Optimizing Data Management at the Advanced Light Source with a Science DMZ

Eli Dart, Network Engineer
ESnet Network Engineering Group

GlobusWorld 2013
Argonne, Il
April 17, 2013
Outline

Science DMZ background

ALS Workflow

Future Work
Science DMZ Background

The data mobility performance requirements for data intensive science are beyond what can typically be achieved using traditional methods

- Default host configurations (TCP, filesystems, NICs)
- Converged network architectures designed for commodity traffic
- Conventional security tools and policies
- Legacy data transfer tools (e.g. SCP)
- Wait-for-trouble-ticket operational models for network performance

The Science DMZ model describes a performance-based approach

- Dedicated infrastructure for wide-area data transfer
  - Well-configured data transfer hosts with modern tools
  - Capable network devices
  - High-performance data path which does not traverse commodity LAN
- Proactive operational models that enable performance
  - Well-deployed test and measurement tools (perfSONAR)
  - Periodic testing to locate issues instead of waiting for users to complain
- Security posture well-matched to high-performance science applications
Science DMZ – Simple Abstract Cartoon

- **WAN** connected to **Border Router** via a **Clean, High-bandwidth WAN path**.
- **Science DMZ Switch/Router** with **Per-service security policy control points**.
- **Enterprise Border Router/Firewall** connected to **Site / Campus LAN**.
- **High performance Data Transfer Node** with high-speed storage.
Science DMZ With Virtual Circuits/Openflow

Border Router

WAN

Clean, High-bandwidth path to/from WAN

10G Routed

10G Virtual Circuit

Dedicated path for virtual circuit traffic

Science DMZ Switch/Router

Per-service security policy control points

High performance Data Transfer Node with high-speed storage

Site/Campus Virtual Circuits

Enterprise Border Router/Firewall

Site / Campus access to Science DMZ resources

10GE

10GE

10GE

perfSONAR

perfSONAR

4/17/13
Science DMZ Supporting Multiple Projects

Diagram:
- **WAN** connected to **Border Router**
- **Border Router** connected to **Enterprise Border Router/Firewall**
- **Border Router** connected to **Science DMZ Switch/Router**
- **Science DMZ Switch/Router** connected to **Project A DTN**, **Project B DTN**, and **Project C DTN**
- **Clean, High-bandwidth WAN path** to **Site/Campus access to Science DMZ resources**
- **Per-project security policy control points**

4/17/13
The Science DMZ in 1 Slide

Consists of **three key components**, all required:

“Friction free” network path
- Highly capable network devices (wire-speed, deep queues)
- Virtual circuit connectivity option
- Security policy and enforcement specific to science workflows
- Located at or near site perimeter if possible

Dedicated, high-performance Data Transfer Nodes (DTNs)
- Hardware, operating system, libraries all optimized for transfer
- Includes optimized data transfer tools such as Globus Online and GridFTP

Performance measurement/test node
- perfSONAR

Details at [http://fasterdata.es.net/science-dmz/](http://fasterdata.es.net/science-dmz/)
Photon Science Data Increase

Many detectors are semiconductors

- Similar technology to digital cameras
- Exponential growth
- Increase in sensor area (512x512, 1024x1024, 2048x2048, …)
- Increase in readout rate (1Hz, 10Hz, 100Hz, 1kHz, 1MHz, …)

Data infrastructure needs significant change/upgrade

- Most photon scientists are not “computer people”
  - Different from HEP, HPC centers
  - They need data issues solved – they don’t want to solve them
  - They should not have to be come network experts!
- Physical transport of portable media has reached breaking point
- Default configs no longer perform well enough

4/17/13
ALS Beamline 8.3.2

Broad science portfolio: Applied science, biology, earth sciences, energy, environmental sciences, geology, cosmological chemistry

Detector upgrade → large increase in data rate/volume (50x)

Detector output: sets of large TIFF files

Beamline scientist Dula Parkinson reached out to LBLnet

LBLnet reached out to ESnet

Infrastructure improvements

- Used perfSONAR to find failing router line card
- DTN built from Fasterdata reference design

NERSC collaboration

- Data workflow (python scripts, etc.)
- Data analysis

Collaboration is ongoing
Original Workflow Infrastructure

Acquisition Computer

CameraLink

Detector

1GE

Drive share

1GE

Drive share

Advanced Light Source Network

Windows File Server
(5TB of disk storage)

LBNL core network

10GE

10GE

perfSONAR

10GE

Workstation

Workstation

Workstation

Data export via portable USB disk

ESnet

WORLD

perfSONAR

perfSONAR

10GE

perfSONAR

perfSONAR
Original Workflow

1. Data acquisition uses LabView
2. Data written to shared drive on firewalled Windows server
3. Analysis done on workstations
   • High-powered Windows hosts
   • Mix of proprietary and open-source tools
   • Scientists can be physically present or use Remote Desktop
4. Data export post-analysis – physical transport of portable media
   • USB hard drives
Improved Workflow

1. Data acquisition uses LabView
2. Data written to SAMBA share on beamline Linux DTN
   • Writing via SAMBA is faster than local disk on acquisition computer
   • Data are TIFF files ~10MB each, ~1000 files per data set
   • Current max performance is ~200MB/sec
3. Automated workflow pushes data to NERSC for analysis
   • Workflow managed using signal files
   • Data set is rolled up into an HDF5 file
   • Python scripts drive Globus Online CLI
   • Data transferred to NERSC DTNs at ~300MB/sec
4. Analysis results pulled back to beamline DTN
   • Additional analysis done on workstations (still use some proprietary tools)
   • Primary data export is via Globus Online from beamline DTN
Future Work

Stop here – hats off to the G.O. folks, esp. Raj and Ian

- Good tools
- Responsive support
- Openness to feature requests
- Increases in scientific productivity

Prioritize data acquisition over other operations on DTN

Generalizable config for windows DAQ and Linux DTN

Lots of AAA questions

- How to integrate with existing systems
- Experiment-specific credentials

Integration with portals

Integration with experiment software
Wrap

Good data mobility tools are a critical part of the Science DMZ model

- Interface to “the network” for many users
- Globus Online provides a good combination of usability and power

Photon science is seeing significant data rate/volume increases

- Increased infrastructure requirements
- Change in workflow
- Need for collaboration with experts: networks, systems, software

One example shown here – ALS beamline 8.3.2

This will need to be replicated – many facilities, many beamlines
Questions?

Thanks!

Eli Dart - dart@es.net
http://www.es.net/
http://fasterdata.es.net/