TRUSTED CI

THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

Science and Networks and Cybersecurity

Von Welch
Director, Trusted CI

Training Workshop for Network Engineers and Educators on Tools and Protocols for High-Speed Networks

July 23, 2019

TRUSTED CI
THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

Thank you to ESnet - Eli Dart, Michael Sinatra, Vlad Grigorescu, and Adam Slagell - for content and concepts throughout this talk.

My Talk

1. A little about Trusted CI
2. Why am I talking about Networks and Science and Cybersecurity?
3. Actionable advice

Trusted CI:
The NSF Cybersecurity Center of Excellence

NSF

Our mission: to lead in the development of an NSF Cybersecurity Ecosystem with the workforce, knowledge, processes, and cyberinfrastructure that enables trustworthy science and NSF's vision of a nation that is a global leader in research and innovation.

CENTER FOR APPLIED CYBERSECURITY RESEARCH
INDIANA UNIVERSITY
Pervasive Technology Institute

NCSA

INTERNET2

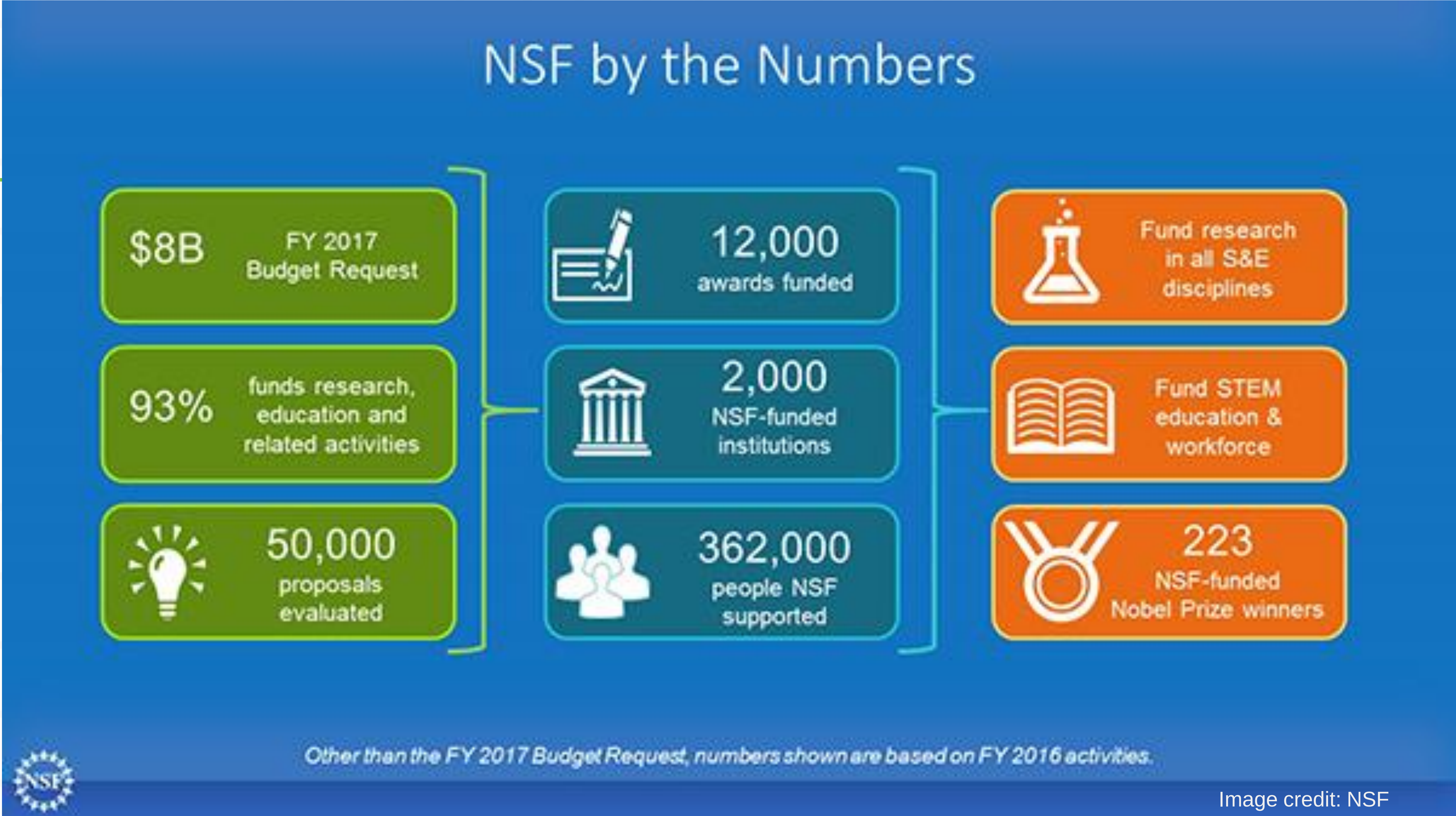
WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

PITTSBURGH SUPERCOMPUTING CENTER

BERKELEY LAB

TRUSTED CI
THE NSF CYBERSECURITY CENTER OF EXCELLENCE

https://trustedci.org/



NSF by the Numbers

$8B FY 2017 Budget Request

93% funds research, education and related activities

50,000 proposals evaluated

12,000 awards funded

2,000 NSF-funded institutions

362,000 people NSF supported

Fund research in all S&E disciplines

Fund STEM education & workforce

223 NSF-funded Nobel Prize winners

Other than the FY 2017 Budget Request, numbers shown are based on FY 2016 activities.

Image credit: NSF

Trusted CI: Impacts

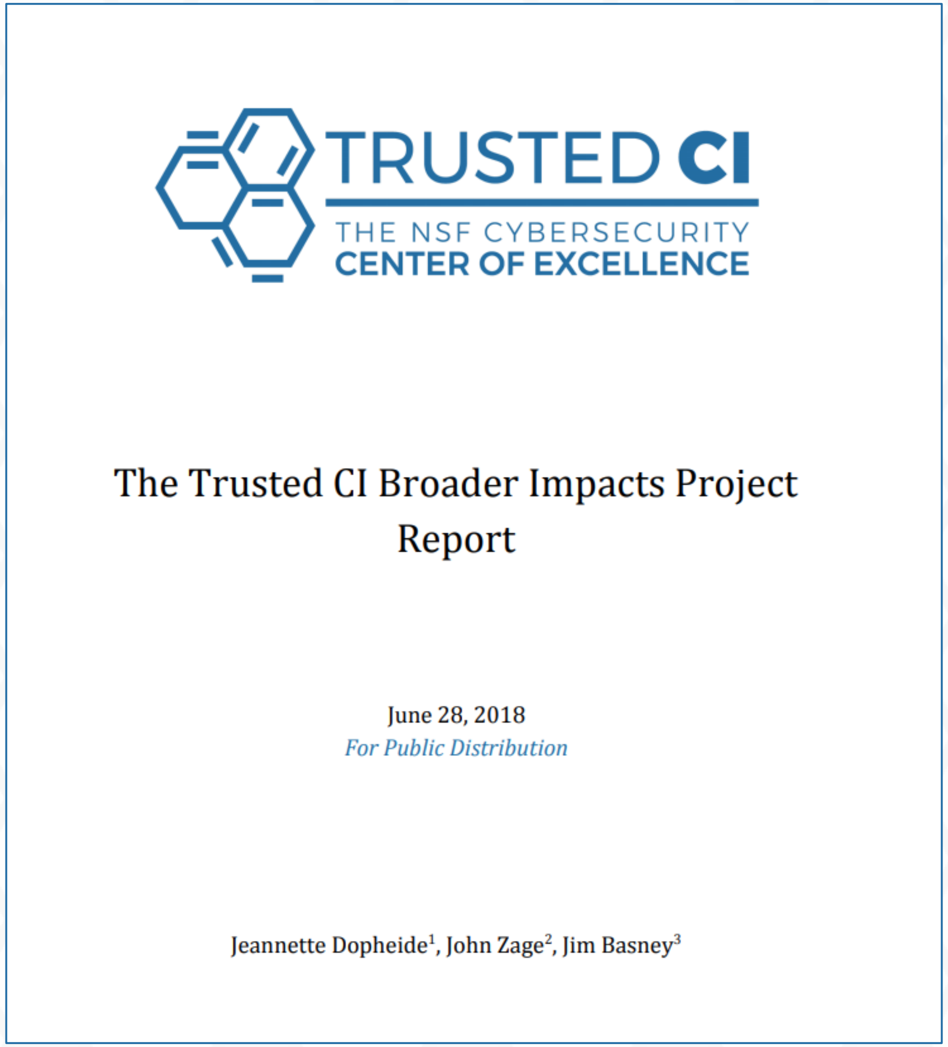
Trusted CI has positively impacted over 260 NSF projects since inception in 2012.

