

Datos disponibles para el monitoreo BGP en España: análisis y resultados

ESNOG 31

Madrid, 25 Abril 2024



Gael Hernandez – Director ISP/Peering @ Catchpoint



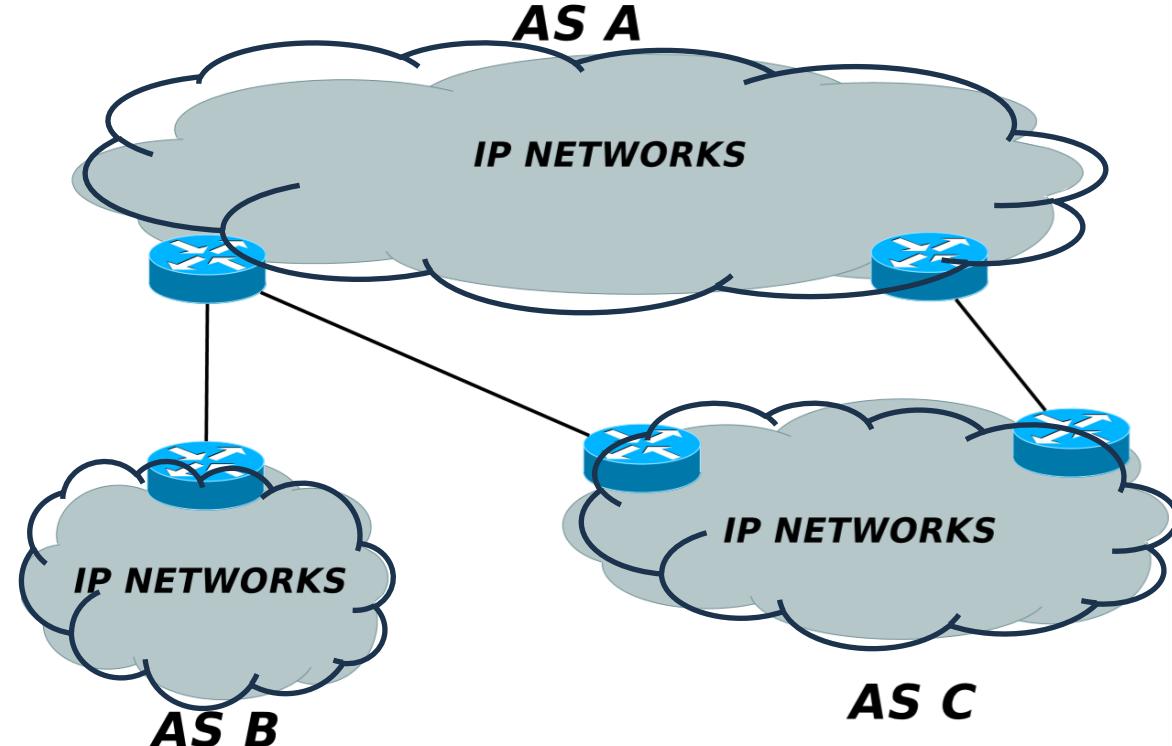
Agenda

- Repaso BGP: AS, routes, announcements, routing table y peers
- Monitoreo BGP: definición, metodología y ejemplos de uso
- Datos disponibles: análisis, comparativas y limitaciones
- Soluciones para mejorar los datos disponibles
- Conclusiones
- Preguntas y respuestas
- Pausa café ☺



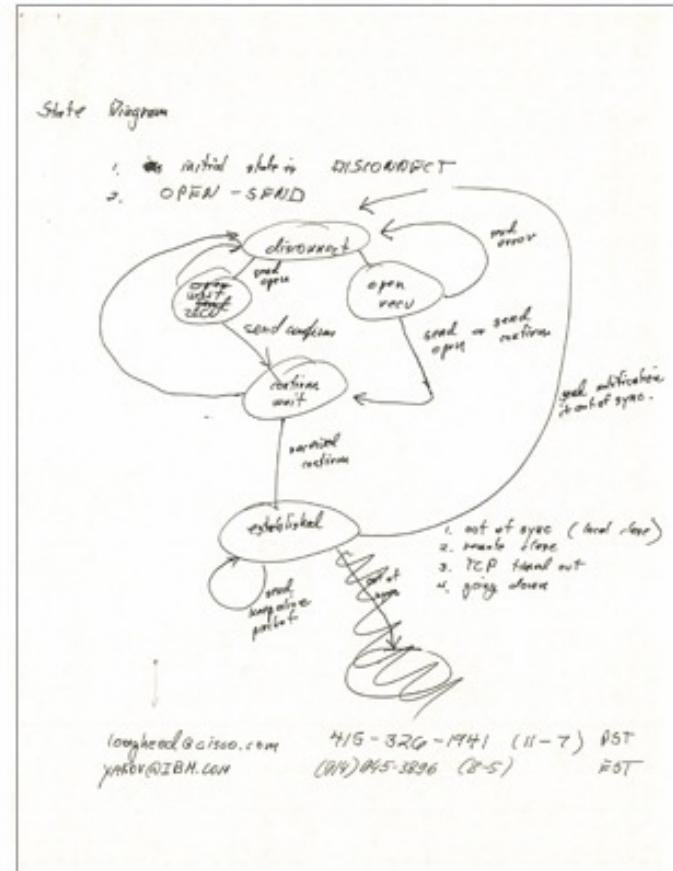
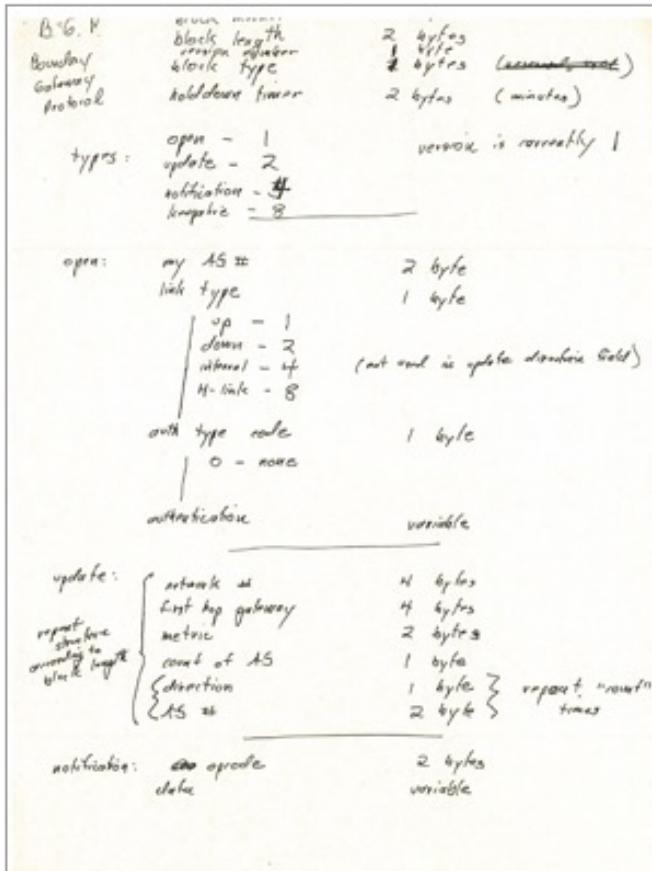
Repaso: Autonomous Systems

- The Internet can be considered as a set of Autonomous Systems connected each other
- Each AS is identified by a unique number (AS number - ASN)
 - Currently the Internet is composed by about 70,000 ASes
- Examples:
 - AS 15954 – Tecnocratica
 - AS 35699 – Adamo Telecom Iberia
 - AS 12430 – Vodafone España
 - AS 15704 - MasMovil
 - AS 15169 - Google
 - AS 54115 – Meta
 - AS 397601 - Catchpoint



Repaso: BGP

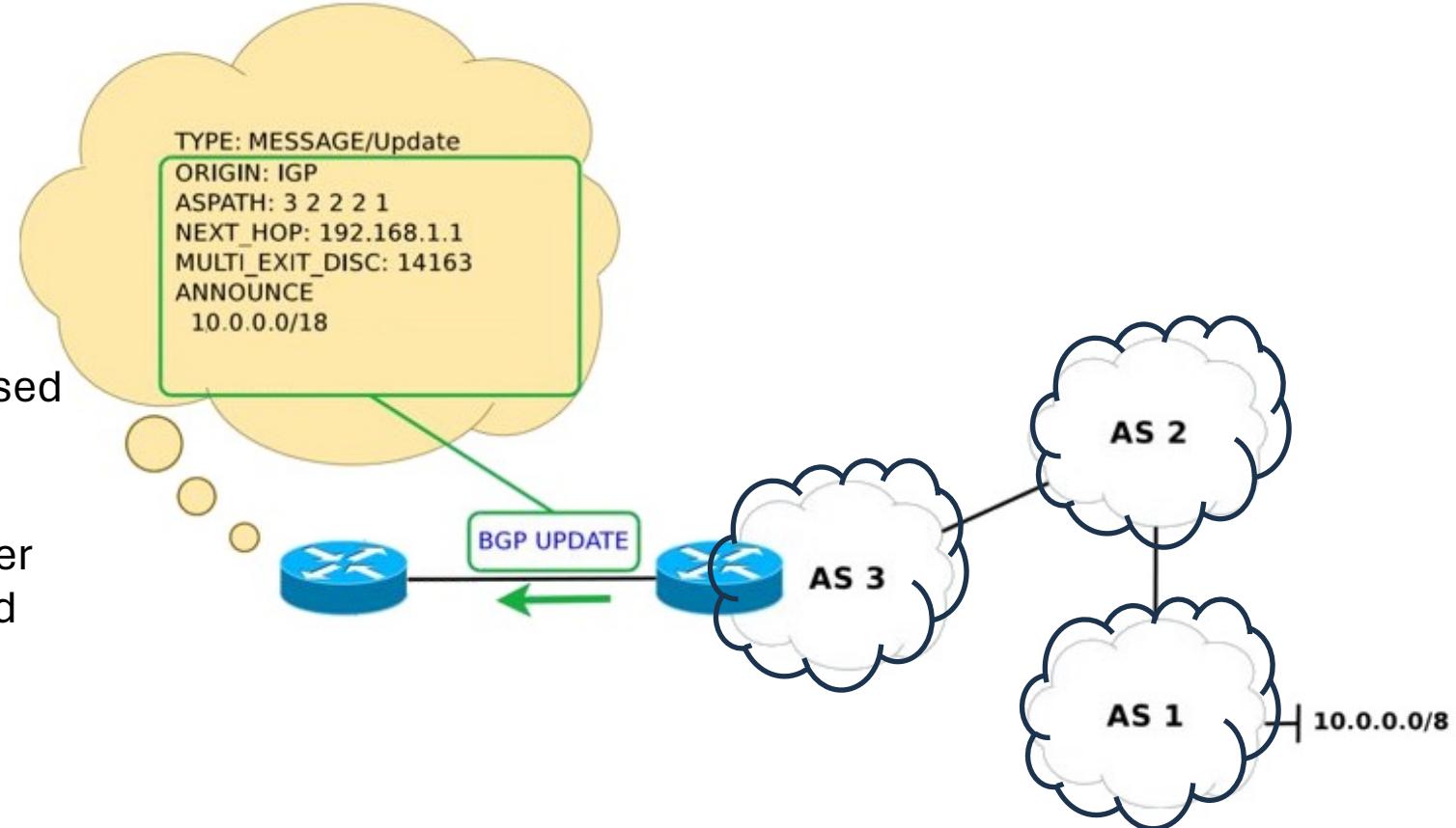
- Conceived by Kirk Lougheed, Len Bosack (Cisco) and Yakov Rekhter (IBM)
- It was intended as a quick fix to EGP... in 1989
- Nowadays, it is still the de-facto standard inter-AS routing protocol



Repaso: BGP route

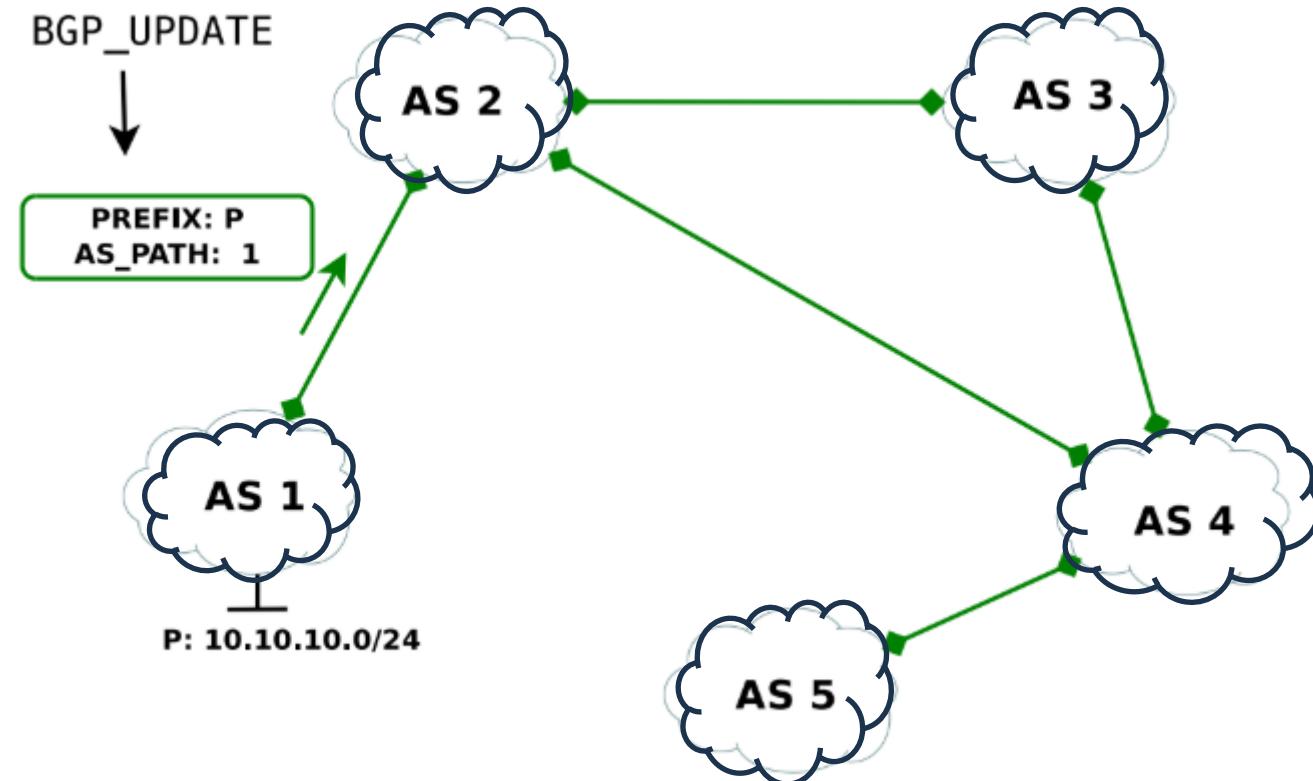
- BGP is intended to exchange routing information (**routes**) with other BGP systems
- “*A **route** is a unit of information that pairs a set of destinations with the attributes of a path to those destinations*” [RFC 4271]

- Attributes examples:
 - AS path
 - The sequence of AS to be crossed to reach the destination
 - Next hop
 - The IP address of the next router to which a traffic packet should be sent in order to reach the destination prefix



