MOREnet

Be better connected.

Performance and Security

Routing safely at high speeds

Shannon Spurling

MOREnet- Core Network Engineer

Basic Overview

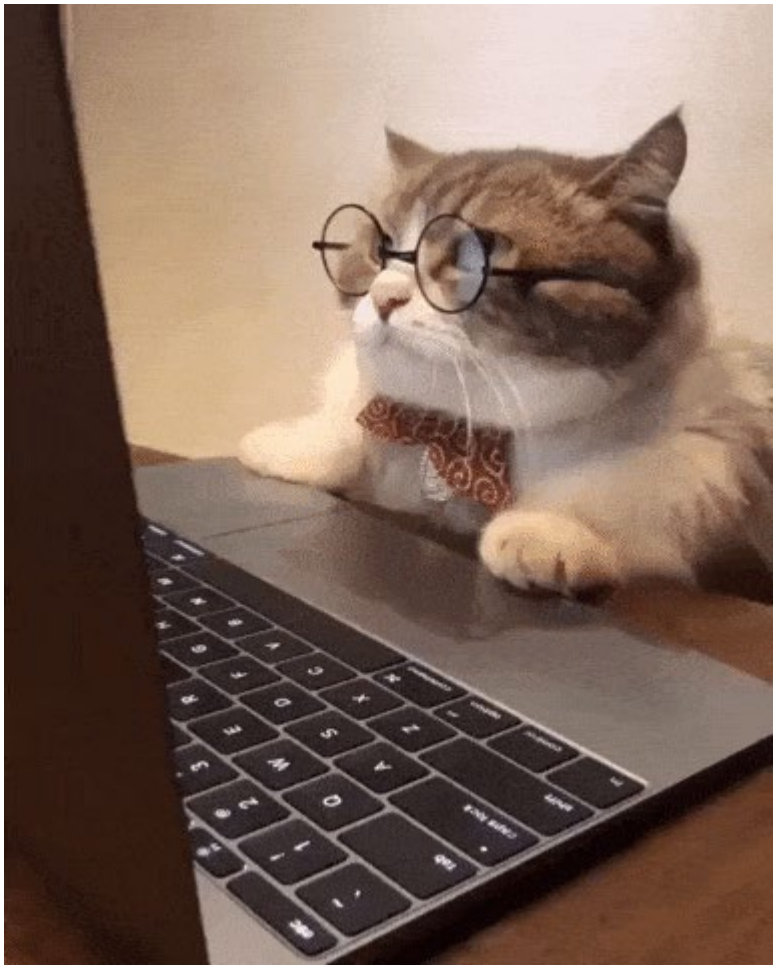
- Define some parts of the network
- Understand how they work
- Performance from a network view
- Security from a network centric view
- How to get it all to work together

Disclaimer…

These are my own biases and observations and do not reflect policy, thoughts, or direction of the other departments or people at MOREnet or the University of Missouri.

My goal is to get you to think about these things….

Gratuitous Cat gif …



Be better connected.

Managing the complexity

Break things down into their components and understand how they work together

Simplify and document how the system is built

Doing the footwork beforehand saves a lot of trouble on the backside

Don’t think you can scam the universe…

The layer models

OSI model

- Physical
- Data Link
- Network
- Transport
- Session
- Presentation
- Application

TCP/IP model

- Physical
- Network
- Internet
- Transport
- Application

But, my layers don't match those…

Transport and tunneling protocols can sometimes lead to confusion because they don't match traditional models…

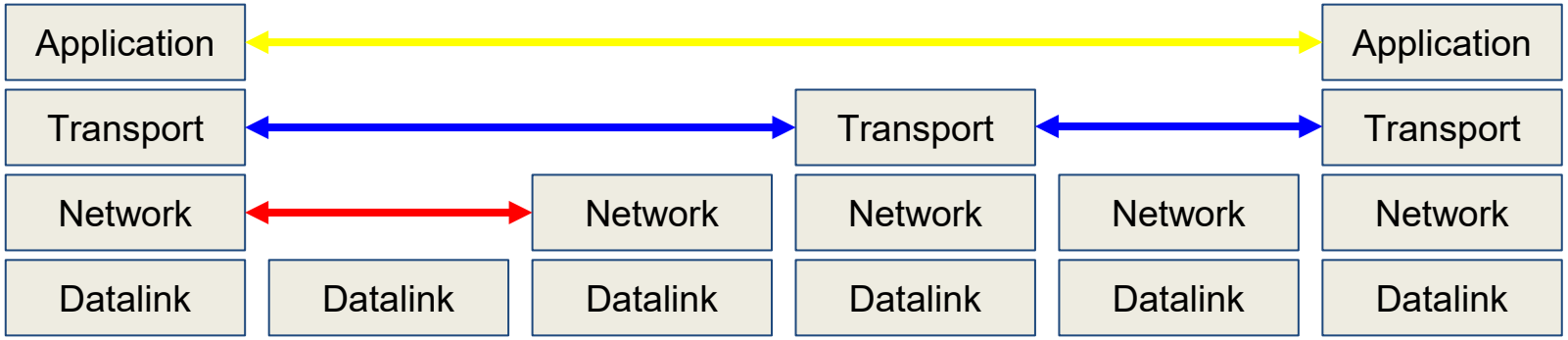
e.g. ATM, MPLS, VXLAN, IPsec, SSL-VPN, Segment Routing/SRV6…

They all still function as a layer between two other network functions.

What the layers really mean

Layers are more of a methodology for connecting things than a rigid structure everything has to fit into

