Phoebus Update

JET Meeting – February 19th, 2013

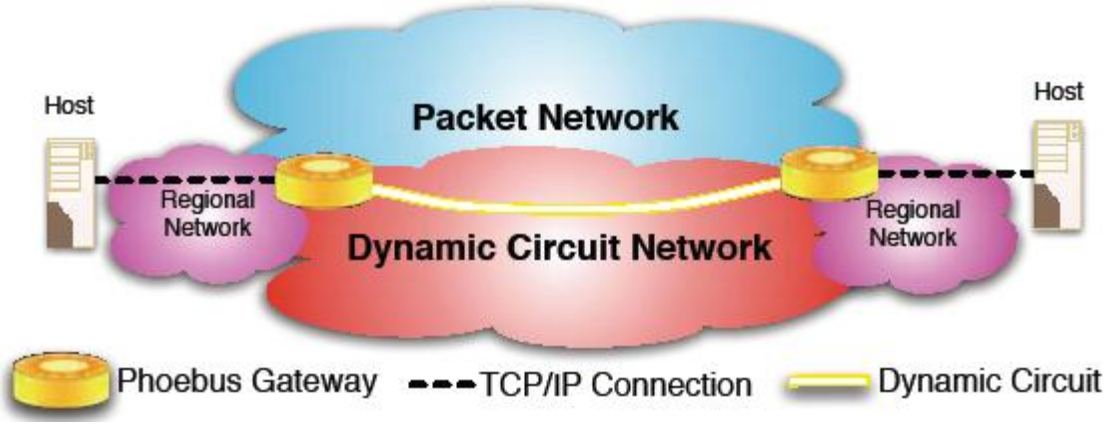
- Martin Swany (swany@indiana.edu)
- Ezra Kissel (ezkissel@indiana.edu)

- Indiana University – School of Informatics and Computing

InCNTRE

Supporting bulk data movement

- Protocol tuning necessary in order to achieve good performance
  - Not always successful for a number of reasons
  - Well- known problems with existing transport protocols
  - "wizard gap"
- "Hybrid" networks
  - Dynamically allocate some links for high-demand flows
  - Virtual circuits
  - SDN
- How to effectively utilize these high-performance paths?



Phoebus and XSP

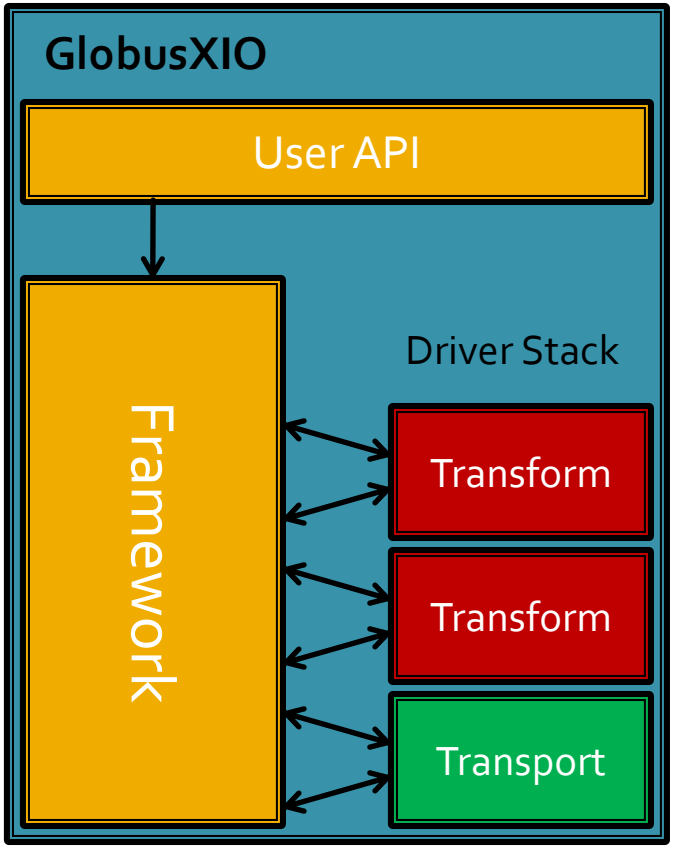
- Designed to help bridge the end-to-end performance gap
  - Shorter RTTs from the edge to PG
  - Automatic tuning on a per-segment basis
  - Use of alternative protocols over the WAN
- eXtensible Session Protocol (XSP), a session-layer protocol for IP networks, provides generalized messaging between gateways and intermediate devices and services
- Phoebus offers a gateway for legacy application to use advanced networks such as the Internet2 ION virtual circuit network
  - Support for AL2S via OESS on the way
- Standard interfaces at the edge, innovation at the core

