Monitoring and Accelerating GridFTP

Ezra Kissel & Martin Swany
Indiana University

Dan Gunter
Lawrence Berkeley National Laboratory
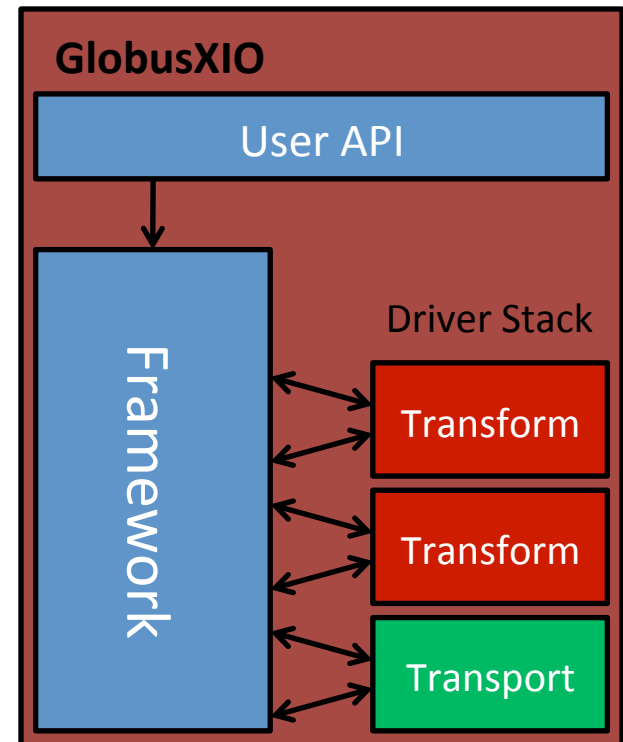
Jason Zurawski
Internet2

