

The Science DMZ

Eli Dart, Energy Sciences Network (ESnet)

TERENA Network Architects and TF-NOC

Prague, Czech Republic

November 13, 2013



Outline



- Motivation
- The Science DMZ Design Pattern
- Security
- Futures
- Wrap

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 networks by scientists has historically been difficult



The Central Role of the Network

The very structure of modern science assumes science networks exist: high performance, feature rich, global scope

What is important?

1. Correctness
2. Consistency
3. Performance

What is “The Network” anyway?

- “The Network” is the set of devices and applications involved in the use of a remote resource
 - This is not about supercomputer interconnects
 - This is about data flow from experiment to analysis, between facilities, etc.
- User interfaces for “The Network” – portal, data transfer tool, workflow engine
- Therefore, servers and applications must also be considered



TCP – Ubiquitous and Fragile

Networks provide connectivity between hosts – how do hosts see the network?

- From an application's perspective, the interface to “the other end” is a socket
- Communication is between applications – mostly over TCP

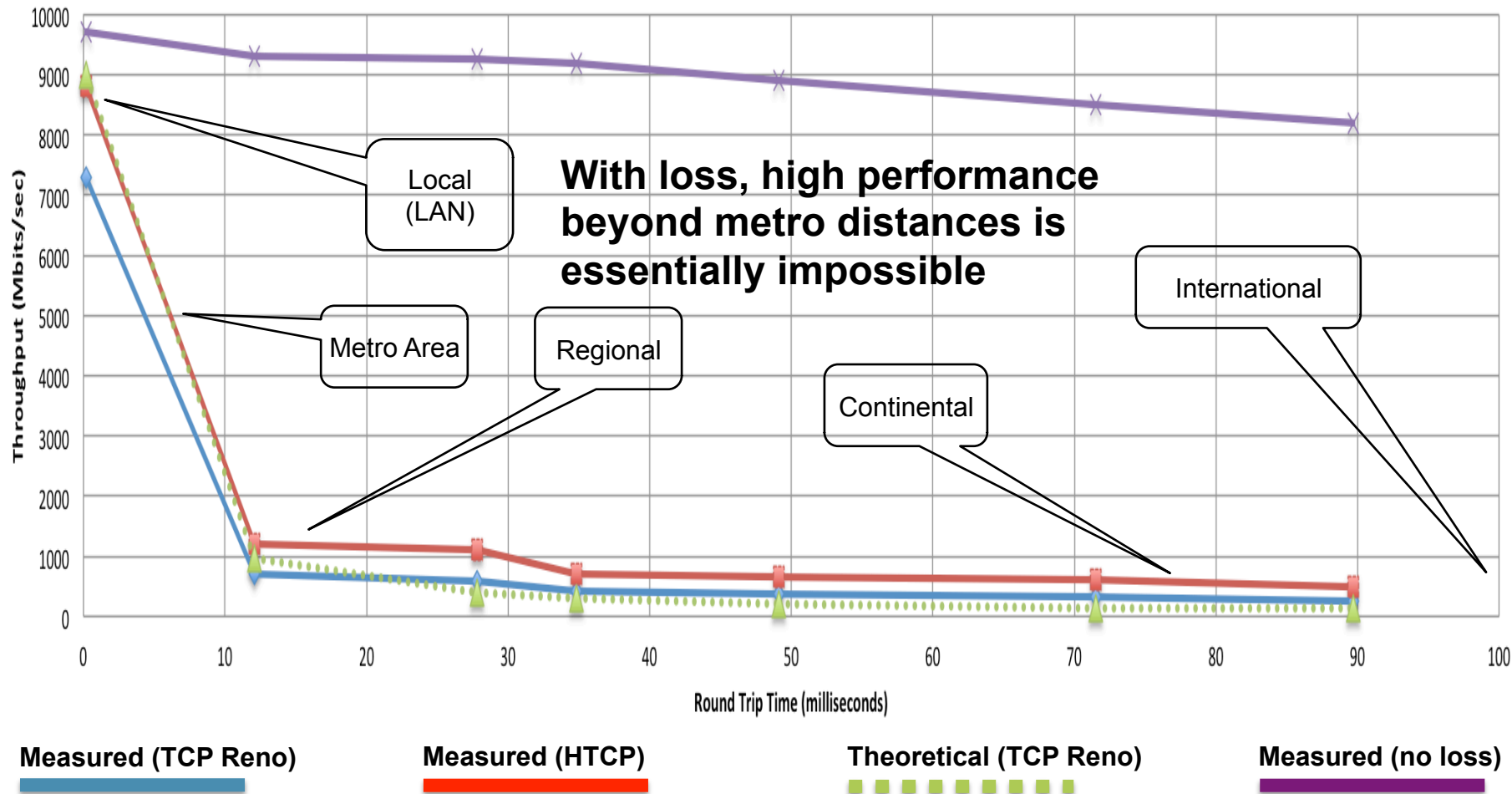
TCP – the fragile workhorse

- TCP is (for very good reasons) timid – packet loss is interpreted as congestion
- Packet loss in conjunction with latency is a performance killer
- Like it or not, TCP is used for the vast majority of data transfer applications

