

# BGP con routers de bajo coste

Cómo levantar sesiones BGP (Full Routing Table) con equipos pequeños

**Amedeo Beck Peccoz**

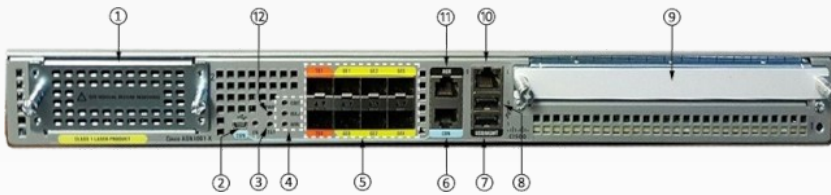
*ESNOG 27 – Madrid*

*16 de noviembre de 2021*

# Objetivo

## ASR1001X-20G-K9

ASR1001-X with 20G license – 6 x 1G and 2x10G



Precio GPL: \$ 53.900

## ASR-920-24SZ-M

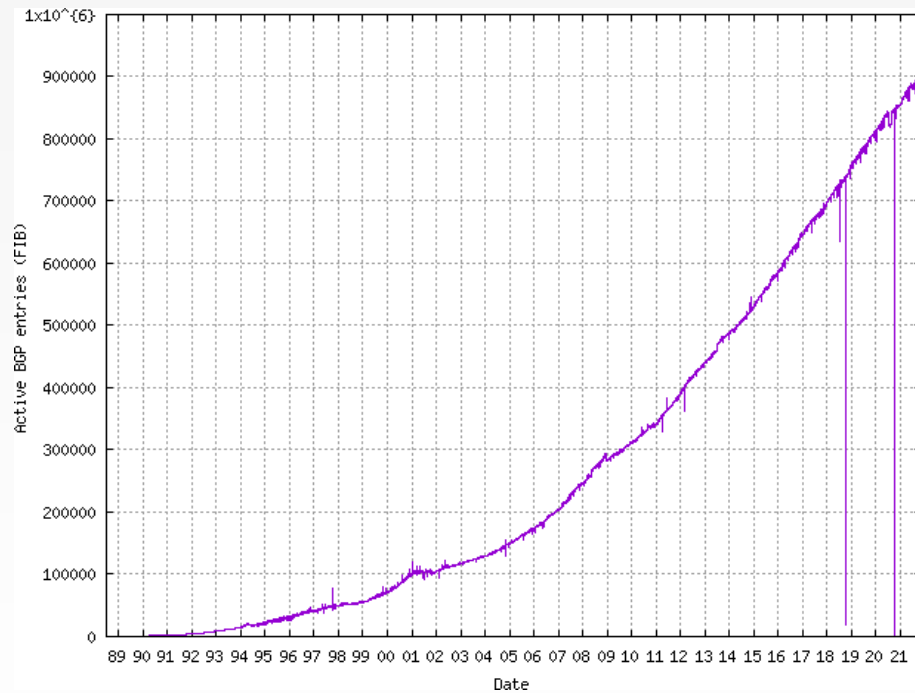
Cisco ASR920 Series – 24 x 1G and 4 x 10G license enabled



