Internet Peering

Concepts and Emerging Trends

Phil Bedard
Distinguished TME

cisco Live!

BRKMSI-2005

Cisco Webex App

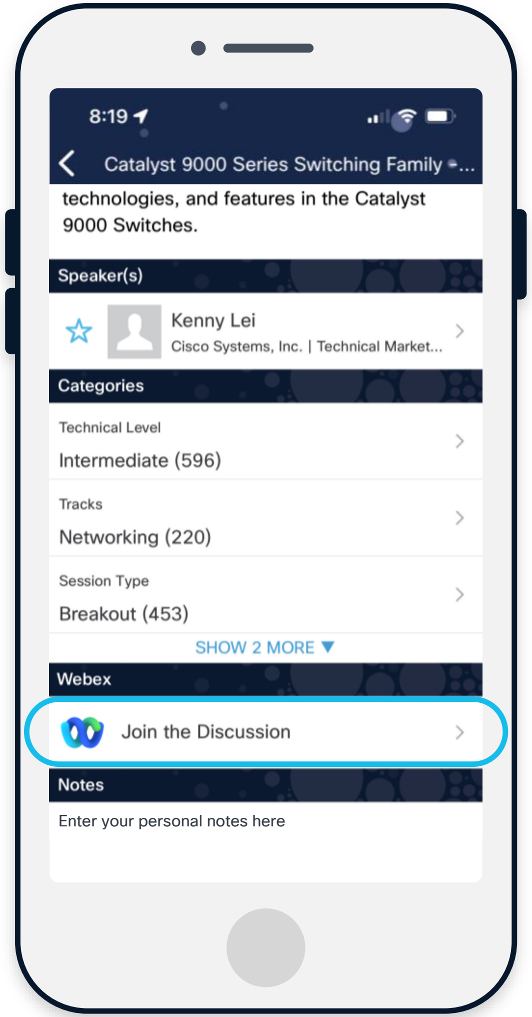
Questions?

Use Cisco Webex App to chat
with the speaker after the session

How

1. Find this session in the Cisco Live Mobile App
2. Click “Join the Discussion”
3. Install the Webex App or go directly to the Webex space
4. Enter messages/questions in the Webex space

Webex spaces will be moderated by the speaker until June 13, 2025.



8:19

Catalyst 9000 Series Switching Family -...

technologies, and features in the Catalyst 9000 Switches.

Speaker(s)

Kenny Lei
Cisco Systems, Inc. | Technical Market...

Categories

Technical Level
Intermediate (596)

Tracks
Networking (220)

Session Type
Breakout (453)

SHOW 2 MORE ▼

Webex

Join the Discussion

Notes

Enter your personal notes here

https://ciscolive.ciscoevents.com/
ciscolivebot/**#BRKMSI-2005**

Agenda – A Peering Story

Peering Intro

Peering Network Design

Peering Network Telemetry

Peering Security

Introduction to Peering

What is Peering?

“Peering is the interconnection and exchange of IP data between two networks under different administrative control.”

Peering is the glue holding together the Internet, without it the flow of data across the Internet would not be possible.

Peering represents an important administrative, operational, and security boundary between IP networks.

”Peering” in 2025 = Interconnection covering Content Delivery, Business to Business Services, Cloud Interconnect, and Traditional Peering

While the fundamental role of peering hasn’t changed, traffic patterns, location, operation, and security requirements have, so peering must evolve as well.

Internet Evolution

“Public” Internet circa 1995

- Low bandwidth clients, dial-up
- Many smaller regional Internet providers
- ~16M users
- Wireline only
- Static content
- More widespread content sources contributed to volume

Cisco Systems
Empowering the Internet Generation

Cisco Connection Online
The Worldwide Leader in Networking for the Internet

HOW TO BUY LOGIN REGISTER SITE MAP COUNTRIES/LANGUAGES

Products
Select a Category Go

Search
Go
Advanced | Help with Search

Solutions
Service Providers, Enterprise SNA, Education, Government, Large Enterprise, Small/Med. Business, Small/Home Office

Corporate News & Info.
About Cisco, Public Relations, Investor Information, Publications

Products & Ordering
General Information, Products, Technology, Online Ordering, Ordering Information, Documentation, More Information

Software & Support
Technical Support, Documentation, Customer Service Programs, Software Center, Support Benefits, Technical Support Contacts

Training, Seminars & Events
Training, Cisco Career Certifications, CCIE Program, Seminars, Events

Top News
Cisco Systems Wins Summary Judgement in Patent Case
More News

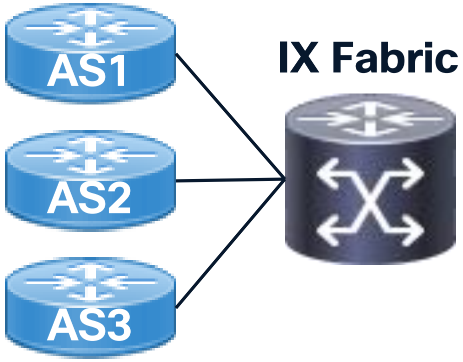
Today’s Internet

- High-speed Internet is widely available
- 100s of millions mobile users
- 4 billion+ users worldwide
- Static content replaced with video
- Traffic **volume** driven by fewer sources
- Leads to “flattening” of Internet: Direct interconnection between producer and consumer networks

YouTube
amazon
NETFLIX
Disney+
tv+

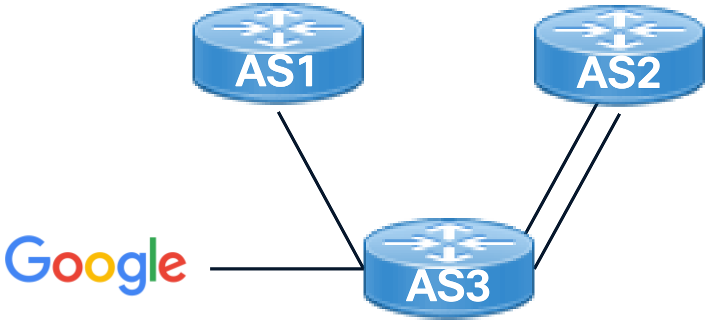
Interconnection Types

Public Fabric



- Easy to connect to many peers
- IXP can provide redundancy

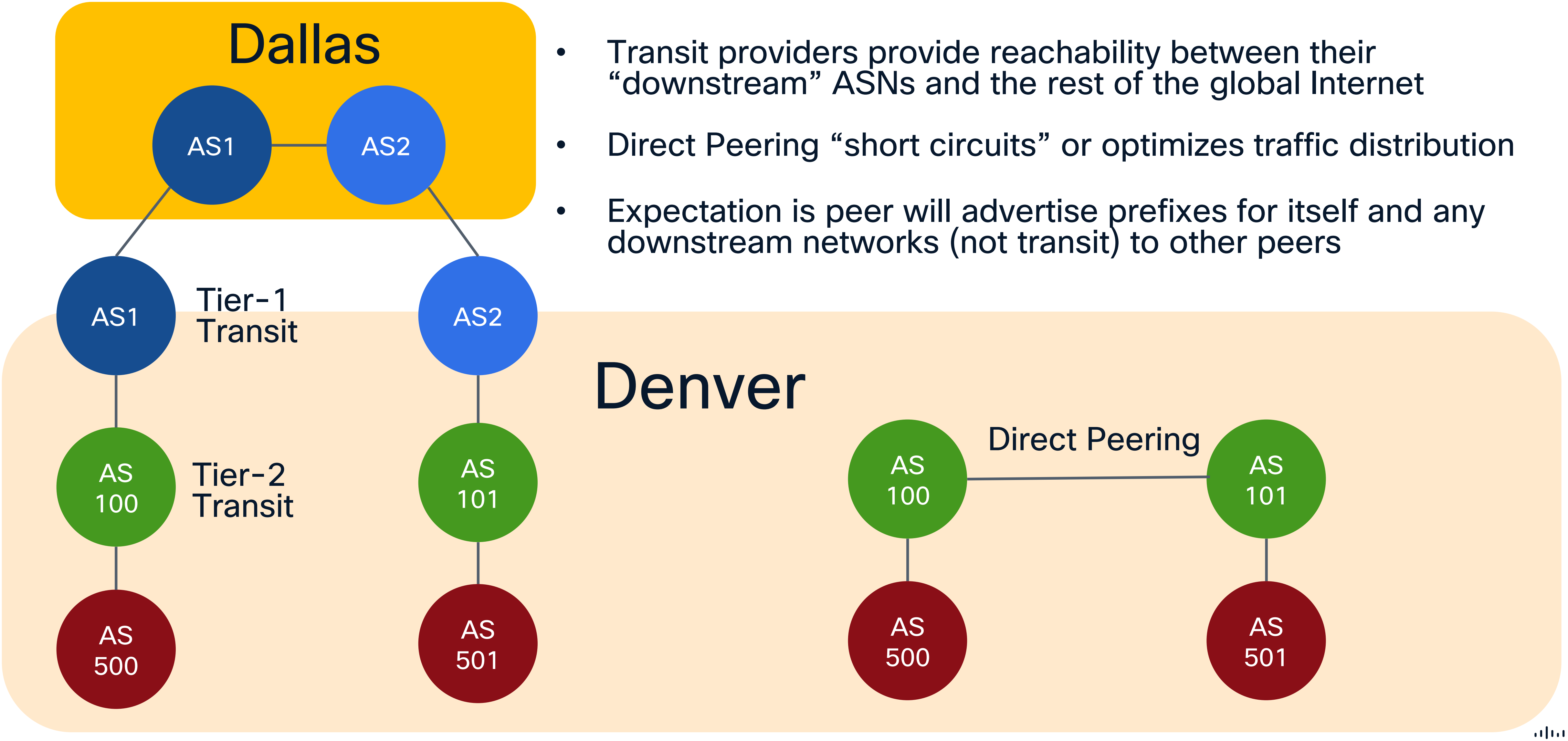
Private Interconnect



- High traffic volume
- Independent capacity

- Public or private fabrics interconnect many networks worldwide
  - AMS-IX and IX.BR two of the largest IX fabrics
- Highest percentage of traffic **volume** today carried over PNI
- Largest SP and content providers trending to more PNI
- CDN is a type of PNI, may or may not include BGP

“Peering” vs. Transit



- Transit providers provide reachability between their “downstream” ASNs and the rest of the global Internet
- Direct Peering “short circuits” or optimizes traffic distribution
- Expectation is peer will advertise prefixes for itself and any downstream networks (not transit) to other peers

Why should you peer?

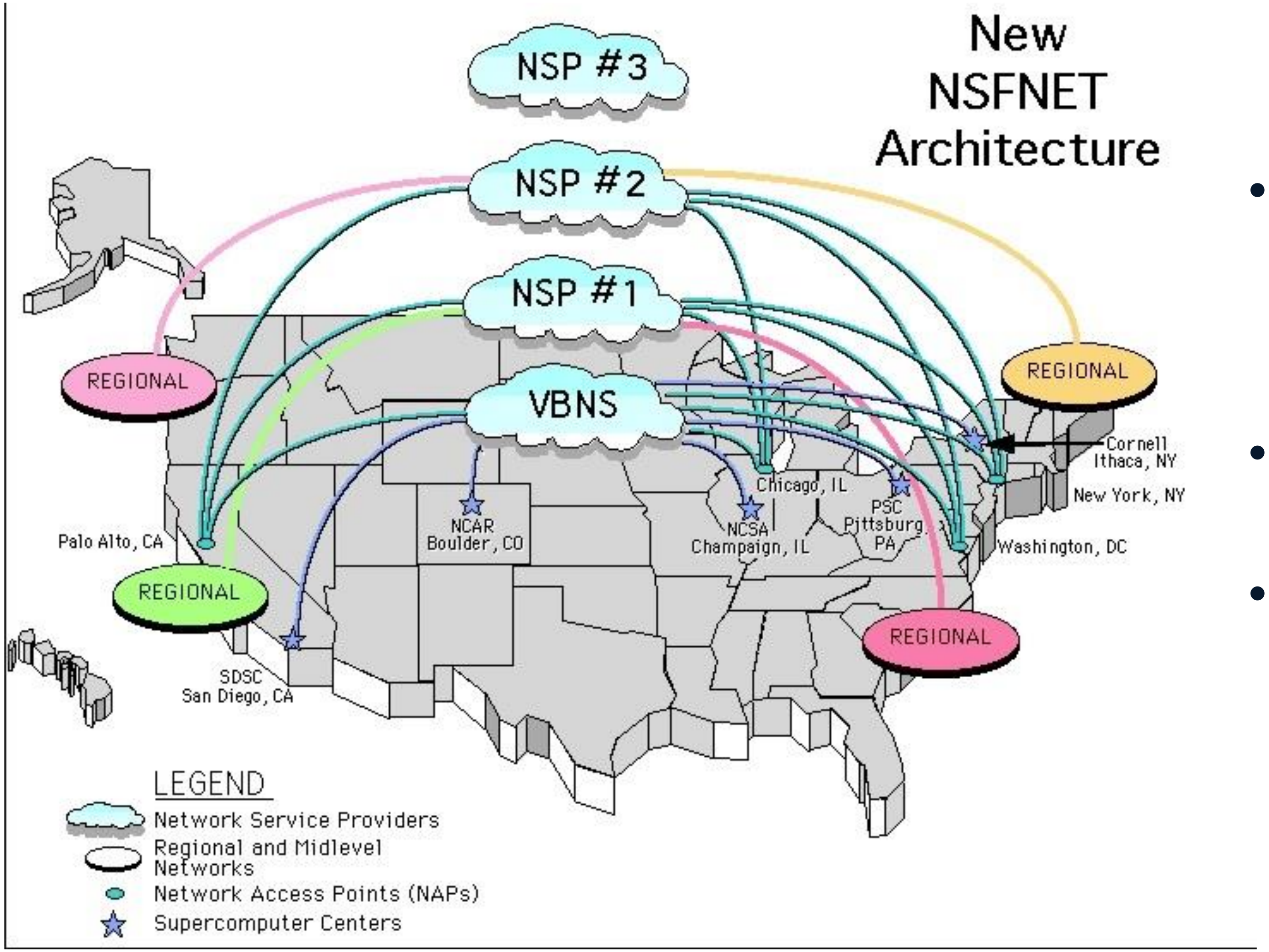
- **Cost** reduction

- **Performance** increase

- **Increased resiliency** by interconnection to multiple providers and multiple points of interconnection

