ESnet
ENERGY SCIENCES NETWORK

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

U.S. DEPARTMENT OF **ENERGY**
Office of Science

BERKELEY LAB

Outline

- Background
- Current Network
- Pilot Network
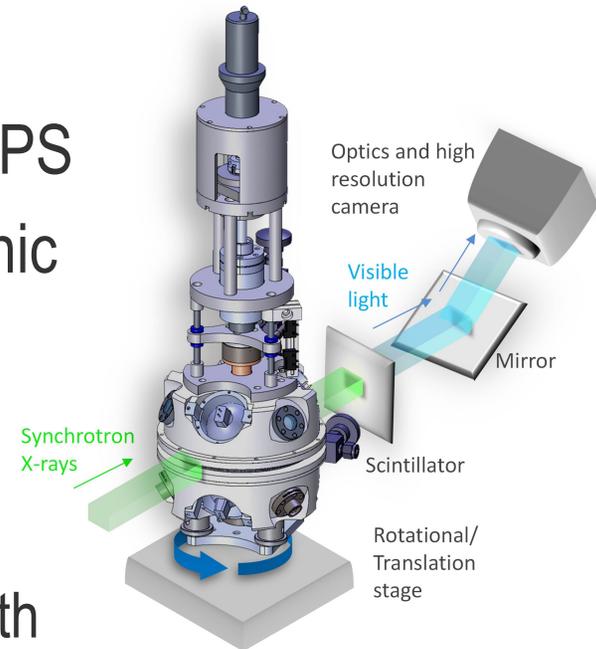- Future Work

ESnet

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

ESnet

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'

Optics and high resolution camera

Visible light

Mirror

Synchrotron X-rays

Scintillator

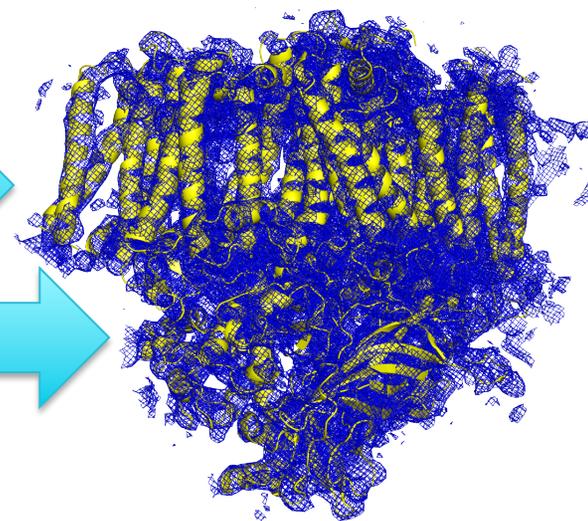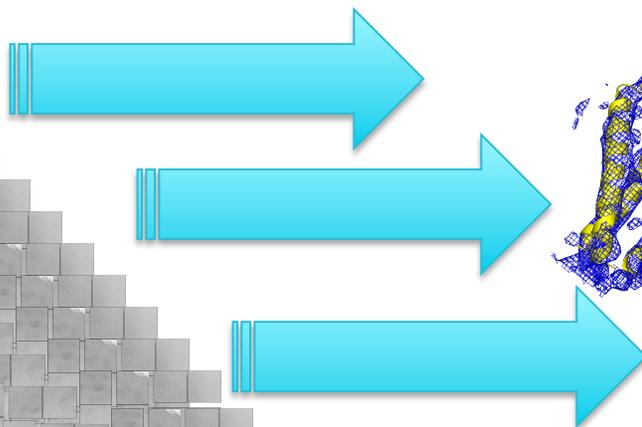Rotational/ Translation stage