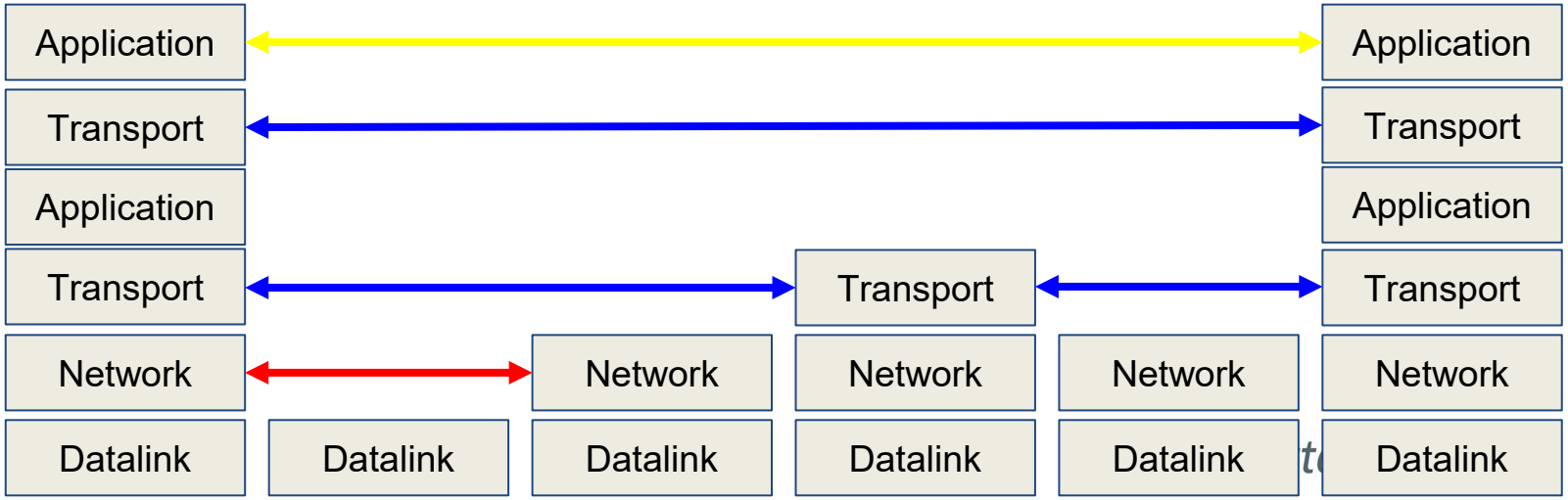
Each layer is a protocol
(connectors, signals, framing, behaviors)



What the layers really mean

Layers are more of a methodology for connecting things than a rigid structure everything has to fit into

Each layer is a protocol
(connectors, signals, framing, behaviors)



Layer 2 - The Network Layer

Normally Ethernet (802.3) these days…

Non-routable.

…means it's localized to an organization/data center or function.

Router required to connect LANs to each other.

Addresses can be globally assigned, or locally assigned depending on protocol and implementation

Tags can be used to segregate LANs within switches. (802.1q or 802.1ad)

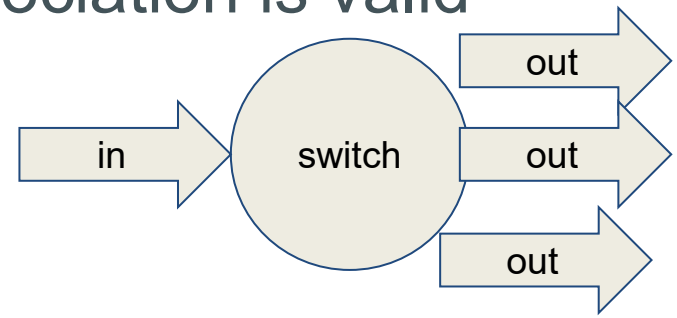
How a switch functions….

Store and forward (hub)

1. Receive Frame on port
2. Forward frame out all other ports

Learning Switch

1. Frame is received on port
2. source address (and vlan's) for frame received is stored as reachable via inbound port (with a timer)
3. switch forwards it out all ports that are associated with the destination address (and vlans), or floods it if no association is valid



A few details on layer 2/Ethernet

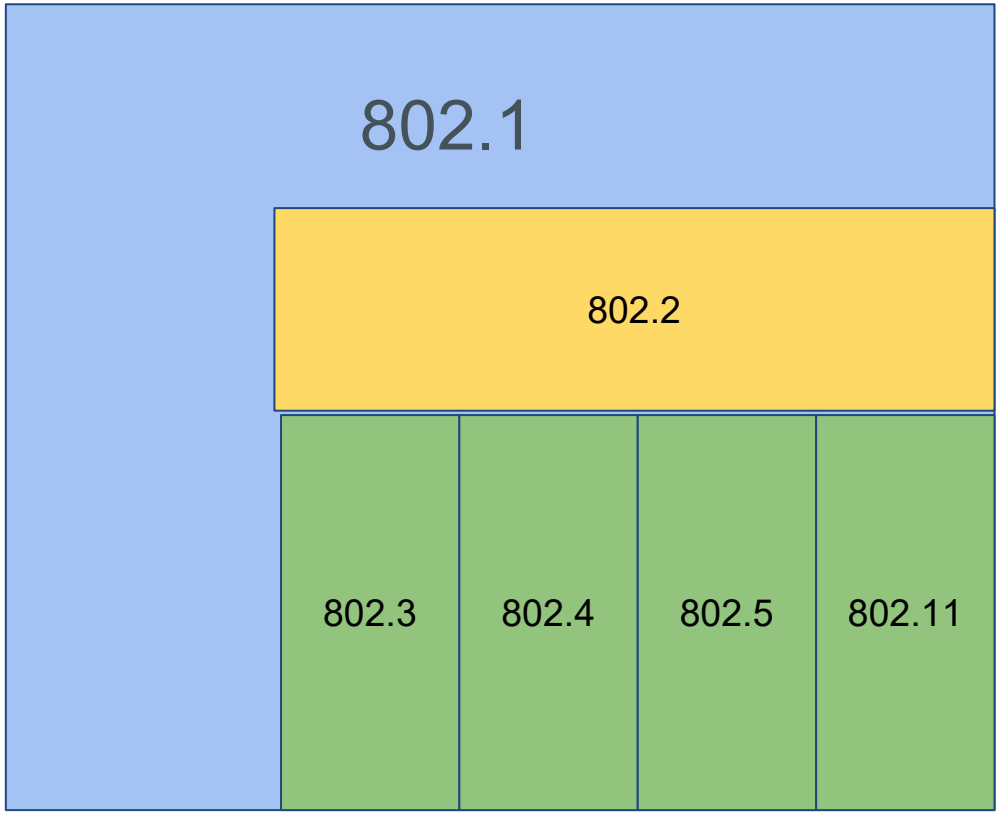
- 2 components to an Ethernet address
  - OUI:ID
  - OUI is assigned by IEEE to vendors
  - ID is assigned by vendors to NIC's
- MAC addresses are not of use outside of a layer 2 domain
- VLAN's group endpoints (802.1q)
- VLAN's can group VLAN's (802.1ad)
- LAN protocols can be chatty…
- Loop detection and prevention (spanning tree) can be complex, slow, and fragile.
  - Special protocols developed to solve this issue

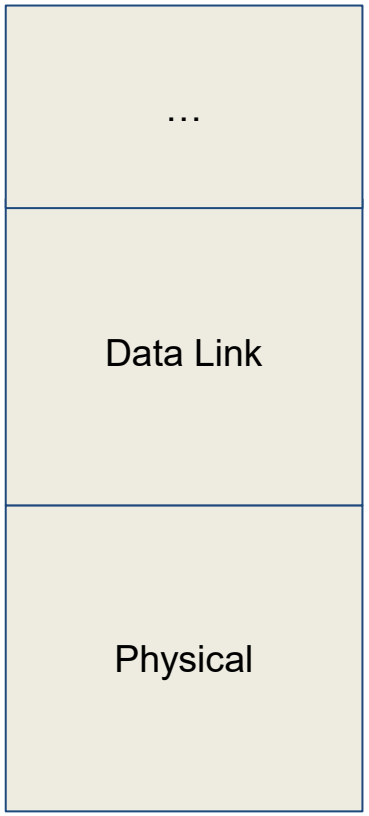
Wait! This isn't a layer2 protocol!

