MOREnet

Be better connected.

# Routing Performance & Security

How did we get here, and why we can’t have nice things…

Shannon Spurling

MOREnet - Network Engineer, Master

# Synopsis

To be safe, we have to understand what we are securing…

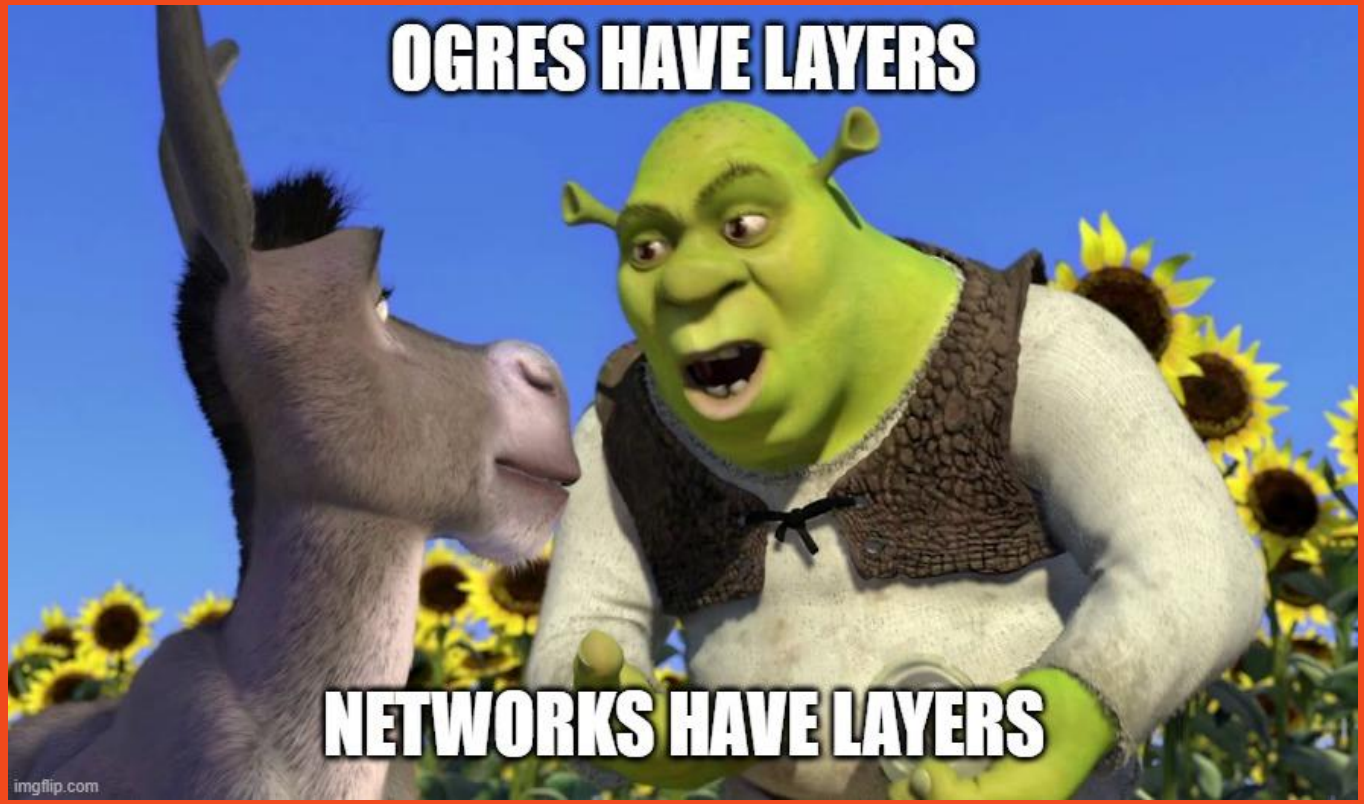
To improve performance, we have to understand how our systems work…

This stuff didn't appear out of nothing
It wasn't designed like this

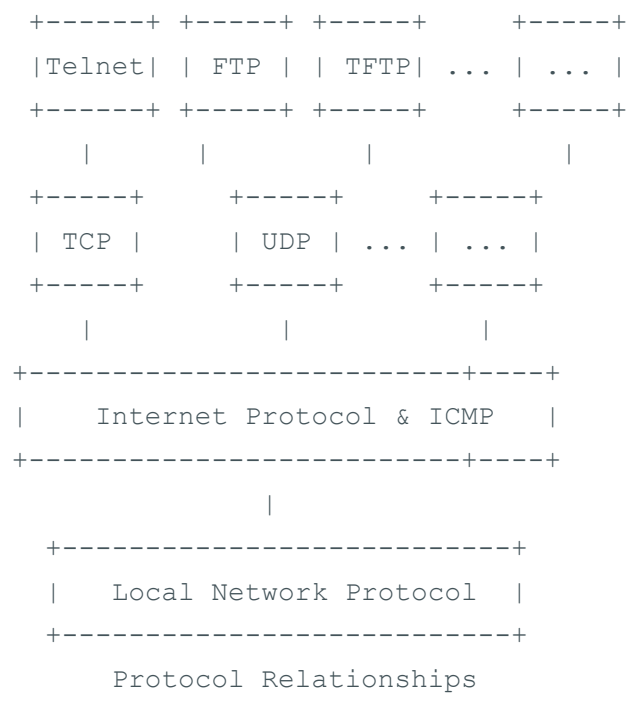
It grew from something else…

WHY?