ESnet

E Pluribus Unum

Processing on this order of magnitude can't be done locally – we need to send (over a network) to a more capable facility
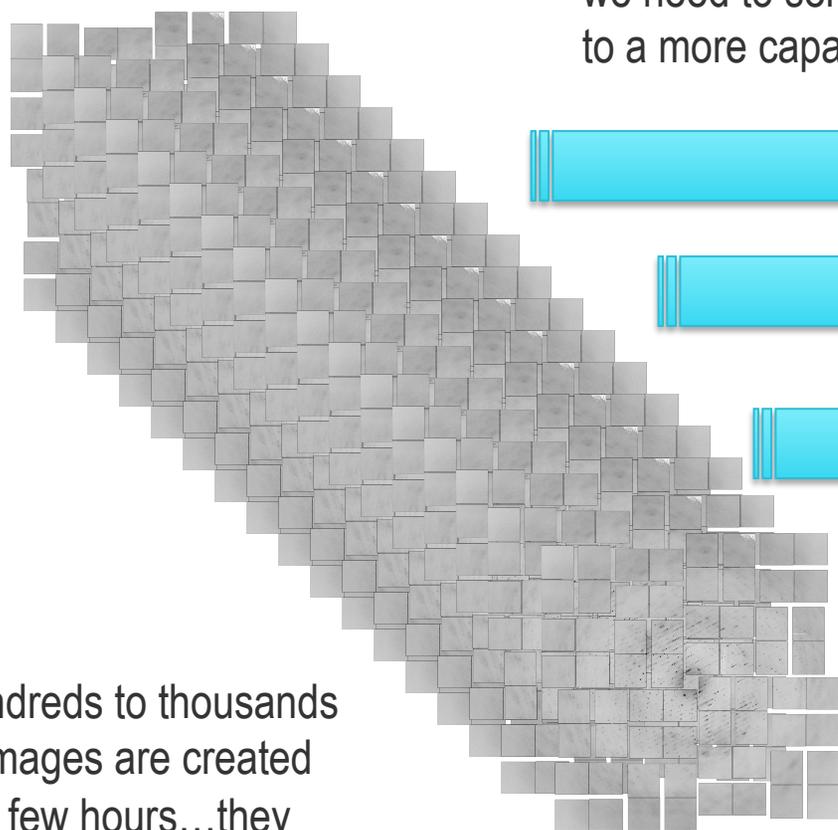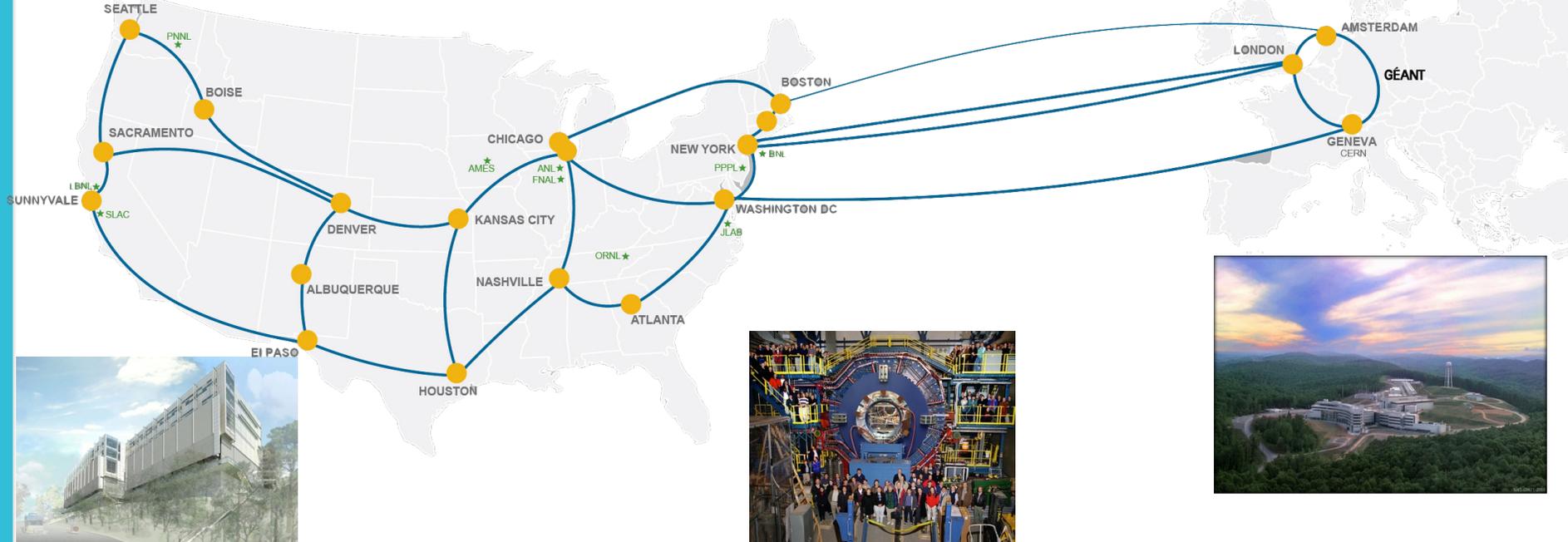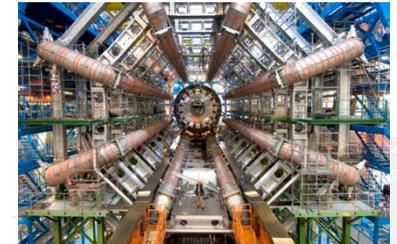
Hundreds to thousands of images are created in a few hours…they can range in size from MB to TB

After processing on a supercomputer, models are created.