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



Throughput vs. Increasing Latency with .0046% Packet Loss





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
 - Must be able to prove a network device or path is functioning correctly
 - Small footprint is a huge win – small number of devices so that problem isolation is tractable



The Science DMZ Design Pattern

Effective support for TCP-based data transfer

- Designed for correct, consistent, high-performance operation
- Easy to troubleshoot
- Cybersecurity – defensible without compromising performance

Borrow ideas from traditional network security

- Traditional DMZ – separate enclave at network perimeter (“Demilitarized Zone”)
 - For WAN-facing services
 - Clean policies
 - Well-supported by proper hardware
- Do the same thing for science – Science DMZ

Science DMZ Design Pattern Components



Dedicated Systems for Data Transfer

Data Transfer Node

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

Network Architecture

Science DMZ

- Dedicated location for Data Transfer Node
- 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 – Network Architecture



Dedicated
Systems for
Data Transfer

Data Transfer Node

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

Network
Architecture

Science DMZ

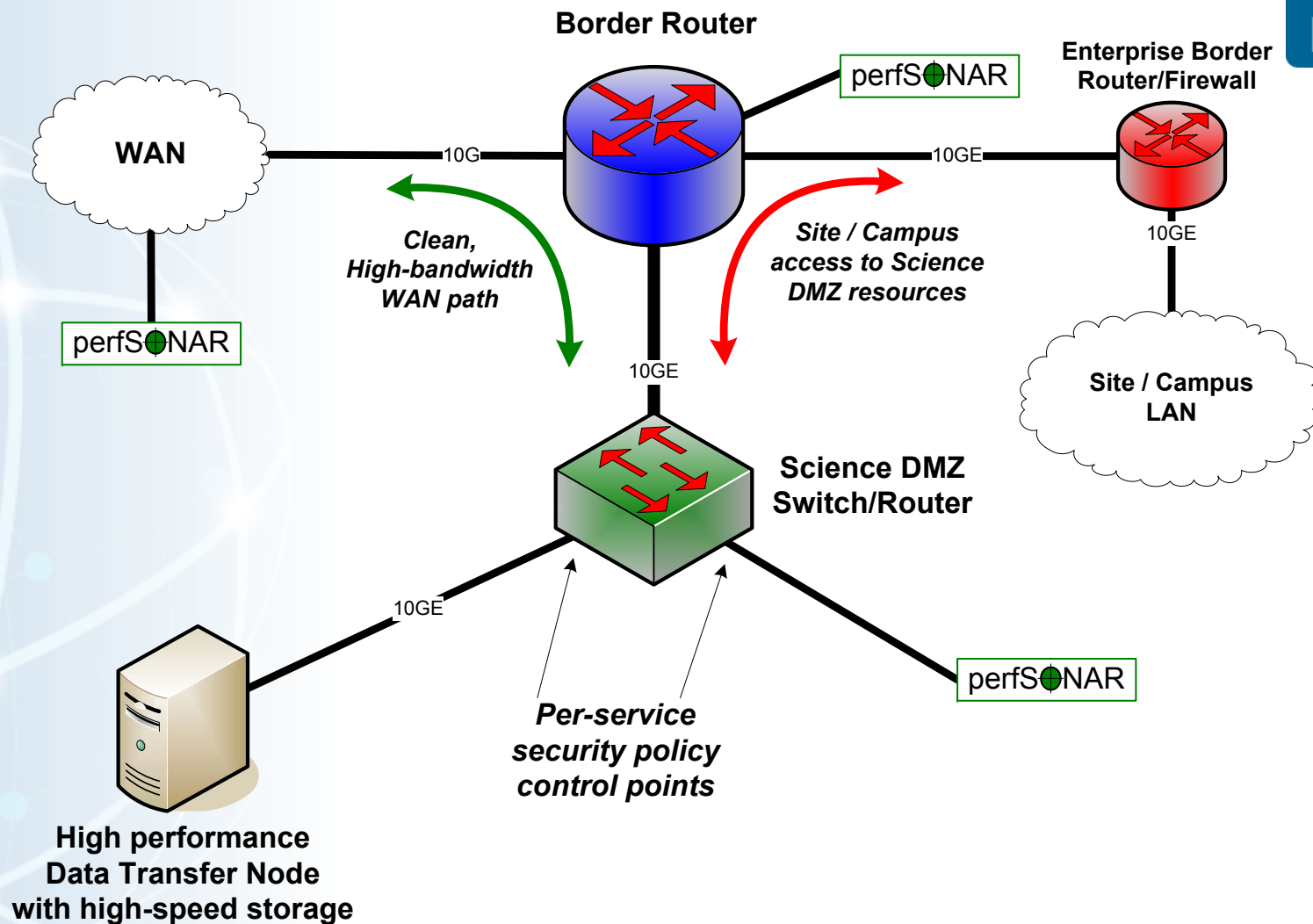
- Dedicated location for Data Transfer Node
- Appropriate security
- Easy to deploy - no need to redesign the whole network

