

# Encaminamiento en Internet: BGP

Sistemas Telemáticos para Medios Audiovisuales

Departamento de Teoría de la Señal y Comunicaciones y  
Sistemas Telemáticos y Computación (GSyC)

Octubre de 2018



©2018 Grupo de Sistemas y Comunicaciones.  
Algunos derechos reservados.  
Este trabajo se distribuye bajo la licencia  
Creative Commons Attribution Share-Alike  
disponible en <http://creativecommons.org/licenses/by-sa/3.0/es>

# Contenidos

- 1 CIDR
- 2 Motivación
- 3 Esquema de Internet: ISPs
- 4 Los anuncios de subredes provocan tráfico en sentido opuesto
- 5 Relaciones entre Sistemas Autónomos
- 6 Exportación de subredes teniendo en cuenta las relaciones entre ASs
- 7 BGP
- 8 Referencias

# Contenidos

- 1 CIDR
- 2 Motivación
- 3 Esquema de Internet: ISPs
- 4 Los anuncios de subredes provocan tráfico en sentido opuesto
- 5 Relaciones entre Sistemas Autónomos
- 6 Exportación de subredes teniendo en cuenta las relaciones entre ASs
- 7 BGP
- 8 Referencias

# Antiguo sistema de clases de direcciones

- Clase A (/8), clase B (/16), clase C (/24), clase D, clase E.
- Inconvenientes:
  - Escasez de direcciones de clase B:
    - Cualquier organización que necesitase más direcciones que las que le proporcionaba una clase C ( $2^8 - 2 = 254$  direcciones de máquinas) solicitaba una clase B (65.533 direcciones de máquinas) o una clase A (16 millones de direcciones). El rango de direcciones asignado quedaba reservado para esa organización, que normalmente no usaba todas esas direcciones.
  - Crecimiento de las tablas de encaminamiento de los *routers* en Internet.
    - Una entrada por cada dirección de clase: más de 2 millones de entradas.

# Classless Inter-Domain Routing (CIDR): RFC 4632 (I)

- Utilización de rangos de direcciones, especificados a través de una dirección de red y un prefijo o máscara.
  - los prefijos pueden ser de cualquier longitud, no sólo /8, /16 y /24.
  - Ejemplo: 194.1.4.0/22

	<i>Network</i>	<i>Host</i>
Subred 194.1.4.0:	1100 0010 . 0000 0001 . 0000 01	00 . 0000 0000
Máscara / 22:	1111 1111 . 1111 1111 . 1111 11	00 . 0000 0000

En este caso las direcciones IP que pertenecen a esta subred serían las siguientes:

	<i>Network</i>	<i>Host</i>
Dirección 194.1.4.1:	1100 0010 . 0000 0001 . 0000 01	00 . 0000 0001
Dirección 194.1.4.2:	1100 0010 . 0000 0001 . 0000 01	00 . 0000 0010
Dirección 194.1.4.3:	1100 0010 . 0000 0001 . 0000 01	00 . 0000 0011
Dirección 194.1.4.4:	1100 0010 . 0000 0001 . 0000 01	00 . 0000 0100
...		
Dirección 194.1.7.254:	1100 0010 . 0000 0001 . 0000 01	11 . 1111 1110

# Classless Inter-Domain Routing (CIDR): RFC 4632 (II)

- Para comprobar si una máquina pertenece a un rango de direcciones se realiza la siguiente operación AND (&&) de bits:

¿Dirección IP de máquina && Máscara = Dirección de subred?

Sólo si se obtiene como resultado la dirección de subred entonces la máquina pertenece al rango de direcciones de esa subred.

- Ejemplo: Comprobar si la dirección 194.1.5.1 pertenece al rango 194.1.4.0/22

	<i>Network</i>								<i>Host</i>				
194.1.5.1:	1100	0010	.	0000	0001	.	0000	01		01	.	0000	0001
Máscara /22:	1111	1111	.	1111	1111	.	1111	11		00	.	0000	0000
AND	<hr/>												
194.1.4.0:	1100	0010	.	0000	0001	.	0000	01		00	.	0000	0000

# Agregación en CIDR

- Agregación: Posibilidad de agrupar varias direcciones de subred **contiguas** en una sola, a través de un prefijo común a todas ellas.

Ejemplo: nos asignan estas clases C. ¿Cómo las podemos representar con CIDR?

```
200.0.0.0/24: 1100 1000 . 0000 0000 . 0000 0000 . 0000 0000
200.0.1.0/24: 1100 1000 . 0000 0000 . 0000 0001 . 0000 0000
200.0.2.0/24: 1100 1000 . 0000 0000 . 0000 0010 . 0000 0000
200.0.3.0/24: 1100 1000 . 0000 0000 . 0000 0011 . 0000 0000
200.0.4.0/24: 1100 1000 . 0000 0000 . 0000 0100 . 0000 0000
200.0.5.0/24: 1100 1000 . 0000 0000 . 0000 0101 . 0000 0000
```



# Ejemplo de agregación en CIDR (I)

	/21															
200.0.0.0/24:	1100	1000	.	0000	0000	.	0000	0	000	.	0000	0000				
200.0.1.0/24:	1100	1000	.	0000	0000	.	0000	0	001	.	0000	0000				
200.0.2.0/24:	1100	1000	.	0000	0000	.	0000	0	010	.	0000	0000				
200.0.3.0/24:	1100	1000	.	0000	0000	.	0000	0	011	.	0000	0000				
200.0.4.0/24:	1100	1000	.	0000	0000	.	0000	0	100	.	0000	0000				
200.0.5.0/24:	1100	1000	.	0000	0000	.	0000	0	101	.	0000	0000				
<hr/>																
200.0.6.0/24:	1100	1000	.	0000	0000	.	0000	0	110	.	0000	0000				
200.0.7.0/24:	1100	1000	.	0000	0000	.	0000	0	111	.	0000	0000				

*Eligiendo prefijo /21, agruparíamos direcciones que no tenemos asignadas: 200.0.6.0/24, 200.0.7.0/24*

## Ejemplo de agregación en CIDR (II)

- No se pueden agrupar todas bajo el mismo prefijo. Es necesario usar varios prefijos.