ESnet

Network as ~~Infrastructure~~ *Instrument*

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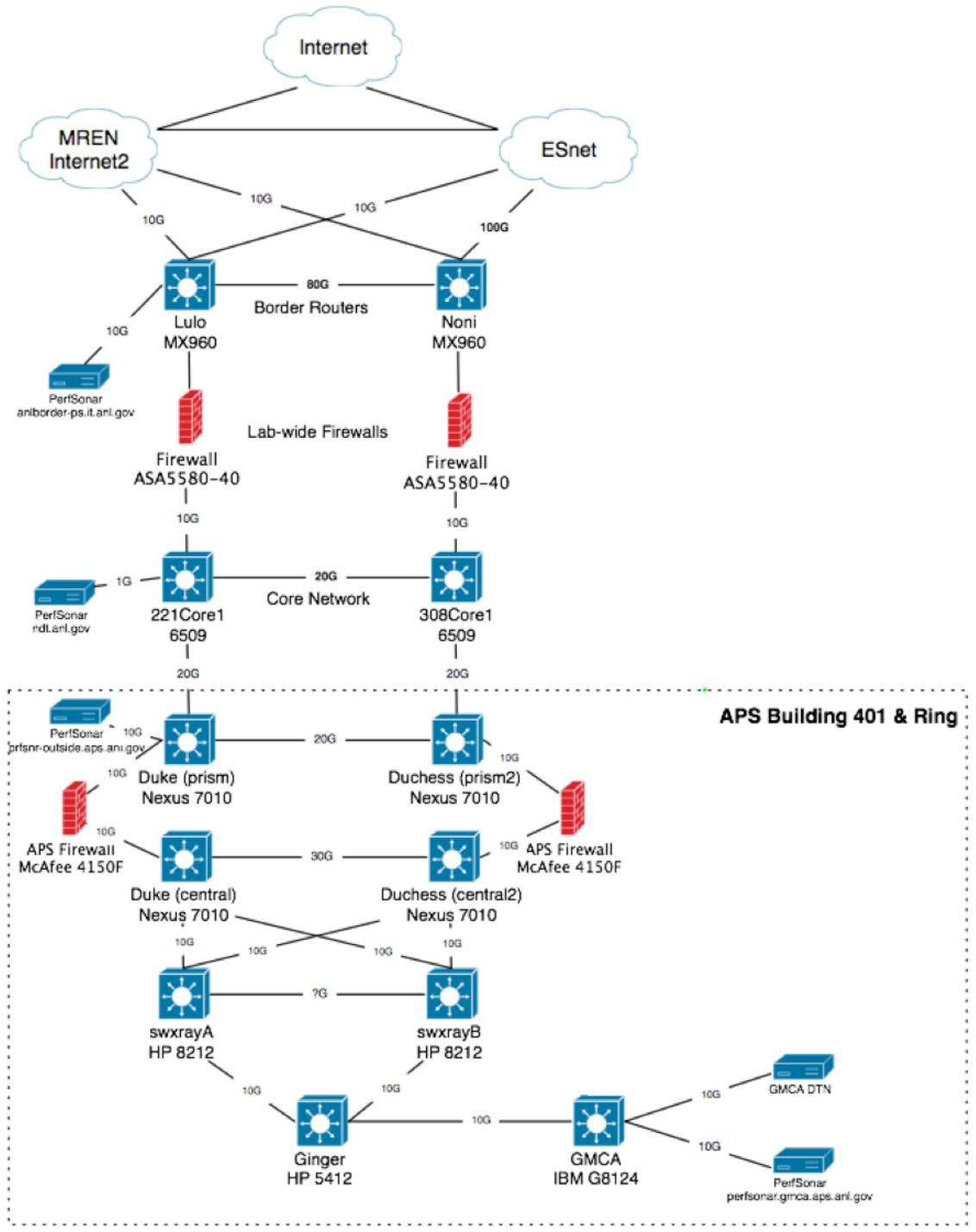
ESnet