Performance
Testing &
Measurement

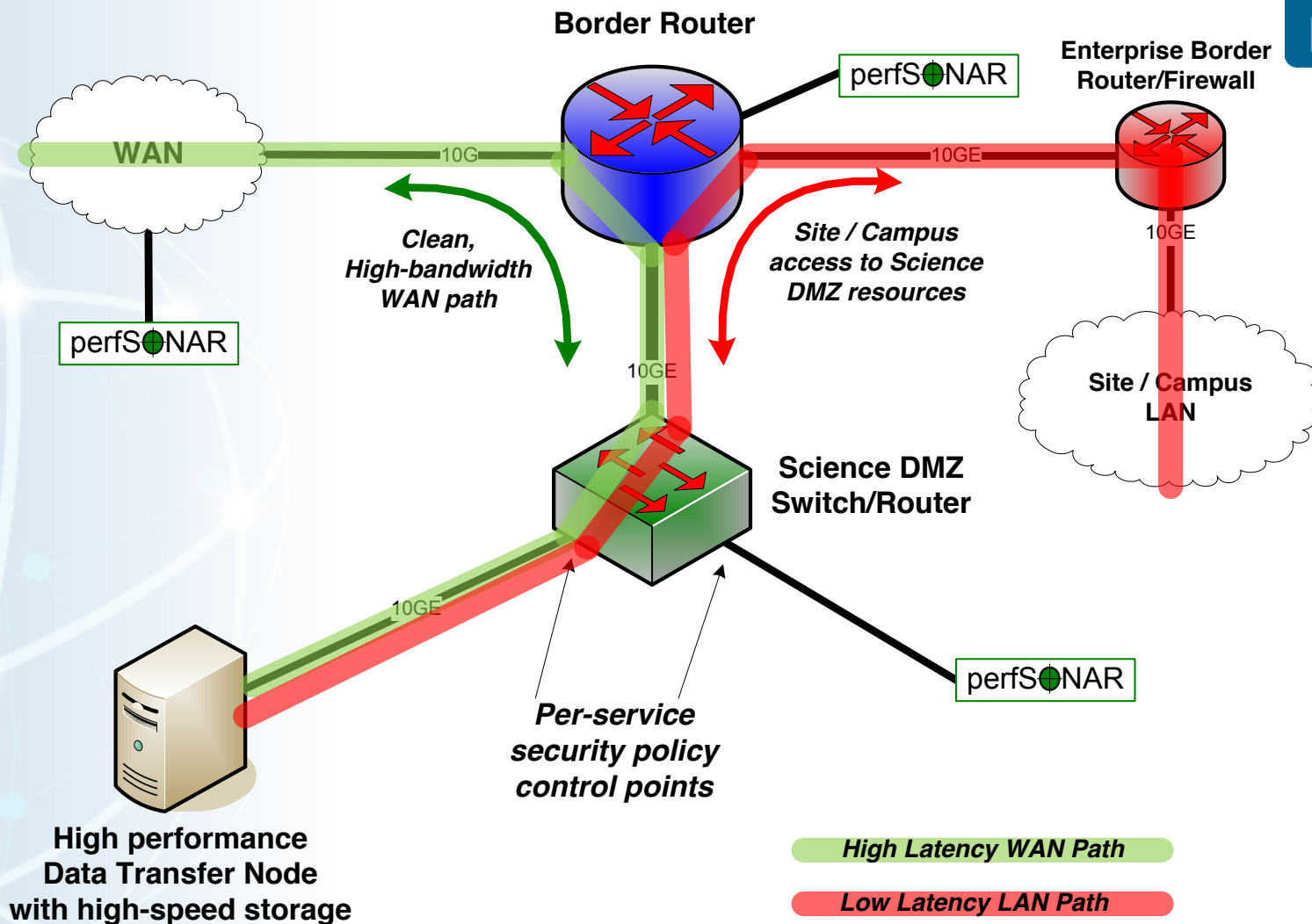
perfSONAR

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

Science DMZ Design Pattern (Abstract)



Local And Wide Area Data Flows



Supercomputer Center Deployment



High-performance networking is assumed in this environment

- Data flows between systems, between systems and storage, wide area, etc.
- Global filesystem often ties resources together
 - Portions of this may not run over Ethernet (e.g. IB)
 - Implications for Data Transfer Nodes

