

¿Qué vemos
esta noche?



BIER

y las
arquitecturas
de distribución
de multicast



Octavio Alfageme Gorostiaga
nopacketloss.es

¿Cómo vemos la “tele”?



OTT – Over The Top



VoD - Video on Demand

IP unicast

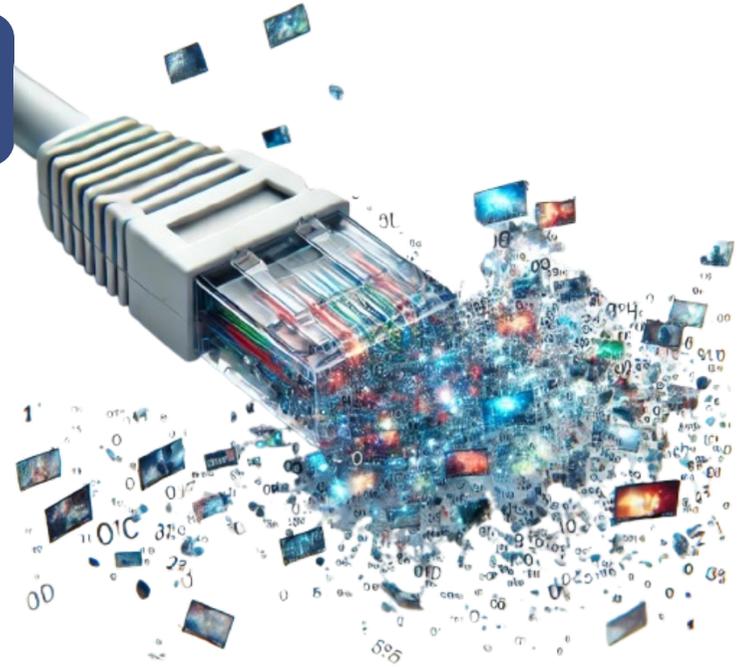
TCP

IP multicast

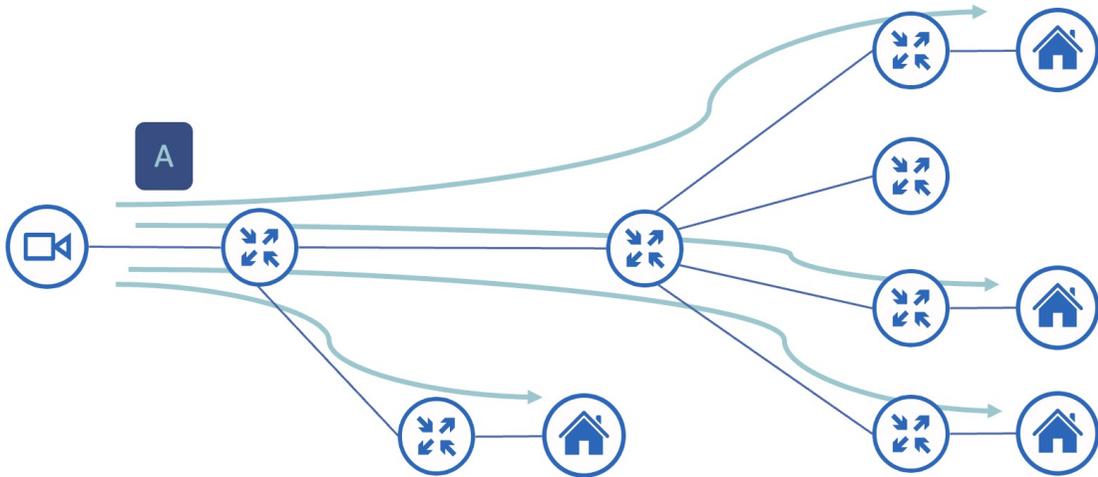
UDP

TV Live

Multicast 101 - *express*



¿Por qué multicast?



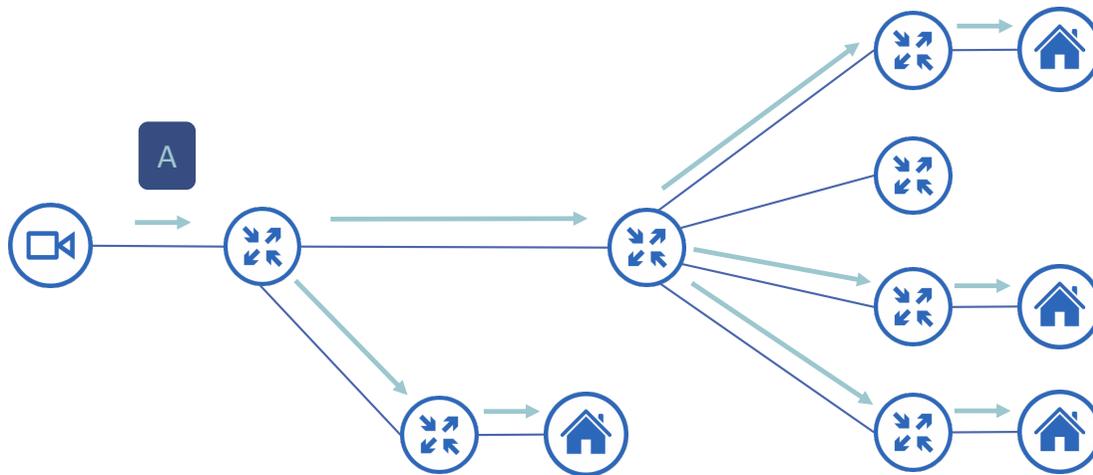
Tráfico en el punto A

UNICAST

$3,5\text{Mbps} \times 1000 \text{ clientes} = 3,5\text{Gbps}$
 $3,5\text{Mbps} \times 10000 \text{ clientes} = 35\text{Gbps}$
...

MULTICAST

¿Por qué multicast?



- La replicación multicast construye "árboles" entre emisor y receptores óptimos desde el punto de vista de consumo de ancho de banda

Tráfico en el punto A

UNICAST

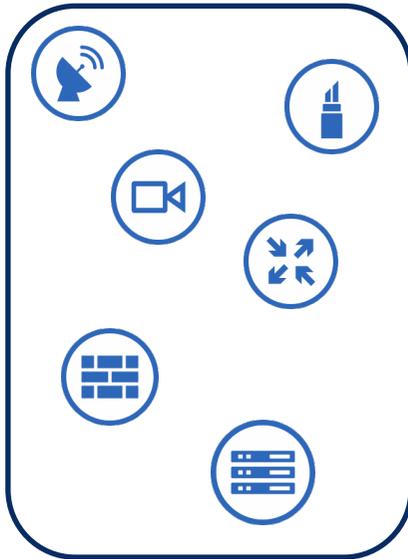
$3,5\text{Mbps} \times 1000 \text{ clientes} = 3,5\text{Gbps}$
 $3,5\text{Mbps} \times 10000 \text{ clientes} = 35\text{Gbps}$
...

MULTICAST

$3,5\text{Mbps} \times 1000 \text{ clientes} = 3,5\text{Mbps}$
 $3,5\text{Mbps} \times 10000 \text{ clientes} = 3,5\text{Mbps}$
...

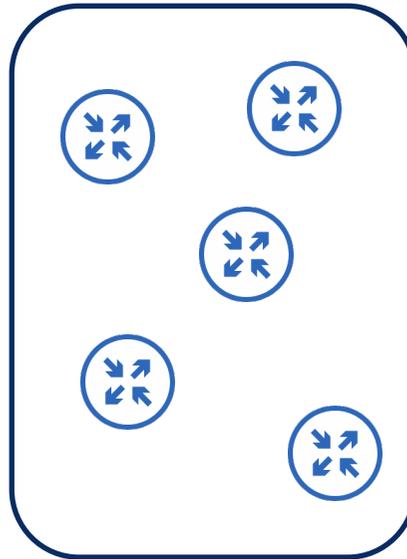
Arquitectura de red IPTV

Cabecera TV



- Captación de señales
- Generación de parrilla de canales (EPG, TSs, ...)
- Encriptación (CAS/DRM)
- IP L3 – PIM SSM/ASM
- ...

Distribución



- Conexión entre cabecera y red de acceso FTTH
- IP/MPLS/SR/BIER
- Alta disponibilidad
- Replicación multicast L3
- ...

Acceso FTTH

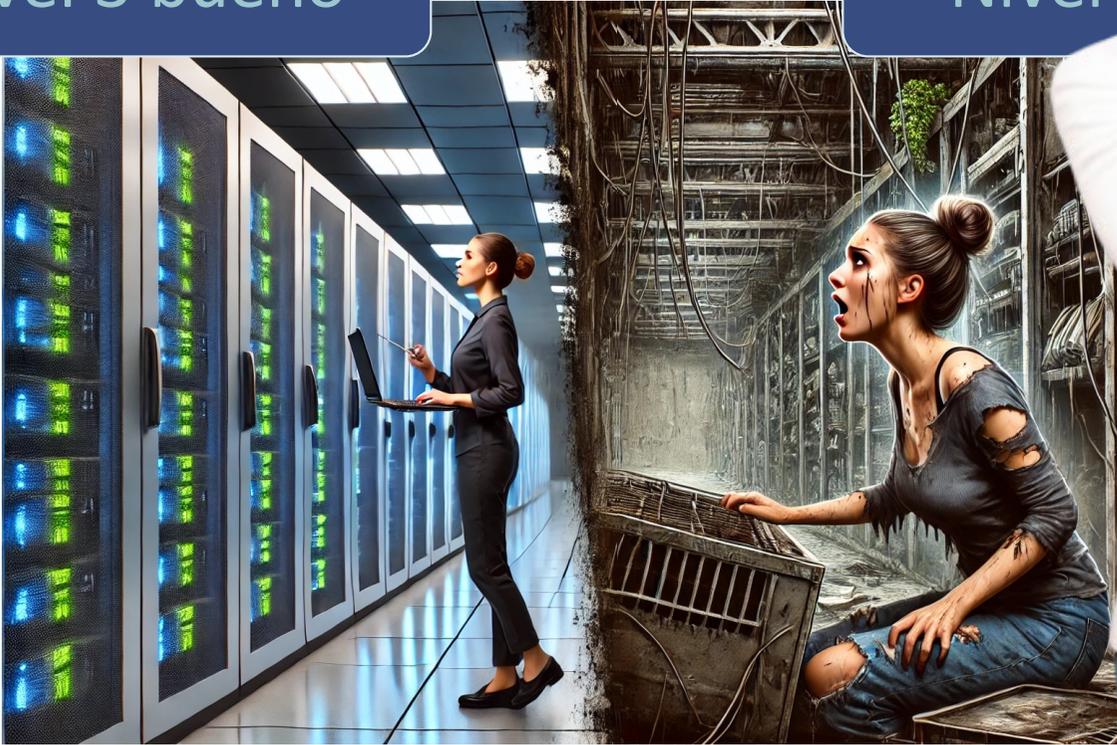


- Red acceso FTTH (OLTs-ONTs)
- STBs
- Control de la replicación multicast L2 (IGMP snooping)
- ...