Using Phoebus

- Transparent wrapper
  - Intercept socket calls (LD_PRELOAD)
  - No code modifications needed
- XSP client library (*libxsp_client)*
  - Authentication with X.509 certificates, SSH subchannel for client connection
- iptables redirection
- Packages for setting up new gateways
  - RPMs for CentOS

Phoebus XIO driver for GridFTP

- A modular Phoebus Transport driver to be used within the Globus Toolkit
- Based on the XIO TCP driver allowing for compatibility with existing configurations and options
- Simplifies use of Phoebus architecture for Globus applications (eliminates need for shim library)
- http://damsl.cs.indiana.edu/phoebus/wiki/GridFTP



Source: Globus XIO developer guide

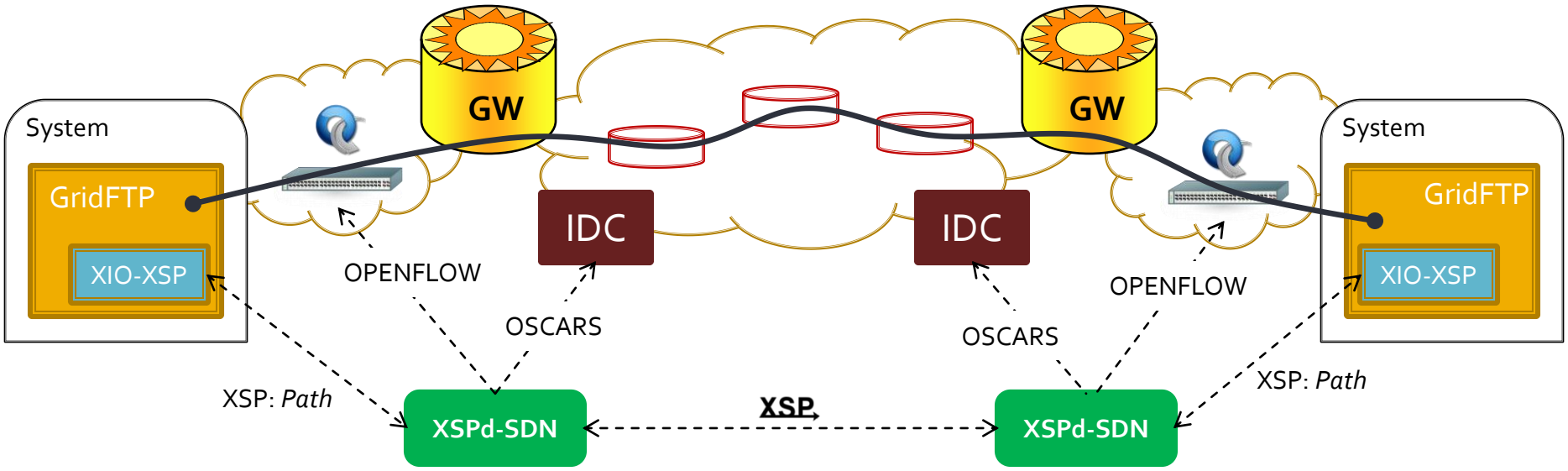
Phoebus and GridFTP

- `globus-gridftp-server` may load the Phoebus XIO driver as requested by a client application
- New connections are made to the specified Phoebus Gateways
- `globus-url-copy` extended to support Phoebus
  - -ph flag
  - -dcstack when explicitly specifying the network stack
- Maintains support for third-party transfers as well as PASV and extended block mode (MODE E) with parallel streams

Building virtual paths for GridFTP

- Phoebus/XSP drivers enable signaling into the network
- Prototype now working with GlobusOnline

