The Science DMZ

Eli Dart, ESnet
Lawrence Berkeley National Laboratory

CC-NIE PI Meeting
Washington, DC
May 1, 2014
Motivation

Networks are an essential part of data-intensive science
• Connect data sources to data analysis
• Connect collaborators to each other
• Enable machine-consumable interfaces to data and analysis resources (e.g. portals), automation, scale

Performance is critical
• Exponential data growth
• Constant human factors
• Data movement and data analysis must keep up

Effective use of wide area (long-haul) networks by scientists has historically been difficult
A small amount of packet loss makes a huge difference in TCP performance

With loss, high performance beyond metro distances is essentially impossible
Working With TCP In Practice

Far easier to support TCP than to fix TCP
- People have been trying to fix TCP for years – limited success
- Like it or not we’re stuck with TCP in the general case

Pragmatically speaking, we must accommodate TCP
- Sufficient bandwidth to avoid congestion
- Zero packet loss
- Verifiable infrastructure
  - Networks are complex
  - Must be able to locate problems quickly
  - Small footprint is a huge win – small number of devices so that problem isolation is tractable
The Science DMZ Design Pattern

Dedicated Systems for Data Transfer

- High performance
- Configured specifically for data transfer
- Proper tools

Network Architecture

- Dedicated network location for high-speed data resources
- Appropriate security
- Easy to deploy - no need to redesign the whole network

Performance Testing & Measurement

- perfSONAR
- Enables fault isolation
- Verify correct operation
- Widely deployed in ESnet and other networks, as well as sites and facilities
Science DMZ Design Pattern (Abstract)

Border Router

Clean, High-bandwidth WAN path

Site / Campus access to Science DMZ resources

Per-service security policy control points

High performance Data Transfer Node with high-speed storage

WAN

Science DMZ Switch/Router

Enterprise Border Router/Firewall

Site / Campus LAN

High performance Data Transfer Node with high-speed storage
Local And Wide Area Data Flows

Border Router

Clean, High-bandwidth WAN path

Site / Campus access to Science DMZ resources

Science DMZ Switch/Router

Per-service security policy control points

High performance Data Transfer Node with high-speed storage

WAN

Enterprise Border Router/Firewall

Site / Campus LAN

High Latency WAN Path

Low Latency LAN Path
Abstract HPC Center With Data Path
Common Threads

Common threads exist in all these examples

Accommodation of TCP
- Wide area portion of data transfers traverses purpose-built path
- High performance devices that don’t drop packets

Ability to test and verify
- When problems arise (and they always will), they can be solved if the infrastructure is built correctly
- Small device count makes it easier to find issues
- Multiple test and measurement hosts provide multiple views of the data path
  - perfSONAR nodes at the site and in the WAN
  - perfSONAR nodes at the remote site

Security policy well-matched to the science workflow
Advanced Services

Elimination of packet loss is a focus of a Science DMZ
- In many cases, that’s the bread and butter
- Huge wins in performance → increased scientific productivity

Science DMZ is also a good fit for advanced services
- Openflow
- Other virtual circuit services (OSCARS, ION, etc.)
- Other things that a converged commodity infrastructure can’t do

In order to be useful for advanced services, the Science DMZ must be flexible
- Hard to support something new if you’re stuck with yesterday’s rigid technology
- This is one of the arguments against enterprise firewalls
Science DMZ Rooted In Pragmatism

The global science complex has several clear and present needs

- Multi-facility workflows
- Dramatic data increases – rate, volume, analysis
- Deterministic behavior at high performance levels

We (“the network people”) are the people who have to enable all this

- Most science communities don’t have the internal capability and need to bring in expertise from outside
- There are exceptions – some collaborations have the scale (hello, LHC…)
  - The LHC experiments are a partner in innovation
  - Many ideas developed for the LHC experiments work in other environments, some do not
- In the general case we must be the knowledge base, the center of excellence, the people who can help
  - After all, we build and run it 😊
The Science DMZ Design Pattern, Enhanced

**Data Transfer Node**
- High performance
- Configured for data transfer
- Proper tools

**Engagement**
- Partnerships
- Education & Consulting
- Resources & Knowledgebase

**perfSONAR**
- Enables fault isolation
- Verify correct operation
- Widely deployed in ESnet and other networks, as well as sites and facilities

**Science DMZ**
- Dedicated location for DTN
- Proper security
- Easy to deploy - no need to redesign the whole network
Human Capability Matters

We are building the cyberinfrastructure together
- We’re all in this together
- Networking is end to end

The thing that matters is that the scientists can use it
- Many don’t know we exist
- Many more only know of problems – they think we’re in the way
- We have work to do

Cyberinfrastructure is a great enabler
We must not neglect the human piece
- People who know how to build and drive the network
- People who use the network for knowledge discovery
Strategic Capabilities

Science DMZ, wide area networks, collaborators all work together

• Strategic capabilities for big team science
• Next-generation capabilities for smaller collaborations with data scale shock

Organizations collaborating internally

• Networking / IT
• Security
• Science

Our job is to make it work, and to show others how to use it
Links

- ESnet fasterdata knowledge base
  - [http://fasterdata.es.net/](http://fasterdata.es.net/)
- Science DMZ paper
  - [http://www.es.net/assets/pubs_presos/sc13sciDMZ-final.pdf](http://www.es.net/assets/pubs_presos/sc13sciDMZ-final.pdf)
- Science DMZ email list
  - [https://gab.es.net/mailman/listinfo/sciencedmz](https://gab.es.net/mailman/listinfo/sciencedmz)
- perfSONAR
  - [http://fasterdata.es.net/performance-testing/perfsonar/](http://fasterdata.es.net/performance-testing/perfsonar/)
  - [http://www.perfsonar.net/](http://www.perfsonar.net/)
- Additional material
  - [http://fasterdata.es.net/science-dmz/](http://fasterdata.es.net/science-dmz/)
  - [http://fasterdata.es.net/host-tuning/](http://fasterdata.es.net/host-tuning/)
Thanks!

Questions?

Eli Dart – dart@es.net
http://www.es.net/
http://fasterdata.es.net/
Thanks!

Questions?

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