El 99% del saber de IPTV multicast

Nivel 3 bueno

Nivel 2 malo



Multicast y nivel 2

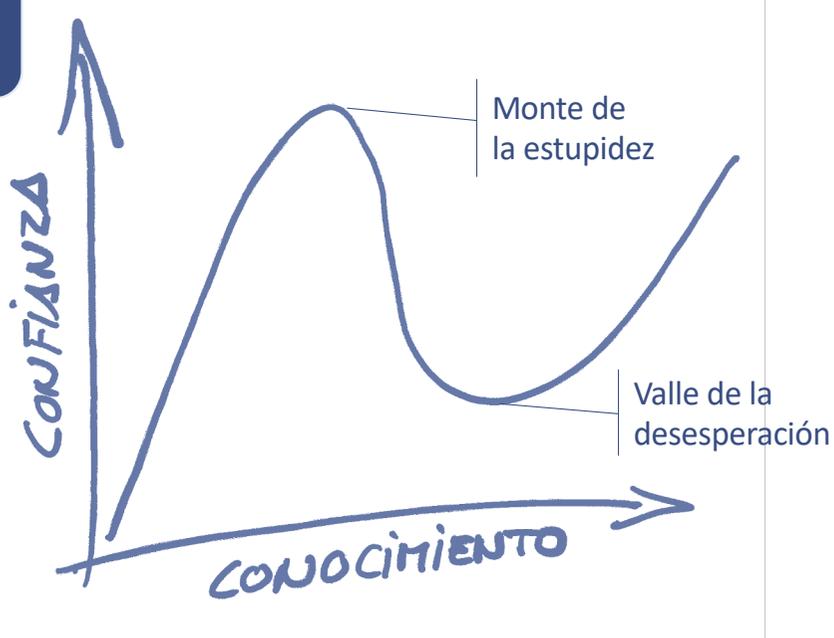
VS

Gurú

Víctima del efecto Dunning-Kruger

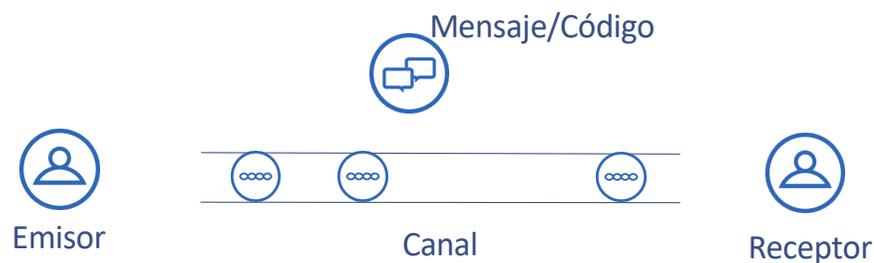
"El peor enemigo del conocimiento no es la ignorancia, es la ilusión del conocimiento"
Stephen Hawking

"Gran parte de las dificultades por las que atraviesa el mundo se debe a que los ignorantes están completamente seguros y los inteligentes, llenos de dudas."
Bertrand Russell



¿Qué requiere una comunicación unicast?

Elementos de la comunicación

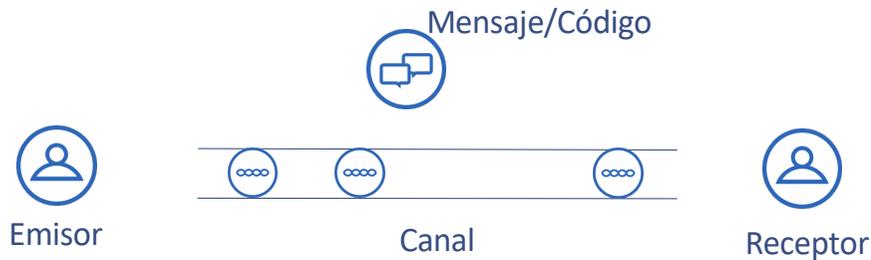


- Comunicaciones punto a punto – no requieren replicación en puntos intermedios
- Emisor y receptor estables – sus IPs los determinan

Emisor	198.51.100.3
Receptor	203.0.113.114
Canal	GRT: IGP, BGP VPN: MP-BGP, T-LDP, ...
	IP, MPLS, SR
Mensaje	PAYLOAD UNICAST

¿Qué requiere una comunicación multicast?

Elementos de la comunicación



- Comunicaciones punto a multipunto – replicación requerida
- Emisor y receptor “no estables”
 - El emisor puede dejar de generar contenido multicast o “moverse” (anycast)
 - Los receptores pueden cambiar de “grupo multicast” solicitado de modo dinámico (“zapping”)

Emisor	192.0.3.34
Receptor	IP en el rango 224/4
Canal	GRT: IGMP, MLD, PIM NG-MVPN: IGMP, MLD, PIM, MP-BGP,, ... IP, MPLS, SR, BIER
Mensaje	PAYLOAD MULTICAST

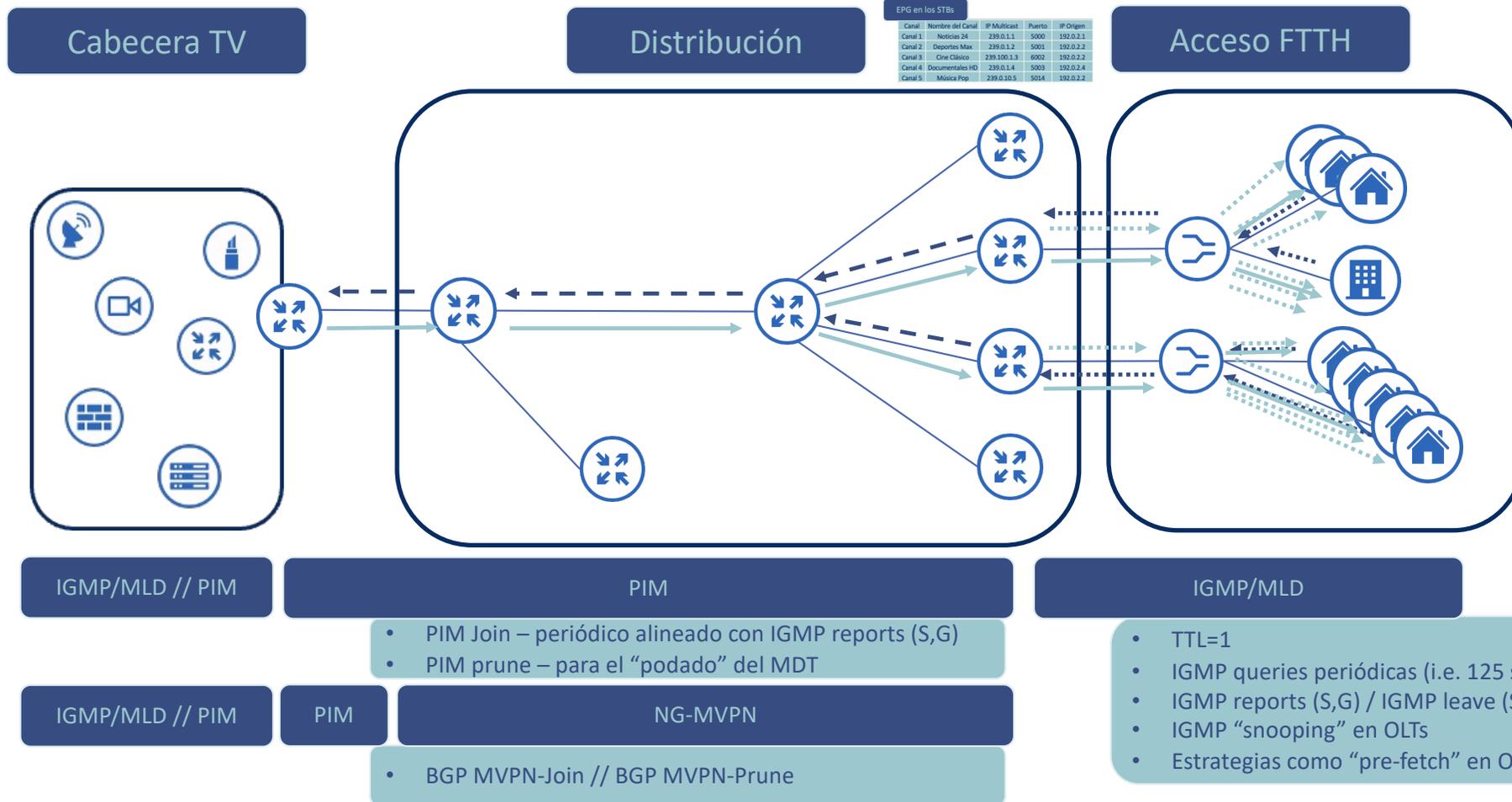
¿Cómo mantenemos control de emisores y receptores?