"You get to drink from the firehose!"


OGRES HAVE LAYERS
NETWORKS HAVE LAYERS
imgflip.com

# Here’s some layers!




Protocol Relationships

*Be better connected.*

# Routing vs Switching

- Switching (Flood/Broadcast then Prune) Layer2
  - Flat topology
  - Fast
  - Dumb, Low Cost, Few Features
  - Learn where things are…
- Routing (Learn then send best adjacency) Layer3
  - Multiple layers
  - More complex, higher cost
  - Tell others what I know! (igp like ISIS OSPF EIGRP RIP)

Paths are unidirectional… packets and frames are going places, not looking at how to return…

# Switching

Repeaters
Bridges
Learning switches

Low latency
Traffic bleeds, very chatty
Low tolerance for bandwidth differences (shallow buffers)
Great for close low cost networks, bad for long distance high cost
Mac addresses and VLAN ID’s

# Routing Principles

Process routed
Express forwarding
CAM forwarding
Network Processors

No route, no path
Deeper buffers, better transitions higher latency
No traffic bleeding
Expensive
IPv4 and IPv6

# But, how do we know where things are?

- ADDRESSES!
  - Workstation Identifier
  - Network locator
  - Organizational Subdivision
  - Organization
  - State
  - Country
  - Region

# What is an IPv4 address

- RFC791 (ca. 1981)
- Prefix assigned by IANA or your local RIR
- Classfull OR Classless (CIDR)

- Network division at the bit boundary…
- 4 sets of 8 bits separated by ‘.’ (32 bits)
- Smallest route-able network on Internet is a /24
- Currently 930K prefixes in global routing table
- Got to brush up on your binary math…

# IPv6 addressing

- RFC 1885 (ca. Dec 1995) Updated since…
- Includes headers useful for transport and path definition.
- Assigned using RFC6177 recomendations
  - /32, /44,/48,.../64,[/127,/128]
- 128 bit addresses
  - 2320:10E0:45FA:342::B213 /64…
- Currently 184K prefixes in global routing table

# Routing between your LAN's

IGP

Pretty  simple

Static/Connected

Distance Vector

Link State


It's your circus! It's your monkeys!