IEEE 802 Standards

Standard	Name	Topic
802.1	Internetworking	Routing,Bridging,and network-to-network Communications
802.2	Logical Link Control	Error and flow control over data frames
802.3	Ethernet LAN	All forms of Ethernet media and interfaces
802.4	Token BUS LAN	All forms of Token Bus media and interfaces
802.5	Token Ring LAN	All forms of Token Ring media and interfaces
802.6	Metropolitan Area Network	MAN technologies,Addressing, and Services
802.7	Broadband technical Advisory Group	Broadband network media,interfaces, adn other Equipments
802.8	Fiber Optic Technical Advisory Group	Fiber Optic media used in token-passing Networks like FDDI
802.9	Integrated Voice/ Data Network	Integration of voice and data traffic Over a single network medium
802.10	Netwok Security	Network access controls,encryption,Certification, and other Security topics
802.11	Wireless Networks	Standards for wireless networking for many different broadcast frquencies and usage techniques
802.12	High-Speed Networking	A variety of 100 Mbps-plus technologies,including 100 BASE-VG
802.14	Cable Broadband LANs and MANs	Standards for designing network over coaxial cable-based broadband connections.
802.15	Wireless Personal Area Networks	The coexistence of wireless personal area networks with Others wireless devices in unlicensed frequency bands.
802.16	Broadband Wireless Access	The atmospheric interface and related functions associated with Wireless Local Loop(WLL)

Ethernet/802 Protocol Stack





Layer 3

Typically IP

- IPv4 32 bit address with a VLSM
- IPv6 128 bit address
  - normally subnet at 32, 44, 48,and 64 bits…

Routable

- Network and Host portions based on mask
- Although source and destination in header, normally bad form to route based on source
- Important to avoid route metrics that induce instability

  i.e. utilization, anything too far down or upstream…

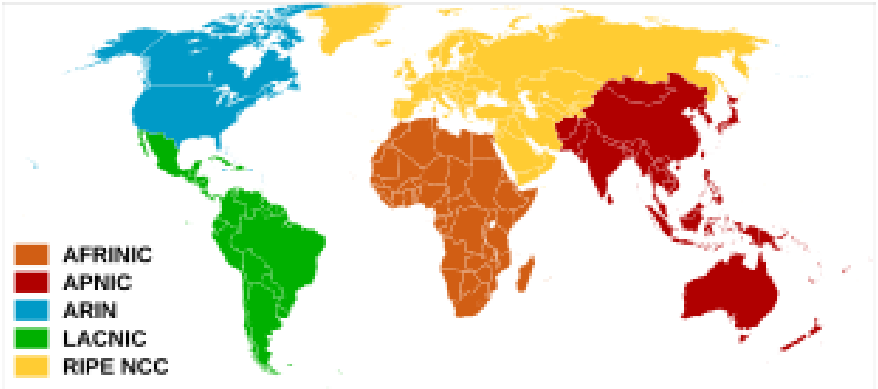
  things that change based on the metric you are setting

Where do layer 3 addresses come from?

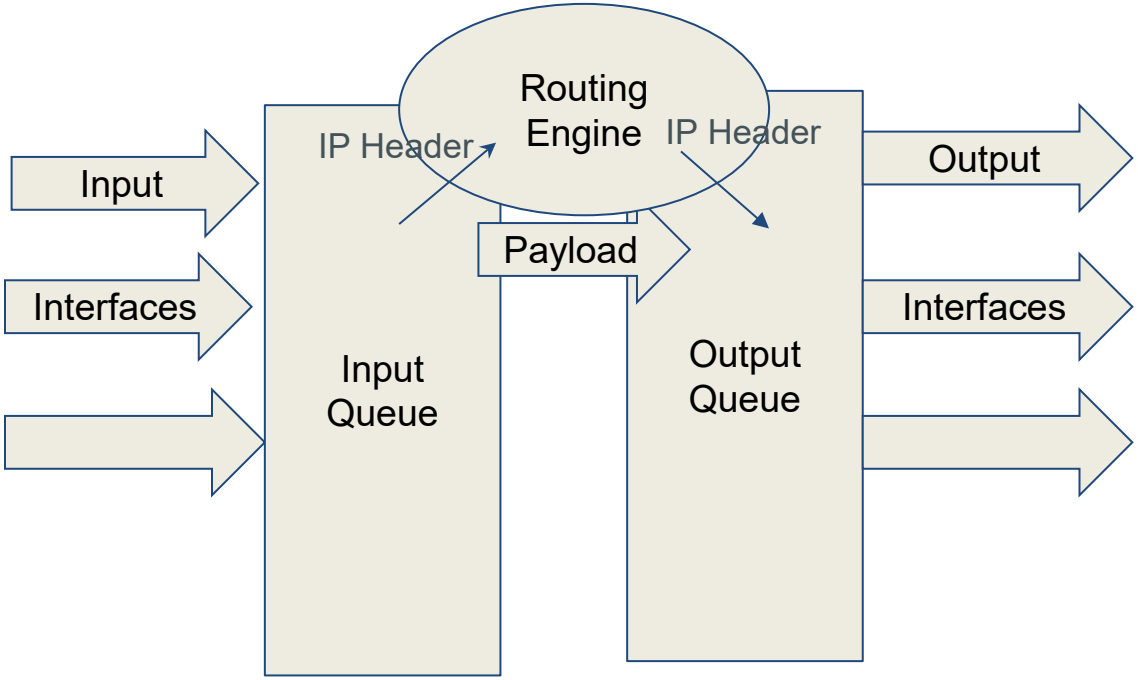
Regional Internet Registries (RIR's)

- Allocated based on need
- From specific blocks assigned for region
  - Legacy allocations don't adhere to post RIR allocations.
- RIR's were established around 1997
- ASN's IPv4 and IPv6 addresses are assigned and recorded