Peering Growth and Distribution

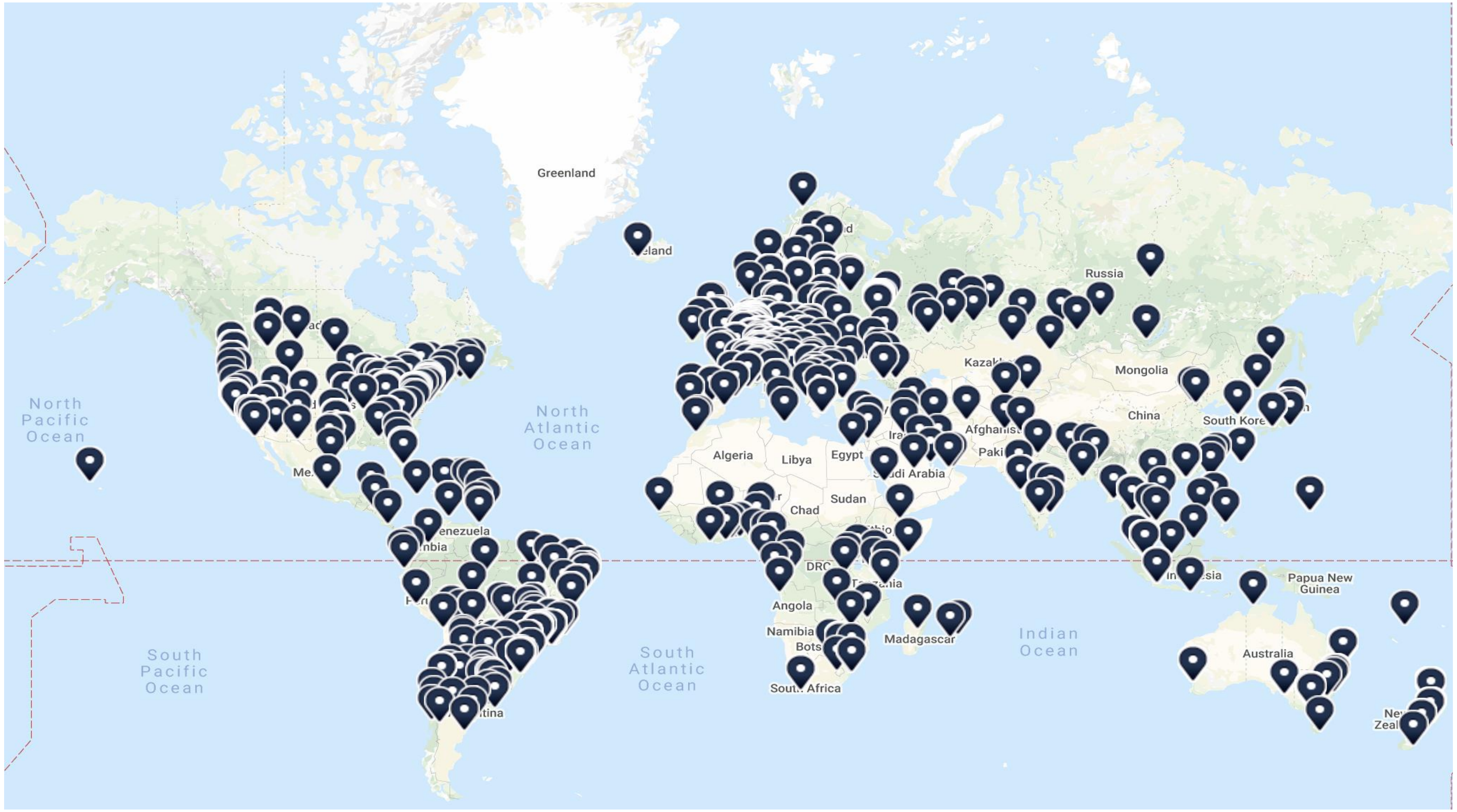
IXP History



- Coincides with Internet transition from public to commercial
- Original 4 US NAPs in Palo Alto (PacBell), New York (Sprint), DC (MFS), and Chicago (Ameritech)
- IXPs were being created in Europe as well
- CIX was first IX in Reston, VA in 1991, not one of the original NAPs

Interconnection growth over the years (2018)

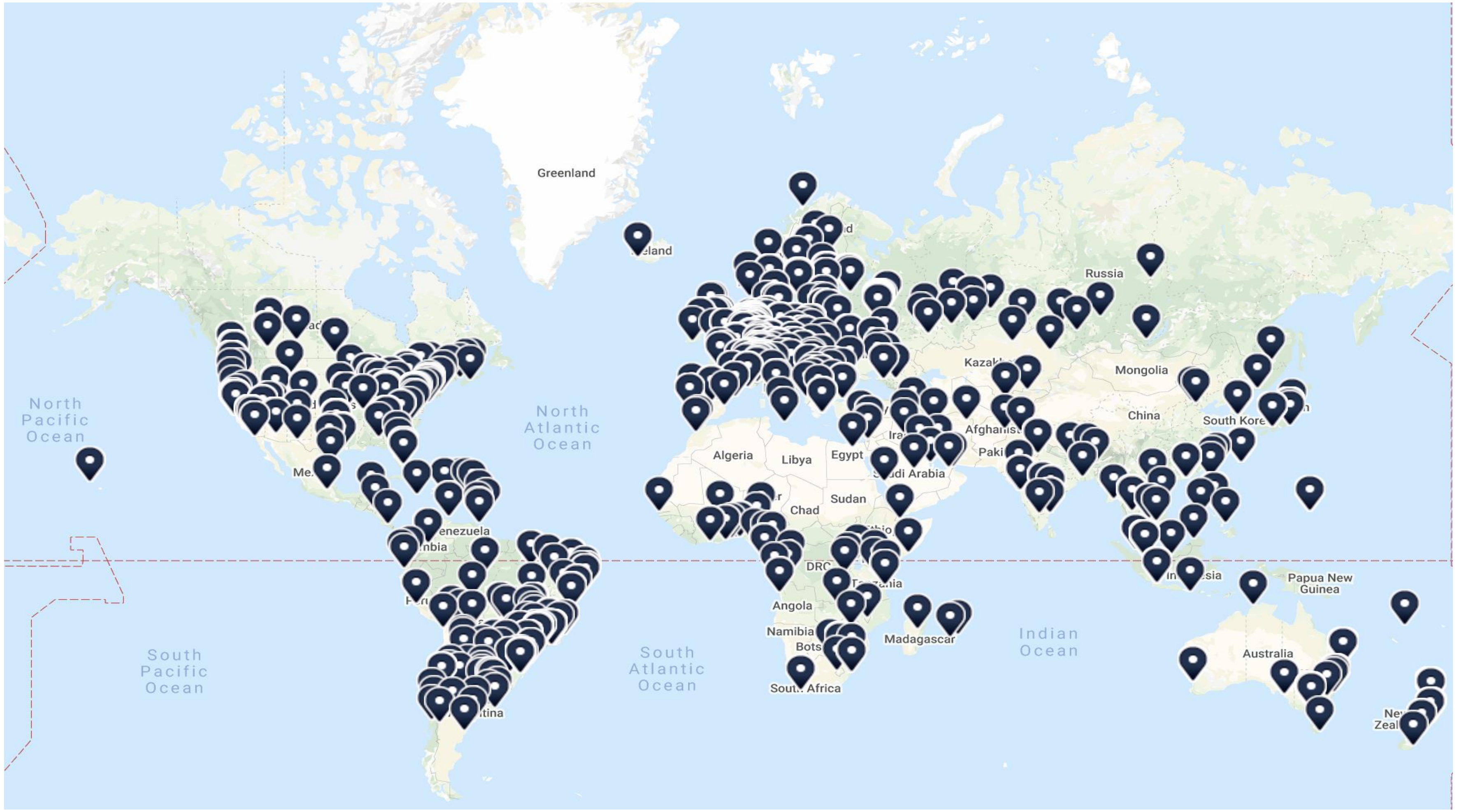
In 1995, ~20 Internet Exchanges, 2018 about 500 worldwide



http://internetexchangemap.com

Interconnection growth over the years (2019)

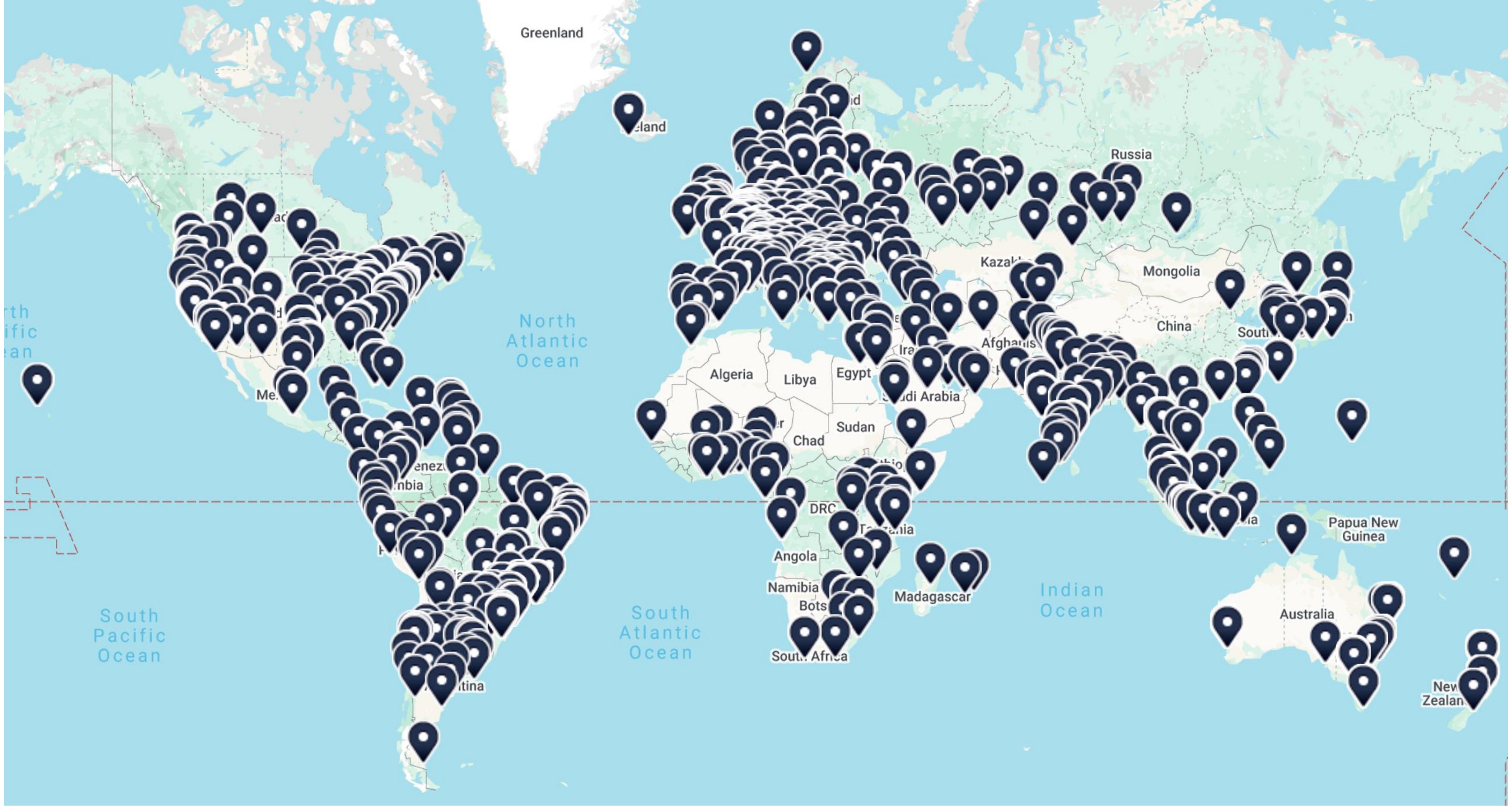
More than 700 IXs worldwide in 2019



http://internetexchangemap.com

Interconnection growth over the years (2025)

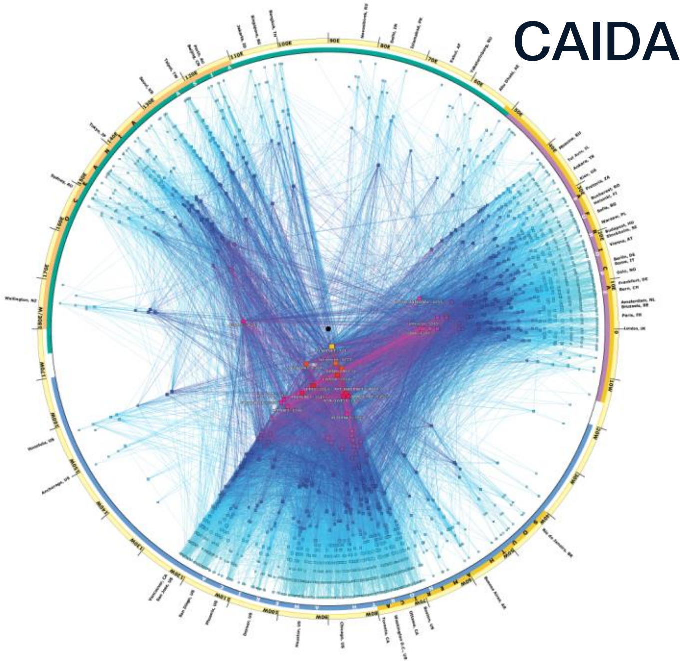
More than 1200 IXs worldwide in 2025



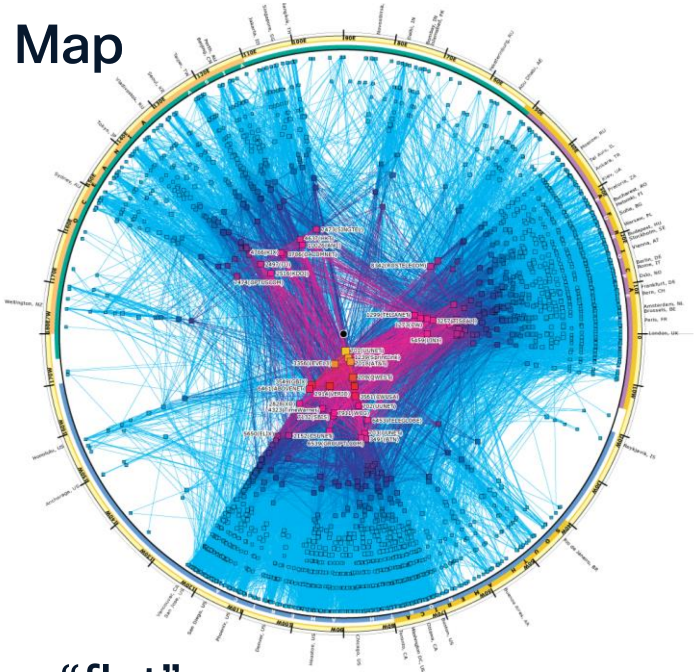
http://internetexchangemap.com

CAIDA Interconnection Map

2000

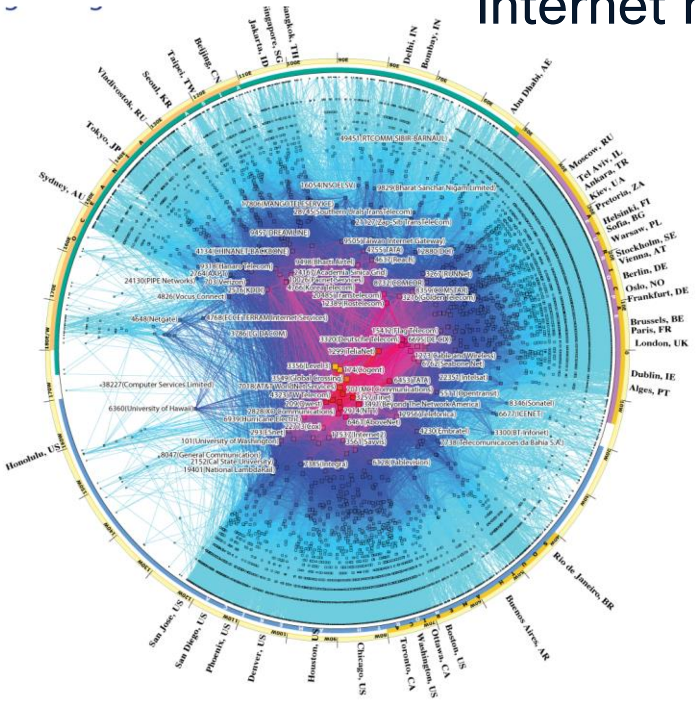


2005

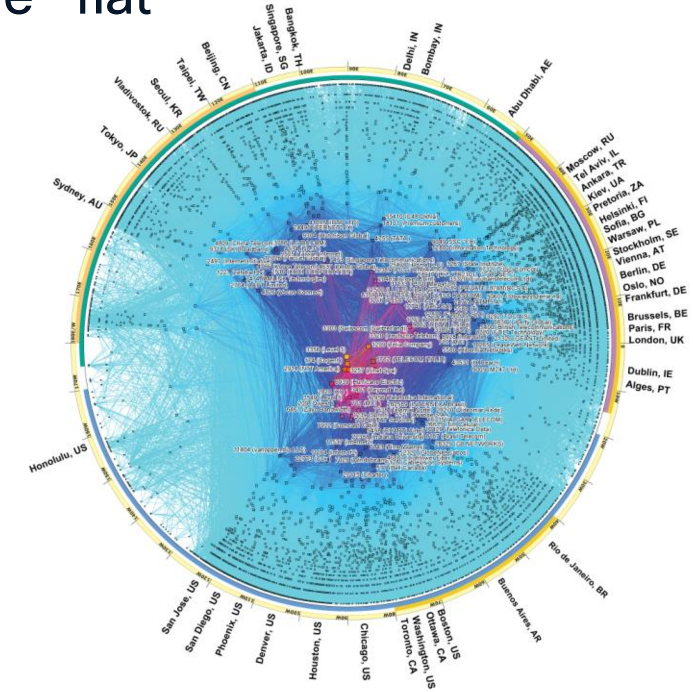


Internet has become more "flat"

