1986–2016
ESnet
30 YEARS OF NETWORKING
AT THE SPEED OF SCIENCE

The Science DMZ Design Pattern

Eli Dart
Network Engineer
ESnet Science Engagement
Lawrence Berkeley National Laboratory

NASA
Mountain View, CA
October 1, 2016

U.S. DEPARTMENT OF ENERGY
Office of Science

BERKELEY LAB

Overview

- Science DMZ Motivation and Introduction
- Science DMZ
  - Architecture
  - Network Monitoring For Performance
  - Data Transfer Nodes & Applications
  - Science DMZ Security
- Larger Context, Platform
  - Science Engagement
  - Pacific Research Platform
  - Data Portal Discussion
  - Petascale DTN Project

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

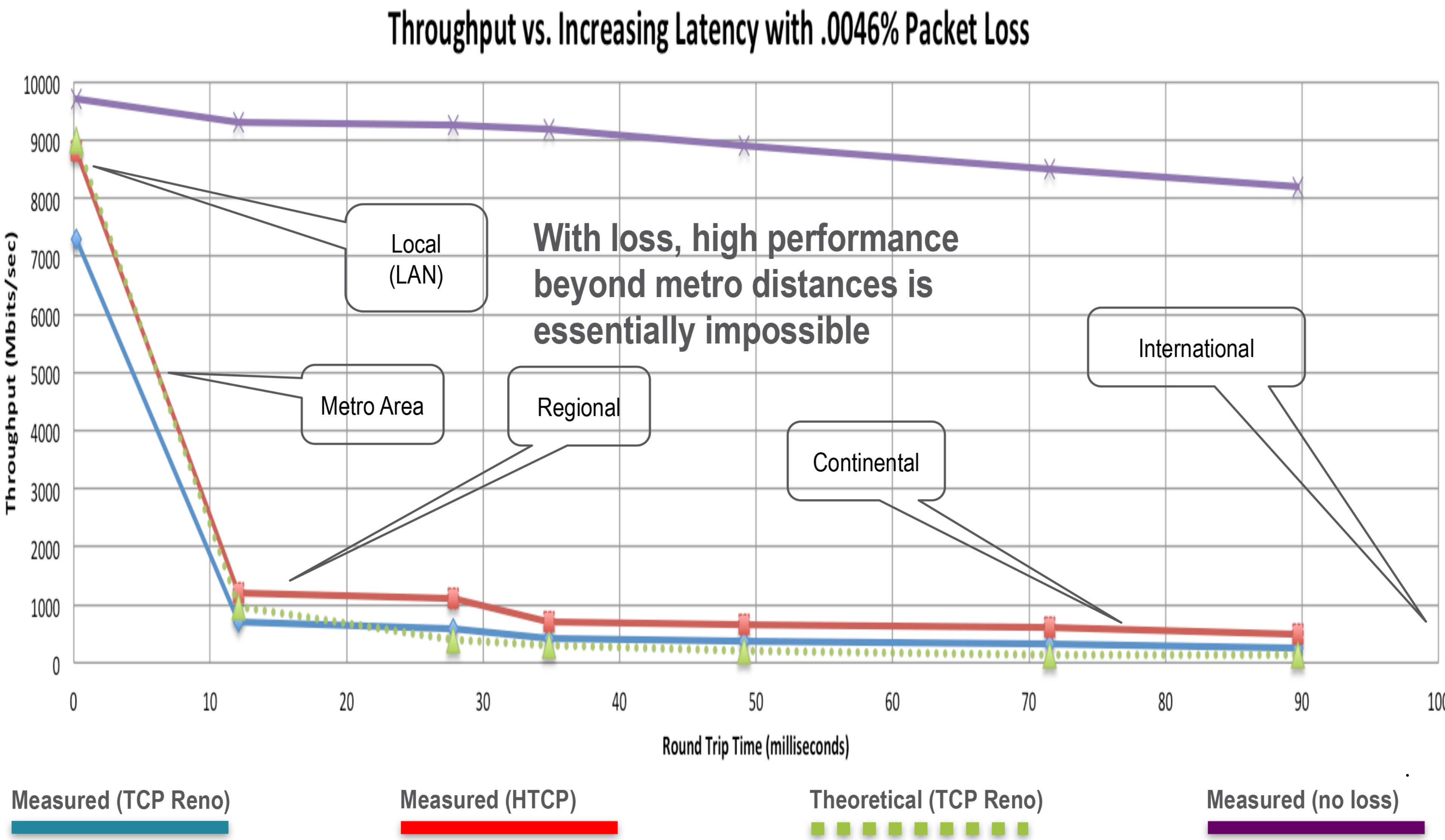
The Central Role of the Network

- The very structure of modern science assumes science networks exist: high performance, feature rich, global scope
- 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
- What is important? Ordered list:
  1. Correctness
  2. Consistency
  3. Performance

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 (more than 95% of ESnet traffic is TCP)

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



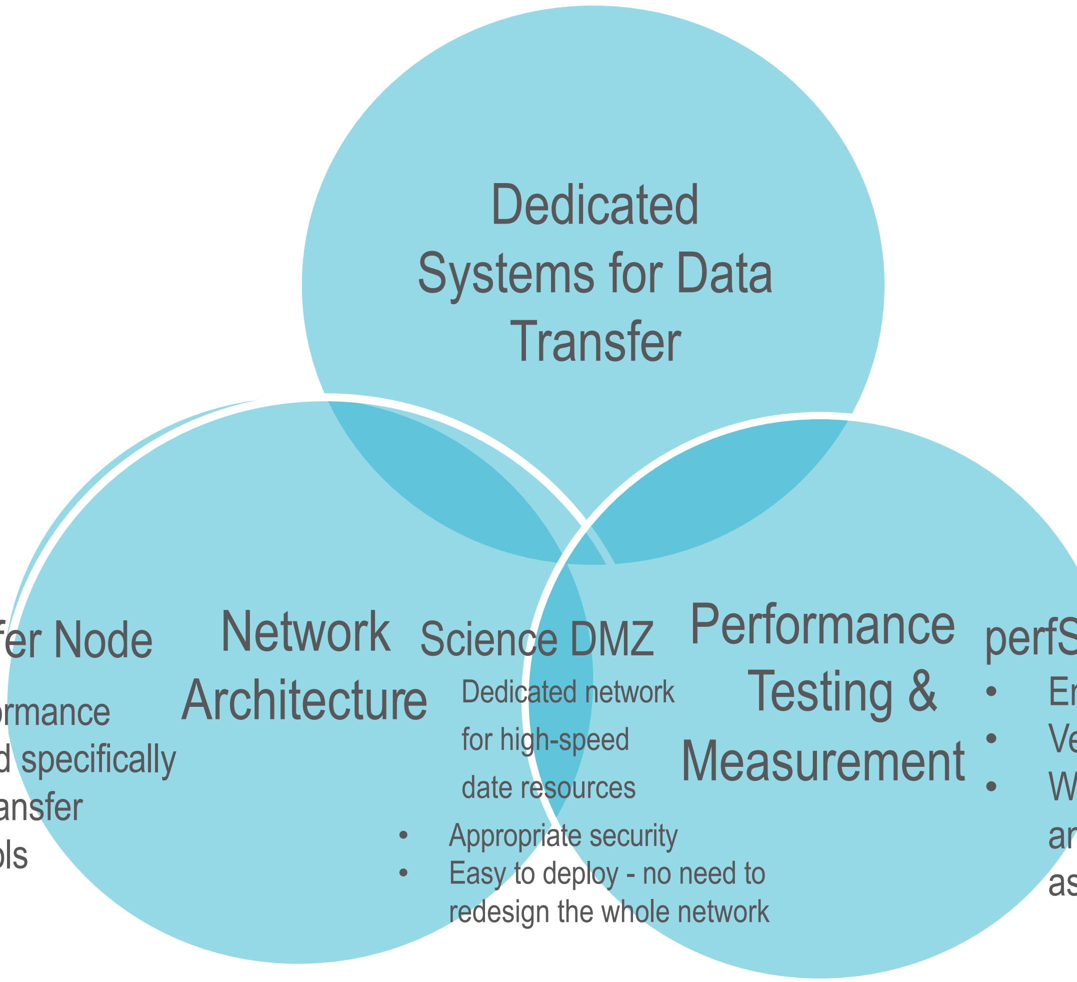
Working With TCP In Practice

- 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

Putting A Solution Together

- Effective support for TCP-based data transfer
  - Design for correct, consistent, high-performance operation
  - Design for ease of troubleshooting
- Easy adoption is critical
  - Large laboratories and universities have extensive IT deployments
  - Drastic change is prohibitively difficult
- Cybersecurity – defensible without compromising performance
- Borrow ideas from traditional network security
  - Traditional DMZ
    - Separate enclave at network perimeter (“Demilitarized Zone”)
    - Specific location for external-facing services
    - Clean separation from internal network
  - Do the same thing for science – ***Science DMZ***

