Jason Zurawski
Senior Research Engineer, Internet2

Performance Working Group:

Firewalls: A Contrabulous Fabtraption That Embiggens Cromulent Networking

2013 INTERNET2 ANNUAL MEETING

BIG IDEAS. BIG COLLABORATION. BIG IMPACT.

Arlington, VA ▪ April 21–24

INTERNET2®

[Focused Technical Workshops]

BRINGING NETWORK EXPERTS TOGETHER—ONE TOPIC AT A TIME

Topic: Networking Issues for Life Sciences Research
July 17- 18, 2013
Lawrence Berkeley National Laboratory
Berkeley, California

- Building on the success of Joint Techs, meeting will bring together technical experts in a smaller setting with domain scientists.
- Workshop will include a slate of invited speakers and panels.
- Format to encourage lively, interactive discussions with the goal of developing a set of tangible next steps for supporting this data-intensive science community
- Four sub-topic areas: Network Architectures, Workflow Engines, Public and Private Cloud Architectures, and Data Movement Tools
- Website: http://goo.gl/v1YL3
- Proposals Due: May 17, 2013, 11:59 PDT

Firewalls: A Contrabulous Fabtraption That Embiggens Cromulent Networking

Contents

- ***State of the Campus***
- When Security and Performance Clash
- “The Science DMZ”, or “The Words You Will Hear 100s of Times This Week”
- Discussion

State of the Campus – A Word Of Caution…

- To be 100% clear – the firewall is a useful tool:
  - A layer or protection that is based on allowed, and disallowed, behaviors
  - One stop location to install instructions (vs. implementing in multiple locations)
  - Very necessary for things that need ‘assurance’ (e.g. student records, medical data, protecting the HVAC system, IP Phones, and printers from bad people, etc.)
- To be 100% clear again, the firewall delivers functionality that can be implemented in different ways:
  - Filtering ranges can be implemented via ACLs
  - Port/Host blocking can be done on a host by host basis
  - IDS tools can implement near real-time blocking of ongoing attacks that match heuristics

State of the Campus - Clarifications

- I am not here to make you throw away the Firewall
  - The firewall has a role; it's time to define what that role ***is***, and ***is not***
  - Policy may need to be altered (pull out the quill pens and parchment)
  - Minds may need to be changed

- I am here to make you think critically about campus security as a system. That requires:
  - Knowledge of the risks and mitigation strategies
  - Knowing what the components do, and do not do
  - Humans to implement and manage certain features – this may be a shock to some (lunch is never free)

State of the Campus – End Game

- The end goal is enabling true R&E use of the network
  - Most research use follows the ‘Elephant’ Pattern. You can’t stop the elephant and inspect it’s hooves without causing a backup at the door to the circus tent
  - Security and performance can work well together – it requires critical thought (read that as time, people, and perhaps money)
  - Easy economic observation – impacting your researchers with slower networks makes them less competitive, e.g. they are pulling in less research dollars vs. their peers

 – J. Zurawski zurawski@internet2.edu

Firewalls: A Contrabulous Fabtraption That Embiggens Cromulent Networking

Contents

- State of the Campus
- ***When Security and Performance Clash***
- "The Science DMZ", or "The Words You Will Hear 100s of Times This Week"
- Discussion

When Security and Performance Clash

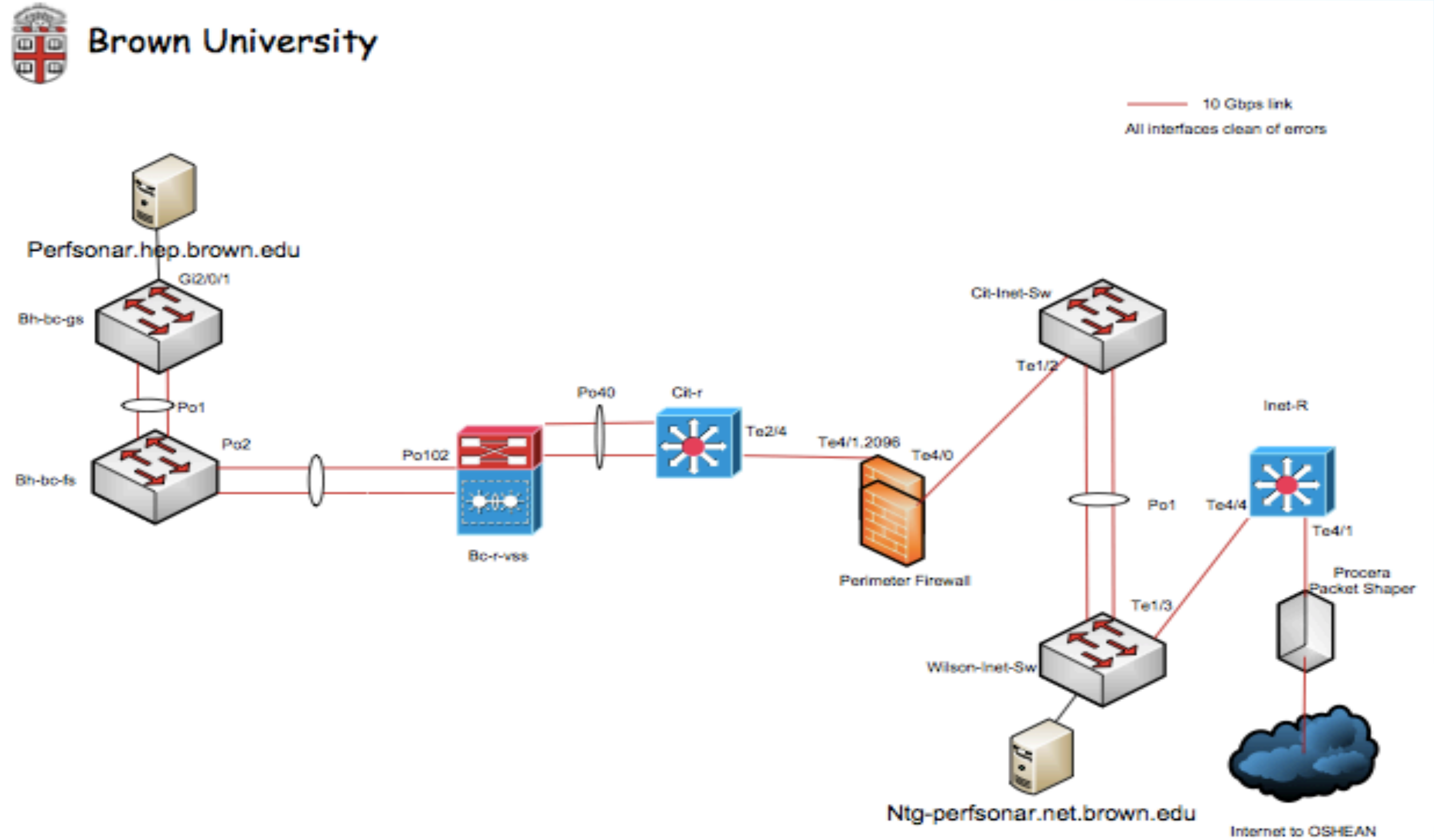
- What does a firewall do?
  - Streams of packets enter into an ingress port – there is some buffering
  - Packet headers are examined. Have I seen a packet like this before?
    - Yes – If I like it, let it through, if I didn't like it, goodbye.
    - No - Who sent this packet? Are they allowed to send me packets? What port did it come from, and what port does it want to go to?
  - Packet makes it through processing and switching fabric to some egress port. Sent on its way to the final destination.
- Where are the bottlenecks?
  - Ingress buffering – can we tune this? Will it support a 10G flow, let alone multiple 10G flows?
  - Processing speed – being able to verify quickly is good. Verifying slowly will make TCP sad
  - Switching fabric/egress ports. Not a huge concern, but these can drop packets too
  - Is the firewall instrumented to know how well it is doing? Could I ask it?