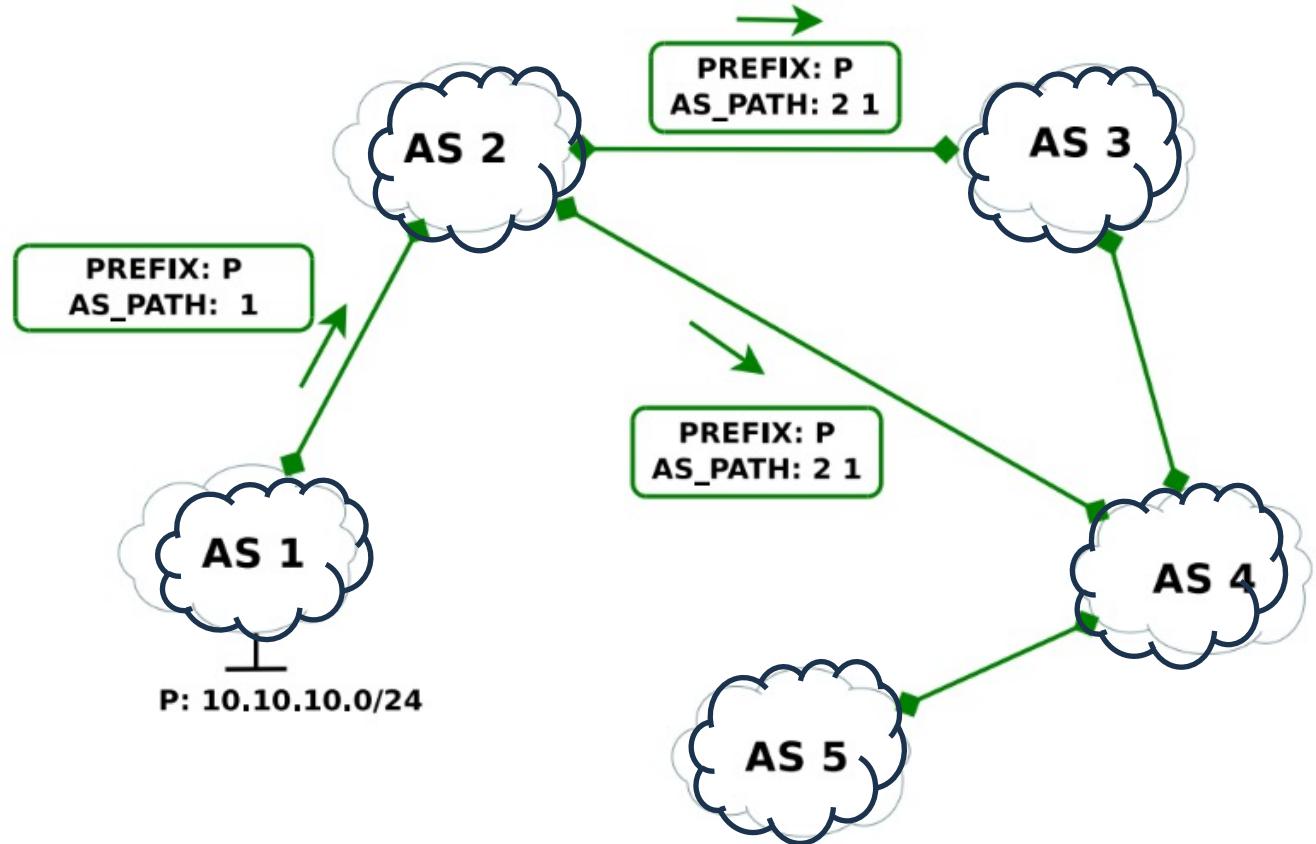
Repaso: Route announcement

- AS1 wants to make its prefix $P=10.10.10.0/24$ reachable by the rest of the Internet, so the services hosted on that network can be globally reached
- AS1 sends an UPDATE message to its neighbors (only AS2) containing the destination prefix and the proper AS_PATH
 - We say that AS1 is the **origin** of prefix P (or that AS1 originates P)



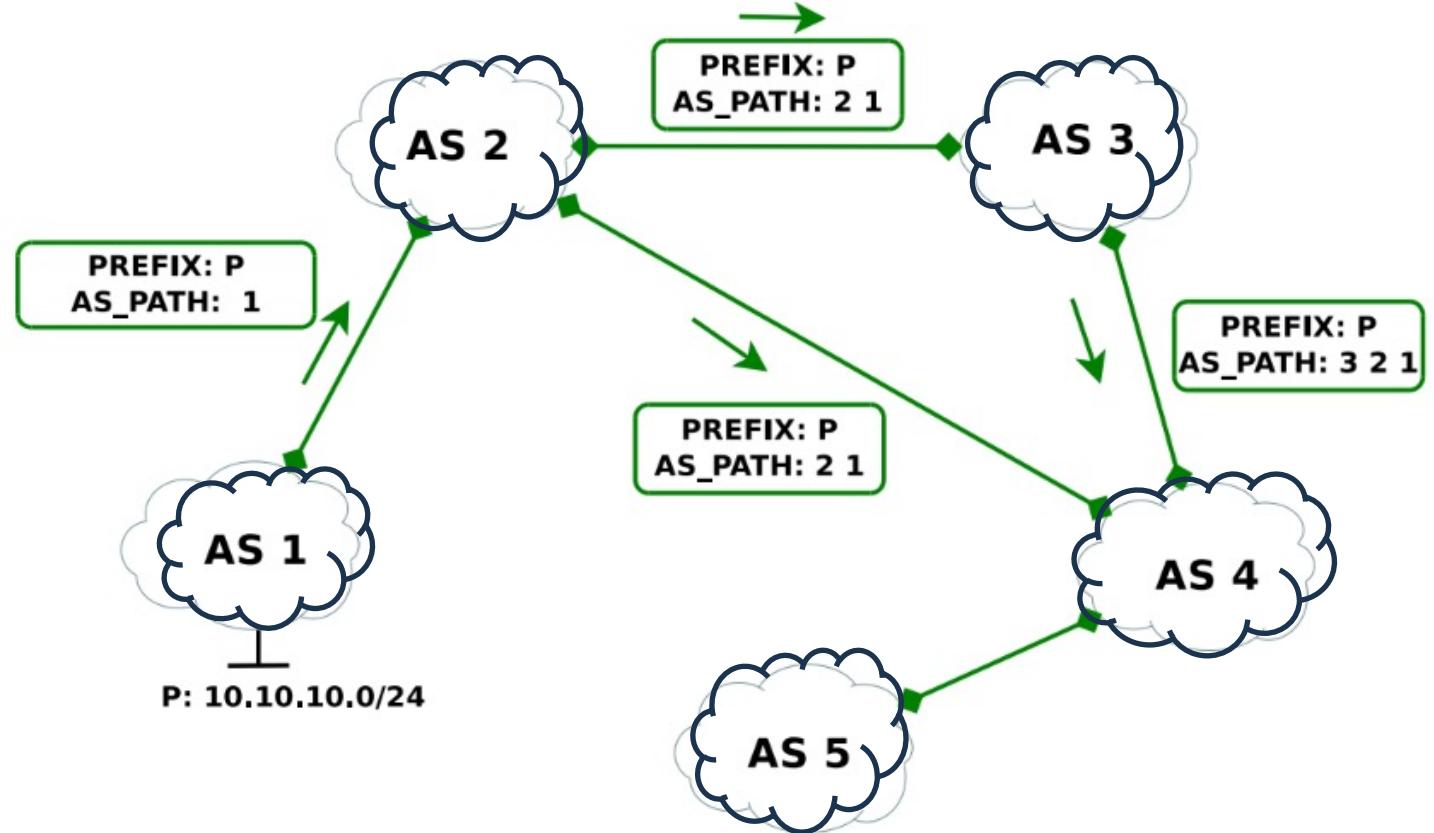
Repaso: Route announcement

- Each AS receiving an UPDATE (only AS2 at this step) propagates the information to its BGP neighbors, prepending to the AS PATH its ASN



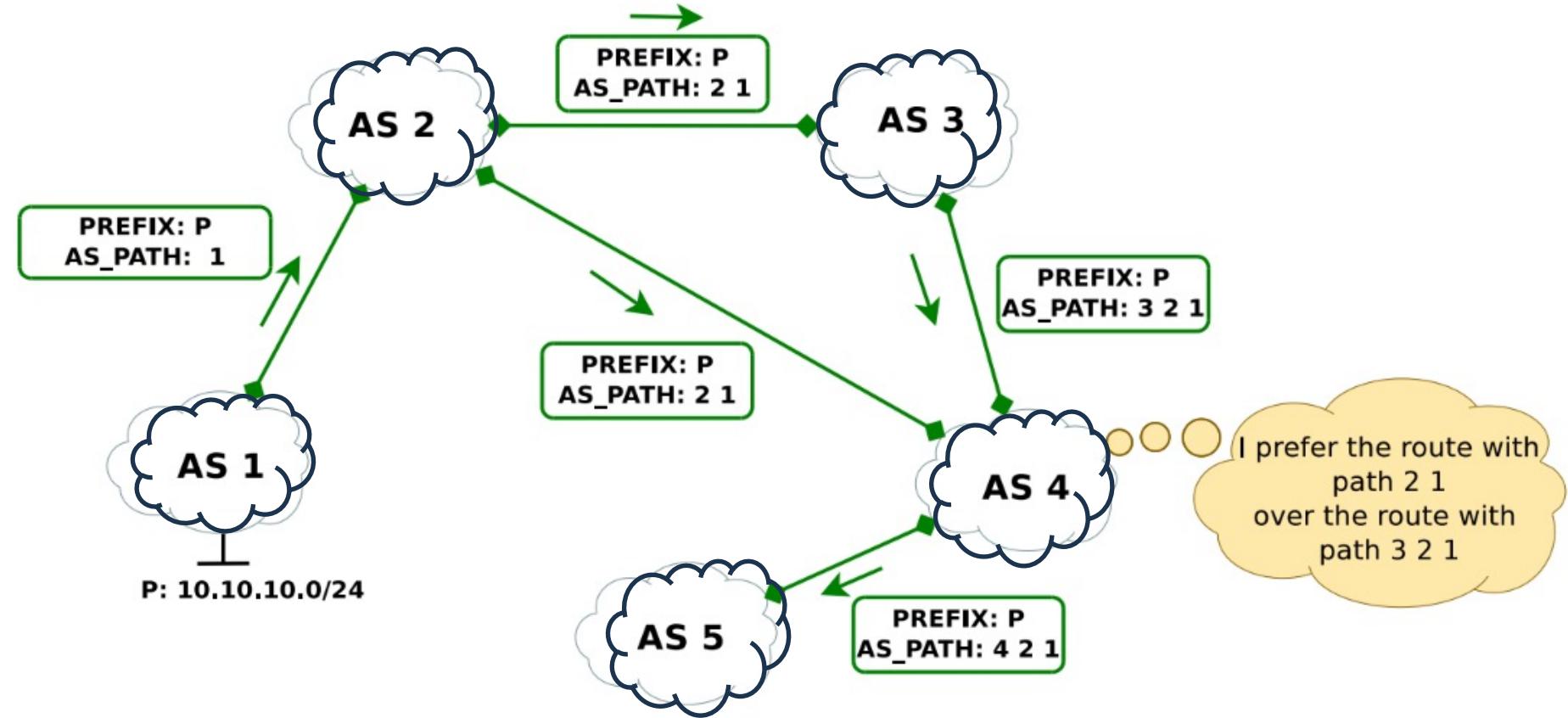
Repaso: Route announcement

- The routing process continues...



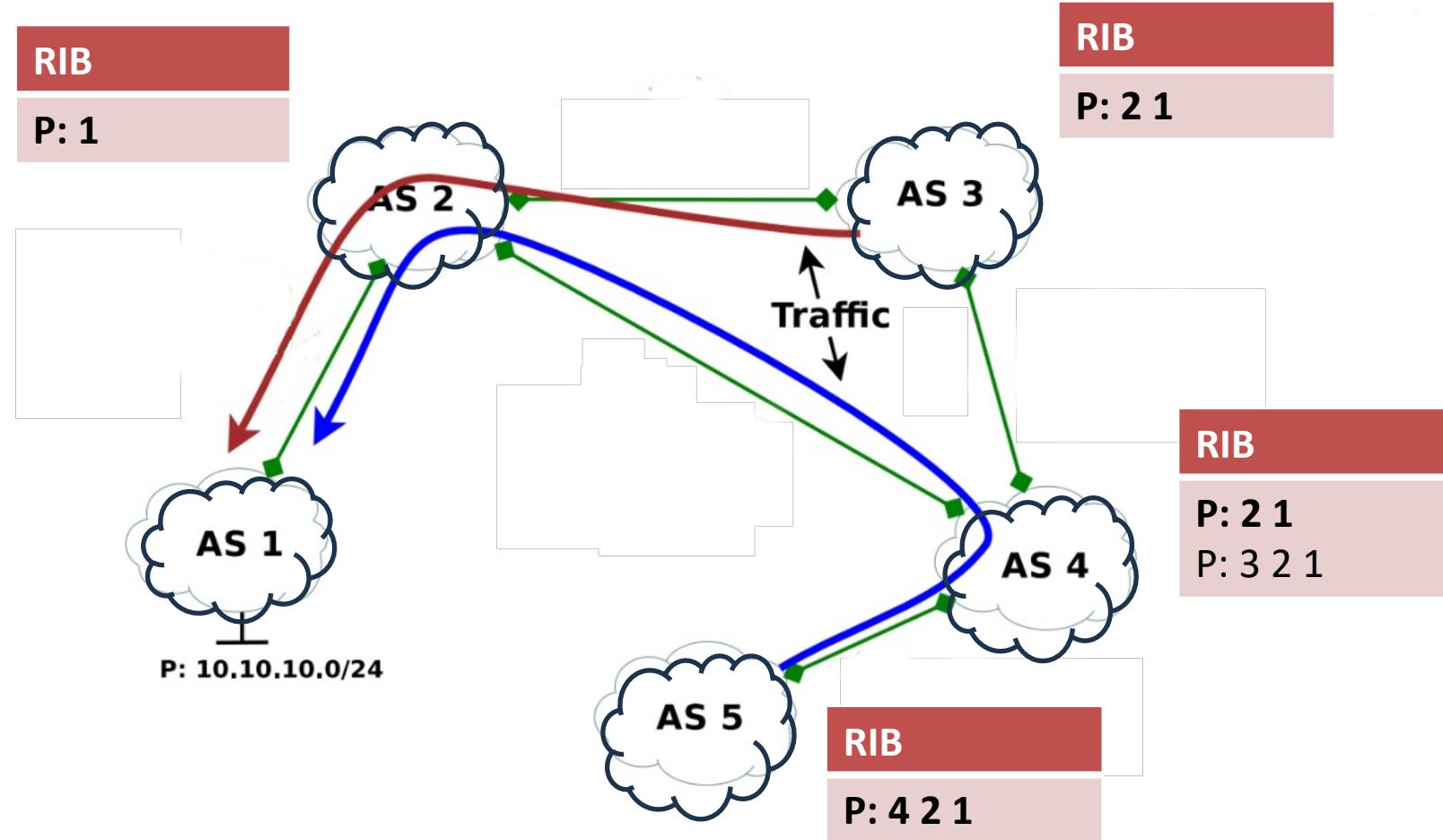
Repaso: Route announcement

- When an AS receives multiple routes to reach a destination, it must apply the BGP decision process to choose the **best route**