GlobusWORLD 2013

Globus XIO

- Modular drivers that can be used within the Globus Toolkit

- We have focused on extending features for GridFTP

- XSP for monitoring and path signaling

- Phoebus for WAN acceleration
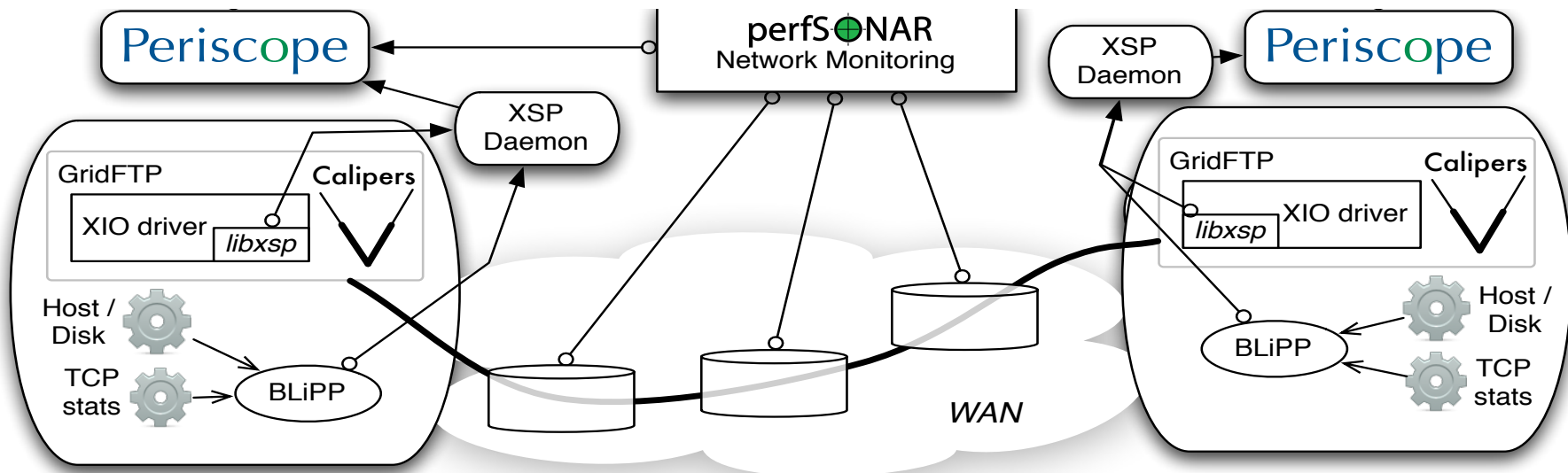

Source: Globus XIO developer guide

XSP - eXtensible Session Protocol

- Session layer (Layer 5 in the OSI model)

- We can also think of a session in the most literal sense:
  - "*a period of time devoted to a particular activity*"

- *xio-xsp* implemented as a Globus XIO driver
  - Loadable with `-dcstack` and `-fsstack`

- Use cases:
  - Monitoring with NetLogger, Calipers, and Periscope
  - Dynamic network provisioning (e.g., OSCARS, OpenFlow)
  - WAN Acceleration with Phoebus Gateways

End-to-end measurement perspective

- NetLogger instruments read and write system calls, and Calipers summarizes these in memory

- The XSP collector daemon collates and forwards to Periscope
  - RESTful Measurement Store and Unified Network Information Service (UNIS)

- BLiPP: **B**asic **L**ightweight **P**eriscope **P**robes
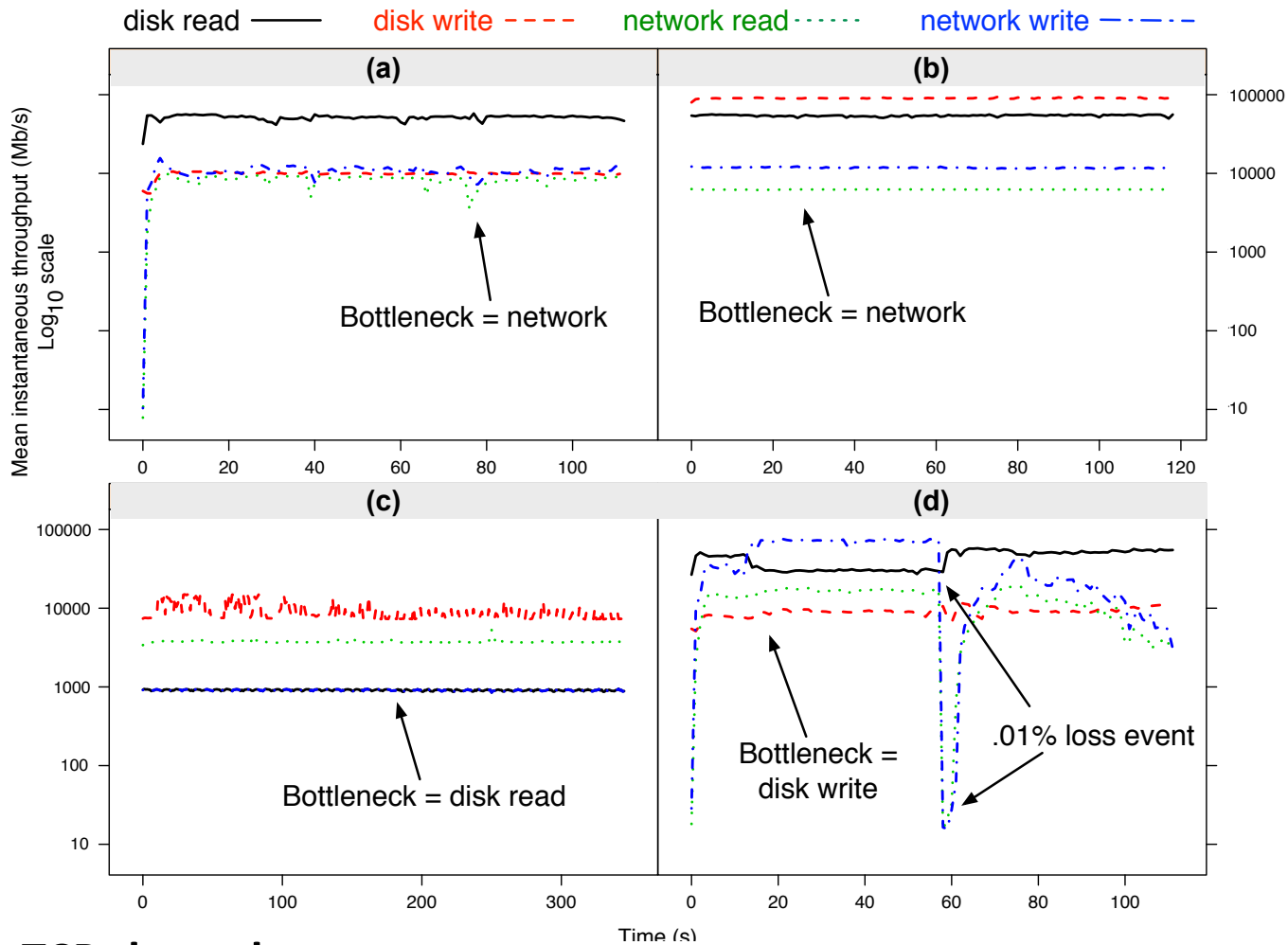
Metadata

```
{
  'eventType': 'nl:tools:calipers:summary:write',
  'id': '8d83ccd4-a708-11e2-b6fa-33d5436b6f2a',
  'parameters': { 'collectionInterval': 5,
                  'datumSchema': 'http://unis.incntre.iu.edu/schema/
   20120709/datum#'},
  'subject': { 'dport': '41186',
               'dst': '127.0.0.1',
               'sport': '60512',
               'src': '127.0.0.1',
               'stream_id': '60512:41186',
               'task_id': 'xsp_test',
               'type': 'network',
               'xfer_id': '8d83a8a8-a708-11e2-b6fa-33d5436b6f2a'}
}
```

