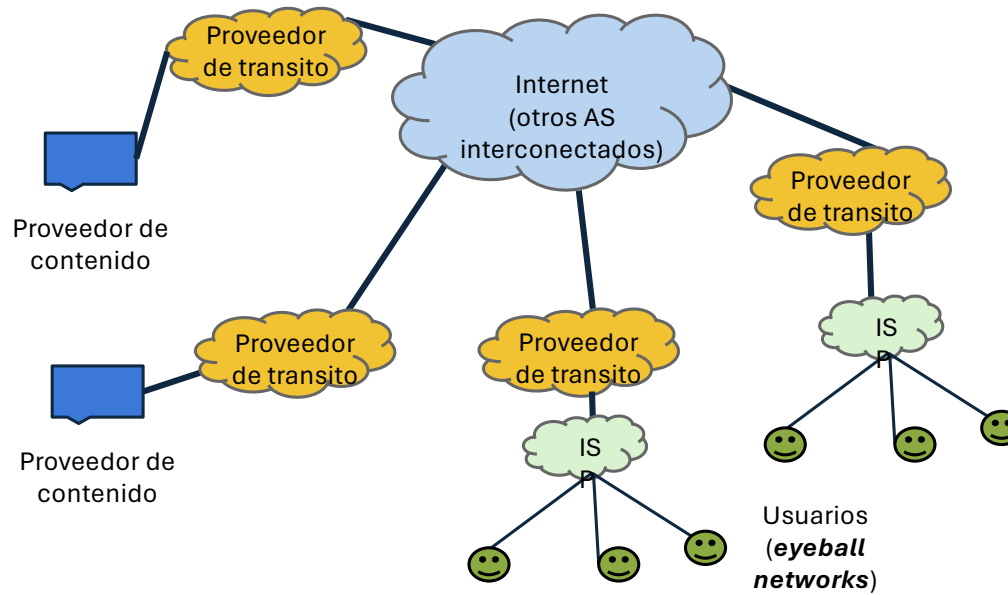


Tutorial de Peering

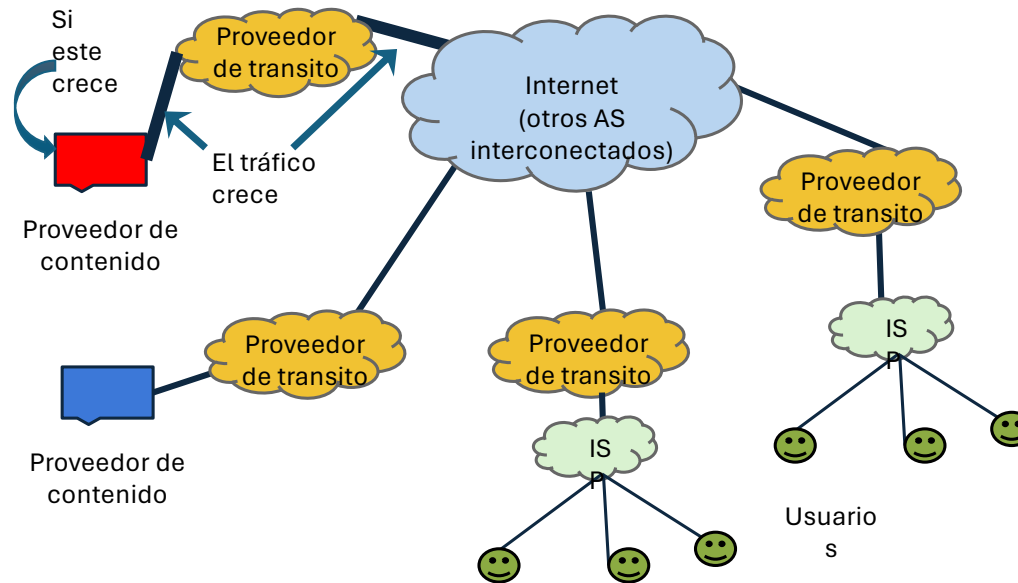
IXSY Meeting 2024

Mauricio Oviedo mauricio@socium.cr

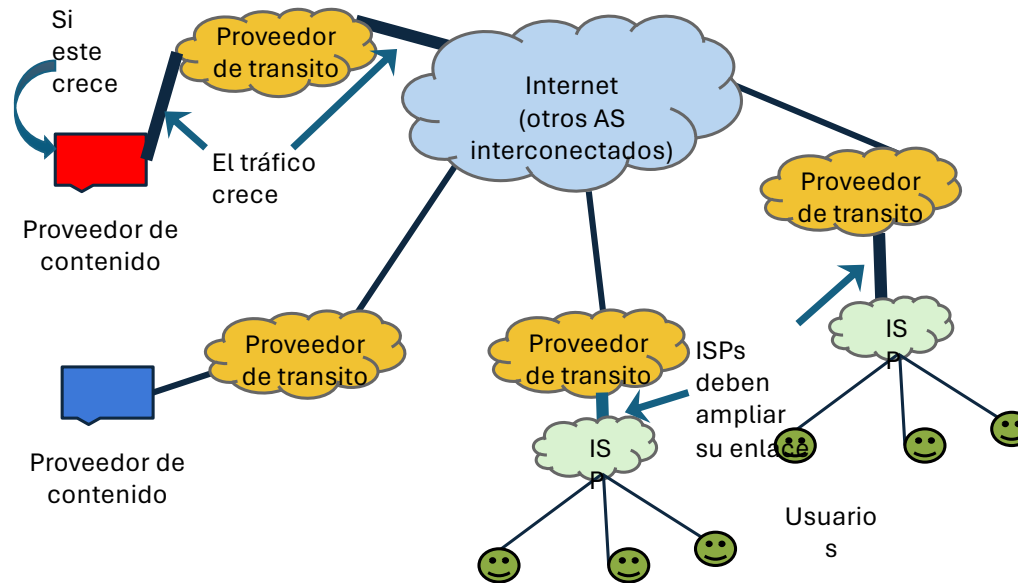
¿Cómo se encamina el tráfico?



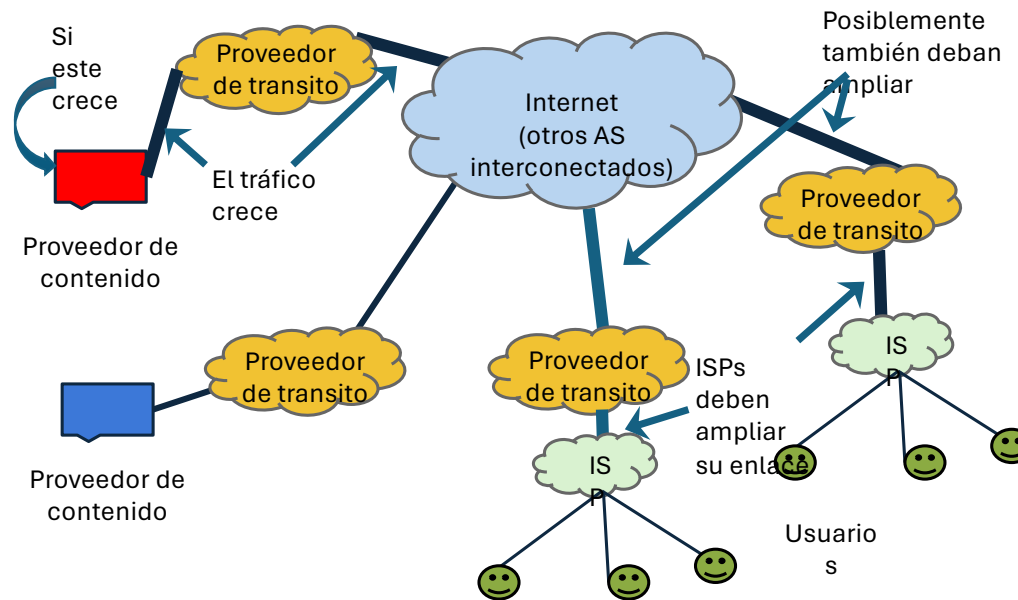
¿Cómo se encamina el tráfico?