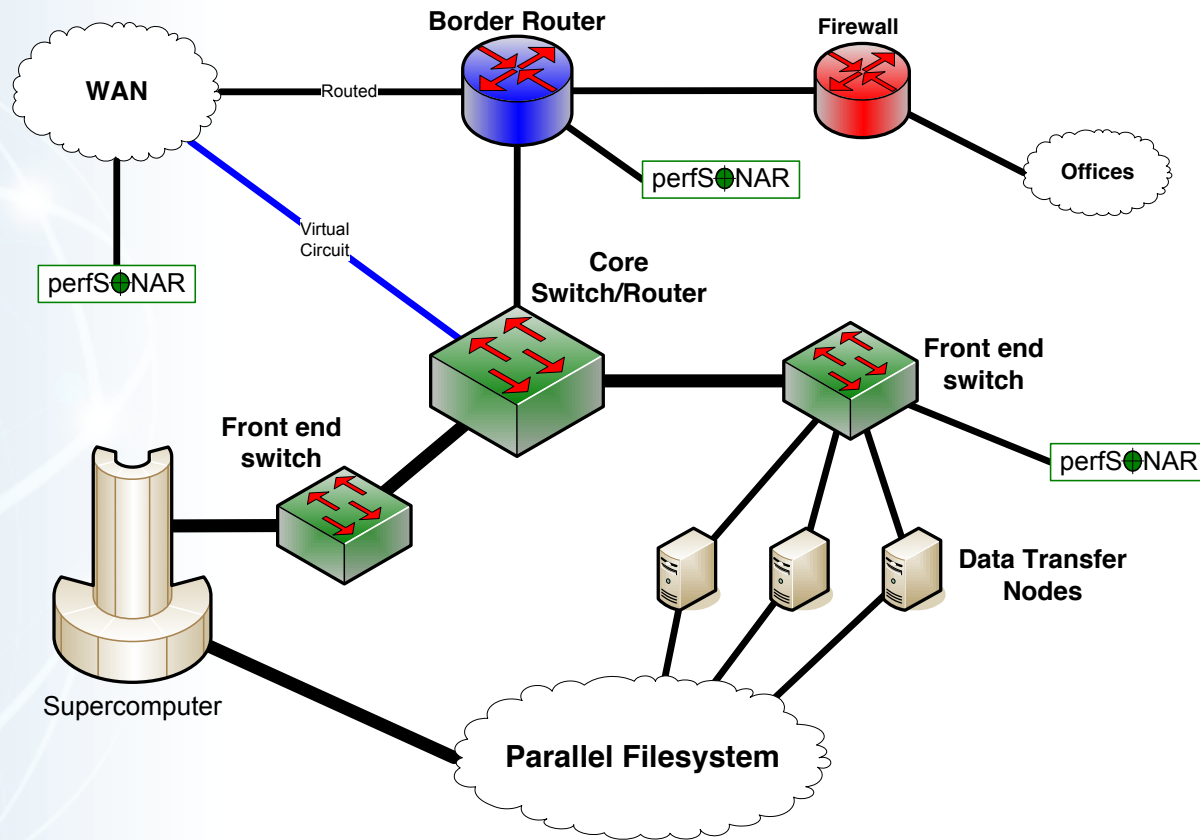
“Science DMZ” may not look discrete

- Most of the network is in the Science DMZ
- This is as it should be
- Appropriate deployment of tools, configuration, policy control, etc.

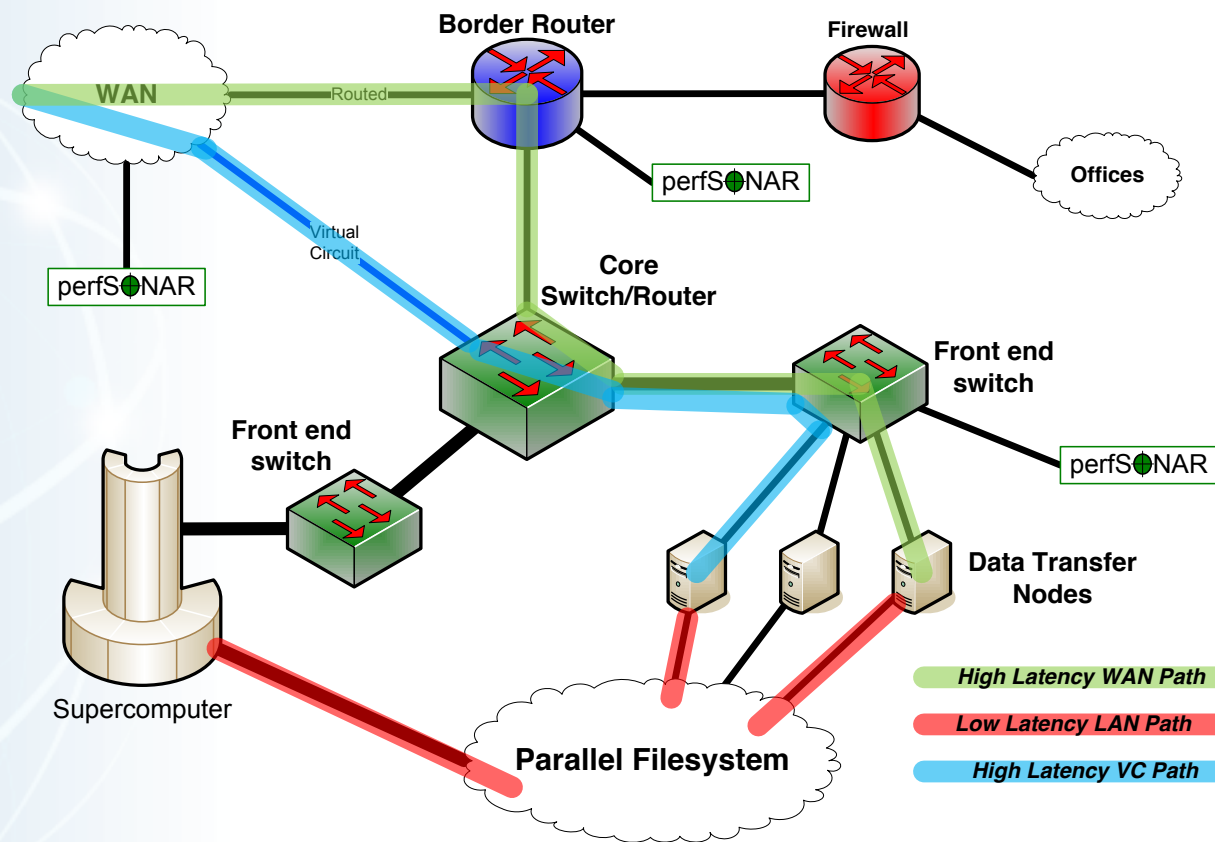
Office networks can look like an afterthought, but they aren't