Outline

- Background
- Current Network
- Pilot Network
- Future Work

ESnet

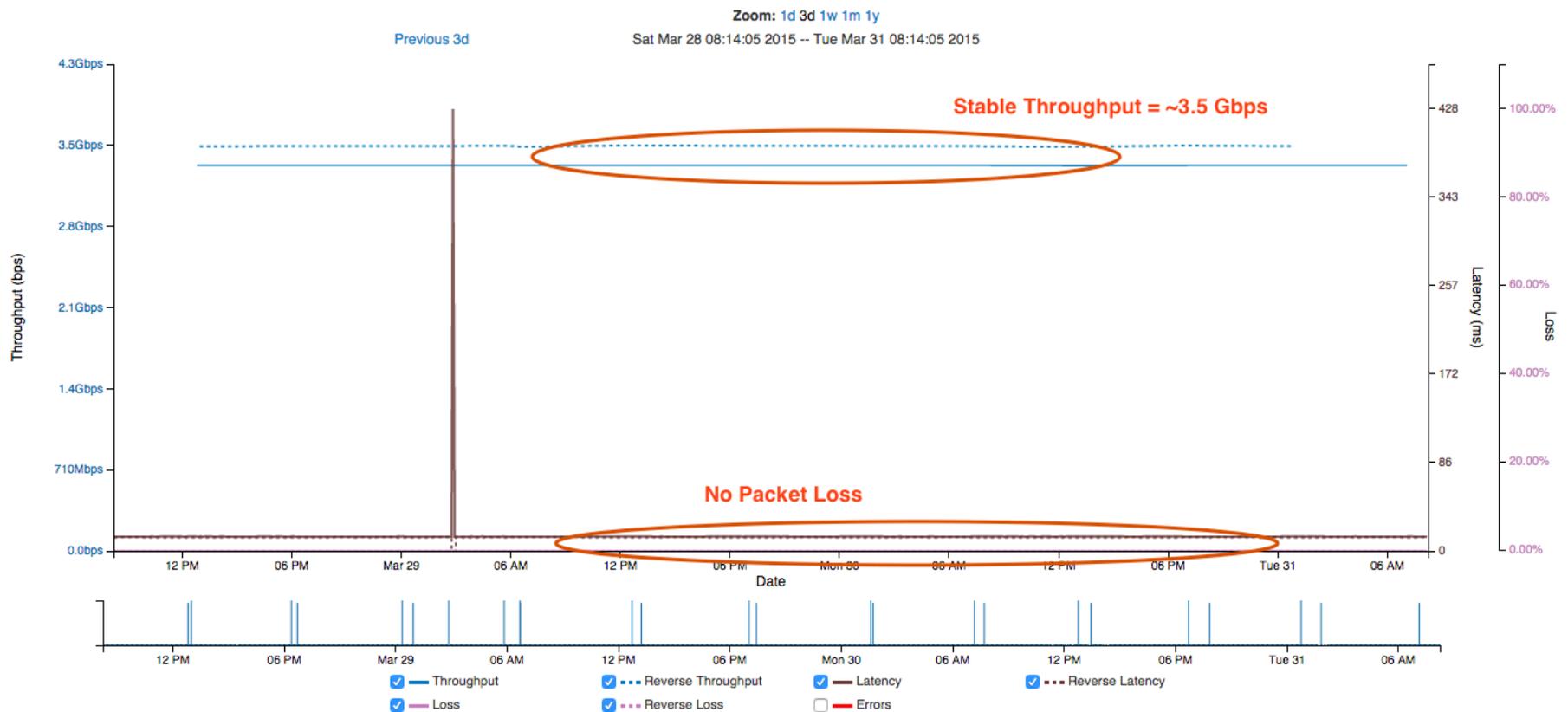# Current Design

ANL Border to ESnet New York (perfSONAR Test)

Source
aniborder-ps.it.anl.gov - 130.202.222.58   --   newy-pt1.es.net - 198.124.238.54
Capacity: 10G   MTU: 9000   Capacity: 10G   MTU: 9000

Destination

aniborder-ps.it.anl.gov - 130.202.222.58   --   newy-owamp.es.net - 198.124.252.158 [traceroute]
Capacity: 10G   MTU: 9000   Capacity: 1.0G   MTU: 1500

Link to this chart

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.

Zoom: 1d 3d 1w 1m 1y

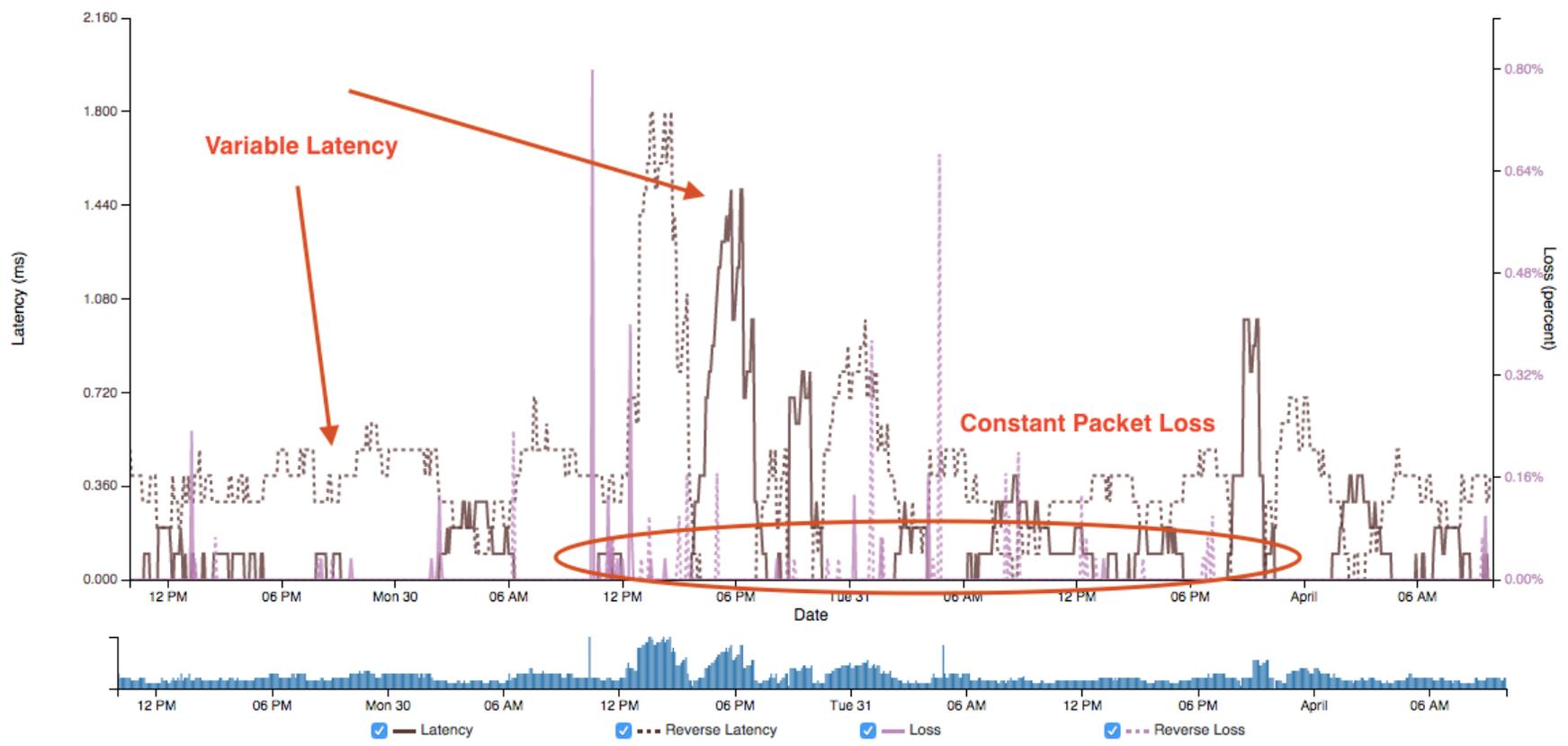Previous 3d   Sat Mar 28 08:14:05 2015 -- Tue Mar 31 08:14:05 2015

