A Science DMZ for Data-intensive Research and Computation at the University of South Carolina

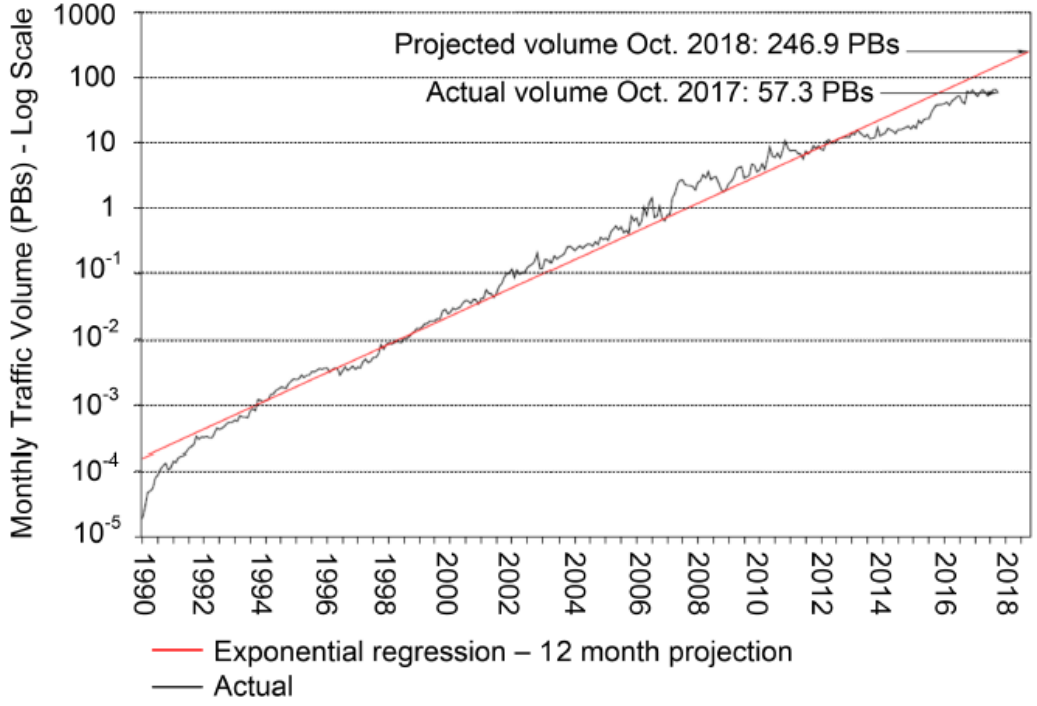
Jorge Crichigno (Presenter), Paul Sagona
College of Engineering and Computing
University of South Carolina
http://ce.sc.edu/cyberinfra

BSRA Computing Exchange
March 30, 2023
Online

Motivation for a High-Speed Science Architecture

UNIVERSITY OF SOUTH CAROLINA

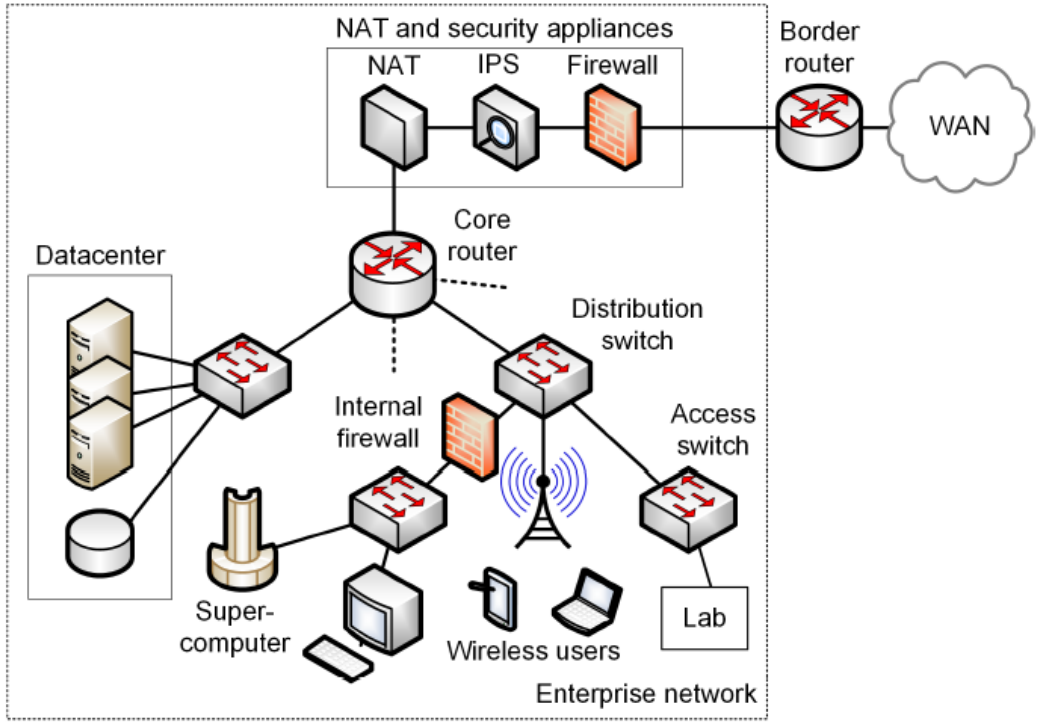
- Science and engineering applications are generating data at an unprecedented rate
- Instruments produce hundreds of terabytes in short time periods (“big science data”)
- Data must be typically transferred across high-bandwidth high-latency Wide Area Networks