¿Cómo controlamos los árboles de distribución multicast o MDTs - grupos (Sx,Gx)?

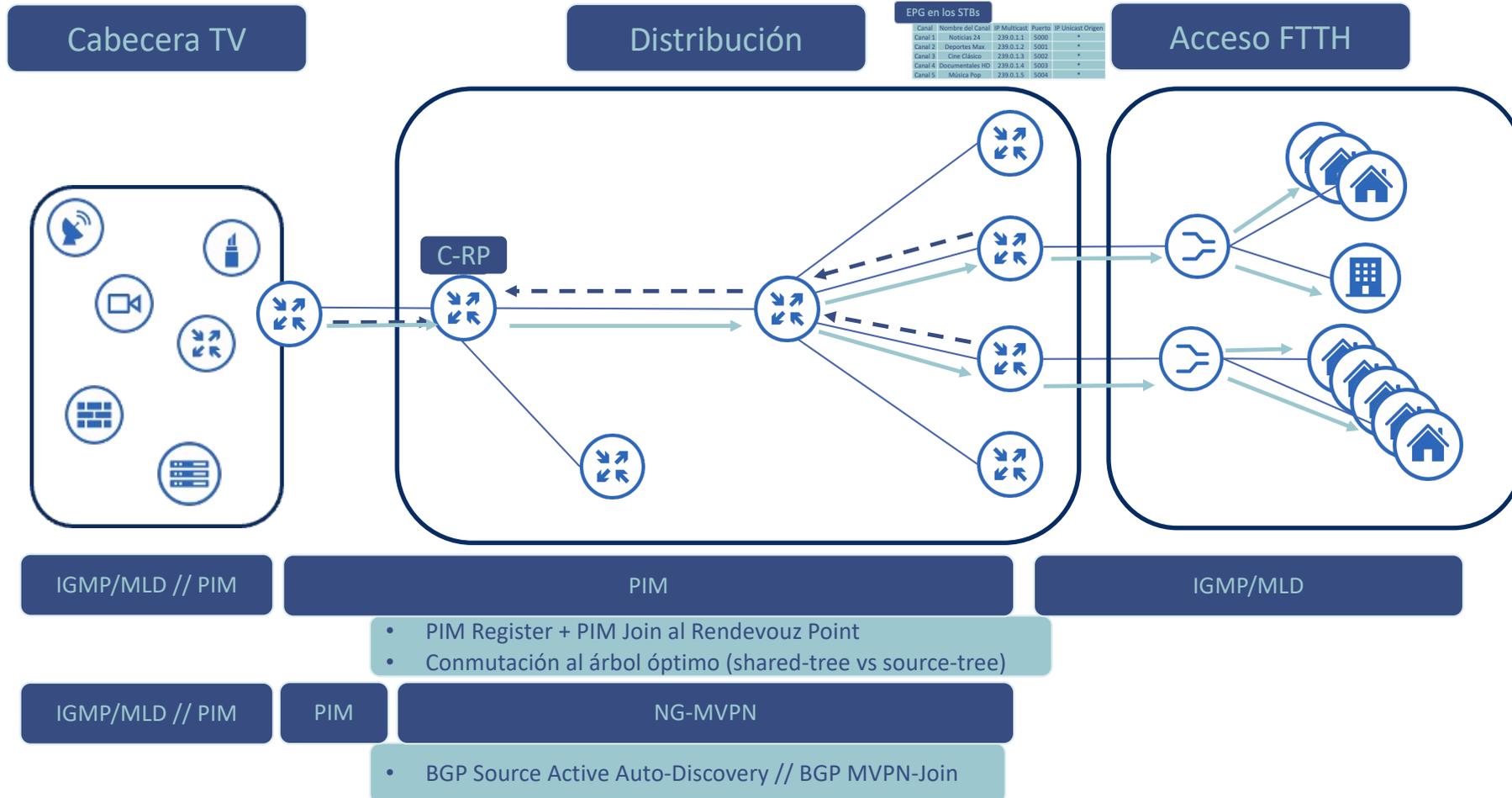
Redes de distribución multicast



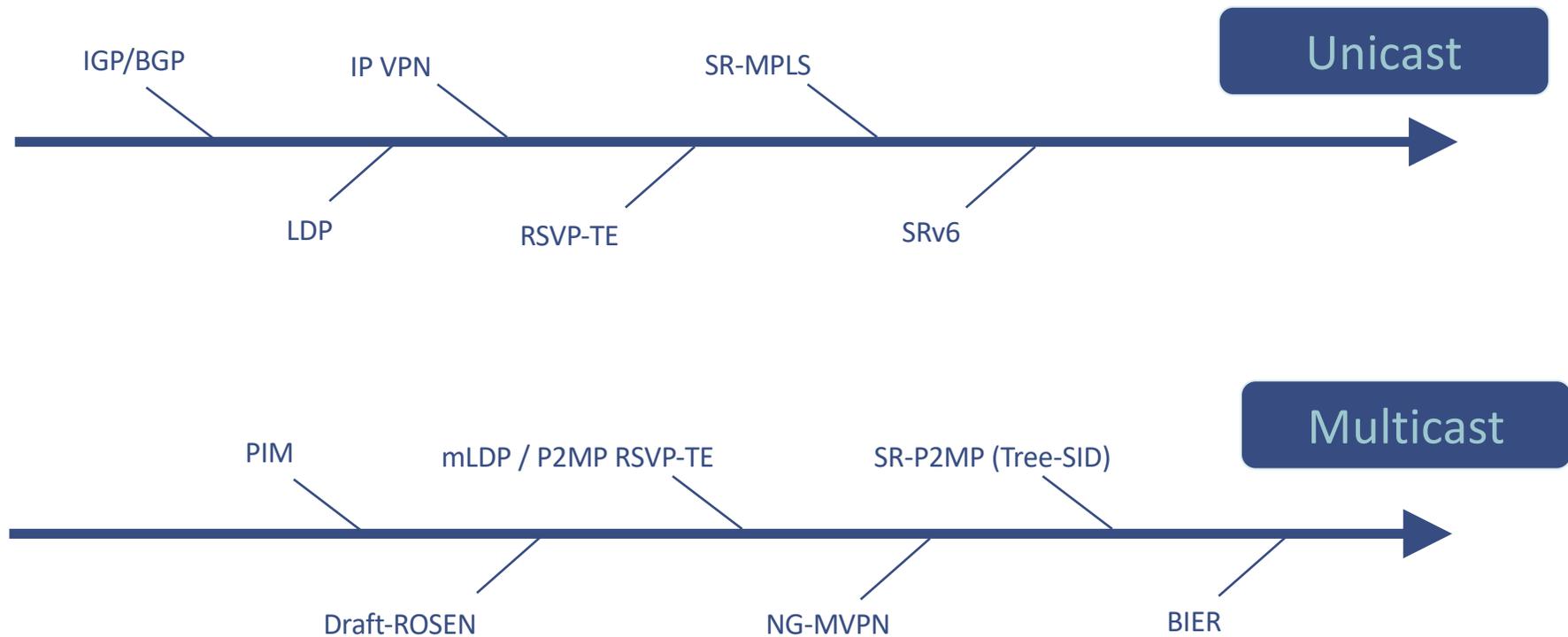
Fuentes fuera de banda - EPG



Descubrimiento de fuentes – EPG sin origen

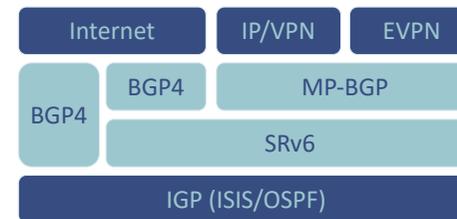
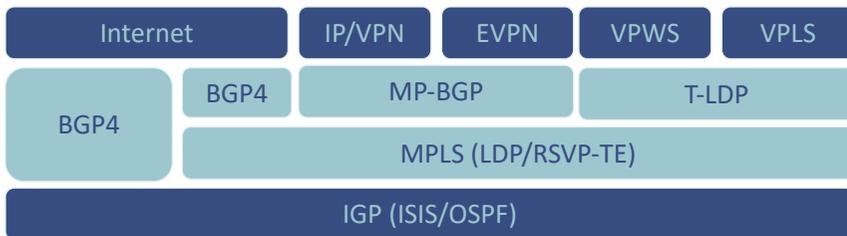


Evolución tecnológica unicast y multicast



Migración hacia redes “sin estado”