# Talking between Networks

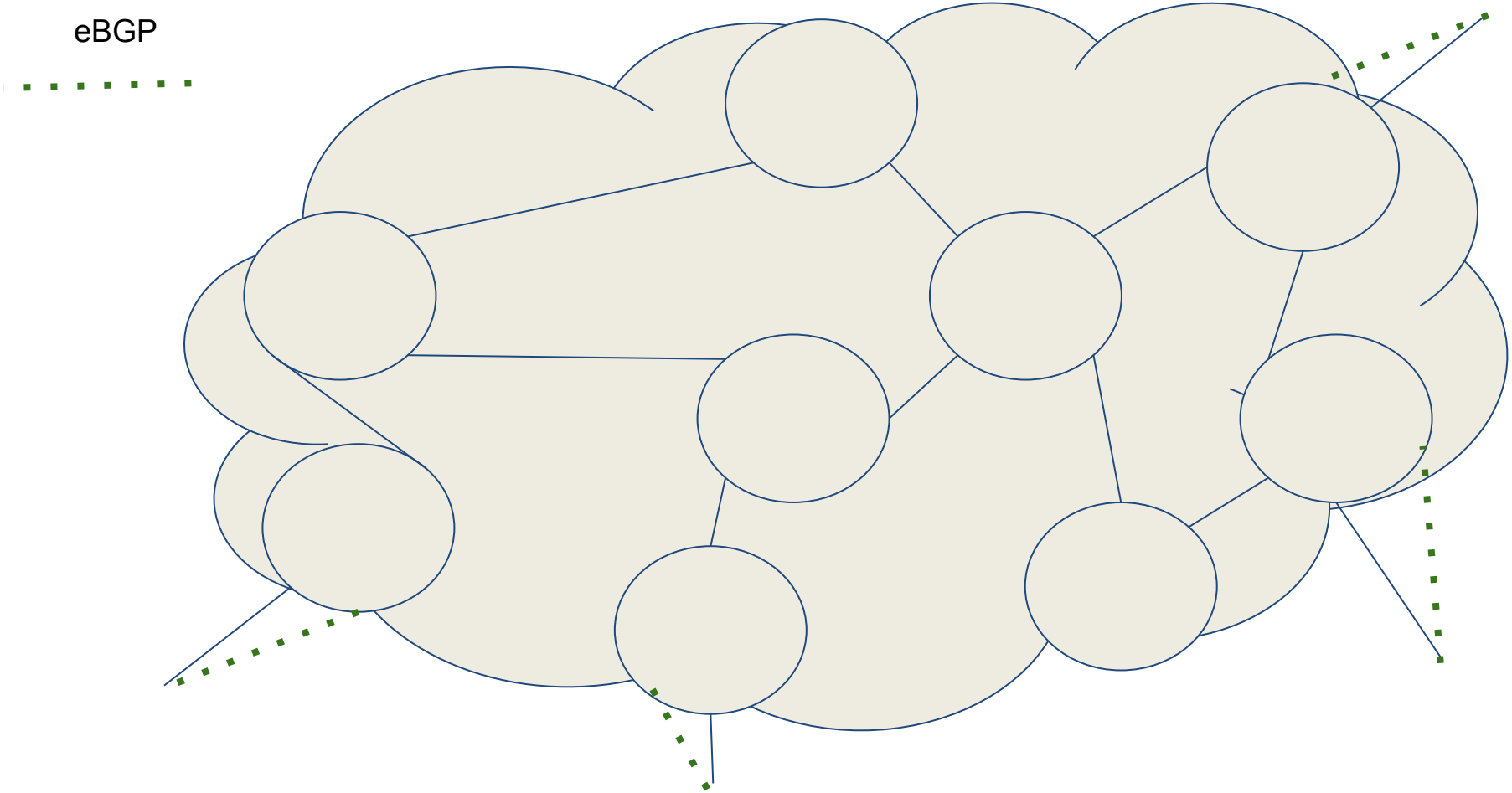
BGP

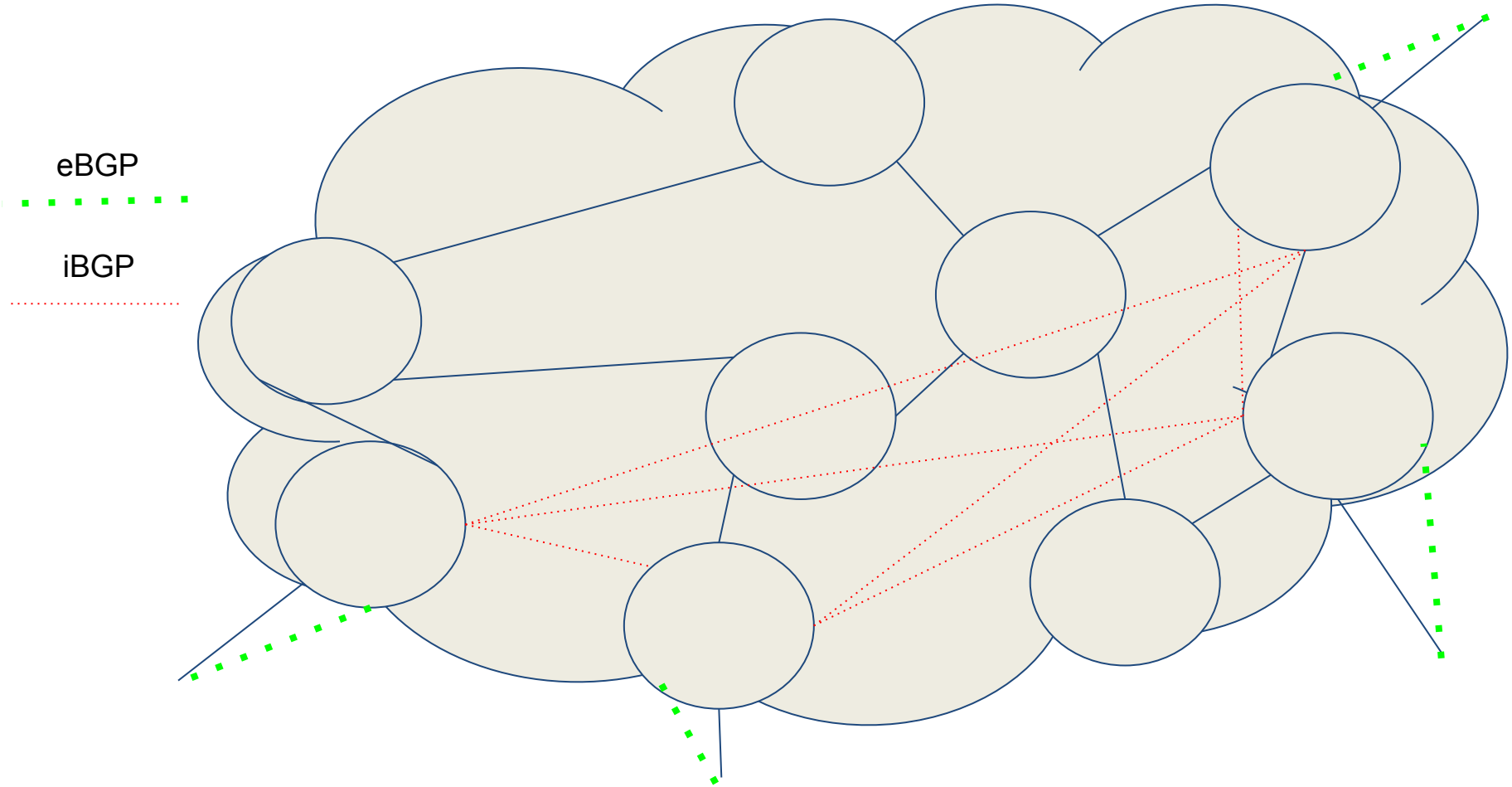
- TCP protocol port 179 (It's an application, like DNS)
- Autonomous System is a group of routers under a common set of policies (Typically a single organization)
- Everything including the kitchen sink (Address Families)
    - Routing
    - Policy
    - Other protocols…

