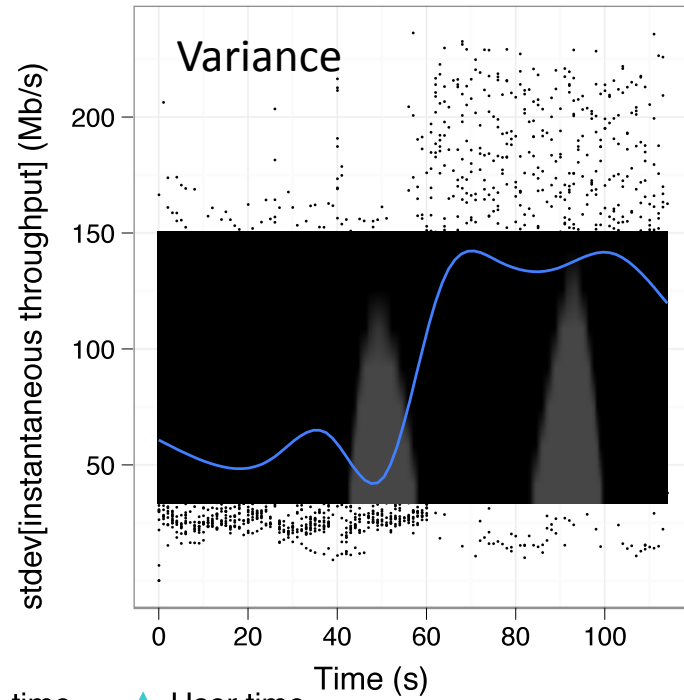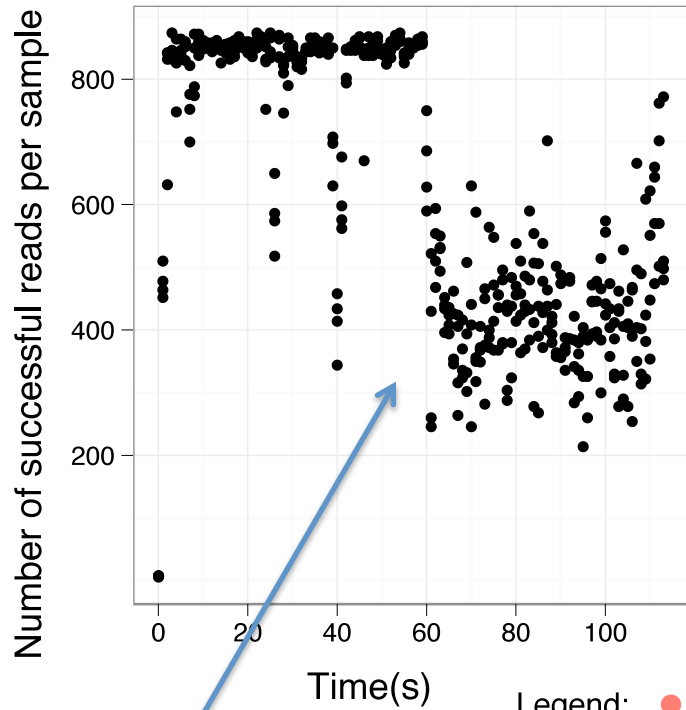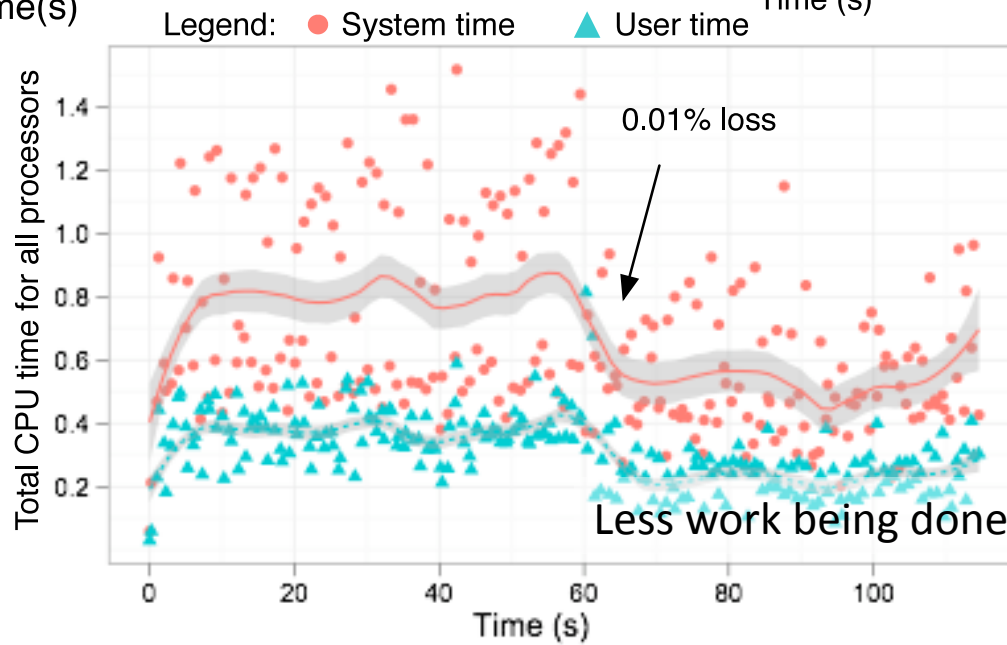
UDT throughput