Data

```
{
  'data': [ 1366166610.043317,        ts
            2,                         _sample
            5866638858,                sum_v
            262161,                    min_v
            262161,                    max_v
            262161,                    mean_v
            0,                         sd_v
            52064.35928057134,         sum_r
            0.1114156396090098,        min_r
            3.318493670886076,         max_r
            2.32658679419838,          mean_r
            0.4392224221097757,        sd_r
            9945.430479743363,         sum_g
            0.3013415420295162,        min_g
            8.975400612600653,         max_g
            0.4444289248254251,        mean_g
            0.09869011618422469,       sd_g
            22378,                     count
            5.000102,                  dur
            2.607303999999963],        dur_inst
  'mid': '8d83ccd4-a708-11e2-b6fa-33d5436b6f2a'
}
```
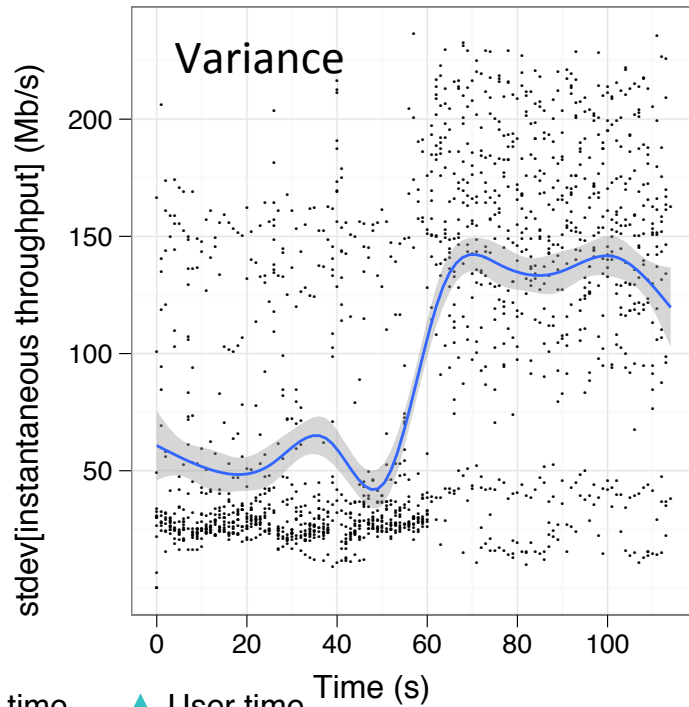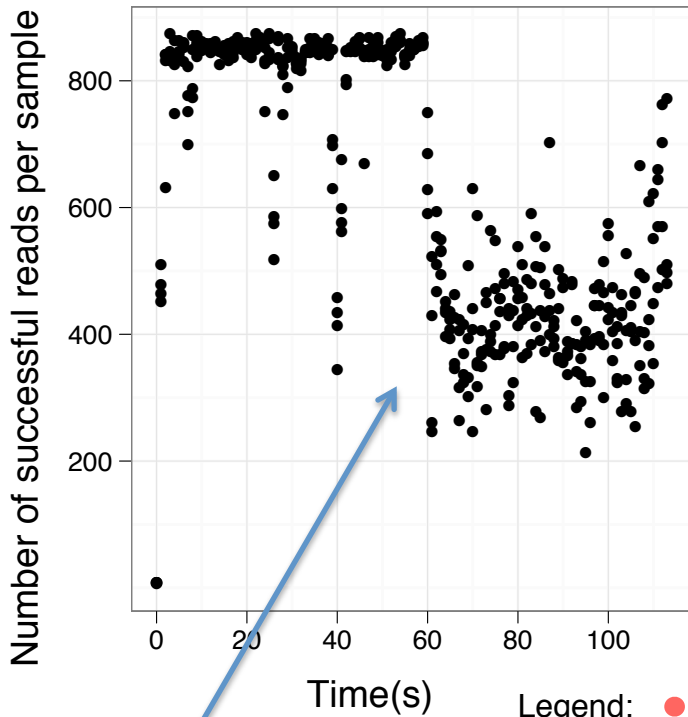
TCP throughput