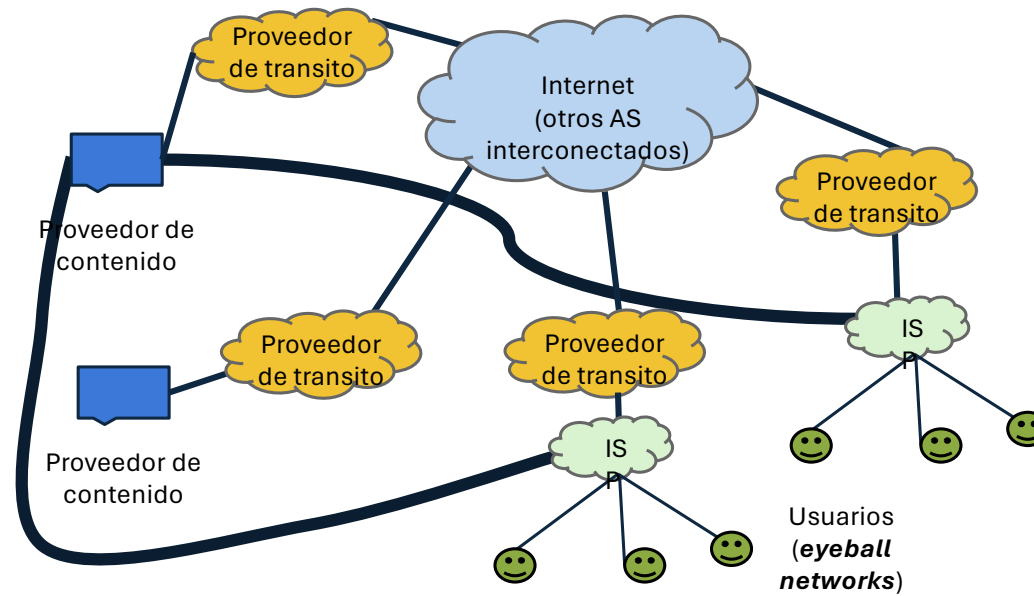
¿Cómo se encamina el tráfico?



¿Cómo se encamina el tráfico?

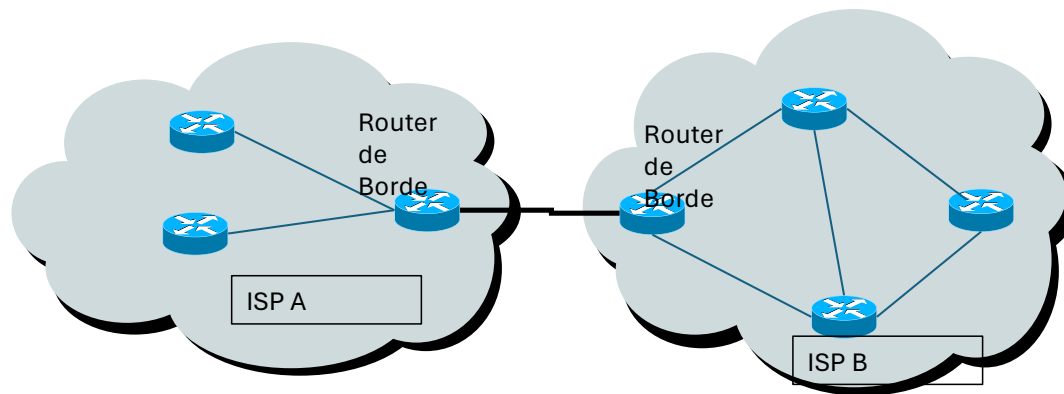


Alternativa: peering

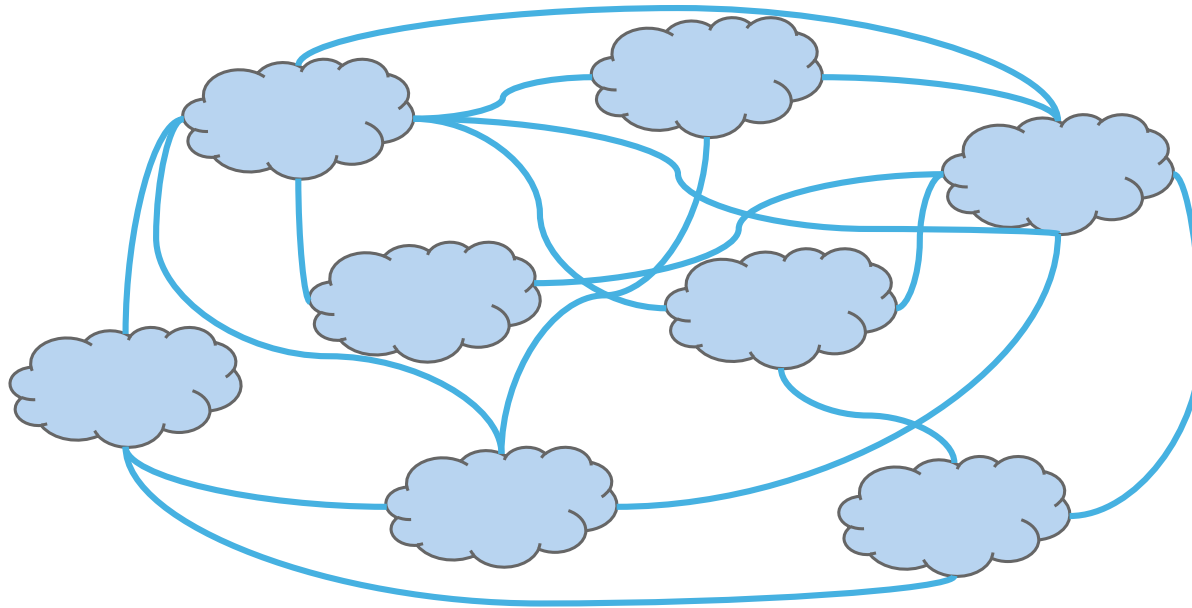


Modalidades de interconexión

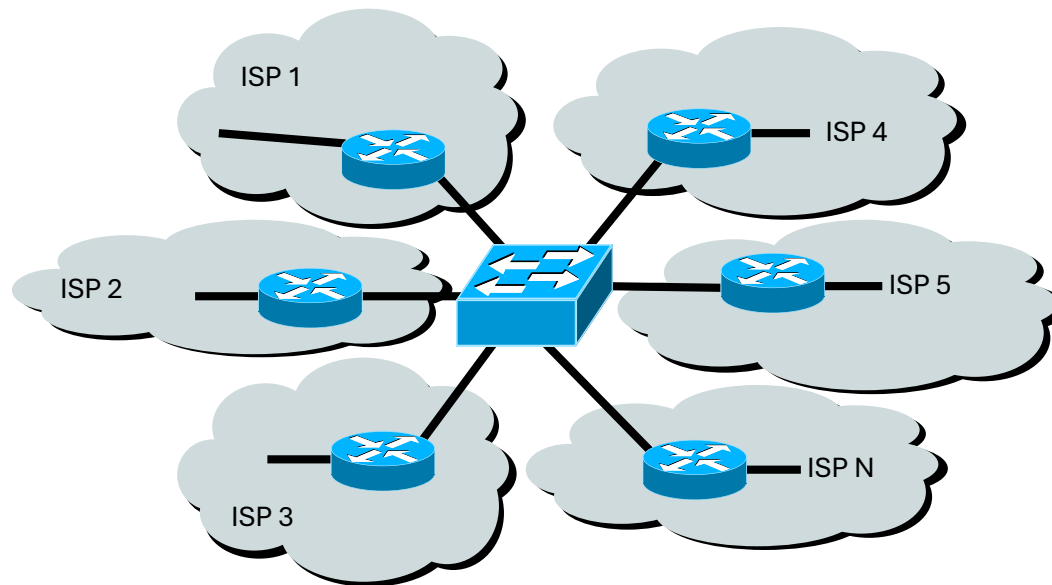
Interconexión directa: Peering



Interconexión directa: puede ser compleja

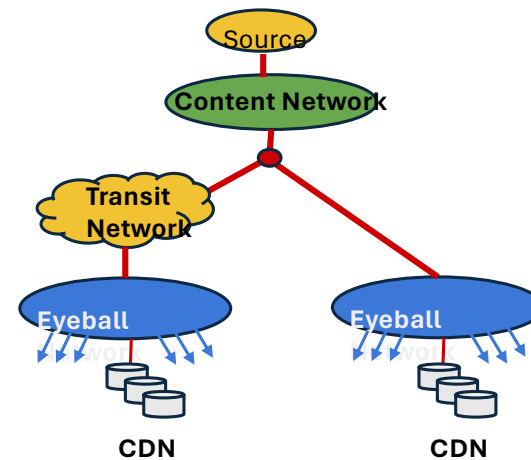


Interconexión pública



Qué es una CDN (Content Delivery Network)?

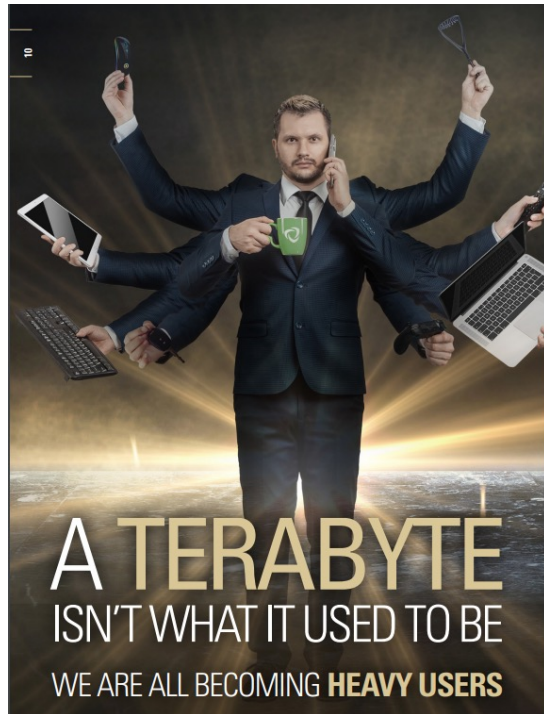
- Plataforma distribuida para entrega de contenido
- Sirve contenido más cerca de los usuarios
- Mejora el desempeño de los servicios a los usuarios
- Menor costo para el proveedor de contenido y el ISP



Ejemplos de CDNs

- CDNs Tradicionales y Telco
 - Akamai
 - Cloudflare
 - Level3
 - Limelight Networks
- Content Provider own-CDNs
 - Google
 - Netflix
 - Facebook

Realidad del tráfico de Internet en la actualidad



- El uso global de BW aumentó un 23% en 2023.
- El ancho de banda de Internet internacional total ahora asciende a 1217 Tbps, lo que representa una tasa compuesta anual de 4 años del 28%.