Members of more than 180 NSF projects have attended our NSF Cybersecurity Summit.

Members of more than 80 NSF projects have attended our monthly webinars.

We have provided more than 300 hours of training to the community.

We've had engagements with 41 projects, including nine NSF Large Facilities.



TRUSTED CI
THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

The Trusted CI Broader Impacts Project Report

June 28, 2018
For Public Distribution

Jeannette Dopheide[1], John Zage[2], Jim Basney[3]

https://hdl.handle.net/2022/22148

Best Practices

Security Best Practices for Academic Cloud Service Providers

https://trustedci.org/cloud-service-provider-security-best-practices/

Operational Security

https://trustedci.org/guide

Identity Management Best Practices

https://trustedci.org/iam

Science Gateways

https://trustedci.org/sgci/

Software Assurance

https://trustedci.org/software-assurance/

Agave Platform

Cornell University Center for Advanced Computing

CYVERSE

Jetstream

TRUSTED CI THE NSF CYBERSECURITY CENTER OF EXCELLENCE

Security Best Practices for Academic Cloud Service Providers

Version 1.0

http://hdl.handle.net/2022/22123

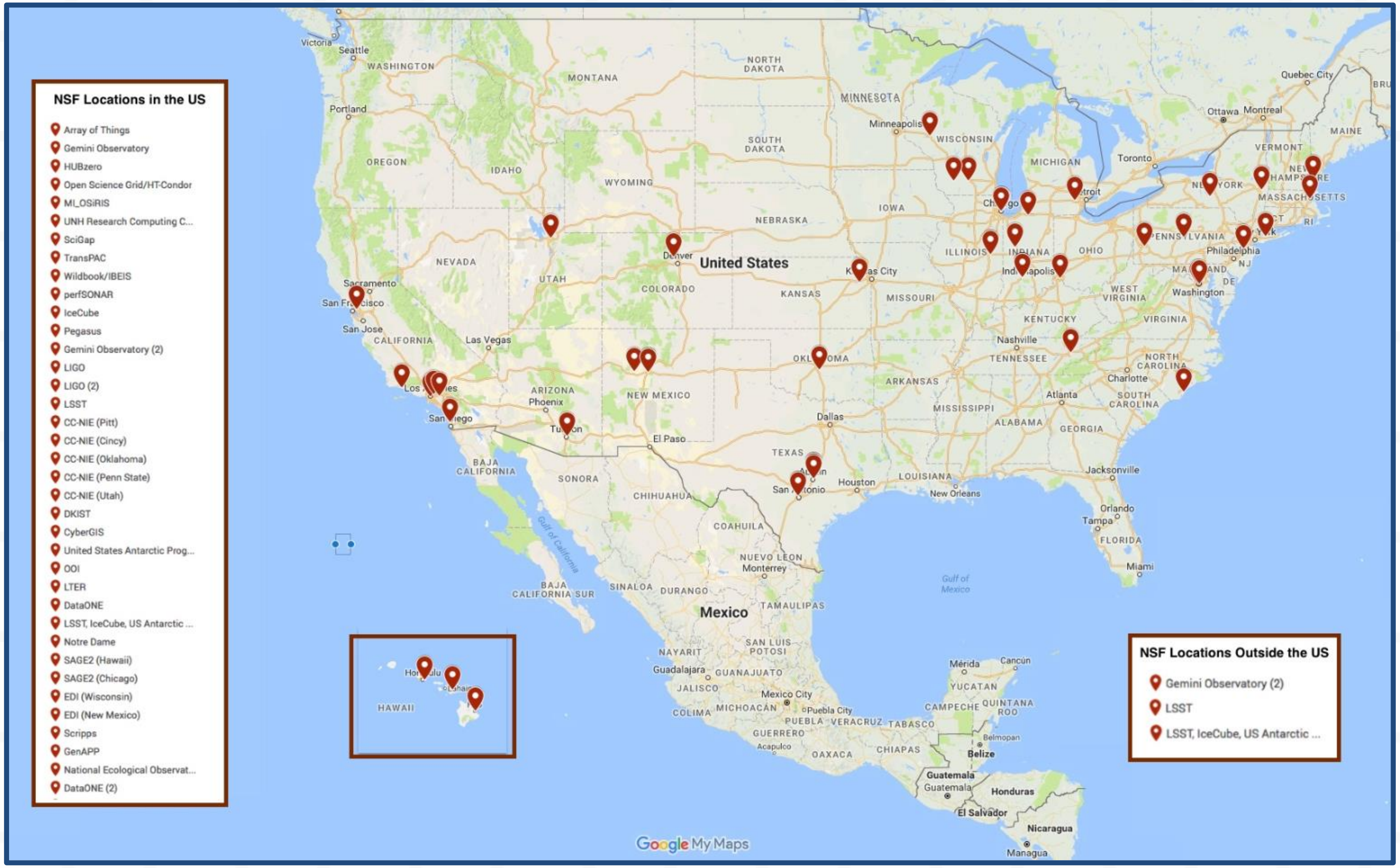
Engagements: One-on-one Collaborations

We take applications every six months.

Accept applications every six months:

https://trustedci.org/application/

Next deadline will be Sep/Oct 2019.



TRUSTED CI
THE NSF CYBERSECURITY CENTER OF EXCELLENCE

Trusted CI Partners

SGCI Science Gateways Community Institute

InCommon®

Open Science Grid

EDUCAUSE

HEISC

Engagement and Performance Operations Center (EPOC)

ResearchSOC

Software Sustainability Institute

XSEDE

REN-ISAC

CI CoE PILOT

ESnet ENERGY SCIENCES NETWORK

CPP SFS

TRUSTED CI
THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

https://trustedci.org/partners

Annual NSF Cybersecurity Summit

One day of training and workshops.

Lessons learned and success from community.