The Energy Science Network (ESnet) is the backbone connecting U.S. national laboratories and research centers

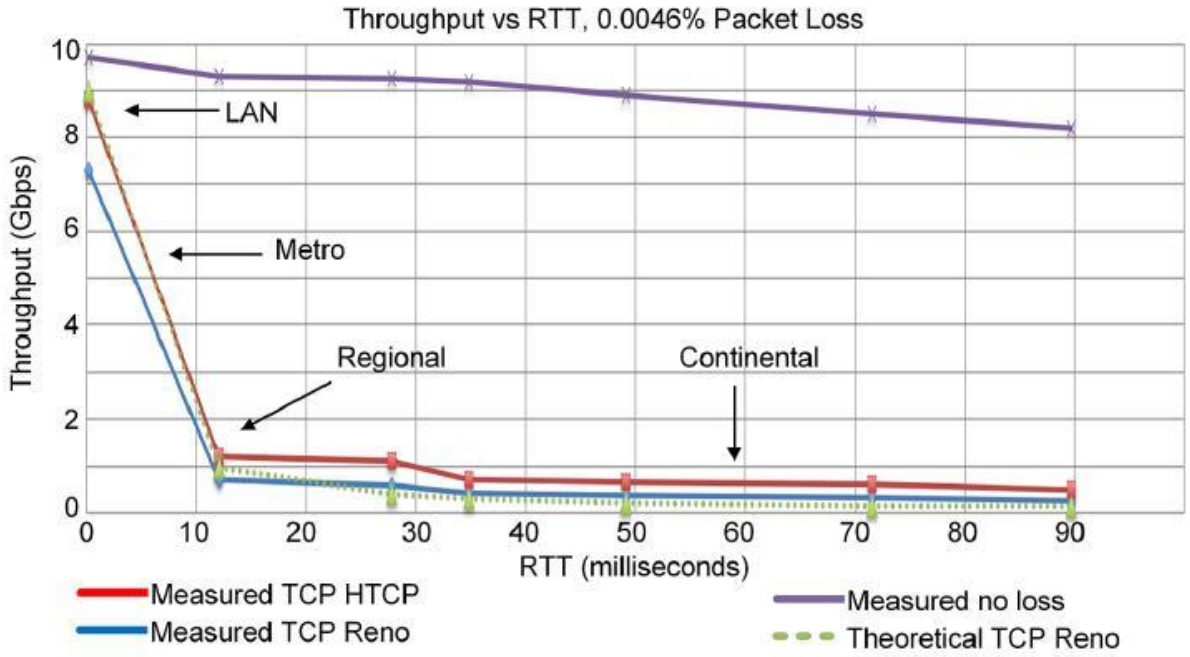
Enterprise Network Limitations

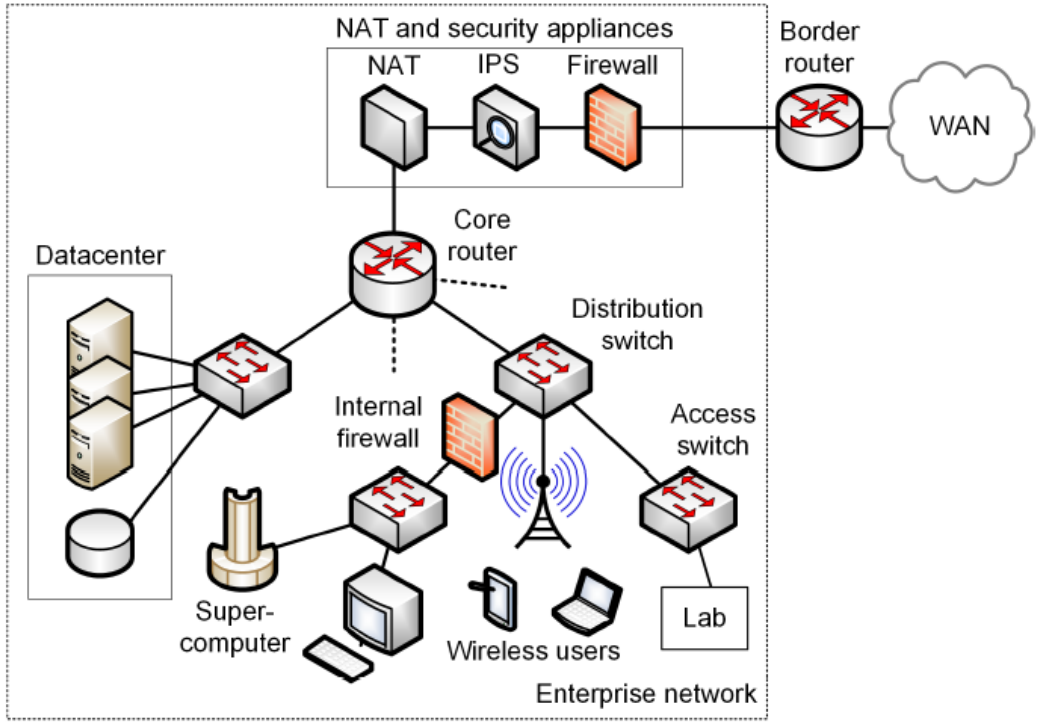
- Security appliances are CPU-intensive
- Inability of small-buffer switches to absorb bursts
- Lack of data transfer apps to exploit available bandwidth



Enterprise Network Limitations

- Security appliances are CPU-intensive
- Inability of small-buffer switches to absorb bursts
- Lack of data transfer apps to exploit available bandwidth