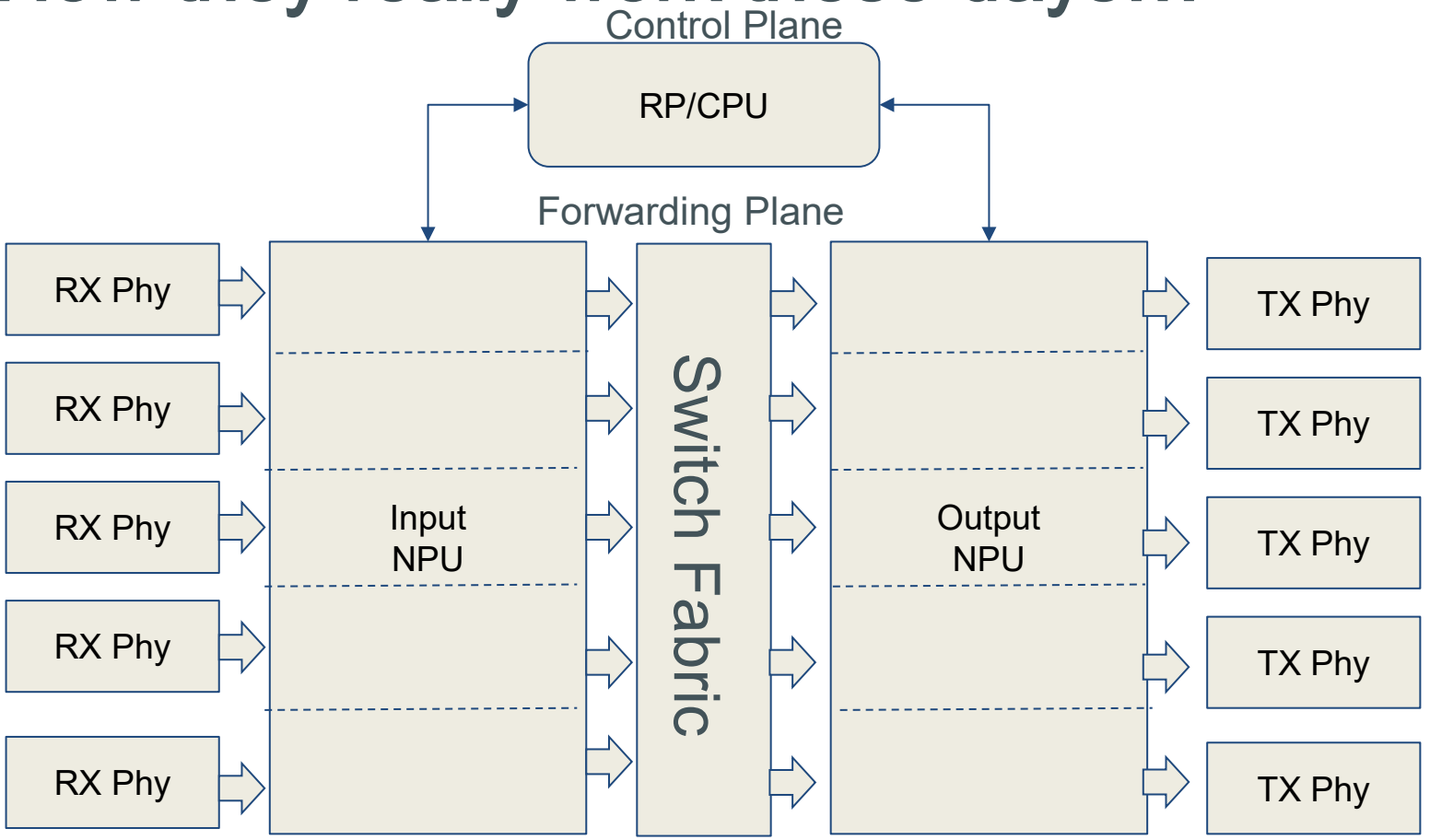
RIPE, ARIN, APNIC, AFRINIC, LACNIC



How a router functions….



How they really work these days…



Connecting networks to networks

How we built up to the Internet…

- BGP is the Internet standard
- ASN's define a group of routers under a common administration system
- AS path is the list of AS's your packet has to pass through to get to its destination

Most defined metrics are not "transitive"

- AS prepending
  - the practice of adding ASN's into the path to make it less desireable
- Community Strings
  - Normally defined as an ASN:Group# (ie 2572:1001)
  - Used to define how an advertisement is to be treated by a network.
  - No defined standard operations

What about higher layers?

Layer 4

- TCP, UDP, and ICMP ports and message types can be used in ACL's
- Are seen but not modified or tracked by routing devices unless they are communications with the control plane

Layer 5+

- Not understood or acted upon at the routing/network level
- I'm a network guy, I deliver the packets…
  I don't want to know what's in them….

How to get better Performance

Make sure the pipes stay full?

Buffering, MTU, Windowing

Are the pipes too full?

Drops, utilization stats

Are we using what we have efficiently?

Too much chatter, scanning, keepalives

Are we trying to do things in the right place on the right stuff?

Routers route, switches switch, Firewalls block and inspect

Are we sure we know what we are looking at?

Is that really traffic dropping, or are they just tired of talking to me?

Performance at the edge

Much of the functions to send and receive data can be offloaded to the modern NIC (Checksums, Syn/Ack, framing and queueing, fragmentation)

Standard buffer sizes are too small to keep large pipes full.

Latency can seriously impact the need for large buffers

Larger MTU's can aid in sending larger flow rates, but you have to make sure the path can handle it.

TCP congestion avoidance and fairness algorithms can be two edged

Some tuning procedures can actually decrease throughput on lower speed links (Large frames and buffers can plug up slower networks)

See https://fasterdata.es.net/host-tuning/

Performance in flight

A circuit with a 50% utilization on a 30 second average isn't half full, it's 100% full for half of that 30 seconds…

.5(1Gbps) x 30sec= 15Gb ->

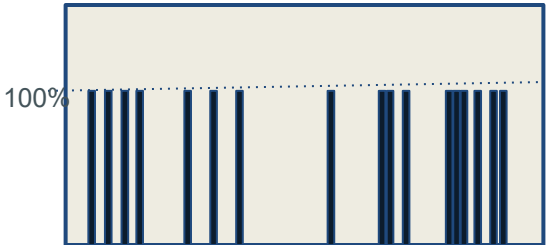
15Gb/100Mbps = 150sec

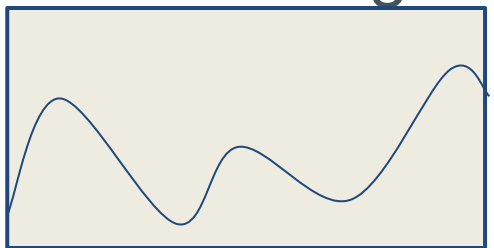
Drops will occur.