2010



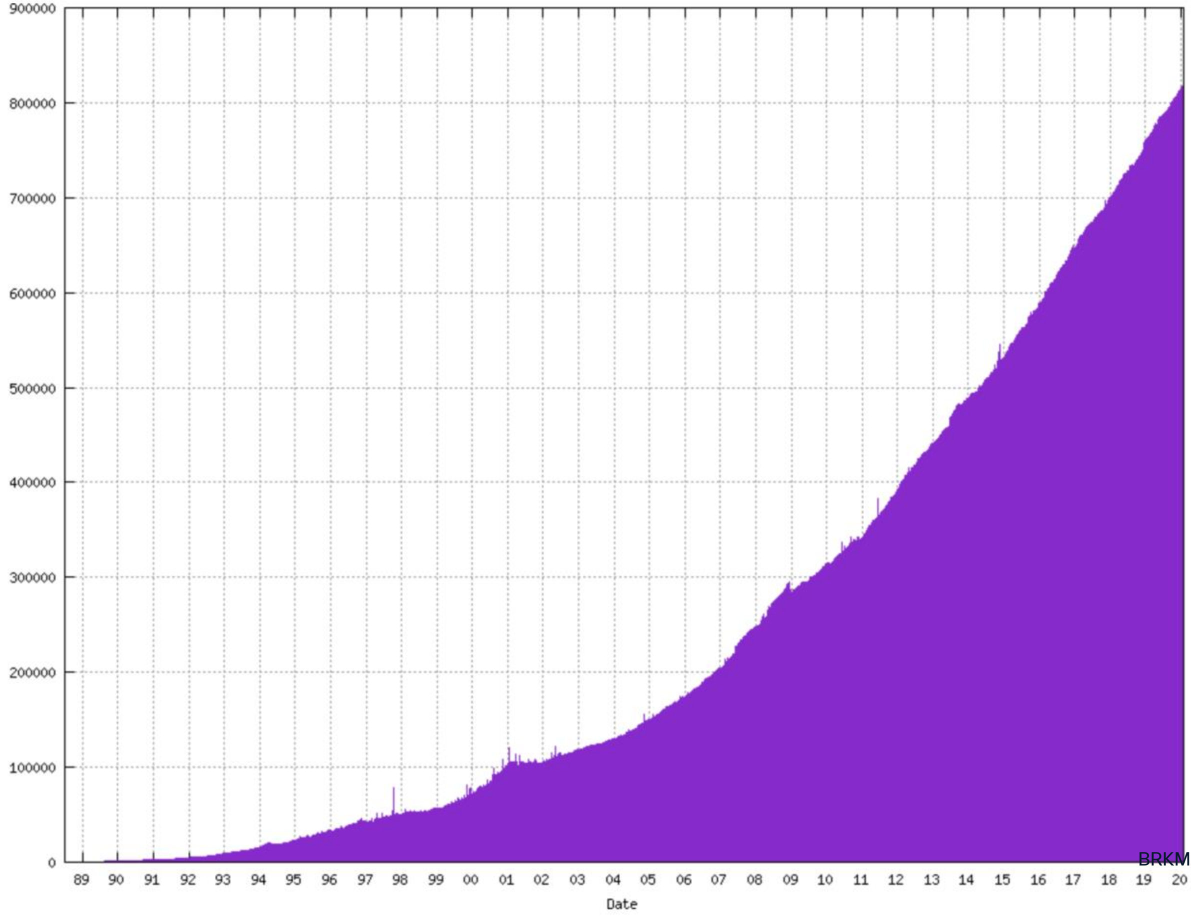
2017



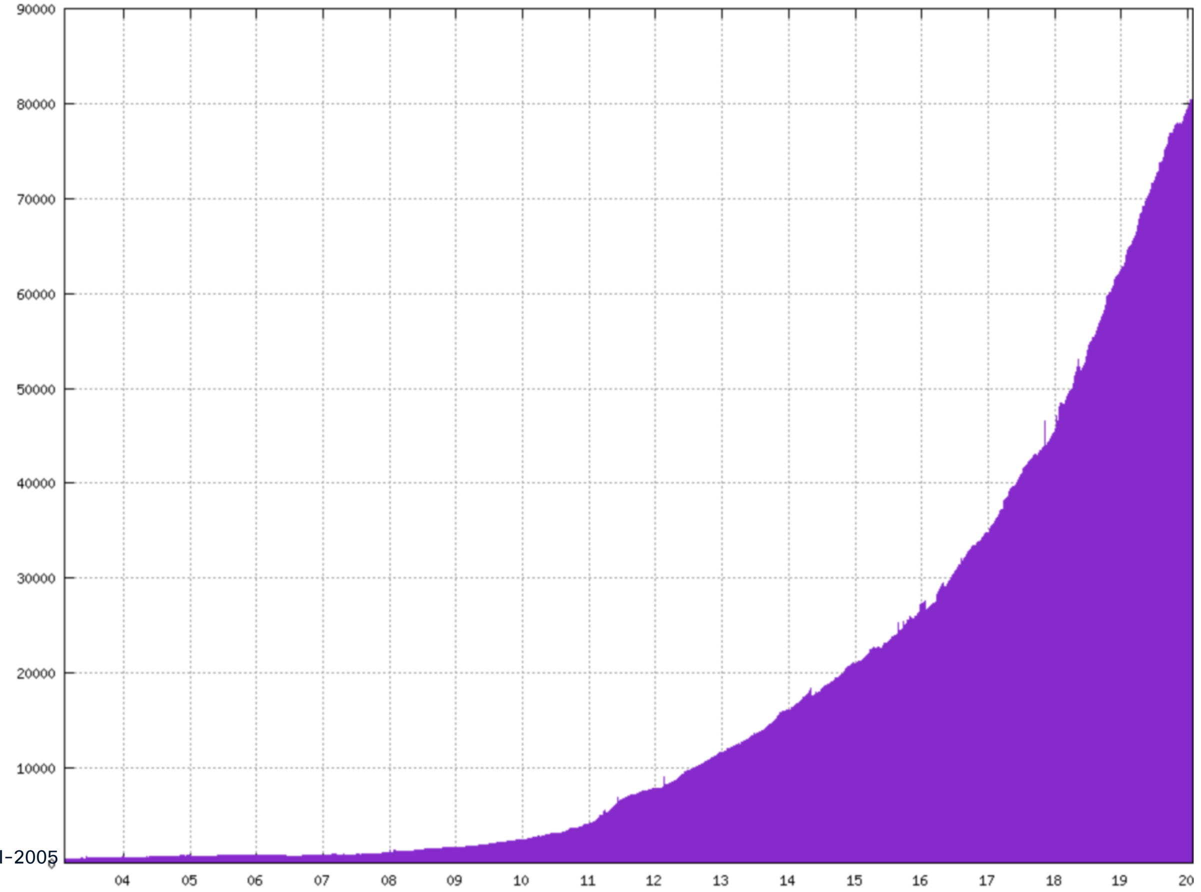
Internet Global Routing Table by numbers (2019)

67057 unique ASNs in global BGP routing table
817505 IPv4 prefixes, 80514 IPv6 prefixes

IPv4 Prefixes



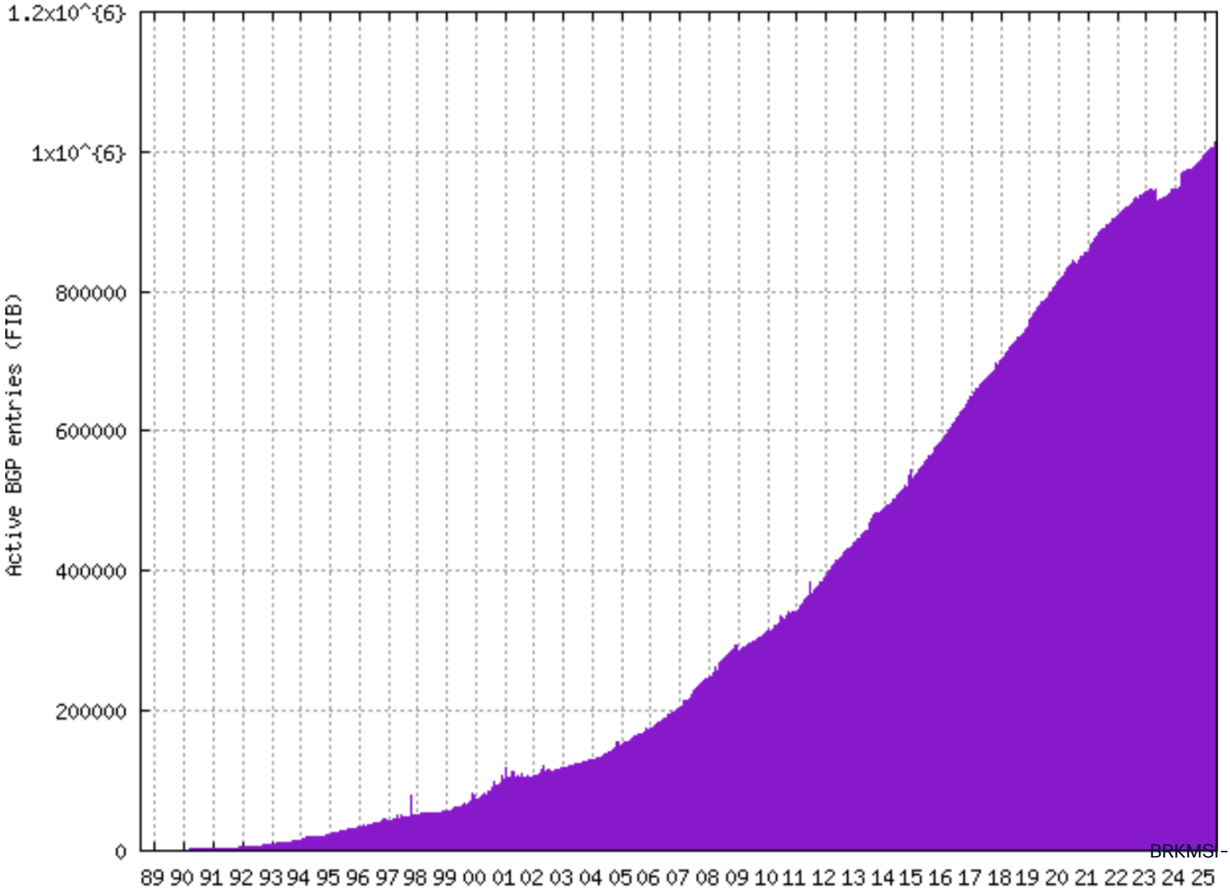
IPv6 Prefixes



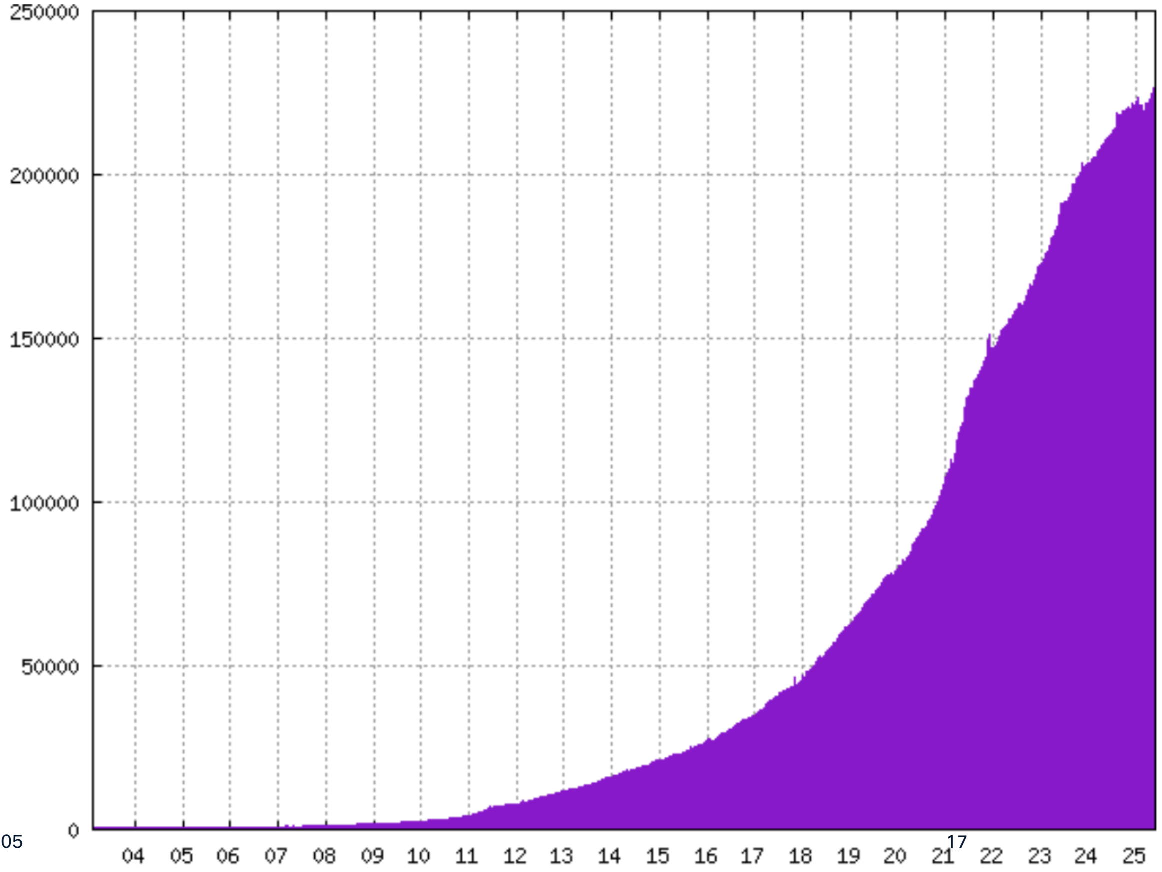
Internet Global Routing Table by numbers (2025)