The Science DMZ Design Pattern



Data Transfer Node

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

Science DMZ

Dedicated network for high-speed date resources

- Appropriate security
- Easy to deploy - no need to redesign the whole network

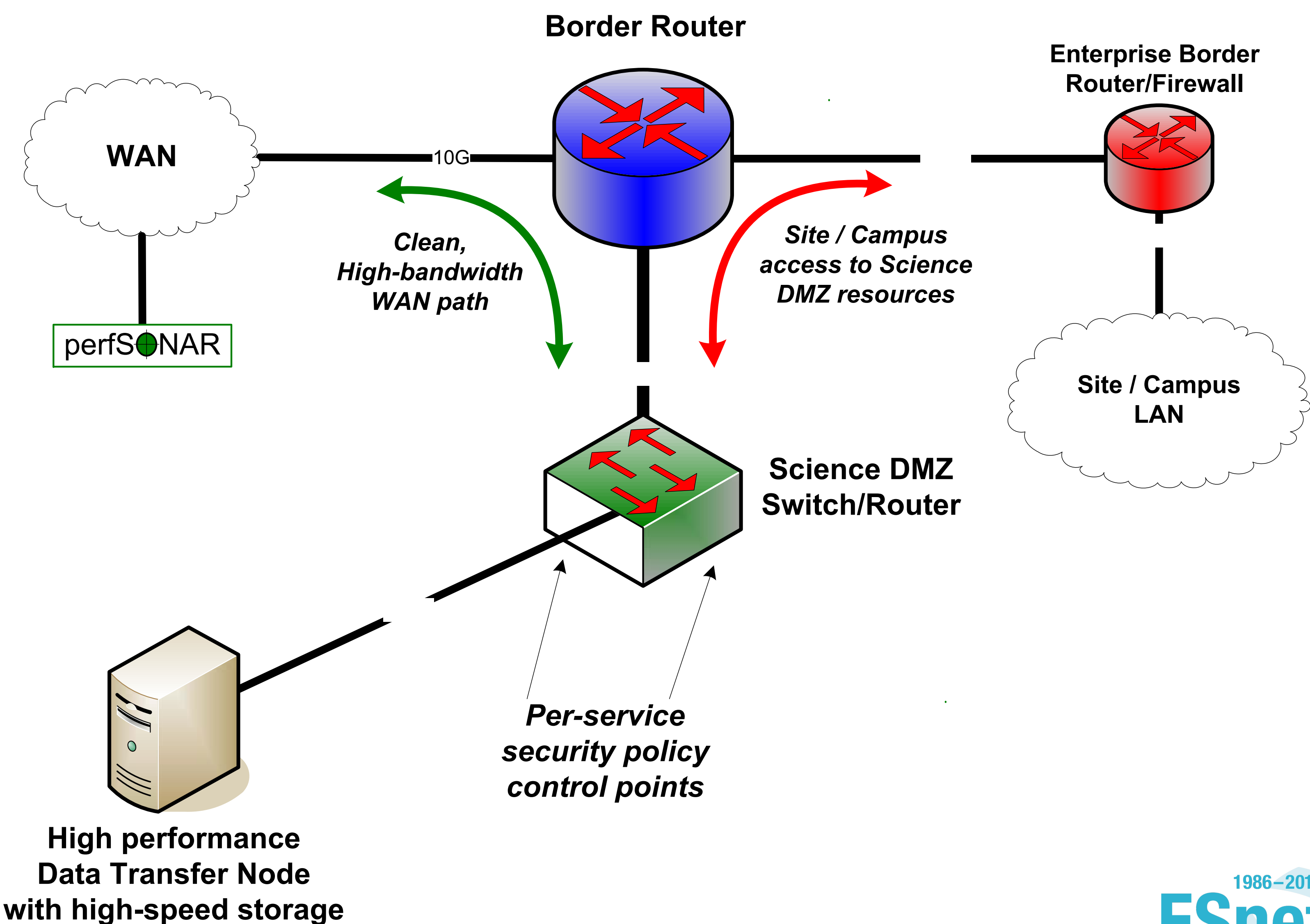
perfSONAR

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

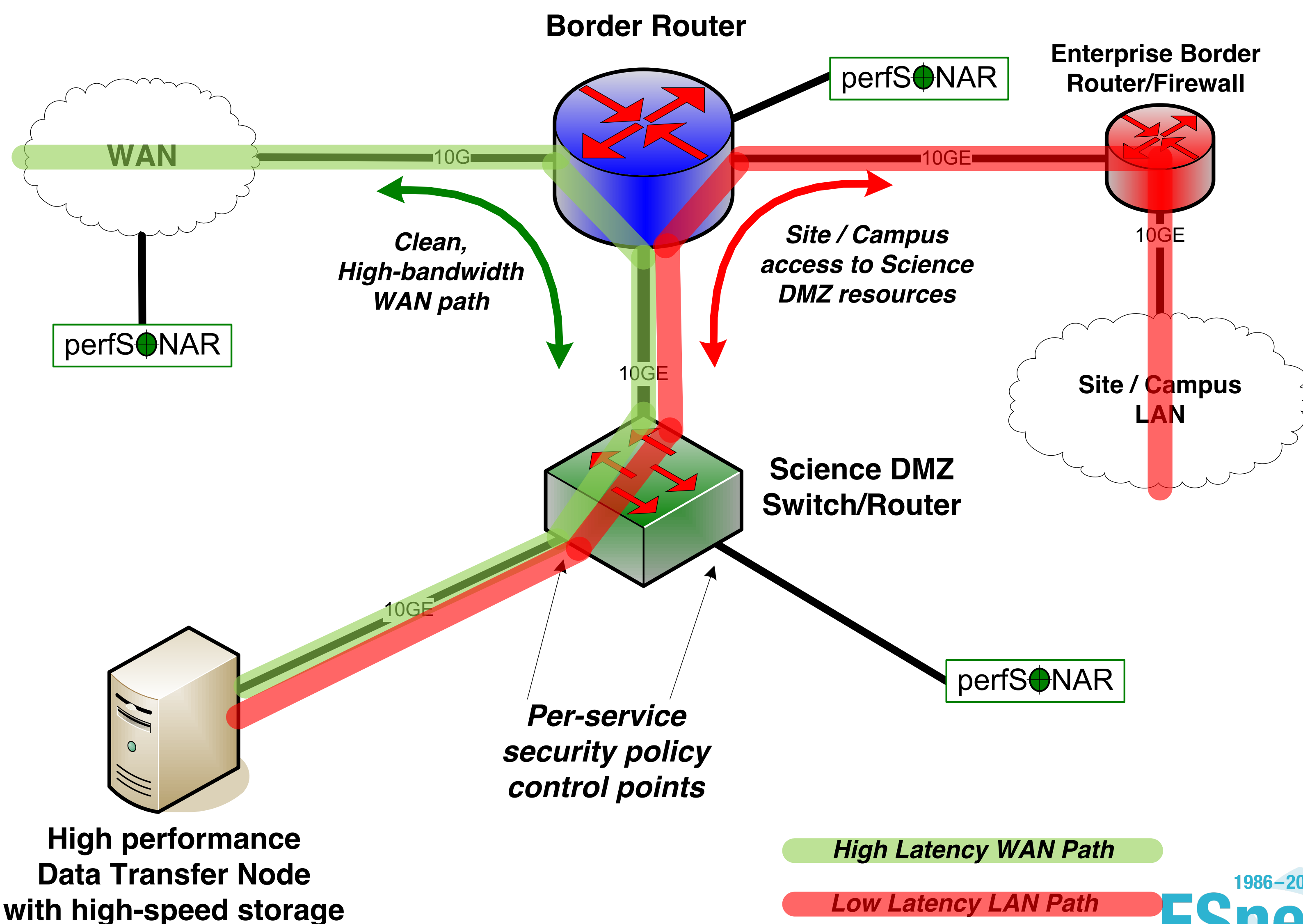
Abstract or Prototype Deployment

- Add-on to existing network infrastructure
  - All that is required is a port on the border router
  - Small footprint, pre-production commitment
- Easy to experiment with components and technologies
  - DTN prototyping
  - perfSONAR testing
- Limited scope makes security policy exceptions easy
  - Only allow traffic from partners
  - Add-on to production infrastructure – lower risk

Science DMZ Design Pattern (Abstract)