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.

1. Ezra Kissel, Dan Gunter, et. al. **Scalable Integrated Performance Analysis of Multi-Gigabit Networks.**
   In 5th International Workshop on Distributed Autonomous Network Management Systems (DANMS), April 2012.
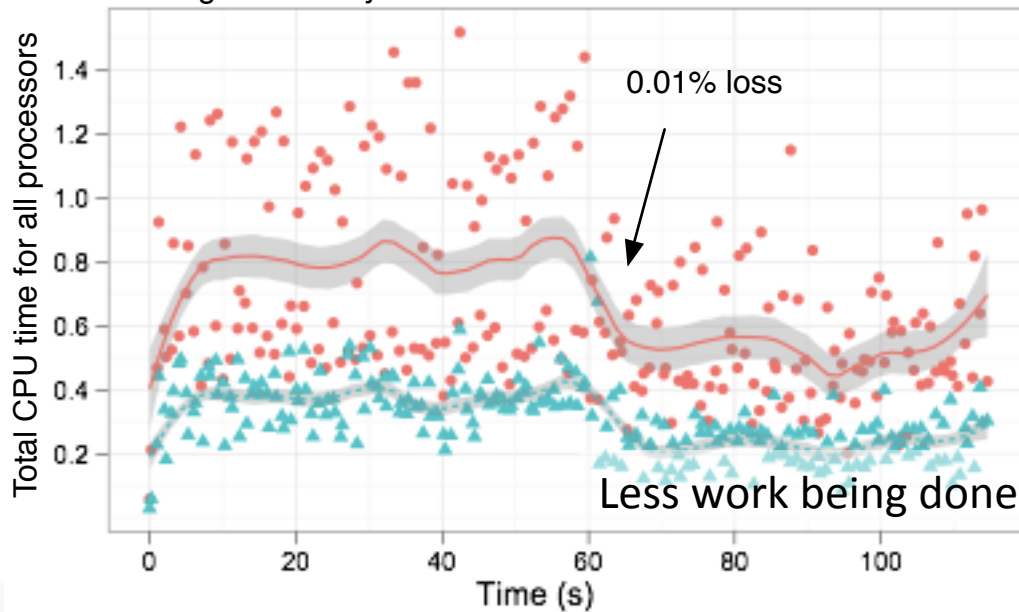
Number of successful reads per sample — Time(s)