Oct 15-17, 2019 in San Diego

https://trustedci.org/summit/

Agenda driven by call for participation, due August 12th

https://trustedci.org/cfp2019

TRUSTED CI
THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

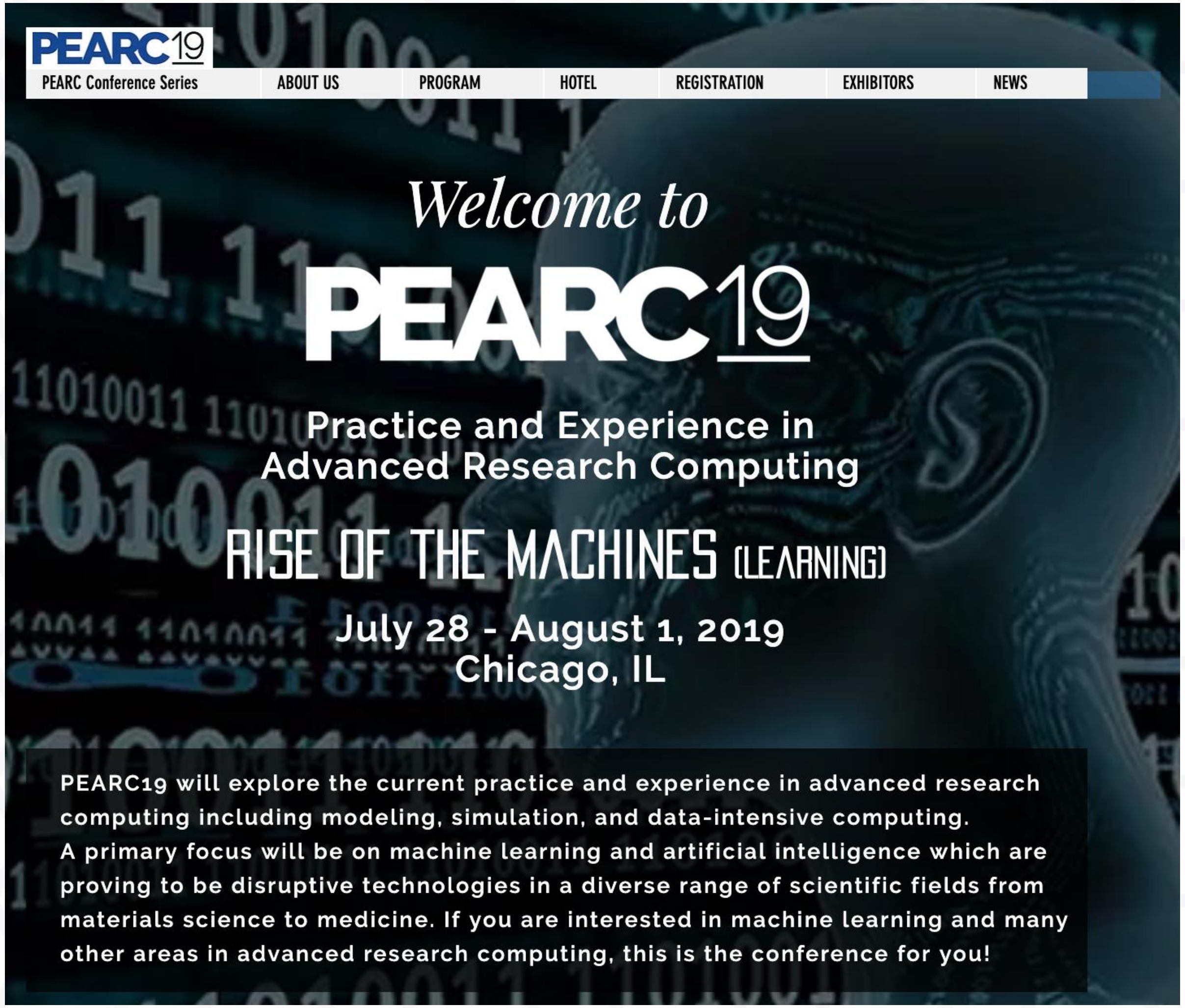
See you at PEARC'19

July 28 - August 1, 2019 in Chicago

https://www.pearc19.pearc.org/

https://blog.trustedci.org/2019/06/many-opportunities-to-meet-with-trusted.html

https://blogs.iu.edu/researchsoc/2019/06/10/join-researchsoc-at-pearc19/



Welcome to

PEARC19

Practice and Experience in Advanced Research Computing

RISE OF THE MACHINES (LEARNING)

July 28 - August 1, 2019
Chicago, IL

PEARC19 will explore the current practice and experience in advanced research computing including modeling, simulation, and data-intensive computing.
A primary focus will be on machine learning and artificial intelligence which are proving to be disruptive technologies in a diverse range of scientific fields from materials science to medicine. If you are interested in machine learning and many other areas in advanced research computing, this is the conference for you!

TRUSTED CI
THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

Trusted CI and Inclusivity

Cybersecurity requires diverse perspectives and cybersecurity community suffers from a lack of diversity.

Trusted CI works to address it through its workforce development, outreach, and community building efforts by explicitly seeking out and encouraging underrepresented groups to apply and striving for inclusive demographics.

2018 NSF Cybersecurity Summit Student Program

TRUSTED CI
THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

Staying Connected with Trusted CI

Trusted CI Webinars

4th Monday of month at 10am ET.

https://trustedci.org/webinars

Follow Us

https://trustedci.org

https://blog.trustedci.org

@TrustedCI 🐦

Email Lists

Announce and Discuss

https://trustedci.org/trustedci-email-lists

Ask Us Anything

No question too big or too small.

info@trustedci.org

Cyberinfrastructure Vulnerabilities

Latest news on security vulnerabilities tailored for cyberinfrastructure community.