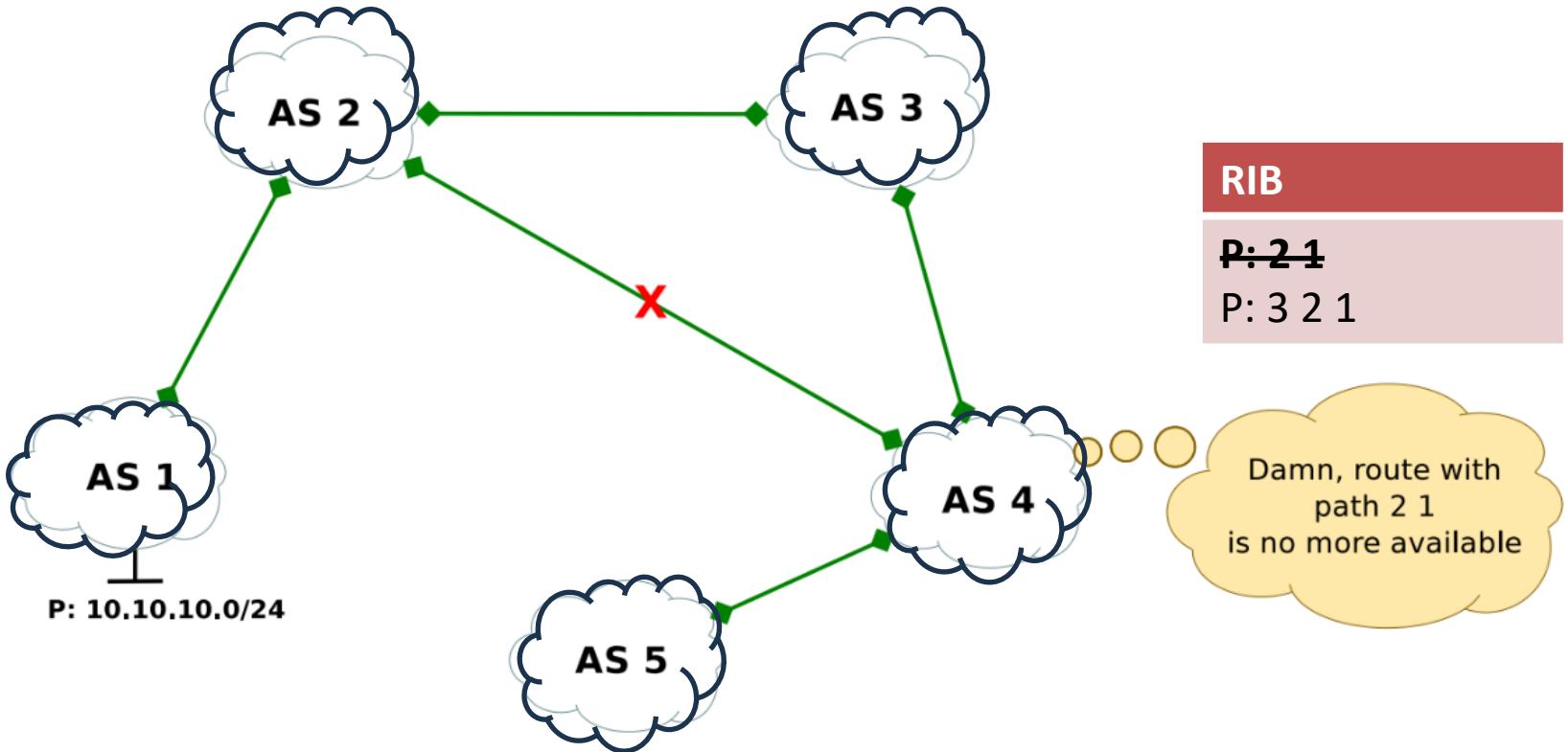
Repaso: Route announcement

- Finally, every AS will know the routes to prefix P



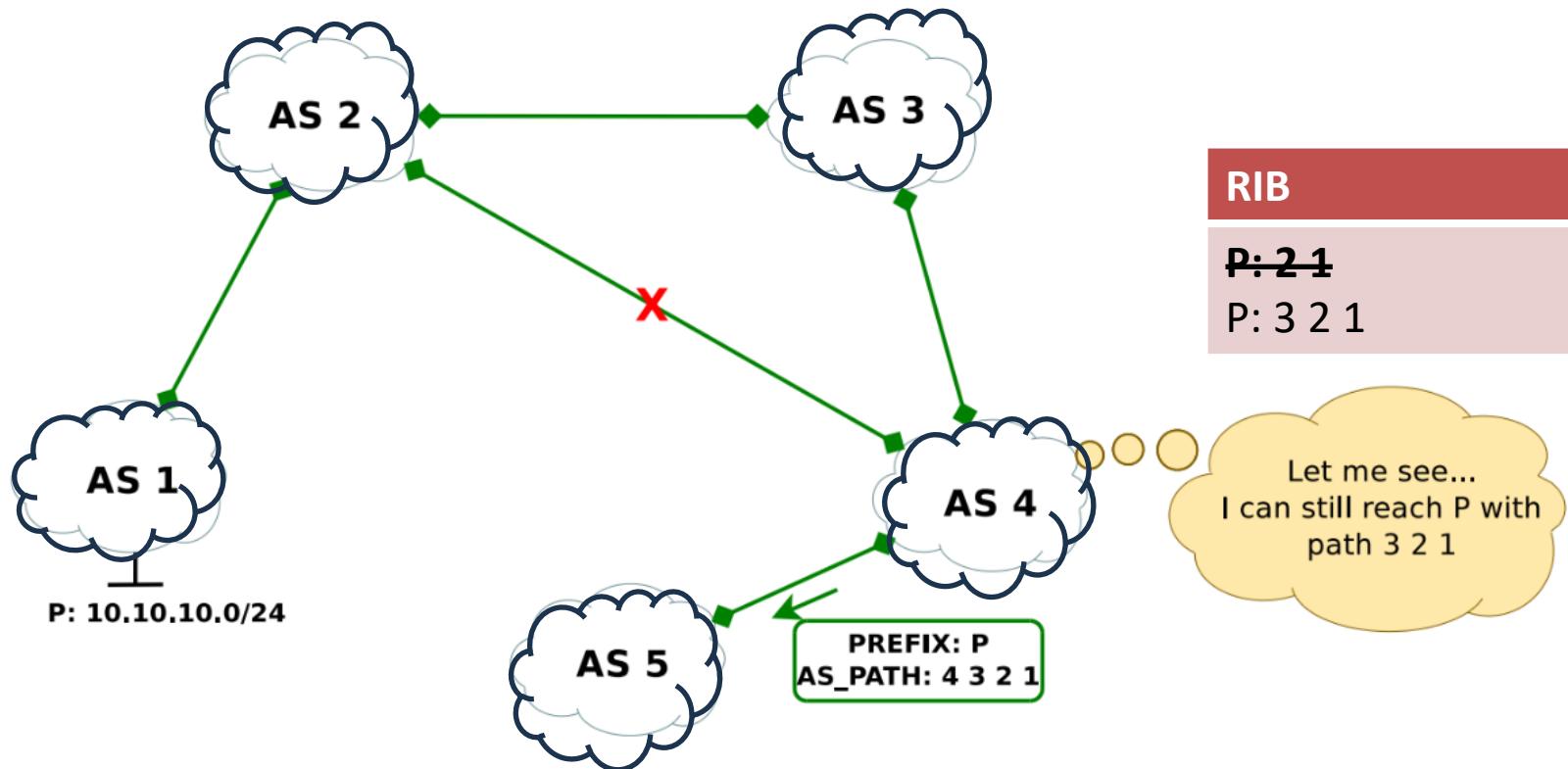
Repaso: Route replacement

- Due to several events (e.g. link failure, router failure, ...) a route may become unavailable



Repaso: Route replacement

- As soon as a change in routing is perceived, the BGP routing process tries to find an alternative way to reach the involved prefixes. If (and when) the preferred route will come back, it will replace the latest. If this happens frequently, we will have a **route flap**



Repaso: BGP routing table

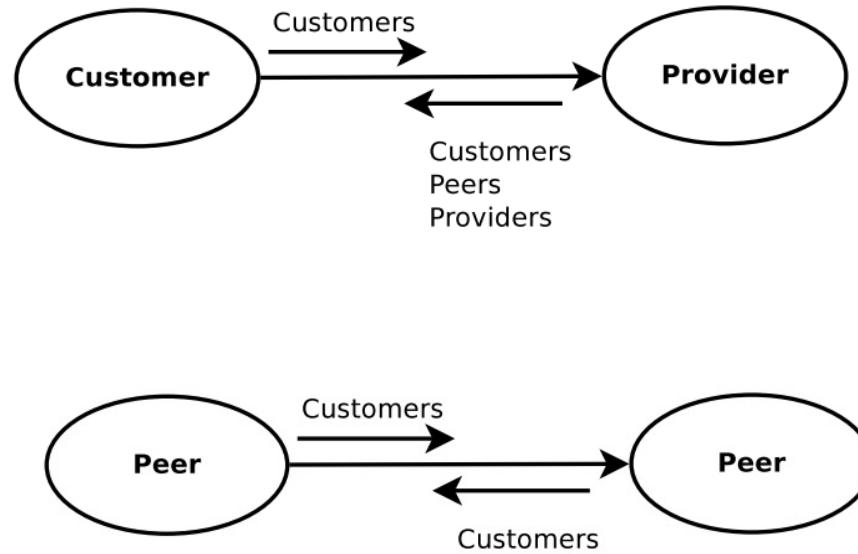
- BGP routes are kept into a routing table called Local Routing Information Base (**Loc-RIB**)
 - A usual IPv4 routing table today consists of about 900k routes
 - A usual IPv6 routing table today consists of about 200k routes
- When a packet is received the router searches into the **Loc-RIB** the route whose network prefix contains the IP destination address
 - Multiple routes could match, in this case the one with the longest prefix is selected (**longest prefix match**)

Prefix	AS_PATH	NEXT_HOP
10.0.0.0/8	6 7 8	3.3.3.3
10.10.0.0/16	3 2 4	2.2.2.2
10.10.10.0/24	3 2 5	1.1.1.1
...

- A packet destined for example to 10.10.10.1 will use the route toward 10.10.10.0/24
- From the AS_PATH we can infer which ASes have established a BGP session together
- Only the best routes are propagated to neighbours – potential hidden routes or events

Repaso: BGP peers

- BGP is governed by commercial agreements between ASes
- Typical agreements are:
 - **customer-to-provider (c2p)**: one of the two ASes (the provider) is providing transit to the whole Internet for the other AS (the customer). Usually the customer pays the provider
 - **peer-to-peer (p2p)**: the two ASes decide to announce each other the networks which each AS can reach without using any transit connection or any other p2p relationship. Usually it is a settlement-free agreement



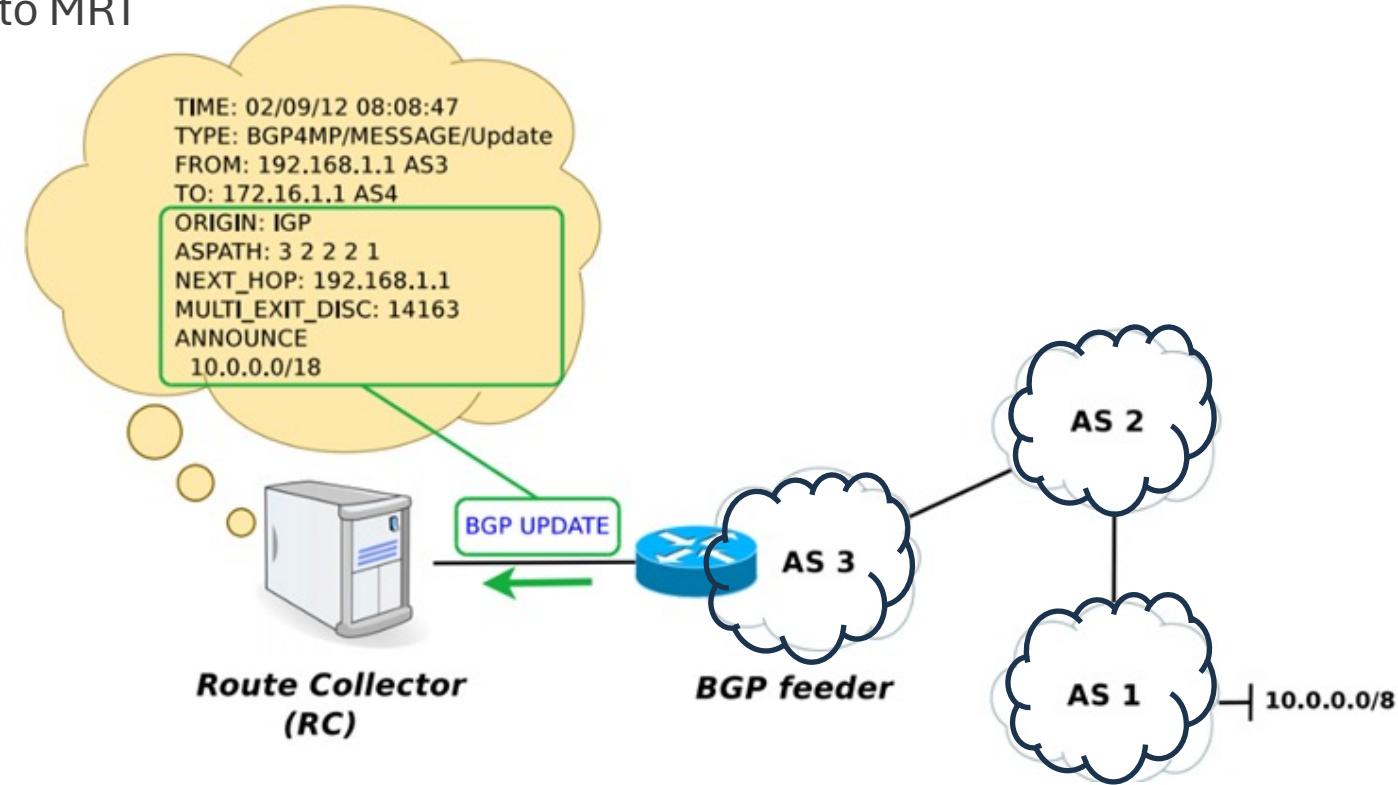


Monitoreo BGP: definición

- El monitoreo de rutas BGP permite detectar de manera pasiva eventos que pueden degradar el rendimiento de la capa de red.
 - Monitoreo del plano de control
 - No necesita enviar paquetes como por ejemplo traceroute – la información BGP ya la tenemos
 - Perspectiva del router del ISP – se pueden agregar la información de muchos routers a la vez
 - Puede expandirse más rápidamente – nuevos ISPs pueden compartir su información
- ¿Qué queremos monitorear?
 - Disponibilidad y accesibilidad de prefijos
 - Cantidad de anuncios y withdrawals
 - Cambios en el AS PATH
 - Secuestro de prefijos y fuga de rutas (prefix hijacks and route leaks)
 - RPKI
- El servicio perfecto:
 - Ofrece resultados en tiempo real
 - Usa fuentes de datos disponibles públicamente
 - Consigue fuentes de datos adicionales (privadas)