Fuente: Telegeography

Estudio Sandvine

Regional Phenomena Americas

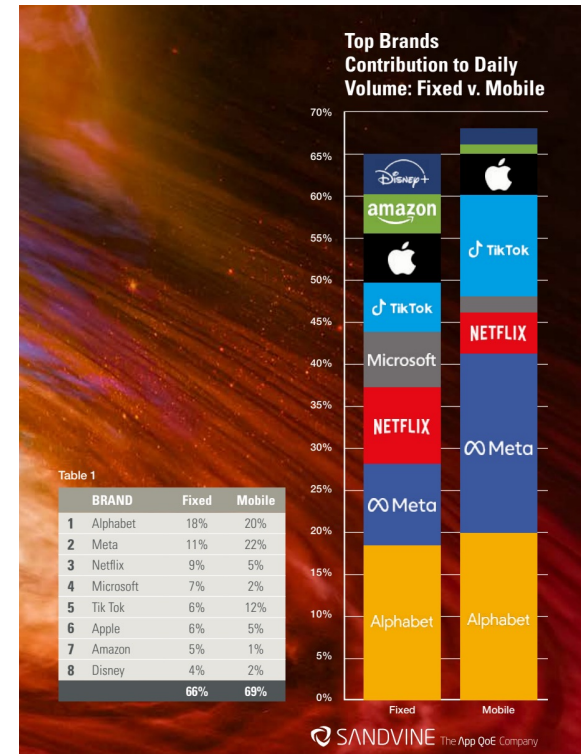
Table 20

Top App Categories by Downstream Volume 2023 – Americas			
Downstream Volume			
Application Category	% DS Vol	Sub. Volume	
1	Video	48%	5.643 GB
2	Television	17%	1.978 GB
3	Social Media	10%	1.122 GB
4	Device Gaming	9%	1.086 GB
5	File Sharing	6%	726 MB
6	General Web Apps	3%	352 MB
7	VPN	1%	173 MB
8	Communication	1%	130 MB
9	Audio	0.9%	110 MB
10	Conferencing	0.3%	36 MB
11	Cloud Gaming	0.1%	13 MB
12	Peer To Peer	0.08%	9 MB
13	IoT	0.04%	5 MB
14	Other Apps	3%	405 MB

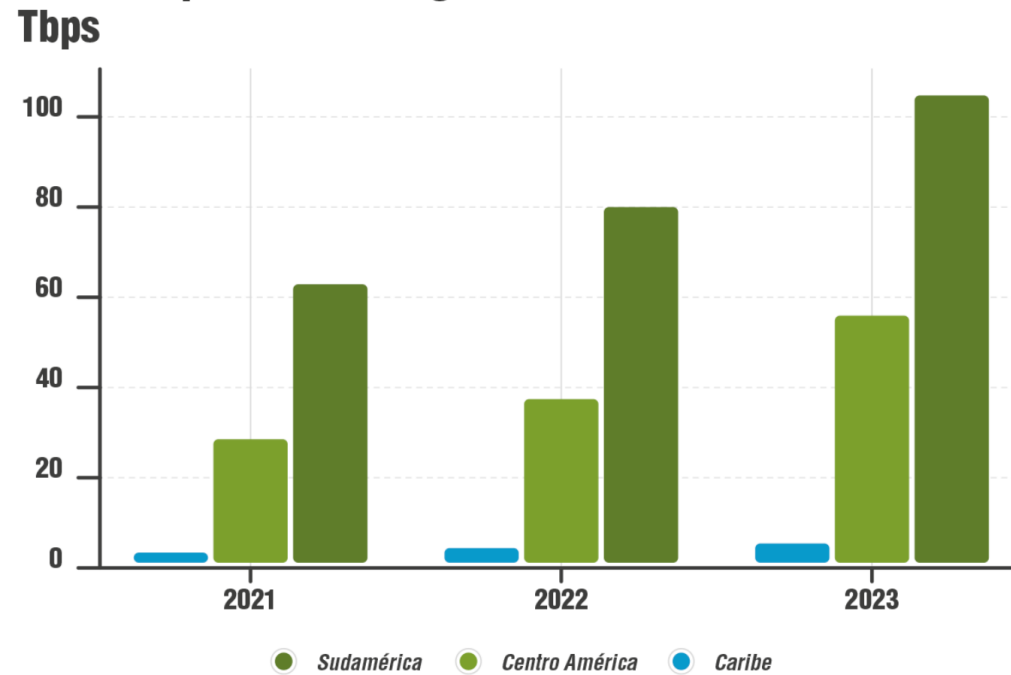
Table 21

Top App Categories by Upstream Volume 2023 – Americas			
Upstream Volume			
Application Category	% US Vol	Sub. Volume	
1	File Sharing	27%	205 MB
2	Communication	14%	101 MB
3	Video	11%	85 MB
4	General Web Apps	7%	54 MB
5	Social Media	6%	44 MB
6	VPN	4%	27 MB
7	Device Gaming	4%	27 MB
8	Television	3%	24 MB
9	IoT	3%	21 MB
10	Conferencing	3%	20 MB
11	Peer To Peer	1%	9 MB
12	Audio	0.4%	3 MB
13	Cloud Gaming	0.04%	323 KB
14	Other Apps	17%	129 MB

https://www.sandvine.com/hubfs/Sandvine_Redesign_2019/Downloads/2024/GIPR/GIPR%202024.pdf



Ancho de banda internacional de Internet para las subregiones latinoamericanas



(Fuente: [Telegeography](#))

Plataformas OTT



Cantidad de usuarios en Internet



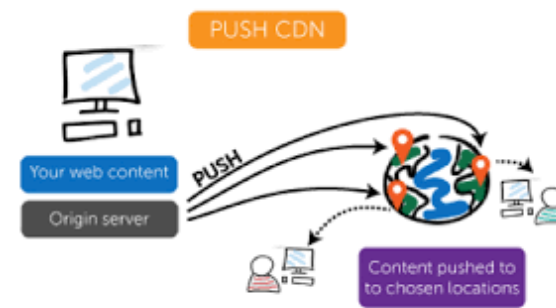
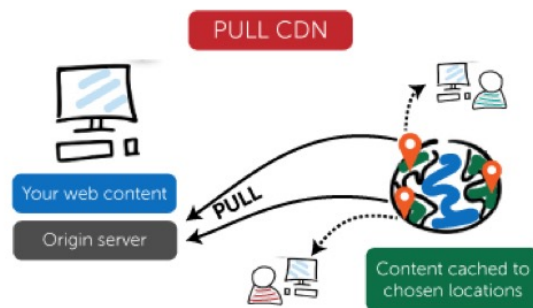
Contenido de Streaming



Calidad de video 4k- 16 k

	Video	Games	Social	Messaging	Enterprise Conferencing
1	YouTube	Player Unknown's Battlegrounds	Facebook	WhatsApp	Zoom
2	Netflix	ROBLOX	TikTok	Discord	Microsoft Teams
3	Facebook video	League of Legends	Instagram	Facebook Messenger	Webex
4	TikTok	Fortnite	Wordpress	LINE	Blackboard Collaborate
5	HTTP media stream	Minecraft	Snapchat	Skype	Amazon Chime
6	Disney+	Garena Free Fire	Twitter	Zoom	Canva
7	Amazon Prime	Call of Duty	Reddit	Microsoft Teams	Udemy
8	Twitch	Mobile Legends	Wattpad	Telegram	Cisco Spark
9	Hulu	Candy Crush	Pinterest	WebEx	GoToMeeting
10	HBO	War Thunder	GIPHY	WeChat	Steam

Modelos de entrega de contenido en las CDN



Definiciones básicas

Definiciones



Tránsito vs Transporte

Tránsito

- Usualmente servicio en capa 3 (IP).
 - Puede ser BGP o no
- Costo en base a Mbps
- Utilizado para enviar tráfico a muchos sitios
- El tráfico depende de quien da el servicio como upstream provider

Transporte

- Usualmente servicio en capa 2: Metro Ethernet, SDH, etc.
- Costo fijo por capacidad de enlace (1Gbps, 10 Gbps).
- Utilizado para conectar dos sitios
- El tráfico queda acotado entre las organizaciones que establecen el transporte