- Deployed with appropriate security controls
- Office infrastructure need not be sized for science traffic

Supercomputer Center



Supercomputer Center Data Path





Major Data Site Deployment

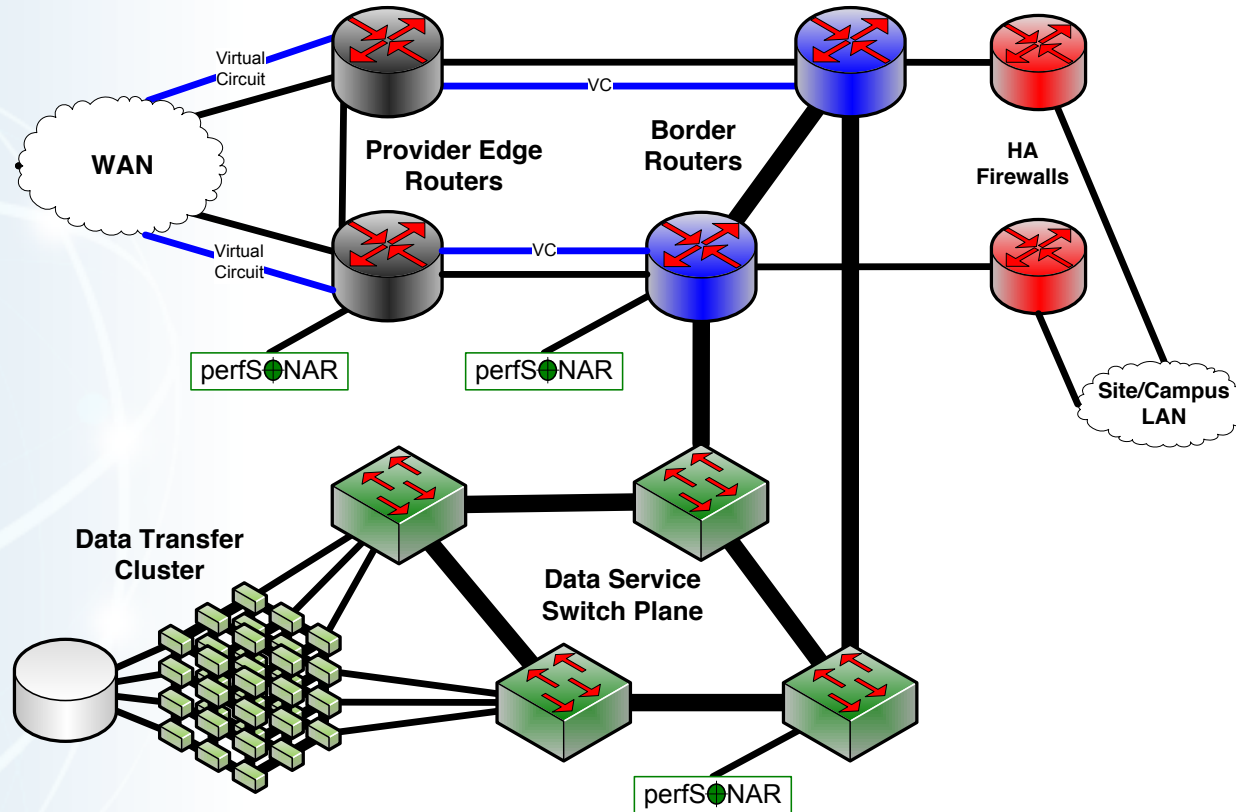
In some cases, large scale data service is the major driver

