

DTN Performance Testing At UMBC

Thanks to Mike Petry, petry@umbc.edu

Jack Suess
Jack@umbc.edu

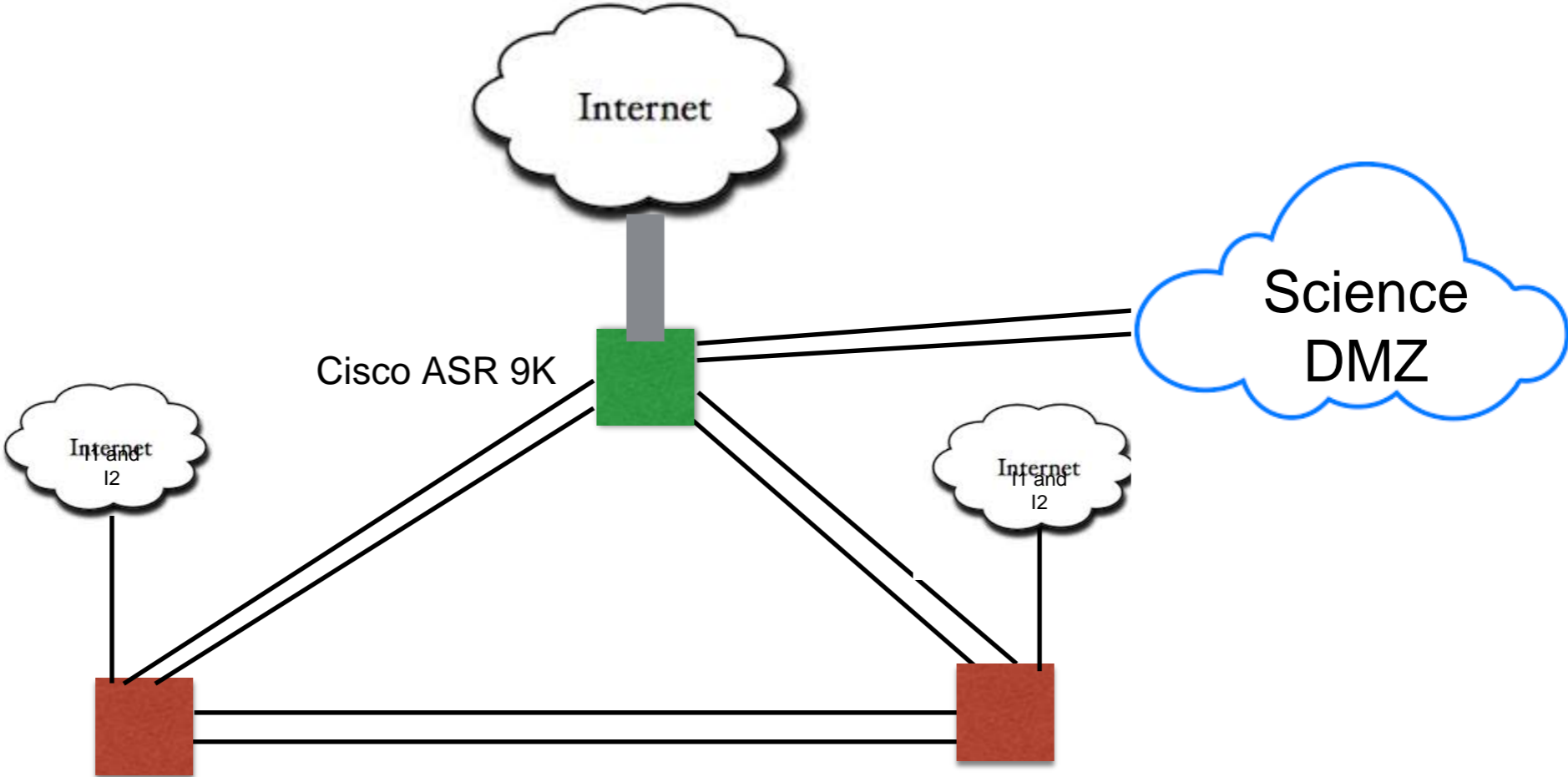
Background

- Last year at this meeting I heard Larry Smarr discuss the Pacific Research Platform and describe FIONA DTN nodes being setup.
- UMBC DTN's are based on the FIONA work at UCSD from NSF-CC-NIE 2012
- <https://fasterdata.es.net/science-dmz/DTN/fiona-flash-i-o-network-appliance/>