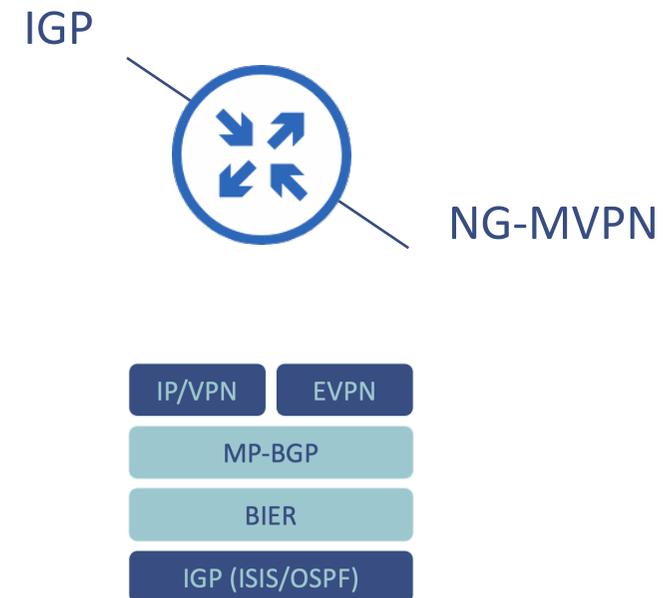
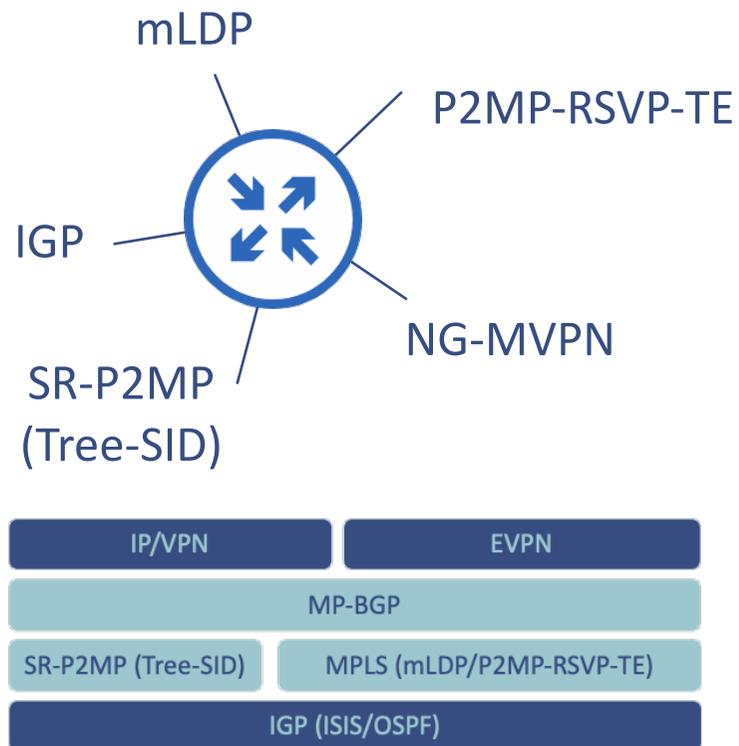
Unicast



- SRv6 elimina el “per flow state” en la red

Migración hacia redes “sin estado”

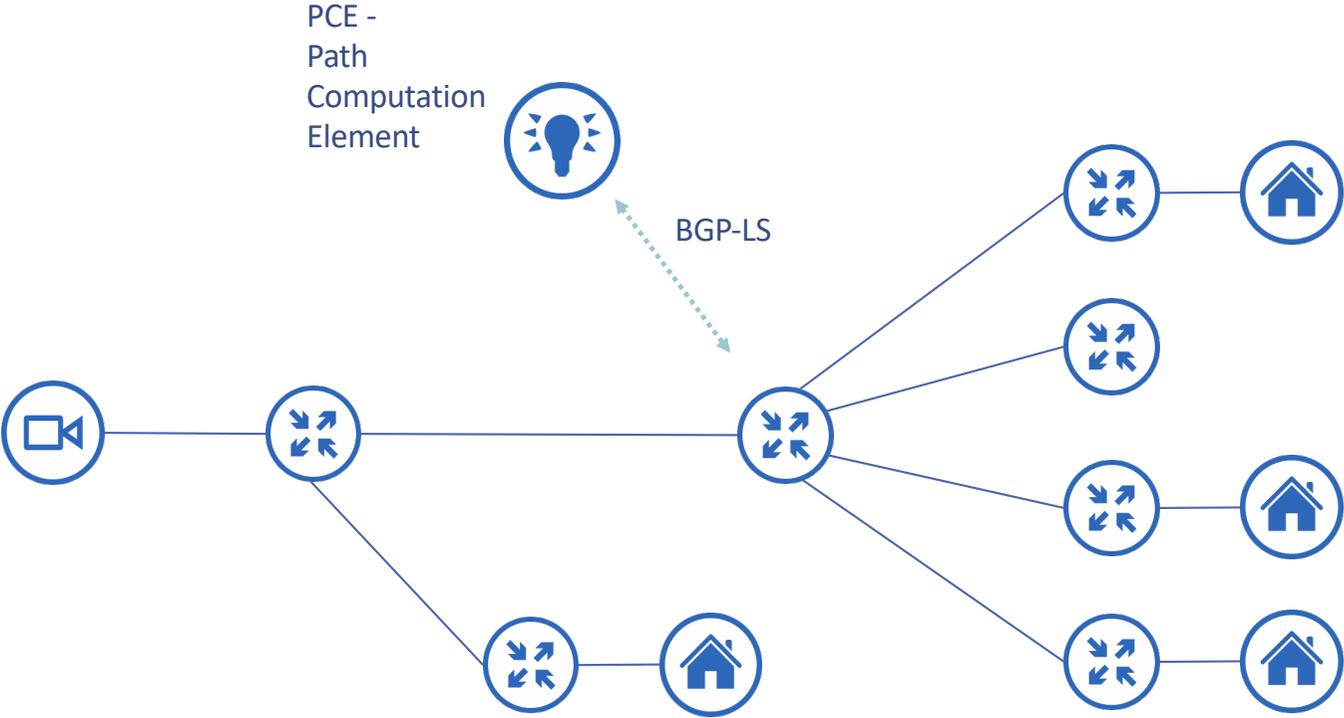
Multicast



- BIER elimina el “per flow/per tree state” de la red

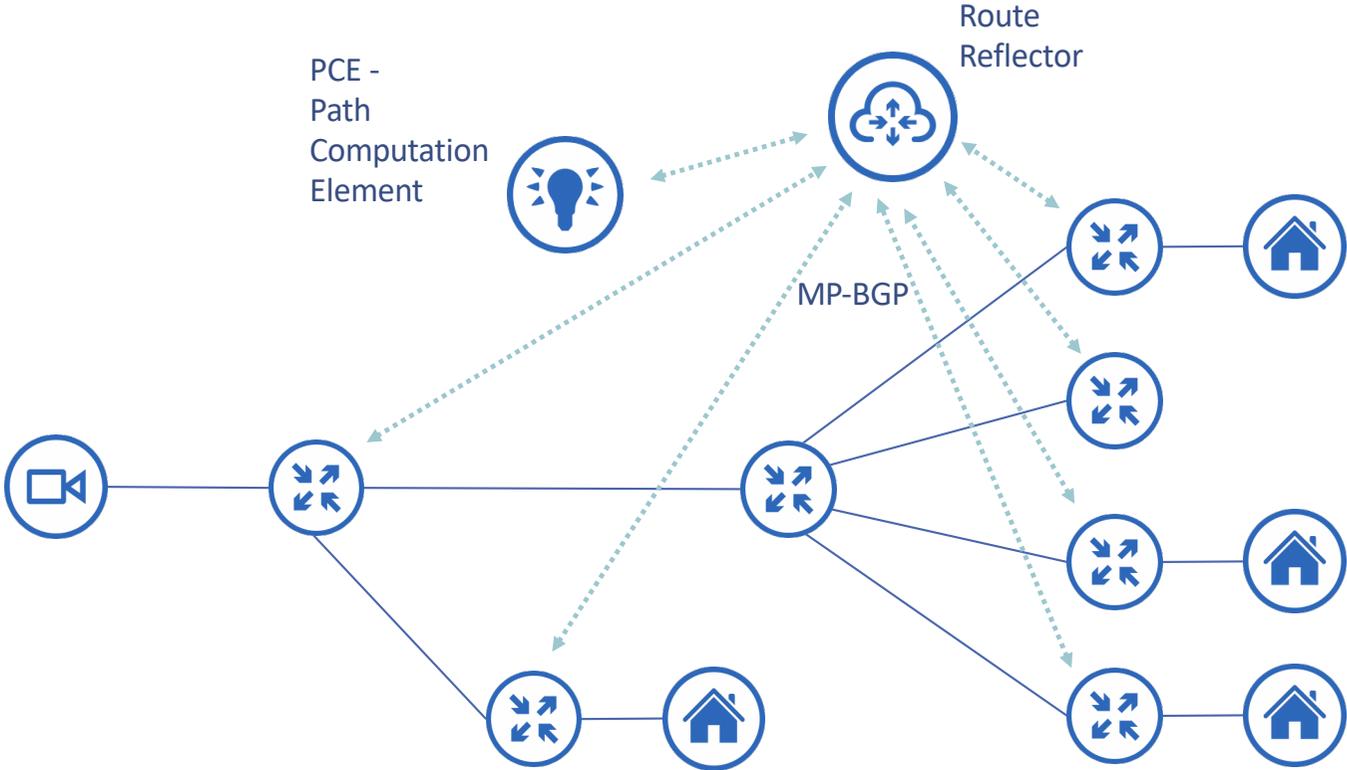
SR-P2MP (Tree-SID)

1. PCE –
Aprendizaje
de topología



SR-P2MP (Tree-SID)