Time series of throughput for representative UDT experiments: (a) 4 streams memory-to-disk with 100ms latency, (b) 4 streams memory-to-disk with 100ms latency and 1% loss added at 60 seconds, (c) 4 streams disk-to-disk with 100ms latency, (d) 4 streams memory-to-memory with 100ms latency.
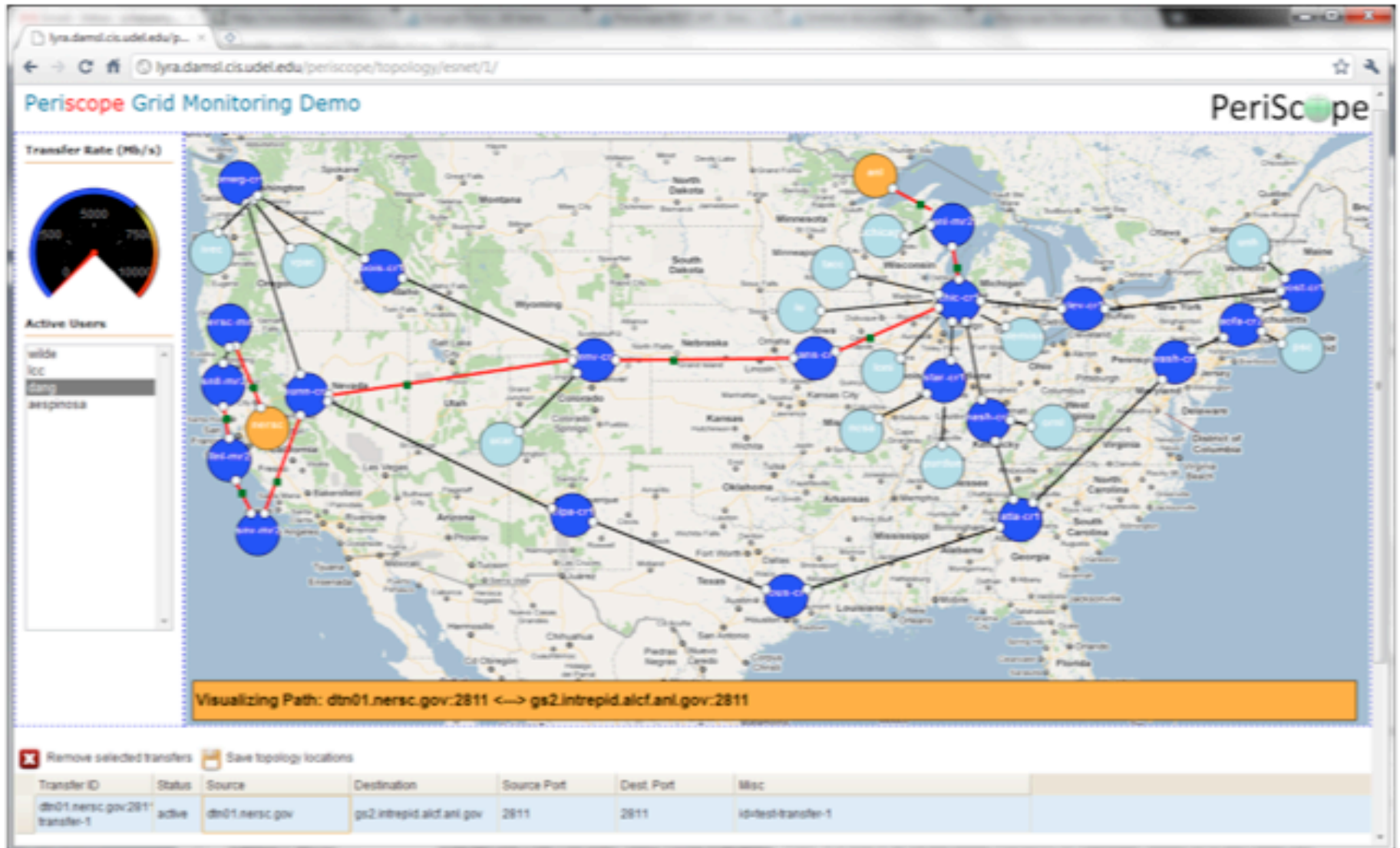
Wait, what?