Agregamos con: 200.0.0.0/22

$$\left\{ \begin{array}{l} 200.0.0.0/24: 1100\ 1000\ .\ 0000\ 0000\ .\ 0000\ 0000\ .\ 0000\ 0000 \\ 200.0.1.0/24: 1100\ 1000\ .\ 0000\ 0000\ .\ 0000\ 0001\ .\ 0000\ 0000 \\ 200.0.2.0/24: 1100\ 1000\ .\ 0000\ 0000\ .\ 0000\ 0010\ .\ 0000\ 0000 \\ 200.0.3.0/24: 1100\ 1000\ .\ 0000\ 0000\ .\ 0000\ 0011\ .\ 0000\ 0000 \end{array} \right.$$

Agregamos con: 200.0.4.0/23

$$\left\{ \begin{array}{l} 200.0.4.0/24: 1100\ 1000\ .\ 0000\ 0000\ .\ 0000\ 0100\ .\ 0000\ 0000 \\ 200.0.5.0/24: 1100\ 1000\ .\ 0000\ 0000\ .\ 0000\ 0101\ .\ 0000\ 0000 \end{array} \right.$$

# Ventajas de CIDR

- Las direcciones (recurso escaso) se asignan de forma más eficiente.
- Al poder agregar direcciones, se ahorra espacio en las tablas de encaminamiento de los *routers*.

# Contenidos

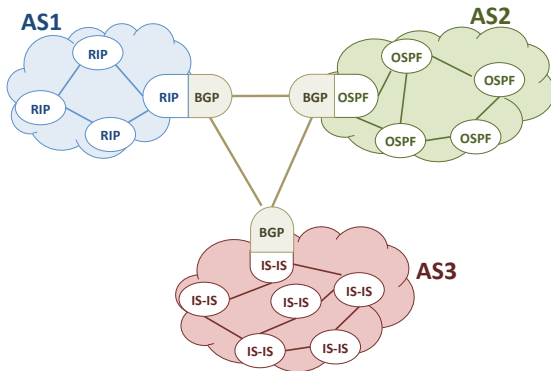
- 1 CIDR
- 2 Motivación**
- 3 Esquema de Internet: ISPs
- 4 Los anuncios de subredes provocan tráfico en sentido opuesto
- 5 Relaciones entre Sistemas Autónomos
- 6 Exportación de subredes teniendo en cuenta las relaciones entre ASs
- 7 BGP
- 8 Referencias

# Dominios de encaminamiento

- La definición de **dominio de encaminamiento** depende del contexto en el que se utilice:
  - Dentro de OSPF, dominio de encaminamiento se refiere a área de encaminamiento, conjunto de routers que inundan la información sobre la topología de encaminamiento.
  - Dentro de BGP (Border Gateway Protocol), dominio de encaminamiento se refiere a conjunto de routers que pertenecen al mismo control administrativo, los gestiona una determinada empresa. Aquellos routers que no los gestiona esa empresa se dice que están fuera de su dominio de encaminamiento. BGP denomina el dominio de encaminamiento **sistema autónomo** (*Autonomous System*, AS).

# Protocolos de encaminamiento interior y exterior

- Los protocolos de encaminamiento interior (OSPF) se centran en encontrar el mejor camino para cada uno de los destinos en función de la métrica número de saltos. Se necesita que converjan rápidamente para conseguir estabilidad en las rutas.
- Los protocolos de encaminamiento exterior (BGP) añaden otros factores, **políticas de encaminamiento**, en el algoritmo de decisión para seleccionar la mejor ruta. Estas políticas están directamente relacionadas con los acuerdos que se establecen en las empresas que administran sus zonas de encaminamiento.



# ¿Por qué no se usa un único protocolo para todo?

- El protocolo debe **escalar a toda la red**, es mejor establecer una jerarquía.
- **No debe exponer información** de una zona de encaminamiento a otras:
  - No se desea que algunos cambios en otras zonas de encaminamiento, tengan impacto en nuestra zona.
  - No se desea exponer información de la topología de nuestra zona a otros competidores.
- Además, la utilización de **políticas para el encaminamiento** permite establecer acuerdos de tránsito de tráfico entre las zonas de encaminamiento. Por ejemplo, encaminar el tráfico:
  - Utilizando el menor número de routers dentro de mi zona (closest-exit o hot potato routing).
  - Utilizando el camino más corto al destino.
  - Utilizando el camino más barato.