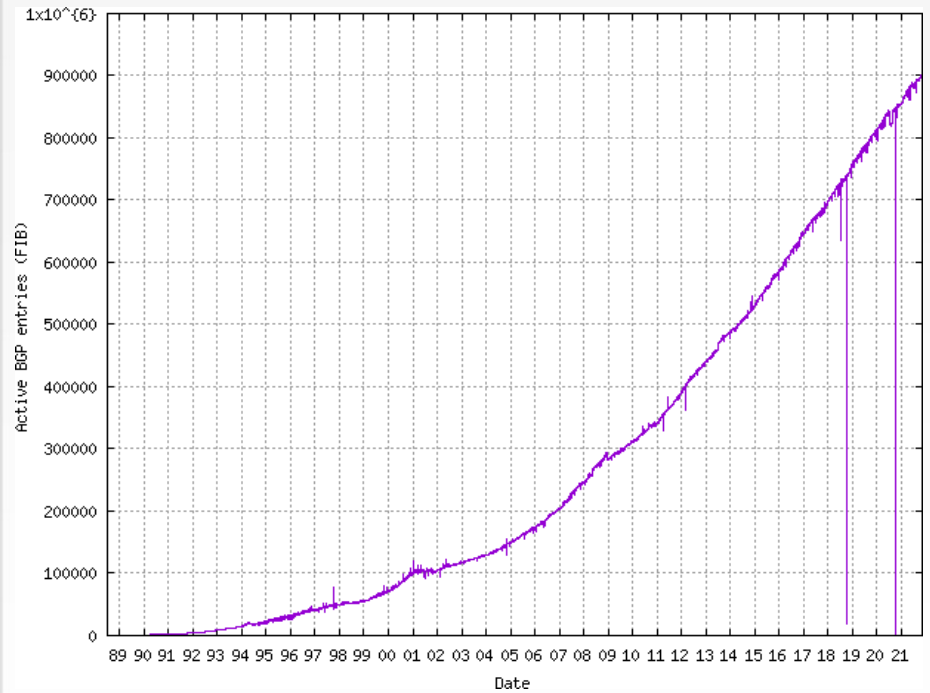
Precio GPL: \$ 11.000



# El tamaño de la FRT



Rutas IPv4: 900k+  
~ +1k / semana



Rutas IPv6: 150k+  
~ +1k / día

# Hardware para la FRT

Routers “bien equipados”

- Pensados para eBGP
  - Mucha memoria
  - Millones de rutas
  - Última generación
    - Licencias