Make sure they aren't excessive…

Beware MTU mismatch

Jumbo MTU's don't work in the general Internet





Ways to test for issues

Path and endpoint reachability tools

- Ping, Trace route, MTR, Smoke Ping, OWAMP, TWAMP
- ICMP or UDP echo and TTL scoped tests

Throughput testing

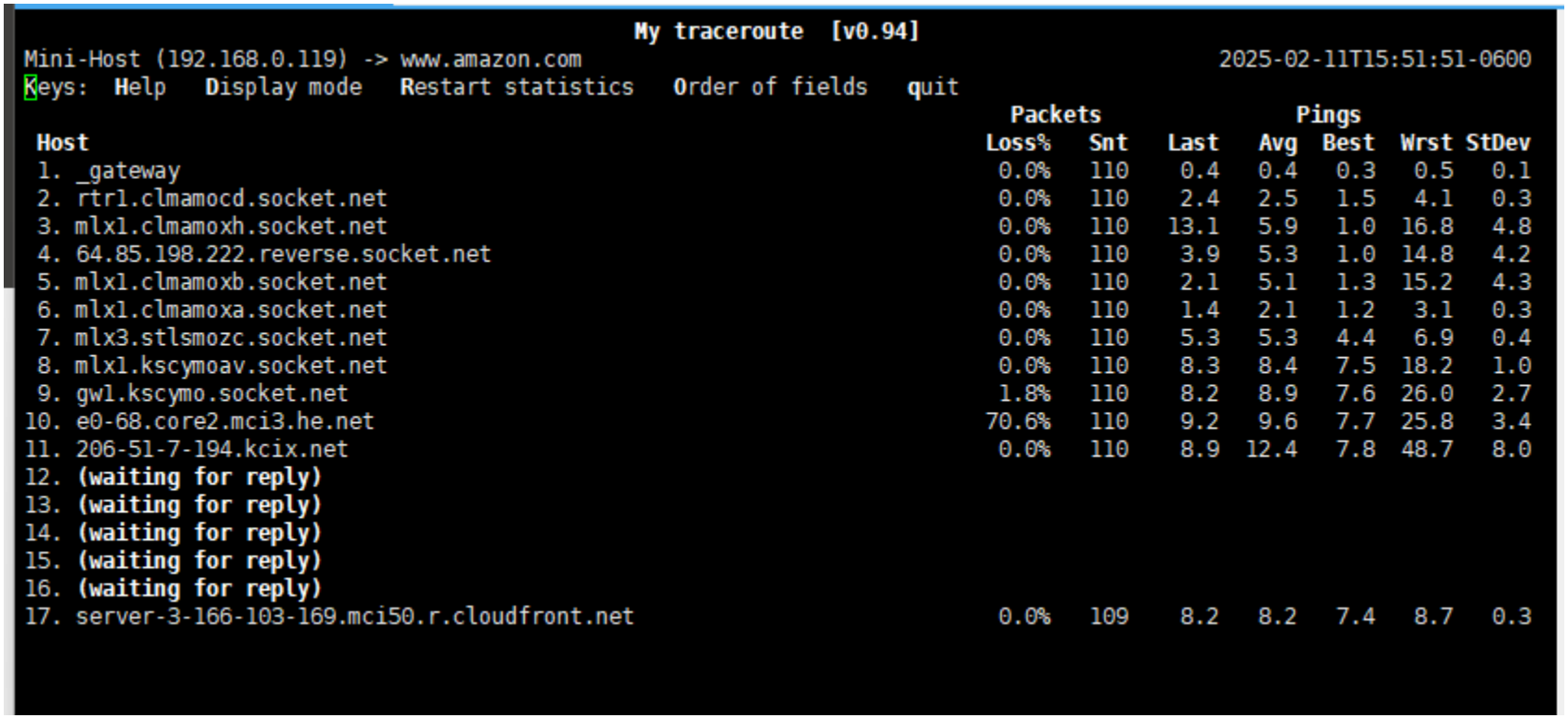
- iPerf, PerfSonar, Thousand Eyes, Ripe Atlas, MLAB tools…
- Normally depends on an endpoint on both ends and some in the middle
- Some are free, some have costs, some have a credit system

Network Mapping and Monitoring

- Don’t do this to others networks unless you have explicit permission
- NMap, SolarWinds, etc…
- Wireshark, Bro, Snort, other IDS/IPS packages…

MTR

mtr -t www.amazon.com



The problem with ICMP…

There's no free lunch…

ICMP requires resources to create and respond to events. Many routing devices rate limit these responses.

We just can't have nice things…

Common open ports and protocol become targets for abuse

Get off my lawn!....

Some security practices require filtering or blocking of ICMP and/or UDP echo

How did you get here?!....