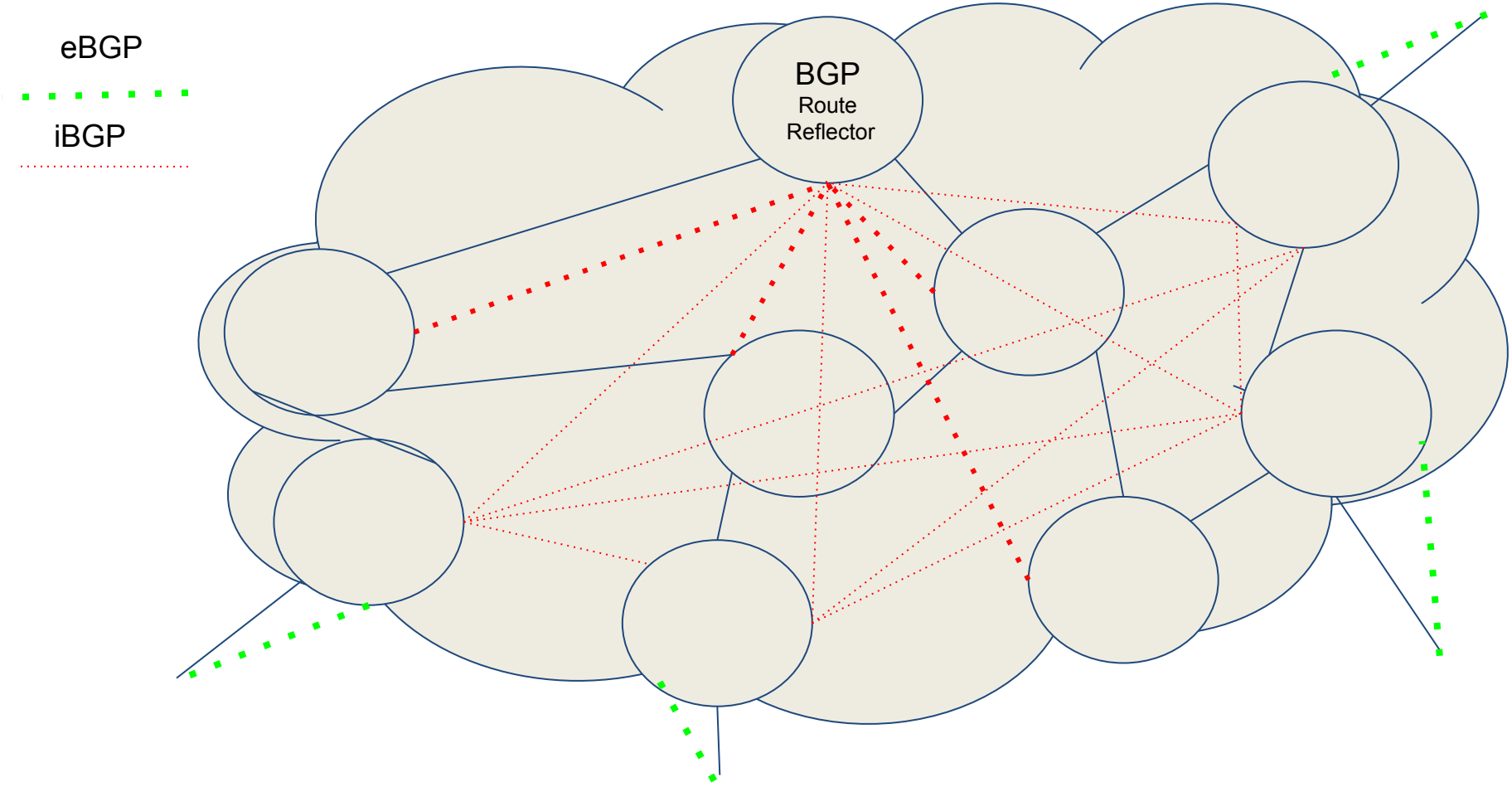
Telling others about your routes (routing), how to treat traffic returning to you (flowspec)... they are recommendations and will be modified in flight… Receivers policy always takes precedence…