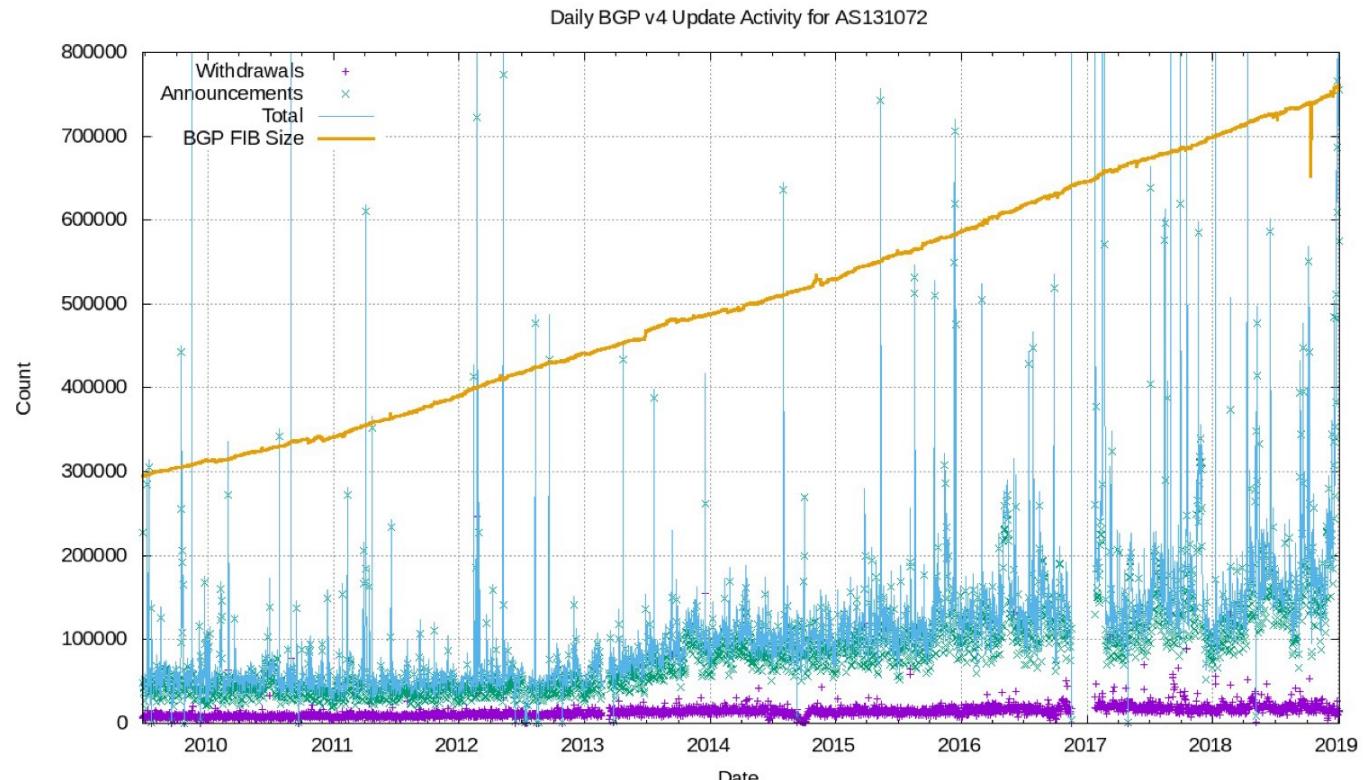
Monitoreo BGP: metodología

- El monitoreo de rutas BGP se realiza a través de route collectors (RC)
- Un RC es un dispositivo que establece sesión(es) BGP con ASes que quieren proporcionar información de enrutamiento
- Normalmente, los RC no anuncian ninguna ruta para no interferir con los procesos de enrutamiento de los ASes participantes
- Los RCs compilan solo rutas, **no tráfico !**
- Datos se almacenan en formato MRT



Monitoreo BGP: tipos de datos recogidos

- RIB snapshots
 - Snapshots of the full tables shared by each peer with the route collector
 - 900k subnets per peer sharing a full routing table in IPv4
 - 200k subnets per peer sharing a full routing table in IPv6
- Collection of UPDATE messages
 - Only subnets announced in the time interval of collection
 - Subnets are ordered by arrival time
 - Contains status messages to understand if a peer shutdown or a new peer showed up





Monitoreo BGP: fuentes de datos

- University of Oregon Route Views project (<http://www.routeviews.org/>)
- Route Views collects data in MRT format since 2001
 - RIB snapshots every 2 hours.
 - MRT updates every 15 minutes
 - **175 peers** sharing a full routing table
- RIPE NCC Routing Information Service (<http://ris.ripe.net/>)
- RIS collects data since 2000 in MRT format
 - RIB snapshots every 8 hours
 - MRT updates every 5 minutes
 - **295 peers** sharing a full routing table
- RIS provides real-time BGP data via websockets
- Packet Clearing House's Routing Information Archive (<http://www.pch.net>)
- PCH collects data in MRT format since 2001
 - RIB snapshots every 24 hours
 - MRT updates every 5 minutes
- Route collectors are hosted at IXPs. By default, **peers don't share full tables**.



Monitoreo BGP: evolución de eventos con datos MRT

- Con un snapshot de la RIB y cada actualización MRT podemos recrear el enrutamiento de cada peer
- Los datos MRT de RIS tienen 5 minutos de retraso y Route Views tiene 15 minutos





Monitoreo BGP: herramientas de manejo MRT

- Existen varias herramientas que leen y extraen informacion de ficheros MRT, como por ejemplo:
 - [bgpdump](#): uno de los primeros traductores MRT, mantenido desde 2005 por RIPE NCC, escrito en C
 - [bgpstream](#): software suite manteido por CAIDA consistente en herramientas linea de commandos, bindings a python, libreria C.
 - [bgpscanner](#): traductor MRT escrito en C. Formaba parte del proyecto Isolario (ahora discontinuado)
 - [BGPKit](#): traductor MRT mantenido por Mingwei Zhang y escrito en Rust
 - [java MRT](#): traductor MRT mantenido por RIPE NCC y escrito en Java
 - [microbgp suite](#): traductor MRT escrito en C



RIPE-NCC/ **bgpdump**

Utility and C Library for parsing MRT files



6
Contributors 4
Issues 56
Stars 9
Forks



BGP  STREAM

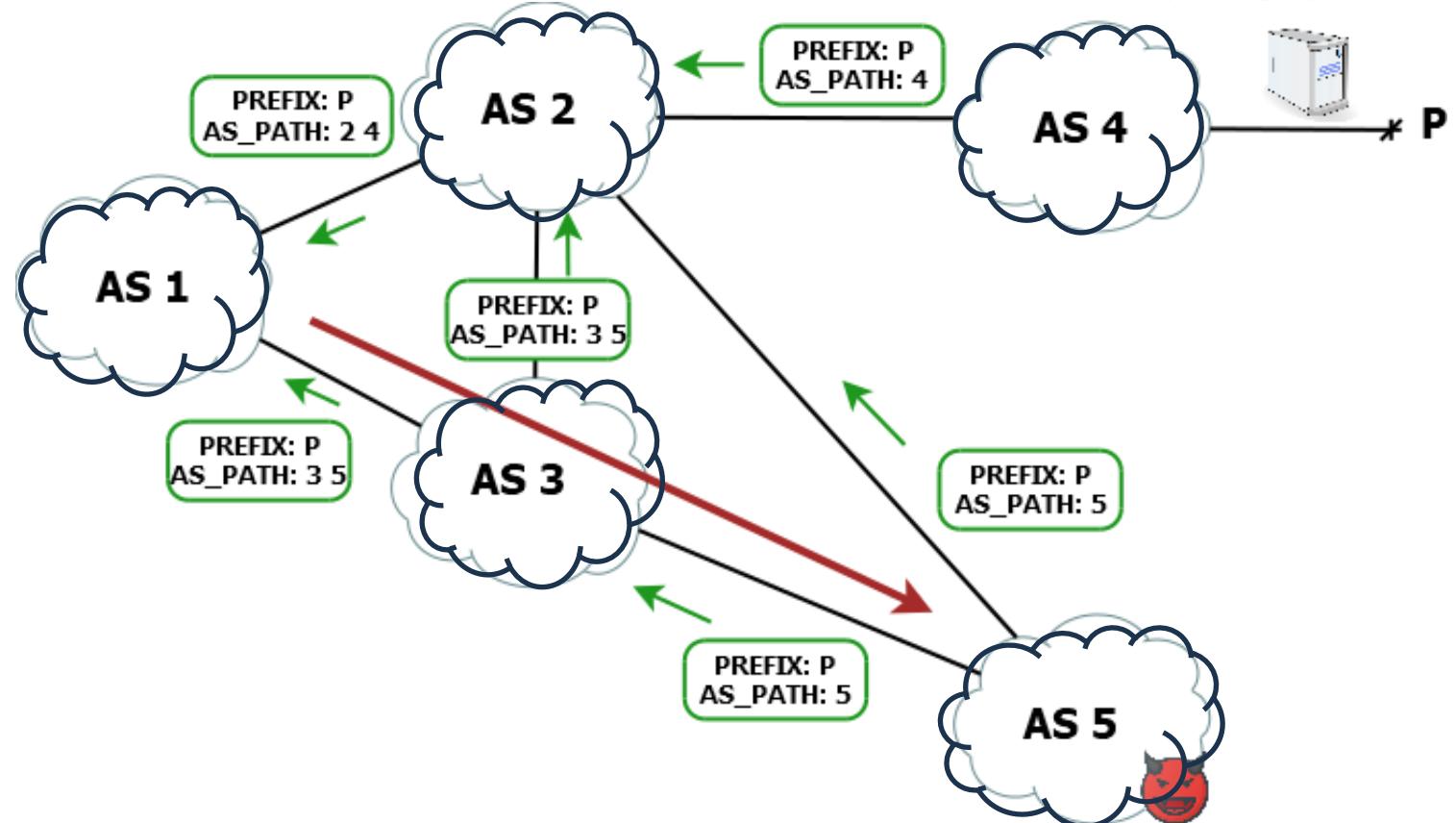
Monitoreo BGP: secuestros/fugas un problema de verdad

- La version actual de BGP fue disenada en los anos 90 sin ninguna caracteristica de seguridad
 - No hay ningun mecanismo incorporado en el protocolo para autenticar el origen del prefijo
- BGP es propenso a ataques y configuraciones erróneas
 - Secuestros de prefijos, con 911 posibles eventos registrados en 2019
 - Fugas de ruta, con 1282 eventos registrados en 2019
- Ejemplo de secuestros/fugas “famosas” :
 - [Junio 2019] Una fuga de BGP afectó a las principales redes (Cloudflare, Amazon, Facebook,...)
 - [Junio 2019] Una fuga de enrutamiento envía tráfico a través de China Telecom
 - [Mayo 2019] El DNS público administrado por la NIC de Taiwán fue secuestrado
 - [Abril 2019] NTEC fugó casi 19.000 prefijos
 - [Nov 2018] Mainone Cable Company filtró los prefijos de Google y Cloudflare
 - [Julio 2018] La empresa de telecomunicaciones iraní filtró 100 prefijos (incluido Telegram)
 - [Junio 2018] Bitcanal secuestra una empresa china de comercio electrónico
 - [Abril 2018] eNet secuestra el DNS de AWS para robar la criptomonedra Ethereum

Monitoreo BGP: secuestro de prefijos (prefix hijack)

- A prefix hijack happens when an AS originates a prefix that has not been allocated to it
 - Often called mis-origination

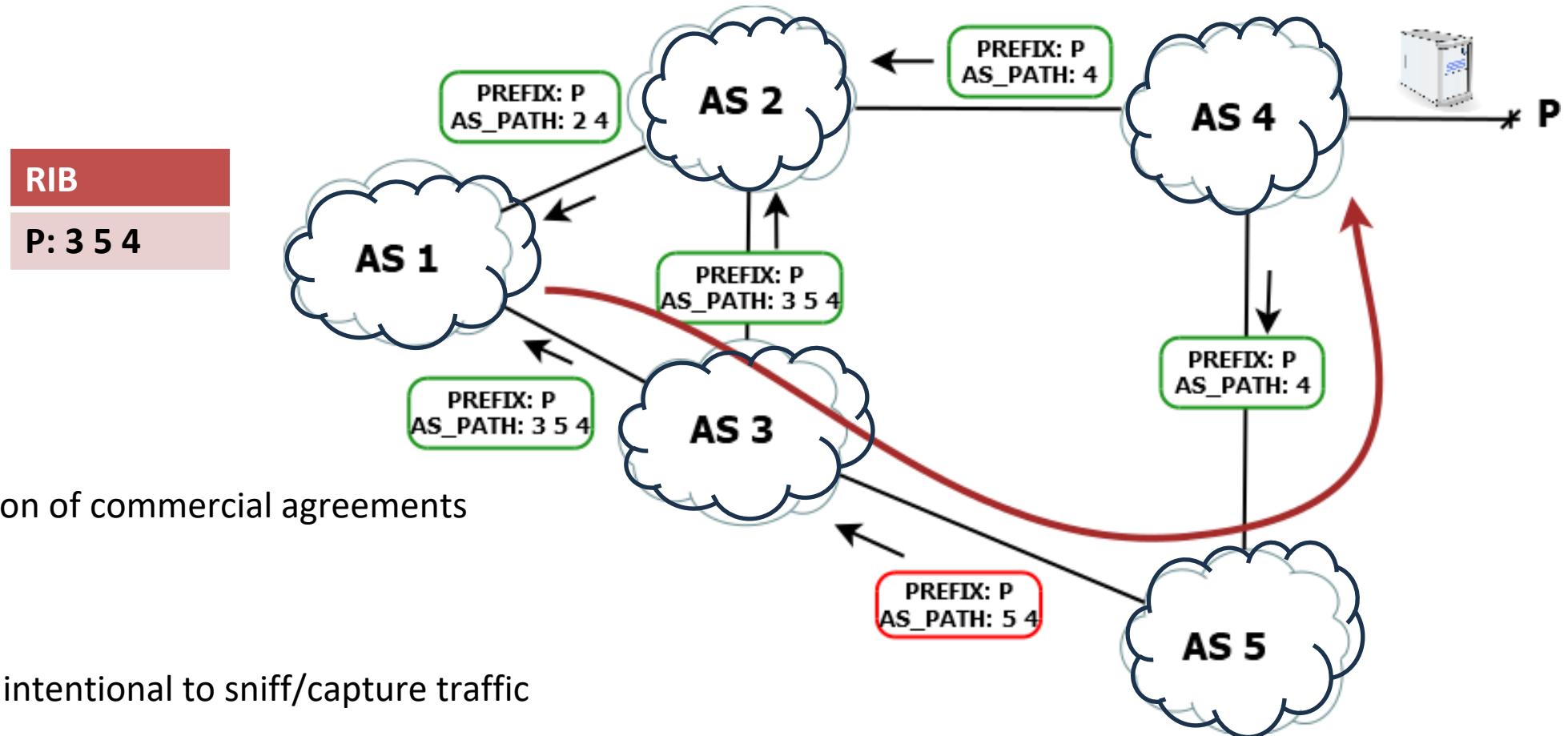
RIB
P: 3 5



- The consequences can be various:
 - Black-holing (DoS)
 - Traffic sniffing
 - Impersonation