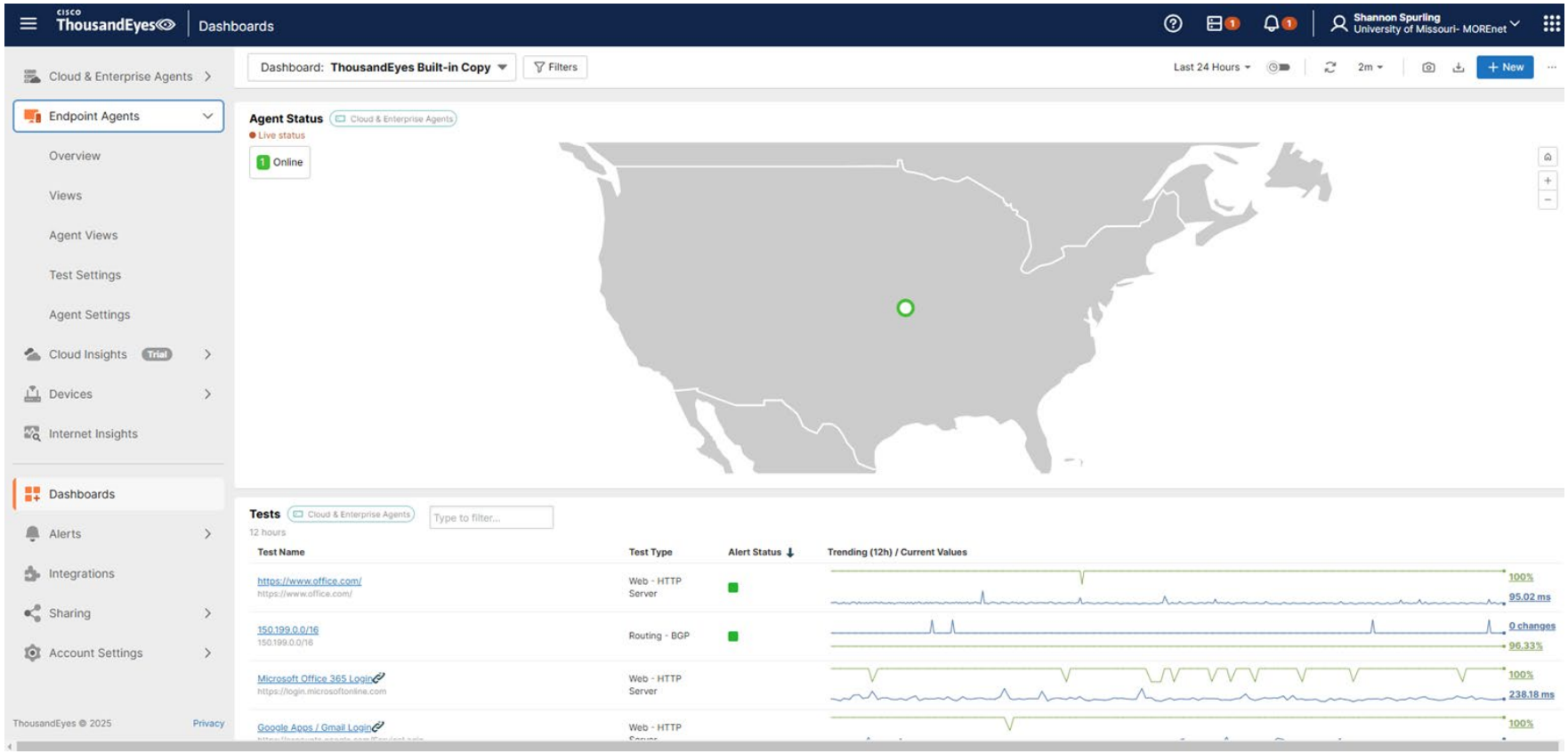
Traceroute is unidirectional. It does not tell you what the return path from any point is, and sometimes that is what is broken. Full Internet routes consume a lot of resources, so some devices only carry what they need to get packets across their local domains and will not know how to return globally routed traffic.

iPerf3 testing

```
shannon@Mini-Host:~$ iperf3 -c 192.168.0.118
Connecting to host 192.168.0.118, port 5201
[  5] local 192.168.0.119 port 53208 connected to 192.168.0.118 port 5201
[ ID] Interval           Transfer     Bitrate         Retr  Cwnd
[  5]   0.00-1.00   sec   113 MBytes   949 Mbits/sec  139    362 KBytes
[  5]   1.00-2.00   sec   111 MBytes   930 Mbits/sec    0    368 KBytes
[  5]   2.00-3.00   sec   112 MBytes   940 Mbits/sec    0    387 KBytes
[  5]   3.00-4.00   sec   108 MBytes   902 Mbits/sec    0    416 KBytes
[  5]   4.00-5.00   sec   111 MBytes   932 Mbits/sec    0    416 KBytes
[  5]   5.00-6.00   sec   103 MBytes   863 Mbits/sec    0    467 KBytes
[  5]   6.00-7.00   sec   111 MBytes   935 Mbits/sec    0    469 KBytes
[  5]   7.00-8.00   sec   107 MBytes   901 Mbits/sec    0    492 KBytes
[  5]   8.00-9.00   sec   112 MBytes   938 Mbits/sec    0    492 KBytes
[  5]   9.00-10.00  sec   109 MBytes   912 Mbits/sec    0    503 KBytes
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bitrate         Retr
[  5]   0.00-10.00  sec  1.07 GBytes   920 Mbits/sec  139             sender
[  5]   0.00-10.04  sec  1.07 GBytes   915 Mbits/sec                  receiver

iperf Done.
```

Thousand Eyes



Be better connected.

PerfSonar Overview

Developed by the research community

PScheduler allows requests from authorized IP's

uses OWAMP/TWAMP, iPerf3, Ping and Traceroute

Several installation options and Linux platforms supported. Important to keep updated…

What about security?

Good security is valuable

Appropriate security is invaluable

Understand what you are protecting and why

Keep policies simple and understandable

Don’t assume something is trustworthy

Does that really need to be on the Internet?

(I am looking at you IoT…)

Safety Considerations…

What

Student Data?

Users?

Medical Records?

Financial Transactions?

Systems?

Why?

Curious people?

Careless People?

Malicious People?

Older systems?

Badly coded software?

Oversights?

Everything over HTTPS?
maybe? Old TLS server?

Understanding context

What is a DDOS (Distributed Denial of Service)?

Functional or Volumetric

Directed at a host or service

Is the goal to disable the network, a service…
…or hide bad actions?

Securing layer 2 and why

Ethernet is leaky, and can be chatty

MacSec (802.1ae) allows Layer2 endpoints to establish encrypted communications (save for untrusted networks)

Locking down ports allows you to prevent unauthorized devices from snooping traffic

Segmenting by purpose and group prevents groups from causing issues with each other (I'm talking about you, IoT…)

Securing layer 3 and why

The Internet is a crazy place

IPsec on the endpoint will keep most data secure to the destination

- Structural VPN's or securing data for regulatory compliance

Route hijacking is possible (register address space properly and implement RPKI)

- Registration under and LRSA/RSA with the RIR and setting up Route Origin Authenticators establishes a trusted anchor for your routes.
- Registration under an Internet Route Registry (ARIN and RAdB) provides additional path information used by networks to build filters

Source address spoofing is problematic

- BCP38 and RPF (Strict and loose) or filtering
- GeoBlocking helps scope traffic for "special" services

What's this security doing to my performance?

Every operation a firewall does induces more latency

More latency means more memory and resources are used per Mbps of traffic

Most firewall vendors will have published numbers for how turning services on impacts throughput, but the sales people won't want you to see them…

Good security takes time and effort, so it will slow things down.

Practical ways to deal with it

Categorize and understand security functions

Layer your security functions and do them where it makes the most sense

Segment groups so that policies can be applied correctly.