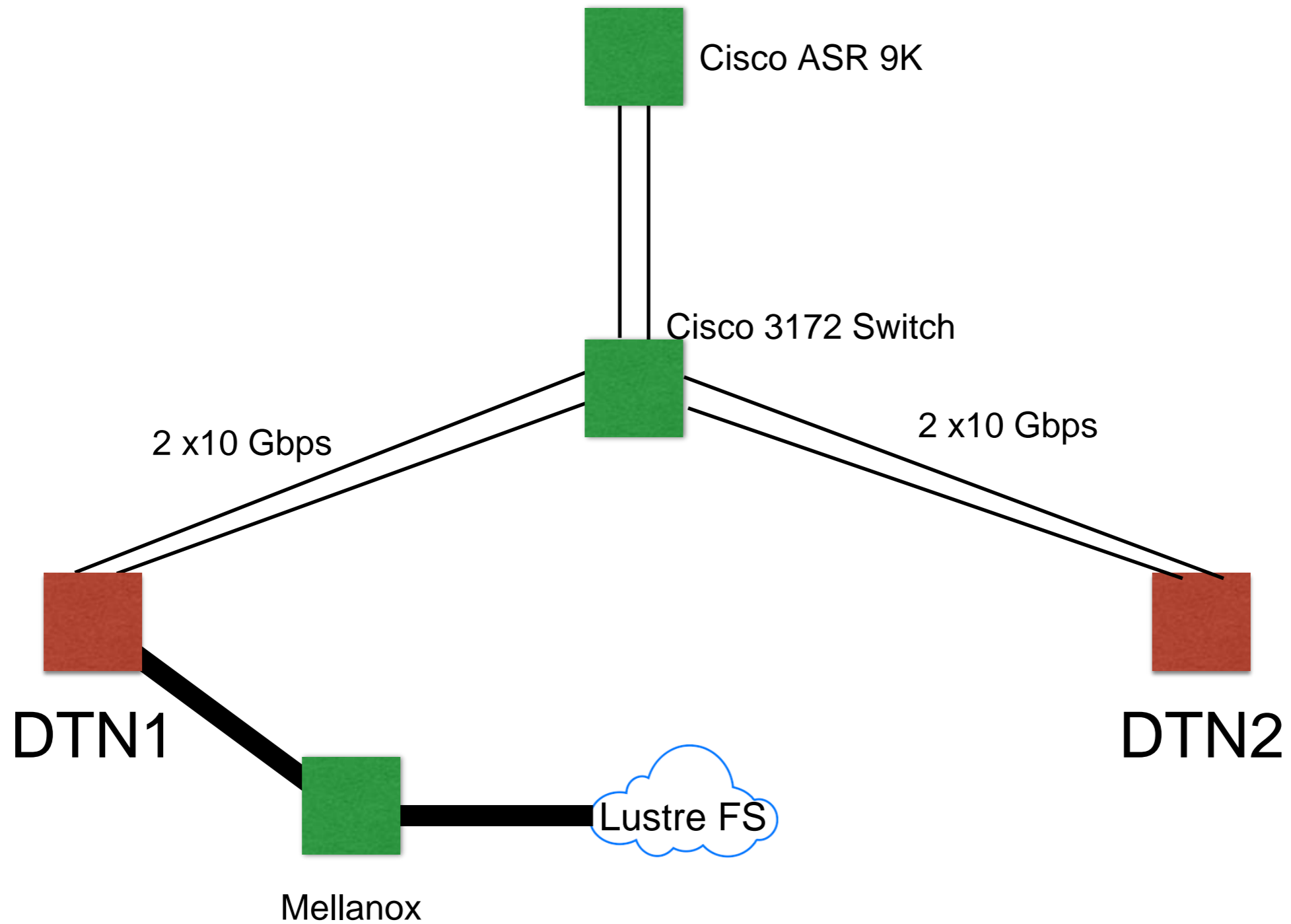
Goals For DTN

1. Work with PI's on our HPC to use DTN to support scientific workflows where we acquire and exchange data for regular processing.
2. Leverage DTN for more ad-hoc large file exchanges. Particularly, pulling down research datasets for use among multiple groups.

