Improving Scientific Outcomes at the APS with a Science DMZ

Jason Zurawski – zurawski@es.net
Science Engagement Engineer, ESnet
Lawrence Berkeley National Laboratory

GlobusWorld 2015
April 15th, 2015
Outline

• Background
• Current Network
• Pilot Network
• Future Work
Why Build A Science DMZ Though?

• What we know about scientific network use:
  – Machine size decreasing, accuracy increasing
  – HPC resources more widely available – and potentially distributed from where the scientists are
  – WAN networking speeds now at 100G, MAN approaching, LAN as well

• Value Proposition:
  – If scientists can’t use the network to the fullest potential due to local policy constraints or bottlenecks – they will find a way to get their done outside of what is available.

• Without a Science DMZ, this stuff is all hard
GM/CA

• General Medical Sciences and Cancer Institutes Structural Biology Facility @ APS
  – Focused on the study of crystallographic structure determination of biological macromolecules by X-ray diffraction

• Typical User:
  – Visits and spends some allocated time with samples and the beamline machinery
  – Mails samples, controls device remotely
  – In either case – data has to go ‘somewhere’
After processing on a supercomputer, models are created. Hundreds to thousands of images are created in a few hours...they can range in size from MB to TB.

Processing on this order of magnitude can't be done locally – we need to send (over a network) to a more capable facility.
Network as Infrastructure *Instrument*

**ESnet Vision**: Scientific progress will be completely unconstrained by the physical location of instruments, people, computational resources, or data.
Outline

• Background
• Current Network
• Pilot Network
• Future Work
Current Design
ANL Border to ESnet New York (perfSONAR Test)

Stable Throughput = ~3.5 Gbps

No Packet Loss
GM/CA to ESnet New York (perfSONAR Test)

Source: perfsonar-gmca.anl.gov - 164.54.103.141
Capacity: 10G
MTU: 1500

Destination: newy-pts.es.net - 198.124.238.54
Capacity: 10G
MTU: 9000

perfsonar-gmca.anl.gov - 164.54.103.141
Capacity: 10G
MTU: 1500

newy-owamp.es.net - 198.124.252.158
Capacity: 1.0G
MTU: 1500

Negative latency values found in the reverse direction. Typically, this occurs when one or both hosts’ clocks are out of sync, or the hosts are very close together.

Unpredictable & Low Throughput
(~350 Mbps -> 1.5 Gbps)

Constant Packet Loss
GM/CA to ANL Border (perfSONAR Test)

Variable Latency

Constant Packet Loss
A small amount of packet loss makes a huge difference in TCP performance

With loss, high performance beyond metro distances is essentially impossible

Throughput vs. Increasing Latency with .0046% Packet Loss

- Local (LAN)
- Metro Area
- Regional
- Continental
- International

Measured (TCP Reno)  Measured (HTCP)  Theoretical (TCP Reno)  Measured (no loss)

Throughput \leq \frac{\text{MSS}}{\text{RTT} \sqrt{P_{\text{loss}}}}
Globus Results

• GM/CA Downloading from:
  – ESnet ANL (~2ms)
    • Average: 285 Mbps
    • Max: 360 Mbps
  – ESnet BNL (~30ms)
    • Average: 26 Mbps
    • Max: 28 Mbps
  – ESnet LBL (~50ms)
    • Average: 16 Mbps
    • Max: 17 Mbps

• GM/CA Downloading from:
  – ESnet CERN (~110ms)
    • Average: 7 Mbps
    • Max: 8 Mbps
  – NERSC HPSS (~50ms)
    • Average: 127 Mbps
    • Max: 134 Mbps

• GM/CA Uploading to:
  – NERSC HPSS (~50ms)
    • Average: 112 Mbps
    • Max: 142 Mbps
Outline

• Background
• Current Network
• Pilot Network
• Future Work
Pilot Network

- **Plan:**
  - Use campus and building fiber resources for 2nd path
  - Applies to only a limited set of resources (perfSONAR, DTN)

- **Benefits**
  - Bypass congested local infrastructure
  - Apply targeted (vs. blanket) security policy

- **Cautions:**
  - Prevent just ‘anything’ from using fast path – policy to control this
  - Still need to figure out cause of local issues (e.g. this isn’t a pave over)
Results (as of 2pm CDT on April 13th)

DMZ Upgrade (~2pm CDT 4/13)

- Throughput (bps)
  - Regular Packet Loss
  - 1.4 Gbps
  - 1.8 Gbps
  - 2.2 Gbps
  - 1.1 Gbps
  - 0.1 Gbps

- Error (ppm)
  - 0.00%
  - 0.06%
  - 0.10%
  - 0.15%

- Source: 192.5.180.130 - 192.5.180.130
  - Capacity: 10G
  - MTU: 1500

- Destination: Engage es.net - 198.129.254.32
  - Capacity: 10G
  - MTU: 1500

- Previous 3d:
  - Sat Apr 11 00:00:00 2015 -- Tue Apr 14 00:00:00 2015

- Zoom: 1d 1w 1m 1y

© 2015, Energy Sciences Network
Results (as of 2pm CDT on April 13th)

Science DMZ Upgrade (~2pm CDT 4/13)
Outline

• Background
• Current Network
• Pilot Network
• Future Work
Future

- Wider use @ APS
  - Solution now goes directly to a specific beamline – how to support the entire facility?
- Wider use @ ANL
  - Service for other research groups (e.g. ALCF, ARM, etc.)
  - Pool of DTN resources w/ Globus, instead of each group manning their own (allows to back up to communal storage)
- Defining Policy
  - ACLs – ports exposed for things like perfSONAR, Globus. Shut off for things that don’t need it. Gray area is defining what is and is not science.
  - Who gets on, who doesn’t
  - Managing routing @ the border to best utilize the available WAN links
Participants & Thanks

• Argonne National Laboratory Networking
  – Corey Hall
  – Linda Winkler
  – Corby Schmitz

• Advanced Photon Source Networking
  – David Leibfritz
  – Mary Westbrook

• GM/CA Beamline
  – Sergey Stepanov
Improving Scientific Outcomes at the APS with a Science DMZ

Jason Zurawski – zurawski@es.net
Science Engagement Engineer, ESnet
Lawrence Berkeley National Laboratory