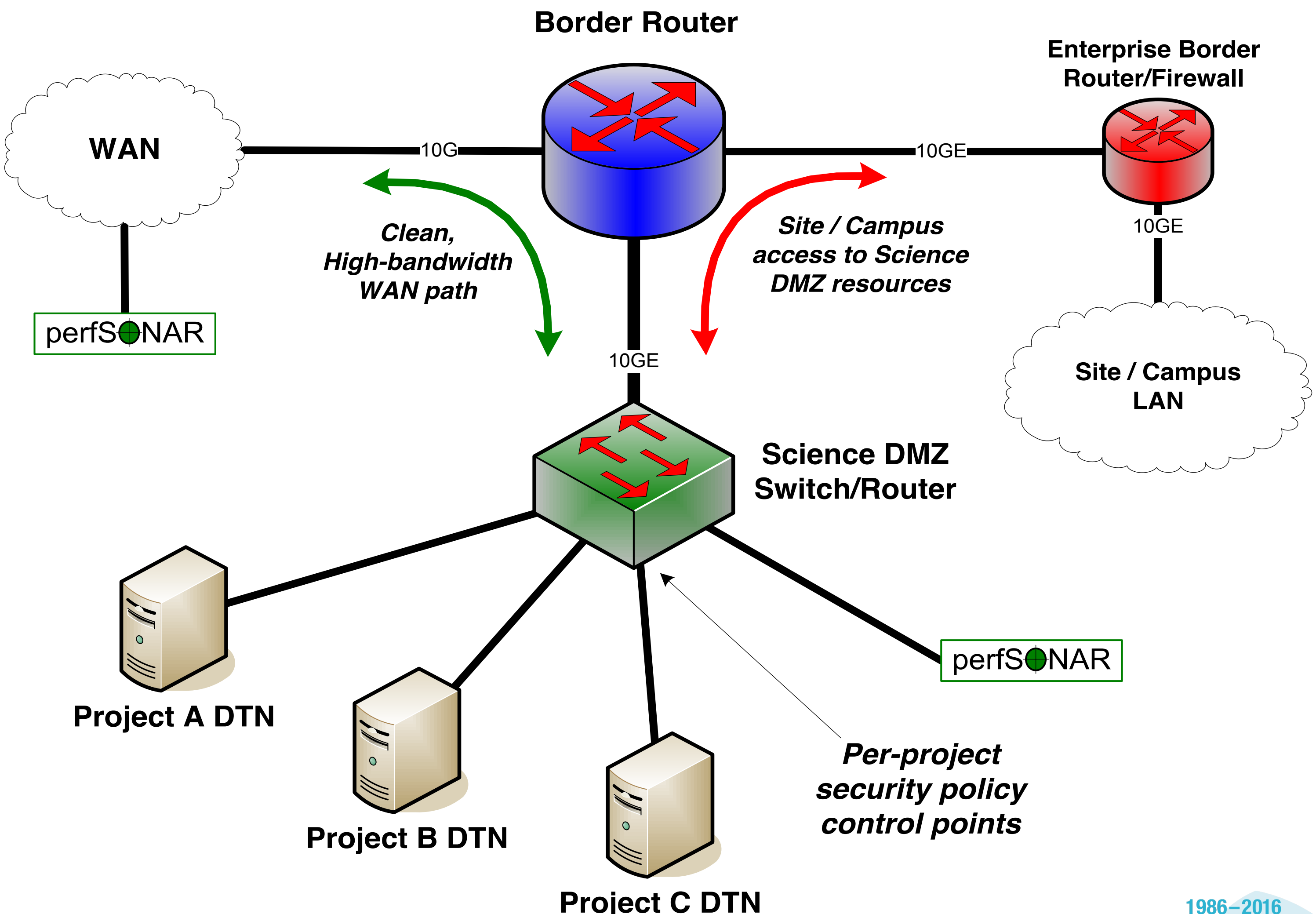
Local And Wide Area Data Flows



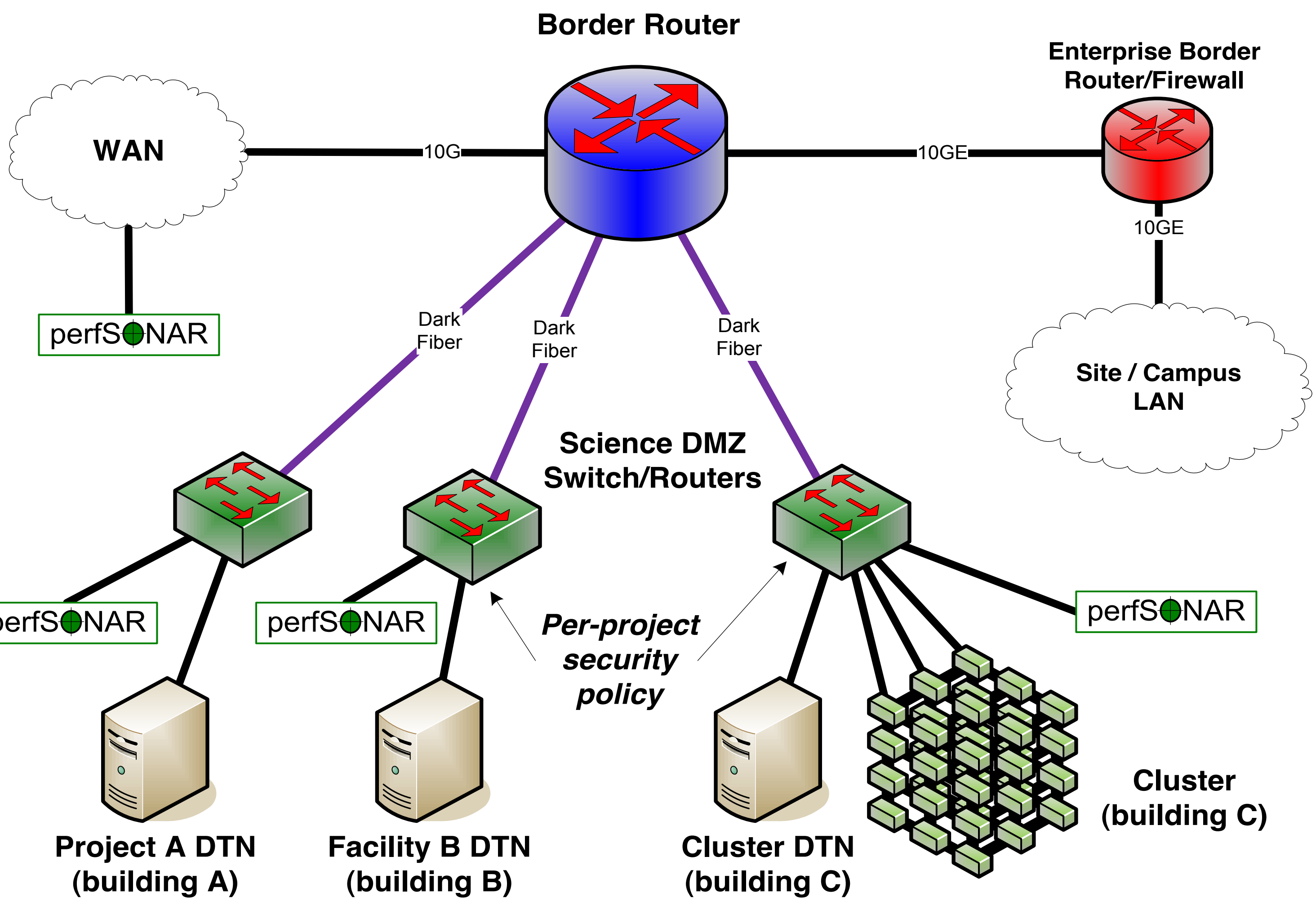
Support For Multiple Projects

- Science DMZ architecture allows multiple projects to put DTNs in place
  – Modular architecture
  – Centralized location for data servers
- This may or may not work well depending on institutional policies
  – Sometimes individual groups deploy their own servers, and centralization is hard
  – Sometimes centralization is a strategic goal
- On balance, this can provide a cost savings – it depends
  – Central support for data servers vs. carrying data flows
  – How far do the data flows have to go?
- Dark fiber asses can be a huge win

Multiple Projects



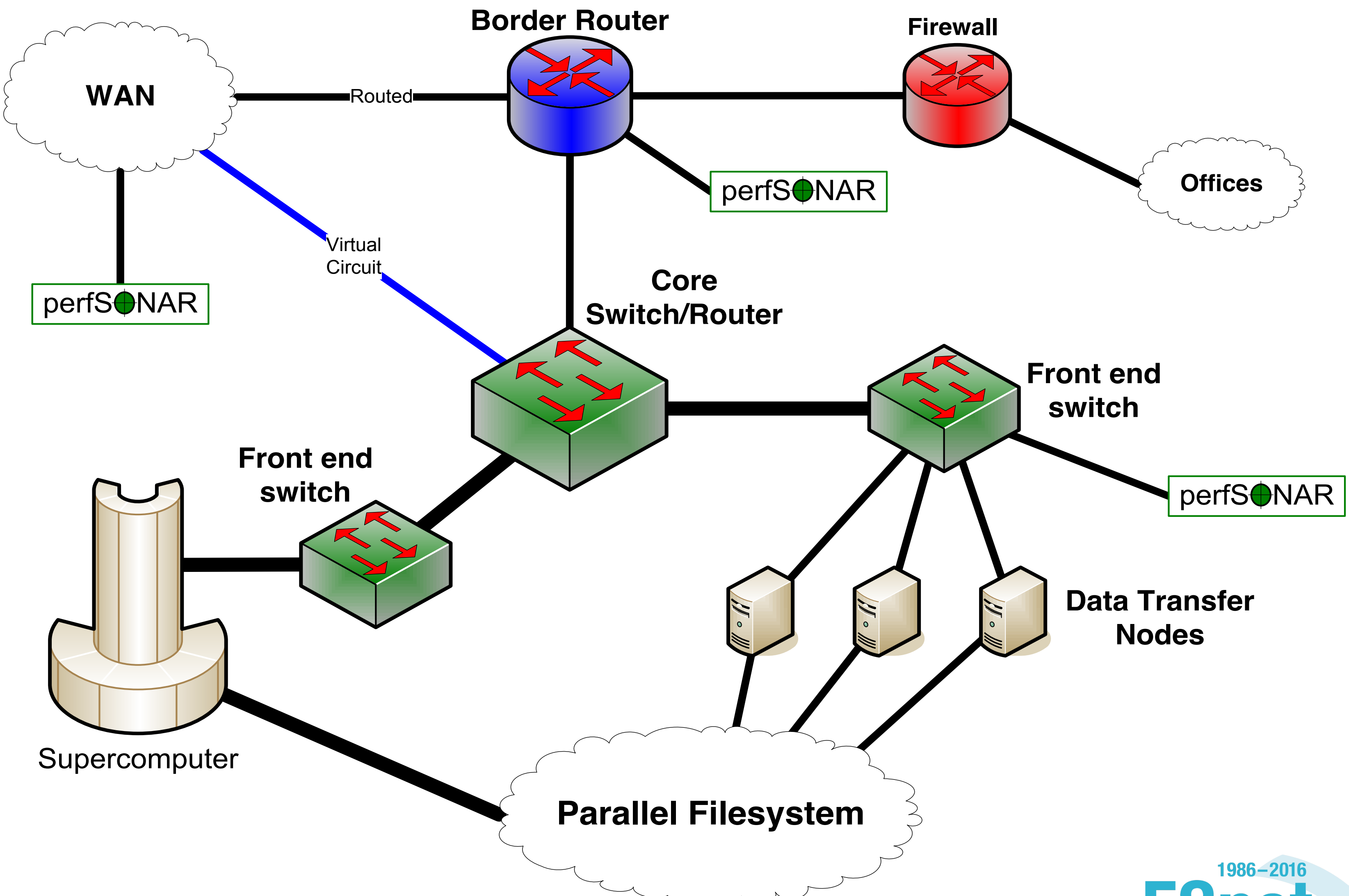
Multiple Science DMZs – Dark Fiber



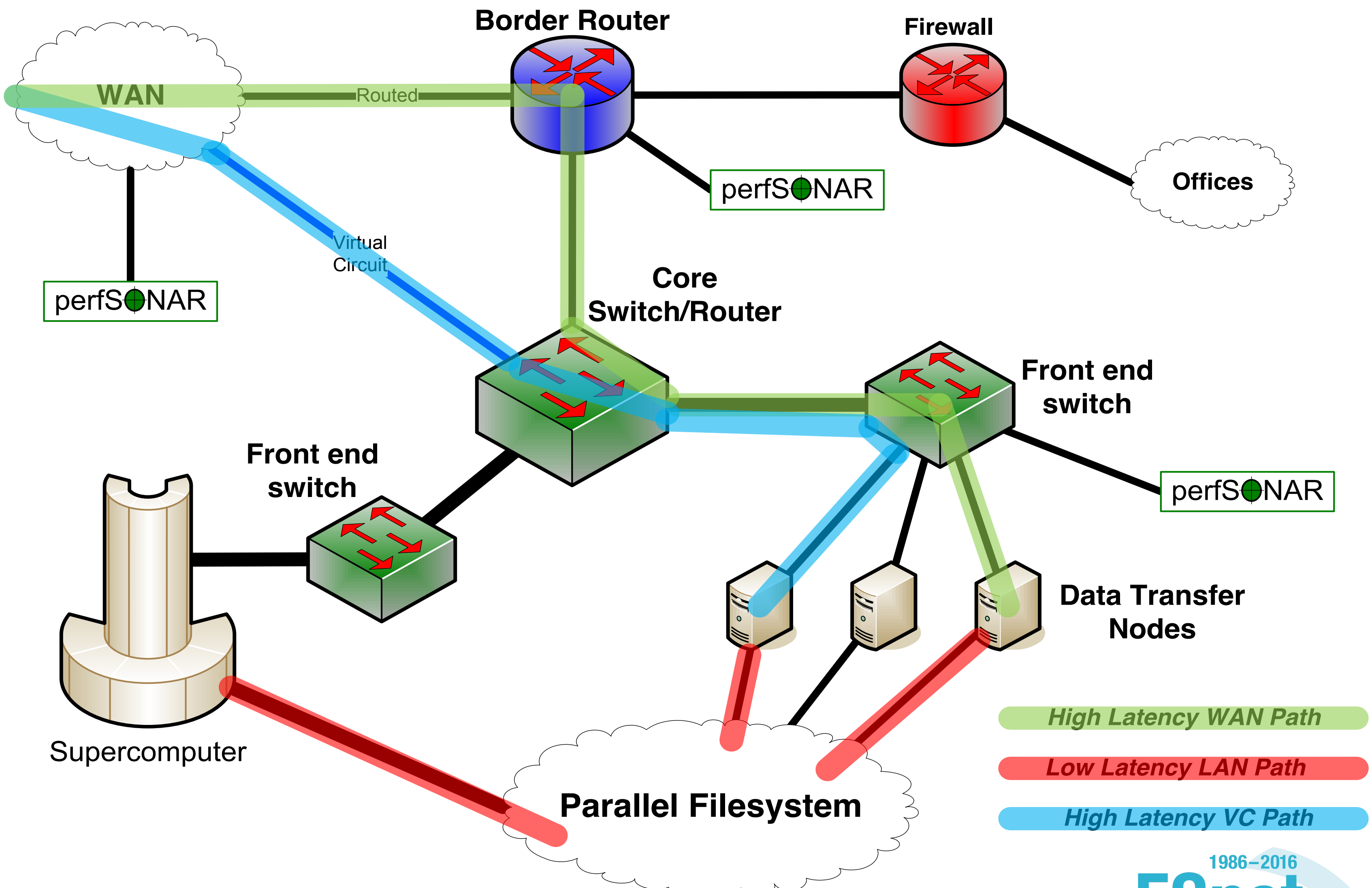
Supercomputer Center Deployment

- 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 like a discrete entity here
  - By the time you get through interconnecting all the resources, you end up with most of the network in the Science DMZ
  - This is as it should be – the point is 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