- Huge volumes of data – ingest, export
- Big infrastructure investment

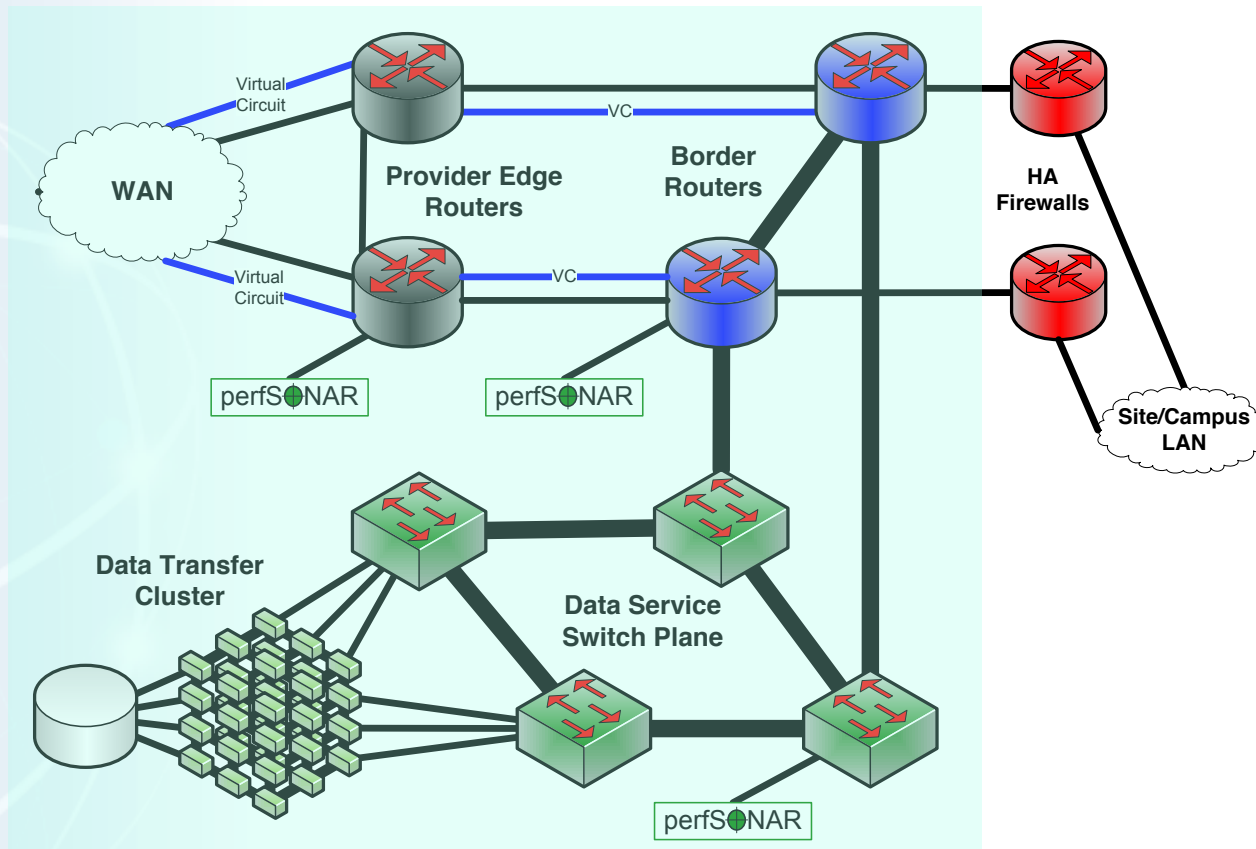
Single-pipe deployments don't work

- Everything is parallel
 - Networks (Nx10G LAGs, soon to be Nx100G)
 - Hosts – data transfer clusters, sets of DTNs
 - WAN connections – multiple entry, redundant equipment
- Any choke point (e.g. firewall) causes problems

Data Site – Architecture



Data Site – Data Path





Common Threads

Two 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

Science DMZ – Test and Measurement



Dedicated
Systems for
Data Transfer

Data Transfer Node

- High performance
- Configured for data transfer
- Proper tools

Network
Architecture

Science DMZ

- Dedicated location for DTN
- Proper security
- Easy to deploy - no need to redesign the whole network
- Additional info:
<http://fasterdata.es.net/>

Performance
Testing &
Measurement

perfSONAR

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



Performance Monitoring

The wide area network, the Science DMZ, and all its systems may be functioning perfectly

Eventually something is going to break

- Networks and systems are complex
- Bugs, mistakes, ...
- Sometimes things just break – this is why we buy support contracts

Must be able to find and fix problems when they occur



Soft Network Failures – Hidden Problems

“Soft failures” result in degraded capability

- Connectivity exists
- Performance impacted
- Typically something in the path is functioning, but not well

Hard failures are easy to detect

- Link down, server down, software crash
- Traditional network/system monitoring tools designed to quickly find hard failures