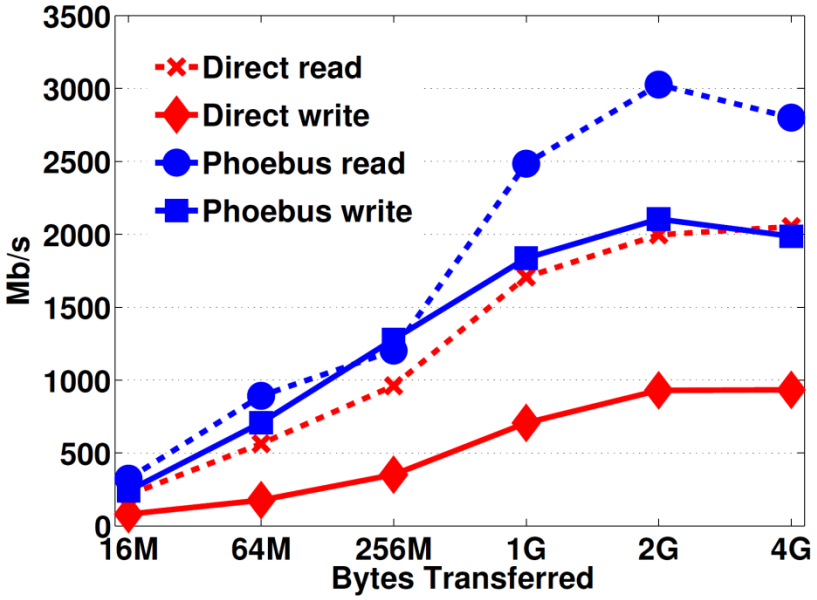
globus-url-copy -vb -p 4 -dcstack
xsp:"xsp_hop=<XSPd-SDN>/5006;xsp_net_path=<TYPE>", phoebus:"phoebus_path=<GW1>/5006#<GW2>/5006 "
ftp://<src host>:2811/dev/zero ftp://<dst host>:2811/dev/null



Phoebus TCP performance and scalability

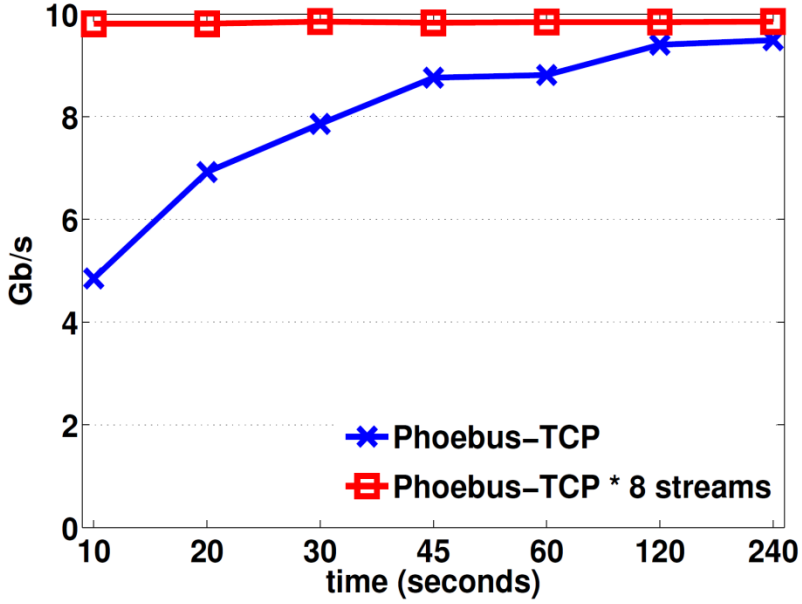
Phoebus in real networks
CERN to Vanderbilt
Transatlantic link