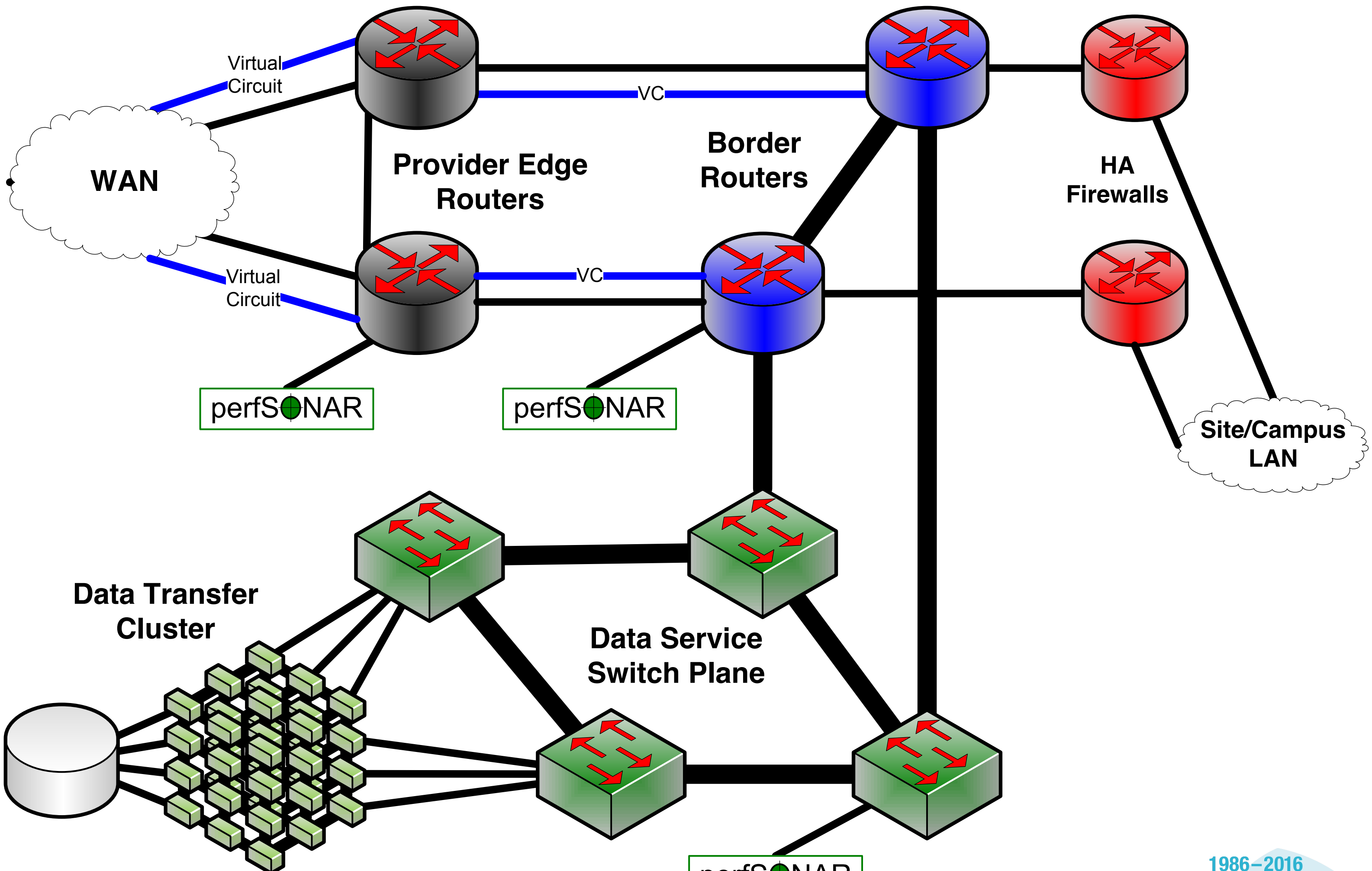
Supercomputer Center Data Path



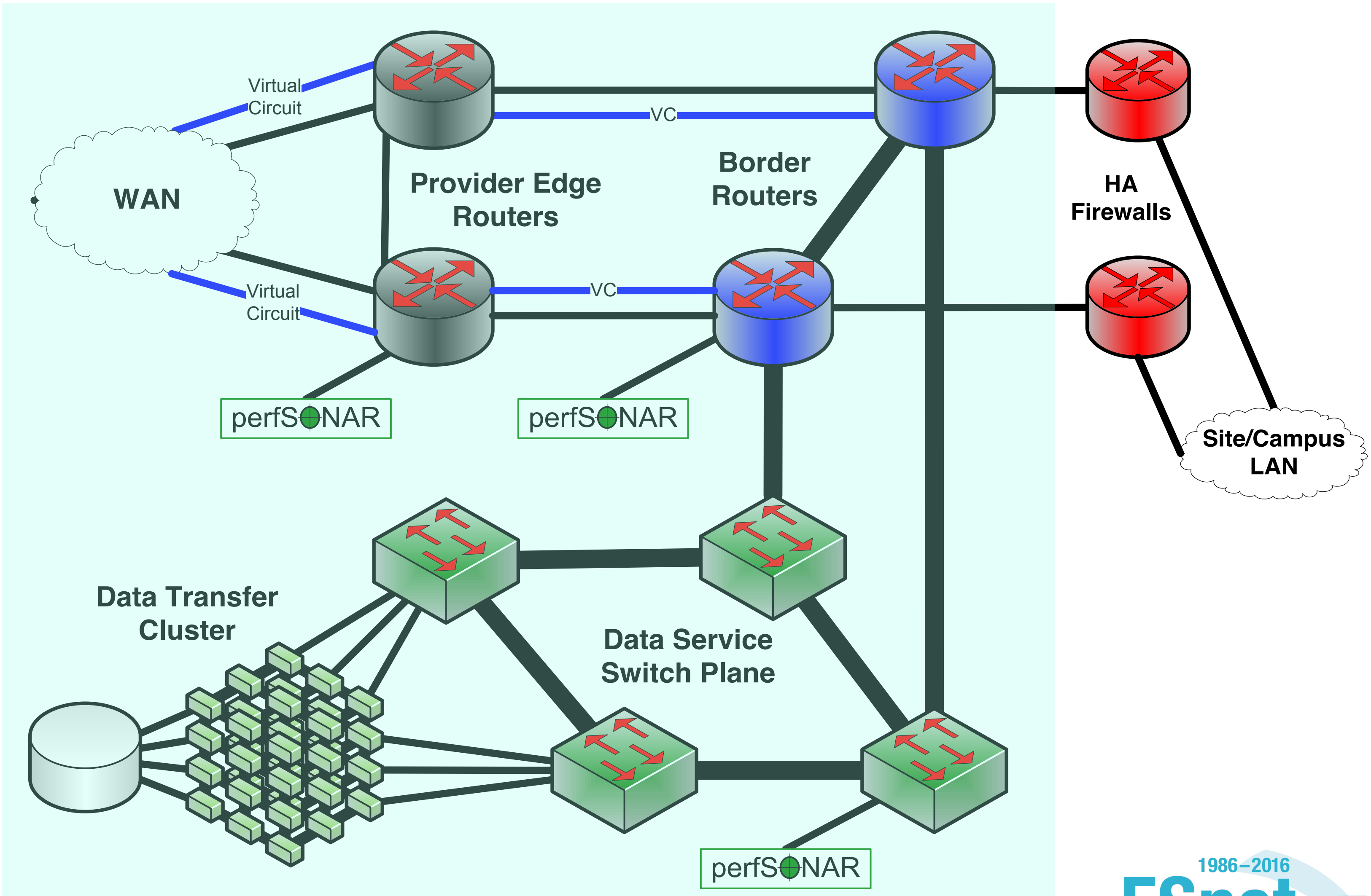
Major Data Site Deployment

- In some cases, large scale data service is the major driver
  - Huge volumes of data (Petabytes or more) – ingest, export
  - Large number of external hosts accessing/submitting data
- Single-pipe deployments don't work
  - Everything is parallel
    - Networks (Nx10G LAGs, soon to be Nx100G)
    - Hosts – data transfer clusters, no individual DTNs
    - WAN connections – multiple entry, redundant equipment
  - Choke points (e.g. firewalls) just cause 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

1986–2016 ESnet 30 YEARS OF NETWORKING AT THE SPEED OF SCIENCE

Performance Monitoring

- Everything may function perfectly when it is deployed
- 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
- Must be able to find problems in other networks (your network may be fine, but someone else's problem can impact your users)

- TCP was intentionally designed to hide all transmission errors from the user:
  - "As long as the TCPs continue to function properly and the internet system does not become completely partitioned, no transmission errors will affect the users." (From RFC793, 1981)

Testing Infrastructure – perfSONAR

- perfSONAR is:
  - A widely-deployed test and measurement infrastructure
    - ESnet, Internet2, US regional networks, international networks
    - Laboratories, supercomputer centers, universities
    - Individual Linux hosts at key network locations (POPs, Science DMZs, etc.)
  - A suite of test and measurement tools
  - A collaboration that builds and maintains the toolkit
- By installing perfSONAR, a site can leverage over 2000 test servers deployed around the world
- perfSONAR is ideal for finding soft failures
  - Alert to existence of problems
  - Fault isolation
  - Verification of correct operation

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.)
  - High performance tools
  - No extraneous software
- ***Limitation of scope and function is powerful***
  - No conflicts with configuration for other tasks
  - Small application set makes cybersecurity easier
    - Limitation of application set is often a core security policy component

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

1986–2016 ESnet 30 YEARS OF NETWORKING AT THE SPEED OF SCIENCE

Science DMZ As Security Architecture

- Allows for better segmentation of risks, more granular application of controls to those segmented risks.
  - Limit risk profile for high-performance data transfer applications
  - Apply specific controls to data transfer hosts
  - Avoid including unnecessary risks, unnecessary controls
- Remove degrees of freedom – focus only on what is necessary
  - Easier to secure
  - Easier to achieve performance
  - Easier to troubleshoot

Performance Is A Core Requirement

- Core information security principles
  - Confidentiality, Integrity, Availability (CIA)
  - Often, CIA and risk mitigation result in poor performance
- In data-intensive science, performance is an additional core mission requirement: CIA → PICA
  - CIA principles are important, but ***if performance is compromised the science mission fails***
  - Not about “how much” security you have, but how the security is implemented
  - Need a way to appropriately secure systems without performance compromises

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***
  - Packet filtering is fine – just don't do it with a firewall
- 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

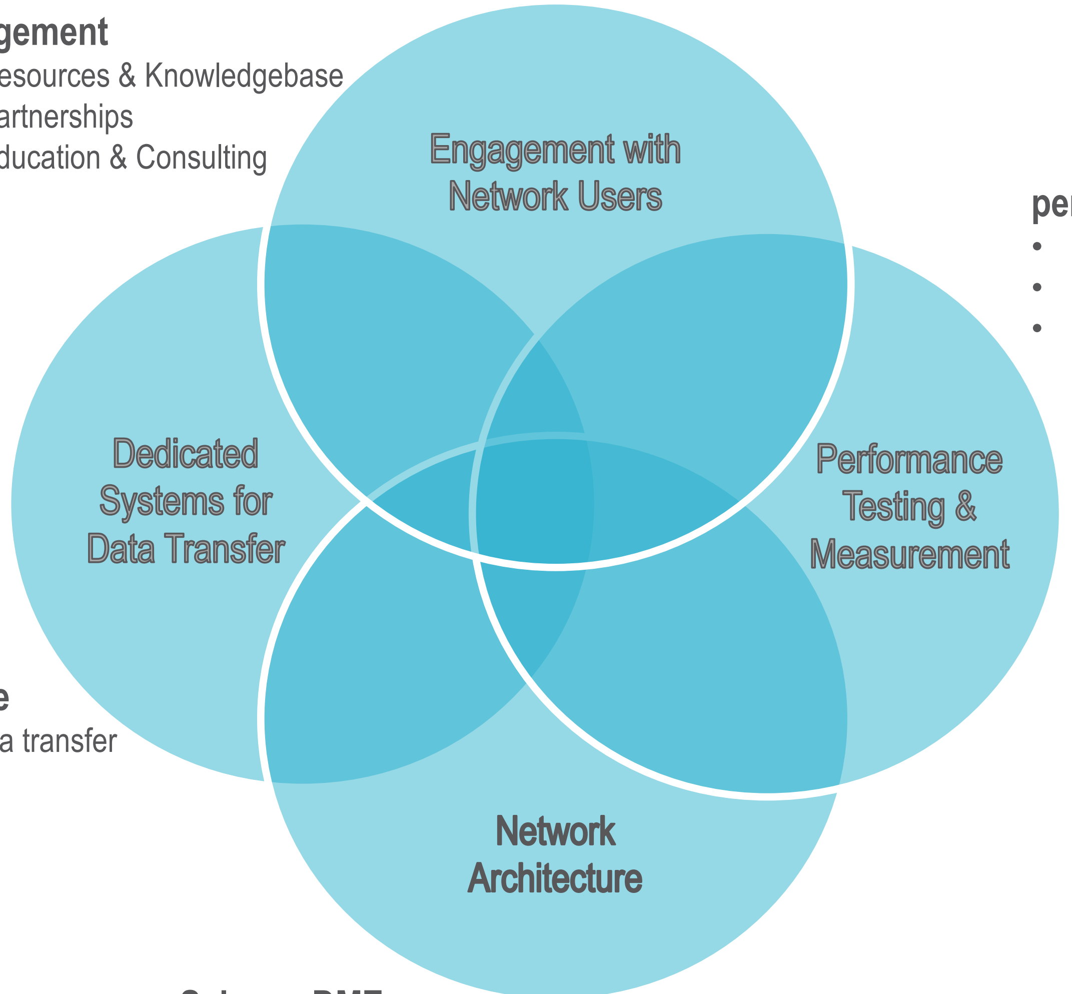
Security Without Firewalls

- Data intensive science traffic interacts poorly with firewalls
- Does this mean we ignore security? ***NO!***
  - We **must** protect our systems
  - We just need to find a way to do security that does not prevent us from getting the science done
- ***Key point – security policies and mechanisms that protect the Science DMZ should be implemented so that they do not compromise performance***
- Traffic permitted by policy should not experience performance impact as a result of the application of policy