When Security and Performance Clash

- Lets look at two examples, that highlight two primary network architecture use cases:
  - Totally protected campus, with a border firewall
    - Central networking maintains the device, and protects all in/outbound traffic
    - Pro: end of the line customers don't need to worry (as much) about security
    - Con: end of the line customers ***must*** be sent through the disruptive device
  - Unprotected campus, protection is the job of network customers
    - Central networking gives you a wire and wishes you best of luck
    - Pro: nothing in the path to disrupt traffic, unless you put it there
    - Con: Security becomes an exercise that is implemented by all end customers

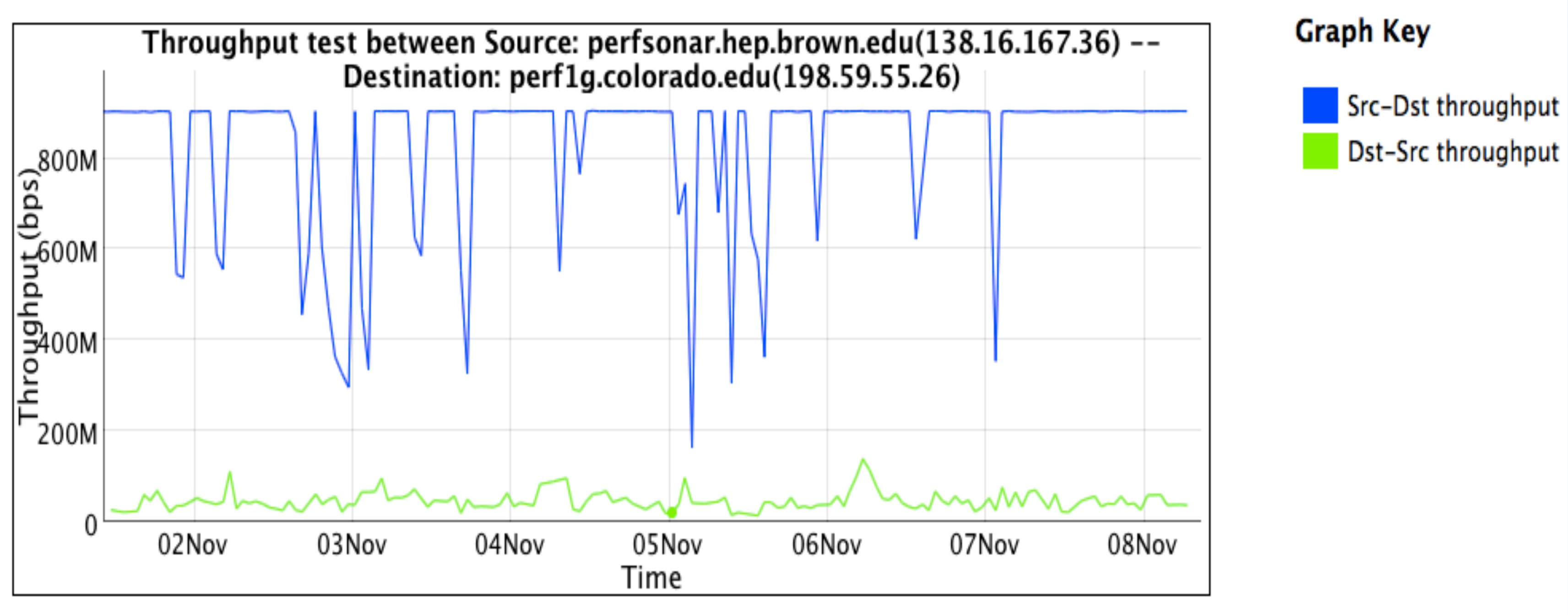
Brown University Example

- Totally protected campus, with a border firewall



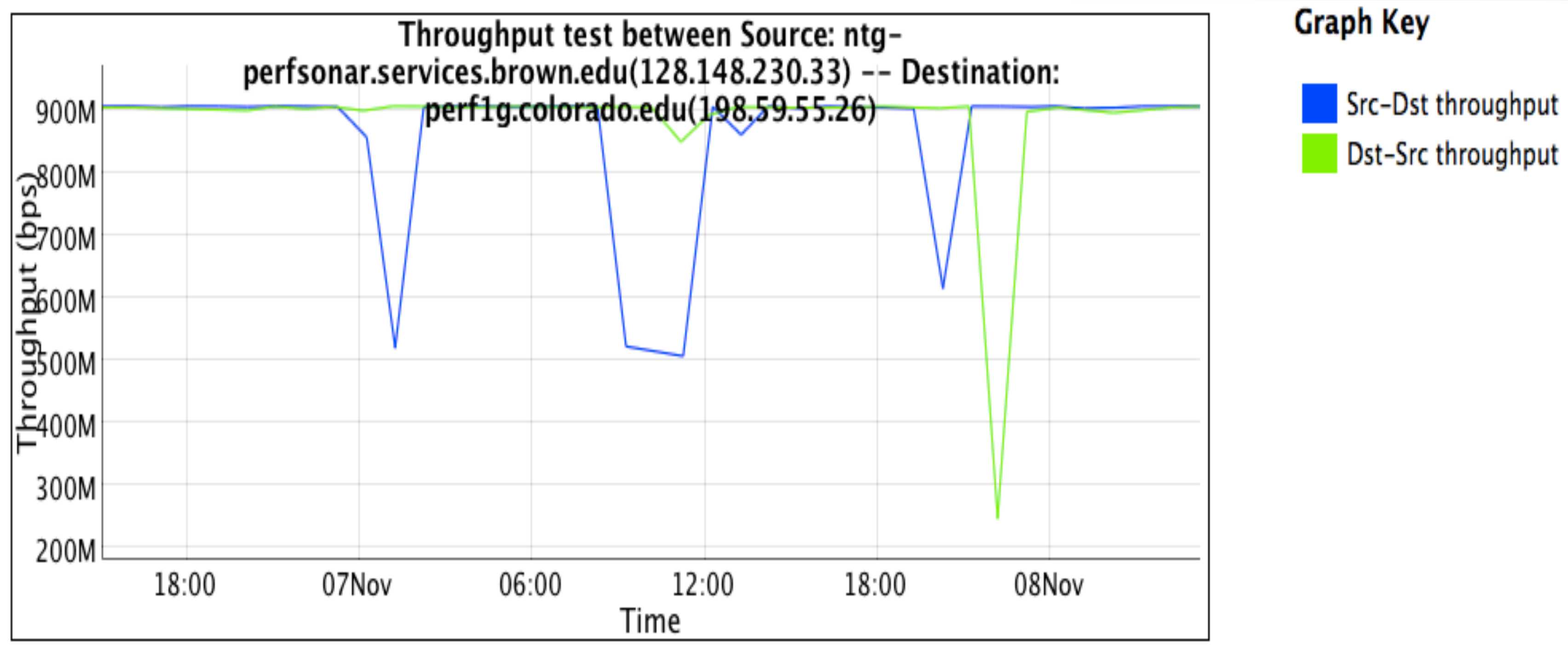
Brown University Example

- Behind the firewall:



Brown University Example