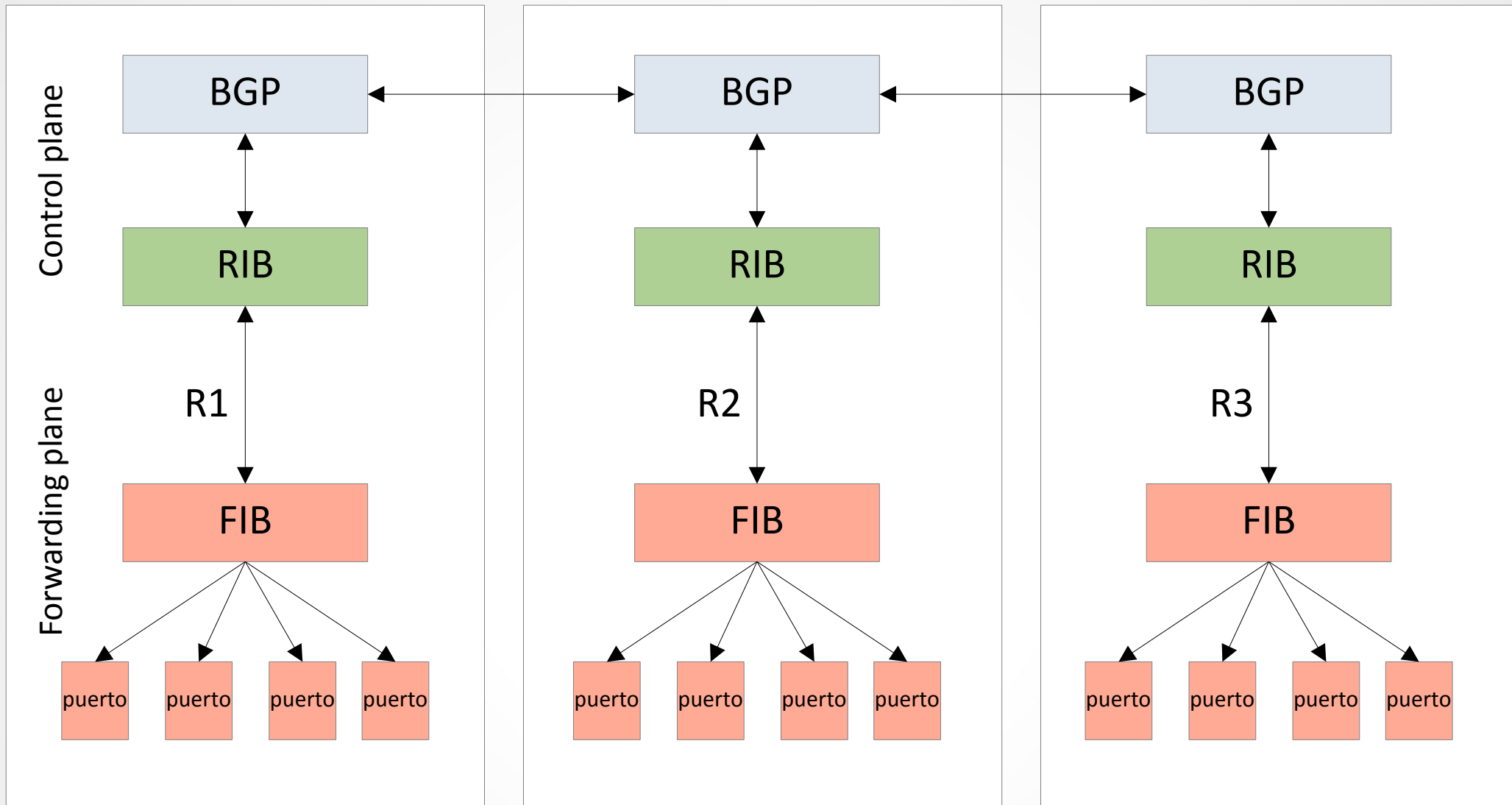


Cisco: ASR 1000 - 9900  
Juniper: MX (grandes) - PTX



*Imagen: dailystar.co.uk*

# BGP, RIB, FIB y las rutas 1/3



RIB: Routing Information Base  
FIB: Forwarding Information Base

# BGP, RIB, FIB y las rutas 2/3

## BGP

```
RouterBGP#show ip bgp
```

```
...  
      Network      Next Hop          Metric LocPrf Weight Path  
*    33.0.0.0      192.1.12.2                0 2000 3000 i  
*>                   192.1.13.3                0           3000 i  
...
```

## RIB

```
RouterBGP#show ip route
```

```
...  
B      33.0.0.0/8 [20/0] via 192.1.13.3, 10:00:16  
...
```

## FIB

```
RouterBGP#show ip cef
```

```
Prefix                          Next Hop                      Interface  
...  
33.0.0.0/24                      attached                       Ethernet0/0  
...
```

# BGP, RIB, FIB y las rutas 3/3

## Tabla BGP

- Almacenada en RAM
- Sirve para intercambiar las rutas con otros BGPs
- Múltiples entradas para cada destino
- “Muchos” bytes para cada destino (100s)
- Regla no escrita: 1G Ram ~ 1M rutas

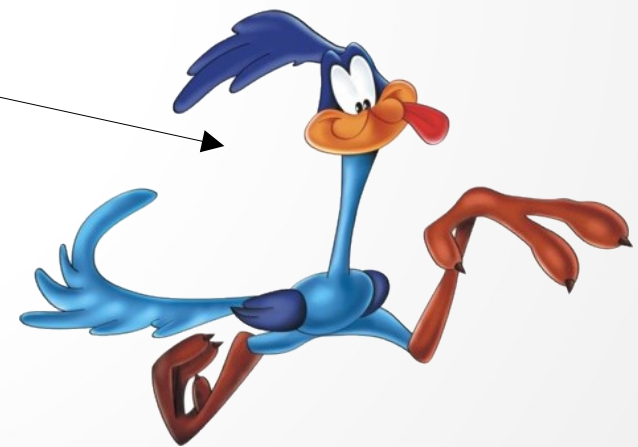
## Tabla RIB

- Almacenada en RAM
- Indica la mejor ruta hacia los destinos
- Una sola entradas para cada destino (excepto multipath)
- “Pocos” bytes para cada destino (10s)

## Tabla FIB

- Almacenada en RAM/TCAM
- Empleada para mover los paquetes por los ASICs (CEF/PFE)
- Una sola entrada para cada destino
- “Muy pocos” bytes para cada destino

ASIC



*Imagen: Warner Brothers Animation*

# La memoria para la FIB

## Tabla FIB

**La TCAM es la parte crítica**



- Cuesta un riñón
- Hay muy poca
- Todos quieren más pero no se puede aumentar

- Almacenada en RAM/TCAM
- Empleada para mover los paquetes por los ASICs (CEF/PFE)
- Una sola entrada para cada destino
- “Muy pocos” bytes para cada destino