Puntos de Intercambio de tráfico: IXPs

Importancia y Beneficios

Características de un IXP

Un IXP es un sitio donde los ***operadores de red*** se interconectan

- Otros nombres: PIT, PTT, NAP (anteriormente)
- Infraestructura compartida intercambiar tráfico:
 - ISPs, Proveedores de Contenido, Universidades, Medios, Bancos, etc.
- Normalmente habrá varios AS que se interconectan, lo que lo distingue de un peering privado que se hace entre dos redes.
- Un IXP es distinto de una red de acceso y de una red de tránsito/carrier
 - La función del IXP es interconectar redes, no proveer acceso ni actuar como un proveedor de tránsito o carrier.
 - Un IXP permite interconectar redes que son organizaciones separadas: sistemas autónomos independientes.
 - Un IXP no requiere que el tráfico entre dos AS pase por un tercero

Algunas ventajas de los IXPs (*estabilidad y resiliencia*)

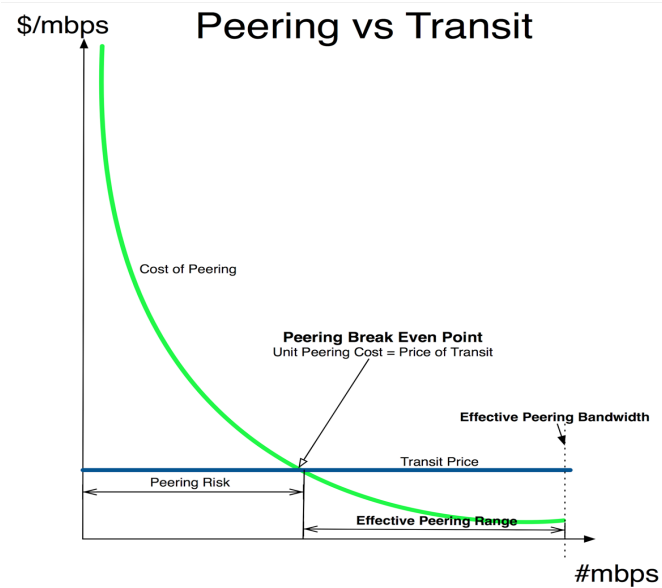
- Tráfico local se rutea localmente
- Menor latencia para las aplicaciones
- Menores costos
- Posibilidad de CDNs
- El tráfico de una región/pais/zona no es visto desde otras regiones/paises
- Introduccion de nuevas tecnologias (IPv6, RPKI, etc)
- Acciones coordinadas ante incidentes de seguridad, problemas técnicos, etc.
- Sentido de "comunidad"
 - Compartir problemas, estrategias, acciones en común

Comparación de costos

Transporte al sitio del IX	Costo fijo por cierta capacidad
Colocation	Fijo
Hardware	Fijo
X-connect	Fijo
IXP fee	Fijo

Transito	Basado en el uso
-----------------	------------------

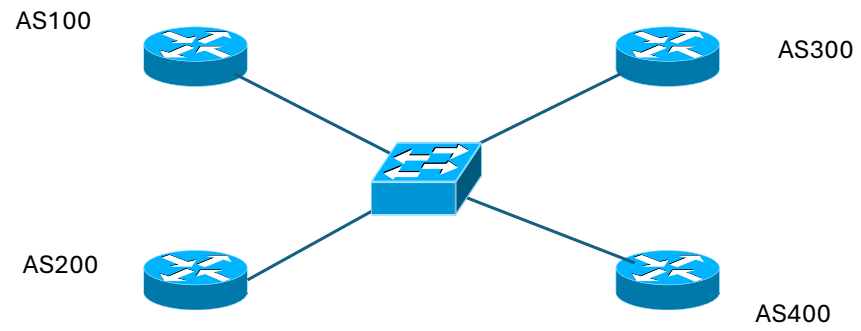
Peering vs. Transit: costos comparados



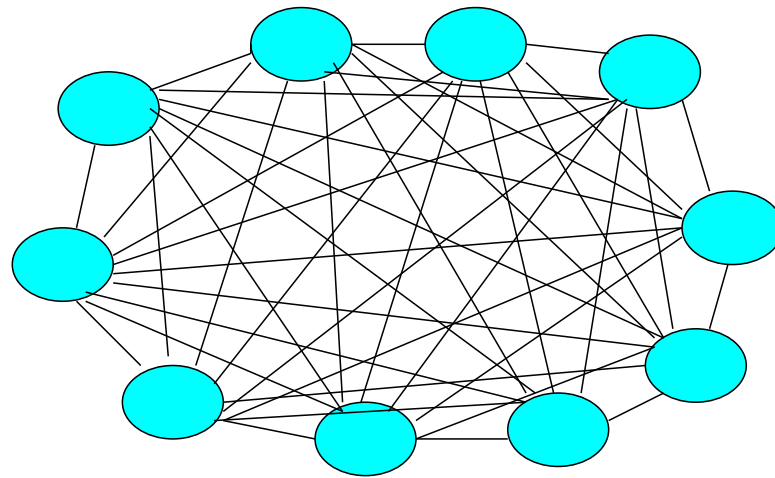
Source: [Dr Peering](#)

Esquema básico de un IXP

Esquema básico de un IXP

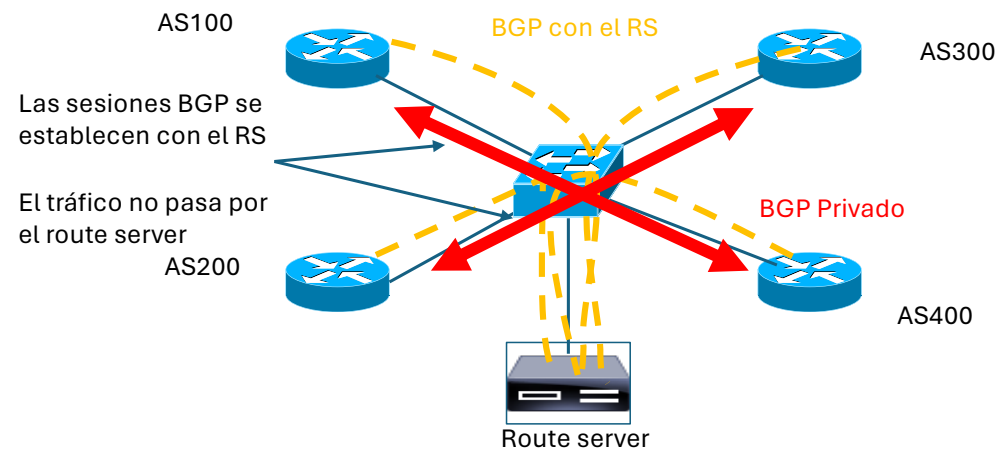


Sin route-server: malla N-cuadrado



ROUTE SERVERS (RS)

Uso de route server en un IXP



Route Servers ¿Qué es?

- Normalmente es un Servidor Unix que corre software de Enrutamiento.
 - Existen soluciones Open Source para esto
- Ruteador que activa la funcionalidad de BGP
- Intercambia la información de ruteo con ruteadores de proveedores de servicio en un IXP basado en políticas
- No envía paquetes – únicamente maneja la lógica de ruteo
- Evita una enorme cantidad de sesiones de BGP
 - Número de sesiones = $n(n-1)$

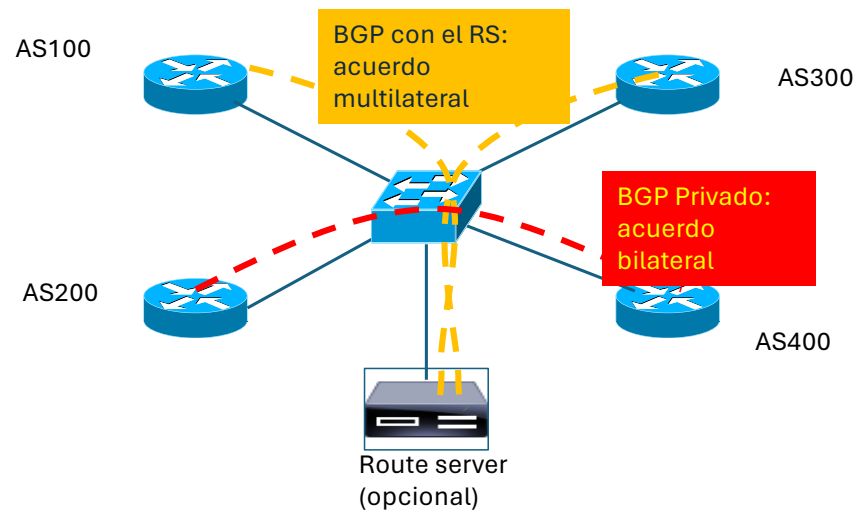
Seguridad: ventajas de un route server

- Medidas básicas: filtrado de ASNs y prefijos bogon, filtros por cliente, etc.
- Evita route-leaks que pueden provenir de errores de configuración
 - Ejemplo: si se filtra una full-table al RS
 - Es un beneficio aún para ISPs que no hacen peering con el RS: sus rutas no se fugarán al resto de los ISPs.
- Posibilidad de implementar filtros por RPKI, por IRR, whois, etc.