- In front of the firewall:



Brown University Example – TCP Dynamics

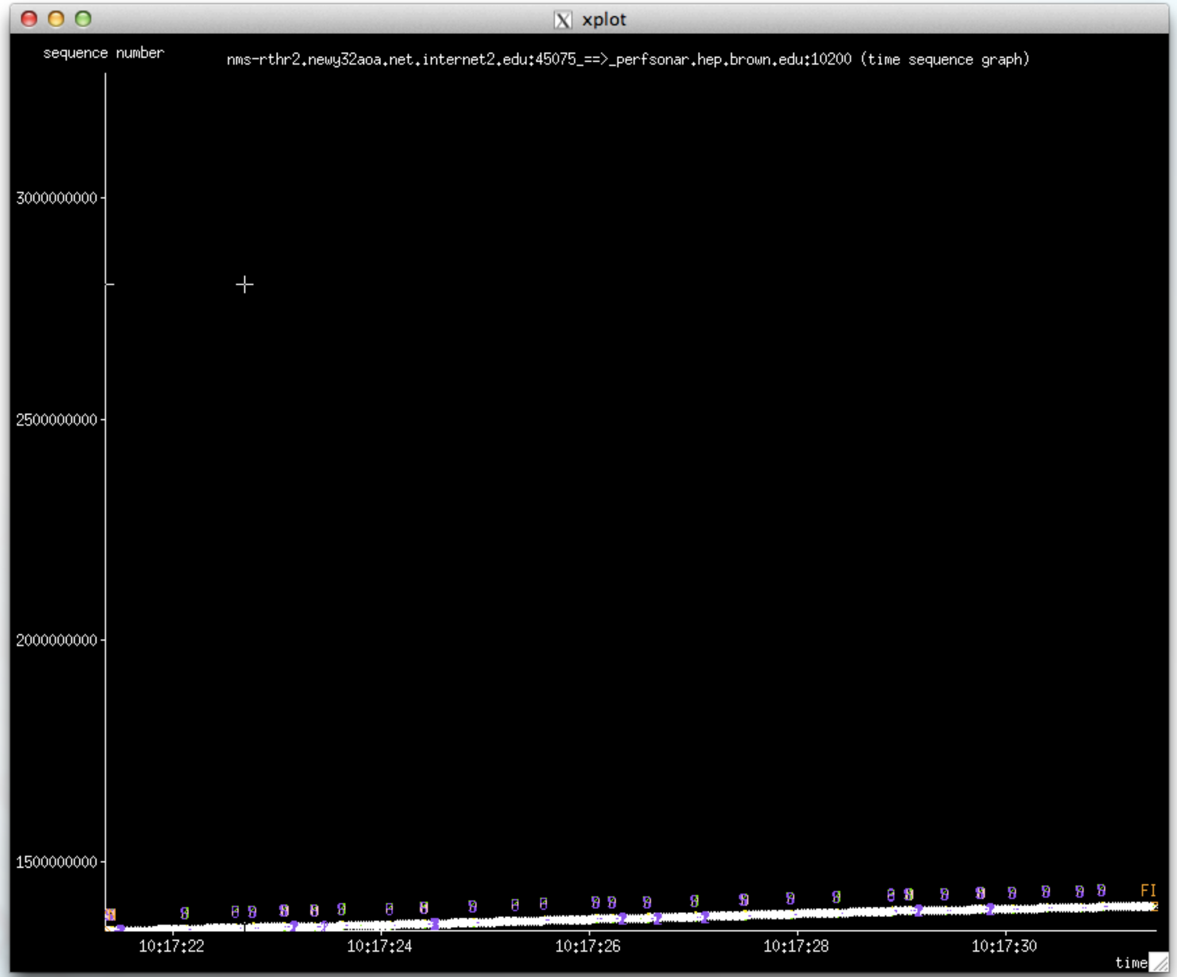
- Want more proof – lets look at a measurement tool through the firewall.
  - Measurement tools emulate a well behaved application
- 'Outbound', not filtered:
  - `nuttcp -T 10 -i 1 -p 10200 bwctl.newy.net.internet2.edu`
  - `92.3750 MB /   1.00 sec =  774.3069 Mbps     0 retrans`
  - `111.8750 MB /   1.00 sec =  938.2879 Mbps     0 retrans`
  - `111.8750 MB /   1.00 sec =  938.3019 Mbps     0 retrans`
  - `111.7500 MB /   1.00 sec =  938.1606 Mbps     0 retrans`
  - `111.8750 MB /   1.00 sec =  938.3198 Mbps     0 retrans`
  - `111.8750 MB /   1.00 sec =  938.2653 Mbps     0 retrans`
  - `111.8750 MB /   1.00 sec =  938.1931 Mbps     0 retrans`
  - `111.9375 MB /   1.00 sec =  938.4808 Mbps     0 retrans`
  - `111.6875 MB /   1.00 sec =  937.6941 Mbps     0 retrans`
  - `111.8750 MB /   1.00 sec =  938.3610 Mbps     0 retrans`