77155 (+6948) unique ASNs in global BGP routing table
1010898 IPv4 prefixes (+193303), 226039 IPv6 prefixes (+145,525)

IPv4 Prefixes



IPv6 Prefixes

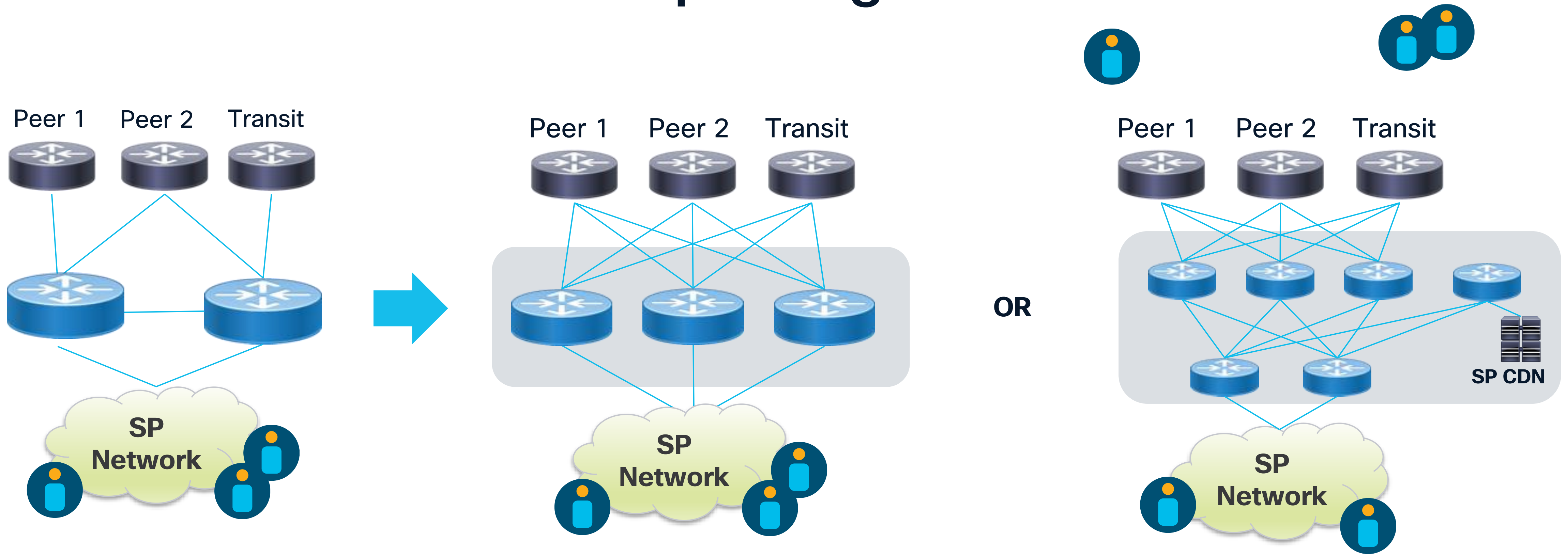


IPv4 prefix exhaustion

- IPv4 space is a commodity, providers are transitioning to IPv6 or renumbering IPv4 networks to have more unused IPv4 space
- Exhaustion and IPv6 transition has NOT happened as fast as originally thought, NAT continues to be widely used to kick the can down the road
- Geoff Huston from APNIC material
  - https://www.potaroo.net/ispcol/2025-01/addr2024.html https://youtu.be/9mSukwT19-U?si=ATO5tmXIQrr_mE_F

Peering network design and traffic engineering

Towards a more resilient peering fabric



Traditional Peering

- Horizontal scaling adds resiliency
- Less reliance on long-haul backup for metro or DC Peering
- Reduced blast radius during maintenance or failure
- Simplified SR control-plane

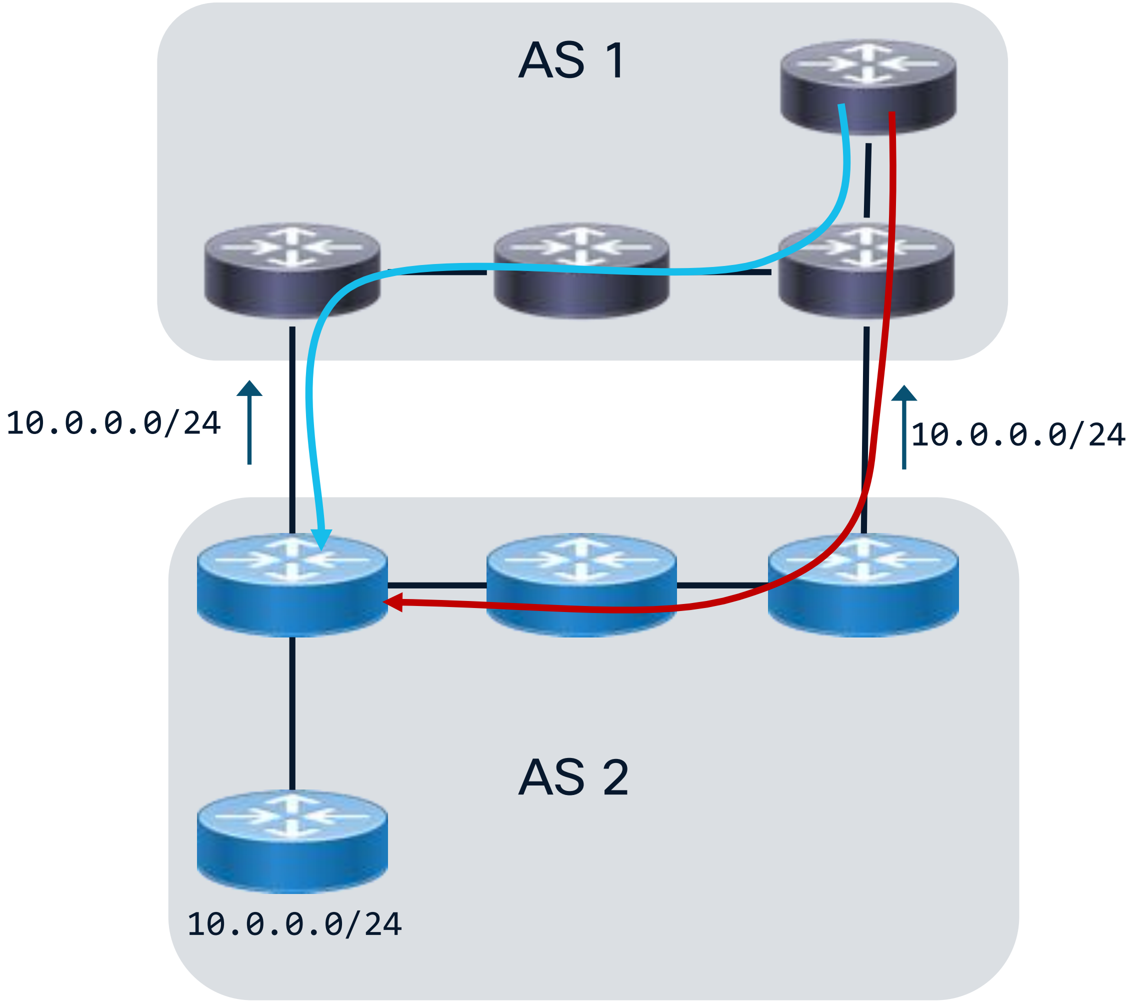
- Greater resiliency and capacity scale
- Optimized feature sets at each layer
- Optimized fabric for both ingress and egress content delivery

How do I influence peering traffic patterns?

Inbound Traffic	Detail
Prefix Advertisement	Suppression, longer prefixes
MED (multi-homed peer)	Some peers (transit) will listen to MEDs and carry traffic over their network to reach yours Typically set to IGP metric
AS Path Length	AS_PATH length influences peer route selection, prepending used for ingress TE

Outbound Traffic	Detail
Local Preference	Highest priority BGP attribute used for path selection
MED	”Metric” attribute also used in outbound path selection
TE Methods (SR-TE, RSVP-TE, EPE)	Steer traffic to specific location or peer using TE overlay methods

Hot potato vs. cold potato routing



- Hot potato (red) has routing policy to always route 10.0.0.0/24 to closest AS1->AS2 interconnect

- Cold potato (blue) carries traffic across AS1 network to AS1→AS2 interconnect point closest to final AS2 destination

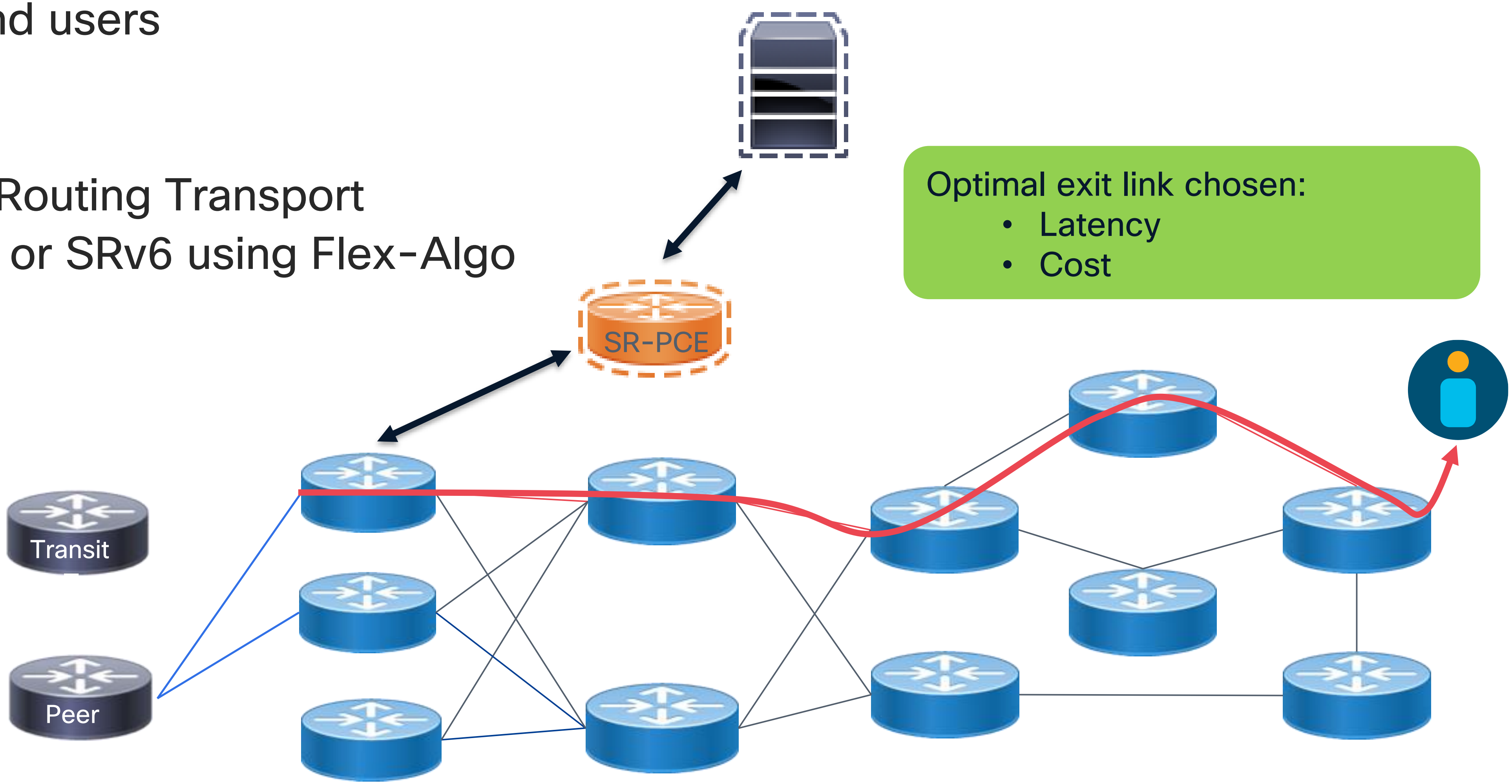
- Transit providers (paid) will typically use cold potato, peers will be use hot potato

SR-TE Ingress Peering Traffic Optimization

Problem: Engineering optimal path across SP network for ingress traffic from peering location to SP end users

Solution:

- ○ Segment Routing Transport
- ○ SR-MPLS or SRv6 using Flex-Algo



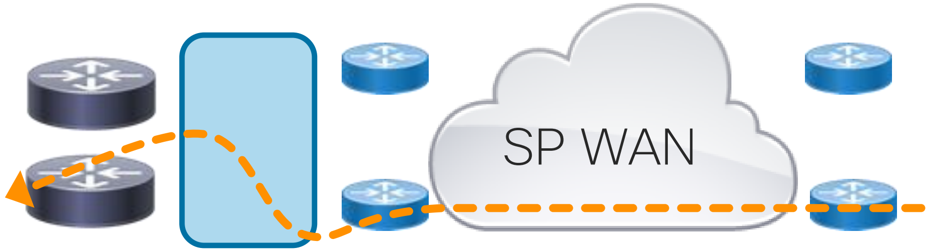
Optimal exit link chosen:

- Latency
- Cost

Optimization and SLAs drive Peering SDN

Egress

Best network exit path that is both cost-efficient and provides good user experience metrics (latency, link utilization & traffic loss)

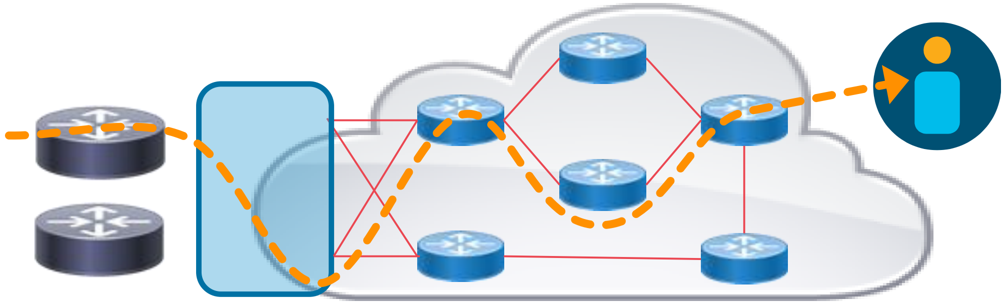


Optimal exit link chosen:

- Low cost (private peer)
- Low utilization link

Ingress

Optimal path across SP network for ingress traffic from peering location to SP end users