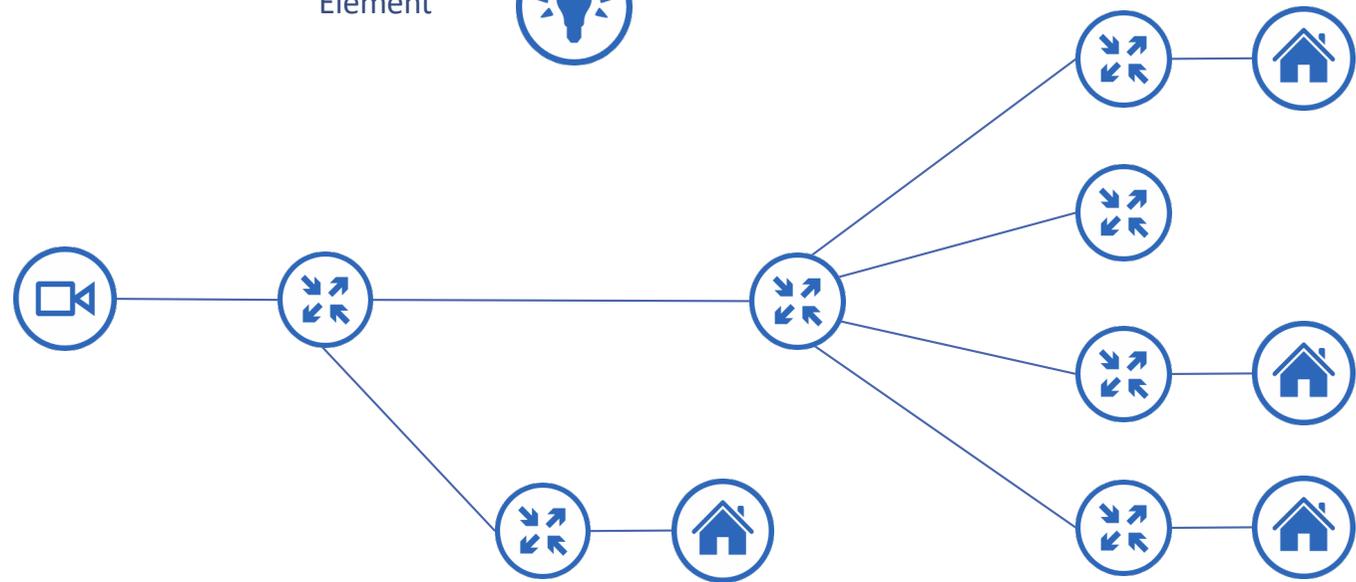
2. PCE –
Aprendizaje
de fuentes,
receptores, ...



SR-P2MP (Tree-SID)

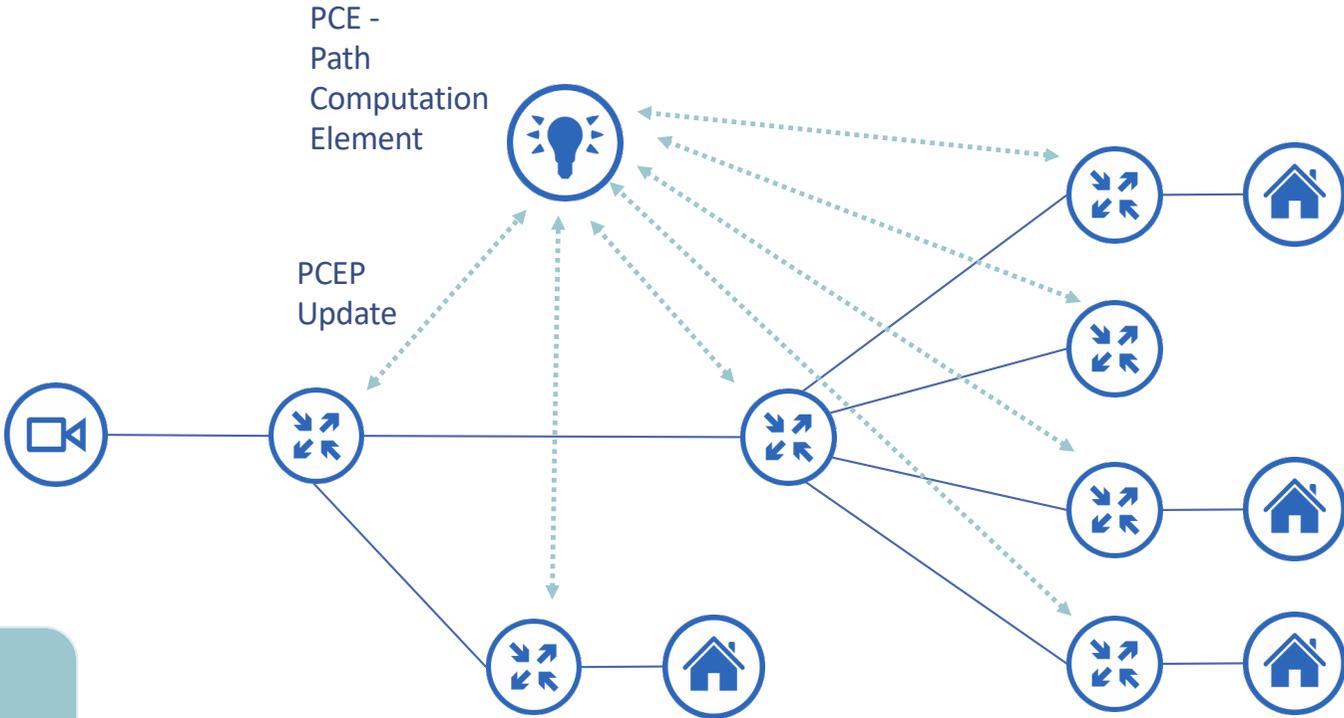
3. PCE -
Computación
del MDT

PCE -
Path
Computation
Element



SR-P2MP (Tree-SID)

4. PCE – Instanciación del Tree-SID



PCEP Update

- Tree-ID
- Tree-SID
- Replication list

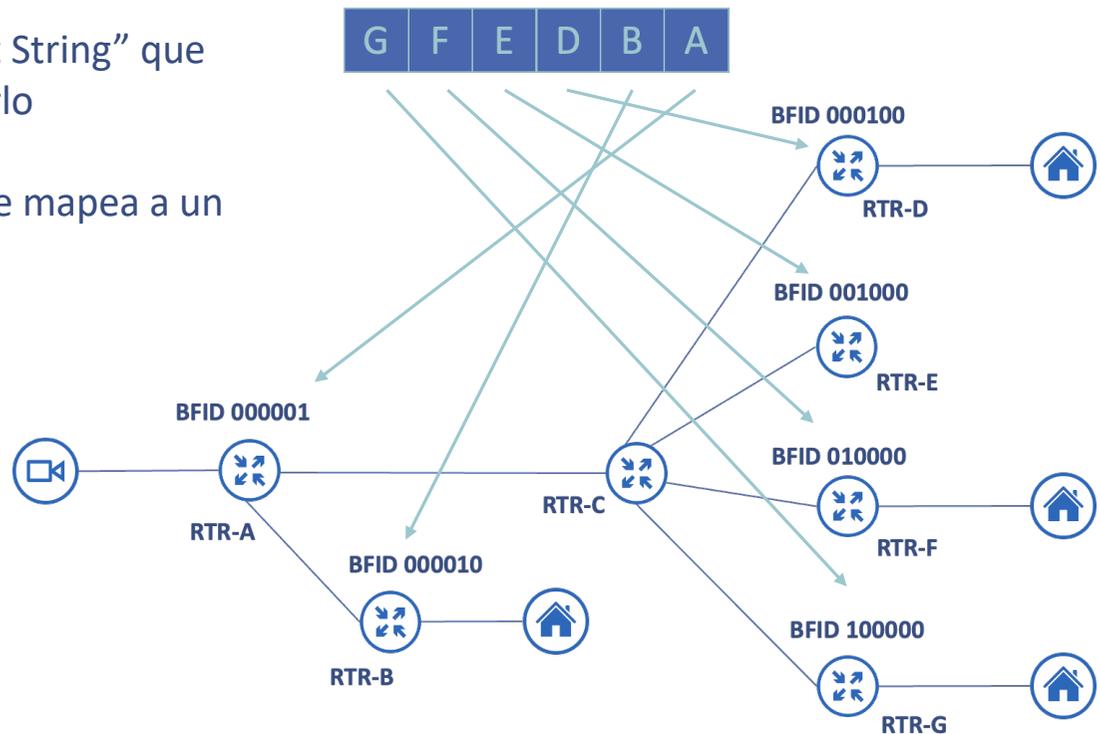
Bit Index Explicit Replication

- Protocolo “Source Based Routing” sin necesidad de mantener estado por cada árbol multicast en la red (no MDT a mantener)
- Emula a Segment Routing en el dominio multicast, incorporando además la replicación de tráfico

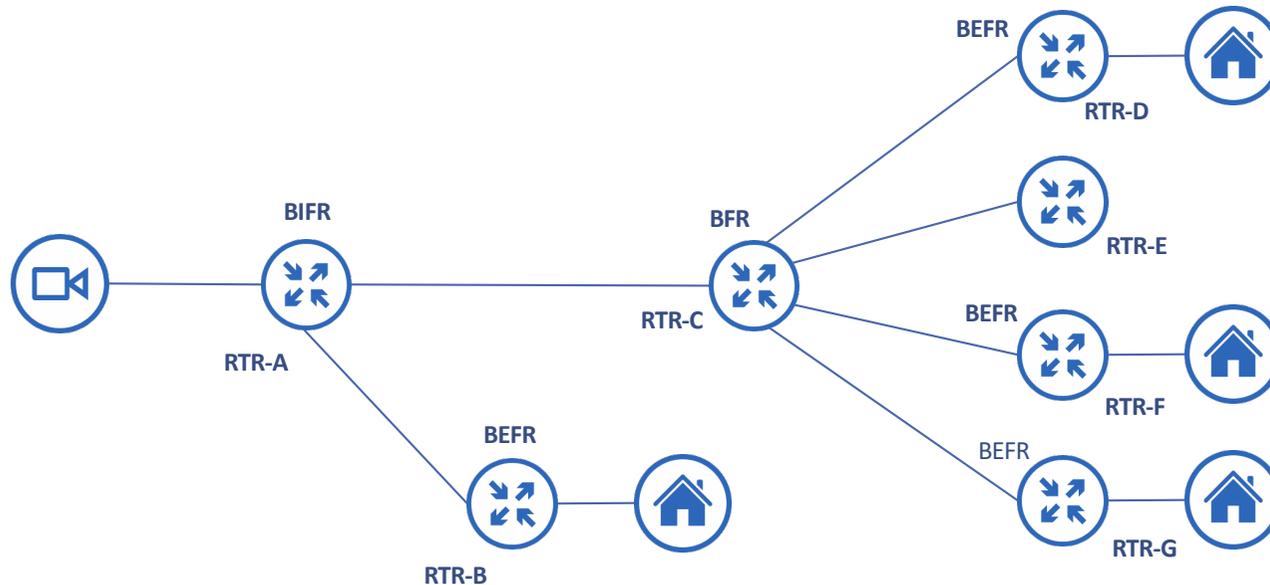


Bit Index Explicit Replication

- Cada paquete multicast BIER lleva un “Bit String” que indica los edge routers que han de recibirlo
- Cada edge router tiene un ID único que se mapea a un bit del “Bit String”
- Cada router en el camino revisa el “Bit String” y replica el paquete hacia sus vecinos en el camino a los “edge routers” destino del tráfico