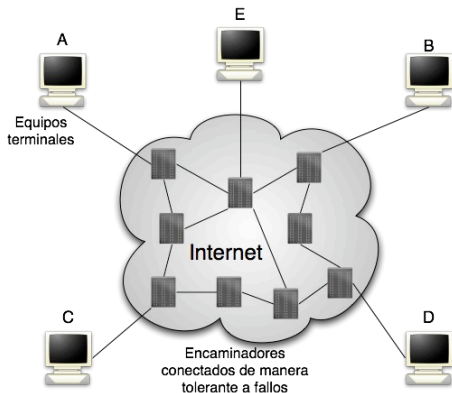
# Contenidos

- 1 CIDR
- 2 Motivación
- 3 Esquema de Internet: ISPs**
- 4 Los anuncios de subredes provocan tráfico en sentido opuesto
- 5 Relaciones entre Sistemas Autónomos
- 6 Exportación de subredes teniendo en cuenta las relaciones entre ASs
- 7 BGP
- 8 Referencias



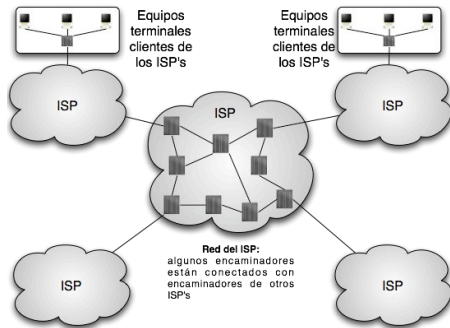
# Esquema idealizado de Internet

- Equipos terminales conectados a encaminadores
- Internet = Grafo conectado de encaminadores “iguales”.  
**FALSO**
- Los encaminadores cooperan intercambiando rutas para obtener los caminos más cortos
- Alto grado de redundancia.  
**FALSO**
- Reparto de carga y recálculo de rutas ante fallos a escala global.  
**FALSO**



# Segunda aproximación

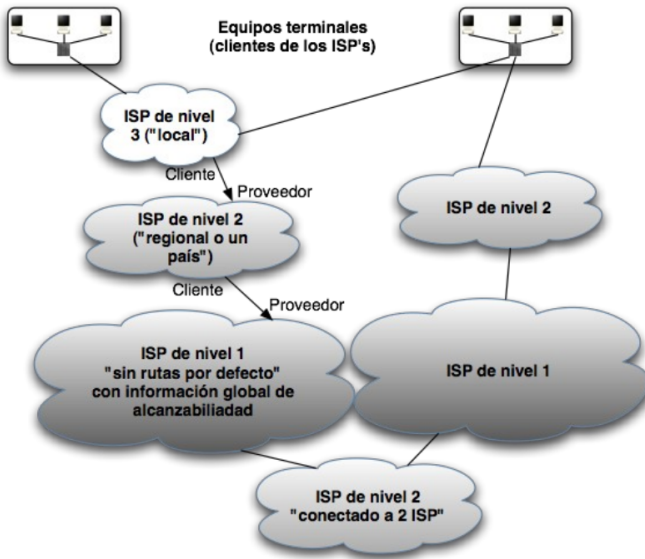
- Proveedores de Servicios de Internet (**ISP: Internet Service Provider**)
- Gran número de empresas (ISPs) que compiten entre sí para dar servicio de encaminamiento global
- **Compiten**, pero tienen que **cooperar**
- La figura sugiere que todos los ISPs son iguales. **FALSO**



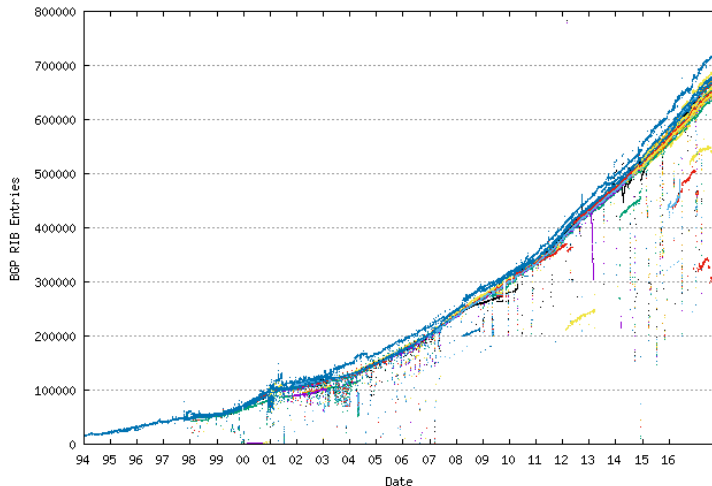
# Esquema real de Internet

- No todos los ISPs son iguales. Hay tres tipos de ISP
- **ISP de nivel 3**
  - Pocos clientes finales, cercanos geográficamente
  - Rutas por defecto hacia sus proveedores de nivel 2
- **ISP de nivel 2**
  - Tamaño intermedio, ámbito regional (estado/país)
  - Rutas por defecto hacia sus proveedores de nivel 1
- **ISP de nivel 1**
  - Ámbito global, sólo unos cuantos en el mundo ( $\approx 15$ )
  - **No hay rutas por defecto** en sus tablas: en sus tablas hay entradas para todas las redes de Internet
- En general consideraremos cada ISP como un Sistema Autónomo (AS, *Autonomous System*).