  - `1107.9867 MB /  10.13 sec =  917.2914 Mbps 13 %TX 11 %RX 0 retrans 8.38 msRTT`

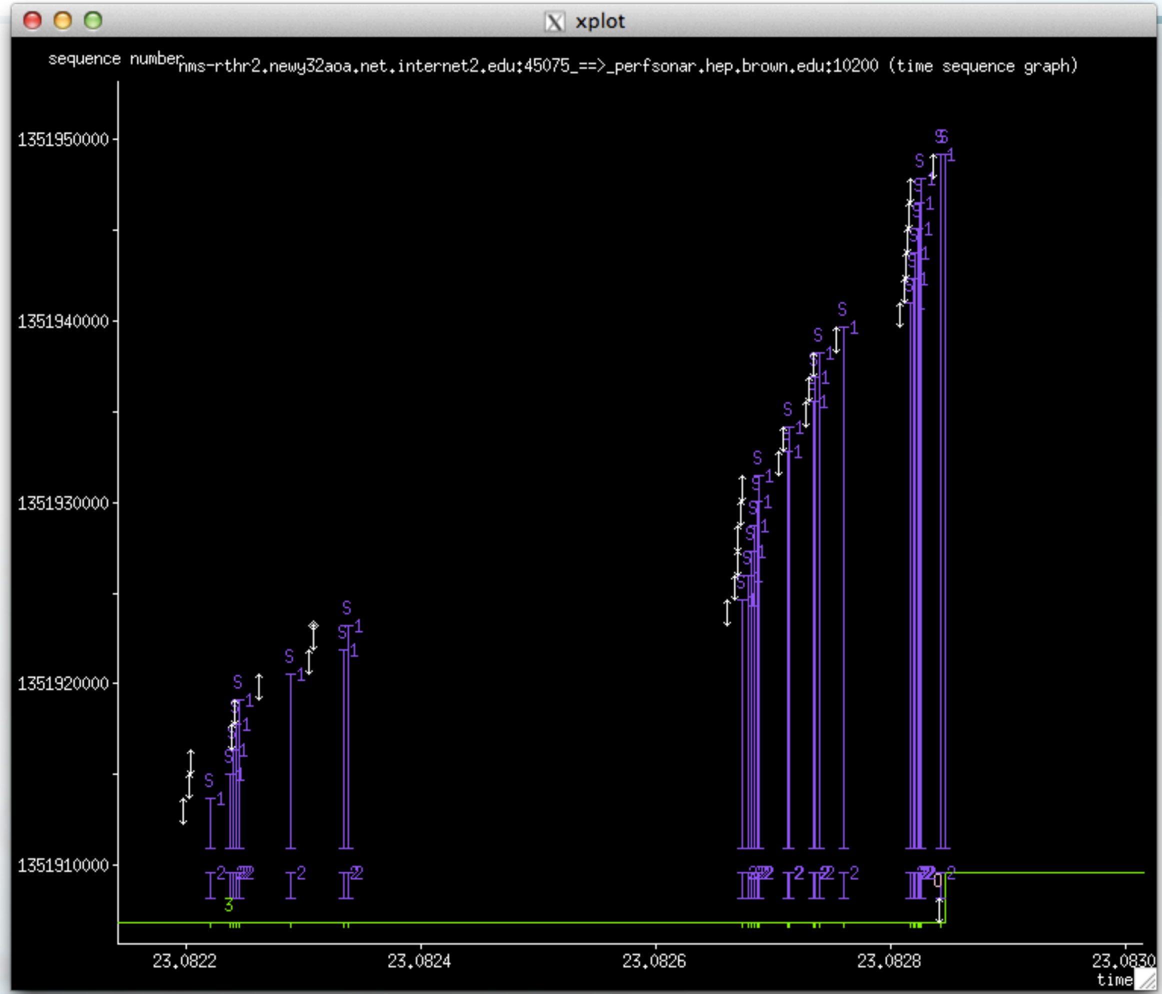
Brown University Example – TCP Dynamics

- ‘Inbound’, filtered:
  - `nuttcp -r -T 10 -i 1 -p 10200 bwctl.newy.net.internet2.edu`
  - `4.5625 MB /   1.00 sec =   38.1995 Mbps    13 retrans`
  - `4.8750 MB /   1.00 sec =   40.8956 Mbps     4 retrans`
  - `4.8750 MB /   1.00 sec =   40.8954 Mbps     6 retrans`
  - `6.4375 MB /   1.00 sec =   54.0024 Mbps     9 retrans`
  - `5.7500 MB /   1.00 sec =   48.2310 Mbps     8 retrans`
  - `5.8750 MB /   1.00 sec =   49.2880 Mbps     5 retrans`
  - `6.3125 MB /   1.00 sec =   52.9006 Mbps     3 retrans`
  - `5.3125 MB /   1.00 sec =   44.5653 Mbps     7 retrans`
  - `4.3125 MB /   1.00 sec =   36.2108 Mbps     7 retrans`
  - `5.1875 MB /   1.00 sec =   43.5186 Mbps     8 retrans`
  - `53.7519 MB /  10.07 sec =   44.7577 Mbps 0 %TX 1 %RX 70 retrans 8.29 msRTT`