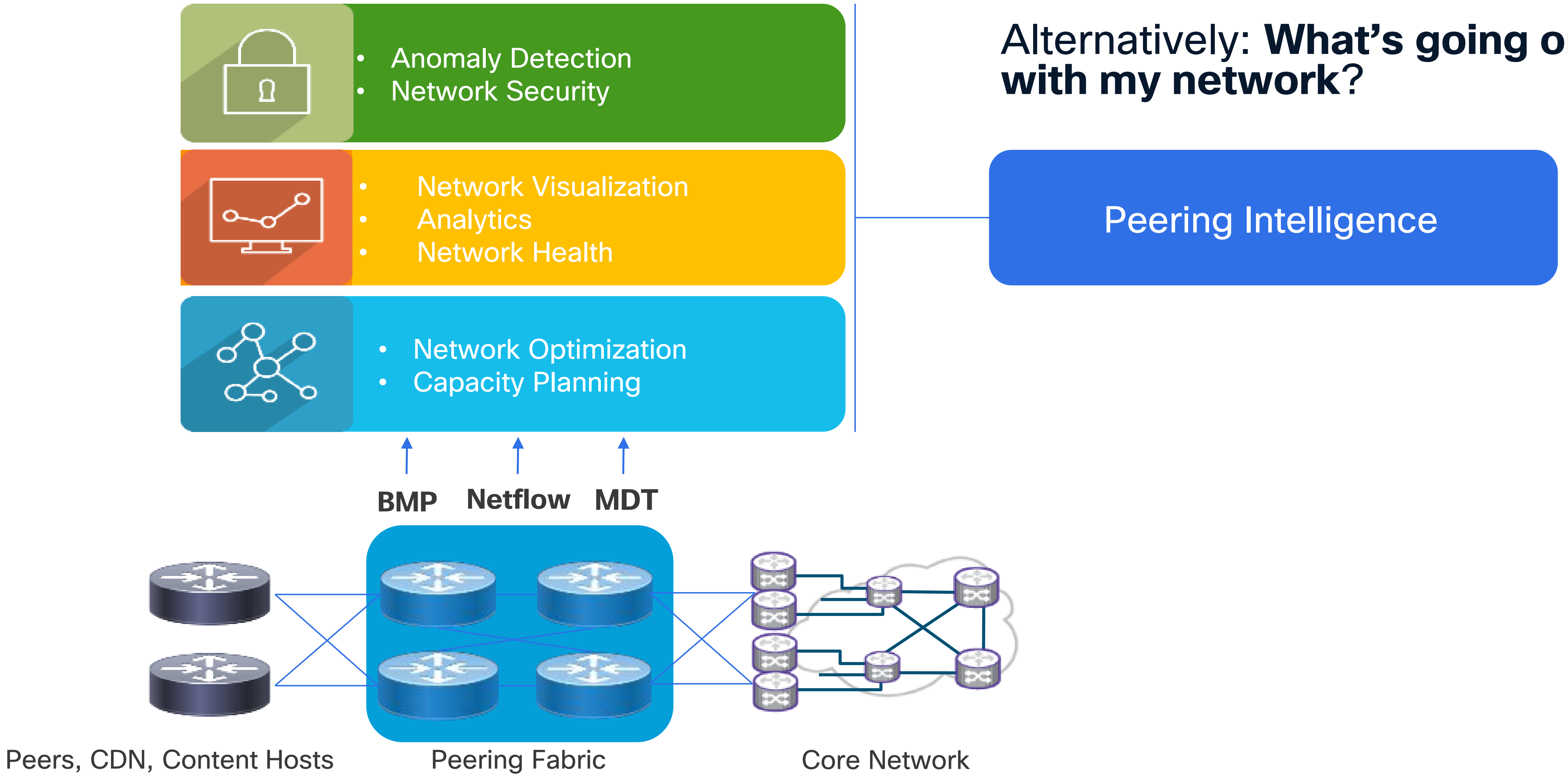
Optimal exit link chosen:

- Latency
- Cost

Peering Telemetry

Peering Data Provides Network Insights for Planning, Policy and Control

Alternatively: **What's going on with my network**?



- Anomaly Detection
- Network Security

- Network Visualization
- Analytics
- Network Health

- Network Optimization
- Capacity Planning

Peering Intelligence

Cisco Model-Driven Telemetry (MDT)

Periodic Streaming Telemetry

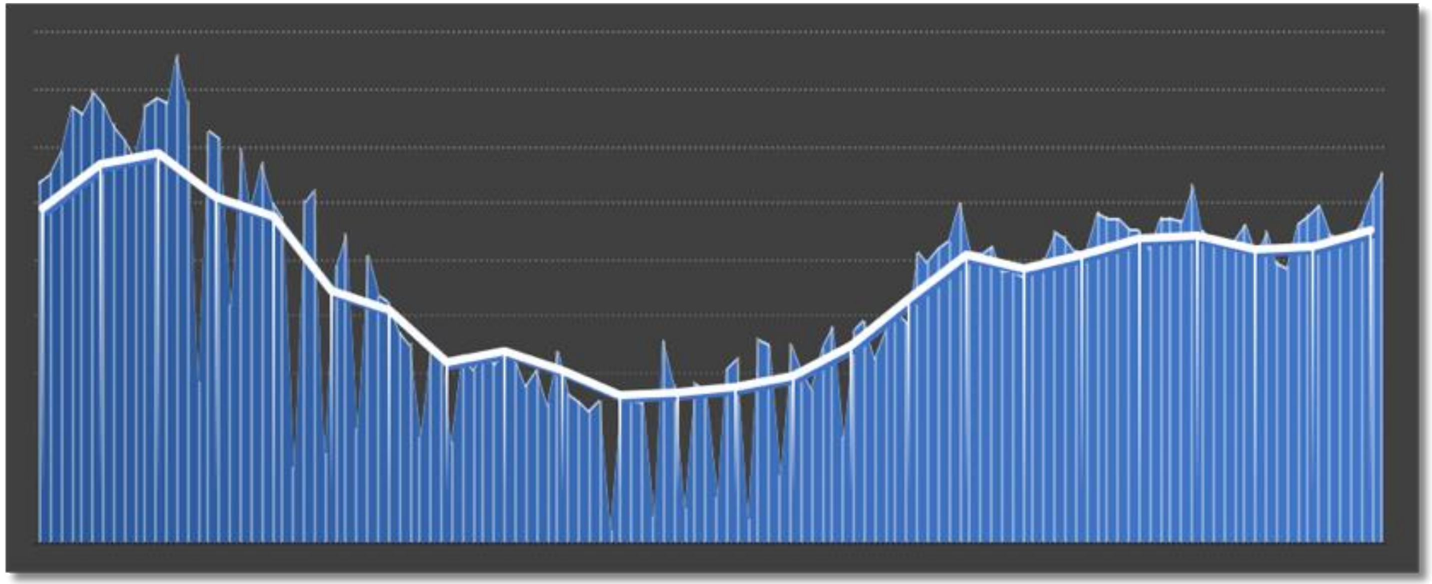
- Data is collected on node, “pushed” to collection entity at periodic intervals
- Cisco calls this model-driven telemetry (MDT)
- Best suited for time-series data, EG: interface statistics, router CPU
- Can also apply to network topology, EG: delay measurement between nodes
- Optimized data collection and optimized transport
- NETCONF/RESTCONF subscriptions can also be considered “streaming telemetry”

Event Driven Telemetry

- Data is pushed asynchronously from node based on state change or monitored event
- SNMP Traps, Syslog, Cisco EEM, Junos event scripts, and RMON are examples of existing event driven telemetry
- Modern approaches use YANG models and same structured encoding as periodic streaming telemetry
- BGP Monitoring Protocol (BMP) can also be thought of as event-driven telemetry

Model-Driven Telemetry for Peering

Higher Resolution Metric Data



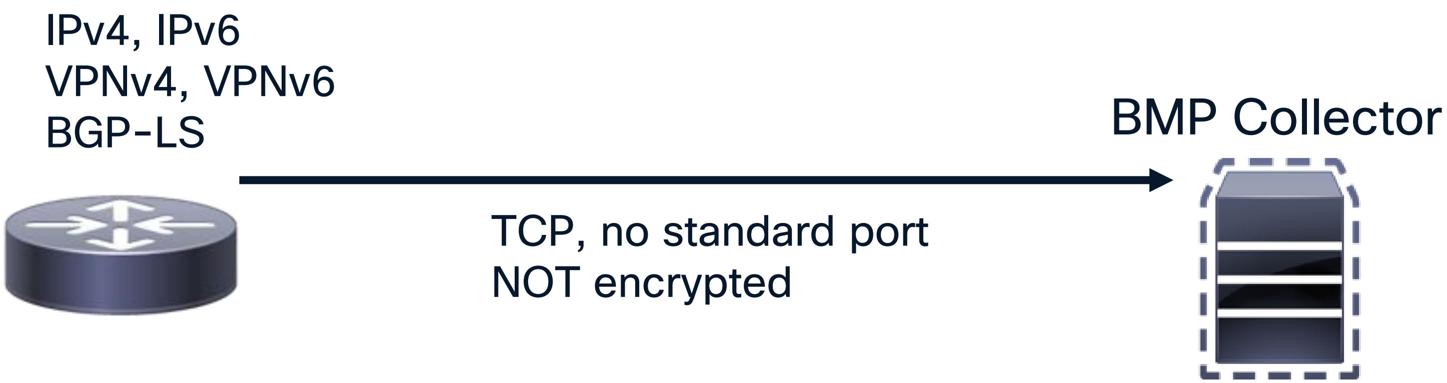
- Quickly detect anomalies when coupled with thresholds or machine learning
- Increased visibility into traffic patterns
- Expose hidden oscillations
- See instant impact of network changes or maintenance events

Network and Device Health Monitoring

- Monitoring queuing resources, can be important across peering or fabric where ingress/egress interfaces are the same speed. Similar in concept to datacenter microburst detection
- Monitor hardware FIB capacity and RIB memory

BGP Monitoring Protocol

Support in NX-OS, IOS-XR, and IOS-XE



BMP Message Type	Data
Route Monitoring	Per-peer NLRI and ongoing NLRI updates
Statistics Report	14 periodic stats values, EG: denied prefixes, RIB counts
Peer Down Notification	Peer down, includes local/remote notification msg
Peer Up Notification	Peer in Established state, includes open msg
Initiation Message	sysName, sysDescr, additional info
Termination Message	Termination reason, additional info
Route Mirroring	Exact copy of BGP message and context

BMP Security Use Cases and Resources

- Monitor peers and prefixes for instability
- Monitor peers for “bad” attributes such as invalid/private ASNs, long ASN lengths, internal communities, bogon prefixes etc.
- Forensic analysis of routing events, having a historical log of routing changes can be invaluable in root cause analysis
- Use diff from pre-policy to easily detect specific rejected prefixes

Netflow / IPFIX

- Has been around for many years
- Cisco Netflow v9 latest Netflow version
- IPFIX – IETF standard flow export

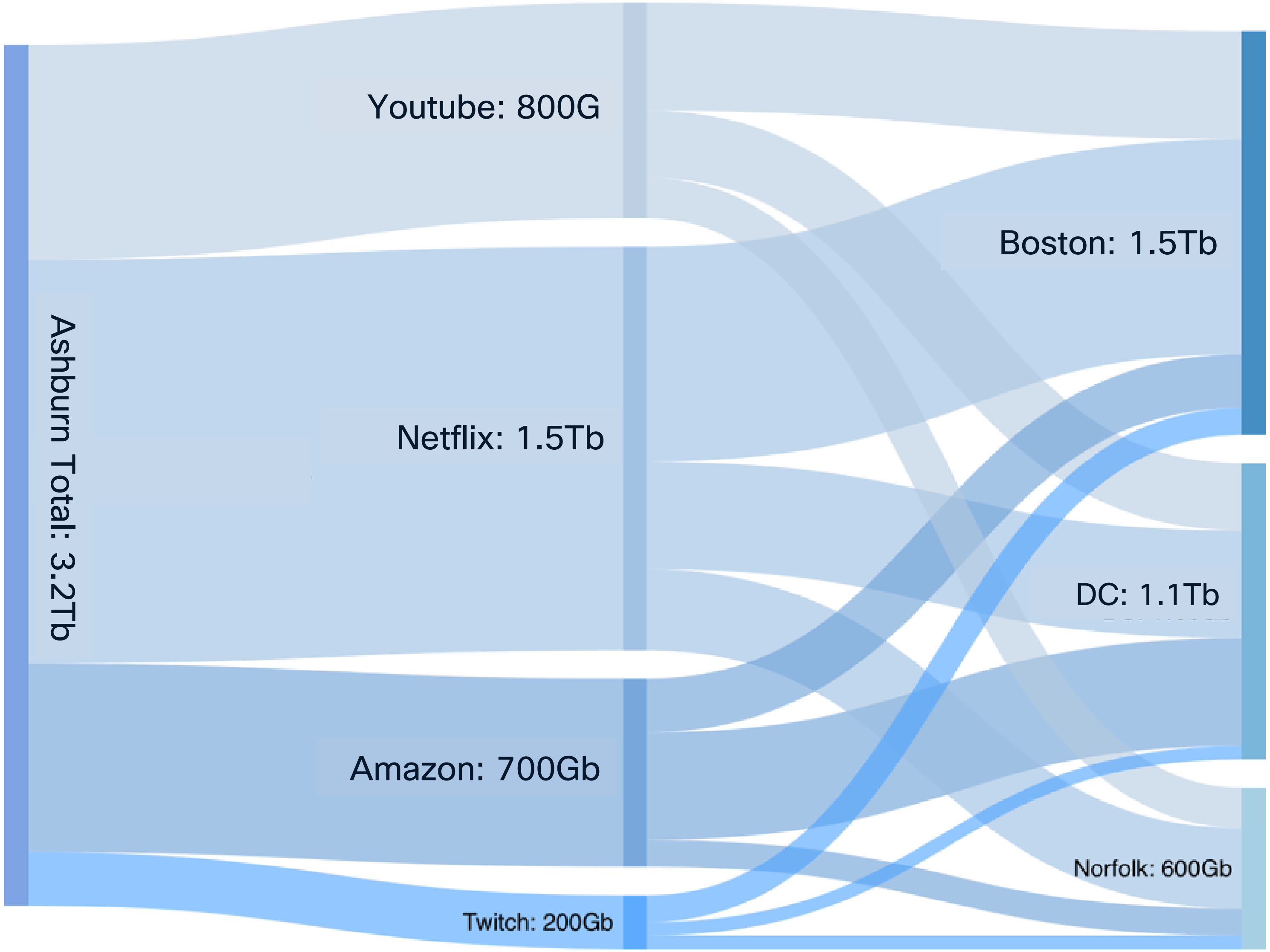
- Peering BGP data must be associated with flow information to be the most meaningful ***bgp attribute-download in XR***

- Modern traffic rates require sampling. **1:4000 is sufficient for accurate traffic modeling, but dimension for your network**
- Application-level visibility is becoming more difficult with encrypted traffic increasing, but peering data is only reliant on SRC/DST IP and still valid

Capacity planning use cases

- “Who should I peer with?”
- “Where should I peer with X,Y,Z?”
- “Should I build local peering or add caching to optimize my network?”
- “Should I change my network topology?”