Ejemplos de route-servers por software

- arouteserver: <http://arouteserver.readthedocs.io>
 - Herramienta en Python para generar configuración para route servers
 - Produce configuraciones para BIRD y OpenBGPD
 - Soporta IRR, RPKI, WHOIS
 - Soporta PeeringDB para obtener los AS-SETS
 - Simple de integrar con otros sistemas
- IXP manager: <https://www.ixpmanager.org>
 - Es un Sistema de administración completo para IX
 - Incluye un portal para administración del IXP y para los miembros
 - Produce configuraciones para BIRD

Interconexión en un IXP



Tipos de Acuerdo

Acuerdos Bilaterales

- Cada proveedor establece la relación que necesite con otros proveedores en el IXP
- Los enrutadores de borde de los ISP establecen sesiones de BGP con los enrutadores de borde de otros proveedores

Acuerdos Multilaterales

- Cada proveedor establece sesiones con el concentrador
- Los enrutadores de borde de los ISP tienen como vecino al IXP

Referencias

- Cursos de Campus de LACNIC: <https://campus.lacnic.net> (BGP y RPKI)
- Tutorial de BGP y RPKI de LACNIC32:
<https://www.lacnic.net/3900/52/evento/tutoriales>
- Internet Exchange BGP Route Server –
<https://tools.ietf.org/html/rfc7947>
- Internet Exchange BGP Route Server Operations -
<https://tools.ietf.org/html/rfc7948>
- A Border Gateway Protocol 4 (BGP-4) -
<https://tools.ietf.org/html/rfc4271>

¿Preguntas hasta acá?

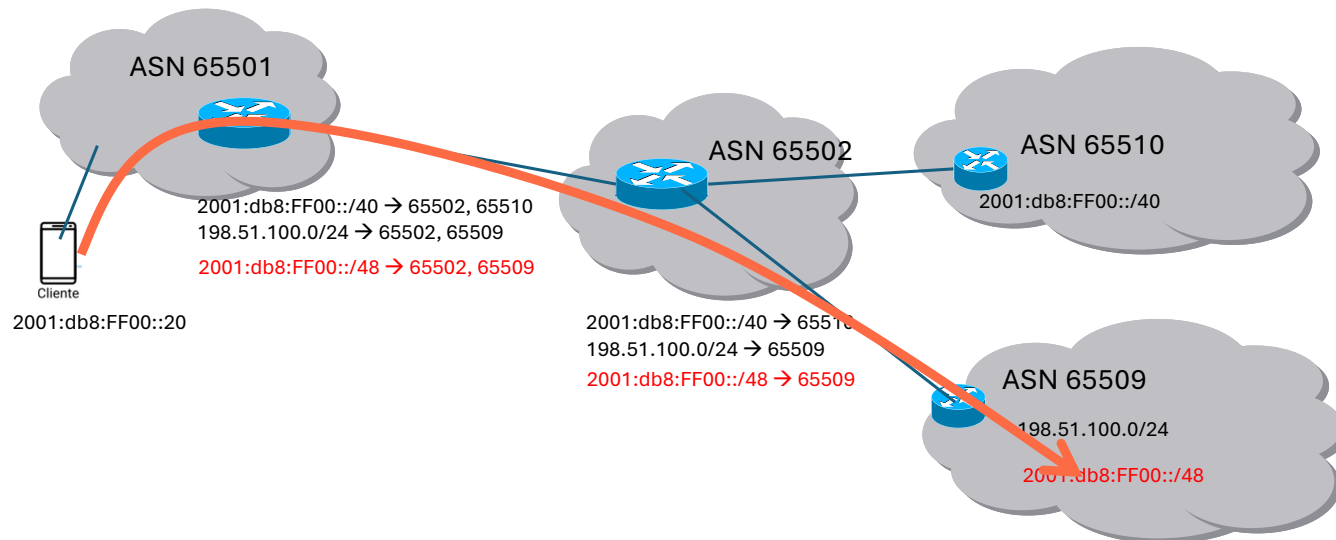


Seguridad en ruteo

Secuestro de rutas

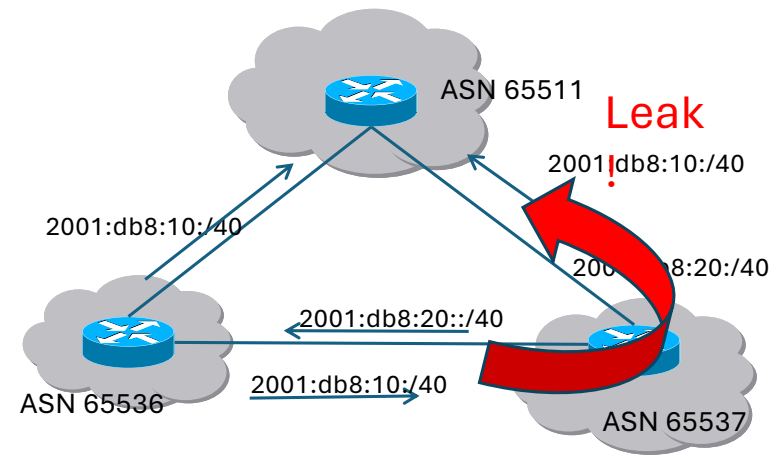
Secuestro de rutas:
Acción de anunciar
prefijos NO
autorizados

Intencional.
Por error en la operación.



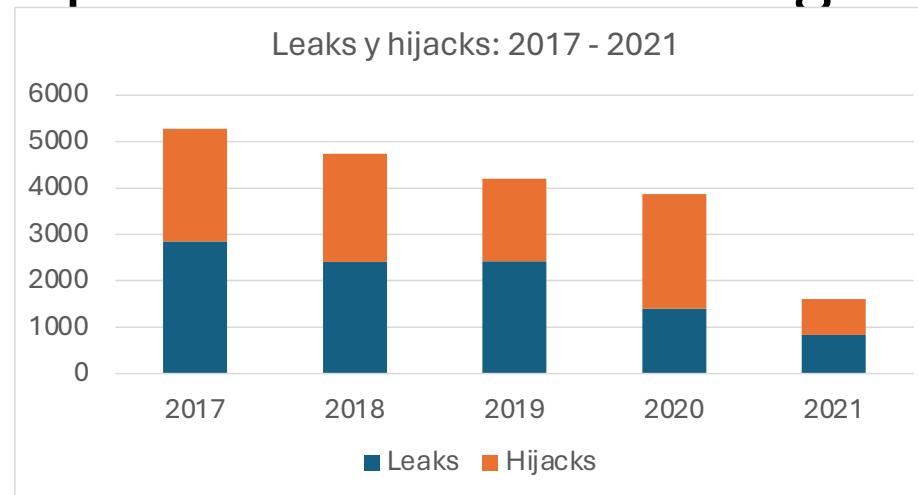
Route leaks – fuga de rutas

- Prefijos aprendidos del **proveedor** no deben anunciarse a otro **peer** o a otro **proveedor**
- Prefijos aprendidos de un **peer** tampoco se anuncian a otros **peers** ni al **proveedor**
- Esos prefijos solo deberían anunciarse a **clientes**



Si no hay filtros configurados, esto trae problemas

Principales incidentes de seguridad



Fuentes:

Informe sobre seguridad en el ruteo de LAC – Augusto Mathurín, 2019

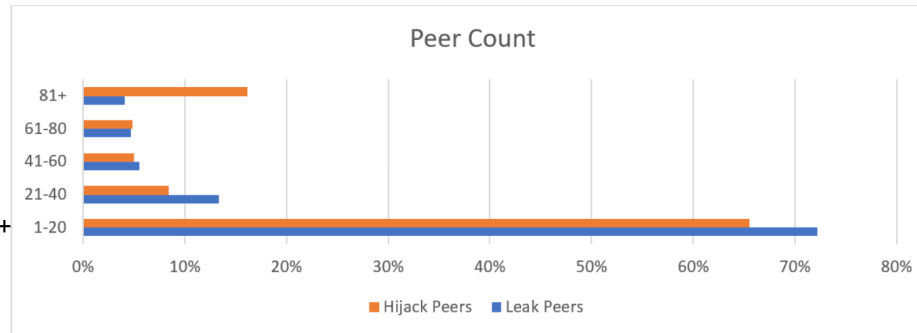
<https://www.lacnic.net/innovaportal/file/4297/1/fort-informe-seguridad-ruteo-es.pdf>

MANRS: <https://www.manrs.org/2021/02/bgp-rpki-and-manrs-2020-in-review/>

MANRS: <https://www.manrs.org/2022/02/bgp-security-in-2021/>

Alcance de los incidentes

(mayor número de peers afectados indica mayor impacto)



En cuántos peers de colectores BGP se detectan estos hijacks/leaks?

- Más del 70% de los incidentes fueron detectados por 1 a 20 peers
- Menos peers recibieron las rutas incorrectas: fueron filtradas antes
- Las medidas de seguridad parecen estar funcionando

<https://www.manrs.org/2022/02/bgp-security-in-2021/>

¿Qué podemos hacer para mitigar los incidentes?