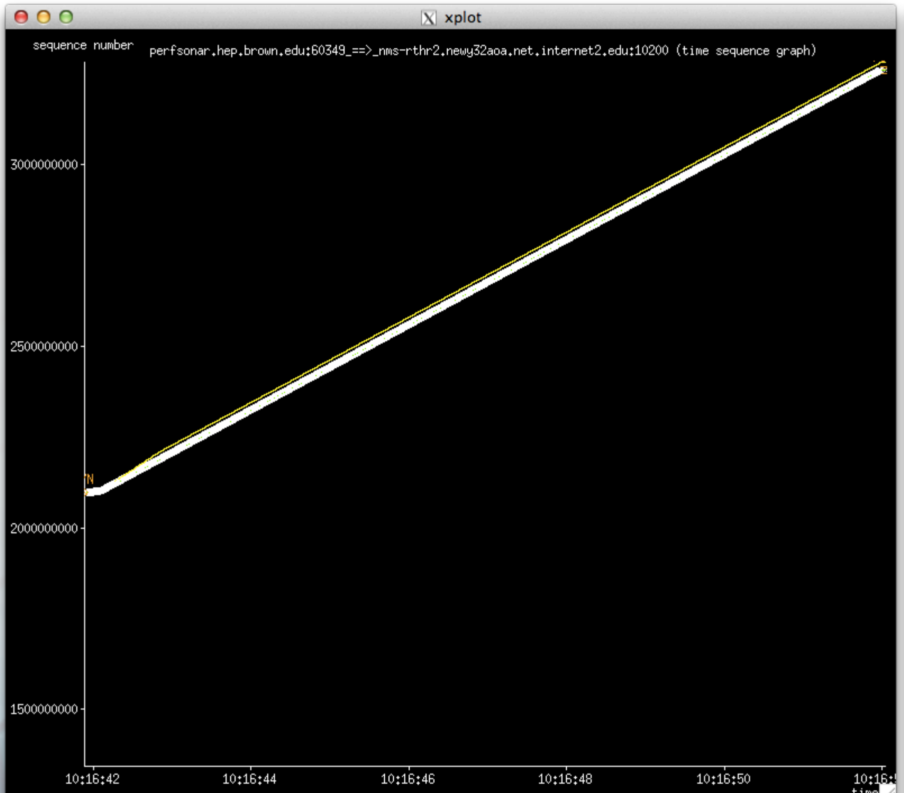
Brown University Example – TCP Plot (2nd)

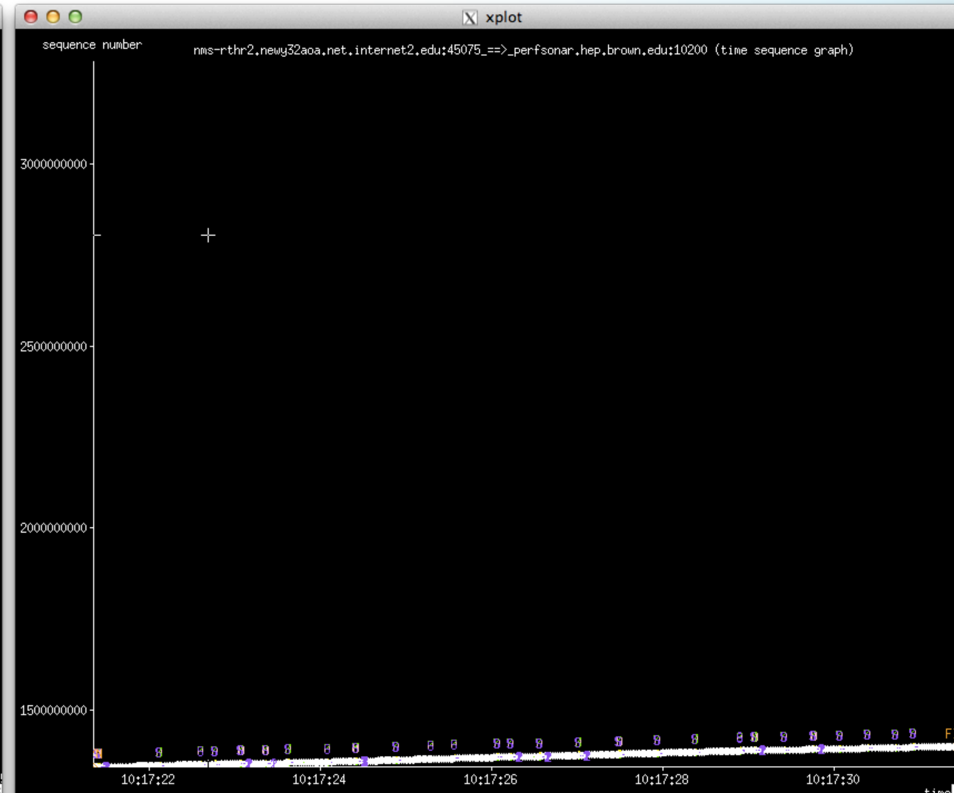


Brown University Example – TCP Plot (2nd)