The Data Transfer Superfecta: Science DMZ Model

Engagement
- Resources & Knowledgebase
- Partnerships
- Education & Consulting

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



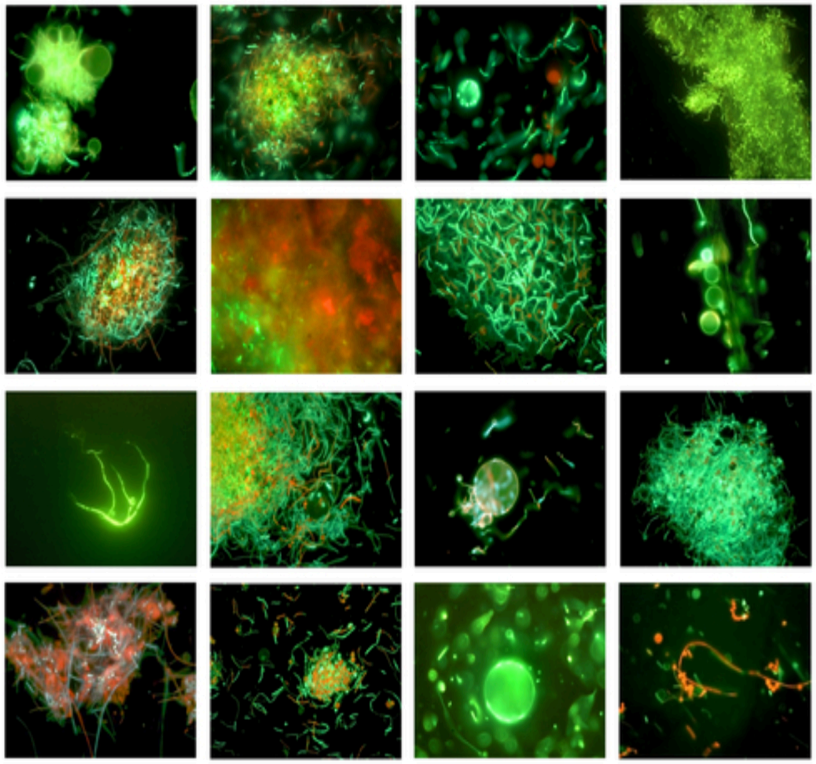
Data Transfer Node
- Configured for data transfer
- High performance
- Proper tools

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

1986–2016 ESnet 30 YEARS OF NETWORKING AT THE SPEED OF SCIENCE

Context Setting

- DOE, NSF, and other agencies are investing billions of dollars in state-of-the-art cyberinfrastructure to support data-intensive science.
- Many researchers do not understand the value of these services and have difficulty using them.
- A proactive effort is needed to drive adoption of advanced services and accelerate science output: **Science Engagement**



1986–2016 ESnet 30 YEARS OF NETWORKING AT THE SPEED OF SCIENCE

ESnet Science Engagement Team Vision

Collaborations at every scale, in every domain, will have the **information and tools** they need to achieve maximum benefit from global networks through the creation of scalable, community-driven strategies and approaches.

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

Science Engagement

- Science Engagement team works in several areas at once
  - Understand key elements which contribute to desired outcomes
    - Requirements analysis – what is needed
    - Also identify choke points, road blocks, missing components
  - Network architecture, performance, best practice
  - Systems engineering, consulting, troubleshooting
  - Collaboration with others
  - Workshops and webinars
- Important bridge between cyberinfrastructure and scientists

Science DMZ 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
  - Science Engagement to foster adoption

Overview

- Science DMZ Motivation and Introduction
- Science DMZ
  - Architecture
  - Network Monitoring For Performance
  - Data Transfer Nodes & Applications
  - Science DMZ Security
  - Science Engagement
- Larger Context, Platform
  - Pacific Research Platform
  - Data Portal Discussion
  - Petascale DTN Project

Context: Science DMZ Adoption

- DOE National Laboratories
  - HPC centers, LHC sites, experimental facilities
  - Both large and small sites
- NSF CC* programs have funded many Science DMZs
  - Significant investments across the US university complex
  - Big shoutout to the NSF – these programs are critically important
- Other US agencies
  - NIH
  - USDA Agricultural Research Service
- International
  - Australia https://www.rdsi.edu.au/dashnet
  - Brazil
  - UK

Strategic Impacts

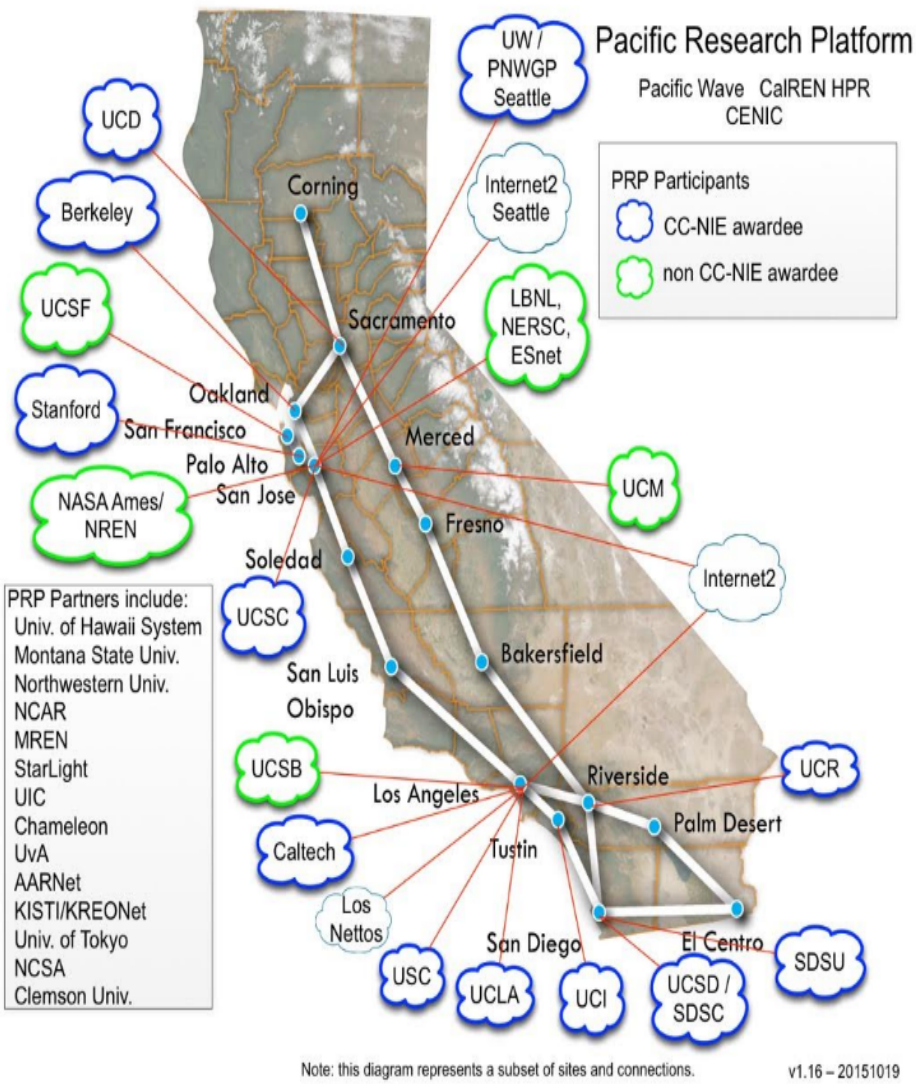
- What does this mean?
  - We are in the midst of a significant cyberinfrastructure upgrade
  - Enterprise networks need not be unduly perturbed ☺
- Significantly enhanced capabilities compared to 3 years ago
  - Terabyte-scale data movement is much easier
  - Petabyte-scale data movement possible outside the LHC experiments
    - ~3.1Gbps = 1PB/month
    - ~14Gbps = 1PB/week
  - Widely-deployed tools are much better (e.g. Globus)
- Metcalfe's Law of Network Utility
  - Value of Science DMZ proportional to the number of DMZs
    - n^2 or $n(\log_n)$ doesn't matter – the effect is real
  - Cyberinfrastructure value increases as we all upgrade

1986–2016 ESnet 30 YEARS OF NETWORKING AT THE SPEED OF SCIENCE

Next Steps – Building On The Science DMZ

- Enhanced cyberinfrastructure substrate now exists
  - Wide area networks (ESnet, GEANT, Internet2, Regionals)
  - Science DMZs connected to those networks
  - DTNs in the Science DMZs
- What does the scientist see?
  - Scientist sees a science application
    - Data transfer
    - Data portal
    - Data analysis
  - Science applications are the user interface to networks and DMZs
- ***The underlying cyberinfrastructure components (networks, Science DMZs, DTNs, etc.) are part of the instrument of discovery***
- Large-scale data-intensive science requires that we build larger structures on top of those components

The Pacific Research Platform Creates a Regional End-to-End Science-Driven "Big Data Freeway System"



Source: John Hess, CENIC

NSF CC*DNI Grant
$5M 10/2015-10/2020

- PI: Larry Smarr, UC San Diego Calit2
- Co-PIs:
  - Camille Crittenden, UC Berkeley CITRIS,
  - Tom DeFanti, UC San Diego Calit2,
  - Philip Papadopoulos, UC San Diego SDSC,
  - Frank Wuerthwein, UC San Diego Physics and SDSC