Monitoreo BGP: fuga de rutas (route leaks)

A route leak is the *propagation of a BGP announcement(s) beyond their intended scope* [RFC 7908]

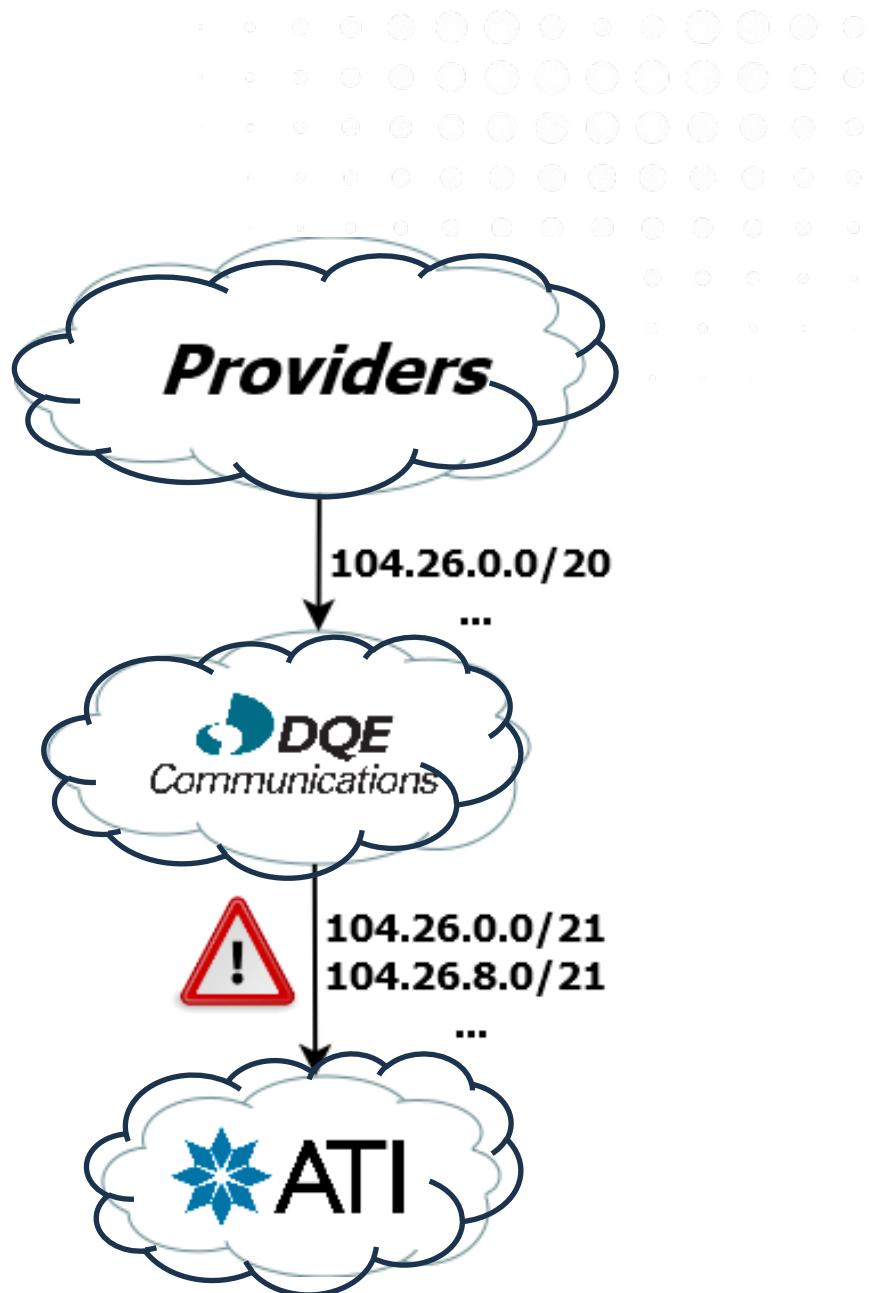


- Unintended violation of commercial agreements
 - Fat finger?
 - Bad filters?
- Also, this could be intentional to sniff/capture traffic



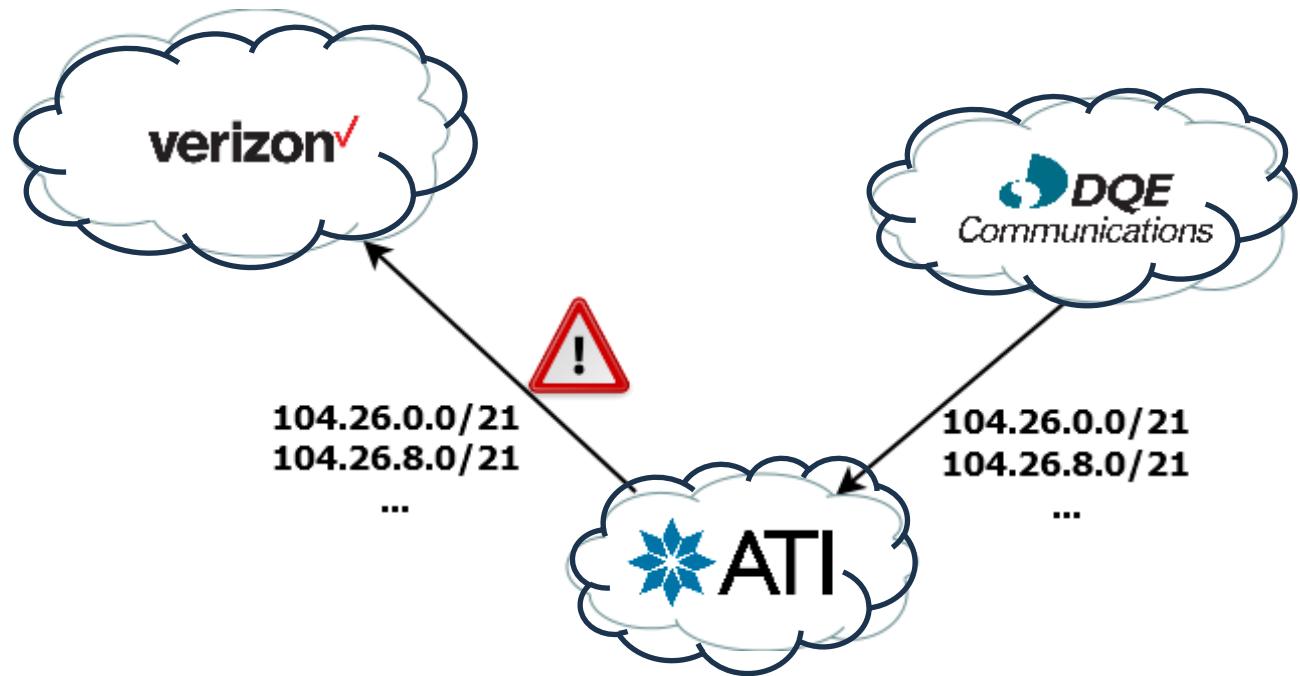
Monitoreo BGP: ejemplos de uso (i)

- **BGP leak: caso Cloudflare**
- On June 24, 2019 around 10:30am UTC AS33154 (*DQE Communications*) sent to **its customer** AS396531 (*Allegheny Technologies Inc.*) more specific prefixes for popular Internet destinations (among which Cloudflare, Amazon and Facebook)
- 104.26.0.0/20 belongs to Cloudflare AS 13335
- ATI received routes like these:
104.26.0.0/21 33154 ... 13335
104.26.8.0/21 33154 ... 13335



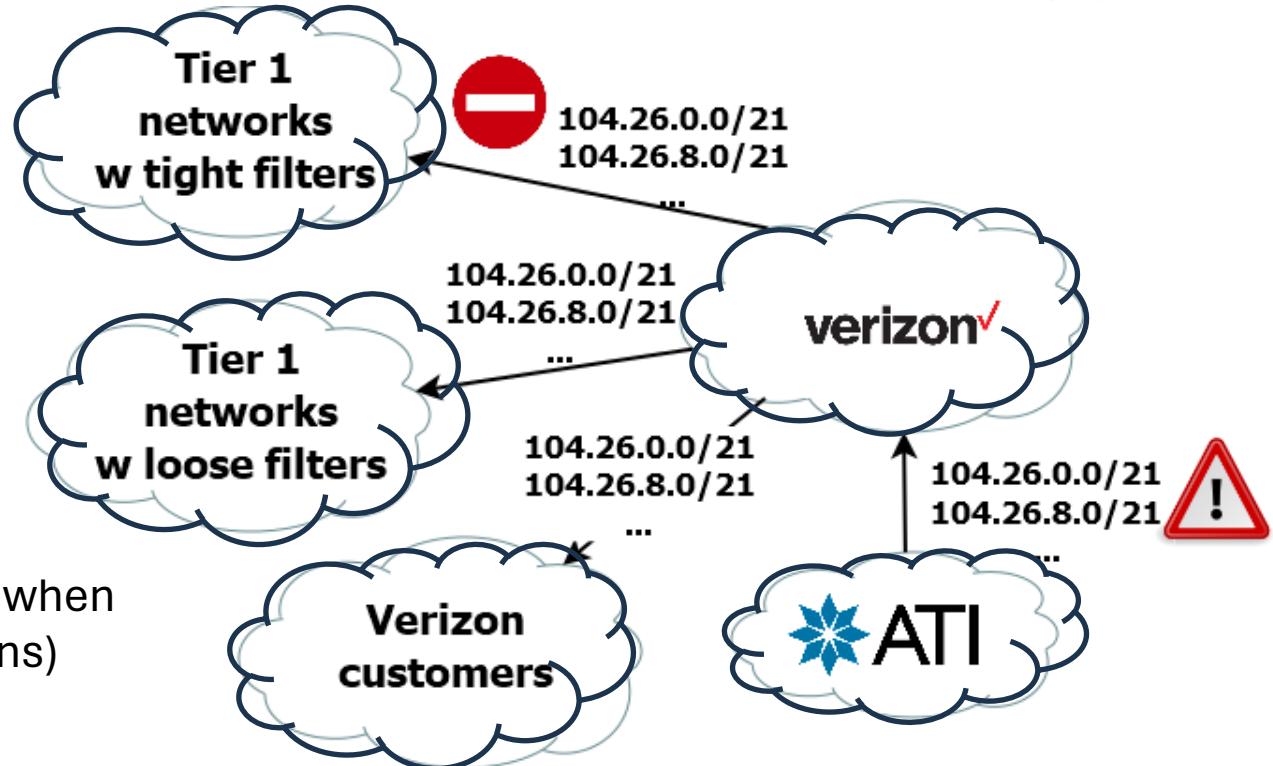
Monitoreo BGP: ejemplos de uso (ii)

- Probably due to a misconfiguration ATI propagated those routes to another of its providers (Verizon – AS701)
 - An AS should send to a provider only the routes learned from customers!
- Verizon received routes like these
 - 104.26.0.0/21 396531 33154 ... 13335
 - 104.26.8.0/21 396531 33154 ... 13335



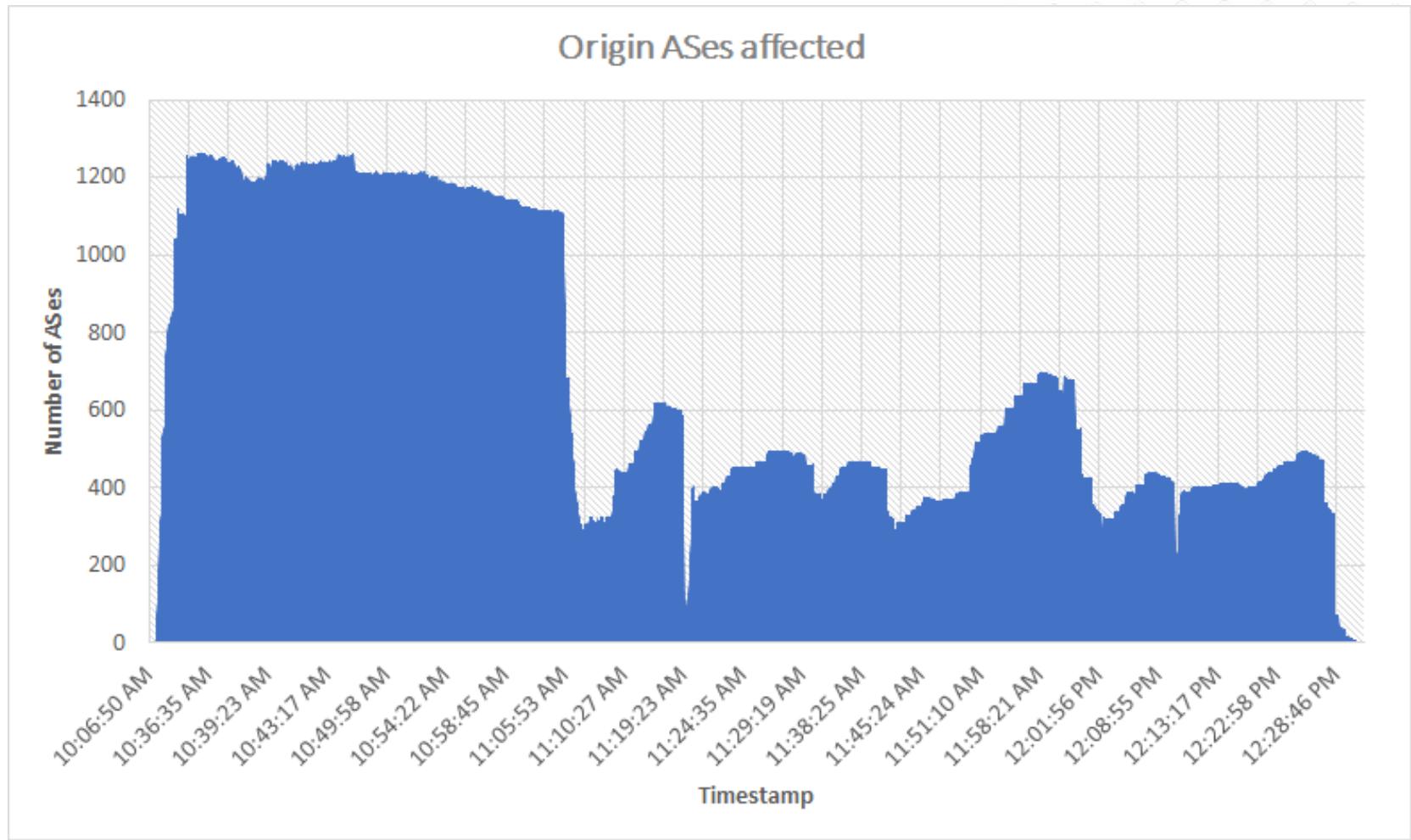
Monitoreo BGP: ejemplos de uso (iii)

- Probably due to a loose filtering mechanism AS701 accepted those routes and propagated them to its neighbours
- Verizon sent out routes like these:
 - 104.26.0.0/21 701 396531 33154 ... 13335
 - 104.26.8.0/21 701 396531 33154 ... 13335
- Some neighbour dropped the routes
- Some neighbour accepted them
- As a result, a part of the Internet used leaked routes when sending packets to 104.26.0.0/20 (longest match wins)
 - Performance degradation
 - Possible loss due to congestion