Nomenclatura BIER



- BFR: BIER Forwarding Router
- BFIR: Ingress BFR
- BFER: Egress BFR

- BFR-ID: cada BFIR/BFER tiene un ID de 16 bits – mapea a un bit en el BitString

- BFR-prefix: loopback en cada BFR para la señalización y el forwarding de BIER

- BIRT: BIER Routing Table
- BIFT: BIER Forwarding Table

- BSL: BitStringLength (min 64 bits)

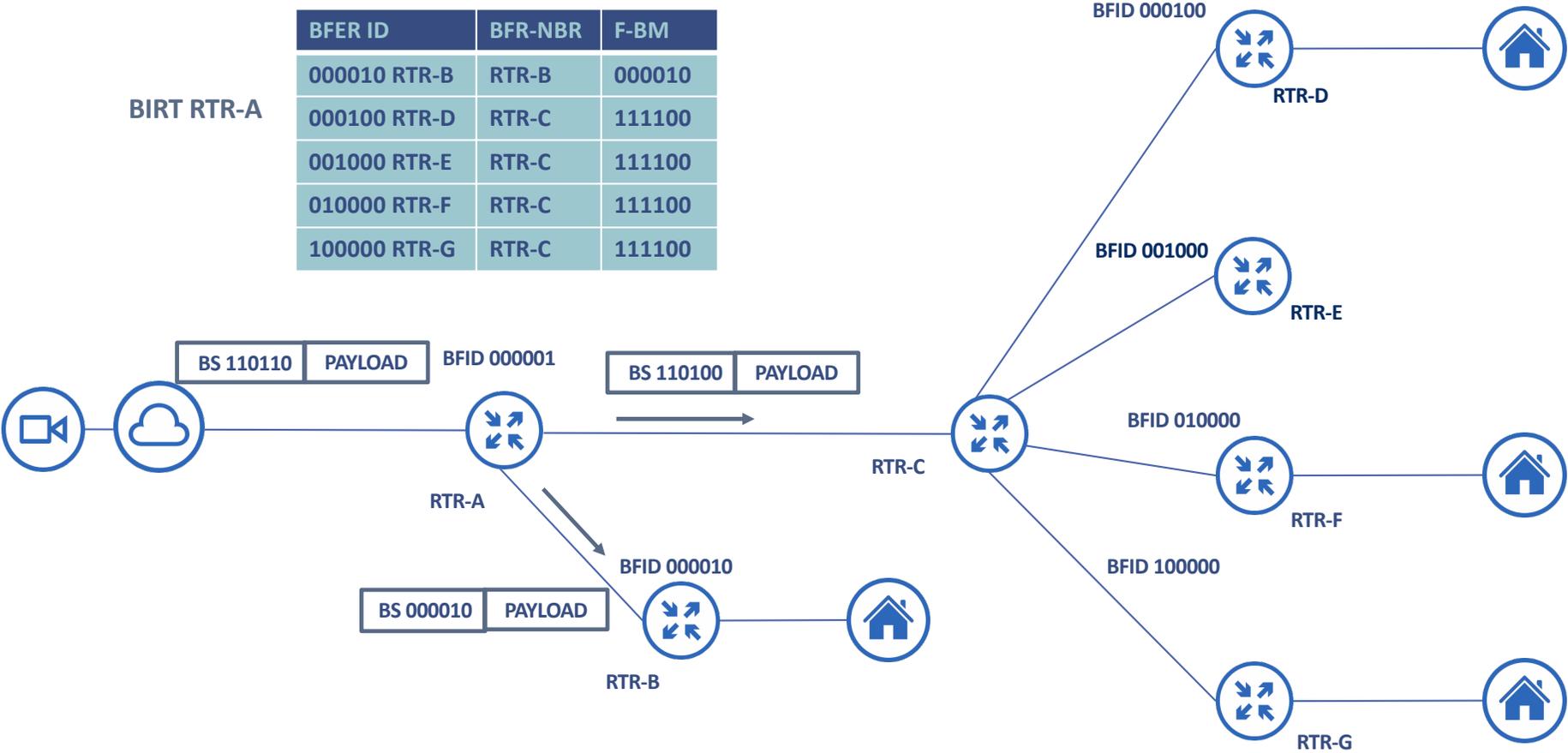
- F-BM: Bit mask asociada a cada vecino de un BFR y que representa a todos los BFERs alcanzables a través de ese vecino

Forwarding BIER

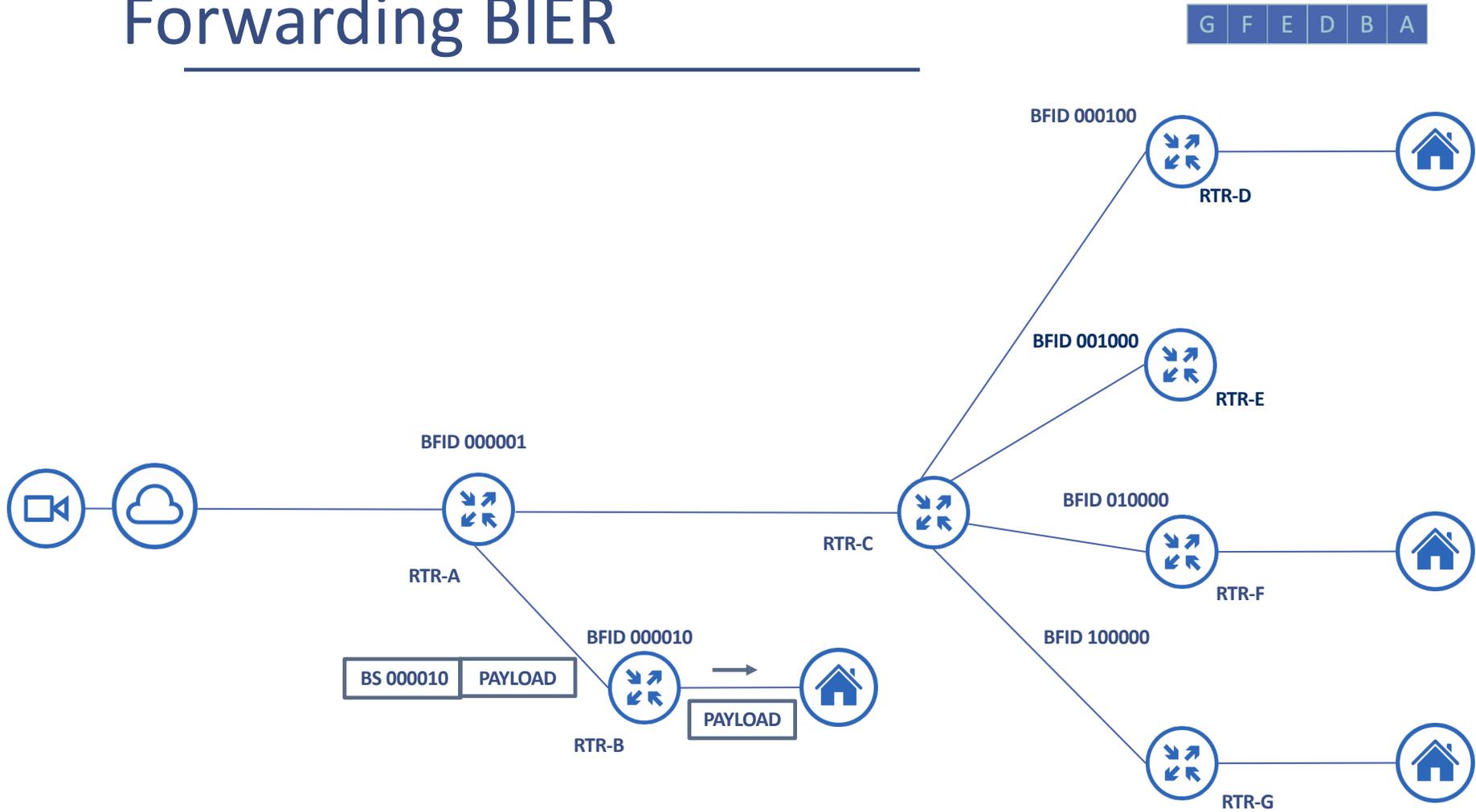
G F E D B A

BIRT RTR-A

BFER ID	BFR-NBR	F-BM
000010 RTR-B	RTR-B	000010
000100 RTR-D	RTR-C	111100
001000 RTR-E	RTR-C	111100
010000 RTR-F	RTR-C	111100
100000 RTR-G	RTR-C	111100