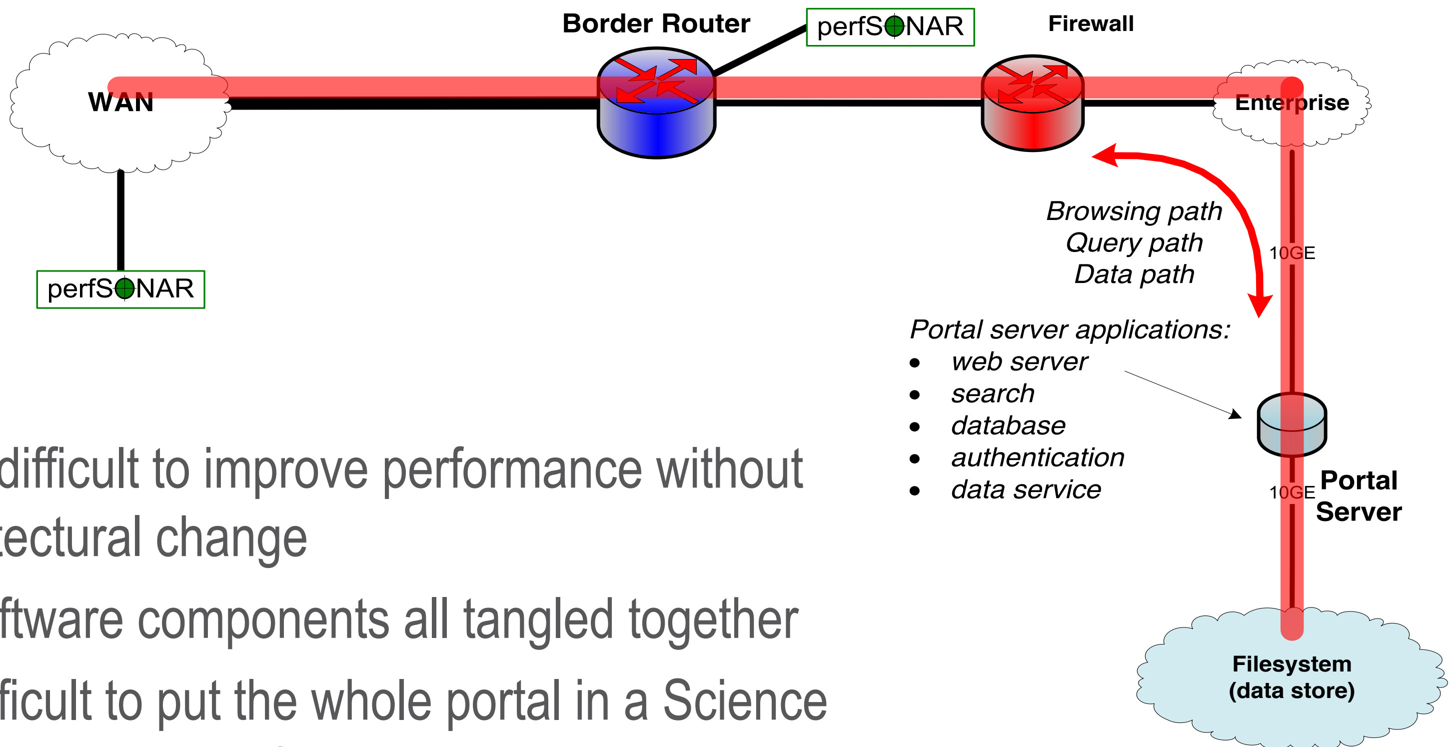
1986–2016 ESnet 30 YEARS OF NETWORKING AT THE SPEED OF SCIENCE

Science Data Portals

- Large repositories of scientific data
  - Climate data
  - Sky surveys (astronomy, cosmology)
  - Many others
  - Data search, browsing, access
- Many scientific data portals were designed 15+ years ago
  - Single-web-server design
  - Data browse/search, data access, user awareness all in a single system
  - All the data goes through the portal server
    - In many cases by design
    - E.g. embargo before publication (enforce access control)

1986–2016 ESnet 30 YEARS OF NETWORKING AT THE SPEED OF SCIENCE

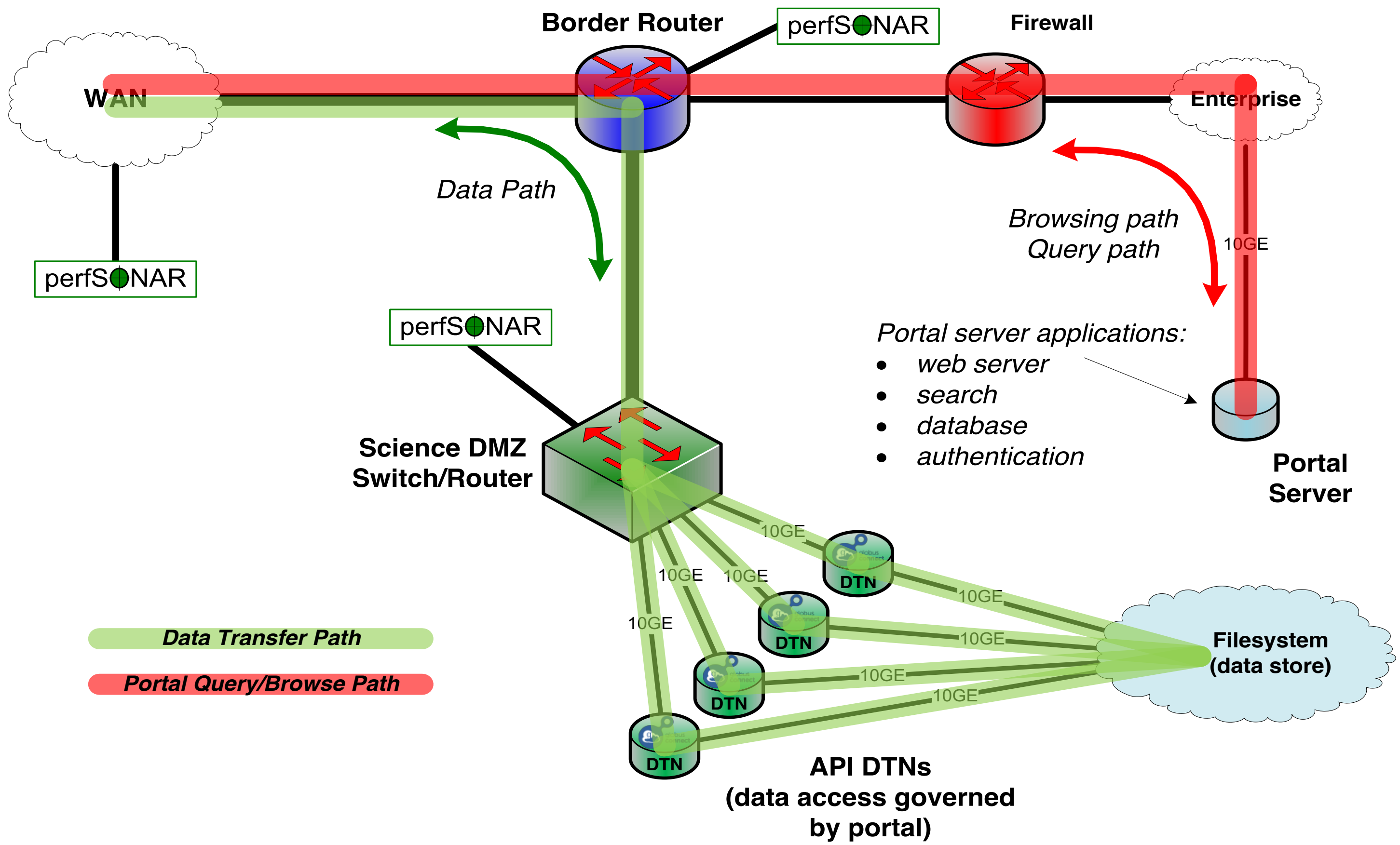
Legacy Portal Design



- Very difficult to improve performance without architectural change
  - Software components all tangled together
  - Difficult to put the whole portal in a Science DMZ because of security
  - Even if you could put it in a DMZ, many components aren't scalable
- What does architectural change mean?

1986–2016 ESnet 30 YEARS OF NETWORKING AT THE SPEED OF SCIENCE

Next-Generation Portal Leverages Science DMZ



Put The Data On Dedicated Infrastructure

- We have separated the data handling from the portal logic
- Portal is still its normal self, but enhanced
  - Portal GUI, database, search, etc. all function as they did before
  - Query returns pointers to data objects in the Science DMZ
  - Portal is now freed from ties to the data servers (run it on Amazon if you want!)
- Data handling is separate, and scalable
  - High-performance DTNs in the Science DMZ
  - Scale as much as you need to without modifying the portal software
- Outsource data handling to computing centers or campus central storage
  - Computing centers are set up for large-scale data
  - Let them handle the large-scale data, and let the portal do the orchestration of data placement

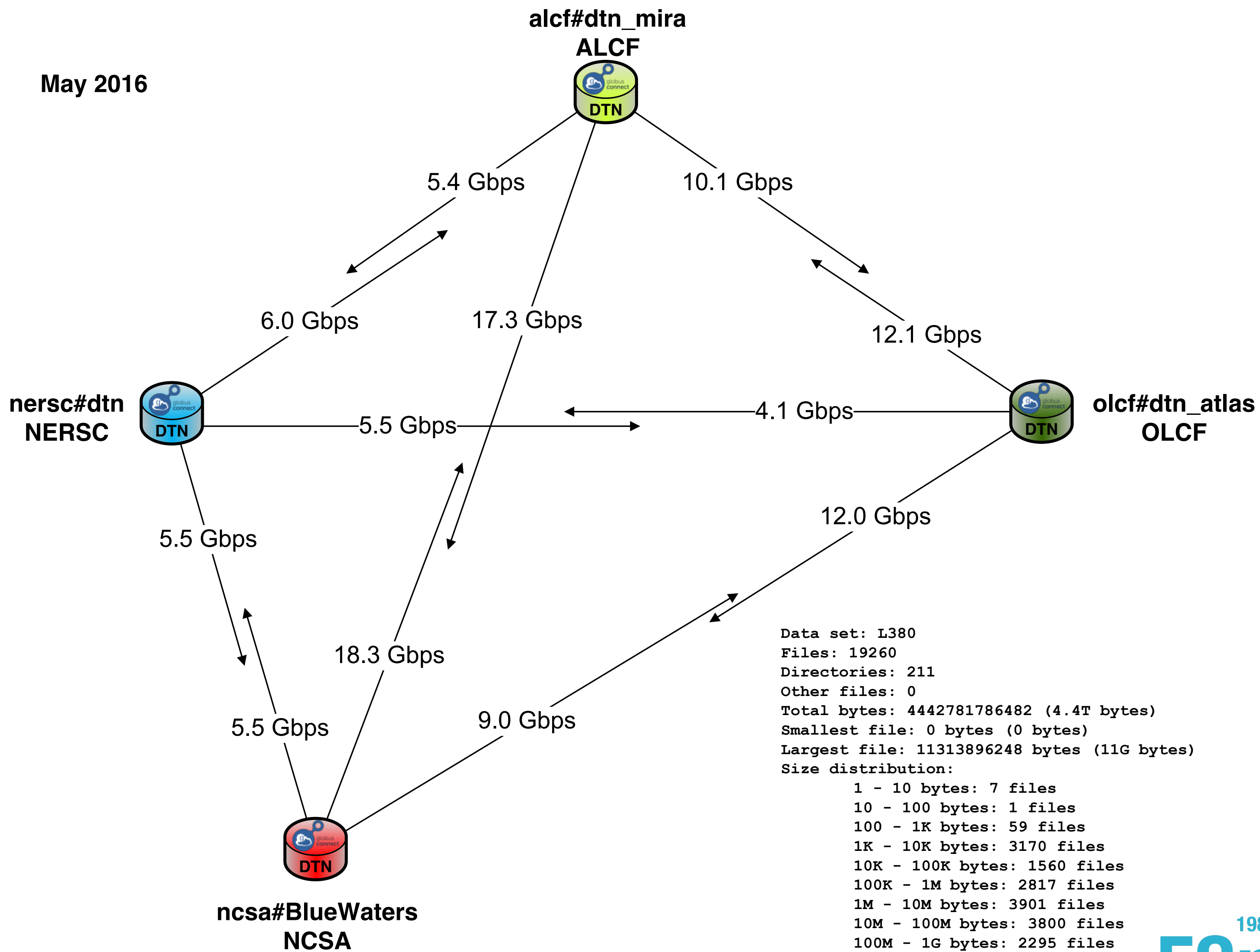
Ecosystem Is Ready For This

- Science DMZs are deployed at Labs, Universities, and computing centers
  - XSEDE sites
  - DOE HPC facilities
  - Many campus clusters
- Globus DTNs are present in many of those Science DMZs
  - XSEDE sites
  - DOE HPC facilities
  - Many campus clusters
- Architectural change allows data placement at scale
  - Submit a query to the portal, Globus places the data at an HPC facility
  - Run the analysis at the HPC facility
  - The results are the only thing that ends up on a laptop or workstation

Petascale DTN Project

- Another example of building on the Science DMZ
- Supports all data-intensive applications which require large-scale data placement
- Collaboration between HPC facilities
  - ALCF, NCSA, NERSC, OLCF
- Goal: per-Globus-job performance at 1PB/week level
  - 15 gigabits per second
  - With checksums turned on, etc.
  - No special shortcuts, no arcane options
- Reference data set is 4.4TB of astrophysics model output
  - Mix of file sizes
  - Many directories
  - Real data!