# Esquema real de Internet



# Entradas BGP activas en las tablas de encaminamiento



Fuente: <http://bgp.potaroo.net>

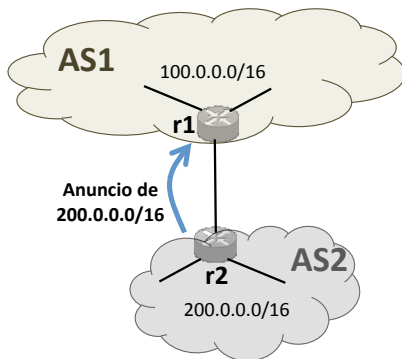
- En 1994 se introdujo CIDR.

# Contenidos

- 1 CIDR
- 2 Motivación
- 3 Esquema de Internet: ISPs
- 4 Los anuncios de subredes provocan tráfico en sentido opuesto**
- 5 Relaciones entre Sistemas Autónomos
- 6 Exportación de subredes teniendo en cuenta las relaciones entre ASs
- 7 BGP
- 8 Referencias

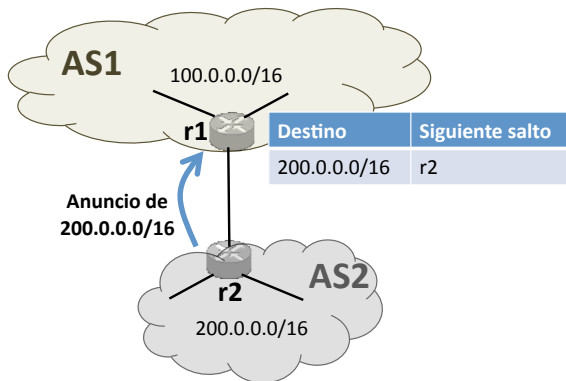
# Tráfico AS1 → AS2: Anuncios de rutas de AS2 a AS1 (I)

- Un sistema autónomo AS2 anuncia sus subredes a otro sistema autónomo AS1 para que las máquinas de AS1 puedan alcanzar las subredes de AS2:



# Tráfico AS1 → AS2: Anuncios de rutas de AS2 a AS1 (II)

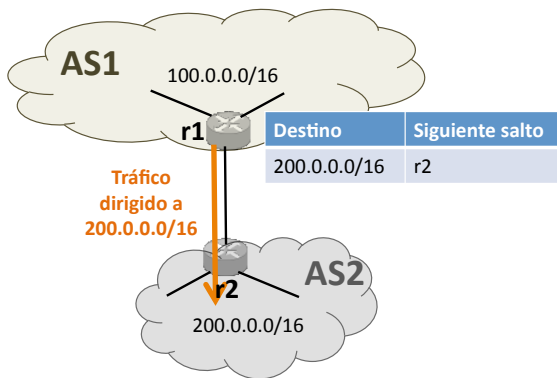
- Un sistema autónomo AS2 anuncia sus subredes a otro sistema autónomo AS1 para que las máquinas de AS1 puedan alcanzar las subredes de AS2:





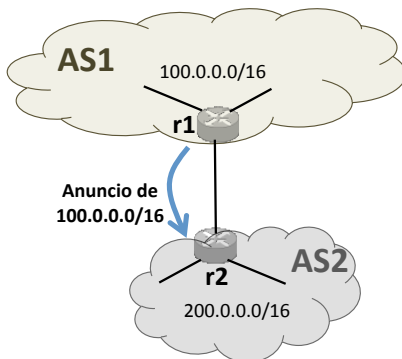
# Tráfico AS1 → AS2: Anuncios de rutas de AS2 a AS1 (III)

- Un sistema autónomo AS2 anuncia sus subredes a otro sistema autónomo AS1 para que las máquinas de AS1 puedan alcanzar las subredes de AS2:



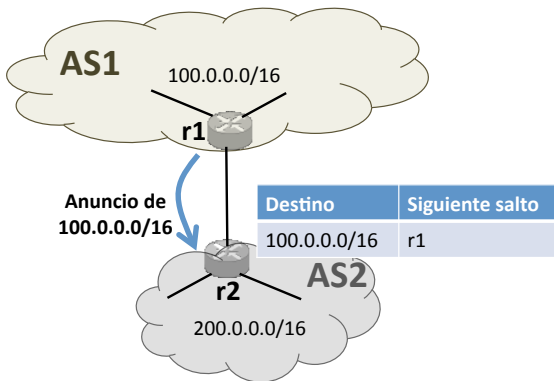
# Tráfico AS2 → AS1: Anuncios de rutas de AS1 a AS2 (I)

- Un sistema autónomo AS1 anuncia sus subredes a otro sistema autónomo AS2 para que las máquinas de AS2 puedan alcanzar las subredes de AS1:



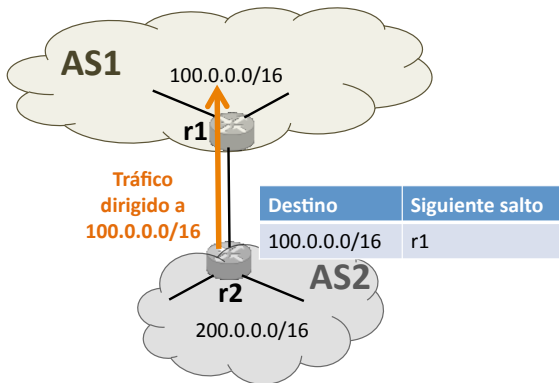
# Tráfico AS2 → AS1: Anuncios de rutas de AS1 a AS2 (II)

- Un sistema autónomo AS1 anuncia sus subredes a otro sistema autónomo AS2 para que las máquinas de AS2 puedan alcanzar las subredes de AS1:



# Tráfico AS2 → AS1: Anuncios de rutas de AS1 a AS2 (III)

- Un sistema autónomo AS1 anuncia sus subredes a otro sistema autónomo AS2 para que las máquinas de AS2 puedan alcanzar las subredes de AS1:



# Contenidos

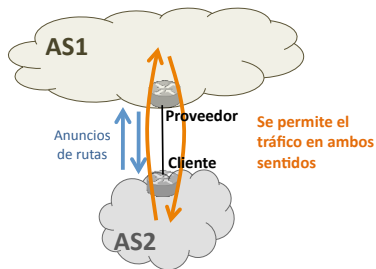
- 1 CIDR
- 2 Motivación
- 3 Esquema de Internet: ISPs
- 4 Los anuncios de subredes provocan tráfico en sentido opuesto
- 5 Relaciones entre Sistemas Autónomos**
- 6 Exportación de subredes teniendo en cuenta las relaciones entre ASs
- 7 BGP
- 8 Referencias

# Relaciones entre ASs

- Las relaciones entre 2 sistemas autónomos determinan qué tráfico va a encaminar cada sistema autónomo y por tanto, qué subredes va a anunciar cada sistema autónomo.
- Dos tipos de relaciones entre ASs:
  - **relación de tránsito** (*transit*)
  - **relación entre iguales** (*peering*)

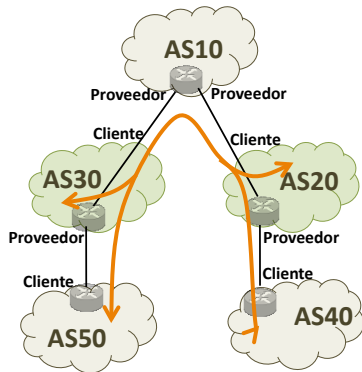
# Relación de tránsito

- Media un contrato económico
- Uno de los ASs es el proveedor de encaminamiento y el otro el cliente
  - El proveedor (p.ej: AS1) proporciona encaminamiento hacia y desde Internet:
  - El proveedor debe anunciar todas las subredes que conoce al cliente.
    - El cliente debe anunciar todas sus subredes al proveedor.
  - El cliente (p.ej.: AS2) paga por ello y debe poder comunicarse con cualquier máquina del planeta.



# Relación entre iguales (I)

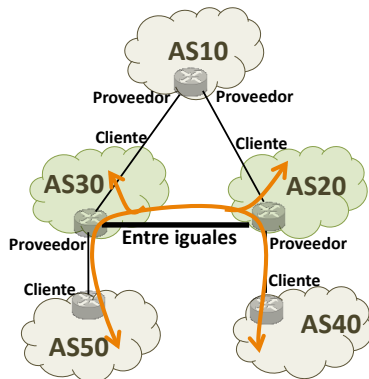
- Si existe mucho tráfico intercambiado entre 2 ASs que no están conectados directamente, puede ser ventajoso para ambos que establezcan una relación más directa.
  - Si AS20 y AS30 utilizan a AS10 para intercambiar su tráfico, AS20 y AS30 estarán pagando a AS10 por todo el tráfico que intercambian.





## Relación entre iguales (II)

- AS20 y AS30 establecen una **relación entre iguales** para no pagar a AS10 por todo el tráfico que intercambien entre estos dos sistemas autónomos y los que son clientes suyos.
- Además AS20 y AS30 consiguen caminos más directos (menor latencia, probablemente menor tasa de pérdida de paquetes. . . )
- Esta relación sólo les vale para comunicarse entre AS20 y AS30 o con sus posibles clientes. Deberán mantener la relación de tránsito para comunicarse con AS10.



# Relación entre iguales

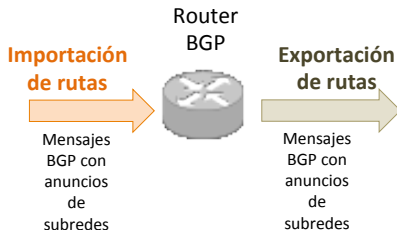
- Las relaciones entre iguales se producen entre empresas competidoras
- No suele mediar contrato económico
- El acuerdo consiste en que cada AS encamine **algunos** de los datagramas IP que van hacia y desde el otro AS
- Mientras no sea muy asimétrico el ratio de tráfico intercambiado ( $< 4:1$ ) no se paga con dinero (“lo comido por lo servido”)

# Contenidos

- 1 CIDR
- 2 Motivación
- 3 Esquema de Internet: ISPs
- 4 Los anuncios de subredes provocan tráfico en sentido opuesto
- 5 Relaciones entre Sistemas Autónomos
- 6 Exportación de subredes teniendo en cuenta las relaciones entre ASs**
- 7 BGP
- 8 Referencias