Stable Throughput = ~3.5 Gbps

No Packet Loss

☑ Throughput   ☑ Reverse Throughput   ☑ Latency   ☑ Reverse Latency
☑ Loss   ☑ Reverse Loss   ☐ Errors

GM/CA to ESnet New York (perfSONAR Test)
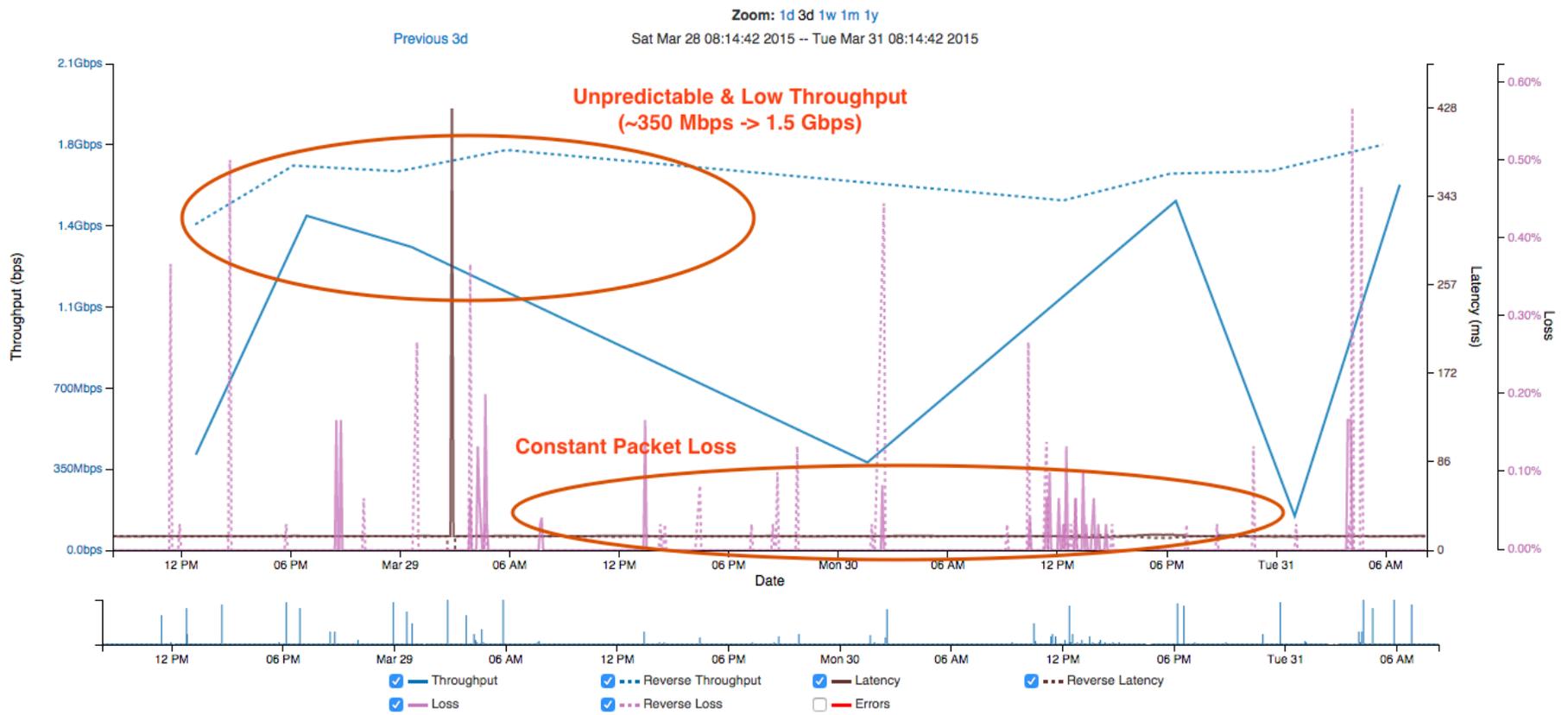
GM/CA to ANL Border (perfSONAR Test)