UMBC External Connectivity



DTN Connection



Two DTN Hosts

- DTN1 – Connected to HPC Lustre File System via Mellanox Switch for infiniband to ethernet and more restricted – limited to HPC users that are need to develop scientific workflows with specific research groups.
- DTN2 – Open to research groups for sharing or acquiring data.

Hardware

- Supermicro 2028-TRTP+ (dual 10GBE SFP+) 24 - 2.5" disk bays
- E5-2623 @ 3.00GHZ - 4 Core - 10MB Cache
- 128GB mem
- Intel 730 480GB SSDs (consumer grade) (20 - DTN1, 5-DTN2) + LSI 9300-16I controller (non-raid)
- 2 x 10Gbps bonded network uplinks
- Infiniband connect for DTN1 Lustre FS

Software

- Linux Red Hat 4.8.3-9
- ZFS - raid z3 (because the data integrity matters)
- Fasterdata tuning for TCP and Interrupt steering
- Grid FTP (Globus)
- Fail2ban, Cisco ACLs, other good practices

Tuning the DTN

- Tuning is changing quickly with newer kernels (like CentOS 7) being pre tuned / self tuning.
- Fasterdata for kernel and BIOS tuning
- Checkout - Brian Tierney's talk at TechEx 2016 on 100G tuning
 - <https://www.es.net/assets/Uploads/100G-Tuning-TechEx2016.tierney.pdf>
- Google's (Bottleneck Bandwidth and RTT) TCP congestion algorithm is seeing great gains in Google's high speed network.
 - A detailed description will be in ACM Queue Sep-Oct 2016 and is available for 4.8 and 4.9 kernels.

DTN Network Tuning

- Network tests - Checked spreading on 2x10Gb bundles.

```
3172_SDN2.noc.umbc.edu# show port-channel traffic
ChanId  Port Rx-Ucst Tx-Ucst Rx-Mcst Tx-Mcst Rx-Bcst Tx-Bcst
-----
  1 Eth1/47 50.85% 40.69% 89.80% 12.48% 100.00%  0.0%
  1 Eth1/48 49.14% 59.30% 10.19% 87.51%  0.0%  0.0%
-----
  2 Eth1/1  48.91% 51.54% 50.00% 48.86%  0.0% 100.00%
  2 Eth1/2  51.08% 48.45% 50.00% 51.13%  0.0%  0.0%
-----
  3 Eth1/3  48.50% 58.70% 50.00% 92.69%  0.0% 100.00%
  3 Eth1/4  51.49% 41.29% 50.00%  7.30%  0.0%  0.0%
```

Chan 1 is the link from the ASR9k
Chan 2 is the link from DTN1
Chan is the link from DTN2

- GridFTP tests (Network, Filesystem IO and CPU)

GridFTP

- `globus-url-copy -vb -fast -p 16`
`ftp://fiona.its.hawaii.edu:2811/export/data/10G.dat`
`file:///export/data/10G.out`
- 16 parallel streams used in out testing
- 10GB test file living at all test nodes