Forwarding BIER

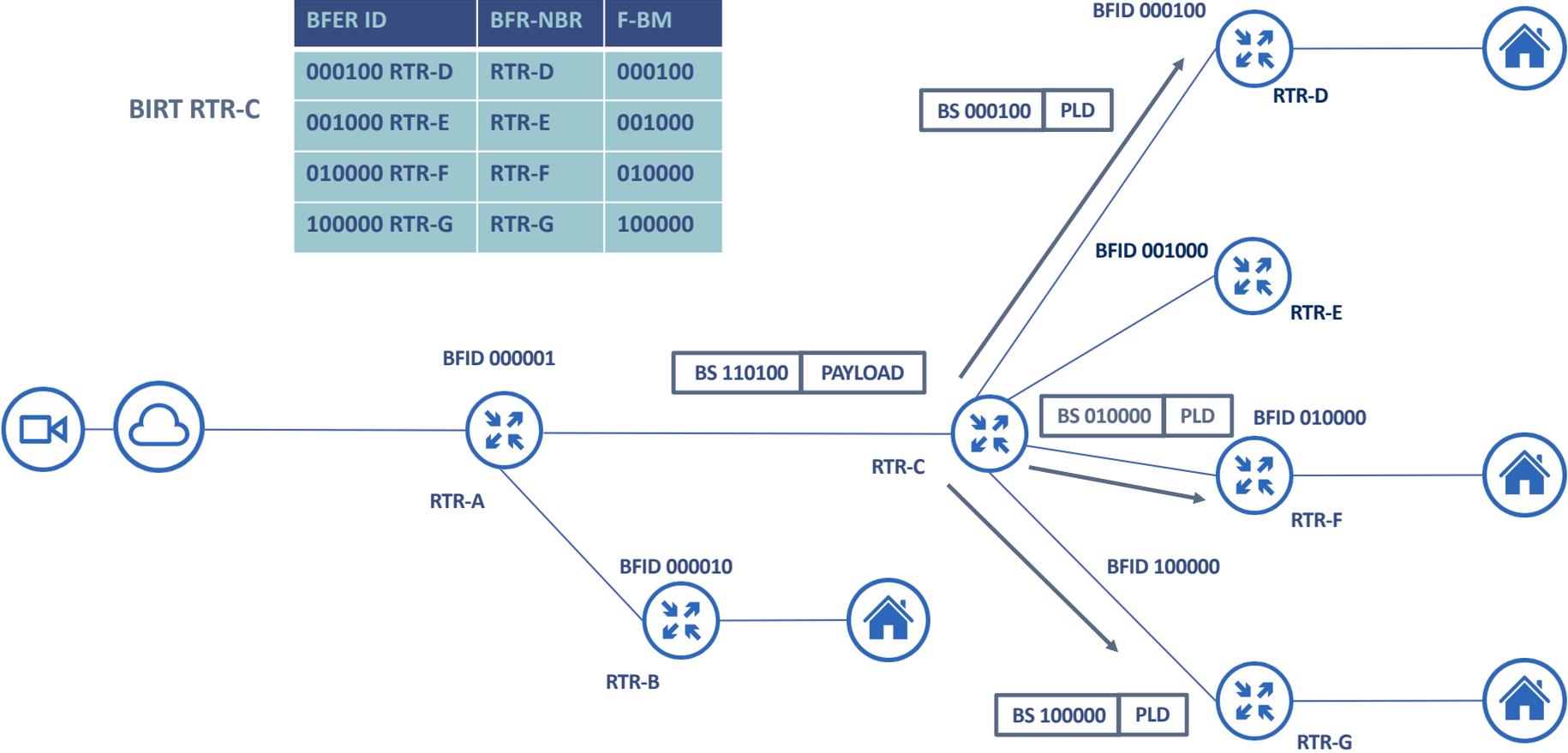


Forwarding BIER

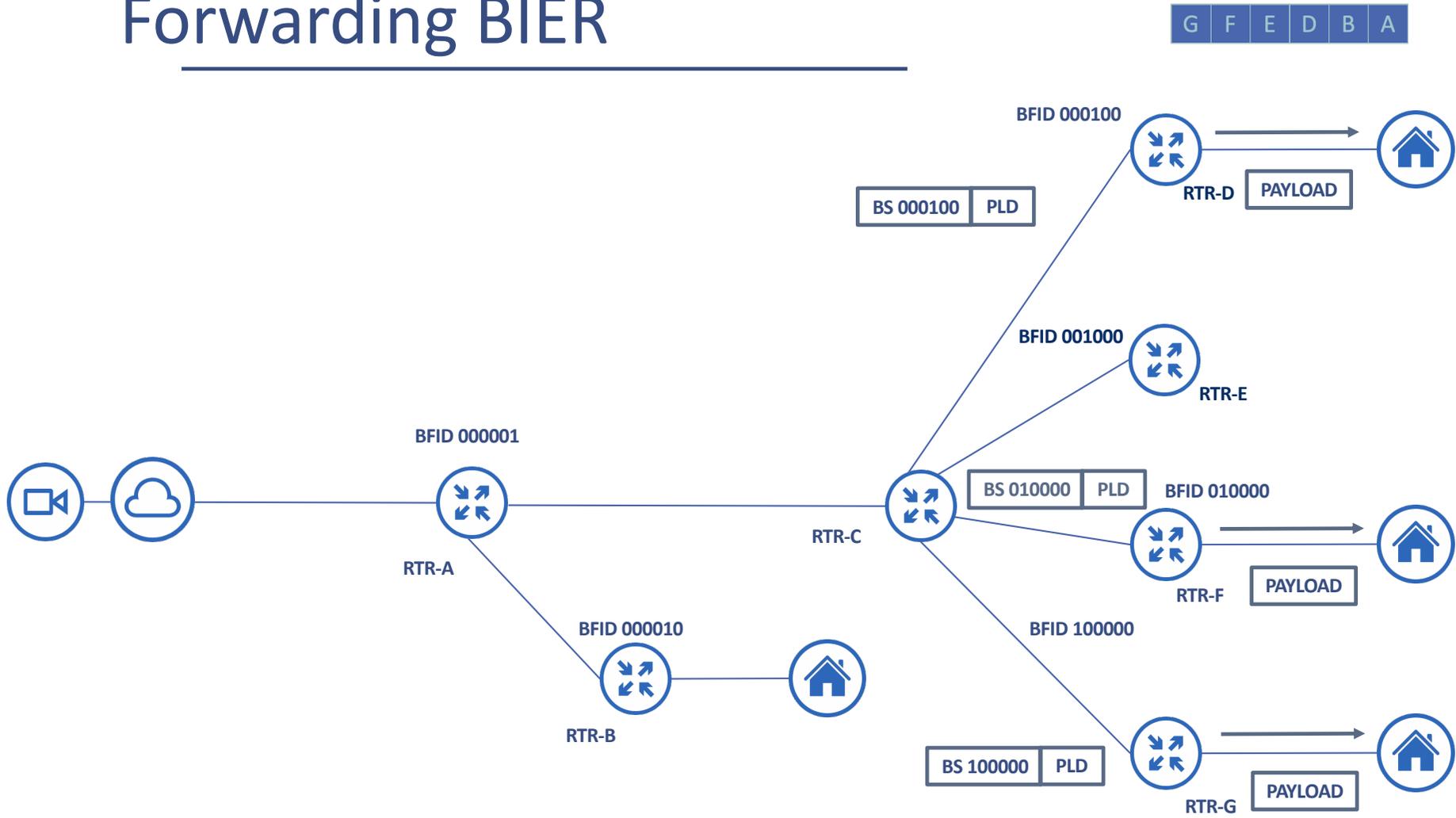
G F E D B A

BIRT RTR-C

BFER ID	BFR-NBR	F-BM
000100 RTR-D	RTR-D	000100
001000 RTR-E	RTR-E	001000
010000 RTR-F	RTR-F	010000
100000 RTR-G	RTR-G	100000

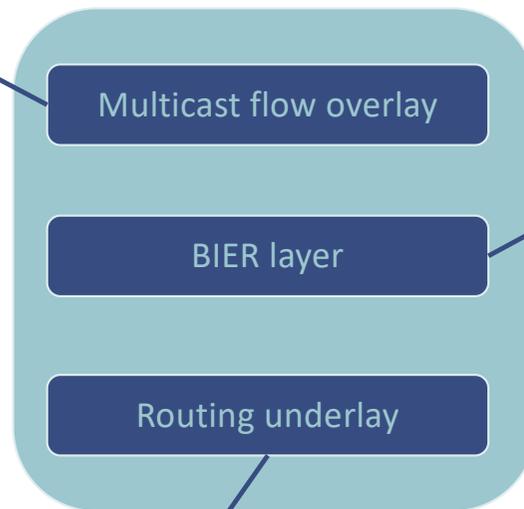


Forwarding BIER



Arquitectura BIER

- Proceso para determinar los emisores y receptores de los flujos multicast en GRT (PIM) o una VPN IP (NG-MVPN)



- Transmisión de datos en un dominio BIER
- Replicación de paquetes BIER
- Encapsulación en el BFIR y desencapsulación en el BFER

- Establecimiento de adyacencias entre BFRs
- Generación del BIFT (Bit Index Forwarding Table)
- ISIS/OSPF/BGP

El uso del “bitstring” para el forwarding y la replicación paquete a paquete hace desaparecer la necesidad del mantenimiento del MDT para cada flujo multicast

BIER en el IETF

- EVPN - RFC8556 – Agosto 2024
- IP VPN – NG-MVPN – Abril 2019

Multicast flow overlay

- ISIS– RFC8401 – Junio 2018
- OSPFv2 – RFC8444 – Noviembre 2018
- OSPFv3 - draft-ietf-bier-te-ospfv3
- BGP4 - draft-ietf-bier-idr-extensions