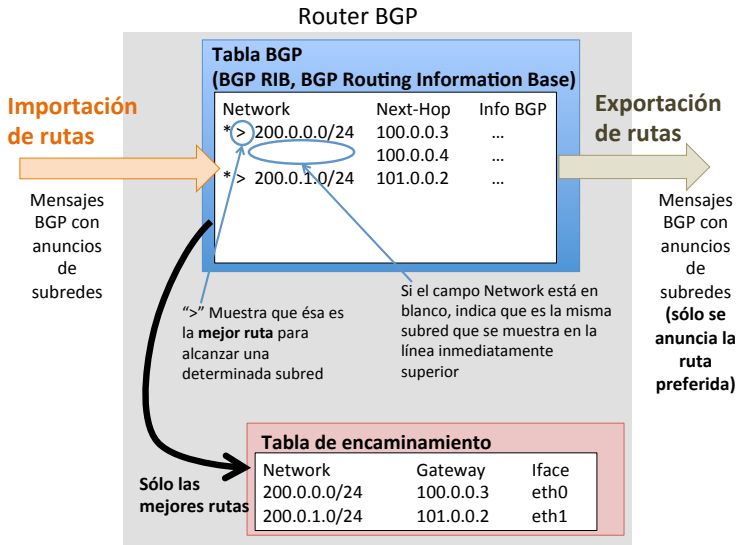
# Motivación de la exportación/importación de subredes

- Las relaciones entre ASs se traducen en la exportación/importación de subredes.



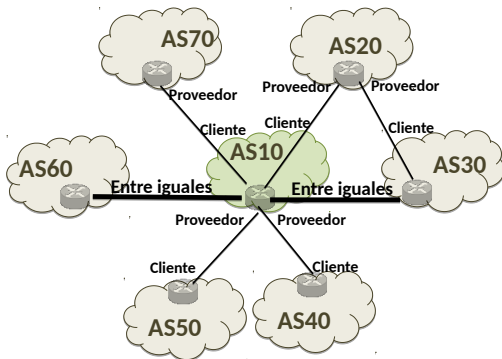
- Exportación de subredes:** subredes que anuncia cada AS. Cada AS debe decidir:
  - Qué subredes anuncia a otros ASs.
- Importación de subredes:** en cada AS se reciben subredes anunciadas por otros ASs. Cada AS debe decidir:
  - Qué subredes incorporar a su tabla BGP.
    - La **tabla BGP** (BGP RIB=Routing Information Base) contiene información de los anuncios BGP recibidos.
  - Si hay varias formas de alcanzar una determinada subred, deberá elegir una ruta preferida. Esa **ruta referida** es la que se incorporará en la tabla de encaminamiento.

# Tabla BGP (BGP RIB) y tabla de encaminamiento



# Exportación de rutas

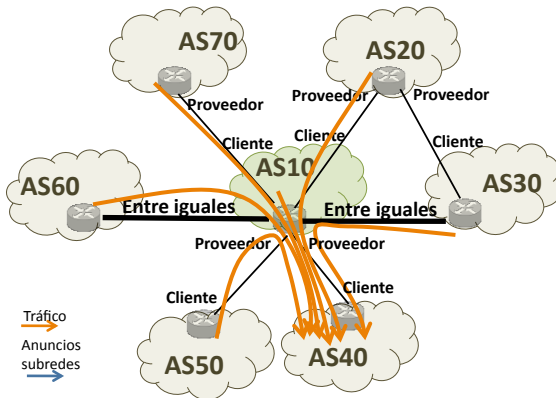
- Qué debe exportar AS10 para que:
  - 1 Sus ASs clientes reciban **todo el tráfico** dirigido a ellos
  - 2 Sus ASs con relación entre iguales reciban el tráfico dirigido a ellos pero sólo el que se genere en AS10 o en los clientes de AS10.
  - 3 Sus ASs proveedores reciban el tráfico dirigido a ellos pero sólo el que genera AS10 o sus clientes.



# Exportación de rutas (I):

## 1. Tráfico hacia un AS cliente (AS10 → AS40)

- Se desea que cualquier máquina pueda alcanzar los ASs que son mis clientes.
- AS10 debe permitir todo el tráfico hacia AS40 (cliente de AS10), desde sus subredes internas a AS10 y desde otros ASs.

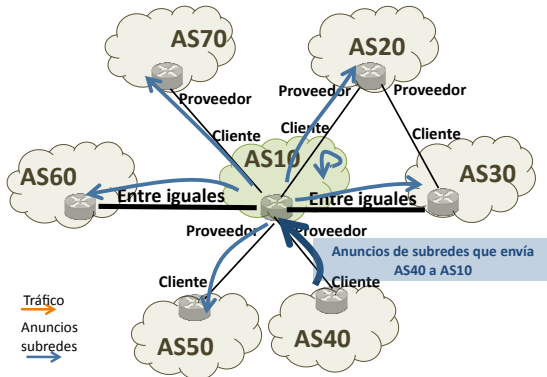


- Igualmente debería hacerlo para AS50 que también es cliente de AS10.

# Exportación de rutas (II):

## 1. Tráfico hacia un AS cliente (AS10 → AS40)

- Para permitir todo el tráfico dirigido a un AS cliente, las subredes que me anuncia dicho AS cliente, debo anunciarlas a todos mis ASs vecinos.
- Ejemplo: AS10 debe anunciar a todos sus ASs vecinos e internamente dentro de AS10, las subredes que ha recibido de AS40.

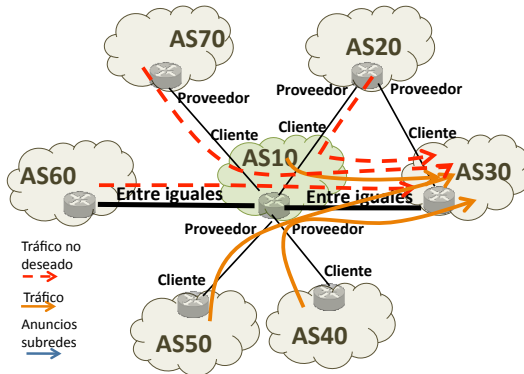




# Exportación de rutas (III):

## 2. Tráfico hacia un AS con relación entre iguales (AS10 → AS30)

- Se desea que sólo las máquinas de mi AS y las de mis ASs clientes utilicen la relación entre iguales.
- Ejemplo: AS10 permitir el tráfico hacia AS30 sólo desde sus subredes internas y desde AS40 y AS50 (sus ASs clientes).