Variance
stdev[instantaneous throughput] (Mb/s) — Time (s)

Legend: ● System time ▲ User time

Half as many read()s.
Others return zero,
not counted

0.01% loss

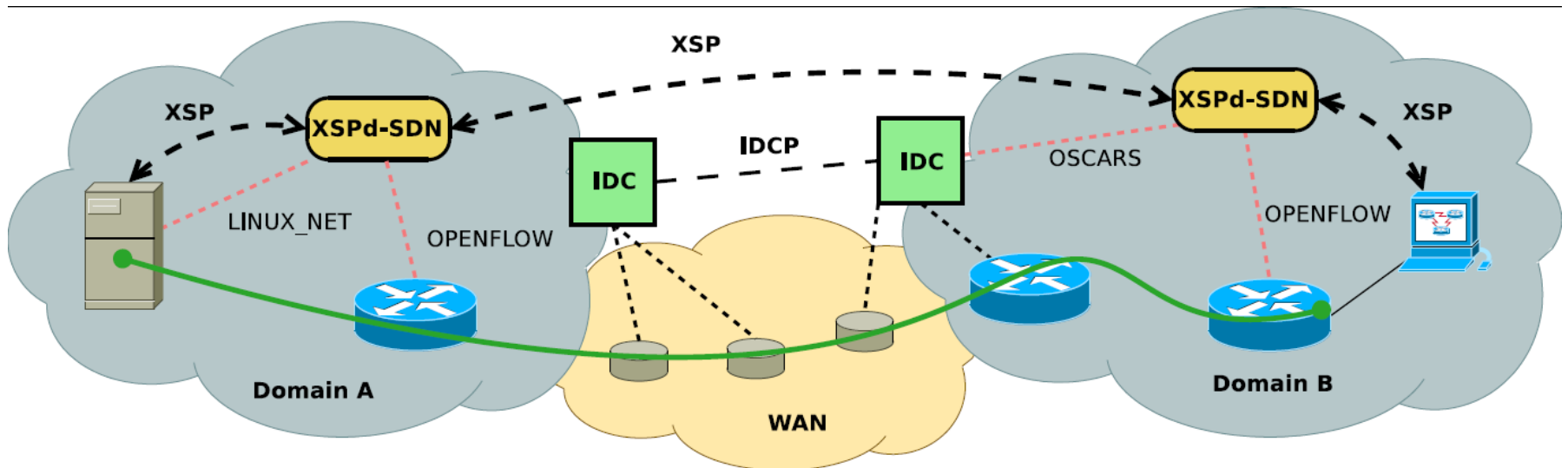Total CPU time for all processors — Time (s)