# BGP Operation

- eBGP is between different ASN's
  - Normally restricted to next hop peering
  - Can be multi-hop with TTL hop count limit
  - Can also MD5 sign BGP messages
- iBGP is when the ASN of the speakers is the same
  - iBGP speakers only tell iBGP neighbors of what was learned via eBGP peers (Full Mesh VS Route Reflectors)
- eBGP peer is maintained as next hop by default
- IGP is needed to provide path to egress
- BGP uses list of "metrics" to choose best path


eBGP


eBGP
iBGP


eBGP
iBGP
BGP
Route
Reflector

# BGP metrics

WEIGHT. (cisco only)
LOCAL_PREF. (Local, internal only)
LOCAL ORIGIN
SHORTEST AS-PATH (Path prepending)
ORIGIN TYPE
LOWEST MED (Transitive if allowed by peer)
eBGP over iBGP (AKA hot potato routing)
Lowest igp metric to BGP next hop
BGP Multipath
Oldest
lowest router ID... etc...

BUT.. Longest prefix length is always preferred (/24 > /21)!
Called LPM or Longest Prefix Match…

# Let’s look at the tools we have…

## Ping

- Ping
- Smoke ping
- MTR

## Traceroute

- Traceroute (UDP vs ICMP)
- MTR

## Bandwidth/Jitter/Latency testing

- iPerf
- PerfSonar
- BW testing sites
- Commercial testing suites

# Barriers to performance

Traffic Congestion on switches and routers

Frame Size (Maximum Transmission Unit- MTU)

Recieve Buffers

In-line Security Nodes/Inspection

# Solving the congestion issue

TX queue might be too small by default?

10G feeding a 10M link will drop frames

Make sure QoS policies are up to date

You might have too much traffic for your link

Microbursts are real.

Just because your traffic is at 30% doesn't mean you won't drop traffic from a large quick burst.

# The problem with MTU

Too large of a packet will be fragmented or dropped forcing a retransmission. (normally less than 1500 bytes)

PMTU (Path Maximum Transmission Unit) negotiation has solved most fragmentation issues.

If there is a MTU mismatch between ends of a link can break PMTU and cause excessive packet drops.

# Client/Server issues

Receive and transmit buffers can be too small

CPU may be too busy with other things

Memory available may be too small to keep pipe full

Excessive interrupts/disruptions

# Security comes at a cost

Deep inspection takes processing time (++$$$$$$)
Session tracking requires processing time (++$$$$)
Classifiers/Signature checks require processing time
SSL decryption takes time and is intrusive

Machine Learning is the solution?

# How do we secure things?

- Assess risk
- Limit questionable sources/protocols
- Monitor ongoing transactions
- Understanding current trends/issues
- Trusted communications channels
- Establish some form of trust
- Centralized agreed upon authority
- Signatures to establish authenticity

# … Are we really secure?

Internet traffic is actually unidirectional.

TCP handshakes are layer 4 and require keeping state to watch/understand.

Many newer protocols are defaulting to UDP to reduce overhead.

# Current Common Attack Methods

## Volumetric DDOS's

- Unidirectional
- Botnet initiated
- Involves DNS or some other amplification service

## Functional DDOS's

- Trigger inspection or other high cost function on the destination or an intermediate security device
- SYN/ACK style attacks
- Injection attacks

## Buffer Overflow/Application and OS level attacks

- Using exceptions to break system controls and gain access to a device

## SPAM/PHISHING/Social Engineering

- Fooling filthy humans

## Man in the Middle

- Involves compromise of middle box, or fraudulent routing

# Fixing the volumetric DDOS

## You don’t want that traffic!

## Identify and drop traffic based on source and protocol?

- Attacker is trying to make traffic look as genuine as possible.
- Bot trigger devices are masked by spoofing

## Divert to a distributed scrubbing service or device

- Scrubbing devices are expensive
- Services are pay up front, or pay per use
- On network devices still require you to burn your bandwidth.
- MOREnet has a consortium agreement to use a GPN contracted service. Up to 20Gbps of scrubbed traffic at no additional cost to members.

# What can be done about functional DDOS's?

Typically use a flaw in security devices to slow or stop processing of incoming traffic.