Netflow / IPFIX – Use Cases



Derived data shows

- Add peering in Boston
- Specifically target Netflix and Youtube in Boston
- People in DC like Amazon Prime, maybe look to add targeted CDN

Crosswork Cloud - Traffic Analysis

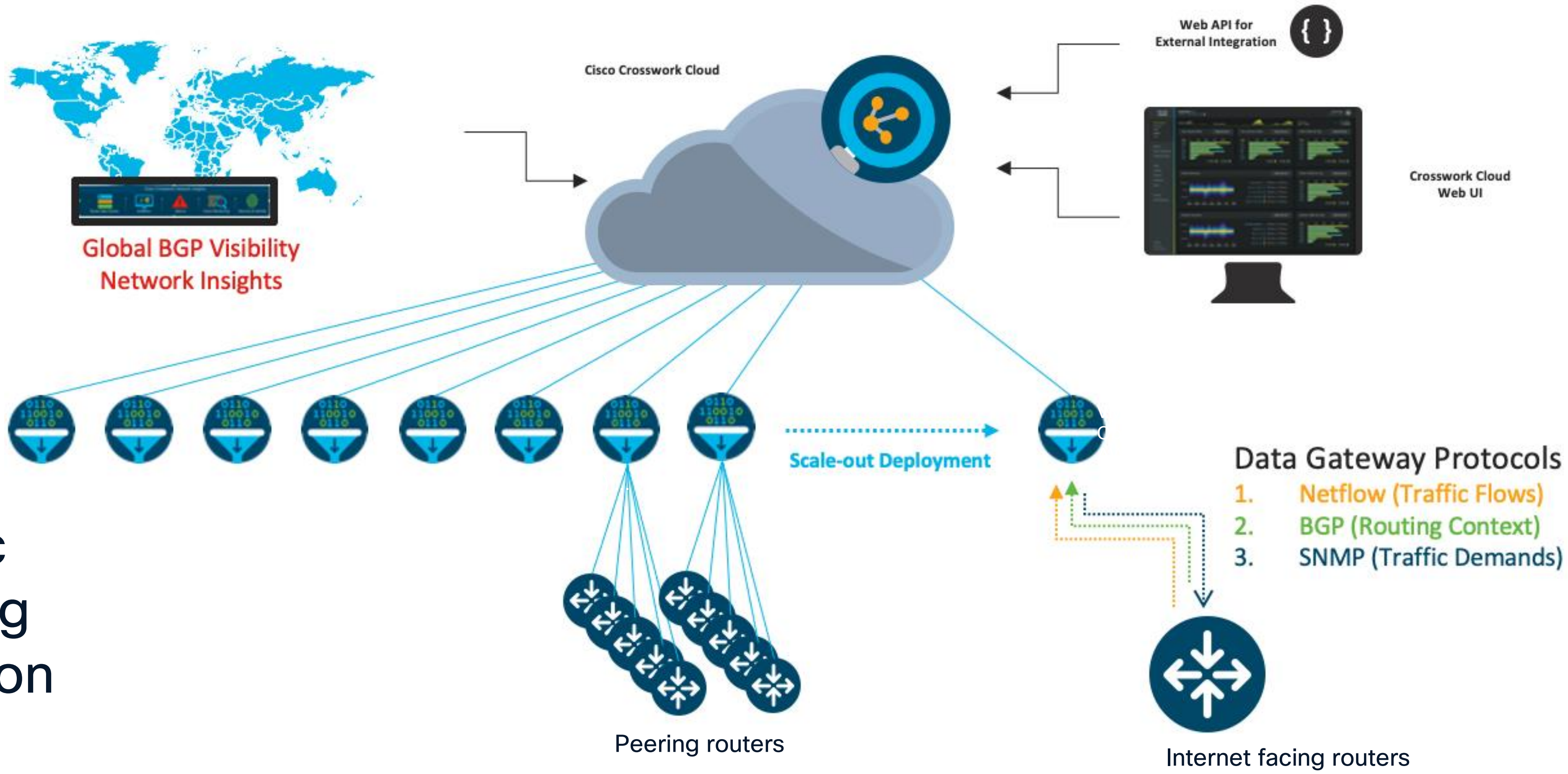
Top talkers

- ASN
- Prefix
- Device
- Interfaces

Tracking

- DoS/DDoS
- High Risk Traffic
- Peer Prospecting
- Edge Optimization

Multivendor Support



Data Gateway Protocols

1. Netflow (Traffic Flows)
2. BGP (Routing Context)
3. SNMP (Traffic Demands)

Peering Capacity Planning

1. Derive traffic matrix
   - SR Traffic Matrix
   - RSVP-TE tunnels
   - Netflow flow source router/interface to egress interface

2. Develop network growth model
   - Use historical data to grow interfaces and links at realistic rates, not the same rate across all links
   - Machine learning, or humans, can add intelligence to the model over time. Filter anomalies and predict seasonal changes

3. Simulate network failures
   - Balance cost vs. consumer experience and SLAs

Crosswork Planning

Cisco Peering Telemetry Open Source

Application	Collection Method	Use Cases
telegraf	Model-Driven Telemetry Open-source collector with Cisco MDT plugins in mainline release	Collect, process, and output router telemetry GNMi or static configuration Input gRPC,JSON telemetry data Output to Telegraf supported streams (Kafka, InfluxDB, etc.)

How do I get started?

Model-Driven Telemetry

- IOS-XRv virtual routers support MDT
- Many good telemetry blogs on https://xrdocs.io
- Telegraf plugins part of master branch for both gNMI and Cisco native telemetry formats

BMP

- SNAS, formerly OpenBMP, available at https://snas.io, no longer actively maintained, but can output to local database or KAFKA
- PMACCT has a BMP collector which can then output to different destination types

Peering Security

Peering Edge Security Threats

Leading Threat Concerns*	Description
DDoS Attack (88%)	Distributed Denial of Service Volumetric traffic to overwhelm network and hosts
Infrastructure Security (55%)	Compromise of network control-plane Compromise of network devices
BGP Route Hijacking (25%)	Man-in-the-middle attack ASN hijacking has also been an issue

Peering Security Events

2024. *Cloudflare 1.1.1.1 BGP hijack and route leak incident affects users of their DNS services*

2024. *Three BGP events affected research networks, two leaks and one hijack event*

2024. *Overall >12,000 BGP route leak incidents, most not malicious*

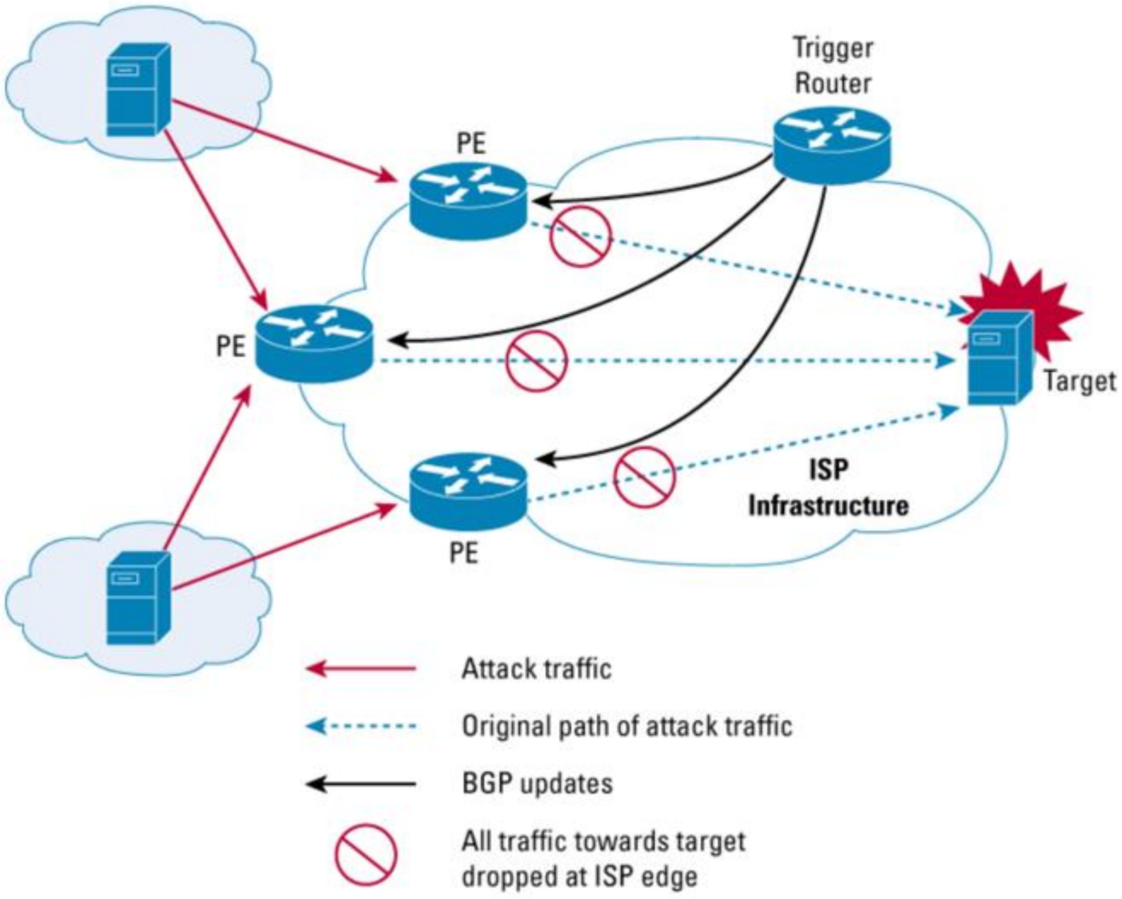
2024. *"Global" route leak incidents are much lower, around 25, but cause more widespread issues*

2024. *Appoximately 50,000 route hijack events*

2024. *Overall DDoS attacks number in the millions, with volumetric attacks still the most prevalent*

Peering DDoS Mitigation - RTBH

- Remote Triggered Black Hole
  - Applicable for content, SP, enterprise
  - Black hole could be sinkhole, honey pot
- S/RTBH
  - Drop based on source address and not destination
  - Uses Unicast RPF with BGP NH set to /32 with static route to Null0
- Upstream providers will often match specific community to allow customers to trigger RTBH (see resources for more info)
- Cymru has UTRS, global RTBH network


Trigger Router
PE
PE
PE
Target
ISP Infrastructure
Attack traffic
Original path of attack traffic
BGP updates
All traffic towards target dropped at ISP edge

https://www.cisco.com/c/dam/en_us/about/security/intelligence/blackhole.pdf

Peering DDoS Mitigation – BGP Flowspec

- New AFI/SAFI NLRI, IPv4 defined in RFC5575, IPv6 nearing RFC status
- Distribute ingress data-plane filtering using MP-BGP
- Match on packet criteria then drop, police, redirect, or mark matched traffic
- XR server, XR / XE clients
- Foundation for scalable distributed DDoS protection
- See BRKSPG-3012 for deep dive
- BGP FSv2 is under active development in the IETF

Server Config

```
class-map type traffic match-all memcached
 match destination-port 11211
 match protocol udp tcp
 end-class-map
!
policy-map type pbr drop-memcached
 class type traffic memcached
  drop
 !
 class type traffic class-default
 !
 end-policy-map
!
  flowspec
   address-family ipv4
    service-policy type pbr drop-memcached
```

Client Config

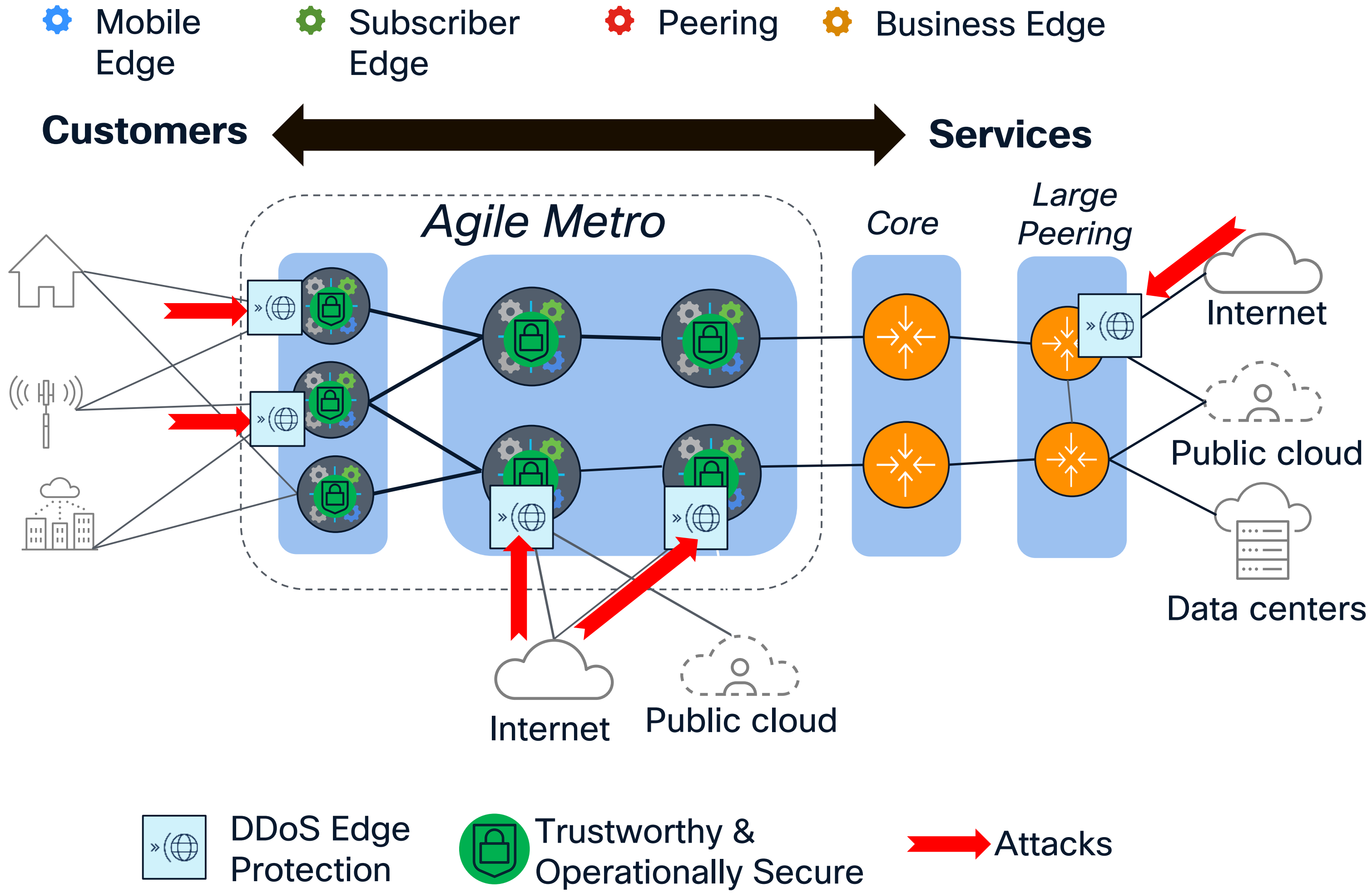
```
flowspec
 address-family ipv4
  local-install interface-all
```