(b)

Bottleneck = network

.01% loss event

Bottleneck = disk write

Variance

Legend: ● System time ▲ User time

0.01% loss

Half as many read()s.
Others return zero,
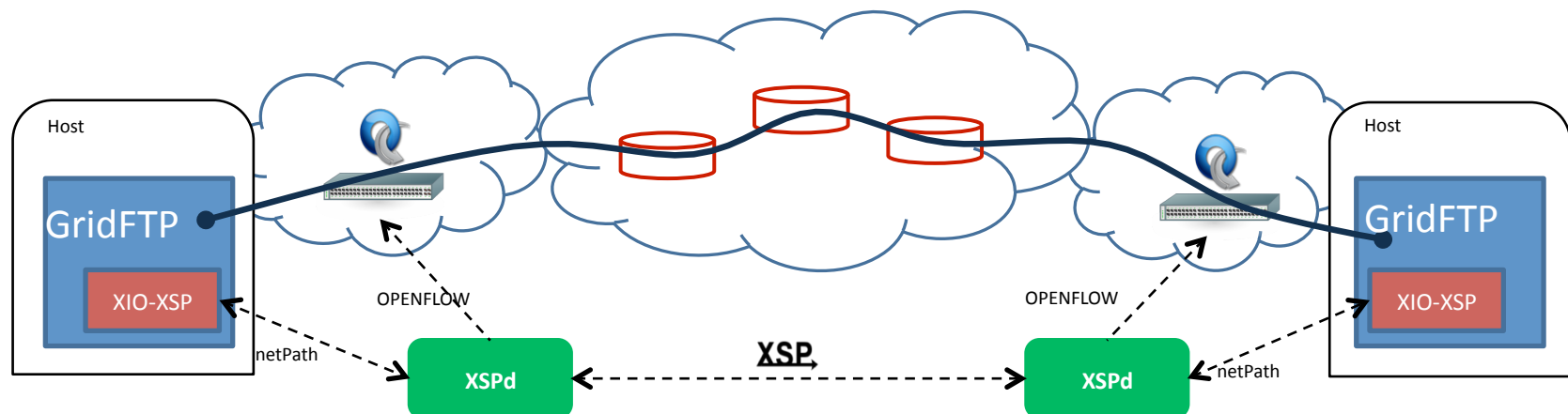not counted
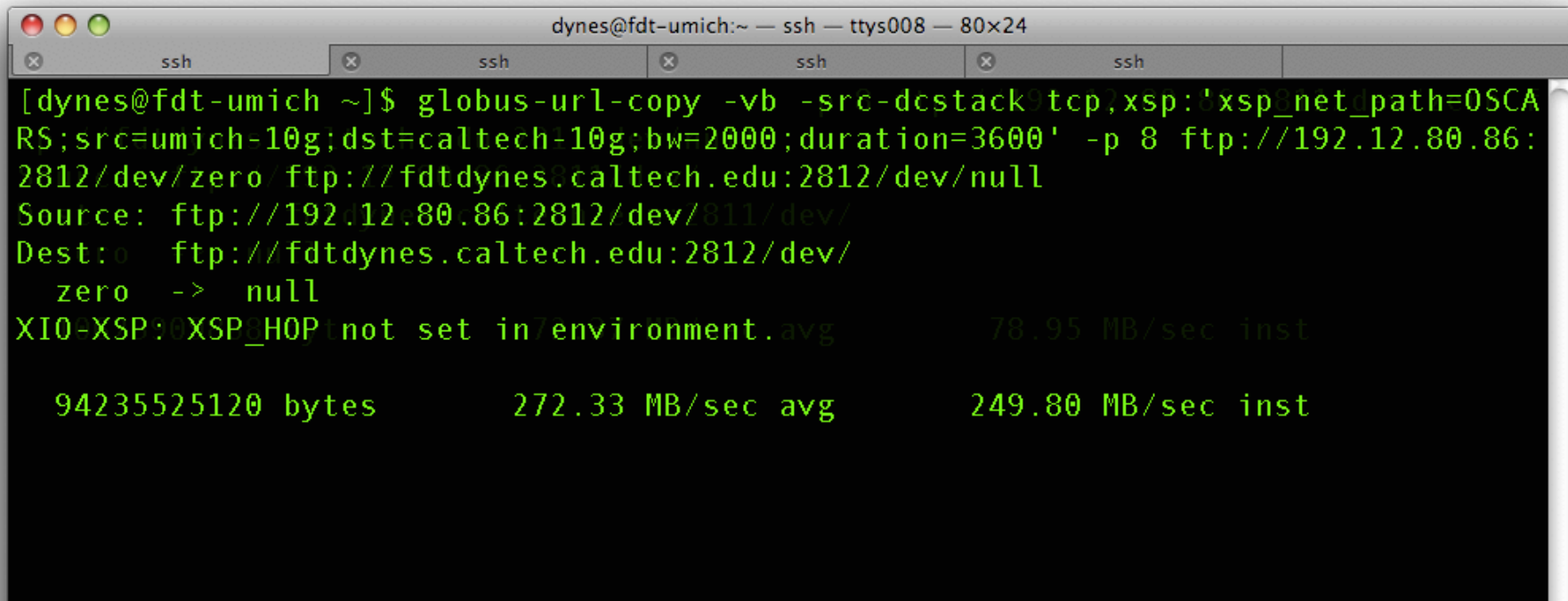
Less work being done

Topology-aware Monitoring

XSP and Dynamic Circuits

- All things old are new again – circuits
  - Still packets over traffic engineered paths
- ESnet OSCARS, Internet2 ION and OS^{3}E/NDDI
  - NDDI is based on OpenFlow, an emerging protocol to allow fine-grained control of network devices
- SC11 Demo: monitor and react

DYNES

- DYNES is an NSF MRI that is distributing storage and OSCARS IDCs for ION to various sites
  - Internet2, Caltech, Vanderbilt, U. Michigan

- Initially using FDT from Caltech, but we are deploying the circuit capable GridFTP with XSP

DYNES Participants

Wed Apr 11 2012

GridFTP+XSP using the Phoebus WAN Accelerator System

- Phoebus is an open source WAN accelerator funded by the DOE and now the NSF
- Phoebus uses XSP to communicate via gateways that can tune, adapt and translate protocols
  - TCP tuning, UDT, RDMA over the WAN

Conclusion

- Quite a few topics focused on the performance of GridFTP

- Flexible and scalable monitoring for troubleshooting

- Adapt performance using emerging network technologies and protocols

- Questions?