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

Throughput vs. Increasing Latency with .0046% Packet Loss

With loss, high performance beyond metro distances is essentially impossible

Labels on chart: Local (LAN), Metro Area, Regional, Continental, International

X-axis: Round Trip Time (milliseconds)
Y-axis: Throughput (Mbits/sec)
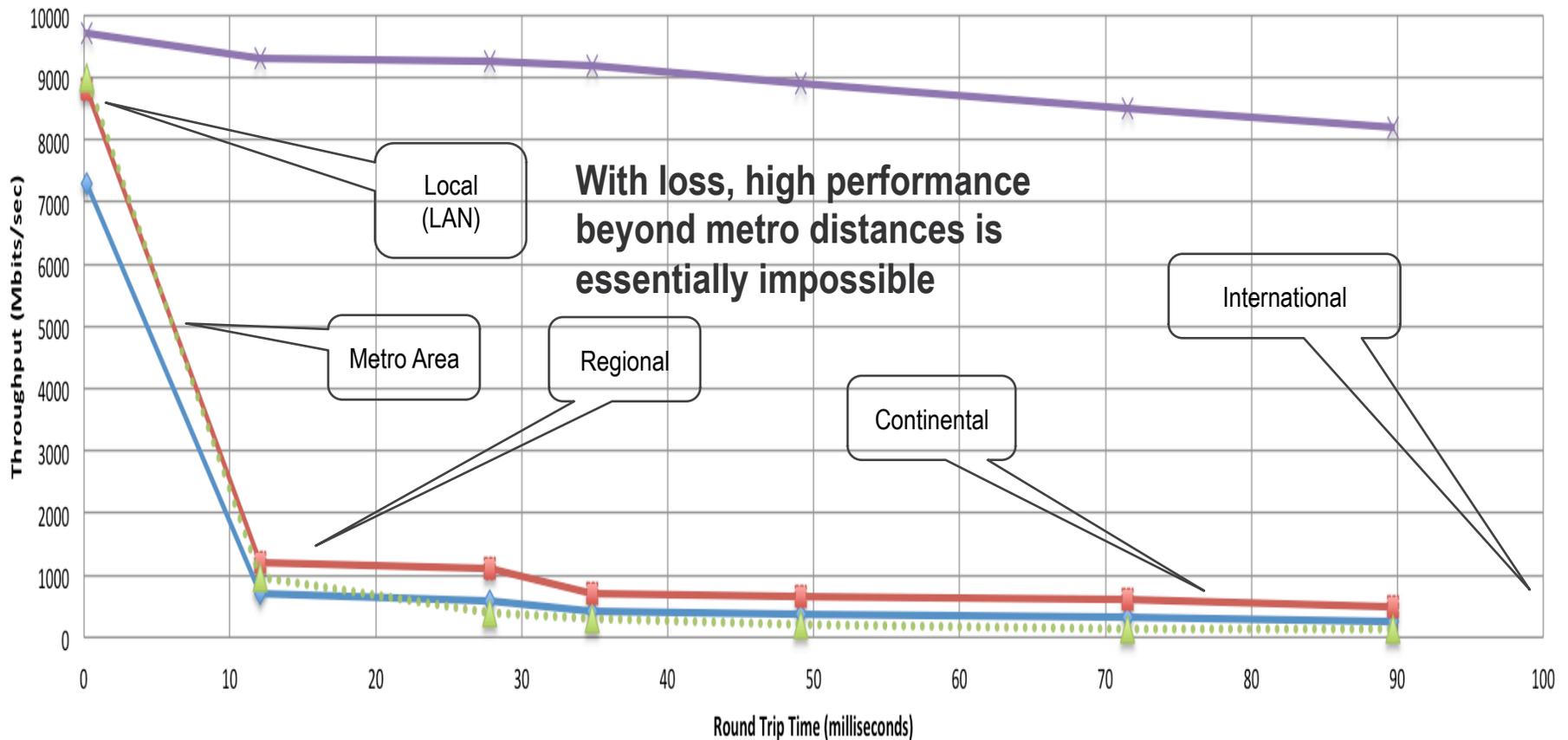
Legend:
- Measured (TCP Reno)
- Measured (HTCP)
- Theoretical (TCP Reno)
- Measured (no loss)