4G max, 2 PGs (Chicago and Atlanta), 154ms RTT

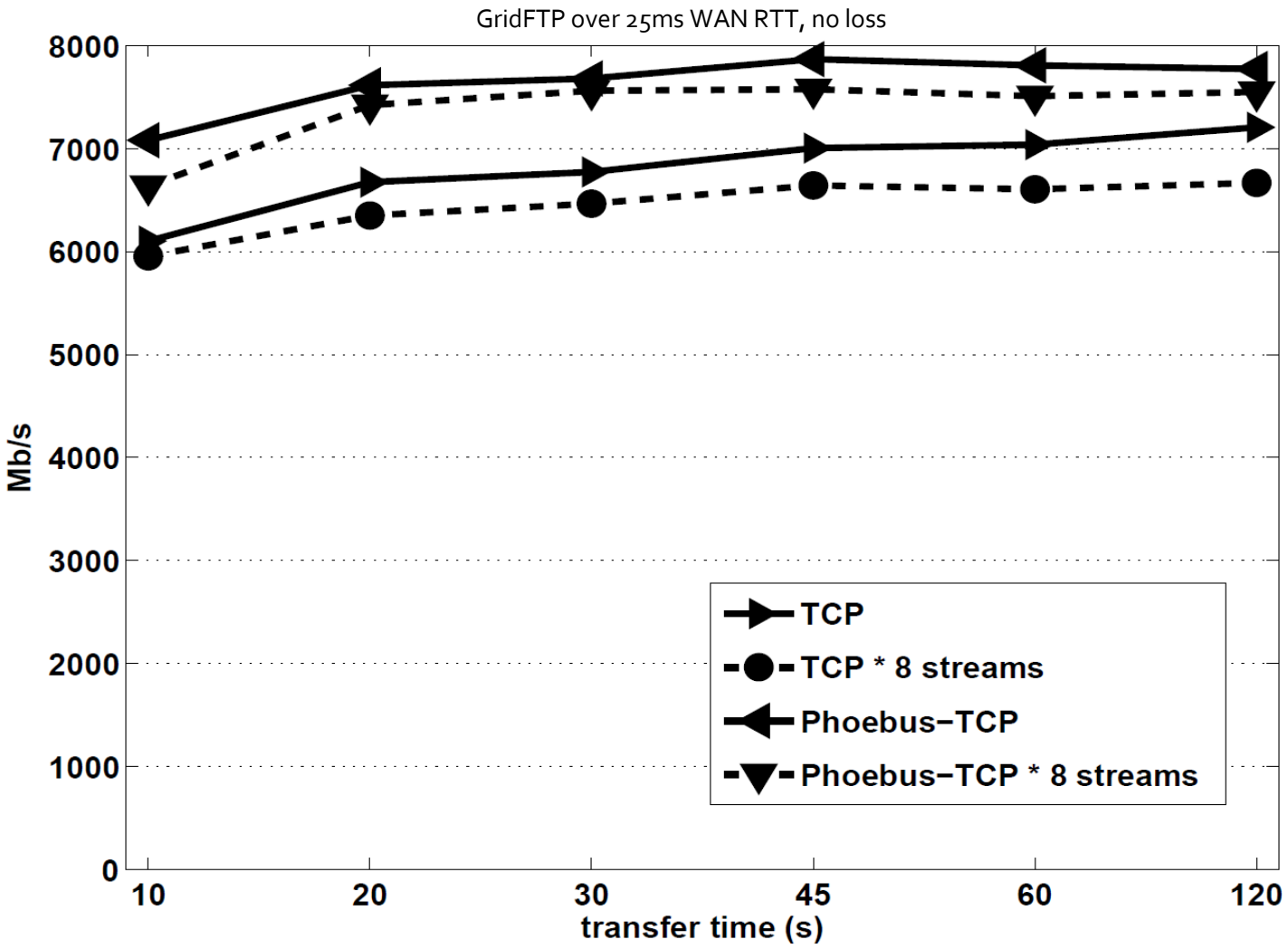