# Hardware para iBGP

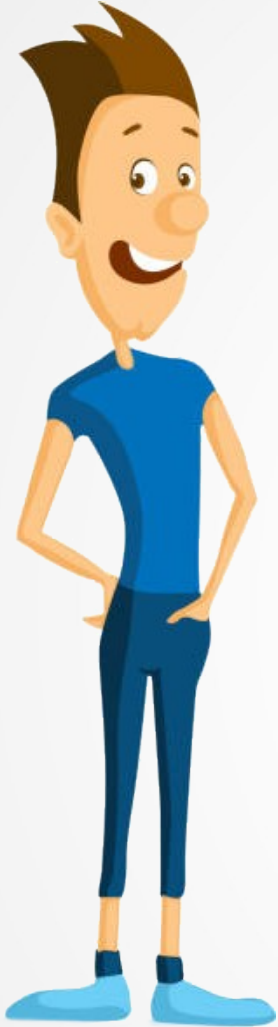


Imagen: istockphoto.com

## Routers “más flacos”

- Pensados para iBGP y no para eBGP
- En ocasiones disponible en algunos switches
- Limitación importante en el número de rutas



Cisco: ASR 900 - 1000  
Juniper: MX (pequeños)

**Es imposible que aguanten la FRT**

# ¿Quién dijo imposible?



Imagen: rtl.de

It always seems impossible  
until it's done

*Nelson Mandela*

Traducción “muy libre”<sup>®</sup>

¿Que no se puede hacer?  
→ ¡Sujetame el cubata y apartate!