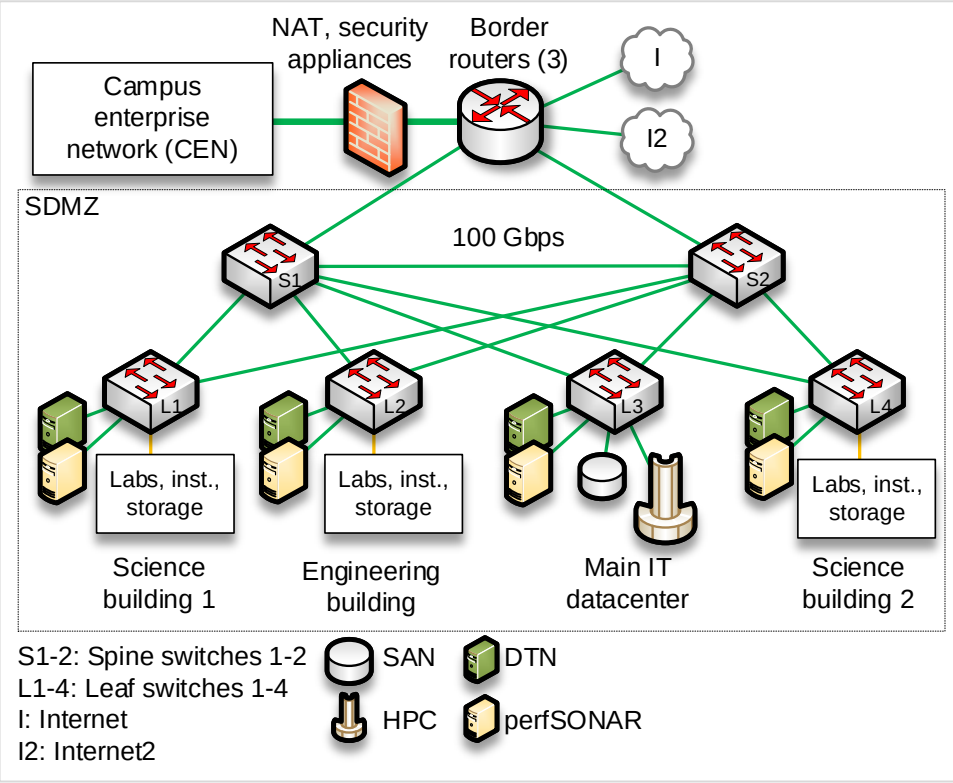


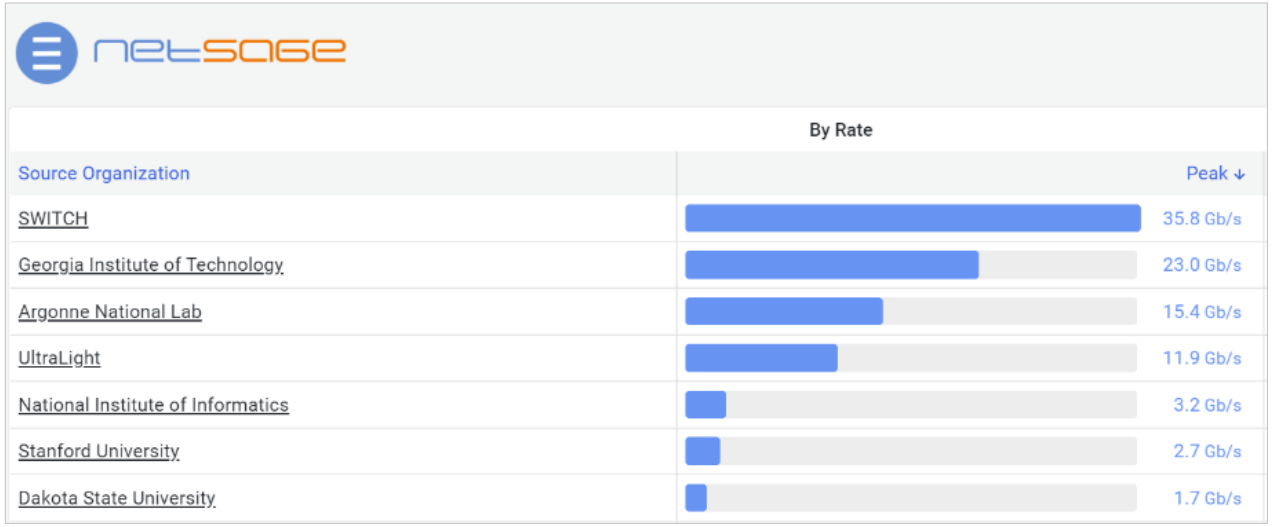


E. Dart, L. Rotman, B. Tierney, M. Hester, J. Zurawski, "The science dmz: a network design pattern for data-intensive science," *International Conference on High Performance Computing, Networking, Storage and Analysis*, Nov. 2013.

Solution

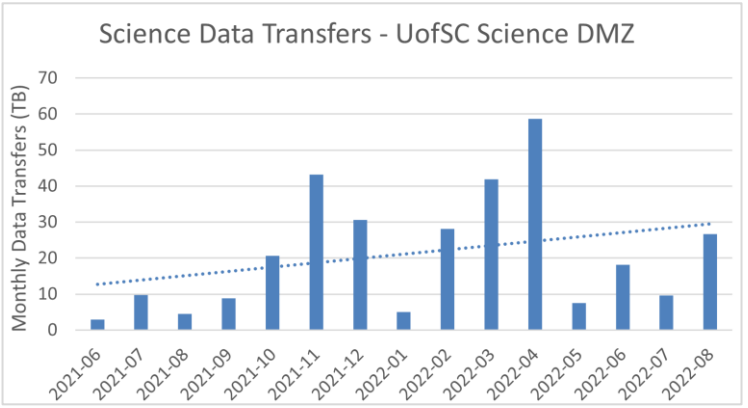
- USC applied to the NSF Campus Cyberinfrastructure program (2019) ($500K)
- It deployed a 100 Gbps Science DMZ (SDMZ), co-located to the campus network
- It increased the bandwidth to Internet2 from ~10 Gbps to 100 Gbps
- It increased bandwidth from key research laboratories on campus to Science DMZ

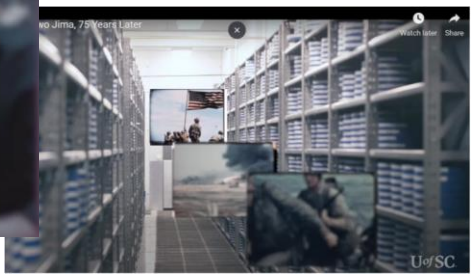




Outcomes

- Increased number of science data transfers
- Research activities on areas including cognitive processes, WWII film digitization, chemical engineering, nuclear physics, and others
- Campus infrastructure suitable for an R1 institution
  - 100 Gbps Science DMZ
  - 100 Gbps connection to Internet2
  - Direct connection to providers (AWS)
- Strengthened collaboration between IT and faculty
- Partnership with agencies and businesses
  - Training on high-speed networks, Science DMZs in collaboration with ESnet / LBNL
  - Internships





WWII film digitization
14,000+ cans of film

https://tinyurl.com/59cwtbs3

CIO SRNL, interns, PI

Outcomes

- Graduate research on cyberinfrastructure: conference papers, journal articles, and one book were published during the 3-year project
- New relations were established
  - Amazon – AWS “Direct Connect” to resources (via I2)
  - Barefoot Networks / Intel
  - Juniper Networks
  - VMware