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

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)
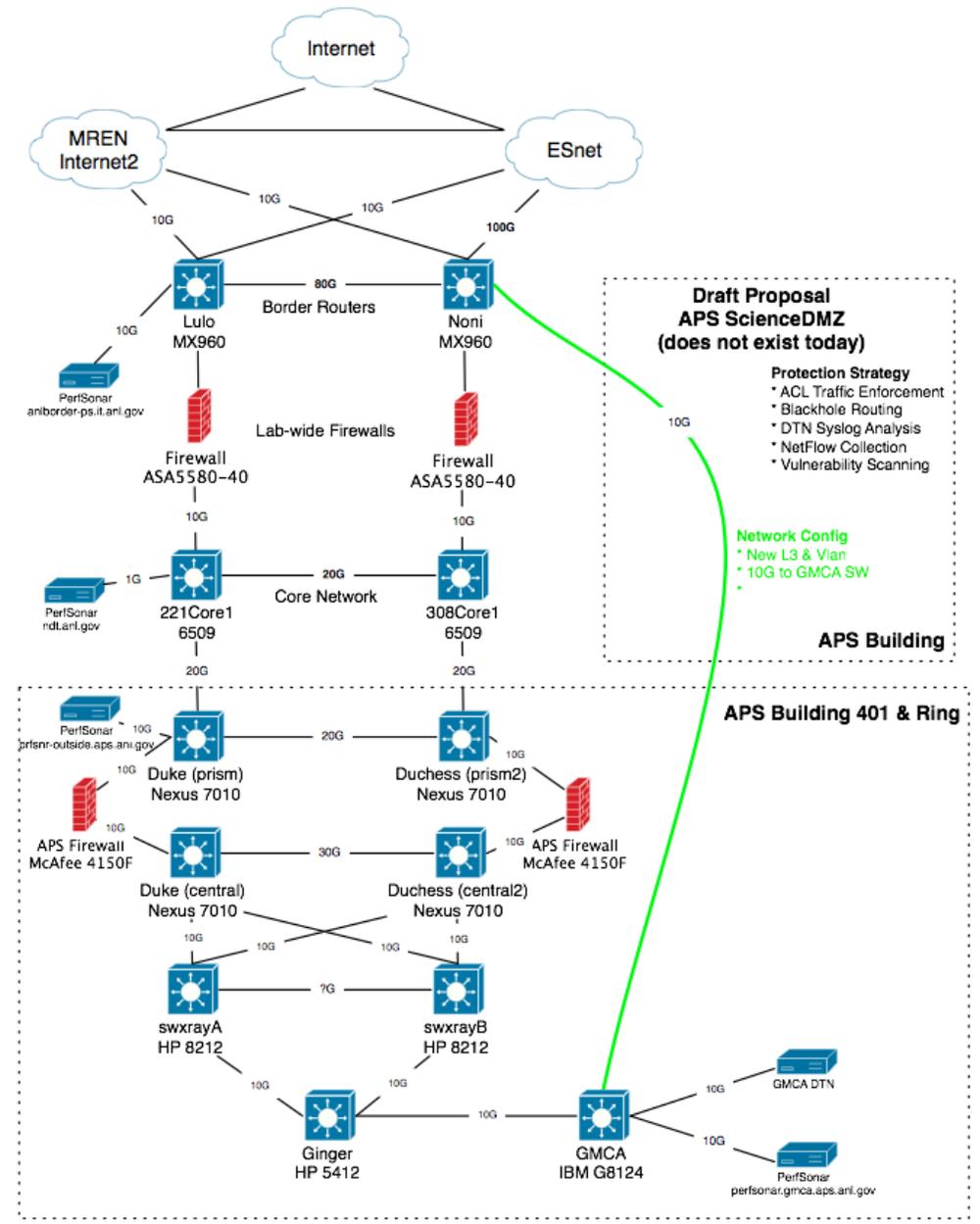   - Average: 112 Mbps
    - Max: 142 Mbps

ESnet

Outline

- Background
- Current Network
- Pilot Network
- Future Work

ESnet

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
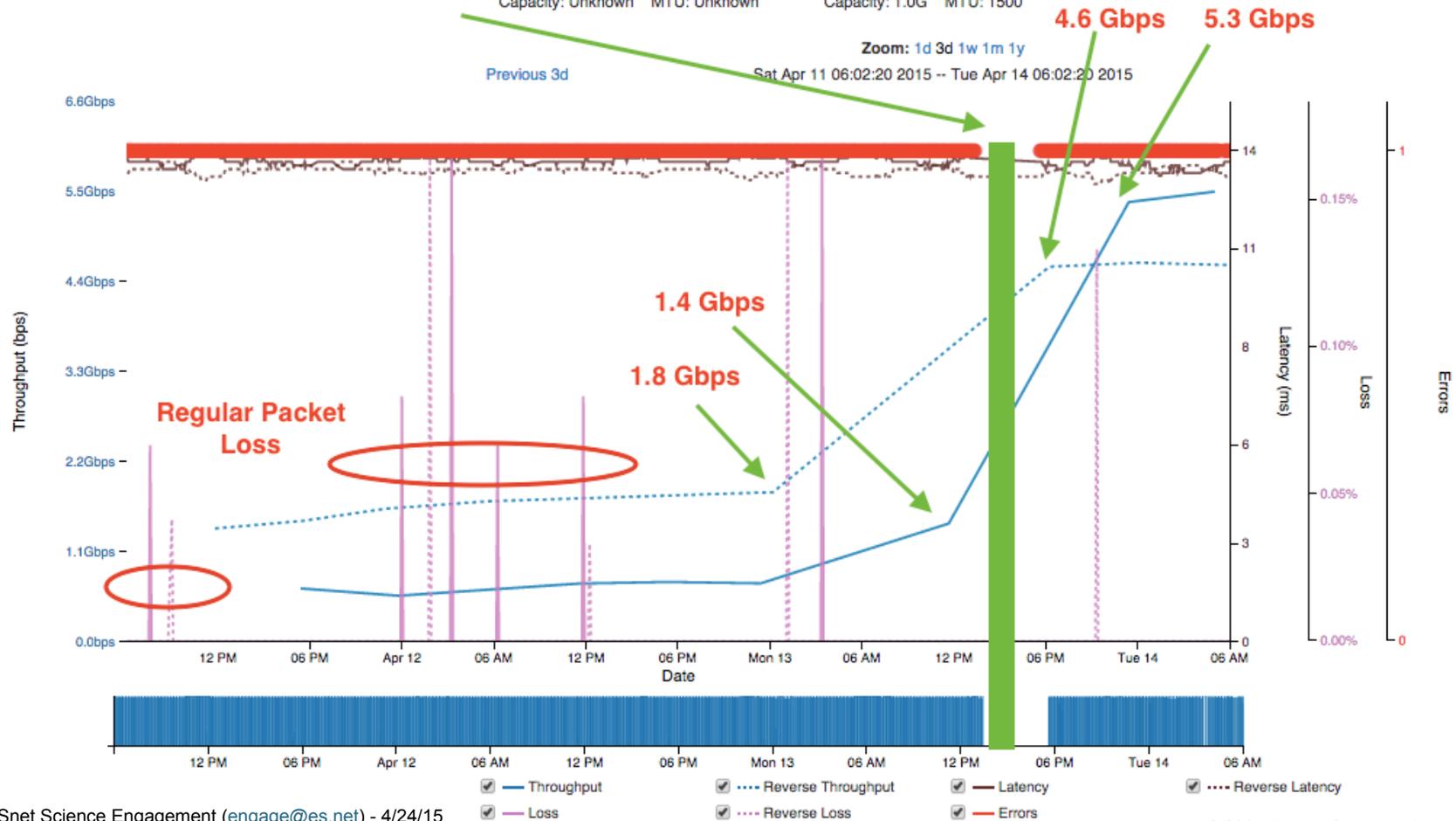 - Still need to figure out cause of local issues (e.g. this isn't a pave over)