Routing underlay

- RFC8279 – Noviembre 2017

BIER layer

- MPLS / Non-MPLS – RFC8296 – Enero 2018
- IPv6 – draft-ietf-bier-bierin6

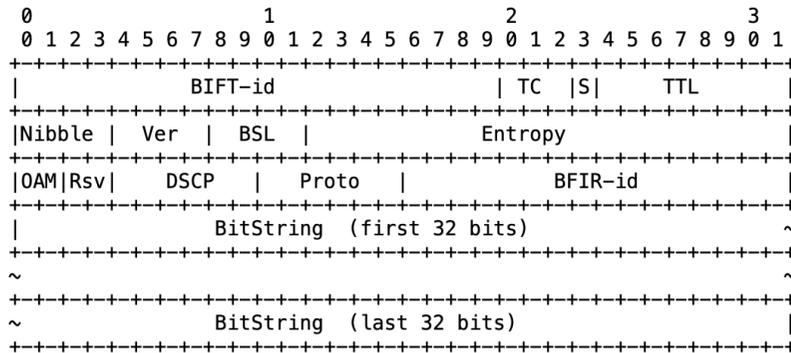
Encapsulación de BIER

- RFC9262 – Octubre 2022

BIER-TE

Encapsulación BIER

RFC8296

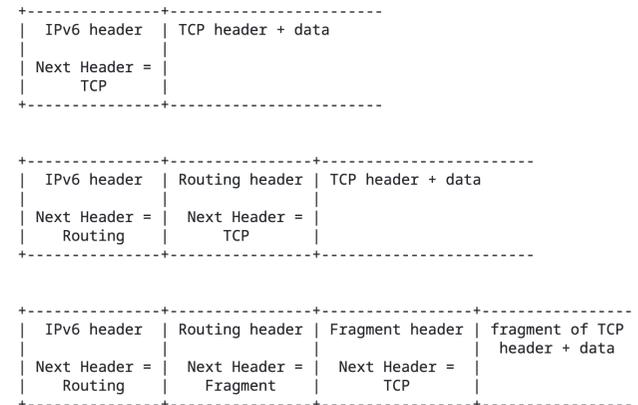


MPLS

Non-MPLS

- Ethernet (ethertype 0xAB37)
- UDP sobre IPv4 o IPv6

draft-ietf-bier-bierin6



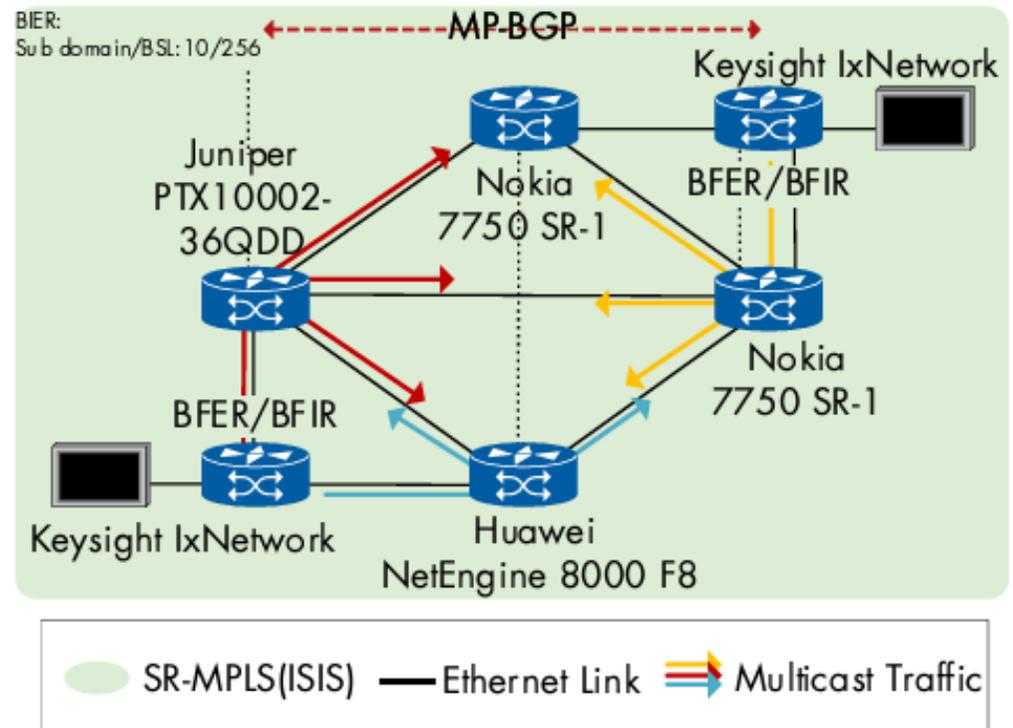
- Encapsulación nativa BIER sobre IPv6 por medio de una “cabecera de extensión IPv6” específica para BIER

- BIERv6 como implementación sobre IPv6 hoy fuera del ámbito del IETF

BIER en “campo”

- Primer interop básico en 2024 en el MPLS SDN World Congress (NG_MVPN sobre BIER sobre red SR-MPLS)

<https://eantc.de/wp-content/uploads/2023/12/EAN-TC-MPLSSDNInterop2024-TestReport-v1.3.pdf>



¿Por qué no existen implementaciones de un protocolo con RFCs de 2017-2018?

Vendors de routers y BIER

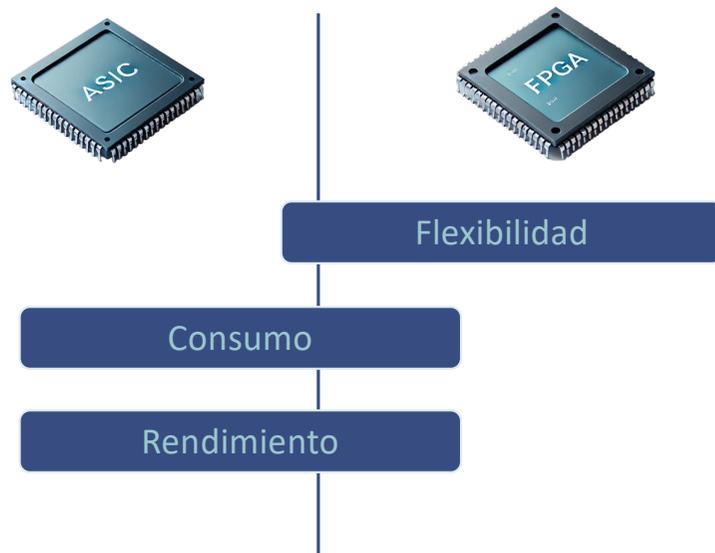
- Existen referencias públicas de documentación de algunos vendors en versiones disponibles en los dos últimos años y con funcionalidades crecientes

<https://www.juniper.net/documentation/us/en/software/junos/release-notes/23.4/junos-evo-release-notes-23.4r2/junos-evo-release-notes-23.4r2.pdf>

https://documentation.nokia.com/cgi-bin/dbaccessfilename.cgi/3HE17155AAABTQZZA01_V1_7450%20ESS%207750%20SR%207950%20XRS%20and%20VSR%20Multicast%20Routing%20Protocols%20Guide%2021.5.R1.pdf

<https://support.huawei.com/enterprise/en/doc/EDOC1100278979/d59dca37/bier-configuration>

Forwarding-engines: FPGAs vs ASICs



- Los routers conmutan paquetes IPv4 e IPv6 y etiquetas MPLS en sus FEs/NPUs
- ¿Qué pasa con las últimas tecnologías unicast?
 - SR-MPLS → forwarding y encapsulación MPLS
 - SRv6 → forwarding y encapsulación IPv6
- ¿Cómo se conmuta el “bit-string” de BIER?
¿Cómo es su encapsulación?

BIER supone una nueva encapsulación y forwarding no disponibles en los ASICs en campo en 2017

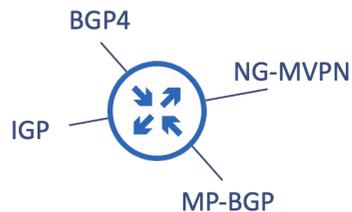
(*) ASICs vs FPGAs y ASICs programables

Uniendo las piezas



BIER y las redes “sin estado”

- SRv6 y BIER nos llevan a redes “sin estado” unicast y multicast



- BIER, como nueva encapsulación y forwarding, impone nuevos requisitos en el hardware



- SR-P2MP (Tree-SID) mantiene estado en todos los routers de la red

- El paso del tiempo favorece la actualización de los equipos en planta y con ello los primeros despliegues de BIER

(*) NG-MVPN es parte de MP-BGP

One more thing...

En IPTV multicast...



Nivel 3 bueno

Nivel 2 malo





ii Muchas gracias !!

www.nopacketloss.es

oalfageme@nopacketloss.es

@oalfageme



Referencias

- Imágenes de hombre con visera roja en transparencias 6 y 38 y mano en la 39 de FreePik (www.freepik.com)
- Resto de imágenes generadas con DALL-E
- NANOG90 Zhang BIER - The Best Multicast Solution for Segment Routing https://storage.googleapis.com/site-media-prod/meetings/NANOG90/4977/20240212_Zhang_Bier_-_The_v1.pdf
- NANOG81 Zhang BIER – Multicast in Segment Routing Networks https://storage.googleapis.com/site-media-prod/meetings/NANOG81/2323/20210202_Zhang_Multicast_In_Segment_v1.pdf
- IETF BIER Working Group <https://datatracker.ietf.org/group/bier/documents/>

Referencias

- APNIC Conference – Bit Index Explicit Replication https://conference.apnic.net/data/41/ripe-bierv4_1446515530_1456002847.pdf
- BIER Juniper Networks <https://www.juniper.net/documentation/us/en/software/junos/multicast/topics/concept/bier-overview.html>