- Keep the security policy as simple as possible
- Make sure to be up to date on patching
- Make sure security device has capacity for its application
- Volumetric scrubbing has little impact. Temporary blocks can be used to ease impact but should not be looked at for long term mitigation.

# Buffer Overflow style attacks

Not much that can be done at the network level.

- These look like normal protocol interactions.
- Looking past the header into the payload is not practical for a router…

Does that port need to be opened?

Are you up to date on your patches?

Who is actually supposed to be talking to that server?

# Social Breaches?

Be practical in your security policies.

Train your people properly for the level of access they have.

Expect issues.

Use common sense, be proactive, be aware, but don’t fall to paranoia…

# Man in the Middle

How does one get in the middle?

- Compromise a router?
- Break into a data center?
- Divert a route…

# So, What's root of the problem?

I.e. BGP is a big game of Telephone…

…Can you trust caller ID?

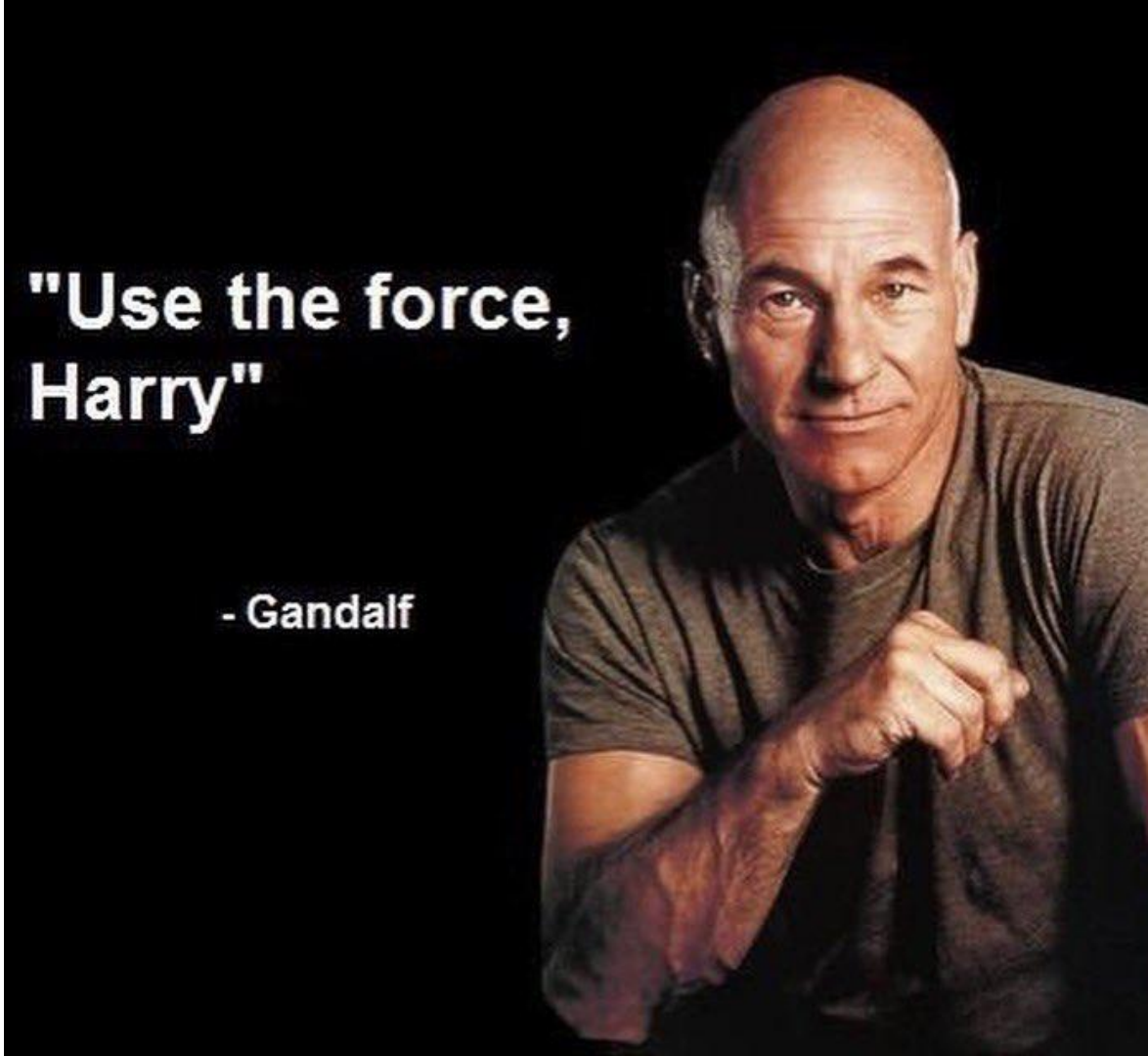
# If they aren't who I think they are, What could happen?