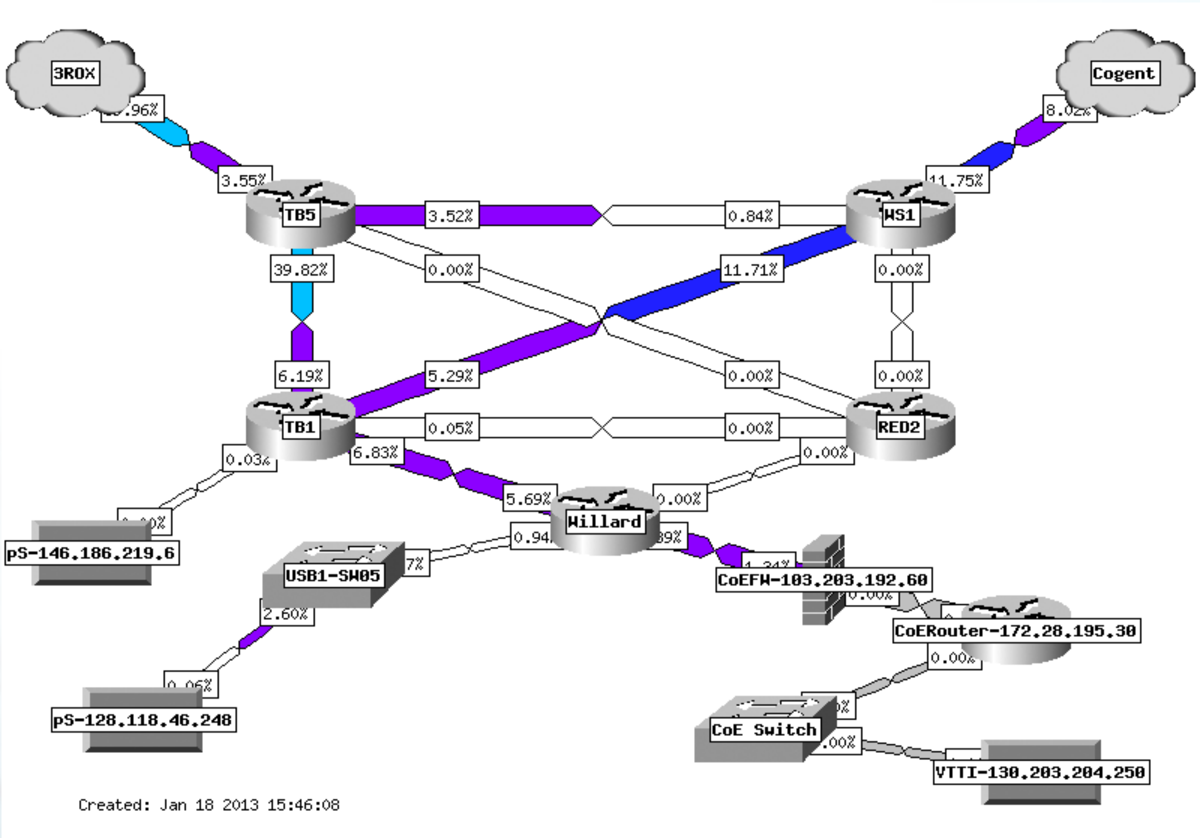
Brown University Example – Side By Side





The Pennsylvania State University Example

- Unprotected campus, protection is the job of network customers



The Pennsylvania State University Example

- Initial Report from network users: performance poor both directions
  - Outbound and inbound (normal issue is inbound through protection mechanisms)
- From previous diagram – CoE firewalll was tested
  - Machine outside/inside of firewall. Test to point 10ms away (Internet2 Washington)

```
jzurawski@ssstatecollege:~> nuttcp -T 30 -i 1 -p 5679 -P 5678 64.57.16.22
    5.8125 MB /   1.00 sec =   48.7565 Mbps     0 retrans
    6.1875 MB /   1.00 sec =   51.8886 Mbps     0 retrans
…
    6.1250 MB /   1.00 sec =   51.3957 Mbps     0 retrans
    6.1250 MB /   1.00 sec =   51.3927 Mbps     0 retrans

  184.3515 MB /  30.17 sec =   51.2573 Mbps 0 %TX 1 %RX 0 retrans 9.85 msRTT
```

The Pennsylvania State University Example

- Observation: `net.ipv4.tcp_window_scaling` did not seem to be working
  - 64K of buffer is default. Over a 10ms path, this means we can hope to see only 50Mbps of throughput:
  - ***BDP (50 Mbit/sec, 10.0 ms) = 0.06 Mbyte***

- Implication: something in the path was not respecting the specification in RFC 1323, and was not allowing TCP window to grow
  - TCP window of 64 KByte and RTT of ***1.0 ms*** <= ***500.00 Mbit/sec.***
  - TCP window of 64 KByte and RTT of ***5.0 ms*** <= ***100.00 Mbit/sec.***
  - TCP window of 64 KByte and RTT of ***10.0 ms*** <= ***50.00 Mbit/sec.***
  - TCP window of 64 KByte and RTT of ***50.0 ms*** <= ***10.00 Mbit/sec.***

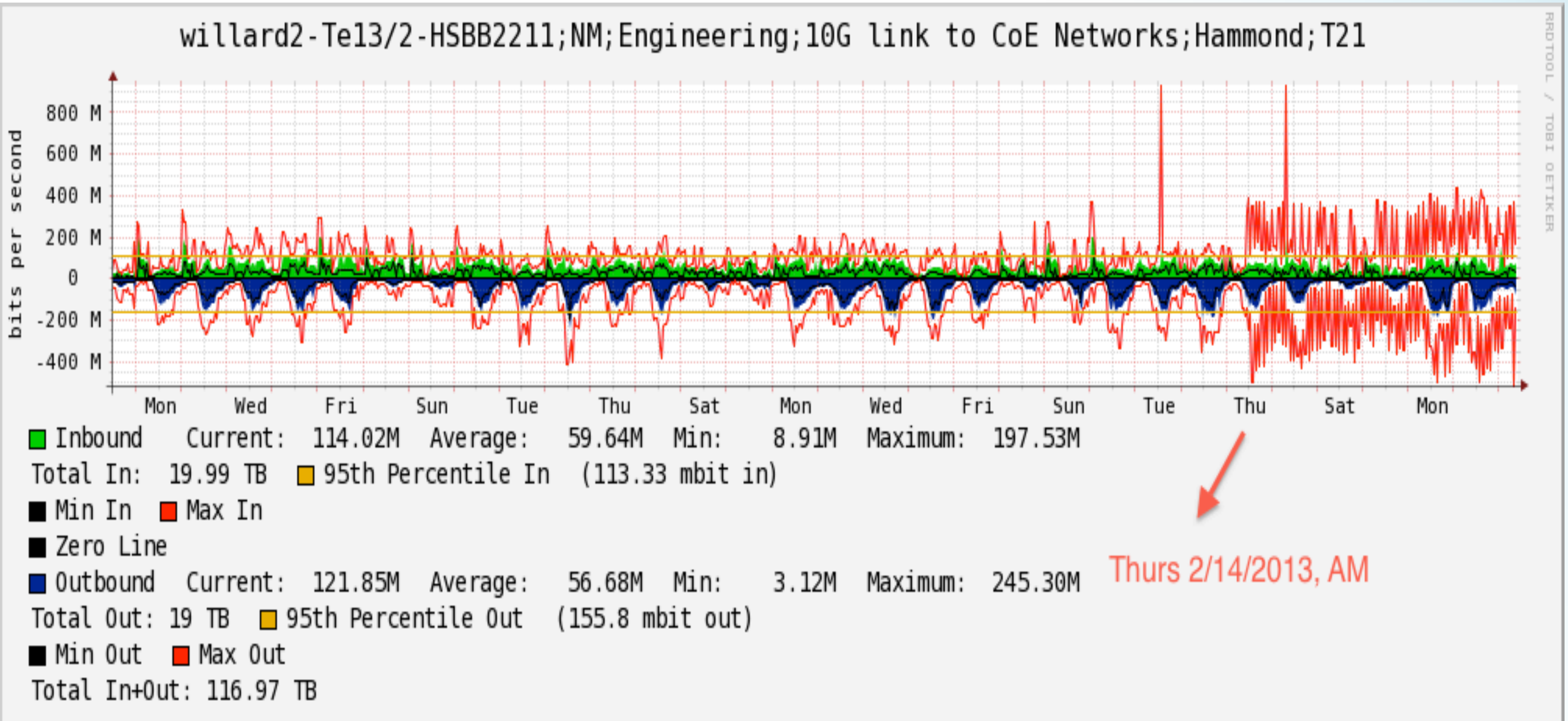
- Reading documentation for firewall:
  - ***TCP flow sequence checking*** was enabled
  - What would happen if this was turn off (both directions?