Petascale DTN Project




```
Data set: L380
Files: 19260
Directories: 211
Other files: 0
Total bytes: 4442781786482 (4.4T bytes)
Smallest file: 0 bytes (0 bytes)
Largest file: 11313896248 bytes (11G bytes)
Size distribution:
        1 - 10 bytes: 7 files
        10 - 100 bytes: 1 files
        100 - 1K bytes: 59 files
        1K - 10K bytes: 3170 files
        10K - 100K bytes: 1560 files
        100K - 1M bytes: 2817 files
        1M - 10M bytes: 3901 files
        10M - 100M bytes: 3800 files
        100M - 1G bytes: 2295 files
        1G - 10G bytes: 1647 files
        10G - 100G bytes: 3 files
```

Links and Lists

- 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
  - Send mail to sympa@lists.lbl.gov with subject "subscribe esnet-sciencedmz"
- perfSONAR
  - http://fasterdata.es.net/performance-testing/perfsonar/
  - http://www.perfsonar.net
- Globus
  - https://www.globus.org/

1986–2016

ESnet

30 YEARS OF NETWORKING

AT THE SPEED OF SCIENCE

Thanks!

Eli Dart
dart@es.net
Energy Sciences Network (ESnet)
Lawrence Berkeley National Laboratory

http://my.es.net/

http://www.es.net/

http://fasterdata.es.net/

U.S. DEPARTMENT OF ENERGY
Office of Science

BERKELEY LAB

Extra Slides

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

1986–2016 ESnet 30 YEARS OF NETWORKING AT THE SPEED OF SCIENCE

Science DMZ As Security Architecture

- Allows for better segmentation of risks, more granular application of controls to those segmented risks.
  - Limit risk profile for high-performance data transfer applications
  - Apply specific controls to data transfer hosts
  - Avoid including unnecessary risks, unnecessary controls
- Remove degrees of freedom – focus only on what is necessary
  - Easier to secure
  - Easier to achieve performance
  - Easier to troubleshoot

Performance Is A Core Requirement

- Core information security principles
  - Confidentiality, Integrity, Availability (CIA)
  - Often, CIA and risk mitigation result in poor performance
- In data-intensive science, performance is an additional core mission requirement: CIA → PICA
  - CIA principles are important, but ***if performance is compromised the science mission fails***
  - Not about “how much” security you have, but how the security is implemented
  - Need a way to appropriately secure systems without performance compromises

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***
  - Packet filtering is great – just don't do it with an enterprise firewall
- 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 enterprise firewalls are typically a poor fit for high-performance science applications

Typical Firewall Internals

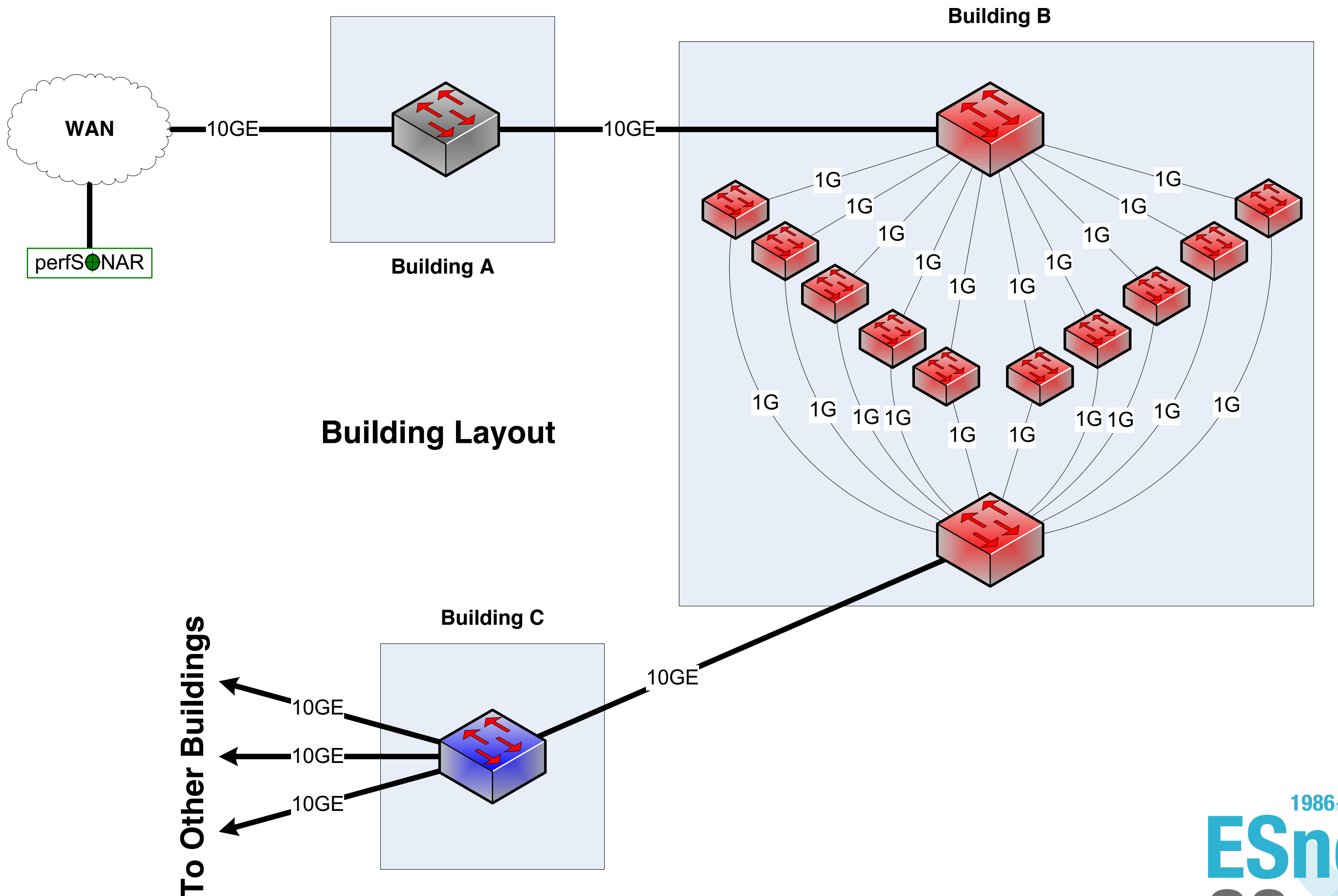
- Typical firewalls are composed of a set of processors which inspect traffic in parallel
  - Traffic distributed among processors such that all traffic for a particular connection goes to the same processor
  - Simplifies state management
  - Parallelization scales deep analysis
- Excellent fit for enterprise traffic profile
  - High connection count, low per-connection data rate
  - Complex protocols with embedded threats
- Each processor is a fraction of firewall link speed
  - Significant limitation for data-intensive science applications
  - Overload causes packet loss – performance crashes

1986–2016 ESnet 30 YEARS OF NETWORKING AT THE SPEED OF SCIENCE

Thought Experiment

- We're going to do a thought experiment
- Consider a network between three buildings – A, B, and C
- This is supposedly a 10Gbps network end to end (look at the links on the buildings)
- Building A houses the border router – not much goes on there except the external connectivity
- Lots of work happens in building B – so much that the processing is done with multiple processors to spread the load in an affordable way, and results are aggregated after
- Building C is where we branch out to other buildings
- Every link between buildings is 10Gbps – this is a 10Gbps network, right???