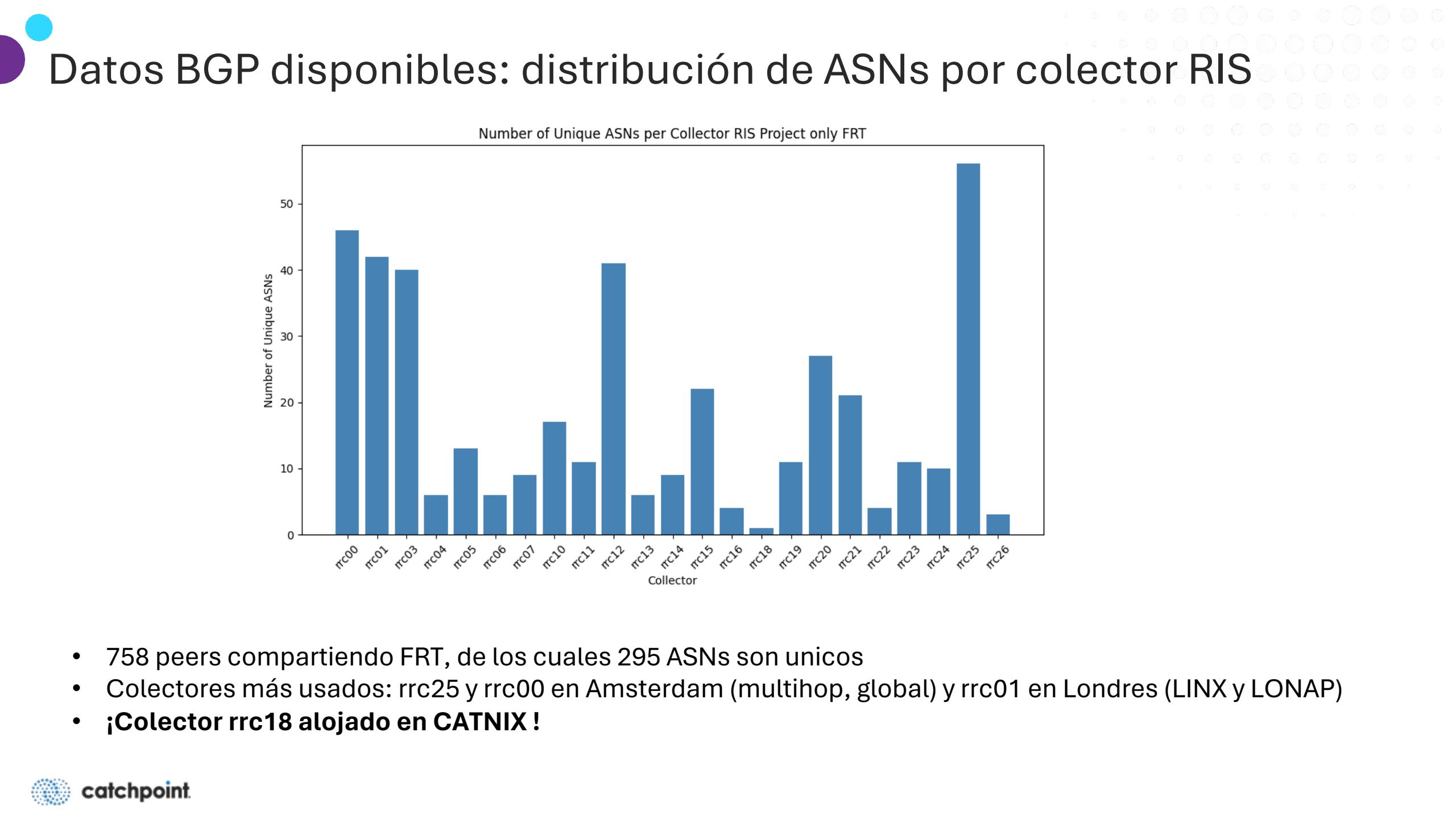
Monitoreo BGP: ejemplos de uso (iv)

- The leak didn't affect only Cloudflare...
- Facebook, Comcast, T-Mobile, Bloomberg
- 9 American banks



Datos BGP disponibles: comparativas y limitaciones

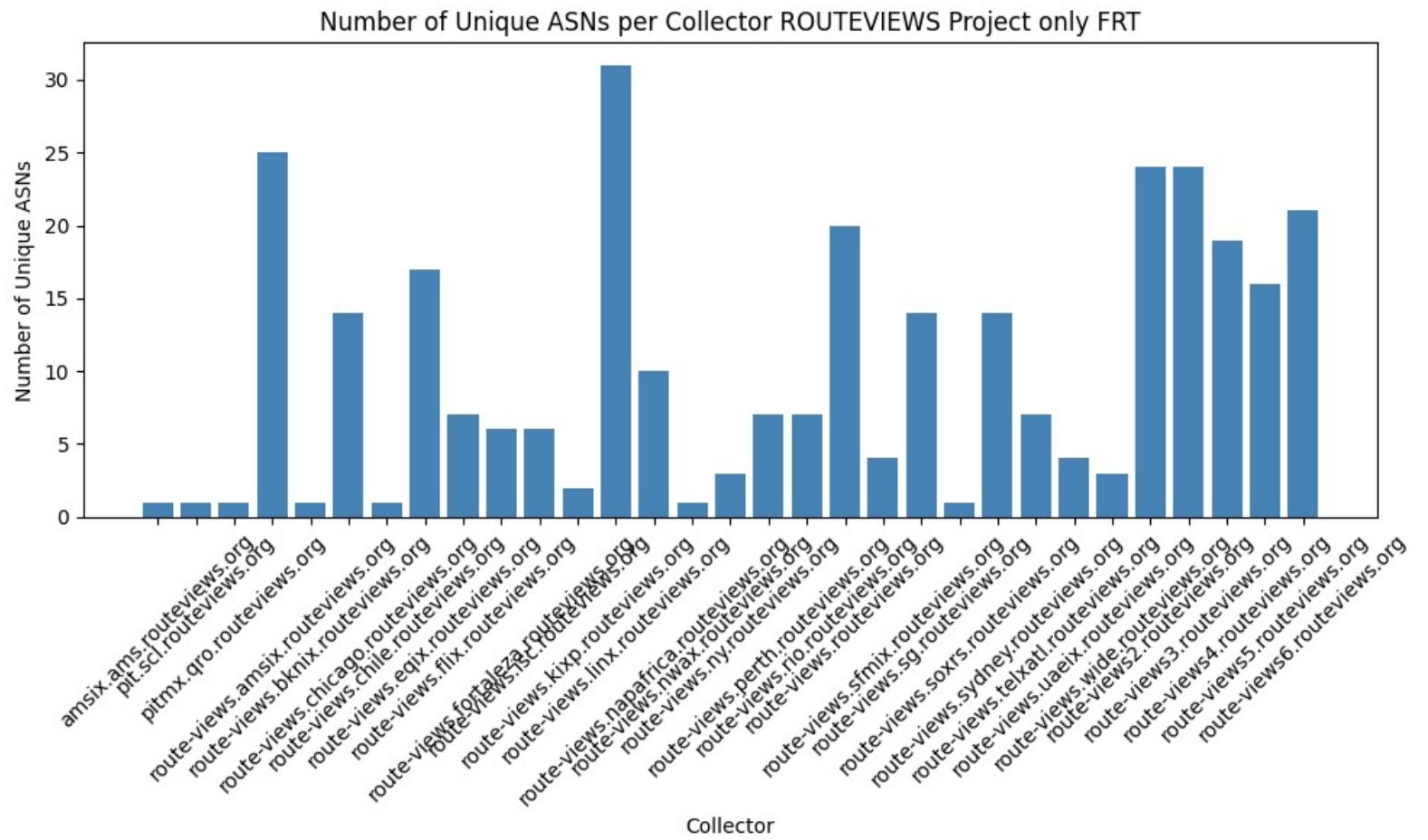
- Análisis de los peers de los proyectos RIPE RIS y RouteViews con la intención de
 - Comparar ASes únicos participantes por RC de cada proyecto
 - Comparar los datos disponibles en Espana con los de otros países vecinos
- Metodología:
 - Información sobre peers/ASNs disponible online
 - <https://www.ris.ripe.net/peerlist/rrc00.shtml> (rrc01, rrc02, etc)
 - <https://www.routeviews.org/peers/peering-status.html>
 - Scripts en python: primera versión escrita por ChatGPT ❤️ y revisión mía
 - recopilar los datos en html y almacenarlos en local en csv
 - mezclar los datos de routeviews y ripe ris en un nuevo formato csv
 - visualizar los datos
 - Disponibles en github <https://gist.github.com/elgaeloHub>



Datos BGP disponibles: distribución de ASNs por colector RIS

- 758 peers compartiendo FRT, de los cuales 295 ASNs son únicos
- Colectores más usados: rrc25 y rrc00 en Amsterdam (multihop, global) y rrc01 en Londres (LINX y LONAP)
- **¡Colector rrc18 alojado en CATNIX !**

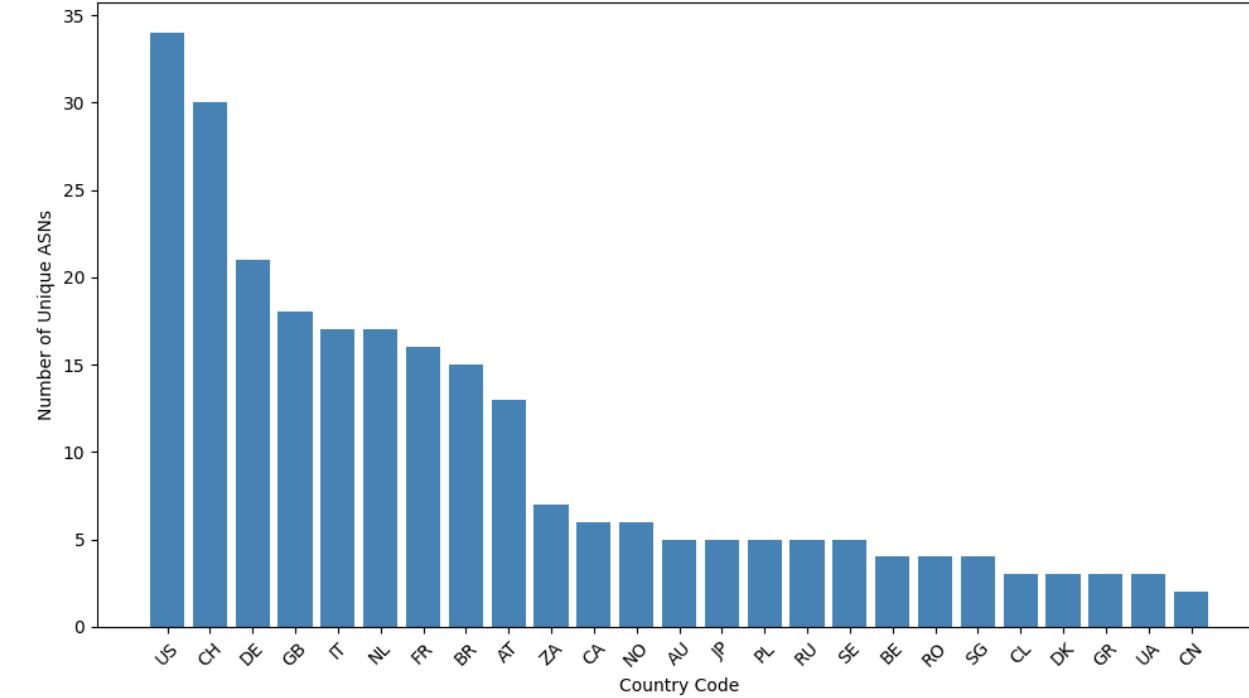
Datos disponibles: distribución de ASNs por colector Route Views



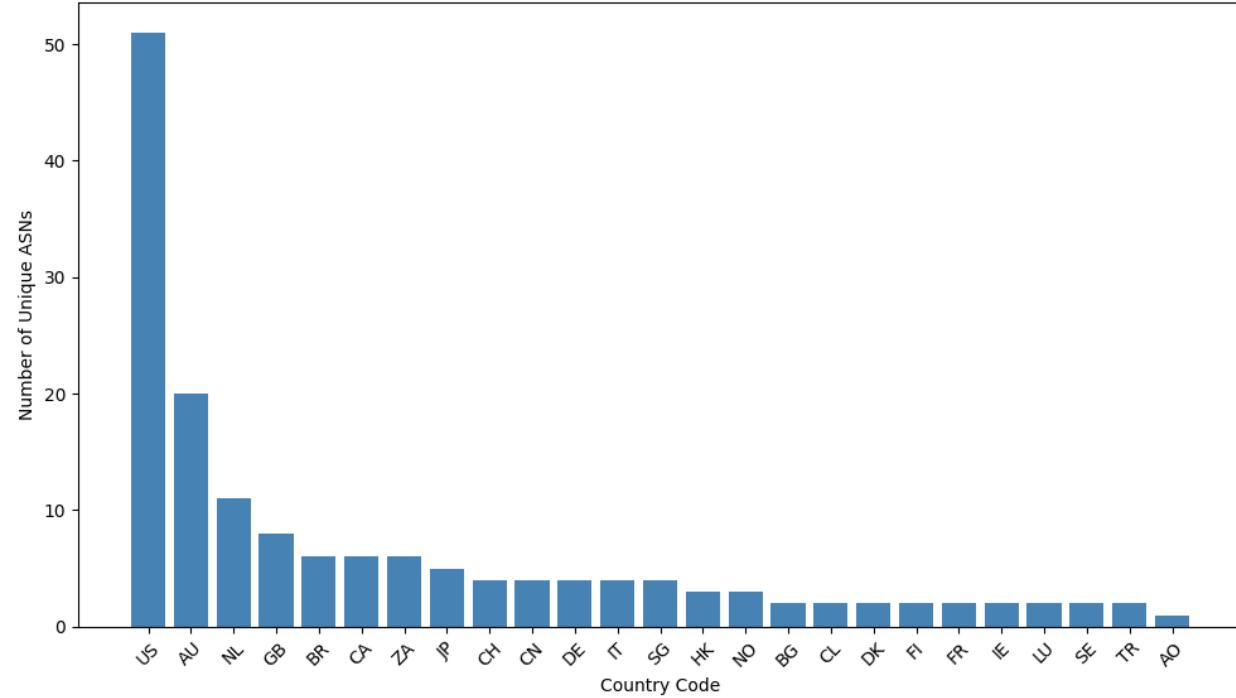
- 537 peers compartiendo FRT, de los cuales 175 ASNs son únicos
- Colectores más usados: LINX (London), AMS-IX (Amsterdam), Routeviews3 (Oregon)