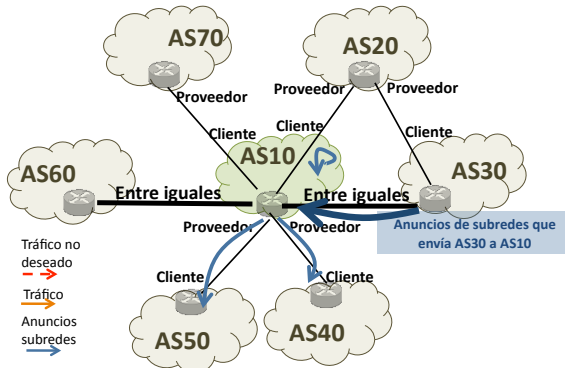


- De forma análoga, AS10 debería permitir hacia AS60 el tráfico desde sus subredes internas y desde AS40 y AS50.

# Exportación de rutas (IV):

## 2. Tráfico hacia un AS con relación entre iguales (AS10 → AS30)

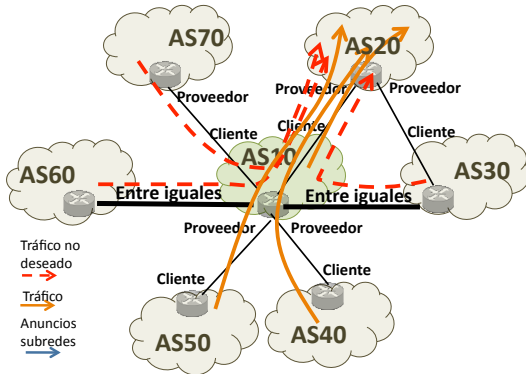
- Para permitir que sólo las máquinas de mi AS y las de mis ASs clientes utilicen la relación entre iguales, las subredes que me anuncia un AS con relación entre iguales:
  - Sólo debo anunciarlas internamente y a mis ASs clientes.
  - No debo anunciarlas a ASs con los que mantengo una relación entre iguales ni a ASs que son mis proveedores.
- Ejemplo: AS10 debe anunciar internamente dentro de AS10 y, a AS40 y AS50 (sus ASs clientes), las subredes que le anuncia AS30.



# Exportación de rutas (V):

## 3. Tráfico hacia un AS proveedor (AS10 → AS20)

- Se desea que sólo las máquinas de mi AS y las de mis ASs clientes utilicen mis relaciones con otros proveedores (relaciones de tránsito).
- Ejemplo: AS10 debe permitir el tráfico hacia AS20 sólo desde sus subredes internas y desde AS40 y AS50 (sus ASs clientes).

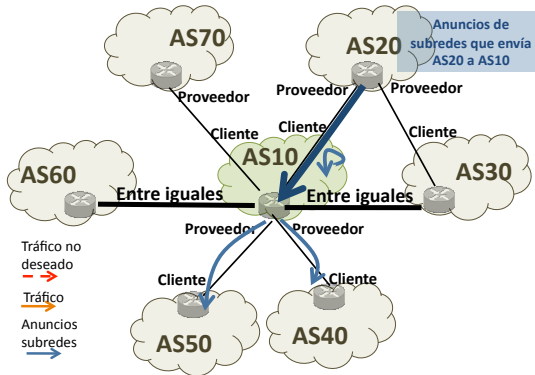


- De forma análoga, AS10 debería permitir hacia AS70 el tráfico desde sus subredes internas y desde AS40 y AS50.

# Exportación de rutas (VI):

## 3. Tráfico hacia un AS proveedor (AS10 → AS20)

- Para permitir que sólo las máquinas de mi AS y las de mis AS clientes utilicen la relación que mantiene una AS con otros proveedores, las subredes que me anuncia un AS proveedor:
  - Sólo debo anunciarlas internamente y a mis ASs clientes.
  - No debo anunciarlas a ASs con los que mantengo una relación entre iguales ni a ASs que son mis proveedores.
- Ejemplo: AS10 debe anunciar internamente dentro de AS10 y, a AS40 y AS50 (sus ASs clientes), las subredes que le anuncia AS20.



# Resumen de exportación de rutas

- Un AS anuncia **las subredes de sus clientes** y las subredes **sus subredes internas** a todos los ASs vecinos.
  - Un AS anuncia **las subredes de sus ASs proveedores** y **las subredes de ASs con relación entre iguales** a sus ASs clientes e internamente dentro del propio AS.
- No** las anuncia **a otros ASs proveedores o a otros ASs con relación entre iguales**

# Contenidos

- 1 CIDR
- 2 Motivación
- 3 Esquema de Internet: ISPs
- 4 Los anuncios de subredes provocan tráfico en sentido opuesto
- 5 Relaciones entre Sistemas Autónomos
- 6 Exportación de subredes teniendo en cuenta las relaciones entre ASs
- 7 BGP**
- 8 Referencias

# BGP: *Border Gateway Protocol* (I)

- Protocolo exterior de encaminamiento, utilizado entre sistemas autónomos
  - BGP se empezó a publicar en 1989. Actualmente se utiliza la versión 4 (RFC-1771, Marzo 1995, actualizado en 2006 por erratas en el documento, ahora RFC 4271).
- Junto con CIDR y NAT ha permitido la subsistencia de IPv4 actual pese a la escasez de direcciones IP y el crecimiento exponencial de las tablas de encaminamiento.
- Dentro de cada sistema autónomo se encamina con un protocolo IGP (RIP, OSPF, IS-IS, IGRP...)