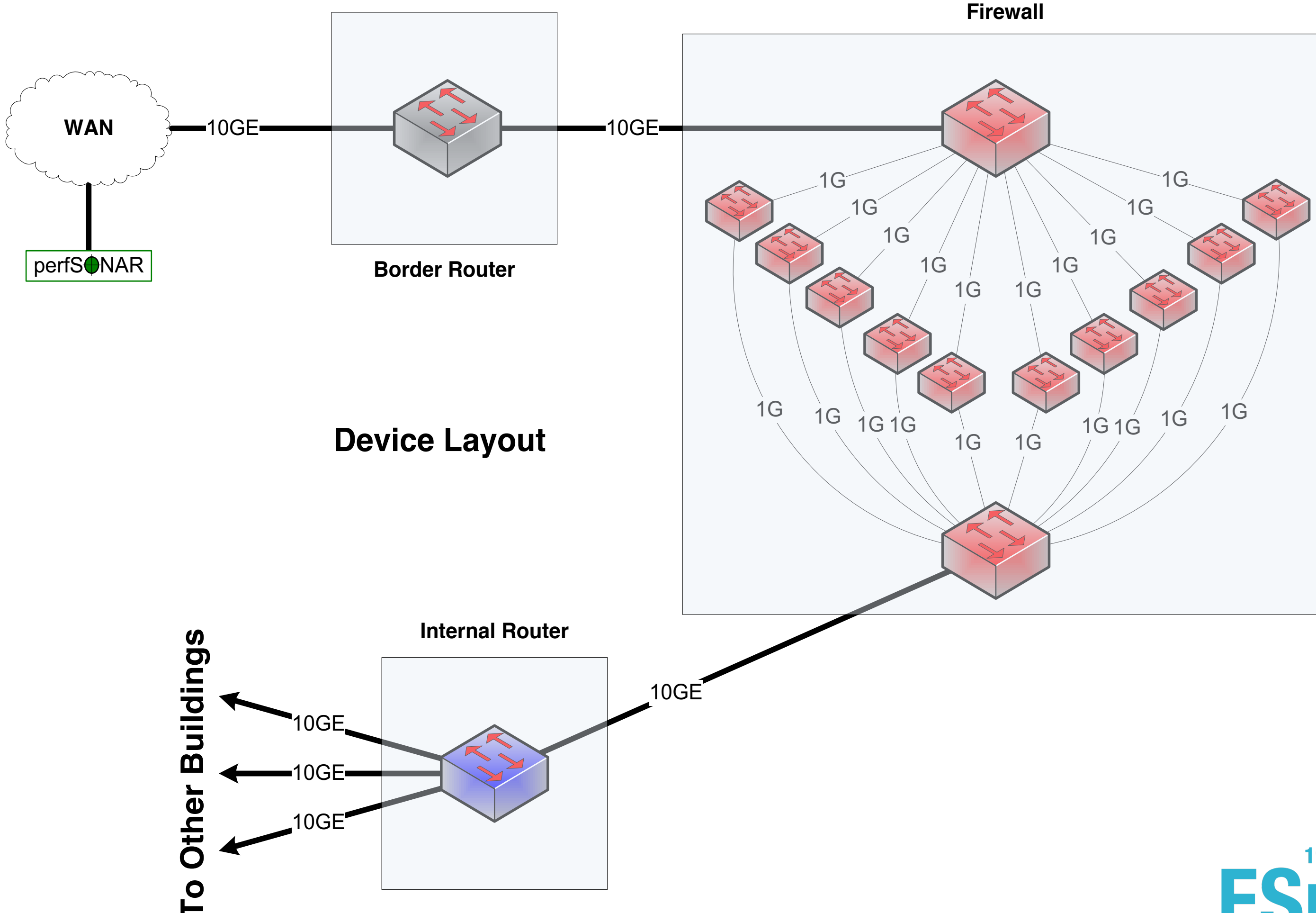
Notional 10G Network Between Buildings



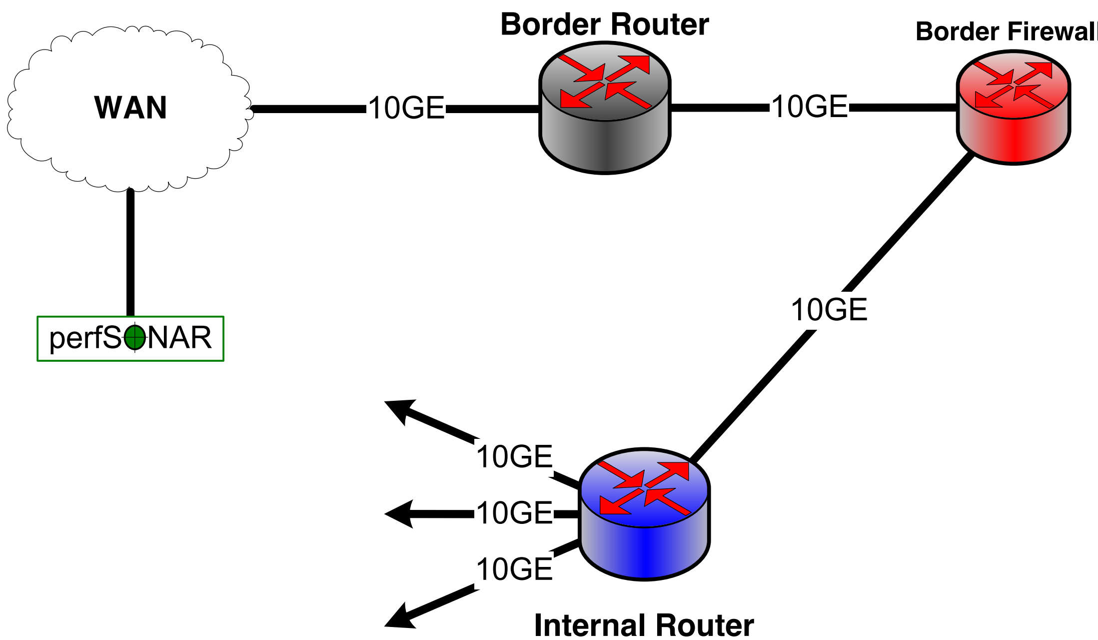
Clearly Not A 10Gbps Network

- If you look at the inside of Building B, it is obvious from a network engineering perspective that this is not a 10Gbps network
  - Clearly the maximum per-flow data rate is 1Gbps, not 10Gbps
  - However, if you convert the buildings into network elements while keeping their internals intact, you get routers and firewalls
  - What firewall did the organization buy? What's inside it?
  - Those little 1G "switches" are firewall processors
- This parallel firewall architecture has been in use for years
  - Slower processors are cheaper
  - Typically fine for a commodity traffic load
  - Therefore, this design is cost competitive and common

Notional 10G Network Between Devices



Notional Network Logical Diagram



Firewall Capabilities and Science Traffic

- Commercial firewalls have a lot of sophistication in an enterprise setting
  - Application layer protocol analysis (HTTP, POP, MSRPC, etc.)
  - Built-in VPN servers
  - User awareness
- Data-intensive science flows typically don't match this profile
  - Common case – data on filesystem A needs to be on filesystem Z
    - Data transfer tool verifies credentials over an encrypted channel
    - Then open a socket or set of sockets, and send data until done (1TB, 10TB, 100TB, …)
  - One workflow can use 10% to 50% or more of a 10G network link
- Do we have to use a commercial firewall?

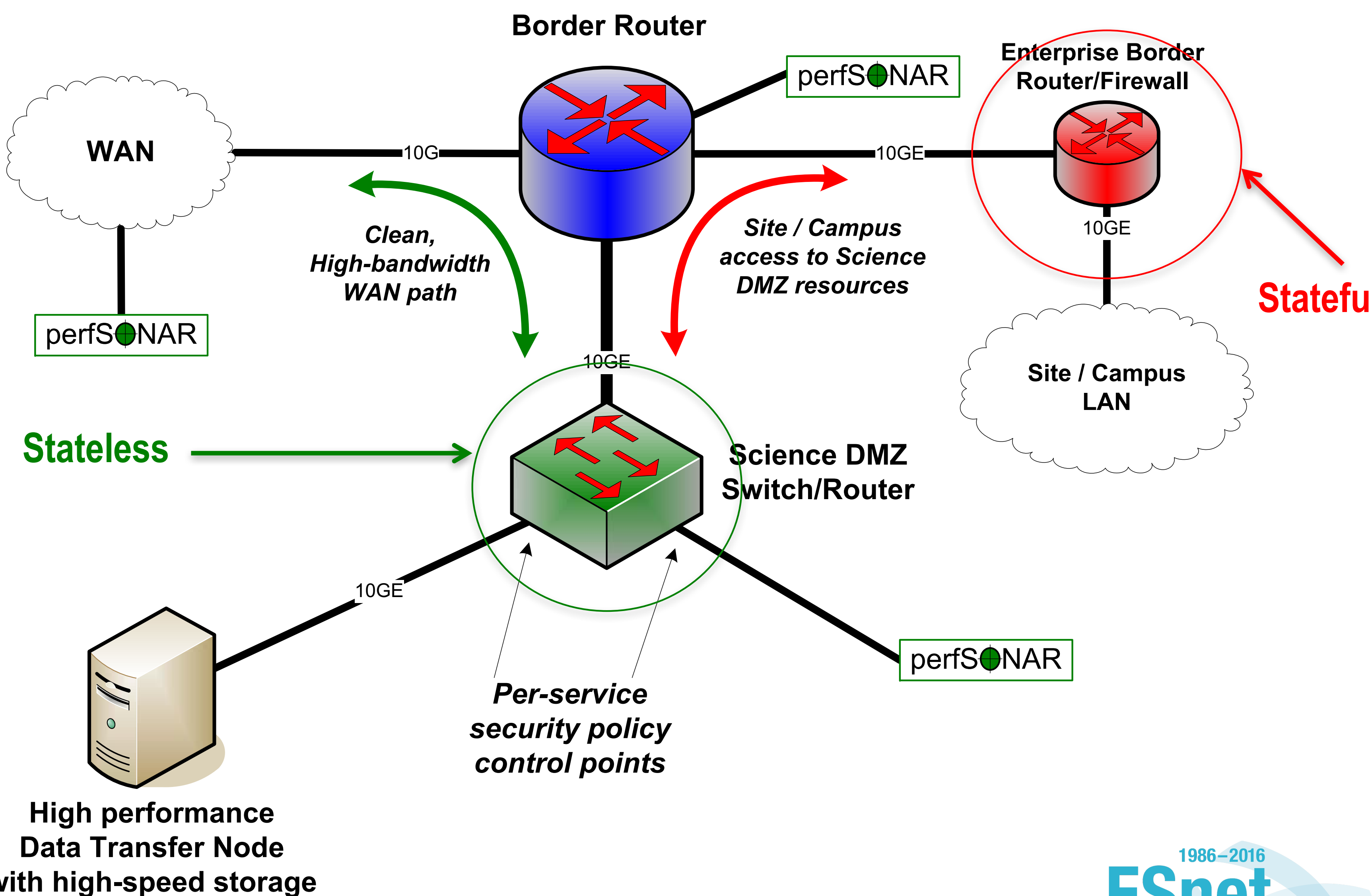
Firewalls As Access Lists

- When you ask a firewall administrator to allow data transfers through the firewall, what do they ask for?
  - IP address of your host
  - IP address of the remote host
  - Port range
  - ***That looks like an ACL to me!***
- No special config for advanced protocol analysis – just address/port
- Router ACLs are better than firewalls at address/port filtering
  - ACL capabilities are typically built into the router
  - Router ACLs typically do not drop traffic permitted by policy

What Is A Firewall?

- Marketplace view
  - Specific security appliance, with “Firewall” printed on the side
  - Lots of protocol awareness, intelligence
  - Application awareness
  - User awareness (VPN, specific access controls, etc.)
  - Designed for large concurrent user count, low per-user bandwidth (enterprise traffic)
- IT Organization view
  - “Firewall” appliance, purchased from the commercial marketplace
  - The place in the network where security policy gets applied
  - Owned by the security group, ***not*** by the networking group
  - Primary risk mitigation mechanism
- NIST view (Publication 800-41 rev. 1, Sep. 2009)
  - “Firewalls are devices or programs that control the flow of network traffic between networks or hosts that employ differing security postures”
  - This is very general, and does not match marketplace view or IT org. view

NIST Sees Two Firewalls, IT Shop Sees One



Stateful Inspection For Science DMZ Traffic?

- Science DMZ traffic profile
  - Small number of connections or flows
  - Large per-connection data rate (Gigabit scale or higher)
  - Large per-connection data volume (Terabyte scale or higher)
- Stateless firewall
  - Address/port filtering (which systems use which service)
  - TCP connection initiation direction (ACK flag)
- Stateful firewall adds
  - TCP sequence number tracking (but Linux stack is as good or better compared to firewall TCP mitigations)
  - Protocol/app analysis (but not for the apps used in DMZ)
  - DoS protection (but the Science DMZ assets are filtered already)

Security Without Enterprise Firewalls

- Data intensive science traffic interacts poorly with enterprise firewalls
- Does this mean we ignore security? ***NO!***
  - We **must** protect our systems
  - We just need to find a way to do security that does not prevent us from getting the science done
- ***Key point – security policies and mechanisms that protect the Science DMZ should be implemented so that they do not compromise performance***
- Traffic permitted by policy should not experience performance impact as a result of the application of policy

Systems View Of Science Infrastructure

- Security is a component, not a gatekeeper
- Think about the workflows
- Think about the interfaces to data (tools, applications)
  - How do collaborators access data?
  - How could they access data if the architecture were different?
- Think about costs/benefits
  - What is a new cancer breakthrough worth?
  - $30k for a few DTNs – what is that in context?
- Think about risks
  - What risks do specific technologies mitigate?
  - What are opportunity costs of poor performance?

Other Technical Capabilities

- Intrusion Detection Systems (IDS)
  - One example is Bro – http://bro-ids.org/
  - Bro is high-performance and battle-tested
    - Bro protects several high-performance national assets
    - Bro can be scaled with clustering: http://www.bro-ids.org/documentation/cluster.html
  - Other IDS solutions are available also
- Netflow and IPFIX can provide intelligence, but not filtering
- Openflow and SDN
  - Using Openflow to control access to a network-based service seems pretty obvious
  - This could significantly reduce the attack surface for any authenticated network service
  - This would only work if the Openflow device had a robust data plane

Other Technical Capabilities (2)

- Aggressive access lists
  - More useful with project-specific DTNs
  - If the purpose of the DTN is to exchange data with a small set of remote collaborators, the ACL is pretty easy to write
  - Large-scale data distribution servers are hard to handle this way (but then, the firewall ruleset for such a service would be pretty open too)
- Limitation of the application set
  - One of the reasons to limit the application set in the Science DMZ is to make it easier to protect
  - Keep desktop applications off the DTN (and watch for them anyway using logging, netflow, etc – take violations seriously)
  - This requires collaboration between people – networking, security, systems, and scientists

Collaboration Within The Organization

- All stakeholders should collaborate on Science DMZ design, policy, and enforcement
- The security people have to be on board
  - Remember: security people already have political cover – it's called the firewall
  - If a host gets compromised, the security officer can say they did their due diligence because there was a firewall in place
  - If the deployment of a Science DMZ is going to jeopardize the job of the security officer, expect pushback
- The Science DMZ is a strategic asset, and should be understood by the strategic thinkers in the organization
  - Changes in security models
  - Changes in operational models
  - Enhanced ability to compete for funding
  - Increased institutional capability – greater science output

1986–2016 ESnet 30 YEARS OF NETWORKING AT THE SPEED OF SCIENCE