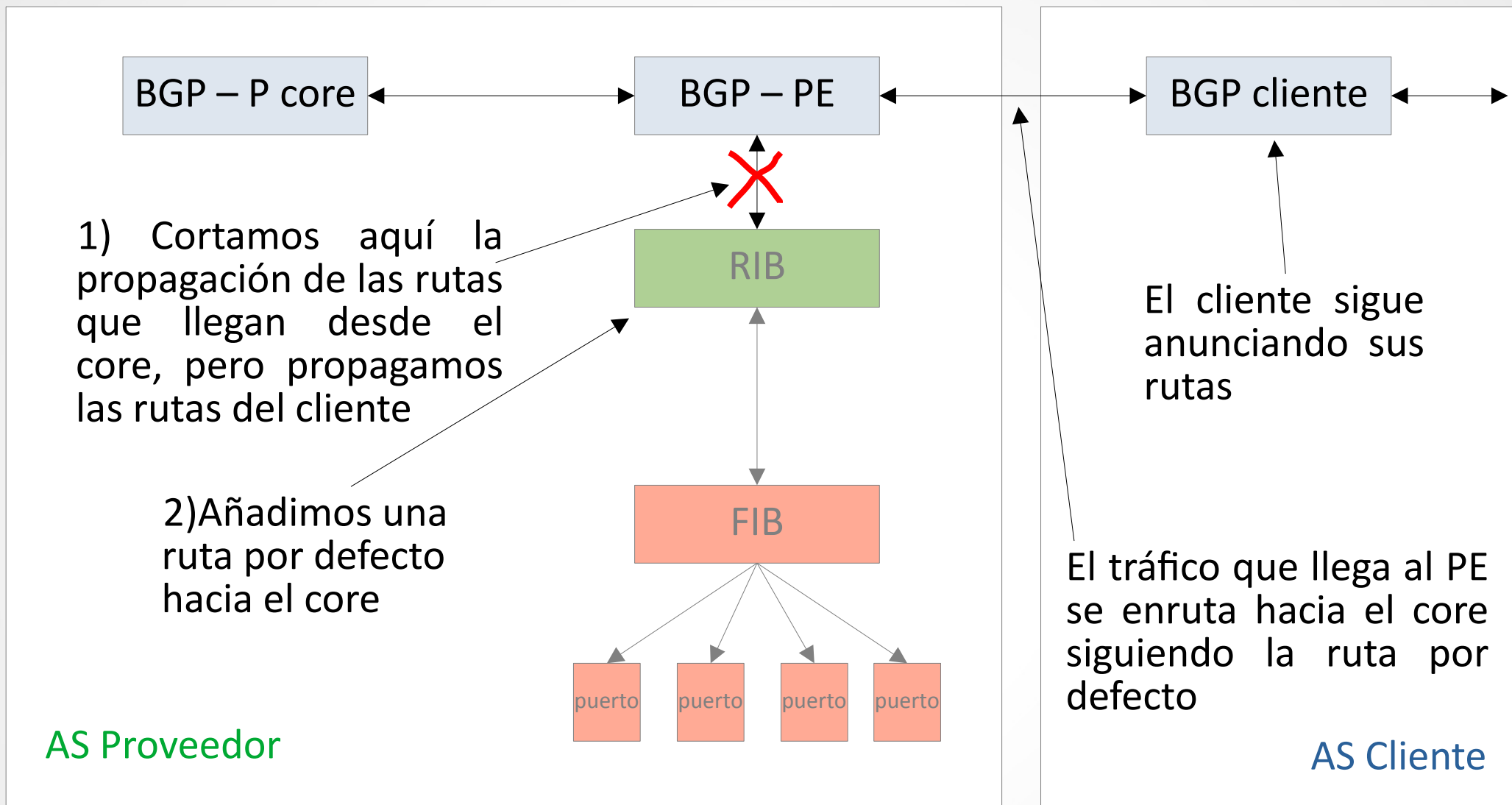
# La idea: teoría



Si no necesitamos mover el tráfico siguiendo las rutas de GBP, podemos mantener sesiones BGP (con toda la FRT) y no propagar las rutas a la RIB y a la FIB.

Eso sí: necesitamos una ruta por defecto, pero solo para el equipo en cuestión.