Source Spoofing

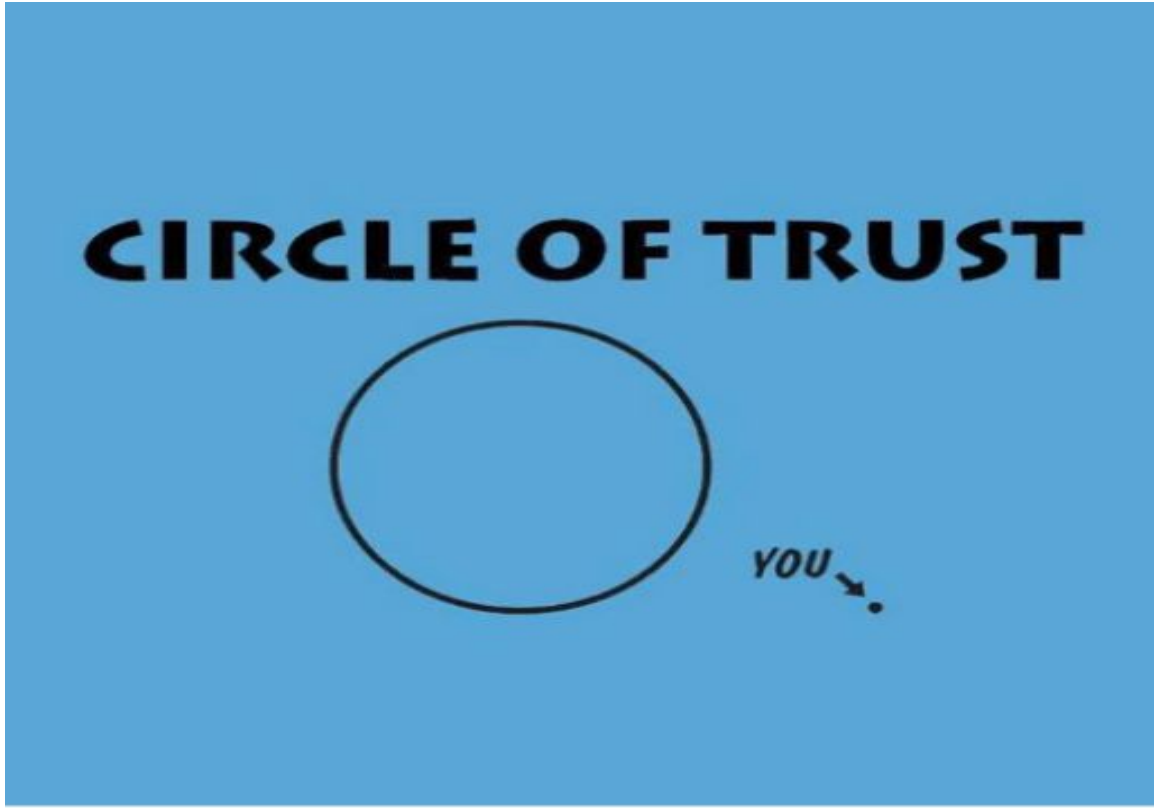
Traffic Diversion

Misconfiguration

If only there were someone we could trust?



# Why can’t people trust you?




*Be better connected.*

# Centralized registries and authorities

ICANN (International Corporation of Assigned Names and Numbers)- IANA (Internet Association for Names and Addresses)

RIR's(Regional Internet Registries) (i.e. ARIN)

IRR's(Internet Route Registries) (i.e. RAdB)

Standards Bodies

OSI/ISO, IEEE, NSF, …

Community groups

PeeringdB, NANOG…

# Internet Route Registries

Typically run by a trusted organization, association, or vendor who has established security and authentication protocols for establishing and updating information

Used to construct routing filters and device policy to keep routing tables clean.

Many Internet Backbone vendors require Route Registry entries so they can appropriately filter routes

# Things recorded in route registries

- IP addresses routed by an organization
  - (Route, Route-set Objects)
- ASN’s that should be learned from an organization
  - (Aut-Num, AS-Set Objects)
- Policies to apply when parsing data from an organization
  - (Peering-set, Filter-set Objects)

# Internet Registries...

ARIN established to delegate Internet Resources so John Postel could take a vacation….

Source for all new IPv4, IPv6 and ASN assignments since December of 1997

Problem is, what about all those numbers delegated before December 1997?...

# ARIN and Legacy resources

- After 1997, IP addresses were regionalized.
- Prior registrations were not assigned under agreement with ARIN and are grandfathered into the Internet.
  - Covers most higher Education and government ranges.
  - Many pre-RIR ranges do not conform to current geographical assignments… (150.199.0.0/16 is in APNIC space)
- ARIN will provide registration services under an LRSA (Legacy Resource Service Agreement).
- Legacy resources are managed at a discounted rate as long as an LRSA is in process before 1/1/2024.

