

Data Intensive Networking at PSC

Michael H Lambert

Internet2 Spring Member Meeting

28 April 2009

History Lesson

- Non-academic department within Carnegie Mellon University
- Originally one of five NSF supercomputing centers
- Now participant in NSF ETF (Extensible Terascale Facility)
- Primary focus is high-end supercomputing
 - Bulk of funding NSF
 - Some NIH, Commonwealth, etc

More History

- NSFnet regional networks created for remote access to supercomputing centers; backbone to interconnect regionals (circa 1987)
 - Grew into modern Internet
 - Did data intensive networking before there was such a notion!
- Node on NSF OC3 vBNS (circa 1995)
 - Before NSF Connections Program
- Connected to Abilene in 1998

Current External Connectivity

- Commodity Internet
- Internet2 (1 Gb/s)
- NLR (10 Gb/s each FrameNet and PacketNet)
- TeraGrid backbone (three 10 Gb/s links)
 - In essence, virtual backplane amongst TeraGrid-approved resources

PSC Sites

- Offices
 - Completely uninteresting
- Machine room
 - As in raised floor, size of several basketball courts
 - Ridiculous amounts of power, cooling
 - Multi-cabinet machines, clusters, file servers, tape archivers, etc
 - At Westinghouse Energy Center

Machine Room Networking

- Underlying assumption: Single flow can fill entire pipe (jumbo frames!)
- Very little use of firewalls
 - Block traffic from “wrong” direction
 - Some IDS, scanning
 - Don’t slow down traffic
- 10 Gb/s switch interconnects
- 40 Gb/s to outside world
- Seeing more 10 Gb/s NICs

Ethernet, IP not Enough

- Historic media: Cray low- and high-speed channels, FDDI, ATM, HiPPI
- Quadrics
 - Very low latency, but proprietary
 - Some older clusters
 - Native, not IP
- InfiniBand
 - Newer clusters, Lustre file systems
 - Native, not IP

Summary

- Obvious statement: High-performance computing can require lots of bandwidth.
- Single flows can fill pipes.
- Interfaces to high-speed networks need to be reasonably open (because of distributed user base).
- Bulk file transfers still constitute majority of data movement.