The Pennsylvania State University Example

- `jzurawski@ssstatecollege:~> nuttcp -T 30 -i 1 -p 5679 -P 5678 64.57.16.22`
- `   55.6875 MB /   1.00 sec =  467.0481 Mbps     0 retrans`
- `   74.3750 MB /   1.00 sec =  623.5704 Mbps     0 retrans`
- `   87.4375 MB /   1.00 sec =  733.4004 Mbps     0 retrans`
- …
- `   91.7500 MB /   1.00 sec =  770.0544 Mbps     0 retrans`
- `   88.6875 MB /   1.00 sec =  743.5676 Mbps    28 retrans`
- `   69.0625 MB /   1.00 sec =  578.9509 Mbps     0 retrans`
-
- ` 2300.8495 MB /  30.17 sec =  639.7338 Mbps 4 %TX 17 %RX 730 retrans 9.88 msRTT`

The Pennsylvania State University Example

- Impacting real users:



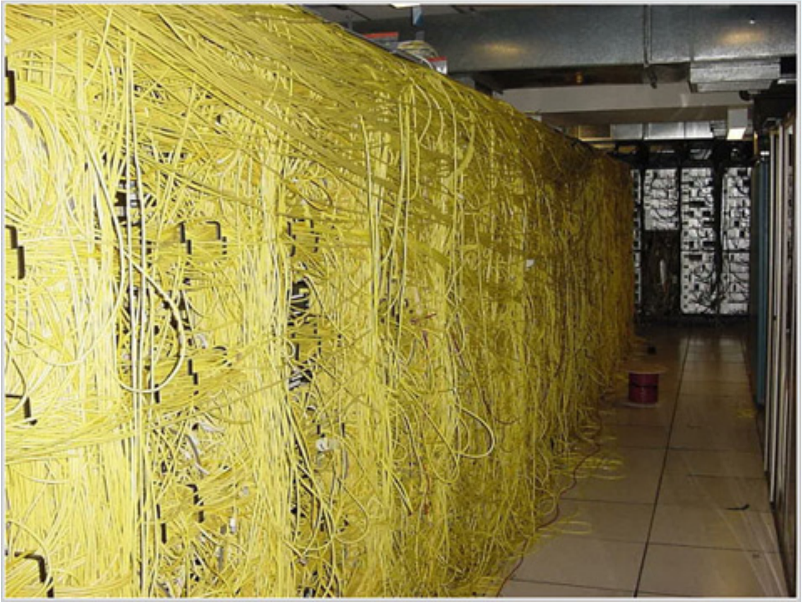
Firewalls: A Contrabulous Fabtraption That Embiggens Cromulent Networking

Contents

- State of the Campus
- When Security and Performance Clash
- ***"The Science DMZ", or "The Words You Will Hear 100s of Times This Week"***
- Discussion

Science DMZ (?)

- A staple of the meeting circuit for several years
- What is it really?
  - “Blueprint”, not a specific design
  - Approach to network architecture that preserves the ability to securely manage two different worlds
    - Enterprise – BYOD, IP Phones, Printers, HVAC, things you don’t know enough about to trust, and shouldn’t
    - Research – Well defined access patterns, Elephant flows, (normally) individuals that can manage their destiny with regards to data protection