Cisco Secure DDoS Edge Protection

Up to

83%

TCO Savings*



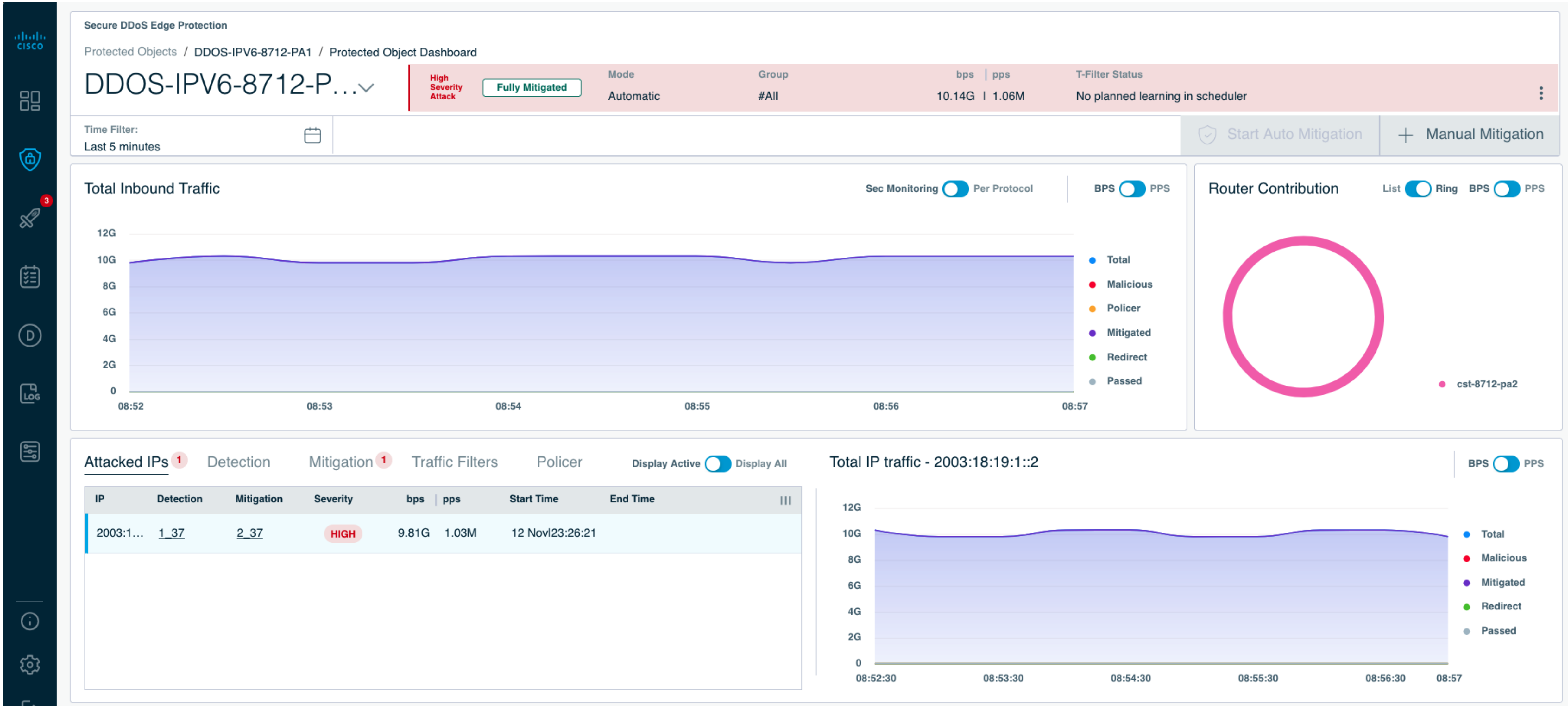
‘Defense in Depth’ – threat free network with secured perimeter

No dedicated scrubber
No traffic diversion
No additional power & space

Faster detection & Zero Day attack mitigation

*Comparison for 4Tbps Peering Network

Cisco Secure DDoS Edge Protection



Secure DDoS Edge Protection

Protected Objects / DDOS-IPV6-8712-PA1 / Protected Object Dashboard

DDOS-IPV6-8712-P...

High Severity Attack | Fully Mitigated

Mode	Group	bps | pps	T-Filter Status
Automatic	#All	10.14G | 1.06M	No planned learning in scheduler

Time Filter: Last 5 minutes

Start Auto Mitigation | + Manual Mitigation

Total Inbound Traffic — Sec Monitoring / Per Protocol — BPS / PPS

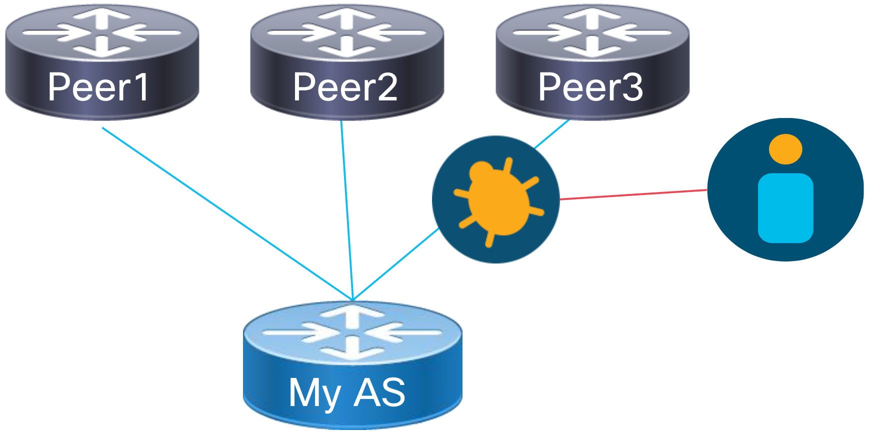
Router Contribution — List / Ring — BPS / PPS

Attacked IPs 1 | Detection | Mitigation 1 | Traffic Filters | Policer — Display Active / Display All

IP	Detection	Mitigation	Severity	bps | pps	Start Time	End Time
2003:1...	1_37	2_37	HIGH	9.81G | 1.03M	12 NovI23:26:21	

Total IP traffic - 2003:18:19:1::2 — BPS / PPS

Increasing BGP Session Security with TCP-AO



- Session threats
  - TCP RST attacks
  - Snooping
  - SYN flooding
  - Peering is being used for more critical applications than just best-effort Internet

- Question: When was TCP MD5 authentication obsoleted?
  - Answer: Obsoleted in 2010

- TCP-AO – TCP Authentication Option – RFC 5925
  - Use HMAC-SHA2-256 hash at minimum
  - Protects BGP TCP connection by authenticating TCP segments
  - Does NOT provide session encryption
  - Supported in IOS-XR in 6.5.3, IOS-XE in 16.12
  - Recommended in RFC 7454 (2015)

TCP-AO IOS-XR Configuration

Key chain and TCP AO Config

```
tcp ao
 keychain TCP-AO-KEY
  key 1 SendID 100 ReceiveID 100
 !
!
key chain TCP-AO-KEY
 key 1
  accept-lifetime 00:00:00 january 01 2018
infinite
  key-string password 0204034B0A131B29
  send-lifetime 00:00:00 january 01 2018 infinite
  cryptographic-algorithm AES-128-CMAC-96
```

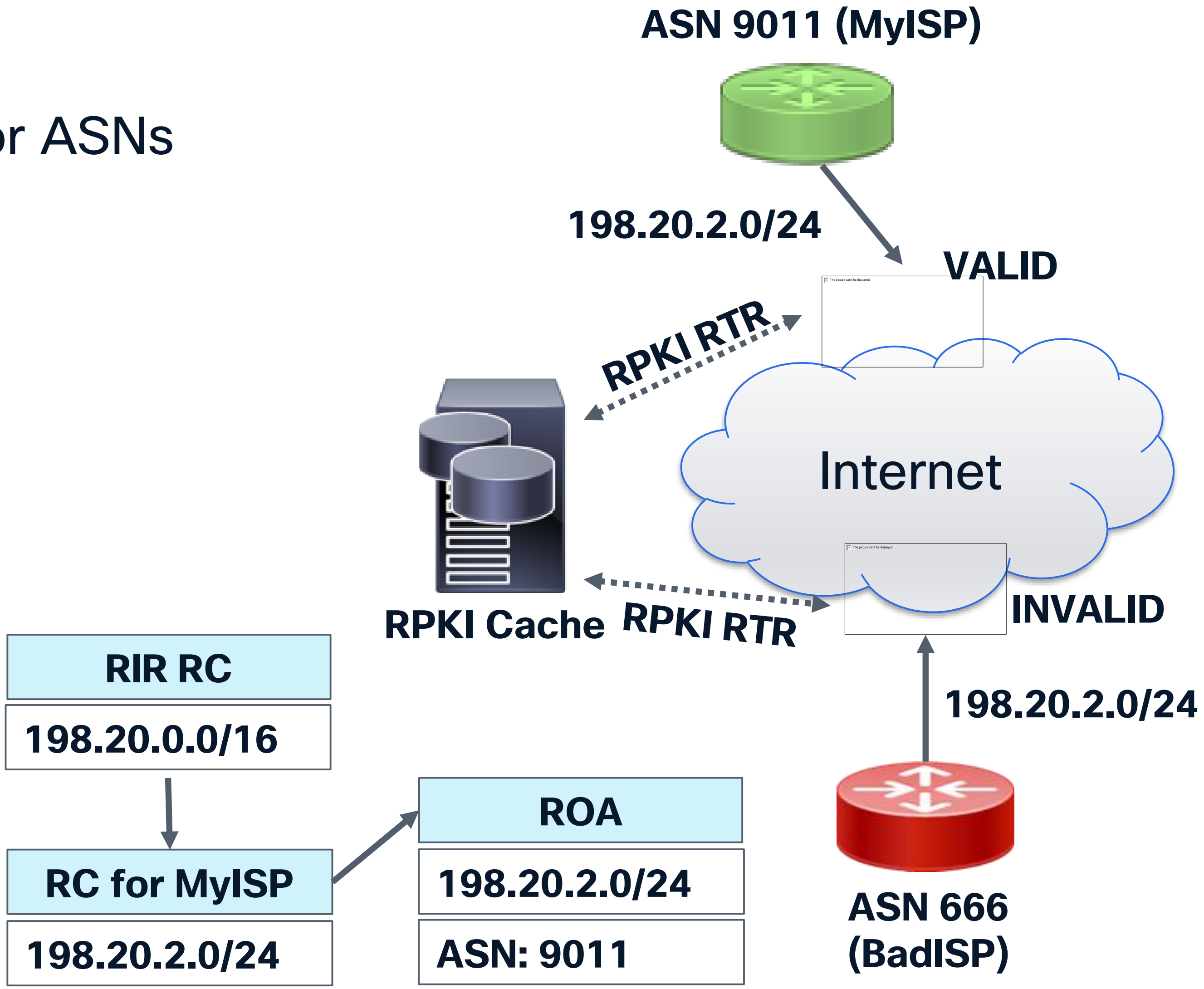
BGP Configuration

```
router bgp 100
 neighbor 1.2.3.4
  remote-as 101
  ao TCP-AO-KEY include-tcp-options enable
```

Troubleshooting: *`show tcp authentication keychain all detail`*

RPKI and Route Origin Validation (RFC 6483)

- Resource Public Key Infrastructure
- Route Origin Authorization is issued for ASNs originating prefix (SP, DDoS service)
- Validates origin ASN to stop hijacking
- Supported in IOS-XR and IOS-XE



IOS-XR RPKI and ROV configuration – using routing policy

```
route-policy rpki
 if validation-state is invalid then
   set local-preference 50
 else if validation-state is valid then
   set local-preference 200
 else
   pass
 endif
end policy
!
router bgp 65536
 bgp router id 192.168.0.1
 rpki cache 172.16.0.254
  transport tcp 32000
  refresh-time 120
!
address-family ipv4 unicast
 bgp origin-as validation signal ibgp

neighbor 192.168.0.254
 remote-as 64555
  address-family ipv4 unicast
   route-policy rpki in
```

- RPKI information is cached on the router
- Periodically polls for new data, RPKI cache also sends notification when it has been updated
- ROA is not automatically checked, requires route policy
- IBGP attribute to convey validity using extended community
  - Not recommended
- Based on scale, enable “soft-reconfiguration inbound always” so full route refreshes are not signaled to peers based on RPKI table changes

IOS-XR RPKI and ROV configuration – best path selection

```
router bgp 65536
 address-family ipv4 unicast
  bgp bestpath origin-as use validity
address-family ipv4 unicast
  bgp bestpath origin-as use validity
```

- Changes best-path selection algorithm to make routes with invalid state less preferred
- Originally also made “unknown” routes less preferred but behavior was changed so valid and unknown have the same preference
- Uses less memory / CPU since refresh is not performed when ROA is updated

https://www.cisco.com/c/en/us/support/docs/ip/border-gateway-protocol-bgp/217020-bgp-rpki-with-xr7-cisco8000-whitepaper.html

RPKI / ROV Status

- 2020 Update
  - ~18% valid prefixes in the IPv4 GRT, up from ~10% in 2017*
  - Larger providers and IXPs are now dropping “Invalid” prefixes

- 2025 Update
  - 56% valid prefixes in the IPv4 GRT
  - 58% valid prefixes in the IPv6 GRT

- Who has generated ROAs? Who is performing Route Origin Validation?

**https://rpki-monitor.antd.nist.gov*

AS Provider Authorization (ASPA)

ASPAs are digitally signed objects that bind a selected AFI Provider AS number to a Customer AS number (in terms of BGP announcements not business), and are signed by the holder of the Customer AS. An ASPA attests that a Customer AS holder (CAS) has authorized a particular Provider AS (PAS) to propagate the Customer's IPv4/IPv6 announcements onward

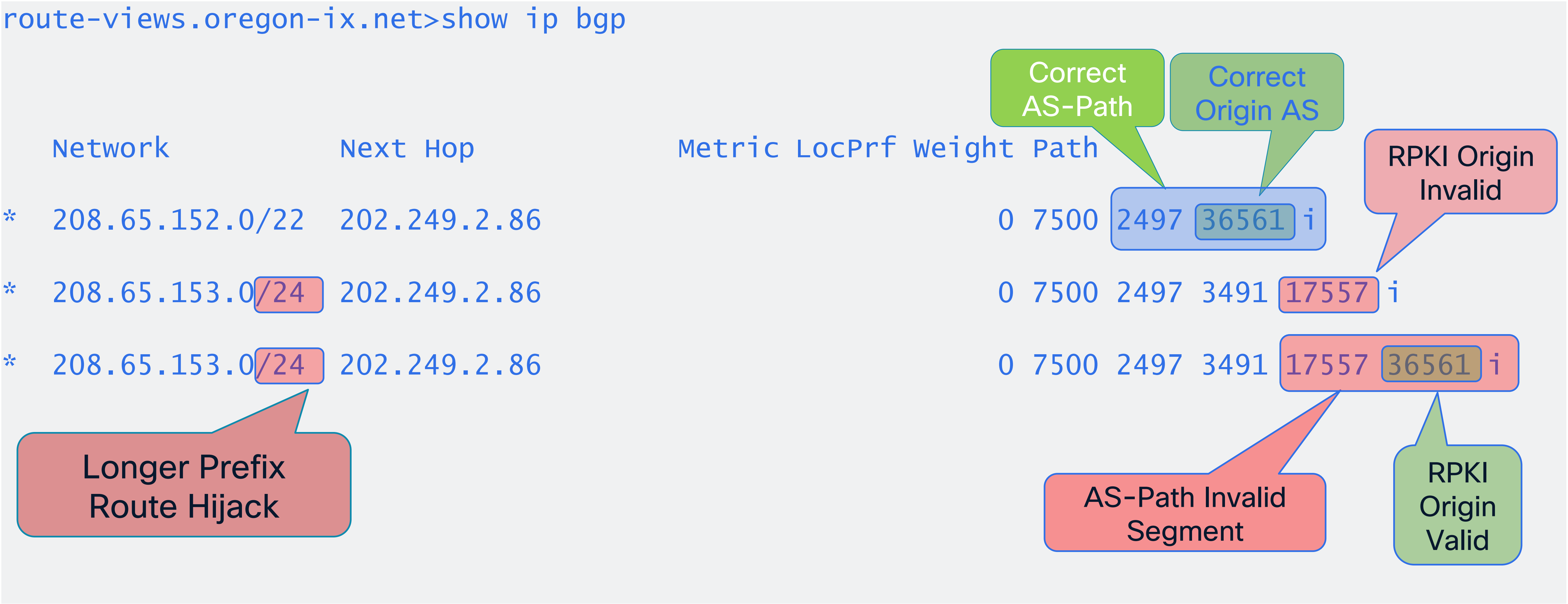
- There are 65000+ ASNs, 55000 are stub ASNs, an AS-PATH should not have a stub in the middle
- An AS that wants protection publishes an attestation of who its transit providers are to RPKI as ASPA object.
- ISPs use these ASPAs to detect the invalid AS in the middle of the AS path similar to ROA

https://tools.ietf.org/html/draft-ietf-sidrops-aspa-verification
https://tools.ietf.org/html/draft-ietf-sidrops-aspa-profile