https://trustedci.org/vulnerabilities/

TRUSTED CI
THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

My Talk

1. A little about Trusted CI
2. Why am I talking about Networks and Science and Cybersecurity?
3. Actionable advice

Think of all the activities on a Campus…

Image Credit: Indiana University

TRUSTED CI
THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

EXIT
EXIT

PREMIUM Credit
1234 5678 9012 3456
07/20
SMITH

INDIANA
Indiana

ECOphlex
BIG RED II
WHAT MATTERS WHERE IT MATTERS

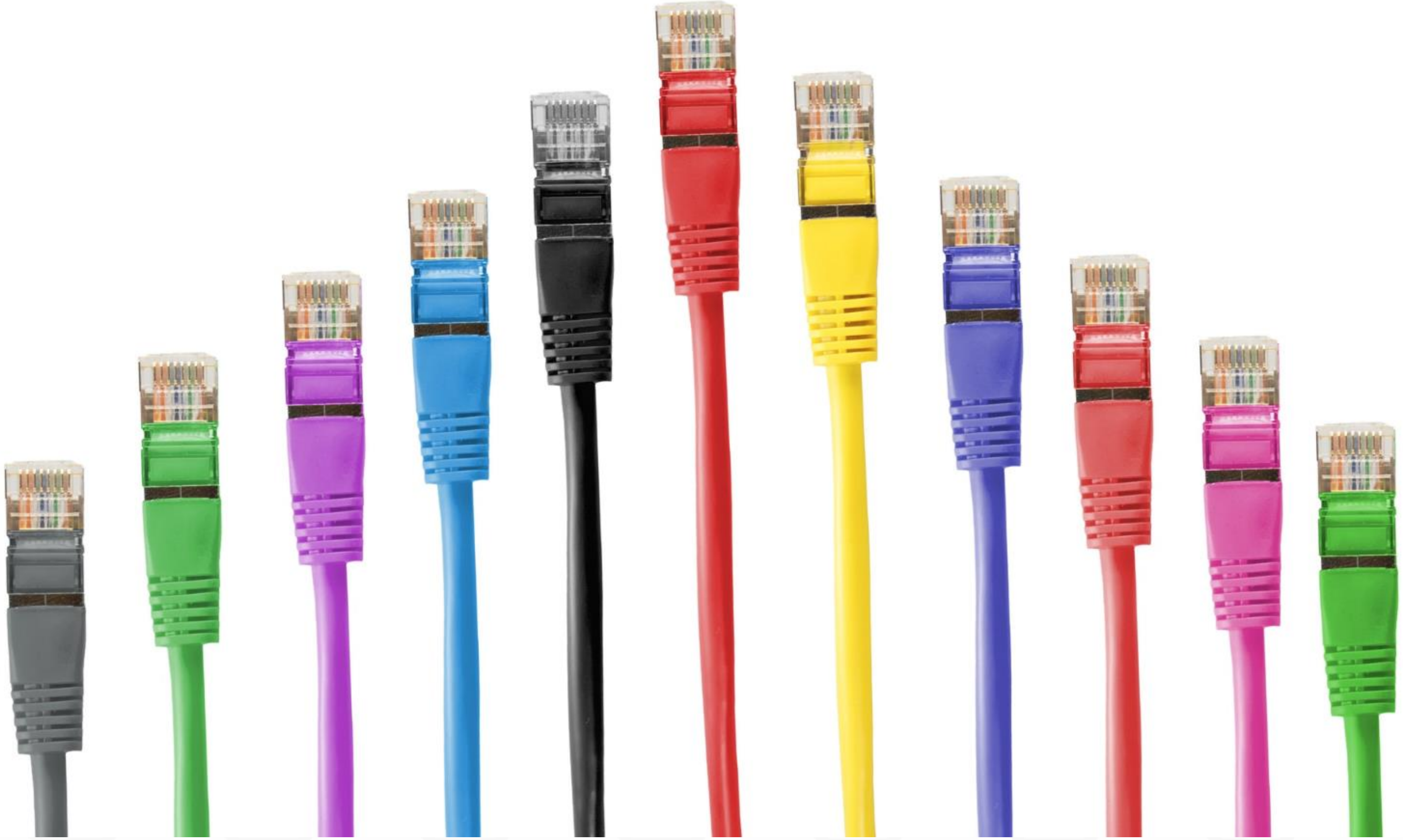
Each Activity Has Different Requirements for Performance...

TRUSTED CI
THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

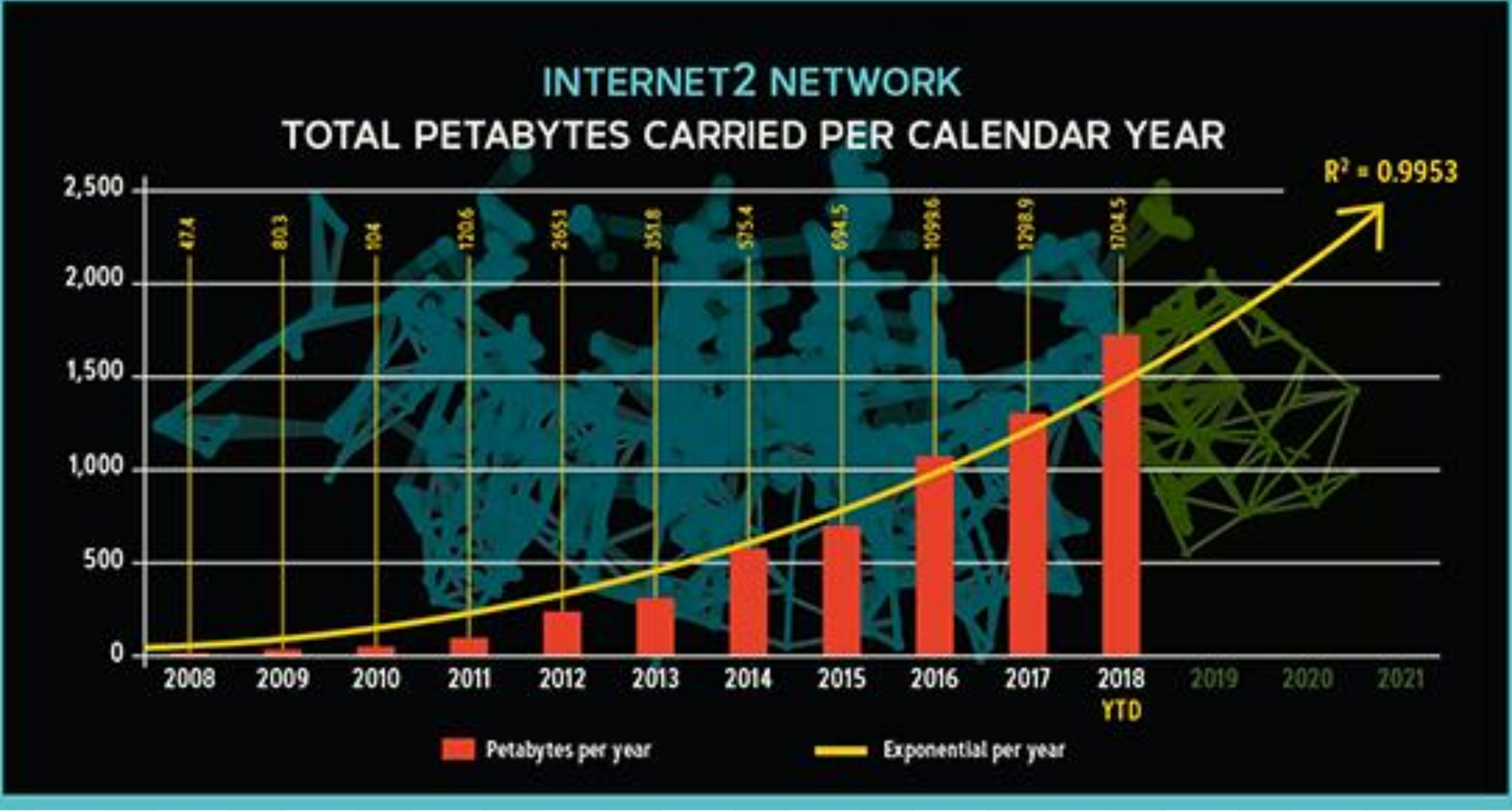
...and access/security.