Datos BGP disponibles: comparativa por país

Top 25 Countries with Most Unique ASNs in RIS Projects only FRT

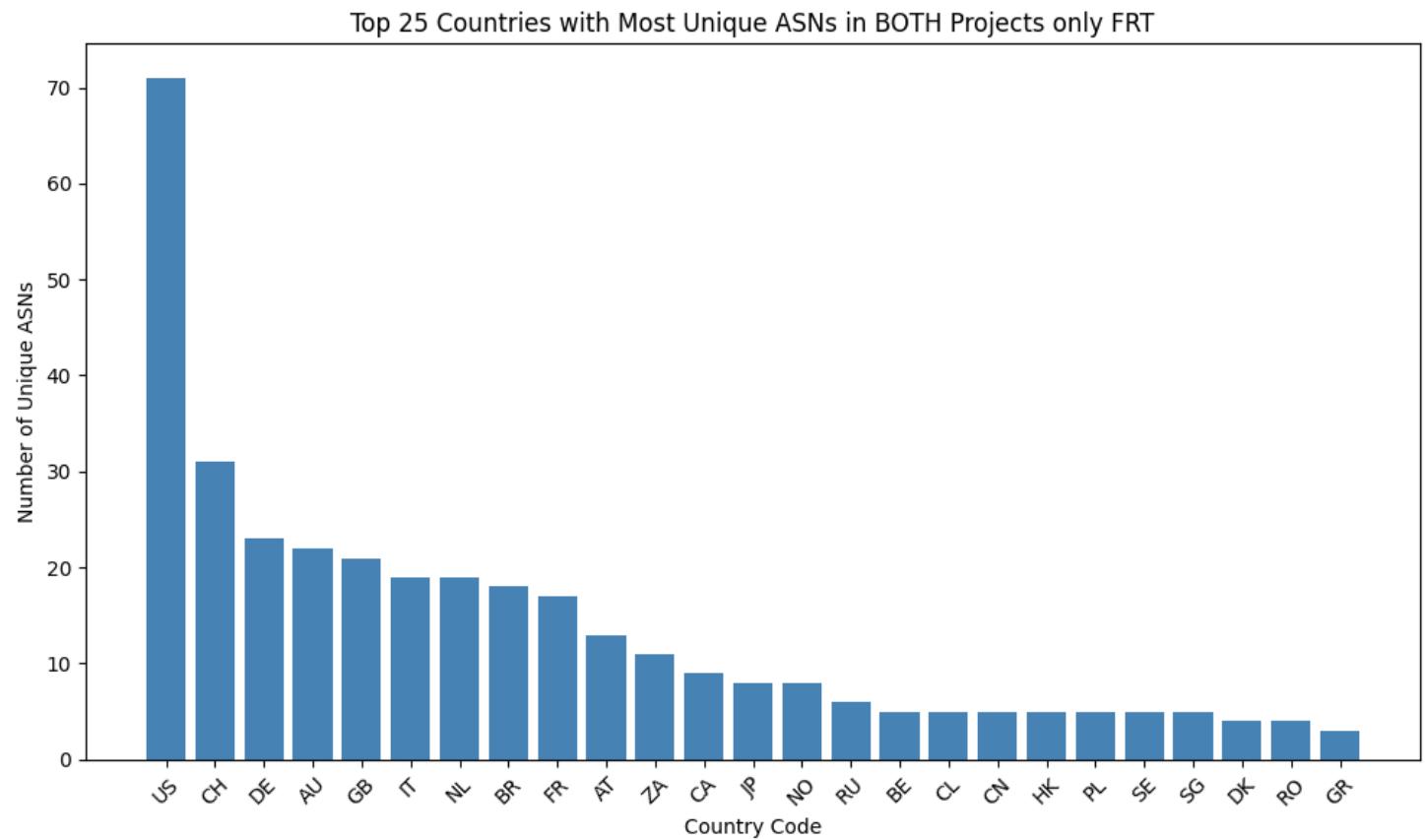


Top 25 Countries with Most Unique ASNs in ROUTEVIEWS Projects only FRT



- Las redes de Estados Unidos son las que más participan en ambos proyectos
- Países como Australia, Reino Unido, Países Bajos, Brasil, Sudáfrica o Canadá aparecen en ambos proyectos. Alemania, Francia e Italia prefieren RIS.
- 60 países representados en RIS y 42 países en Route Views

Datos BGP disponibles: comparativa por países



- Feeds de 69 países diferentes
- Los 5 países que más comparten (Estados Unidos, Suiza, Alemania, Australia y Reino Unido) contribuyen la mitad de las feeds en ambos proyectos >> Alta concentración y poca diversidad

Country	# of Peers	% total	Unique ASNs
US	308	23.78%	71
CH	95	7.34%	31
NL	95	7.34%	19
AU	83	6.41%	22
GB	63	4.86%	21
FR	56	4.32%	17
BR	54	4.17%	18
DE	53	4.09%	23
LU	48	3.71%	2
AT	47	3.63%	13
IT	47	3.63%	19
SG	37	2.86%	5
ZA	28	2.16%	11
JP	23	1.78%	8
CA	23	1.78%	9
NO	19	1.47%	8
SE	14	1.08%	5
RU	13	1.00%	6
DK	12	0.93%	4
HK	10	0.77%	5
CL	9	0.69%	5
FI	9	0.69%	2
CO	8	0.62%	8
BE	8	0.62%	5
CN	7	0.54%	5
ES	4	0.31%	2
	1169		342

Datos BGP disponibles: peers registrados en España

project	collector	FRT	asn	desc	cc	rir	ip	v4_prefix	v6_prefix
ris	rrc01	-	60171	AFR-IX T	ES	ripencc	2001:7f8:4::eb0b:1	0	39
ris	rrc01	-	60171	AFR-IX T	ES	ripencc	195.66.226.123	315	0
ris	rrc03	v6	12956	TELXIUS	ES	ripencc	2001:7f8:1::a501:2956:1	0	189375
ris	rrc03	v4	12956	TELXIUS	ES	ripencc	80.249.208.208	932382	0
ris	rrc18	-	766	RedIRIS	ES	ripencc	193.242.98.19	140	0
ris	rrc18	-	766	RedIRIS	ES	ripencc	2001:7f8:2a:0:1:1:0:766	0	17
ris	rrc18	-	13041	CESCA	ES	ripencc	193.242.98.38	9	0
ris	rrc18	-	13041	CESCA	ES	ripencc	2001:7f8:2a:0:1:1:1:3041	0	2
ris	rrc18	-	15699	ADAM	ES	ripencc	2001:7f8:2a:0:2:1:1:5699	0	5
ris	rrc18	-	15699	ADAM	ES	ripencc	193.242.98.137	27	0
ris	rrc18	-	16030	ALTECOM	ES	ripencc	193.242.98.4	71	0
ris	rrc18	-	16030	ALTECOM	ES	ripencc	2001:7f8:2a:0:1:1:1:6030	0	0
ris	rrc18	-	24592	NEXICA	ES	ripencc	2001:7f8:2a:0:2:1:2:4592	0	2
ris	rrc18	-	24592	NEXICA	ES	ripencc	193.242.98.133	9	0
ris	rrc18	v6	29680	VOZTELECOM	ES	ripencc	2001:7f8:2a:0:2:1:2:9680	0	189904
ris	rrc18	v4	29680	VOZTELECOM	ES	ripencc	193.242.98.141	941541	0
ris	rrc18	-	30892	DEUTSCHE T	ES	ripencc	193.242.98.130	12	0
ris	rrc18	-	30892	DEUTSCHE T	ES	ripencc	2001:7f8:2a:0:2:1:3:892	0	2
ris	rrc18	-	35699	ADAMO	ES	ripencc	193.242.98.143	476	0
ris	rrc18	-	43578	bitNAP	ES	ripencc	193.242.98.160	63	0
ris	rrc18	-	49835	GUIFINET	ES	ripencc	2001:7f8:2a:0:2:1:4:9835	0	9
ris	rrc18	-	49835	GUIFINET	ES	ripencc	193.242.98.144	57	0
ris	rrc18	-	60082	CATNIX RS	ES	ripencc	2001:7f8:2a:0:1:1:6:82	0	882
ris	rrc18	-	60082	CATNIX RS	ES	ripencc	193.242.98.98	7282	0
ris	rrc21	-	60171	AFR-IX T	ES	ripencc	2001:7f8:54::239	0	39
ris	rrc21	-	60171	AFR-IX T	ES	ripencc	37.49.236.239	314	0

Datos BGP disponibles: comparativa entre España y sus vecinos

Country	# of Peers	Unique ASNs	Country's ASNs	% Country ASNs
NL	95	19	1087	1.75%
GB	63	21	2300	0.91%
FR	56	17	1532	1.11%
DE	53	23	2431	0.95%
IT	47	19	1097	1.73%
ES	4	2	991	0.20%
PT	2	1	125	0.80%
	320	102	-	-

- Aunque las comparaciones son odiosas, en esta tabla se observa una dura realidad: el número de peers y ASNs registrados en España que comparten datos es alarmantemente bajo.
- No conozco las razones: ¿desconocimiento de la necesidad/importancia? ¿Miedo a compartir?
- Recordatorio: ¡el colector rrc18 se encuentra alojado en CATNIX !

Datos BGP disponibles: demografía y limitaciones

- Existen limitaciones a lo que podemos observar a través de los peers de RIS y Route Views puesto que los datos están sesgados e incompletos
 - 2659 peers en ambos proyectos
 - de los cuales, solo 1295 comparten FRT
 - provenientes de 403 ASNs únicos – total de 70,000 ASNs en Internet
 - 23.78 % de los ASNs están registrados en Estados Unidos, 7.34% en Suiza y China, y 6.41% en Australia.
 - Los 5 países que más feeds comparten (Estados Unidos, Suiza, Alemania, Australia y Reino Unido) contribuyen la mitad de las feeds en ambos proyectos
 - Hay feeds de ASNs registrados en 69 países – o dicho de otro modo, no existen datos de cerca de 130 países.
- En conclusión:
 - Datos de Route Views son muy “US-centric” por lo que ofrecen una vista polarizada y muy diferente de la que se ve desde Europa.
 - Datos de RIS son más “European-centric” pero no todos los países aportan feeds.
 - Debido a que Europa son países con diferente lengua y cultura, el tráfico es más regional y usa IXPs, lo que contribuye a no tener vistas disponibles dado que el enrutamiento se escapa de los route collectors

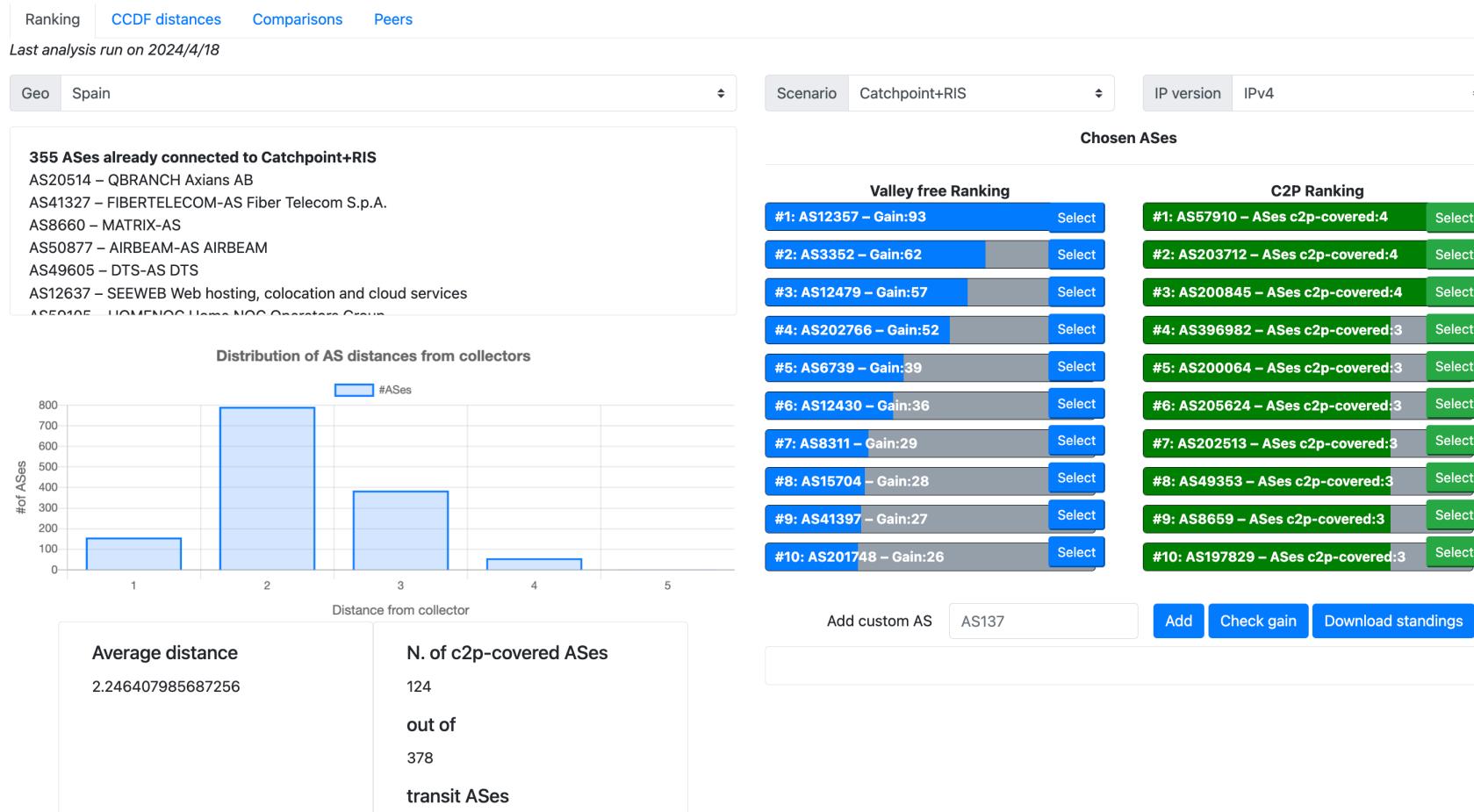


Soluciones para mejorar datos BGP disponibles

- El sistema **ideal u óptimo** es aquel en el que todos los ASes que componen Internet comparten su información BGP
 - No es realista puesto que es inviable llegar a todos los ASes
- El sistema **sub-óptimo** pero realista e implementable es aquel en el que maximizamos el número de ASes que podemos observar usando el mínimo número de RCs.
- **Solución:** desplegar infraestructura propia para mejorar los datos de RIS y Routeviews
 - Beneficios
 - selección de que ASNs se quieren añadir para obtener mejores datos
 - obtención de información en tiempo real
 - Desventajas
 - desarrollar nuestro colector de rutas
 - obtener nuevos peers y gestionar las sesiones
 - coste de operación €€

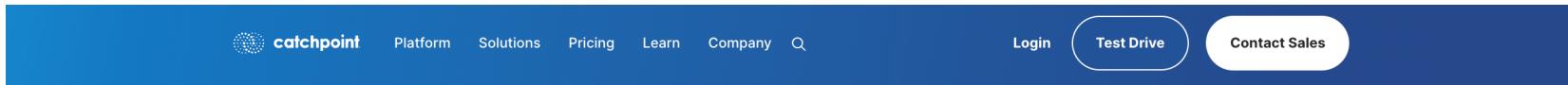
Soluciones: metodología de selección de ASNs

- Desarrollo de una metodología (patente obtenida) y herramienta asociada que nos permite seleccionar aquellos ASNs que nos proporcionan mayores ganancias



Soluciones: programa de colaboración BGP

- Desarrollo de un programa de colaboracion BGP con ASNs en el que se ofrece servicios de monitoreo a cambio de feeds BGP. Mas info en <https://www.catchpoint.com/bgp/partner>



Catchpoint BGP Partner Program

At Catchpoint, we run a BGP partner program with network operators around the world with the purpose of receiving real-time routing information updates to support our BGP Monitoring solution.



As of February 2023, we receive and analyze routing data from more than 140 peers in all five continents. Collected BGP data from the program is combined with RIPE RIS and Route Views data and presented to our customers through Catchpoint's Internet Performance Monitoring (IPM) platform.