# La idea: práctica



**Las sesiones BGP (iBGP con el core, eBGP con el cliente) siguen levantadas y mueven toda la FRT**

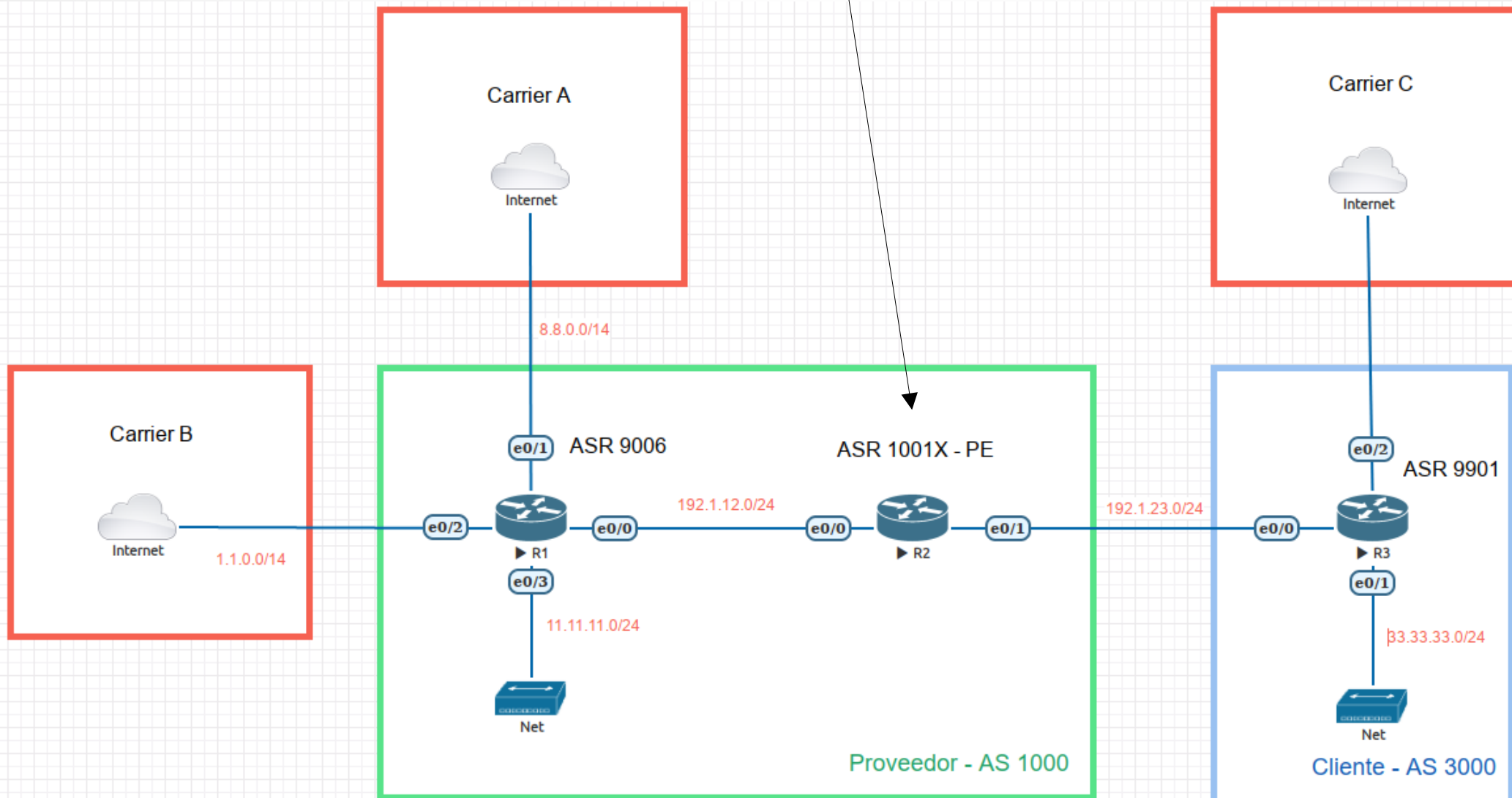
# De la teoría a la práctica



*Imagen: strutmasters.com*

# El taller

Vamos a sustituir este router  
De ASR 1001X a ASR 920



# Antes

La RIB de R2 contiene toda la FRT:

```
R2-ASR1001X#show ip route bgp
```

```
...
```

```
    1.0.0.0/16 is subnetted, 4 subnets
B       1.0.0.0 [200/0] via 10.1.1.1, 00:01:42
B       1.1.0.0 [200/0] via 10.1.1.1, 00:01:42
B       1.2.0.0 [200/0] via 10.1.1.1, 00:01:42
B       1.3.0.0 [200/0] via 10.1.1.1, 00:01:42
    8.0.0.0/16 is subnetted, 4 subnets
B       8.8.0.0 [200/0] via 10.1.1.1, 00:01:42
B       8.9.0.0 [200/0] via 10.1.1.1, 00:01:42
B       8.10.0.0 [200/0] via 10.1.1.1, 00:01:42
B       8.11.0.0 [200/0] via 10.1.1.1, 00:01:42
    11.0.0.0/24 is subnetted, 1 subnets
B       11.11.11.0 [200/0] via 10.1.1.1, 00:01:42
    33.0.0.0/24 is subnetted, 1 subnets
B       33.33.33.0 [20/0] via 192.1.23.3, 00:01:42
```



**FRT**

# Después

La RIB de R2 contiene solo la ruta por defecto y las rutas del cliente:

```
R2-ASR1001X#show ip route bgp
...
B*    0.0.0.0/0 [200/0] via 10.1.1.1, 00:10:06
      33.0.0.0/24 is subnetted, 1 subnets
B      33.33.33.0 [20/0] via 192.1.23.3, 00:07:51
```

...y el tráfico sigue fluyendo:

```
R3-ASR9901#ping 11.11.11.11 source 33.33.33.33
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 11.11.11.11, timeout is 2
seconds:
Packet sent with a source address of 33.33.33.33
!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max =
1/1/1 ms
```



# Resumen de los comandos

## **Router R1:**

```
router bgp 1000
  neighbor 10.2.2.2 default-originate
```

## **Router R2:**

```
ip prefix-list default-only seq 10 permit 0.0.0.0/0
ip prefix-list default-only seq 65535 deny 0.0.0.0/0 le 32
!
ip as-path access-list 1 permit 3000
!
route-map BGP2RIB permit 10
  match ip address prefix-list default-only
!
route-map BGP2RIB permit 20
  match as-path 1
!
route-map BGP2RIB deny 65535
!
router bgp 1000
  address-family ipv4
    table-map BGP2RIB filter
```

# Casos de uso

El router tiene que observar algunas normas:

- Tiene que poder mover el tráfico hacia un solo destino dentro de su propio AS (ruta por defecto)
- Si tiene enlaces con otros ASs, éstos no pueden ser de tránsito

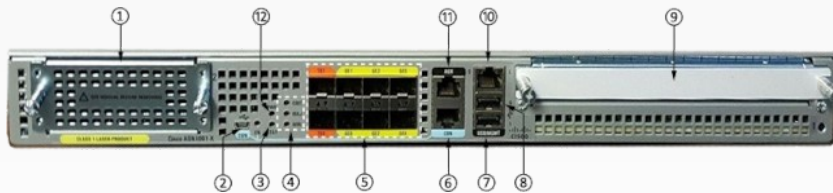
...y cómo siempre:

Se recomienda un estudio previo (comprobar efectiva viabilidad de los equipos)

# ¡Objetivo alcanzado!

## ASR1001X-20G-K9

ASR1001-X with 20G license – 6 x 1G and 2x10G



Precio GPL: \$ 53.900

## ASR-920-24SZ-M

Cisco ASR920 Series – 24 x 1G and 4 x 10G license enabled



Precio GPL: \$ 11.000