TRUSTED CI
THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

Network segmentation is a well accepted solution



The Challenge for Science



INTERNET2 NETWORK

TOTAL PETABYTES CARRIED PER CALENDAR YEAR

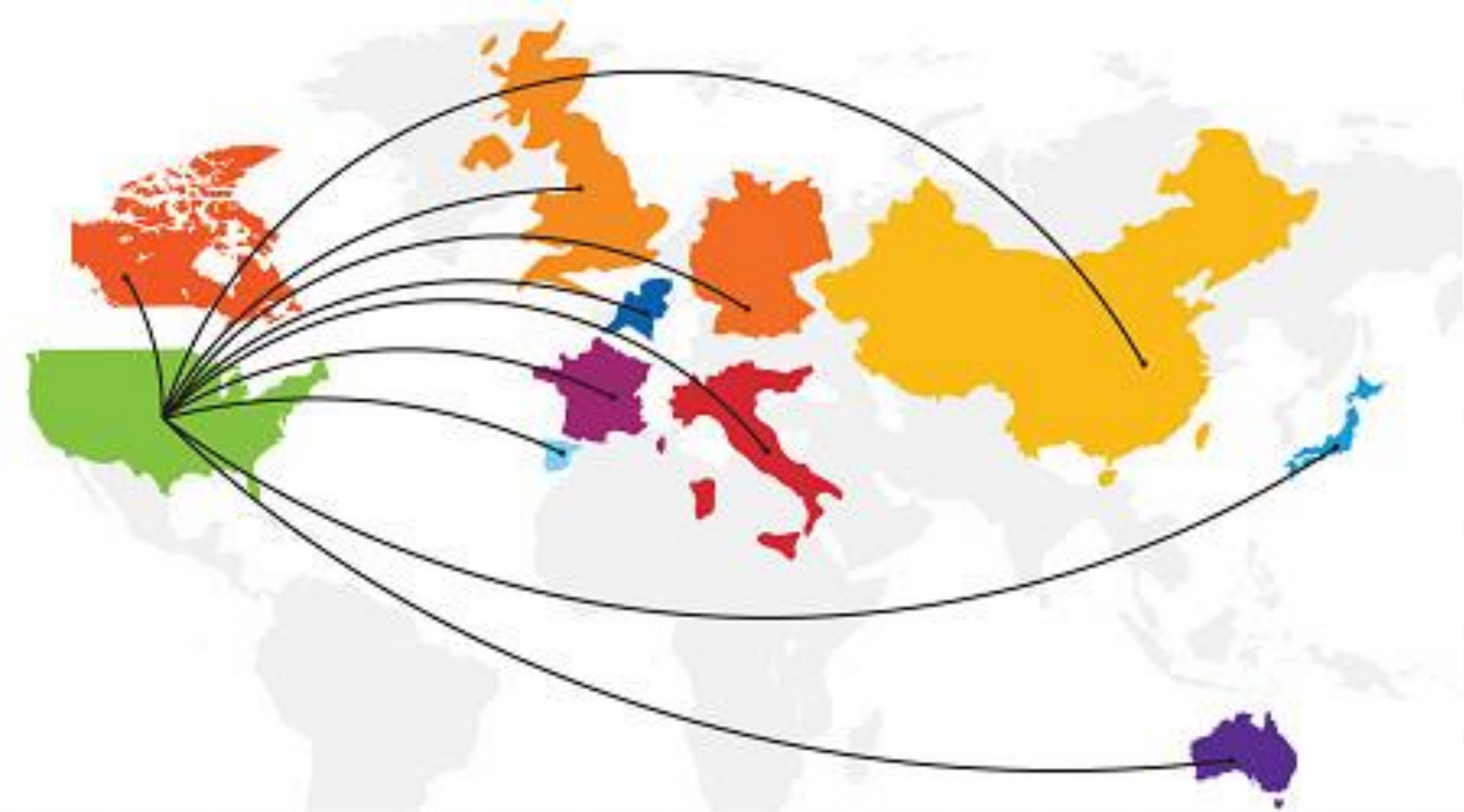
https://www.internet2.edu/blogs/detail/16934

Rapid, Collaborative Projects

Often short-lived (3-5 years).

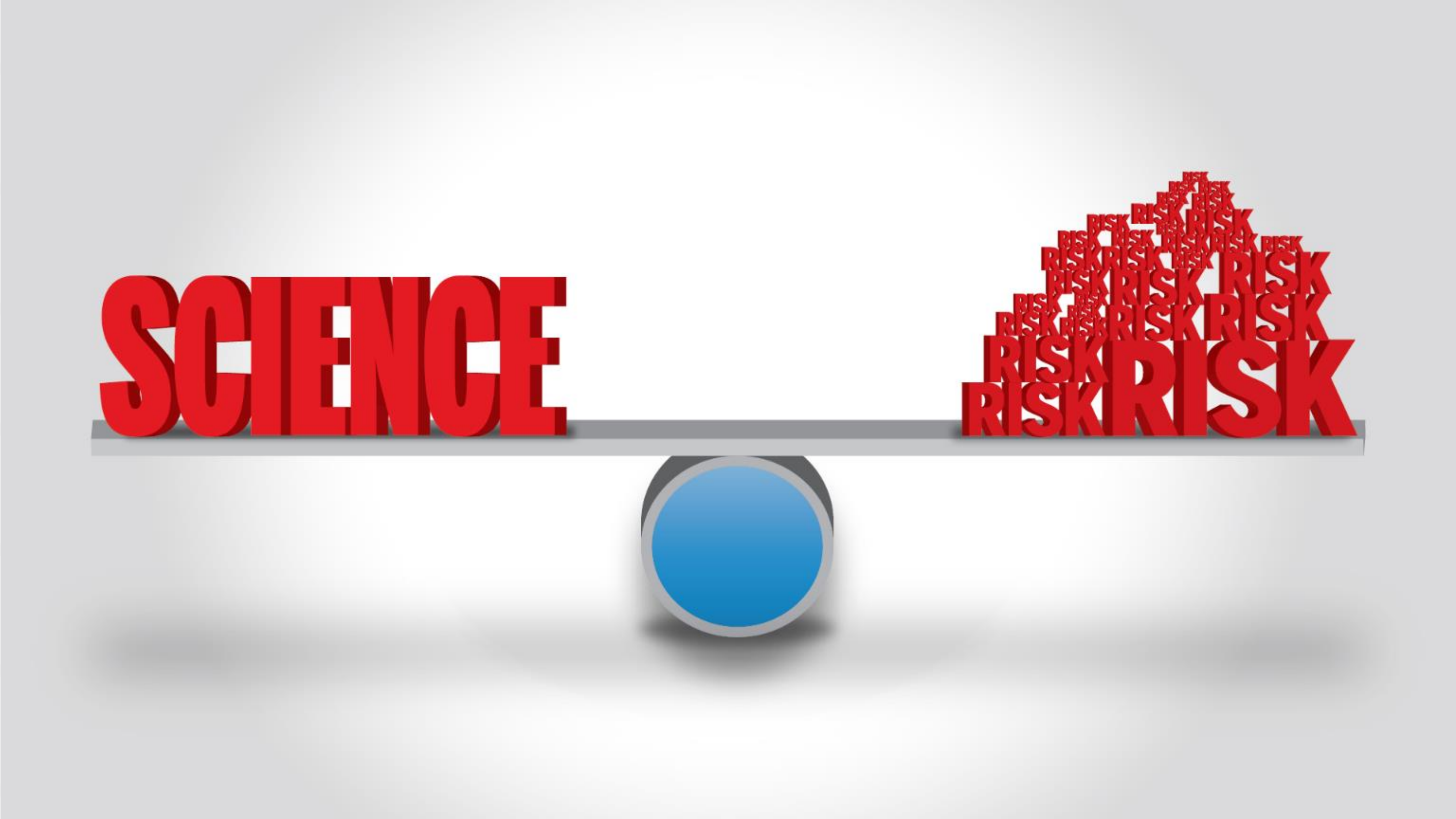
Start and progress quickly.