Less work being done
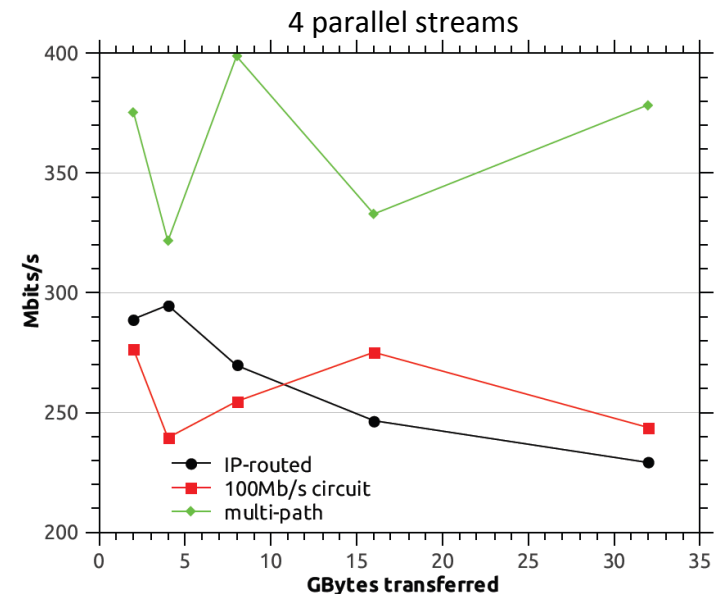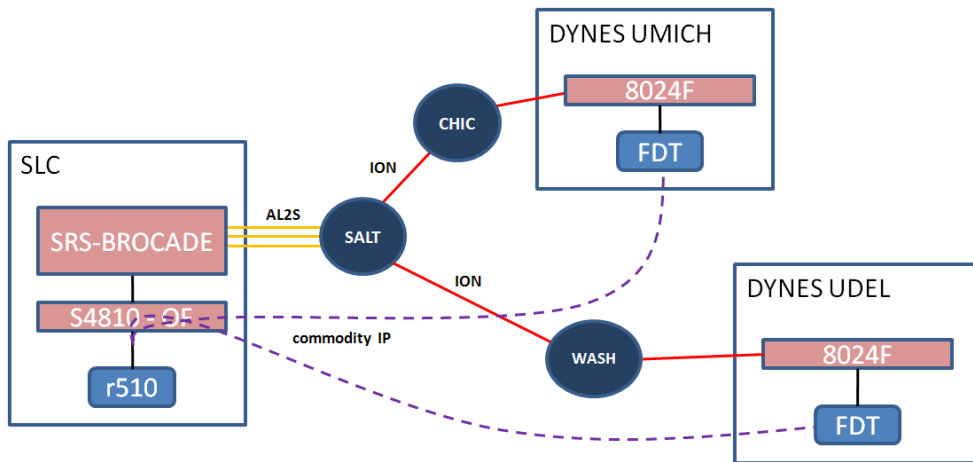
Topology-aware monitoring

XSP and Dynamic Circuits

- Common interface for path provisioning
  - ESnet OSCARS, Internet2 ION, OpenFlow, Linux end-hosts, OS^{3}E/NDDI
- Prototype installation for DYNES
- DYNES is an NSF MRI that is distributing storage and OSCARS IDCs for ION to various sites
  - Internet2, Caltech, Vanderbilt, U. Michigan
- Recent testing with multi-path GridFTP at SC12 (Science Research Sandbox)