# If I am grandfathered in…

Why do I need an LRSA?

- Ensures current contact information filed with the authoritative organization…
- Access to ARIN services…
  - ARINdB
  - Hosted RPKI

# What is this RPKI you speak of?

Resource Public Key Infrastructure

- Cryptographically signed resource Certificates (ROA's- Route Origin Authenticator's) are held at Internet Registries
  - These contain origin AS, Prefixes and prefix lengths that can be expected in BGP advertisements.
- ROA's are downloaded and cached by a validator
- Prefixes are compared and marked
  - Valid - prefix and length, as well as AS match
  - Unknown - There is no ROA that covers this prefix
  - Invalid - There is an ROA, but the source and/or length don't match. Indicates config issue or hijack.

# What does all this do for me?

- Path information can be derived from IRR data
  - Prevents route leakage
  - AS paths could be validated in future when combined with RPKI origin signing
- Source ASN and prefix limits derived from RPKI
  - Prevents misconfiguration from causing outages
  - Prevents malicious foreign actors from diverting traffic
- If you sign an LRSA, and use MOREnet as a provider, we can help manage IRR and RIR (sign ROA's etc…) if we are set as a routing Point Of Contact (POC)

# MANRS Community

Mutually Agreed Norms for Routing Security (MANRS)
is a global initiative that helps reduce the most common routing threats.

# Science DMZ

Runs in parallel to production environment

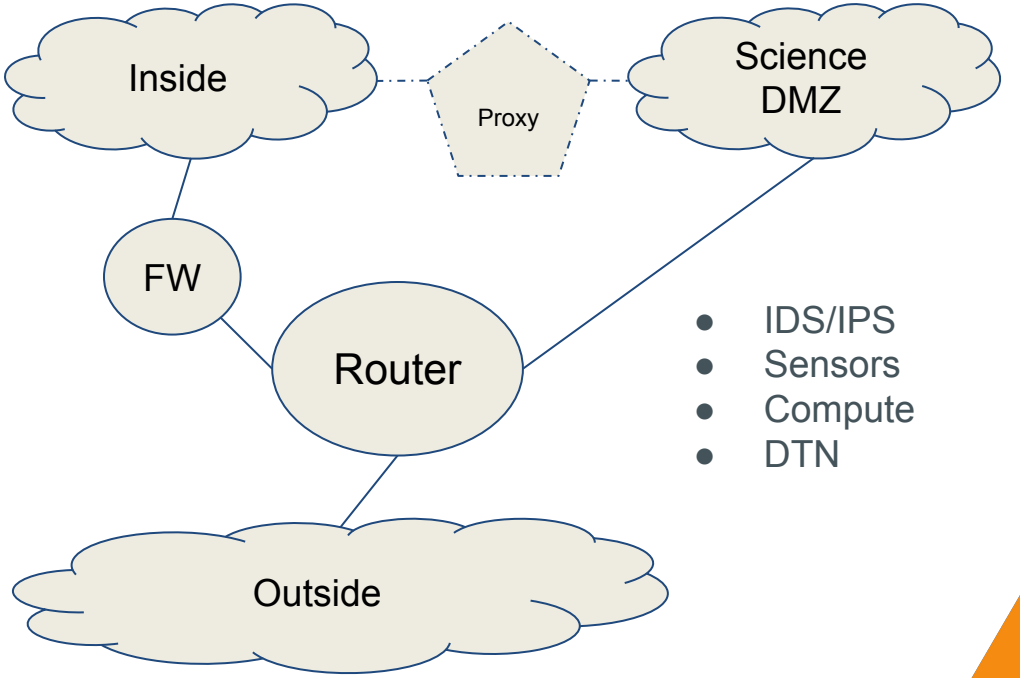
Reduces performance drain due to firewall

Improves security by removing soft targets

Only houses DTN’s, instrumentation, and Compute nodes

Relies on IPS/IDS and packet filtering for security policies

# What's it look like?



# Fin...

## Questions?

Oh, yeah, here are some fun resources…

# Resources...

ARIN: https://www.arin.net/
RAdB: https://www.radb.net/
Radia Perlman: https://youtu.be/qXz_RxBFQ20?si=sBYRDymdjiULarlg
RFC Editor: https://www.rfc-editor.org/
MANRS: https://www.manrs.org/
History of IANA: https://www.internetsociety.org/ianatimeline/
RIPE Tools:
https://www.ripe.net/manage-ips-and-asns/resource-management/rpki/tools-and-resources
RIPE Atlas:https://atlas.ripe.net/
MTR tutorial: https://chemicloud.com/kb/article/how-to-perform-a-mtr-in-windows-mac-os-and-linux/
PerfSonar: https://www.perfsonar.net/
SmokePing: https://github.com/oetiker/SmokePing
iPerf: https://github.com/esnet/iperf