AS Provider Authorization (ASPA) Operation

- BGP ASes have 2 types of relationships: transit-customer or peer-peer (sibling).
- A neighbor of an AS can be either transit provider, peer or customer.
- If an AS receives a route from a non-customer and sends it to a non-customer, then it is leaking that route.
- **Cisco has working code to support ASPA today**
- Routinator has ASPA support in latest builds
- If you are interested in a demo or IOS-XR with PoC version email iosxr-aspa@cisco.com or reach out to me

BGP Route Hijack Mitigated by ASPA



```
route-views.oregon-ix.net>show ip bgp

   Network          Next Hop            Metric LocPrf Weight Path
*  208.65.152.0/22  202.249.2.86                       0 7500 2497 36561 i
*  208.65.153.0/24  202.249.2.86                       0 7500 2497 3491 17557 i
*  208.65.153.0/24  202.249.2.86                       0 7500 2497 3491 17557 36561 i
```

Correct AS-Path

Correct Origin AS

RPKI Origin Invalid

Longer Prefix Route Hijack

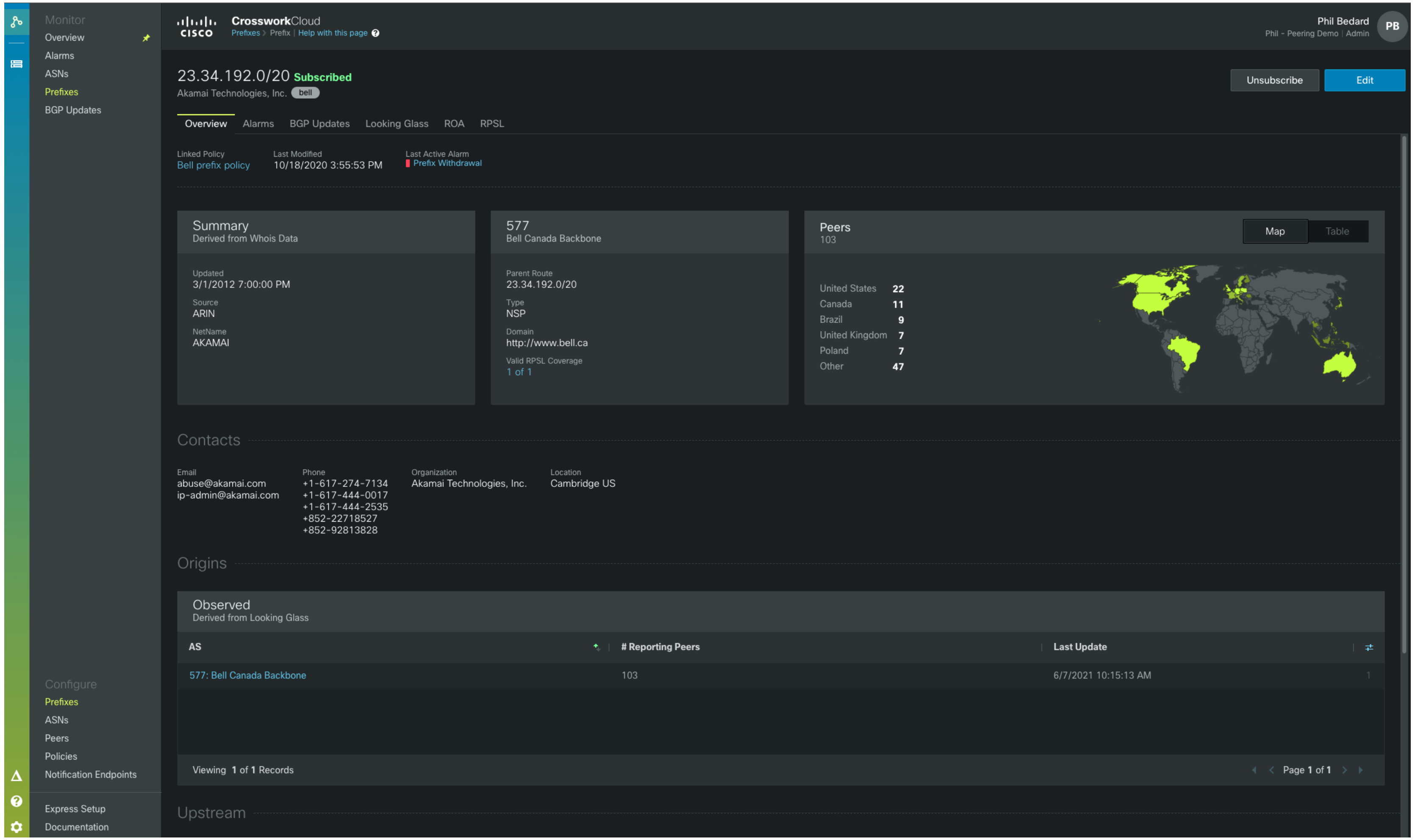
AS-Path Invalid Segment

RPKI Origin Valid

- Customer AS 36561 attests to 2497 as its provider via ASPA entry, but NOT 17557
- Prefix is INVALID

Crosswork Cloud Network Insights
External BGP prefix security

- Prefix and ASN monitoring to identify BGP prefix anomalies including **prefix hijacking**
- Global Internet BGP monitoring reveals scope and impact of Internet BGP routing events
- Monitor and alert on many types of reachability and violation criteria
- Advanced looking glass with historical data



23.34.192.0/20 Subscribed

Akamai Technologies, Inc. bell

Linked Policy: Bell prefix policy

Last Modified: 10/18/2020 3:55:53 PM

Last Active Alarm: Prefix Withdrawal

Summary
Derived from Whois Data

Updated: 3/1/2012 7:00:00 PM

Source: ARIN

NetName: AKAMAI

577
Bell Canada Backbone

Parent Route: 23.34.192.0/20

Type: NSP

Domain: http://www.bell.ca

Valid RPSL Coverage: 1 of 1

Peers
103

United States	22
Canada	11
Brazil	9
United Kingdom	7
Poland	7
Other	47

Contacts

Email: abuse@akamai.com
ip-admin@akamai.com

Phone: +1-617-274-7134
+1-617-444-0017
+1-617-444-2535
+852-22718527
+852-92813828

Organization: Akamai Technologies, Inc.

Location: Cambridge US

Origins

Observed
Derived from Looking Glass

AS	# Reporting Peers	Last Update
577: Bell Canada Backbone	103	6/7/2021 10:15:13 AM

Viewing 1 of 1 Records — Page 1 of 1

Upstream

Route Leak Mitigation – Outbound

- Use common sense when writing egress peering policies
  - Explicitly match elements for permit (specific prefixes, communities, AS path) with an explicit deny
  - XR does not allow empty prefix-lists, but other vendors do which may implicitly match all prefixes
- Require validation for policies with widespread distribution
  - Automation isn't always your friend
  - Use of BMP data and network simulation can show effect of policy changes
- SPs with downstream customers
  - Require IRR registration to generate strict prefix-lists
  - Maintain your own database of customer prefixes to generate strict prefix-lists
- https://www.manrs.org/
  - Mutually Agreed Norms for Routing Security, best practices for being a good enterprise or SP in the global Internet routing domain
- draft-ietf-grow-route-leak-detection-mitigation - Methods for Detection and Mitigation of BGP Route Leaks Usage of special RLP community to identify and mitigate leaks
  - In the same vein as methods to define peer "relationships" to determine propagation "cones", RFC 9234: Route Leak Prevention and Detection Using Roles in UPDATE and OPEN Messages (rfc-editor.org)

Route Leak Mitigation – Inbound

Strict filtering on peer connections

- Can derive data from sources like RADB and other IRRs
- Not particularly feasible with the size of tables today

Filter your transit ASNs from “peering” connections

- Use AS_PATH filtering
- Will keep your transit provider routes stable and keep a valid path through that transit provider

Filter common Tier-1 transit ASNs from ”peering” connections

- Use AS_PATH filtering
- A step further by making sure your transit carriers are only used for transit to other Tier-1 carriers, and not peers who should never advertise Tier-1 prefixes

Filter peering and your transit ASNs from all other peering connections

- Built using automation
- Use localpref, dropping routes could lead to traffic blackholes in some instances

Data Plane Boundary Concerns

- Scanning vulnerability probes and botnet C&C
- Volumetric and application-layer DoS
- CoS value retention
- Spoofed traffic
- Infrastructure attack traffic to peering edge, DNS, and other critical services

Ingress Traffic

- What should I do at the edge?
  - Filter control-plane traffic to internal infrastructure
  - Filter well-known bad traffic that won't cause user issues (chargen, etc.)
  - Fragments? Source of many attacks but may not be feasible
  - Explicitly reset CoS values on ingress
  - Monitor everything, characterize steady-state and rate-limit if you can
  - Follow security alerts from US-CERT (https://www.us-cert.gov/ncas/alerts), CVE feeds and other security organizations

- CDN is still an unsecured edge device
- Use BGP-FS for transient dynamic events, use stateless ACLs for well-defined long-term filters
- Route dark (unused) space to honeypot servers for threat inspection and research

Egress Traffic Filtering – Much the same as ingress

- Follow BCP 38 for ingress filtering on downstream connections ☺
    - Use strict filtering based on well-maintained data

- Known bad protocols with no current legitimate Internet use

- Automation is key to deploying filters quickly so your customers are not actors in attacks

Complete your session evaluations

Complete a minimum of 4 session surveys and the Overall Event Survey to be entered in a drawing to win 1 of 5 full conference passes to Cisco Live 2026.

Earn 100 points per survey completed and compete on the Cisco Live Challenge leaderboard.

Level up and earn exclusive prizes!

Complete your surveys in the Cisco Live mobile app.

Continue your education

Visit the Cisco Showcase for related demos

Book your one-on-one Meet the Engineer meeting

Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs

Visit the On-Demand Library for more sessions at [www.CiscoLive.com/on-demand]()

Contact me at: phbedard@cisco.com

Thank you

CISCO Live!

Additional slides

Best Practices Summary

- TCP-AO session authentication with strong encryption (AES)
  - TCP-AO available in IOS-XR 6.5.1 w/stronger crypto algorithms
  - MD5 as a lowest common denominator
- Control-plane policing per-peer, default in IOS-XR
- Reset IPP, DSCP, EXP on inbound peering traffic, **except for newer L4S traffic using DSCP 46**
- Delete inbound communities, especially if doing VRF peering, some vendors may accept routes with an RT set from an EBGP neighbor
- Limit BGP control-plane to only configured peers
- Data-plane filters inbound and outbound
  - If feasible whitelist your own IP space at edge
  - Automation is key in maintaining accuracy
- Review BCP 84,194, and BCP 38 if you are providing Internet service

Additional Peering Resources

- Cisco Peering Fabric HLD
  - https://xrdocs.io/design/blogs/latest-peering-fabric-hld
    - Details on best practices, validated model driven telemetry
- https://github.com/cisco-ie/anx to explore NETCONF and telemetry paths
- http://www.team-cymru.com/
  - Resource for security best practices, BOGON API feed
- https://onestep.net/communities/
  - List of communities supported by SPs to trigger route behavior
- IETF working groups
  - IDR (Inter-Domain Routing)
  - SIDR (Secure Inter-Domain Routing, now closed)
  - SIDROPS (Secure Inter-Domain Routing Ops)
  - GROW (Global Routing Operations)

Resources for Finding Peers

- Peering DB
  - www.peeringdb.net
  - Database of peering locations, who is peering at those locations, and what their peering policies are

- Networking and Peering Conferences
  - NANOG, RIPE, APRICOT, etc.
  - Meet other providers and IXP organizers
  - Negotiate peering terms and interconnection cost

- Content cache providers
  - Netflix OpenConnect
  - Google Global Cache
  - Akamai
  - Apple

Example of Important Peering MDT

```
sensor-group peering
  sensor-path Cisco-IOS-XR-shellutil-oper:system-time/uptime
  sensor-path Cisco-IOS-XR-mpls-te-oper:mpls-te/tunnels/summary
  sensor-path Cisco-IOS-XR-infra-xtc-agent-oper:xtc/policy-summary
  sensor-path Cisco-IOS-XR-pfi-im-cmd-oper:interfaces/interface-summary
  sensor-path Cisco-IOS-XR-ethernet-lldp-oper:lldp
  sensor-path Cisco-IOS-XR-wdsysmon-fd-oper:system-monitoring/cpu-utilization
  sensor-path Cisco-IOS-XR-nto-misc-oper:memory-summary/nodes/node/summary
  sensor-path Cisco-IOS-XR-clns-isis-oper:isis/instances/instance/topologies
  sensor-path Cisco-IOS-XR-ipv4-acl-oper:ipv4-acl-and-prefix-list/oor/access-list-summary/details
  sensor-path Cisco-IOS-XR-ipv6-acl-oper:ipv6-acl-and-prefix-list/oor/access-list-summary/details
  sensor-path Cisco-IOS-XR-clns-isis-oper:isis/instances/instance/neighbor-summaries/neighbor-summary
  sensor-path Cisco-IOS-XR-infra-statsd-oper:infra-statistics/interfaces/interface/latest/generic-counters
  sensor-path Cisco-IOS-XR-ipv4-bgp-oper:bgp/instances/instance/instance-active/default-vrf/process-info
  sensor-path Cisco-IOS-XR-ip-rib-ipv4-oper:rib/vrfs/vrf/afs/af/safs/saf/ip-rib-route-table-names/ip-rib-route-table-name/protocol/isis

sensor-group peering-openconfig
 sensor-path openconfig-bgp:bgp
 sensor-path openconfig-acl:acl
 sensor-path openconfig-mpls:mpls
 sensor-path openconfig-rib-bgp:bgp-rib
 sensor-path openconfig-bgp:bgp/neighbors
 sensor-path openconfig-platform:components
 sensor-path openconfig-interfaces:interfaces
```