Researcher-managed teams.



Top ten countries with most U.S. co-authors in 2015
https://www.aje.com/arc/collaboration-2015/

Hence, the Need:



A Network Segment That Allows Secure And Fast Data Access to Distributed Collaborators

My Talk

1. A little about Trusted CI
2. Why am I talking about Networks and Science and Cybersecurity?
3. Actionable advice

____Following slides taken with little change from ESnet

TRUSTED CI

THE NSF CYBERSECURITY CENTER OF EXCELLENCE

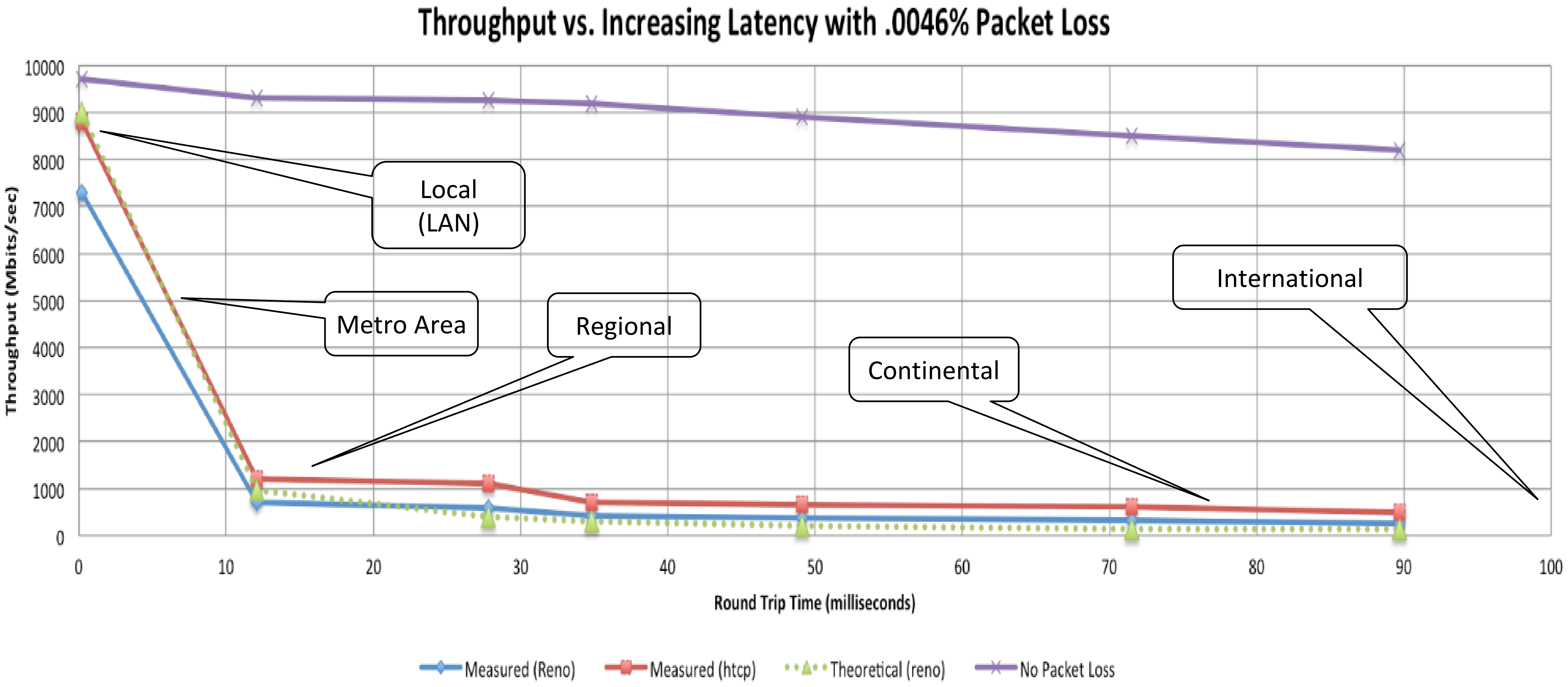
How is Science Data Being Transferred?

- A small number of (very) large flows
  - 10 Gigabit minimum, 100 and 400 Gigabit in production, 1 Terabit in planning stages
- GridFTP is the de-facto standard

> *“GridFTP is a **high-performance**, **secure**, **reliable** data transfer protocol optimized for **high-bandwidth wide-area networks**. The GridFTP protocol is based on FTP, the highly-popular Internet file transfer protocol.”*

- Focus on “data transfer nodes” (DTNs)
  - Systems designed from the ground up for lightning-fast disk-to-network transfers
  - https://fasterdata.es.net/science-dmz/DTN/

Effects of Packet Loss



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

Science DMZ Security Myth

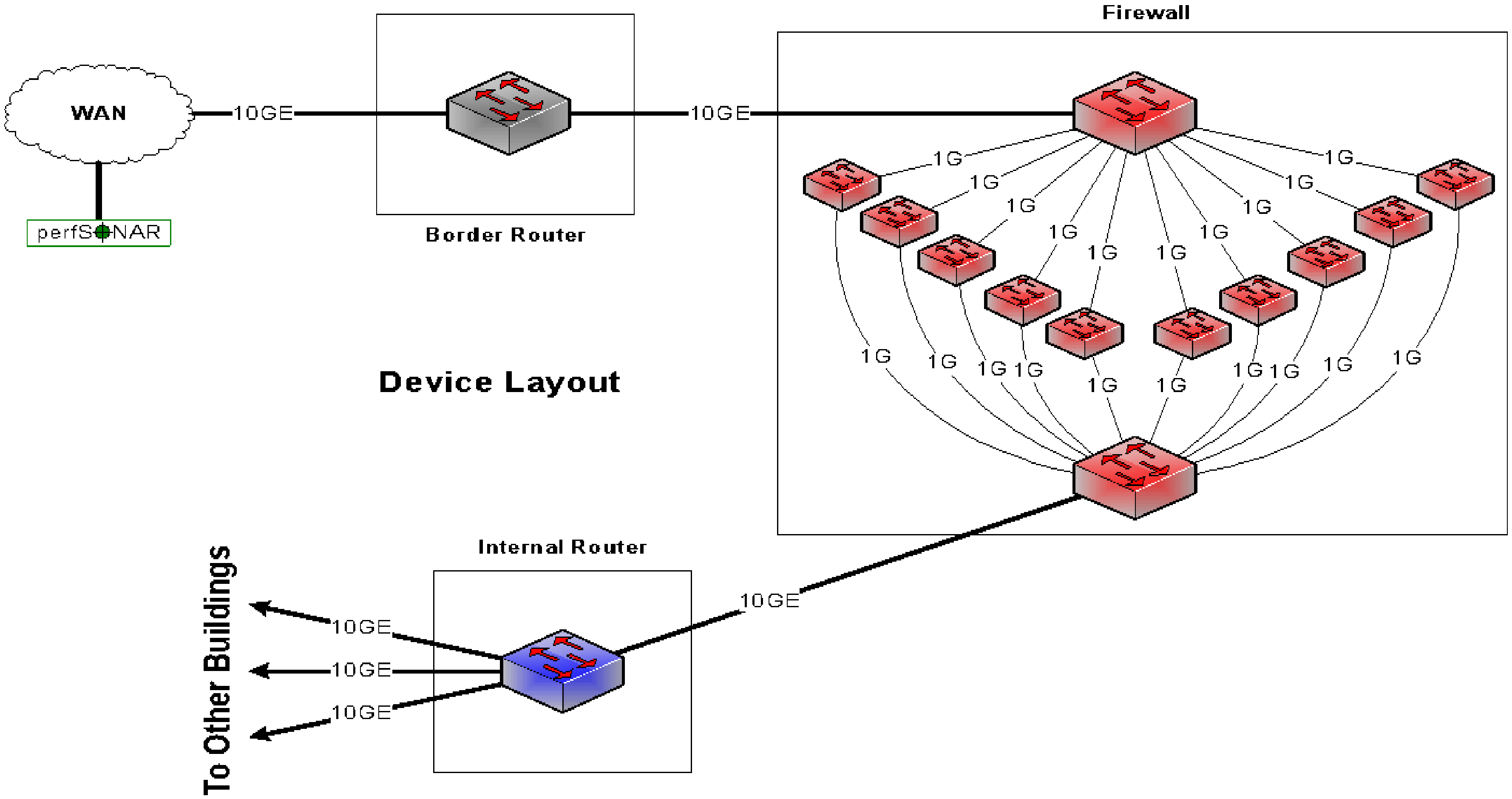
- **The big myth:** The main goal of the Science DMZ is to avoid firewalls and other security controls.
  - Leads to all sorts of odd (and wrong) claims like:
    - “Our whole backbone is a Science DMZ because there is no firewall in front of the backbone.”
    - “The Science DMZ doesn’t allow for **any** security controls.”
    - “The Science DMZ requires a default-permit policy.”