Acciones acordadas para promover la seguridad del ruteo

MANRS – Mejores prácticas

MANRS es un conjunto de "Normas Mutuamente Acordadas para la Seguridad del Enrutamiento"

Acciones propuestas por MANRS para **operadores**:

- Filtrado
- Anti-spoofing
- Coordinación
- Validación global

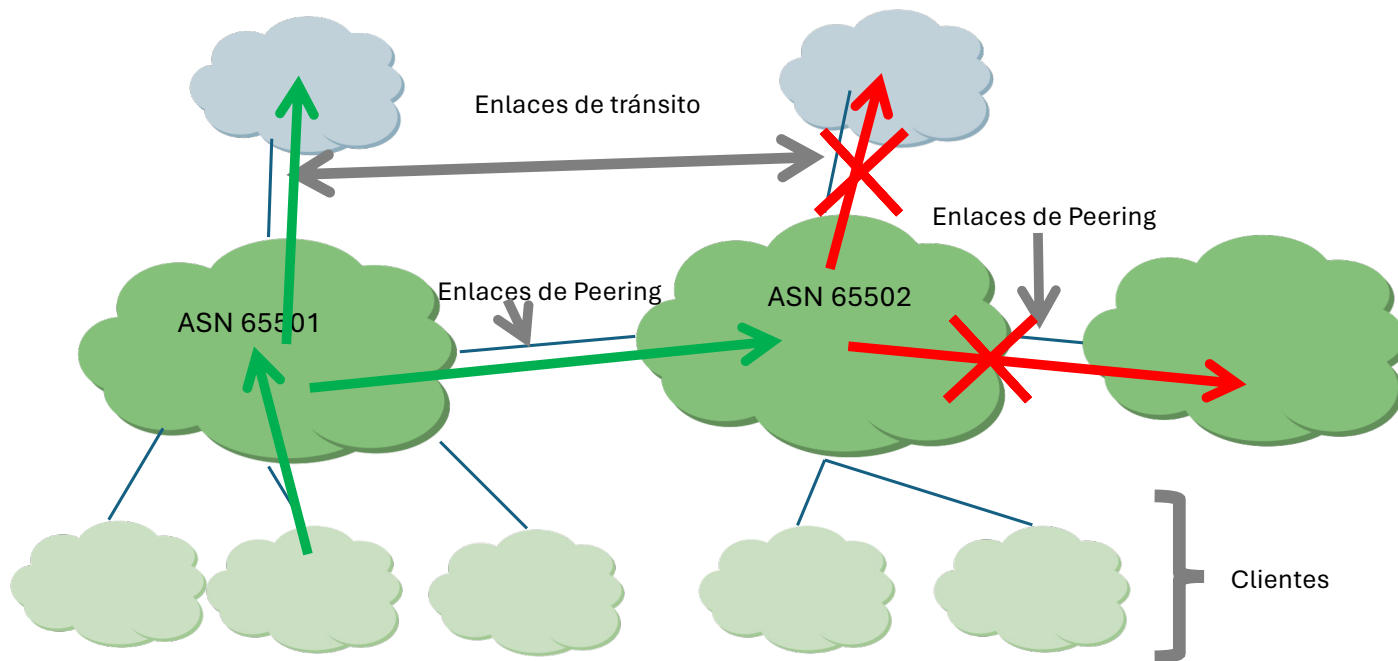
Veremos estas acciones en más detalle a continuación



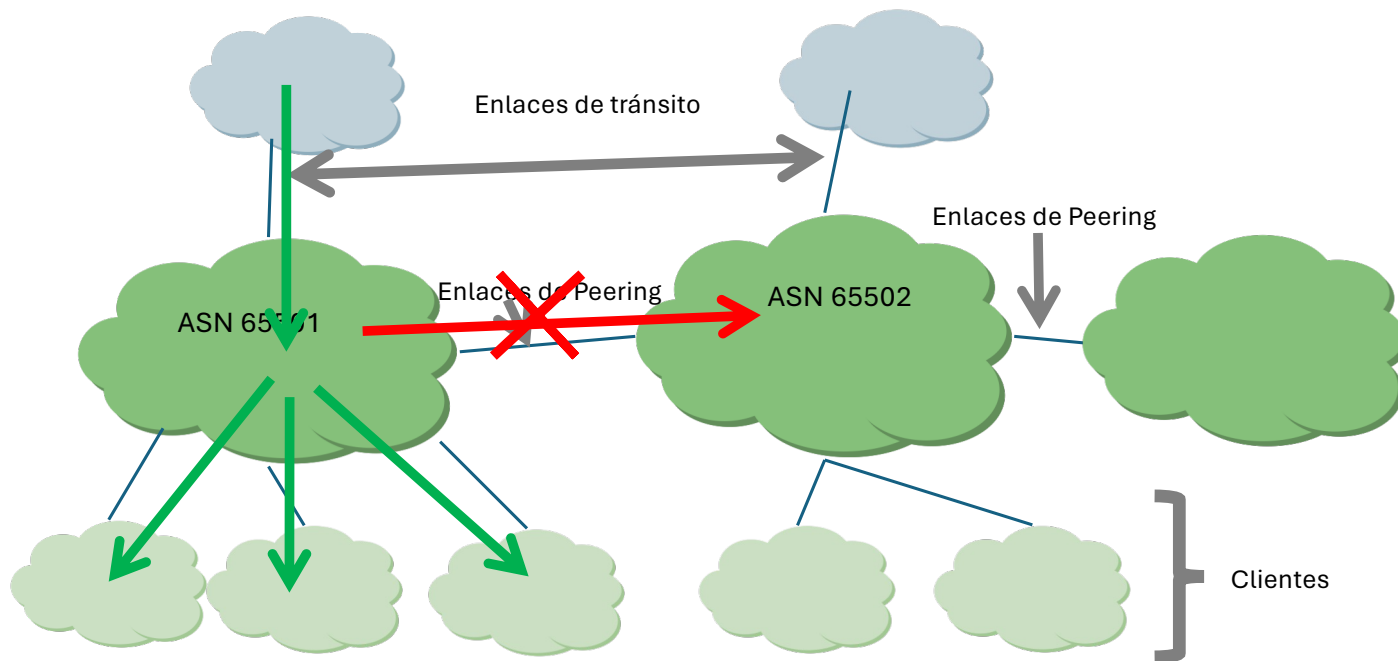
Hay también un programa específico para **IXPs** y para **CDNs**