Phoebus scaling to 10G

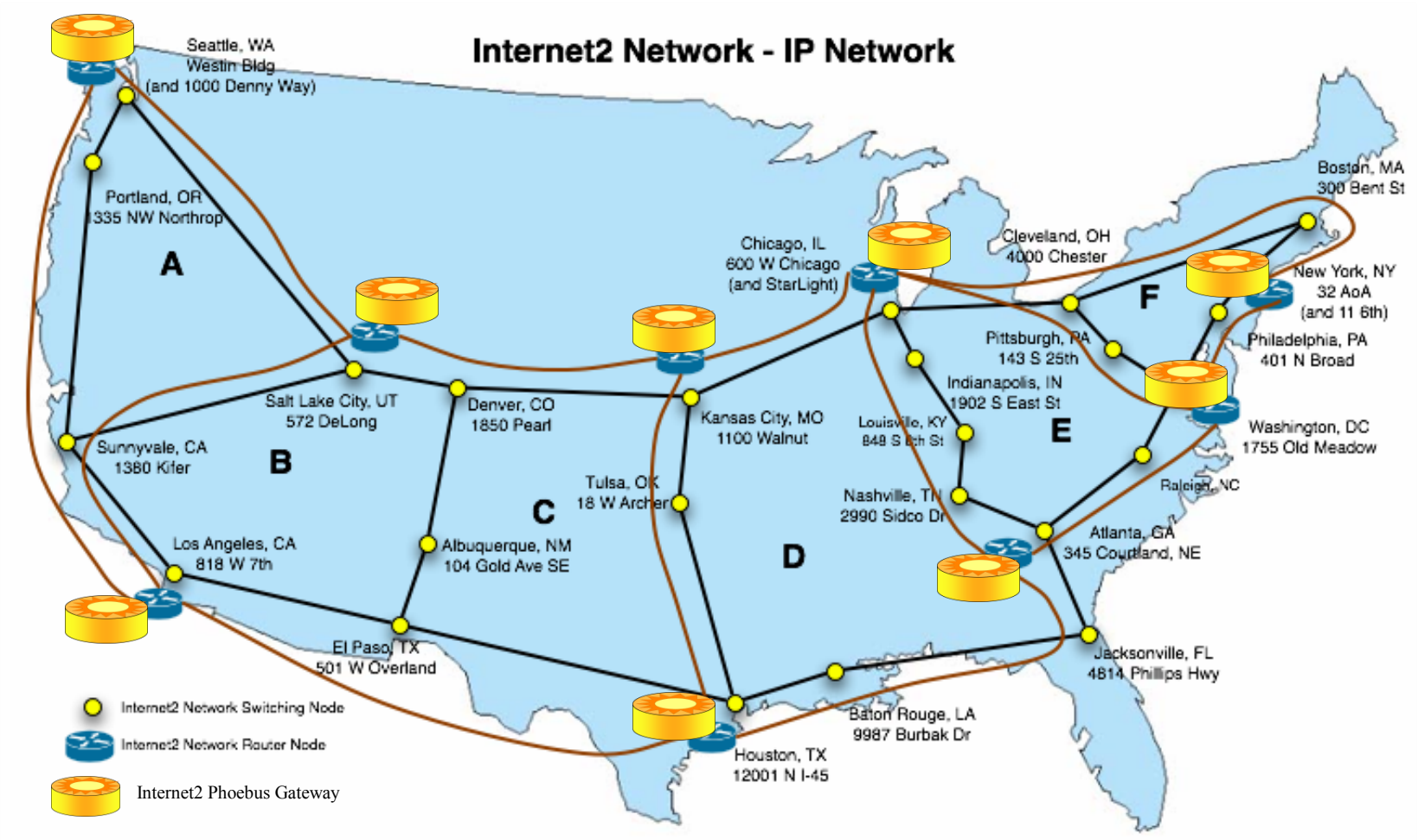
10G network testbed with clean links, 2PGs