Local DTN Performance

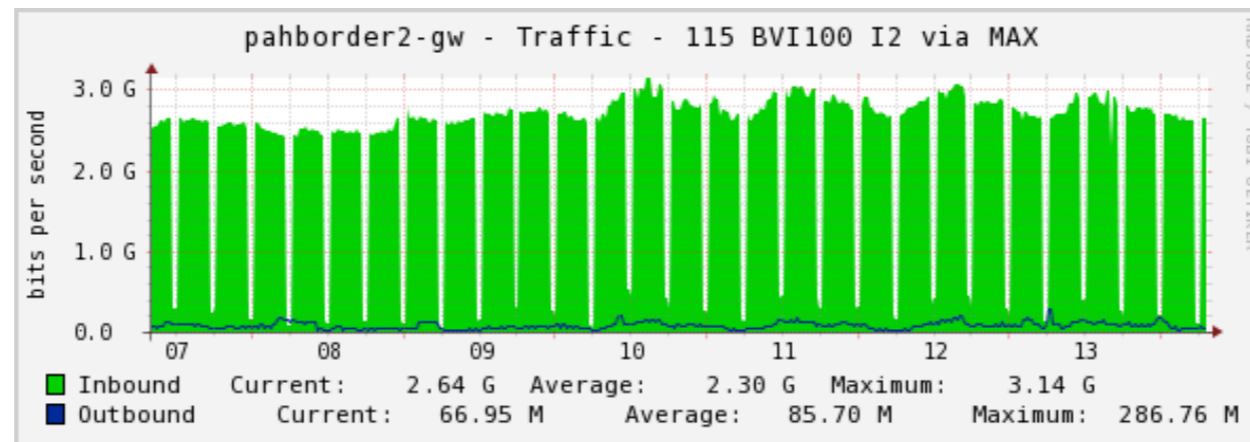
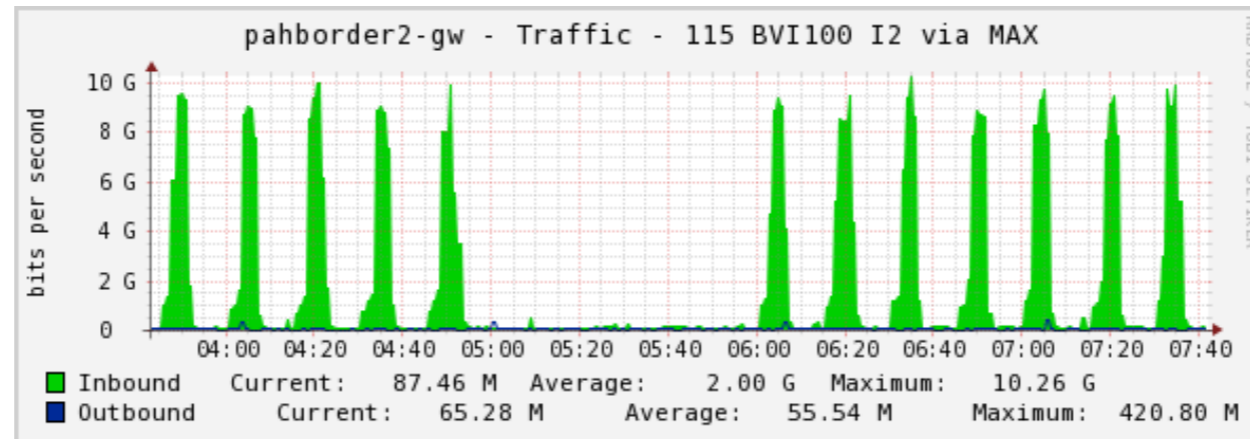
- local DTN2 10G file(in cache) to DTN2 /dev/null
 - 2330.41 MB/sec avg. 2217.24 MB/sec inst.
- Network - DTN1 10G file to DTN2 /dev/null
 - 1173.27 MB/sec avg. 1189.49 MB/sec inst.
- Net + FS - DTN1 10G file to DTN2 File system
 - 771.35 MB/sec avg. 688.99 MB/sec inst.

Performance Tuning

- MB/s to distant schools
- Mostly West Coast
- 771.35 - was our local upper bound

Site	P16	P1
hawaii.edu	91.57	23.95
uic.edu	871.17	654.27
washington.edu	653.28	223.79
ucdavis.edu	686.59	179.92
stanford.edu	215.02	184.78
ucmerced.edu	445.99	29.19
ucsc.edu	545.91	183.31
ucla.edu	295.59	108.40
ucr.edu	520.72	144.63
uci.edu	696.22	205.34
sdsu.edu	446.75	85.84

Testing Impact on I2 Link



- Testing bursts over 10Gb/s on I2 link
- Consistent over days of the week

Lessons Learned

- Hardware is a moving target for DTN performance
- Network dynamics are subtle (balancing ethernet LAGS was very configuration/hardware specific)
- Get help from a friend (John Graham at UCSD helped mentor our team)
- Tracking DTN-DTN performance is a good proxy for faculty research network performance. It is also useful for showing faculty the benefits of working with us to tune regular scientific workflows.

Great ESNET Links

- <https://fasterdata.es.net/science-dmz/DTN/>
- <https://fasterdata.es.net/science-dmz/DTN/tuning/>
- <https://fasterdata.es.net/science-dmz/DTN/hardware-selection/>
- <https://fasterdata.es.net/science-dmz/DTN/reference-implementation/>