<https://www.manrs.org>

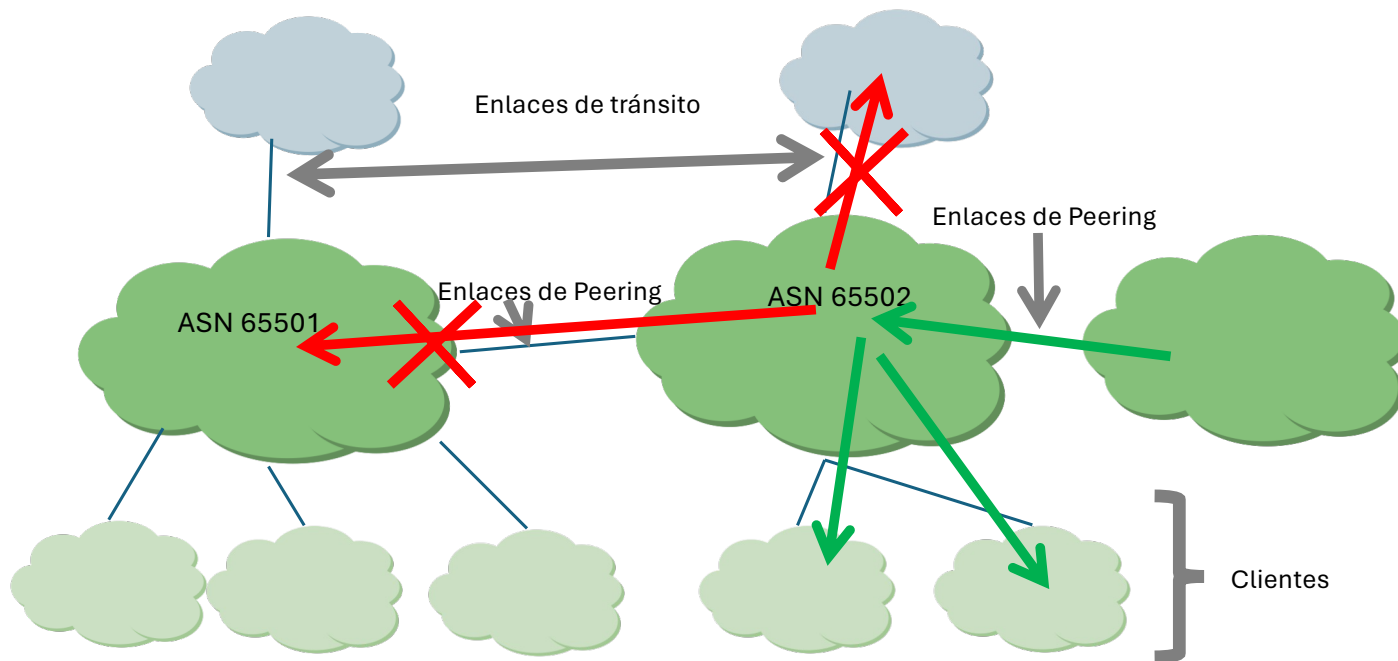
BGP: qué debemos anunciar y qué no



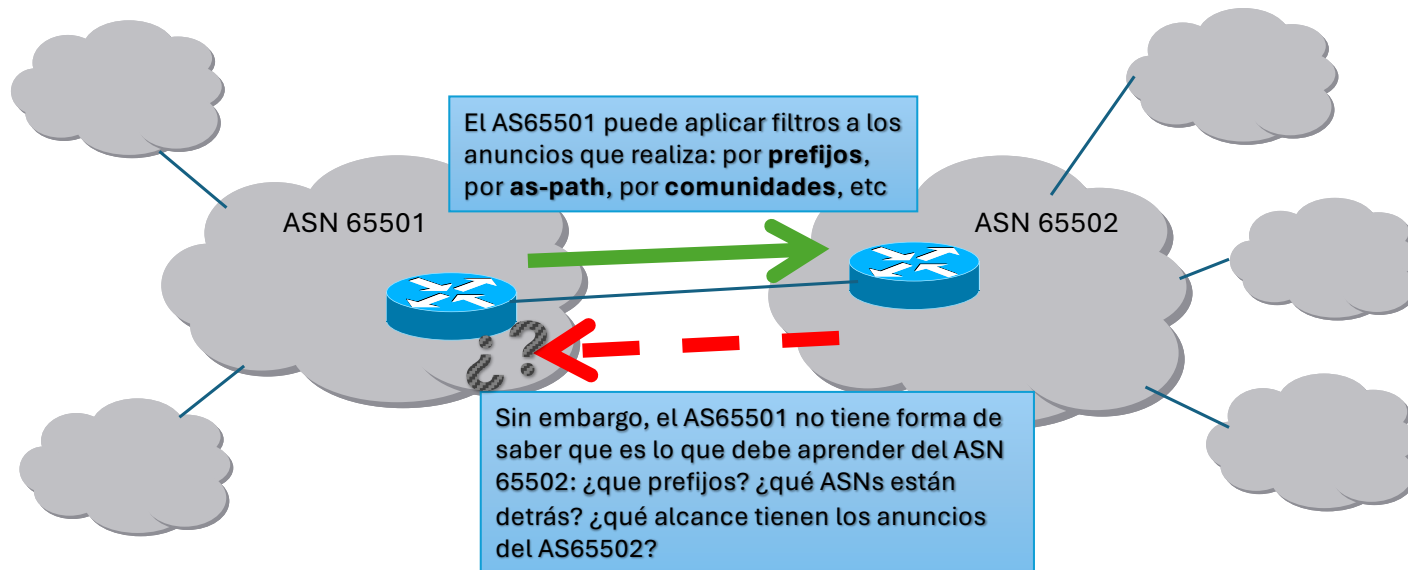
BGP: qué debemos anunciar y qué no



BGP: qué debemos anunciar y qué no

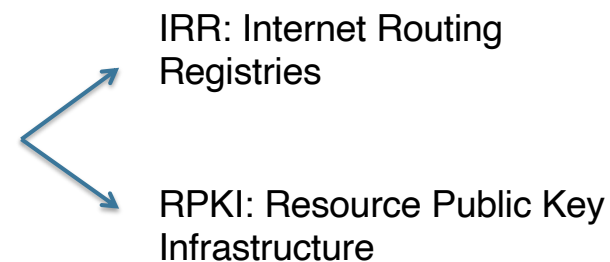


BGP: filtros de salida y entrada



IRRs vs RPKI

- Cómo chequear que la información que recibimos por BGP es correcta?
 - BGP no tiene mecanismos intrínsecos que permitan verificar esto
 - Se deben contrastar los anuncios recibidos por BGP contra fuentes externas
- Existen dos formas:



IRR – Internet Routing Registries

- Existe una gran cantidad de IRRs
 - El más conocido es RADB
 - RADB replica todos los demas IRRs
- Las organizaciones definen sus políticas de ruteo en un IRR
- Los operadores (ISP) utilizan esa información para generar filtros para BGP, muchas veces en forma automática
- Existen herramientas para utilizar esa información y configurar los routers: bgpq3/bgpq4, etc.

- AFRINIC
- ALTDB
- AOLTW
- APNIC
- ARIN
- BELL
- BBOI
- CANARIE
- EASYNET
- EPOCH
- GT
- HOST
- JPIRR
- LEVEL3
- NESTEGG
- NTTCOM
- OPENFACE
- OTTIX
- PANIX
- RADB
- REACH
- RGNET
- RIPE
- RISQ
- ROGERS
- TC

- Ahora también LACNIC

Ejemplos de registros

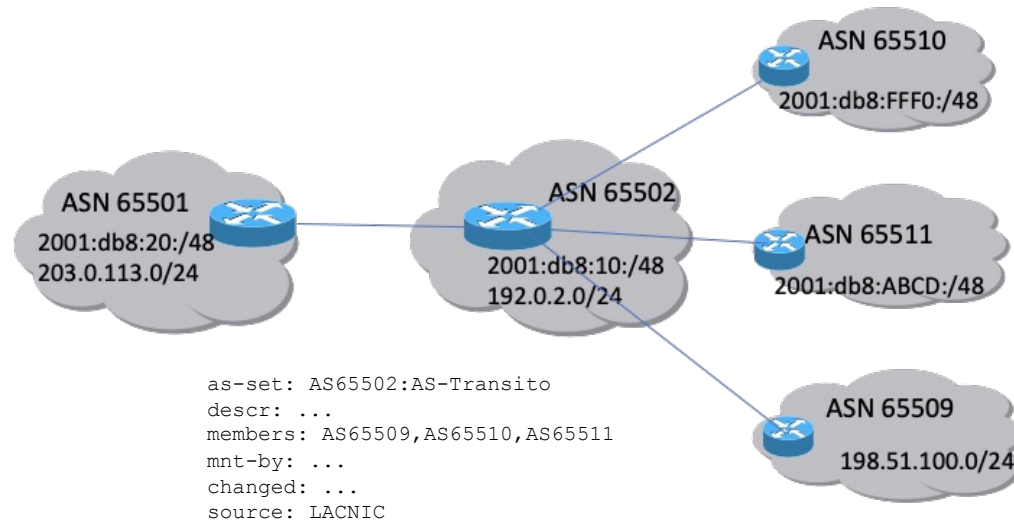
```
whois -h whois.radb.net -- '-s radb -i mnt-by MAINT-AS6057'
```

```
route:      201.221.32.0/19
descr:      ANTEL
origin:     AS6057
notify:     noc@antel.net.uy
mnt-by:     MAINT-AS6057
changed:    nantoniello@antel.net.uy 20080903
changed:    nantoniello@antel.net.uy 20080903 #19:20:32Z
source:     RADB
```

```
route:      201.217.128.0/18
descr:      ANTEL
origin:     AS6057
notify:     noc@antel.net.uy
mnt-by:     MAINT-AS6057
changed:    nantoniello@antel.net.uy 20080903
changed:    nantoniello@antel.net.uy 20080903 #19:21:34Z
source:     RADB
```

Cómo usar la información

Ejemplo de tránsito



Utilizando bgpq3/bgpq4

- En este caso, usamos el as-set:
- Prefijos IPv4

```
$ bgpq4 -h irr.lacnic.net -l clientes-as65502 AS65502:AS-Transito  
no ip prefix-list clientes-as65502  
ip prefix-list clientes-as65502 permit 198.51.100.0/24
```

- Prefijos IPv6

```
$ bgpq4 -h irr.lacnic.net -6 -l clientes-as65502 AS65502:AS-Transito  
no ipv6 prefix-list clientes-as65502  
ipv6 prefix-list clientes-as65502 permit 2001:db8:FFF0:/48  
ipv6 prefix-list clientes-as65502 permit 2001:db8:ABCD:/48
```

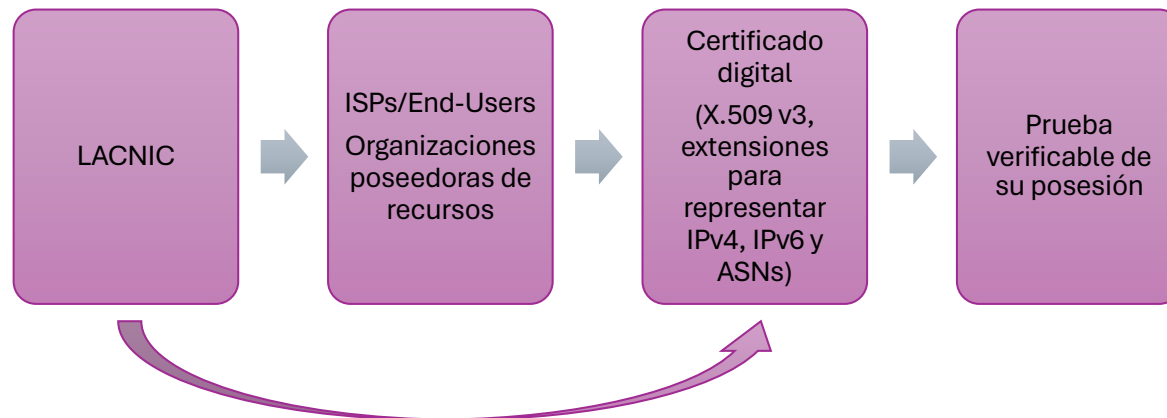
- Ver más información sobre bgpq4 en <https://github.com/bgp/bgpq4>