Phoebus at 10G



Phoebus I2 deployment

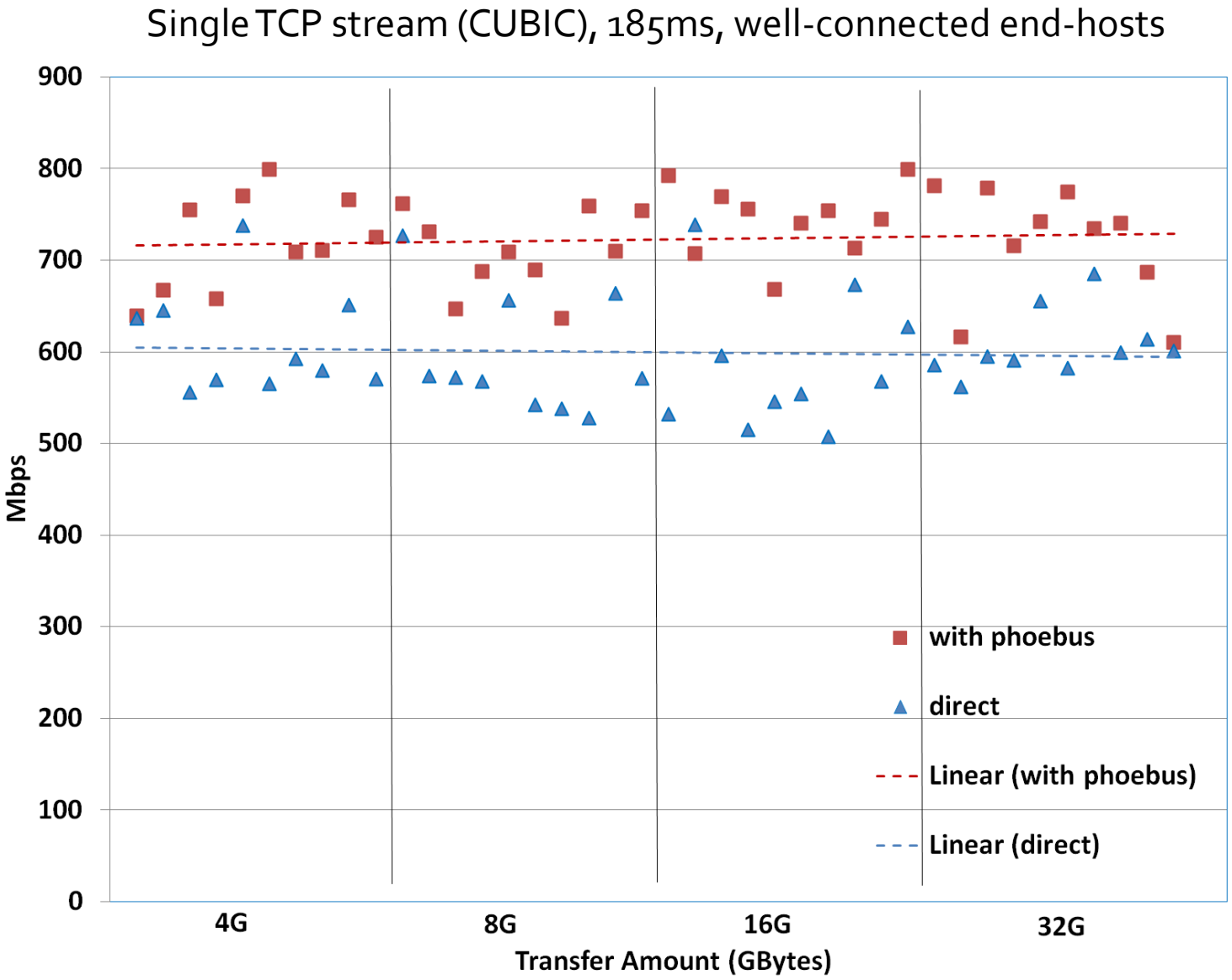


TransPAC testing

- IU to Tokyo
- Using Phoebus Gateways on I2 and TP3 host in LA

```
$ traceroute 203.181.249.202
traceroute to 203.181.249.202 (203.181.249.202), 30 hops max, 60 byte packets
 1  c-1.mgmt-sw.bldc.testlab.grnoc.iu.edu (156.56.5.1)  2.821 ms  2.791 ms  2.774 ms
 2  ge-4-0-6.576.cr3.bldc.net.uits.iu.edu (156.56.249.73)  0.188 ms  0.174 ms  0.170 ms
 3  ae-13.0.br2.bldc.net.uits.iu.edu (134.68.3.58)  0.227 ms  0.246 ms  0.220 ms
 4  ae-0.0.br2.ictc.net.uits.iu.edu (134.68.3.34)  1.072 ms  1.060 ms  1.066 ms
 5  xe-0-1-0.9.rtr.ictc.indiana.gigapop.net (149.165.254.229)  1.115 ms  1.117 ms  1.094
    ms
 6  et-3-0-0.1.rtr.ll.indiana.gigapop.net (149.165.255.194)  1.224 ms  1.213 ms  1.233 ms
 7  xe-0-0-1.110.rtr.atla.net.internet2.edu (149.165.254.21)  15.736 ms  15.750 ms  15.728
    ms
 8  xe-1-0-0.0.rtr.hous.net.internet2.edu (64.57.28.112)  39.117 ms  39.069 ms  39.095 ms
 9  * * *
10  transpac-1-lo-jmb-702.lsanca.pacificwave.net (207.231.240.136)  70.062 ms  70.002 ms
    70.077 ms
11  tokyo-losa-tp2.transpac2.net (192.203.116.146)  185.641 ms  185.727 ms  185.726 ms
12  jm-e600-v34.jp.apan.net (203.181.248.109)  198.723 ms  185.922 ms  185.908 ms
13  nms8.jp.apan.net (203.181.249.202)  185.845 ms !X  185.819 ms !X  185.764 ms !X
```

IU-Tokyo with and without Phoebus



Ongoing development

- Efficient gateway discovery
  - Find nearest available gateway based on network location
- Improved client support
  - Windows and OSX widgets
- Openstack appliance
- Phoebus RoCE/SLaBS backend
  - "buffer and burst" model with protocol adaptation
  - http://damsl.cs.indiana.edu/projects/phoebus/phoebus_rdma.pdf

Thank you!

- Support:
  - DOE – DE-FG02-04ER25642
  - NSF – OCI-1127349
- Always looking for new users and use-cases
  - Happy to help integrate with trial applications
  - Big Data Testbed
- Questions?

http://damsl.cs.indiana.edu/phoebus