Results (as of 2pm CDT on April 13ᵗʰ)

Results (as of 2pm CDT on April 13th)

Outline

- Background

- Current Network

- Pilot Network

- Future Work

ESnet

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
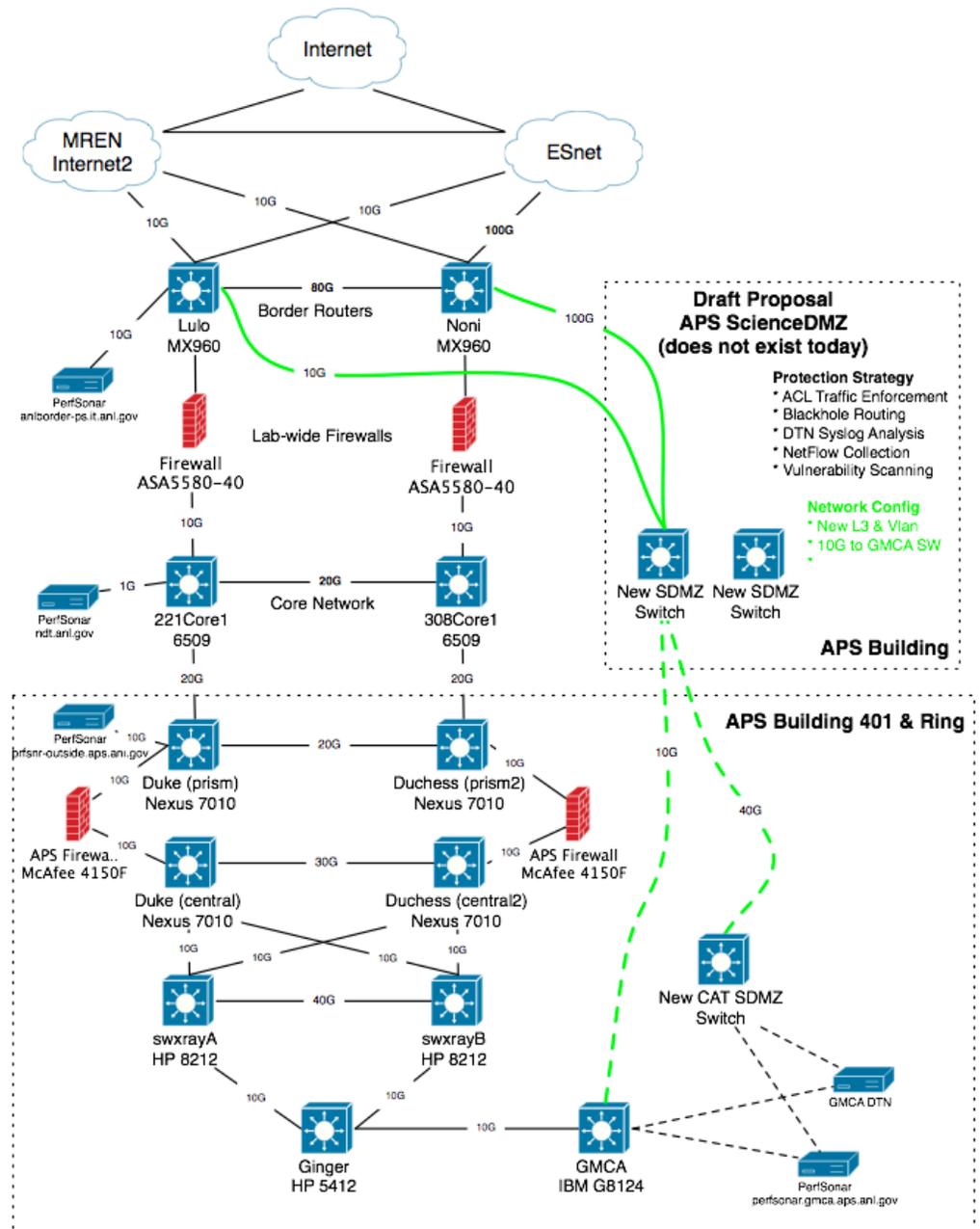
Participants & Thanks

- Argonne National Laboratory Networking
  - Corey Hall
  - Linda Winkler
  - Corby Schmitz

- Advanced Photon Source Networking
  - David Leibfritz
  - Mary Westbrook

- GM/CA Beamline
  - Sergey Stepanov

ESnet

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

U.S. DEPARTMENT OF **ENERGY**
Office of Science

BERKELEY LAB