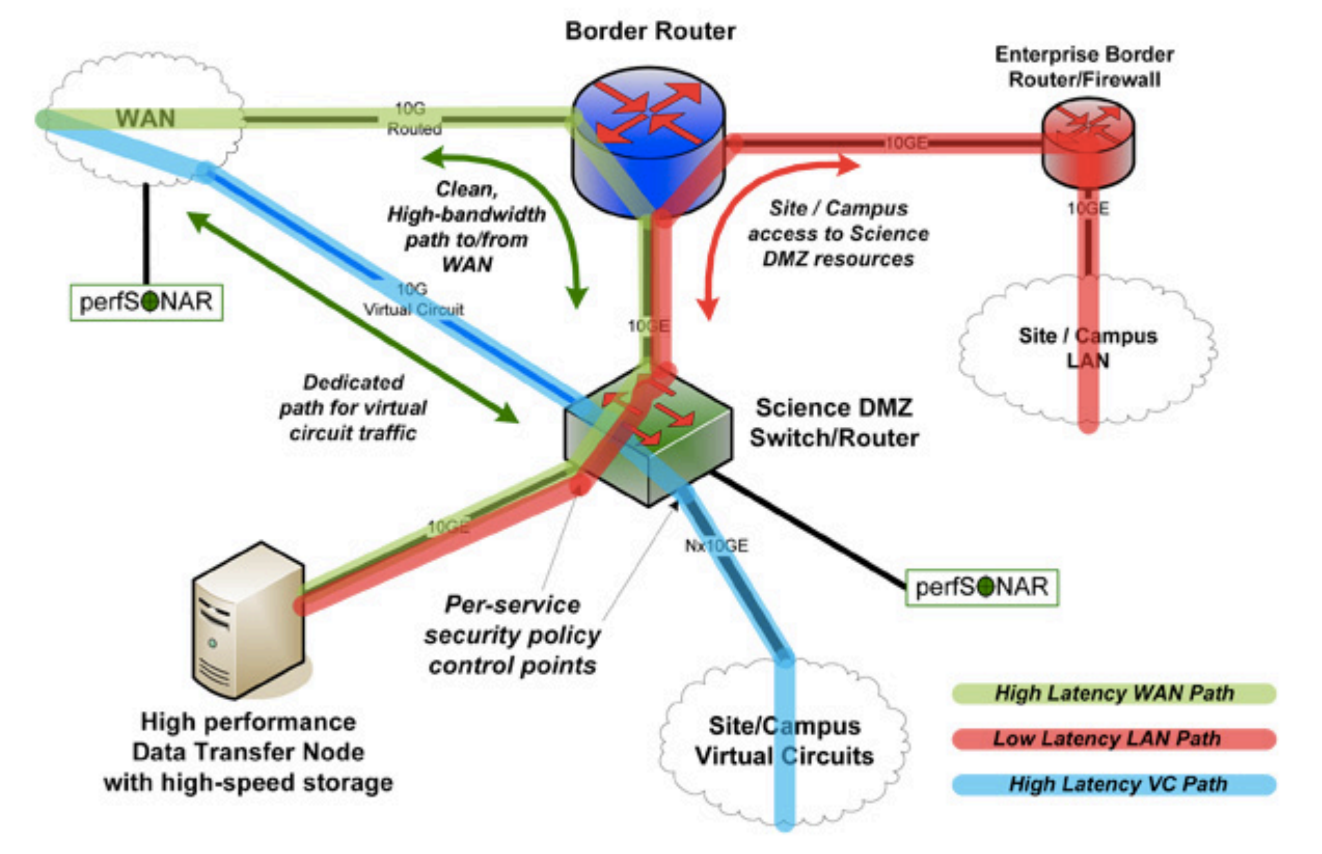


Science DMZ – Pro/Con on Generalities

- Pro:
  - Unspecified nature makes the pattern fungible for anyone to implement
  - Hits the major requirements for major science use cases
  - A concept that “anyone” should be able to understand on a high level
- Con:
  - Unspecified nature implies you need your own smart person to think critically, and implement it for a specific instantiation
  - Those that don’t do heavy science (or don’t know they do) may feel “its not for us”
  - A concept easy to treat as a ‘checkbox’ (hint: CC-NIE schools – are you stating ‘we have perfSONAR’ and moving on?)

When Rubber Meets the Road

- Lets start with the generic diagram:



When Rubber Meets the Road

- There are 4 areas I am going to hit on, briefly (note the last one is not ‘pictured’):
  - Network Path
  - Adoption of “New” Technology
  - Security
  - User Outreach

Network Path

- Engineers 'get it'
  - No one will dispute that protected and unprotected path will have benefits (and certain dangers).
  - ***$***, 100G isn't cheap (10 and 40 are). You don't ***have*** to go 100, implementing the architecture with existing technology is a perfectly good way forward
  - You still need a security professional (if you don't have one already) for the secured and non-secured paths. Learn to love your IDS just as much as your firewall and shapper …
- Tuning is important. Small buffers (as seen previously) make data movement sad. This means servers, and network devices
- Ounce of prevention – you need monitoring, and you certainly need training in how to use the performance tools to debug. You will be debugging (bet me a $1 if you honestly think you won't be…)

Adoption of "New" Technology

- SDN, perfSONAR, etc. etc.
  - We will keep making acronyms, don't worry
- What matters in all this? Being able to make your job easier
  - perfSONAR = insurance policy against risky behavior.
    - Will tell you if you have done things wrong, and warn you if something breaks.
    - Crucial for your campus, and costs only the price of a server, and getting an engineer up to speed on how to use it
  - SDN will be a game changer. Is it ready for production (?) – hard to say. The ability to afford more control over the network to the end user relies on applications (and end users) getting caught up. Hint.
- There will be more changes in the future, it's the nature of the game. R&E needs to be about certain risky moves away from the norm

Security

- I can spend an entire deck on this, but to keep it short:
  - Component based security is wrong. Needs to be a system.
  - System:
    - Cryptography to protect user access and data integrity
    - IDS to monitor before (and after) events
    - Host-based security is better for performance, but takes longer to implement. Firewalls are bad on performance but easy to plot down in a network. Attack vector from the “inside” is prevented.
    - Let your router help you – if you know communication patterns (and know those that should be disallowed), why not use filters?
  - Campus CI Plan. Make one, update it often. Shows funding bodies you know what is going on and have plans to address risks, and foster growth
- Economic argument – if you are non-competitive for grants because you cheaped out on security, are you better in the long run?

User Outreach

- The unstated factor:
  - Could you name your top 10 (5? 3?) network users? Do you know where their traffic is going? Do you know why? Should you care?
  - Simple solution – (net | s)flow monitoring (pick a brand, many are good).
    - Top 10 src/dst for some period of time, go and talk to the researchers.
    - Ask them what they are doing, how they are doing it, and if its going ok.
  - Campus CI days – was a sponsored thing, but why not have one ‘just because’?
    - Gets IT and research talking.
    - Identifies areas of growth; areas of friction
  - Requires an outgoing person – hire a research engineer.
    - Someone who knows what a network is, and can translate statements like “the beamline will be firing at 200Khz 2 times a week and generating 2PB of data a year” into “they need 40Gbps and a clear path to 4 international sites as well as the domestic routing table”

Firewalls: A Contrabulous Fabtraption That Embiggens Cromulent Networking

Contents

- State of the Campus
- When Security and Performance Clash
- "The Science DMZ", or "The Words You Will Hear 100s of Times This Week"
- ***Discussion***

?

Firewalls: A Contrabulous Fabtraption That Embiggens Cromulent Networking

Jason Zurawski – zurawski@internet2.edu
Senior Research Engineer, Internet2

http://www.internet2.edu/research

2013 INTERNET2 ANNUAL MEETING

BIG IDEAS. BIG COLLABORATION. BIG IMPACT.

Arlington, VA ▪ April 21–24