We are looking to establish additional BGP sessions and expand the number of network operators from which we collect their full routing tables. In return, you will get access to our Internet Performance Monitoring platform for BGP monitoring and other performance related measurements. More information can be found in the [BGP partner program information datasheet](#).

If you are interested in joining the program, please complete the form or email us at peering@catchpoint.com and we'll be in touch.

Become a BGP Partner

First Name*

Last Name*

ASN*

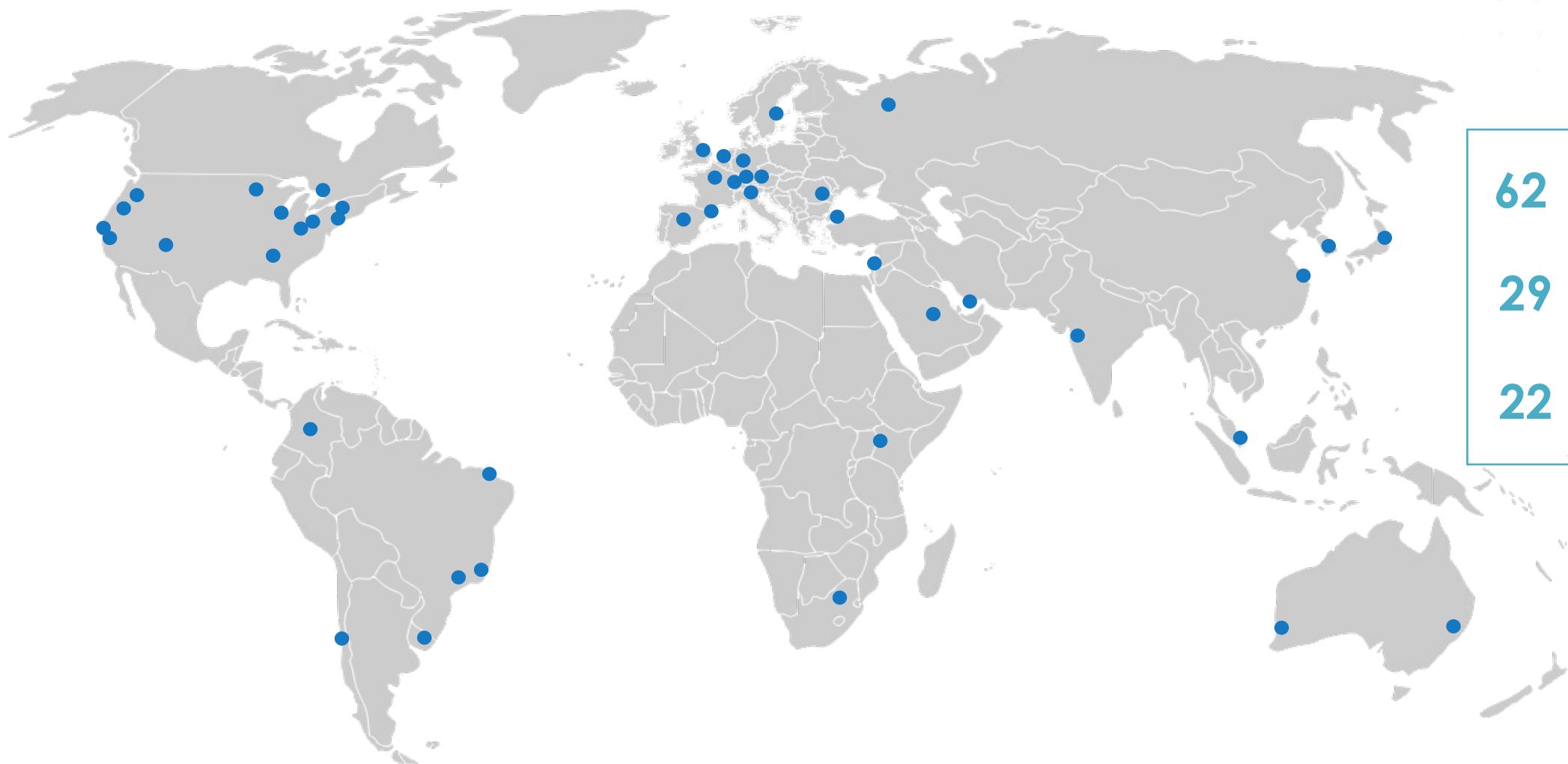
Company*

Email*

By signing up, you agree to our [Privacy Policy](#) and [Terms and Conditions](#).*

Submit

BGP monitoring infrastructure – April 2024



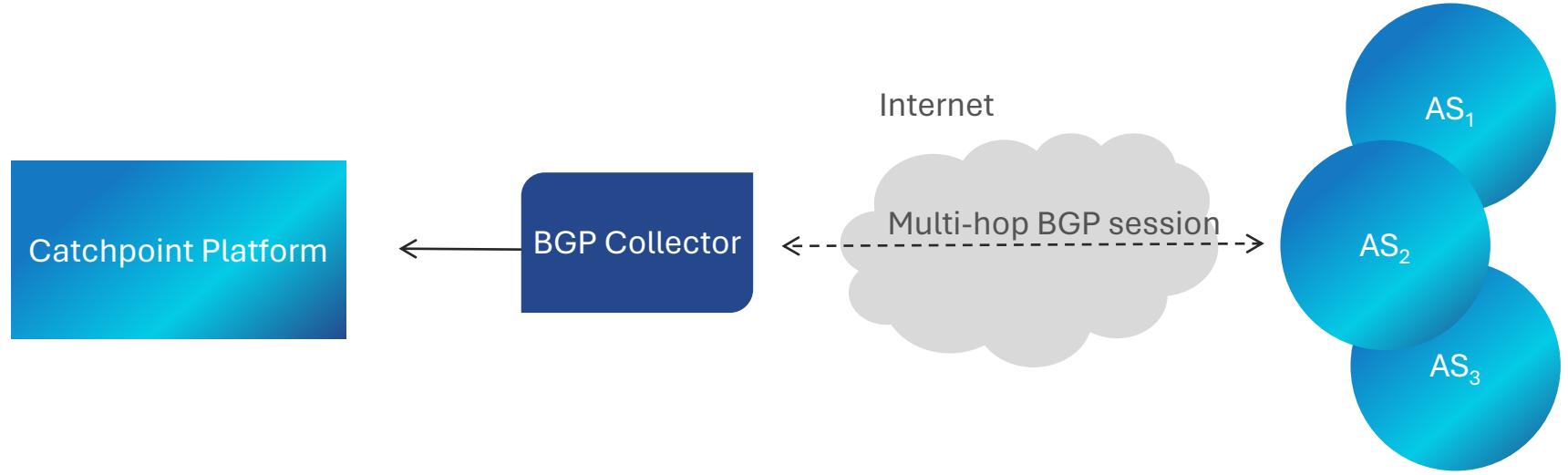
62	Catchpoint route collectors
29	Route Views route collectors
22	RIPE RIS route collectors



Type 1: Multi-hop collectors



Type 1: Multi-hop collectors



- 3 global multi-hop collectors located in US, Europe and Asia.
- With multi-hop, no need to be on the same subnet.

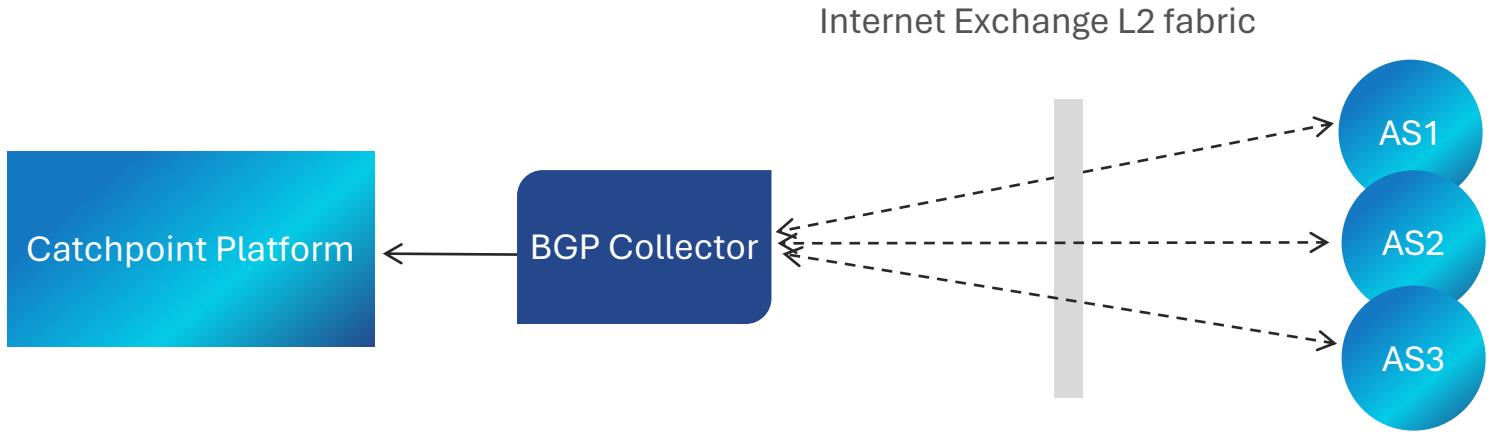
Type 2: Collectors at Internet Exchanges



11 Active IXP connections

AMS-IX Amsterdam
DE-CIX Frankfurt
DE-CIX Marseille
JINX Johannesburg
LoNAP London
NAPAfrica JNB
PIT Chile Santiago
Piter-IX St Petersburg
SOX Belgrade
SGIX Singapore
TOPIX Turin

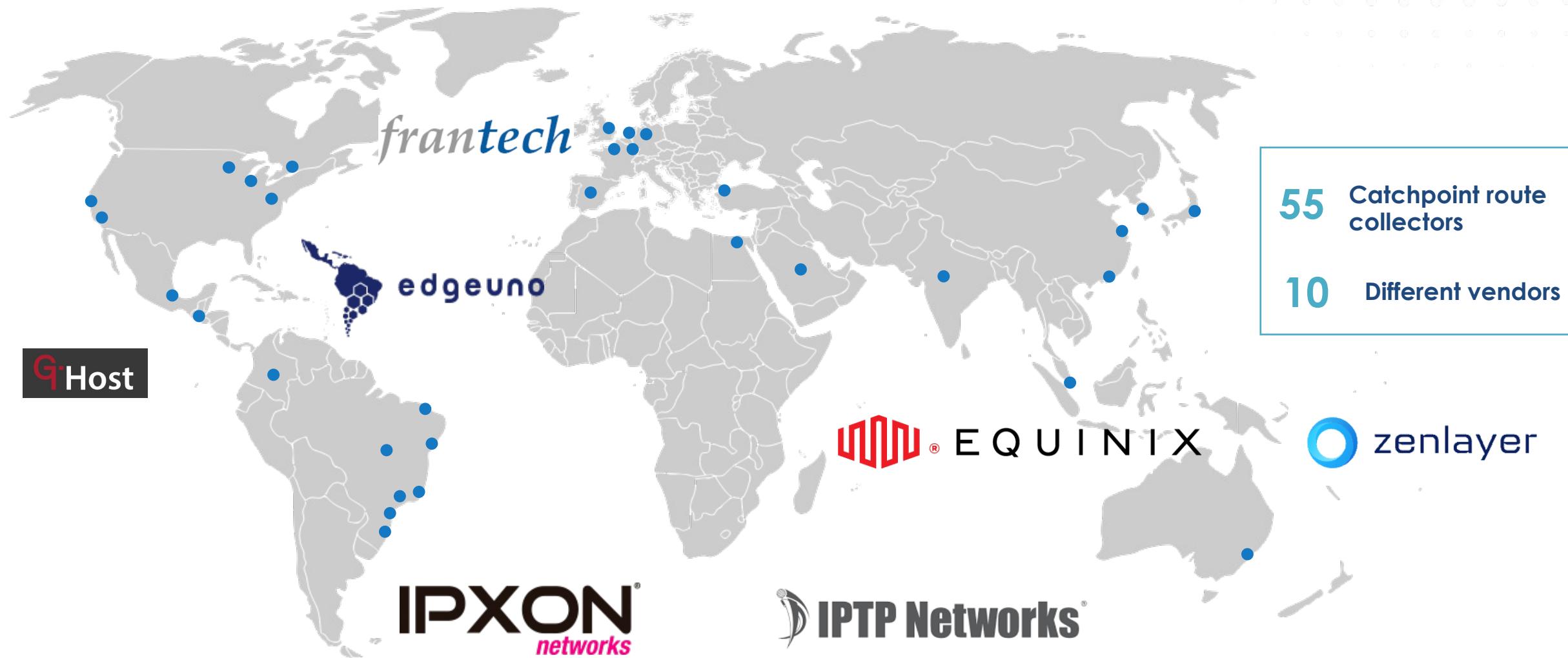
Type 2: Collectors at Internet Exchanges



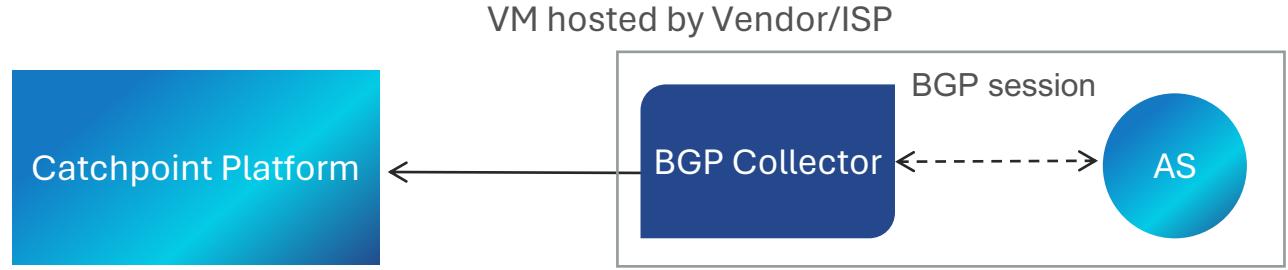
- We can setup multiple BGP sessions over an Internet Exchange
- In some PoPs, we connect to multiple IXs from the same server (eg. London)



Type 3: Collectors hosted by xSPs



Type 3: Collectors hosted by xSPs

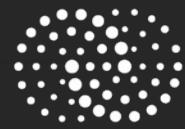


- The Vendor/ISP hosts our BGP collector on a VM.
- We setup a direct BGP session against their router.
- Need to deploy a new VM to consume each feed.



Conclusiones: ¿Porque compartir datos BGP?

- Monitorear el enrutamiento en Internet es un asunto global que requiere de la participación de los operadores de Internet a escala global (si, si, tú también).
- Se necesitan más datos de operadoras en España que den una vista “desde dentro del país”.
- Alimentar herramientas usadas por operadores de red, investigadores, etc.
- Poner tu granito de arena y añadir la vista de Internet desde tu AS.
- No cuesta nada (o muy poco).
- Vale, me has convencido. ¿Como comarto los datos?
 - **RIS**: habla con María Isabel de CATNIX para conectarte al RRC18
 - **Route Views**: envía un email a help@routeviews.org
 - **PCH**: envía un email a peering@pch.net
 - **Catchpoint**: ¡habla conmigo durante la pausa café!



catchpoint

ghernandez@catchpoint.com