- **The reality:** The Science DMZ emphasizes reducing degrees-of-freedom, reducing the number of network devices (including middleboxes) in the path, eliminating devices that can’t perform, and ensuring that the devices that remain in the path are capable of large-scale data-transfer caliber performance.

From Myth to Reality

- Contrary to myth, the Science DMZ *is a security architecture.*
- The Science DMZ is a form of security *control,* not something to be controlled.
- At the same time, the Science DMZ enables us to do a better job of risk-based security through segmentation.
- 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***

How Do Firewall Appliances Work?



What is a Firewall?

NIST Answer (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”

What is a Firewall?

Vendor Answer

- Specific 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)

What is a Firewall?

Security Group Answer

- “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

Problems with Firewall Appliances

- 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 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 firewall?

Firewalls as Access Lists

- What does a firewall admin ask for when asked to allow data transfers?
    - IP address of your host
    - IP address of the remote host
    - Port range
    - ***That looks like an ACL to me – I can do that on the router***

- No special config for advanced protocol analysis – just address/port

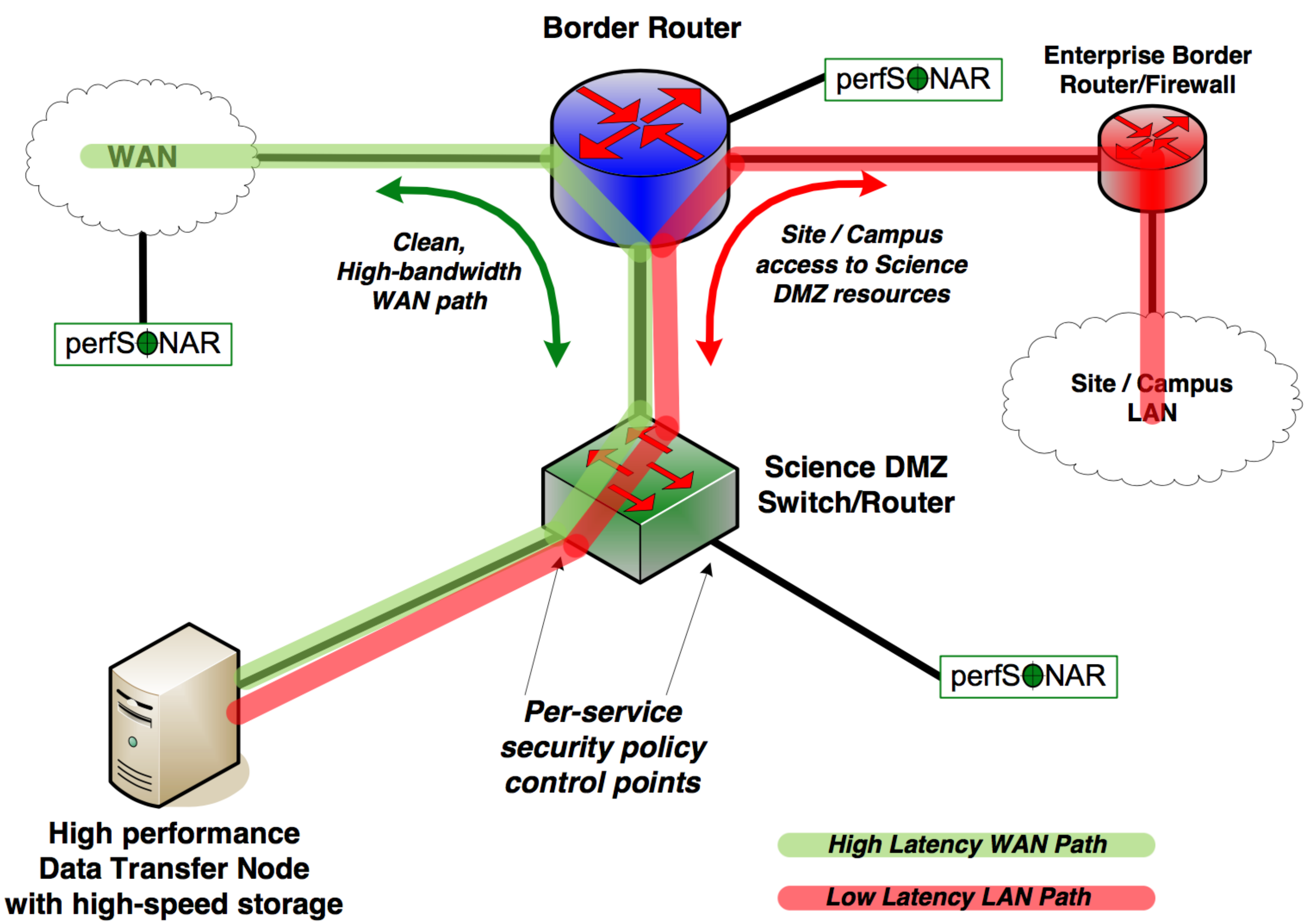
Security Without Enterprise Firewalls

- Data intensive science traffic interacts poorly with enterprise firewalls
- Does this mean we ignore security? ***NO!***
    - We **must** protect our systems
    - We 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***