Referencias

- IRR de LACNIC: <https://labs.lacnic.net/Uso-de-IRR-LACNIC/>
- Peering, IRR y AS-SET: <https://www.labs.lacnic.net/Peering-IRR/>
- Bgpq4: <https://github.com/bgp/bgpq4>
- IRRd v4: <https://irrd4.readthedocs.io/en/master/users/queries.html>
- Documentación Mi LACNIC:
 - General: <https://lacnic.zendesk.com/hc/es/categories/360002625214-Internet-Routing-Registry>
 - RPKI: <https://lacnic.zendesk.com/hc/es/sections/206490008-RPKI>
 - IRR: <https://lacnic.zendesk.com/hc/es/categories/203940327-Soporte-Mi-LACNIC>

RPKI

- Define una infraestructura de clave pública especializada para ser aplicada al enrutamiento
 - En particular, para BGP

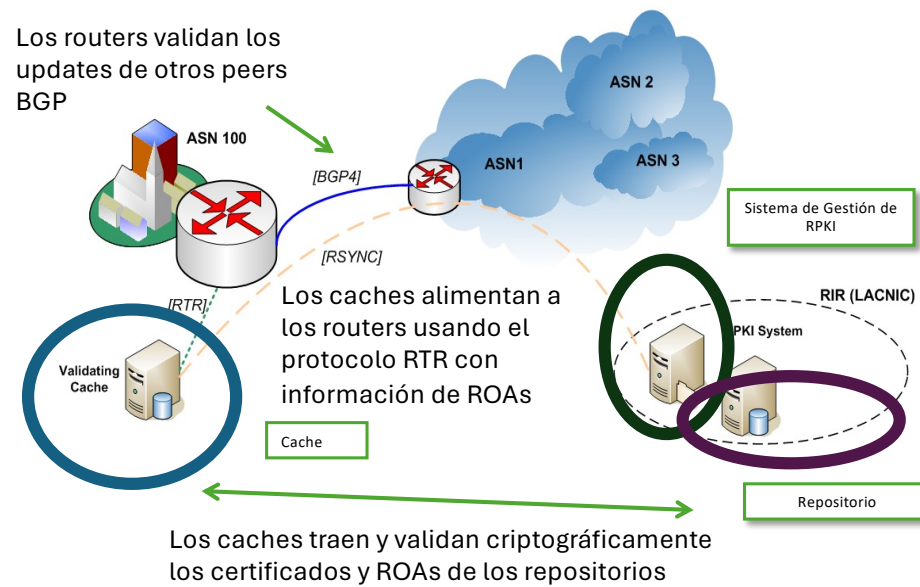


¿Qué compone la solución RPKI?

- **ROA:** Objetos firmados digitalmente para soportar seguridad del enrutamiento
 - Equivalentes a route o route6 objects de un IRR
 - Los ISPs u organizaciones pueden ***definir y certificar los anuncios de rutas que autorizan*** realizar
 - Los **ROAs** permiten definir el AS de origen para nuestros prefijos
 - **Firmados** con la clave privada del certificado
 - Toda la información es copiada en un **repositorio públicamente accesible**
- Un **mecanismo de validación** de prefijos
 - Validación de origen

Validación de Origen

RPKI en acción



Validación de Origen

- Una vez que los routers reciben la información de los caches, tendrán una tabla con:

Prefix	Length	Max length	Origin-AS
200.0.112.0	22	24	65501

- Con esto es posible asignar un **estado de validez** a cada UPDATE de BGP
- El estado de validez puede ser:
 - Válido: El AS de origen y el Largo Máximo coinciden con la información del ROA
 - Inválido: La información del ROA no coincide
 - No encontrado: No hay un ROA para el prefijo dado

RPKI en la práctica

¿Cómo definir los ROA?

- Un ROA es semánticamente equivalente a un route(6) object:
 - **Asocia un prefijo a un ASN de origen**
 - Con esta información es posible hacer chequeo de un anuncio BGP
- Quienes tienen recursos IPv4, IPv6, ASN:
 - Pueden hacerlo desde el sistema de administración de recursos de LACNIC (MiLACNIC)
 - Se necesita para eso los datos de usuario y contraseña de administración de recursos
- Quienes no tienen recursos propios, dependerán del ISP
- Puede haber organizaciones con recursos IP pero no ASN
 - Deben crear los ROA permitiendo a cada ASN (upstream) anunciar los prefijos
 - La creación la realiza quien posee los recursos (diferente modelo que en el IRR en el que lo hace el que posee el ASN)

¿Qué tener en cuenta?

- Verificar cómo estamos realizando los anuncios
- Ejemplo: red 203.0.112.0/22
 - La estamos publicando sumariada?
 - La estamos publicando desagregada?
 - En bloques de qué tamaño? /23? /24?
 - Con qué sistema autónomo se originan las publicaciones?
 - Siempre es el mismos ASN?
 - Los distintos bloques se anuncian siempre con un mismo ASN?
- Importante: los ROA que creamos deben respetar esta política
- De lo contrario, estaremos invalidando nuestras publicaciones

Validadores

Software disponible

- RIPE NCC's RPKI Validator 3
 - Uno de los primeros validadores disponibles, muy utilizado, buena interfaz gráfica
 - RIPE ha dejado de mantenerlo desde Julio 2021
- Cloudflare: OctoRPKI & GoRTR
 - Soporte para uso en CDNs, separación clara entre la validación y el protocolo RTR
- NLnetLabs: Routinator 3000
 - Una versión con soporte profesional, muy eficiente en términos de RAM y CPU
- RPKI-client
 - Implementación libre para facilitar la validación de origen de los anuncios BGP. Genera configuración para OpenBGPD o BIRD, pero también otros formatos como CSV o JSON para ser consumidos por otros programas
- LACNIC y NIC.MX: Validador FORT
 - Proyecto FORT incluye el validador y el Monitoreo FORT. El Validador está desarrollado en C y es muy eficiente, muy liviano para ejecutar en una VM

Validador FORT

El validador FORT es un validador RPKI de código abierto

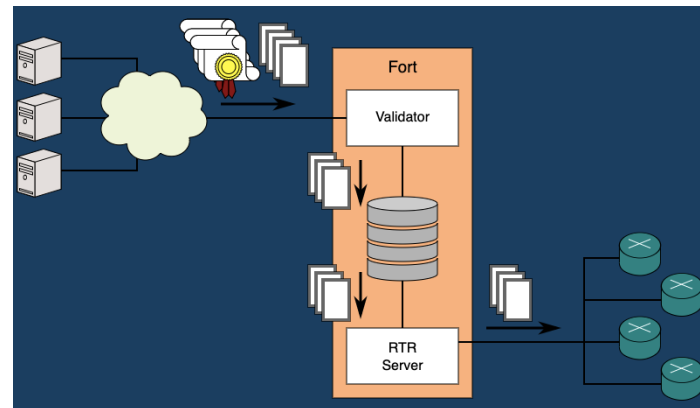
- Es parte del Proyecto FORT, iniciativa conjunta entre **LACNIC** y **NIC.MX**
- Soporte para Linux y BSD
- Desarrollado en C

Documentación general:

<https://nicmx.github.io/FORT-validator/>

Descargar el validador:

<https://github.com/NICMx/FORT-validator/releases>



Herramientas útiles

- Mi LACNIC: <https://milacnic.lacnic.net>
- LACNIC Tools: <https://tools.labs.lacnic.net/>
 - Información de los repositorios de RPKI, consultas a RDAP, WHOIS y preguntas directas a servidores de nombres
- Inforedes: <https://inforedes.labs.lacnic.net/>
 - Información de recursos de numeración, ruteo, conectividad, DNS, RPKI
- Monitoreo FORT: <https://monitor.fortproject.net/>
 - Cobertura de ROAs, validez de los updates BGP, anomalías en la información de ruteo, etc
- RIPE RIS: <https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris>
- BGP HE.NET <https://bgp.he.net>
- Cursos de Campus de LACNIC: <https://campus.lacnic.net> (BGP y RPKI)
- Documentación RPKI: <https://rpki.readthedocs.io/en/latest/>

Créditos de la Presentación

Fuente principal: Tutorial de Peering IXSY Meeting 2022

Alejandro Acosta – LACNIC
Mauricio Oviedo - SOCIUM.CR

Material creado por LACNIC