Document and monitor your network

(IDS/IPS, Netflow, Interface stats)

The concept of Science DMZ

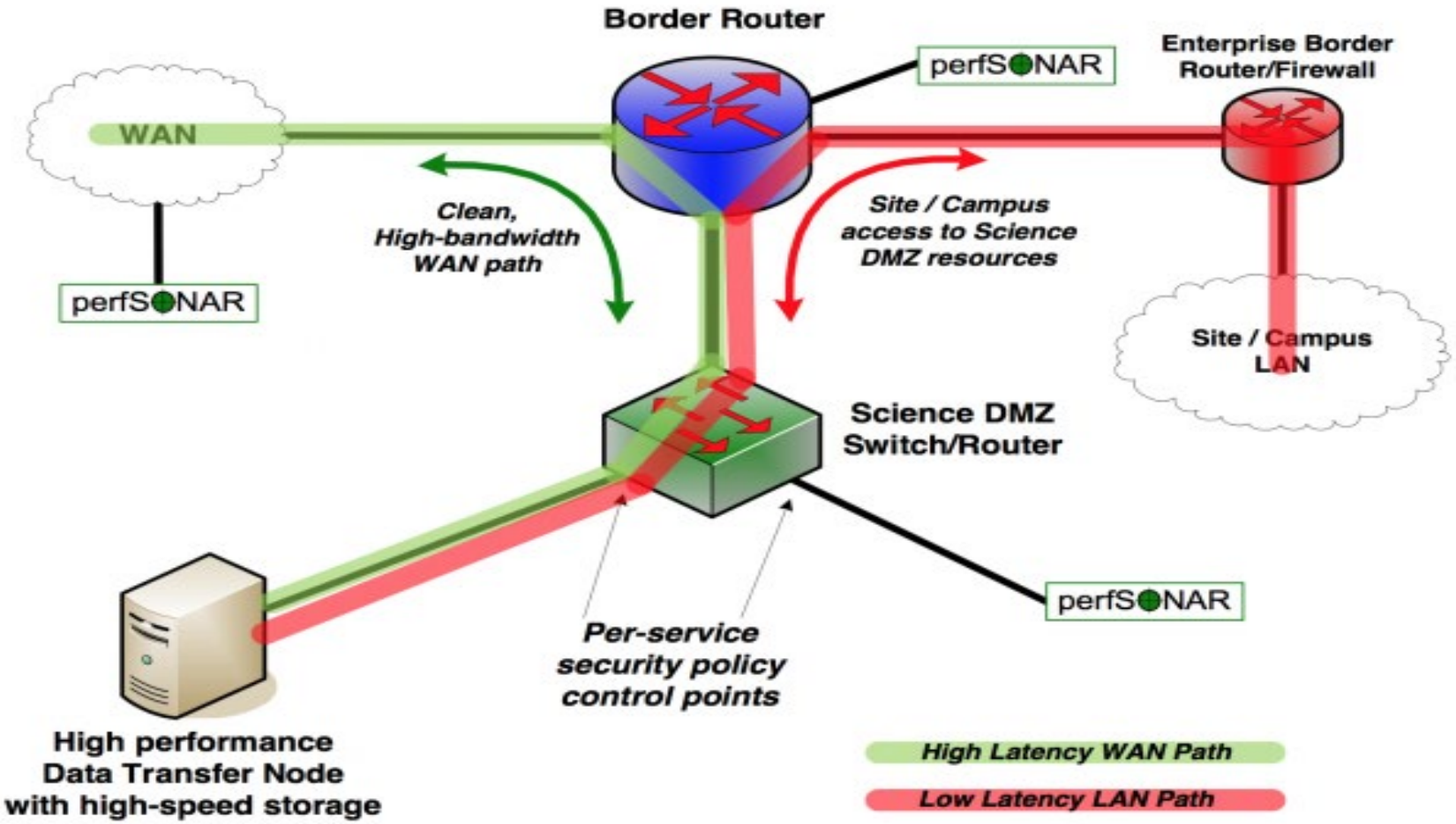
High speed large data transfers are important to researchers.

Create a zone outside of the site security parameter where a set of DTN (Data Transfer Nodes), and compute nodes can be hosted in a scoped and monitored network.

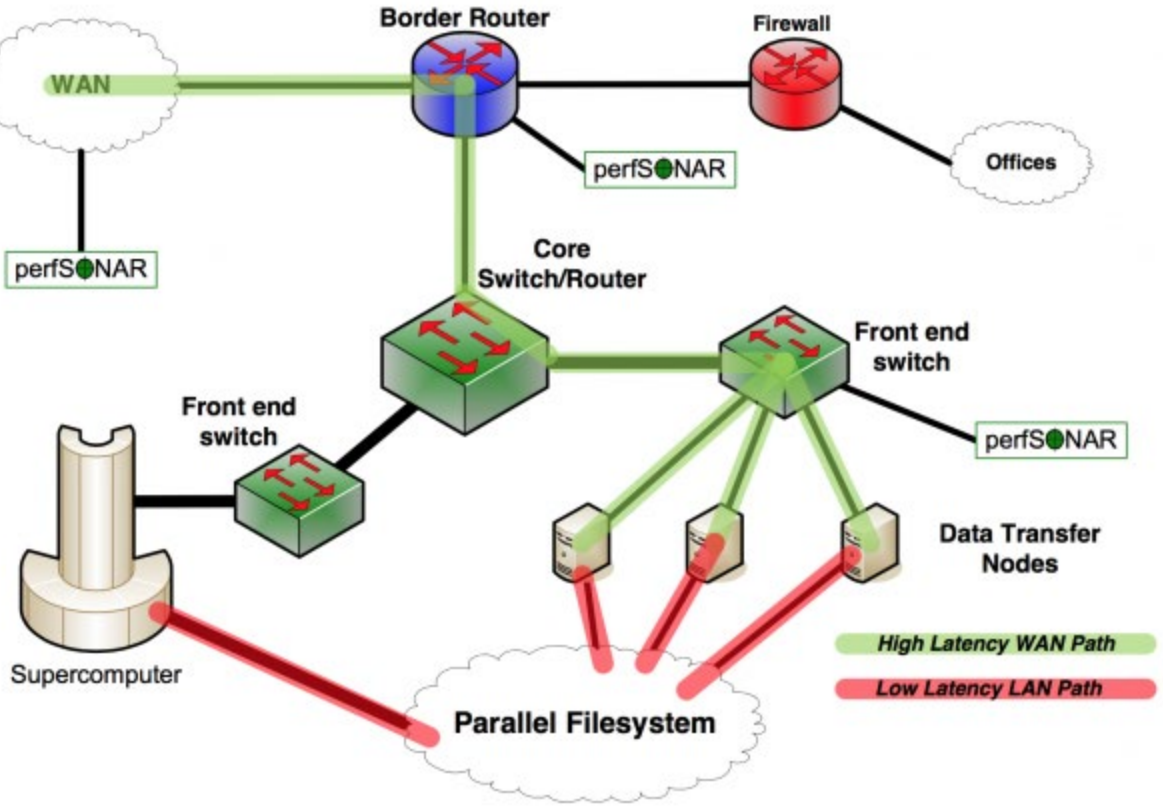
No users or questionable systems allowed!