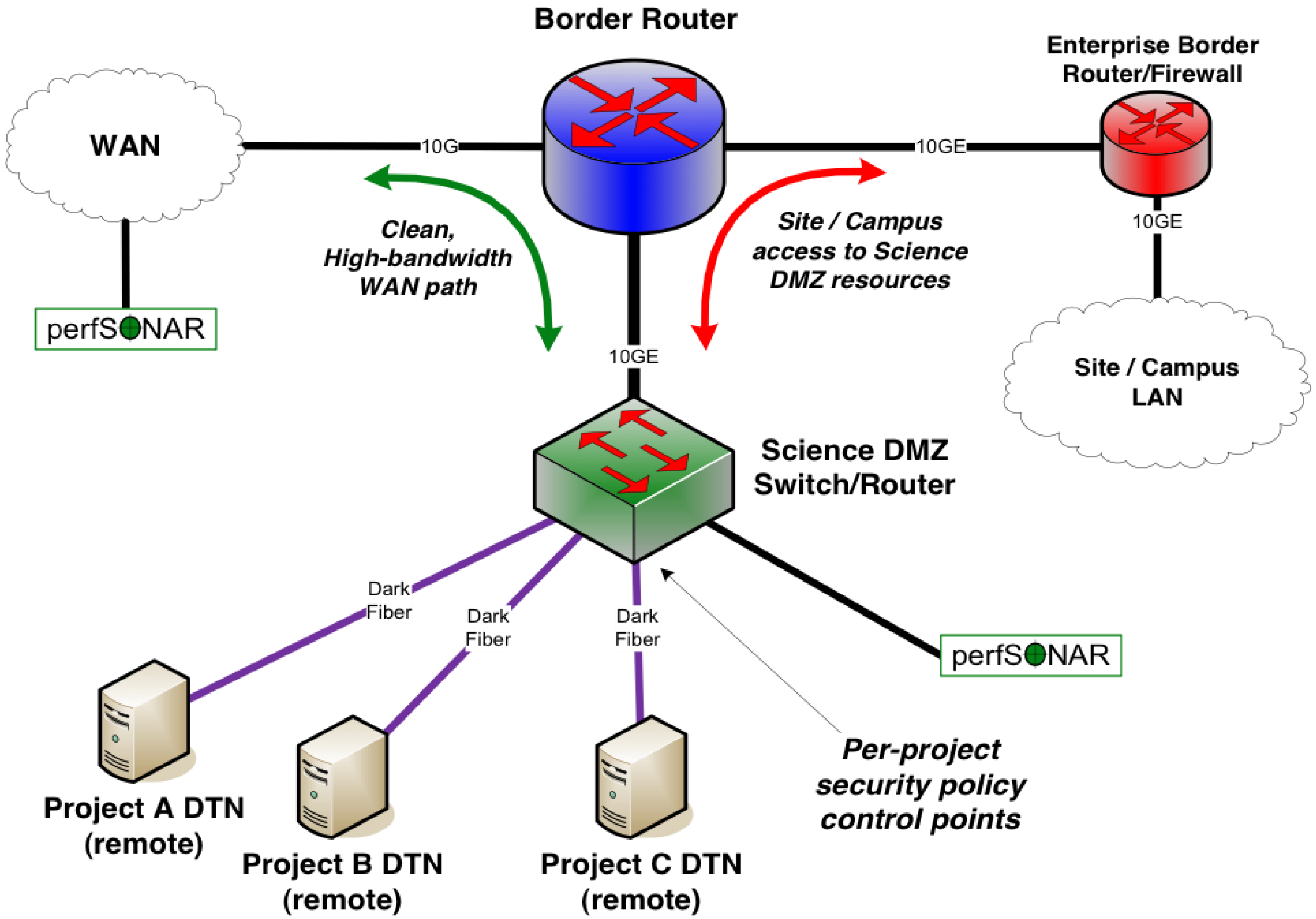
New and Emerging Firewall Designs

- Several organizations are working on ways to make firewalls better
- Some use SDN to dynamically switch approved flows around the firewall
- Some allow the firewall to control a switch directly
- Some vendors are now building firewalls to accommodate elephant flows
- ESnet hasn’t directly tested these approaches, though they look promising
- Some have significant cost

Science DMZ Example 1


Border Router
perfSONAR
Enterprise Border Router/Firewall
WAN
Clean, High-bandwidth WAN path
Site / Campus access to Science DMZ resources
perfSONAR
Site / Campus LAN
Science DMZ Switch/Router
perfSONAR
Per-service security policy control points
High performance Data Transfer Node with high-speed storage
High Latency WAN Path
Low Latency LAN Path

Science DMZ Example 2: Multiple Projects


Border Router
WAN
10G
10GE
Enterprise Border Router/Firewall
10GE
perfSONAR
Clean, High-bandwidth WAN path
Site / Campus access to Science DMZ resources
Site / Campus LAN
10GE
Science DMZ Switch/Router
Dark Fiber
Dark Fiber
Dark Fiber
perfSONAR
Per-project security policy control points
Project A DTN (remote)
Project B DTN (remote)
Project C DTN (remote)

Other Security Mechanisms: ACLs and Applications

- **Aggressive access lists**
  - More useful with project-specific DTNs
  - Exchanging data with a small set of remote collaborators = ACL is fairly easy to manage
  - Large-scale data distribution servers = difficult/time consuming to handle (but then, the firewall ruleset for such a service would be, too)
- **Limitation of the application set**
  - Makes it easier to protect
  - Keep unnecessary applications off the DTN (and watch for them anyway using a host IDS – take violations seriously)

Other Security Mechanisms: Network Monitor

- Network Security Monitors
  - One example is Bro – https://bro.org/
  - Bro is high-performance and battle-tested
    - Bro protects several high-performance national assets
    - Bro can be scaled with clustering: https://docs.zeek.org/en/stable/cluster/
  - Other IDS/NSM solutions also available

Other Security Mechanisms: Host IDS

- Using a Host IDS is recommended for hosts in a Science DMZ
- Several open source solutions exist:
    - OSSec: http://www.ossec.net/
    - Rkhunter: http://rkhunter.sourceforge.net (rootkit detection + FIM)
    - chkrootkit: http://chkrootkit.org/
    - Logcheck: http://logcheck.org (log monitoring)
    - Fail2ban: http://www.fail2ban.org/wiki/index.php/Main_Page
    - denyhosts: http://denyhosts.sourceforge.net/

Collaboration Within the Organization

- **All stakeholders should collaborate on Science DMZ design, policy, and enforcement**
- **The security people have to be on board**
  - Political cover for security officers
  - 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

Conclusions and Implications

- Think about what the Science DMZ is trying to do.
  - Improve performance, both by removing impediments and improving the performance of the devices that must be in line
  - Apply security policies appropriate for the data and the applications being protected
  - Ease troubleshooting
  - In general, reduce degrees of freedom from science networks to increase security flexibility/options
  - Maximize performance **and** security **and** resiliency

Wrapping Up

For More Information

ESNet

http://fasterdata.es.net/science-dmz/science-dmz-security/

NSRC

https://learn.nsrc.org/science-dmz/security

Trusted CI

https://trustedci.org/useful-links

Acknowledgments

Trusted CI is supported by the National Science Foundation under Grant ACI-1547272. The views expressed do not necessarily reflect the views of the National Science Foundation or any other organization.

Trusted CI activities are made possible thanks to the contributions of a multi-institutional team:

https://trustedci.org/who-we-are/

TRUSTED CI
THE NSF CYBERSECURITY
CENTER OF EXCELLENCE

Additional Acknowledgments

ESnet - Eli Dart, Michael Sinatra, Vlad Grigorescu, and Adam Slagell - for content and concepts throughout this talk.

pexels.com, images.iu.edu, Wikipedia Commons, IU IT Communications Office for images used in this presentation.

Trusted CI License Statement

TRUSTED CI
THE NSF CYBERSECURITY
CENTER OF EXCELLENCE