Soft failures are hard to detect with traditional methods

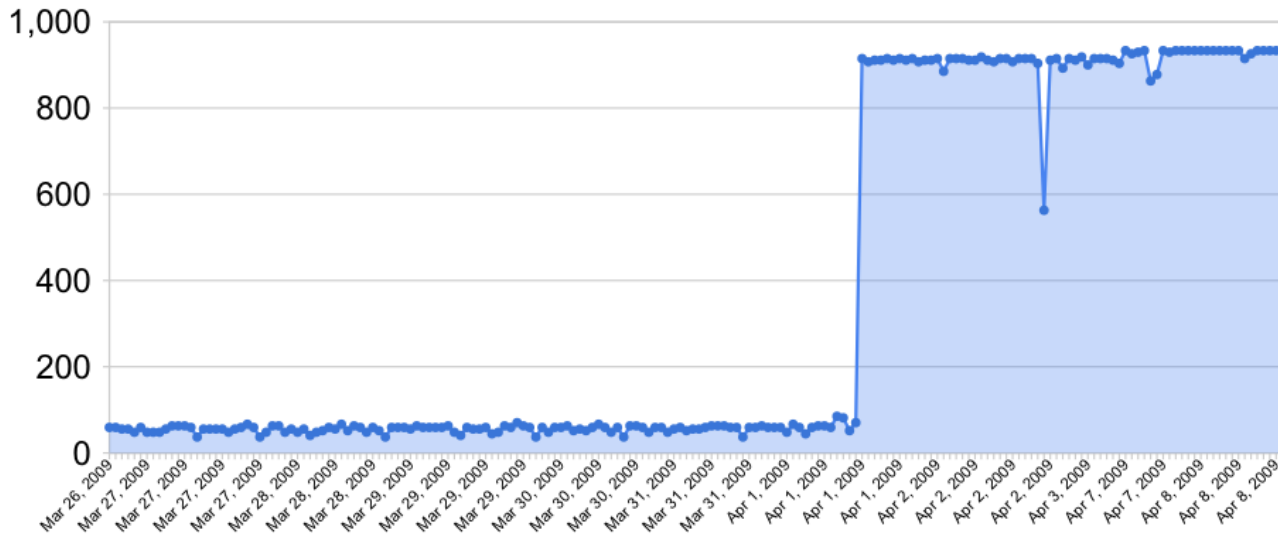
- No link down event
- Sometimes no error counters

Independent testing is the only way to reliably find soft failures

Sample Results: Finding/Fixing soft failures

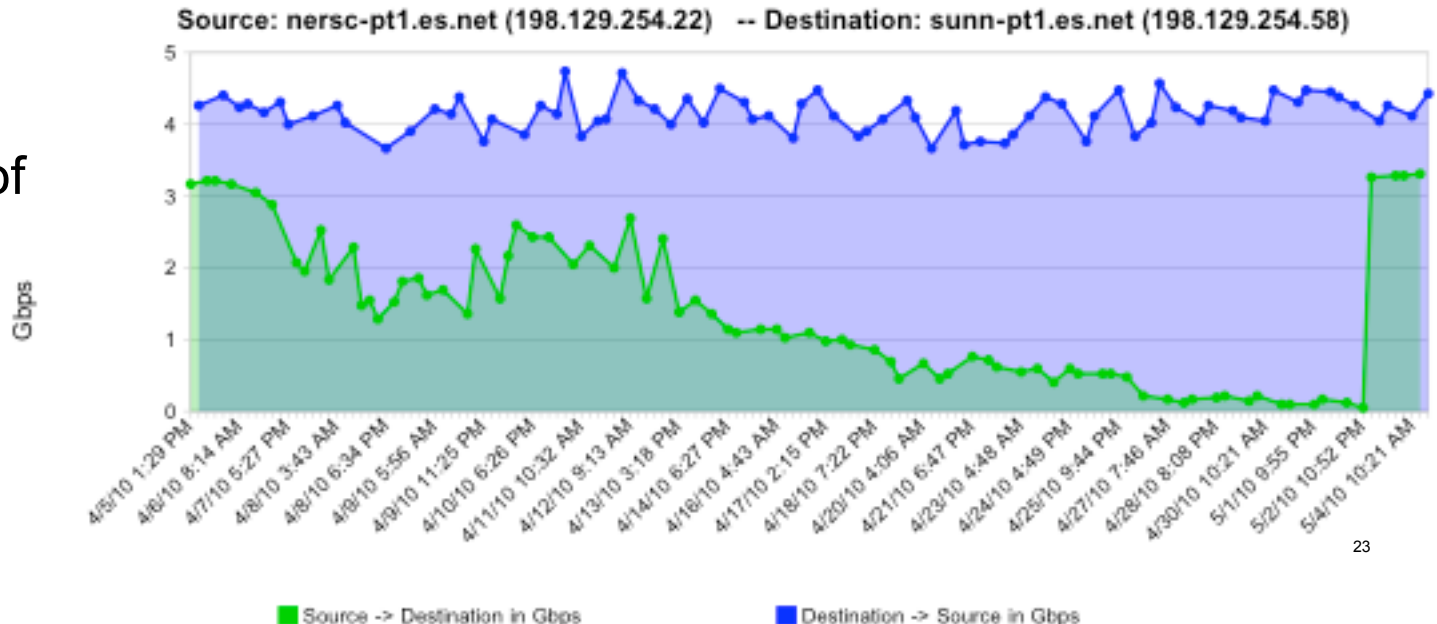


Bandwidth (Mbits/sec)



Rebooted router
that was process
switching after
route table overload

Gradual failure of
optical line card



Testing Infrastructure – perfSONAR



perfSONAR is:

- A widely-deployed test and measurement infrastructure
 - ESnet, Internet2, US regional networks, international networks
 - Laboratories, supercomputer centers, universities
- A suite of test and measurement tools
- A collaboration that builds and maintains the toolkit

By installing perfSONAR, a site can leverage over 900 test servers deployed around the world

perfSONAR is ideal for finding soft failures

- Alert to existence of problems
- Fault isolation
- Verification of correct operation



Science DMZ – Data Transfer Systems



Dedicated Systems for Data Transfer

Data Transfer Node

- High performance
- Configured for data transfer
- Proper tools

Network Architecture

Science DMZ

- Dedicated location for DTN
- Proper security
- Easy to deploy - no need to redesign the whole network
- Additional info:
<http://fasterdata.es.net/>