SC12 SRS experiments

- Multi-path GridFTP transfers to DYNES end-sites
- *xio-xsp* and OpenFlow used to dynamically redirect parallel streams over circuit services
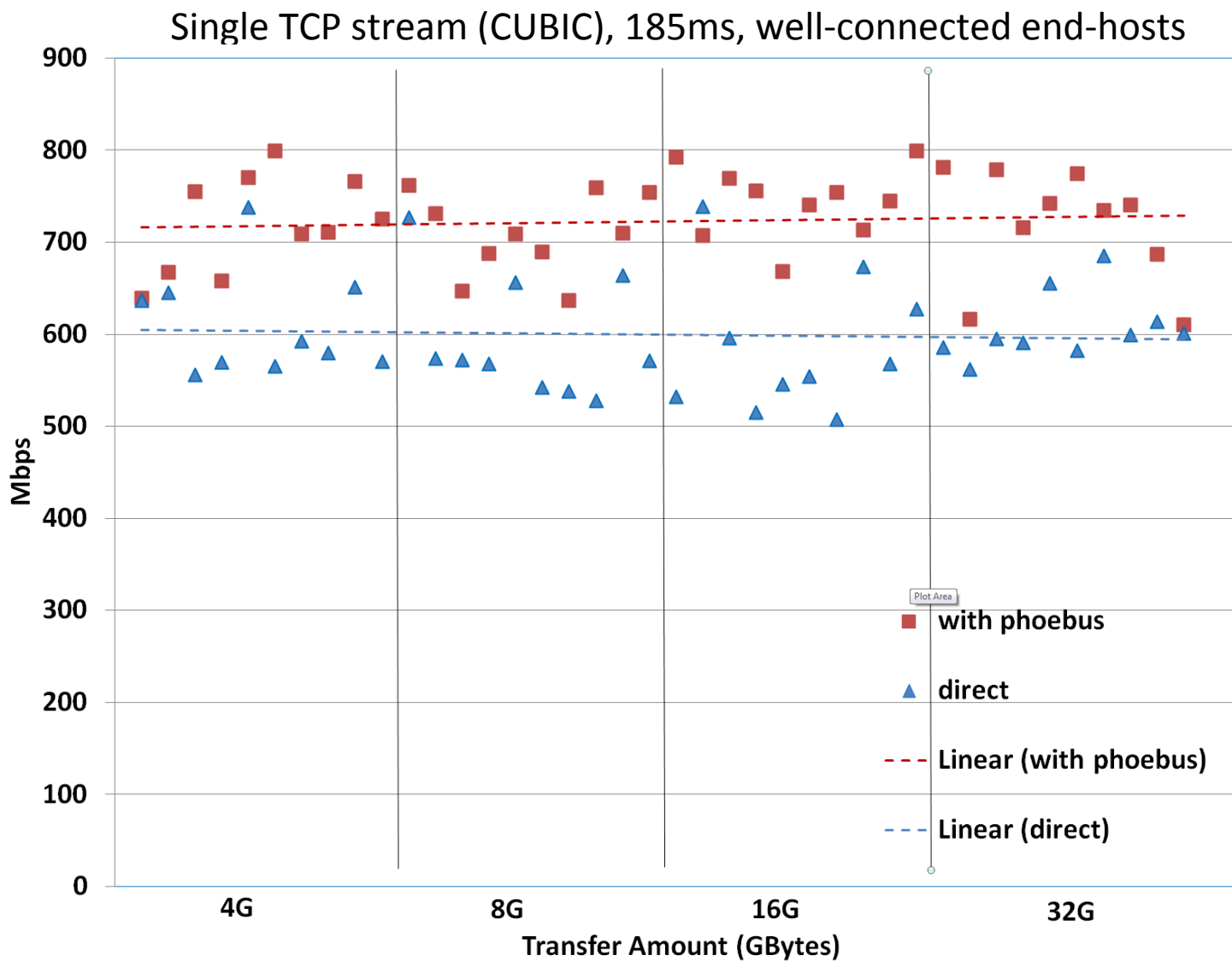- Many lessons learned at SC

GridFTP+XSP using Phoebus

- 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, UDP, RDMA over the WAN

```
globus-url-copy -vb -p 4 -dcstack
        xsp:"xsp_hop=<host/5006>;xsp_net_path=<TYPE>",
        phoebus:"phoebus_path=<GW1>/5006#<GW2>/5006 "
        ftp://<src host>:2811/dev/zero
        ftp://<dst host>:2811/dev/null
```

IU-Tokyo with and without Phoebus

Single TCP stream (CUBIC), 185ms, well-connected end-hosts

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

- Ongoing work: early prototype of *xio-xsp* for GlobusOnline (GCMU installs)

Thank you for your time

- Thanks to our colleagues at IU, Internet2, ESnet

- Support: NSF OCI-0910812, OCI-1127349, and CNS-105011, DOE DE-SC0001421

  - https://damsl.cs.indiana.edu/projects/phoebus
  - Email: ezkissel@indiana.edu

- Questions?