Monitored by an IDS/IPS and scoped by proper ACL rules to limit attack surface.

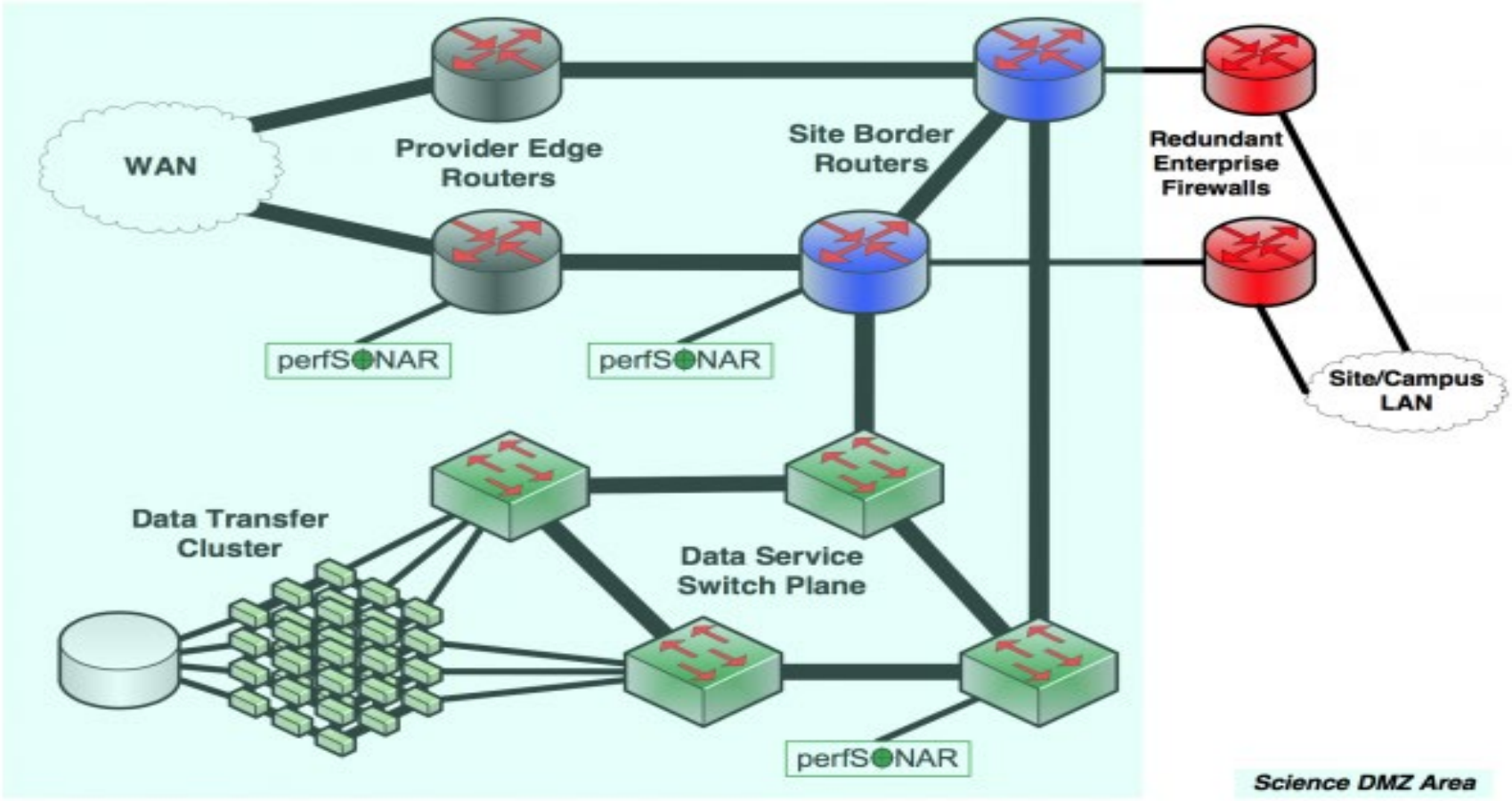
What it looks like…



But, what if I have a Supercomputer?



… Or a big data site?



Using the RIR's and IRR's

RIR's are the authority where all resource ownership for Layer 3 addresses and ASN's are recorded.

ARIN also offers an RPKI ROA registry

ROA's (Route Origin Autenticator) cryptographicly associate route blocks with the BGP ASN they should be originated from.

IRR's (Internet Route Registries) Are used to allow BGP speaking networks to say what ASN's and IP networks they provide service for.

The MANRS project

Mutually Agreed Norms for Routing Security (MANRS) is a global initiative…that provides crucial fixes to reduce the most common routing threats.

Encourages the registration of routing information with RIR's and IRR's, as well as the use of RPKI.

https://observatory.manrs.org/#/overview

Resources

Information on tuning and science DMZ

https://fasterdata.es.net

https://atlas.ripe.net/

https://radar.cloudflare.com/

https://irrexplorer.nlnog.net/

https://console.internet2.edu/#/

https://www.routeviews.org/routeviews/

https://www.measurementlab.net/tests/

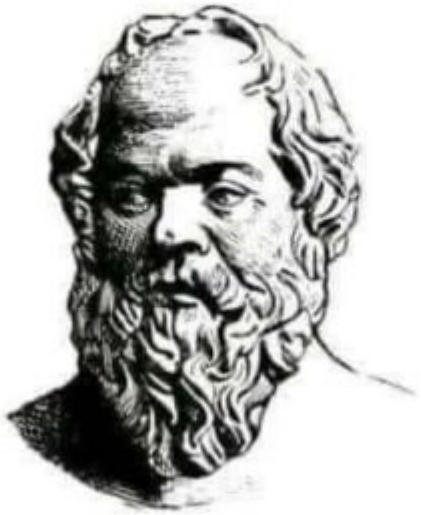
https://docs.perfsonar.net/index.html

https://www.speedguide.net/downloads.php

Done…

Questions?

"ehh, good enough"



- Mediocrates

Be better connected.