Performance Testing & Measurement

perfSONAR

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



Dedicated Systems – The Data Transfer Node

The DTN is dedicated to data transfer

Set up specifically for high-performance data movement

- System internals (BIOS, firmware, interrupts, etc.)
- Network stack
- Storage (global filesystem, Fibrechannel, local RAID, etc.)
- Tools

Limitation of scope and function is actually powerful

- No conflicts with configuration for other tasks
- Small application set makes cybersecurity easier



Data Transfer Tools For DTNs

Parallelism is key

- It is much easier to achieve a given performance level with four parallel connections than one connection
- Several tools offer parallel transfers, including GridFTP/Globus Online

Latency interaction is critical

- Wide area data transfers have much higher latency than LAN transfers
- Many tools and protocols assume a LAN
- Examples: SCP/SFTP, HPSS mover protocol

Workflow integration is important

Science DMZ Security



Goal – disentangle security policy and enforcement for science flows from security for business systems

Rationale

- Science data traffic is simple from a security perspective
- Narrow application set on Science DMZ
 - Data transfer, data streaming packages
 - No printers, document readers, web browsers, building control systems, financial databases, staff desktops, etc.
- Security controls that are typically implemented to protect business resources often cause performance problems

Separation allows each to be optimized



Performance Is A Core Requirement

Core information security principles

- Confidentiality, Integrity, Availability (CIA)
- Often, CIA and risk mitigation result in compromised performance

In data-intensive science, performance is an additional core mission requirement

- CIA principles are important, but ***if the performance isn't there the science mission fails***
- Not about “how much” security you have, but how the security is implemented
- Must be able to appropriately secure systems in a way that does not compromise performance or limit advanced services



Placement Outside the Firewall

The Science DMZ resources are placed outside the enterprise firewall for performance reasons

- The meaning of this is specific – ***Science DMZ traffic does not traverse the firewall data plane***
- This has nothing to do with whether packet filtering is part of the security enforcement toolkit

Lots of heartburn over this, especially from the perspective of a conventional firewall manager

- Lots of organizational policy directives mandating firewalls
- Firewalls are designed to protect converged enterprise networks
- Why would you put critical assets outside the firewall???

The answer is that firewalls are typically a poor fit for high-performance science applications



Security Without Firewalls

Data intensive science traffic interacts poorly with firewalls

Does this mean we ignore security? **NO!**

- We **must** protect our systems
- Just do security without preventing science

Key point – security policies and mechanisms that protect the Science DMZ should be implemented so that they do not compromise performance

Example – firewall rules for science traffic use address/port

- Implement that filtering on a high-performance router
- Science wins – increased performance
- Business network wins – no need to size the firewall for science data deluge

Futures



The Science DMZ design pattern is highly adaptable to new technologies

- Software Defined Networking (SDN)
- Non-IP protocols (RDMA over Ethernet)

Deploying new technologies in a Science DMZ is straightforward

- The basic elements are the same
 - Capable infrastructure designed for the task
 - Test and measurement to verify correct operation
 - Security policy well-matched to the environment, application set strictly limited to reduce risk
- Change footprint is small – often just a single router or switch
- The rest of the infrastructure need not change

Wrapup



The Science DMZ design pattern provides a flexible model for supporting high-performance data transfers and workflows

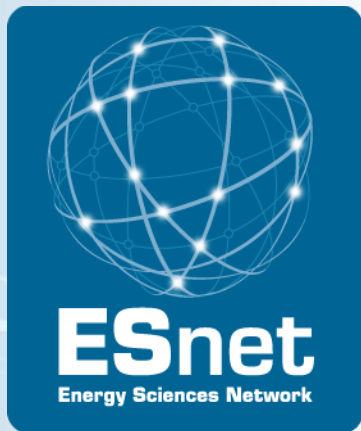
Key elements:

- Accommodation of TCP
 - Sufficient bandwidth to avoid congestion
 - Loss-free IP service
- Location – near the site perimeter if possible
- Test and measurement
- Dedicated systems
- Appropriate security

Support for advanced capabilities (e.g. SDN) is much easier with a Science DMZ

Links

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



Thanks!

Questions?

Eli Dart – dart@es.net

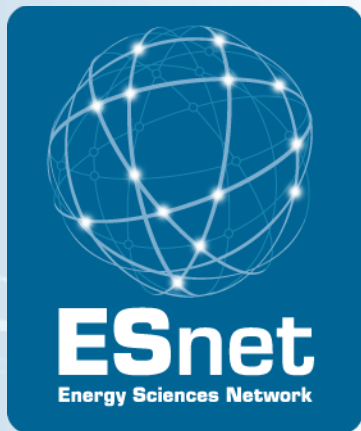
<http://www.es.net/>

<http://fasterdata.es.net/>



U.S. DEPARTMENT OF
ENERGY
Office of Science





Thanks!

Questions?

Eli Dart – dart@es.net

<http://www.es.net/>

<http://fasterdata.es.net/>



U.S. DEPARTMENT OF
ENERGY
Office of Science