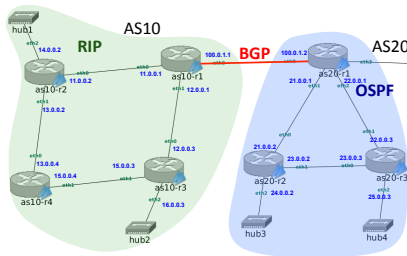
# BGP: *Border Gateway Protocol* (II)

- Internet se organiza en varios Sistemas Autónomos (AS): cada AS pertenece y es administrado por una única entidad comercial (ISP).
- **Cada AS tiene asignado un identificador (ASN=AS Number)**. Originalmente era un número de 16 bits, pero tras estar a punto de agotarse, desde 2009 se utilizan 32 bits (ej: AS100, AS1500, etc). Actualmente hay asignados unos 70.000.
- Es un protocolo de **vector de caminos** (*path vector*): los mensajes de encaminamiento que intercambian los routers contienen las subredes que conocen junto con datos adicionales, p. ej. información del camino necesario para alcanzarlas, de esta forma pueden evitarse bucles de encaminamiento:
  - Un router de un AS incluye el identificador de su AS junto a la subred que anuncia. Si vuelve a recibir este anuncio, podrá detectar que hay un bucle.



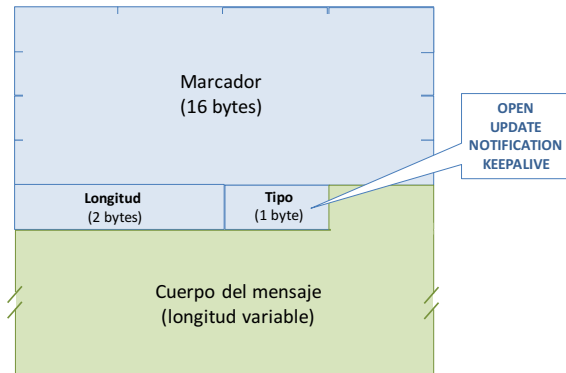
# BGP

- Con BGP **los encaminadores en la frontera** de un sistema autónomo **intercambian prefijos** de redes hacia las que saben encaminar
- **Las rutas aprendidas son inyectadas en el IGP** para distribuir las entre los encaminadores interiores al AS
  - Ej. Si un AS tiene un único encaminador R de salida hacia Internet, todos los encaminadores del AS tienen que tener una ruta que acabe llevando al encaminador R para salir de ese AS.



# Sesión BGP

- Los *routers* BGP establecen **sesiones** entre ellos, utilizando por debajo conexiones TCP en el puerto 179. TCP se encarga de que el envío de mensajes BGP sea fiable.
- Todos los mensajes BGP llevan esta cabecera común, en la que se puede identificar el tipo de mensaje BGP:



# Contenidos

- 1 CIDR
- 2 Motivación
- 3 Esquema de Internet: ISPs
- 4 Los anuncios de subredes provocan tráfico en sentido opuesto
- 5 Relaciones entre Sistemas Autónomos
- 6 Exportación de subredes teniendo en cuenta las relaciones entre ASs
- 7 BGP**
  - Mensajes BGP
    - Selección de la mejor ruta
    - Atributos: ORIGIN, AS\_PATH, NEXT\_HOP, LOCAL\_PREF
    - iBGP y eBGP
- 8 Referencias

# Mensajes BGP:

## OPEN, UPDATE, NOTIFICATION, KEEPALIVE

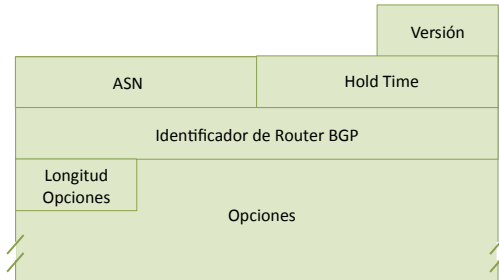
- Tras establecer la conexión TCP, los *routers* BGP se envían mensajes **OPEN** para que se conozcan.
- Posteriormente se aplican las políticas de exportación y se intercambian la información de encaminamiento utilizando mensajes **UPDATE**.
  - En la tabla de BGP pueden almacenarse varias rutas para un mismo destino. Sólo la **ruta preferida** se incorporará a la tabla de encaminamiento del *router*, se inyectará en el IGP, y se anunciará a otros *routers* vecinos BGP.
  - Pueden tardar minutos en el intercambio
- Durante una sesión los *routers* se intercambian periódicamente (cada 60 segundos) mensajes **KEEPALIVE** en ambos sentidos para comprobar que la sesión BGP sigue activa.
- Si se detecta algún problema en la sesión BGP se envía un mensaje **NOTIFICATION** y se cierra la sesión.

# Mensaje OPEN

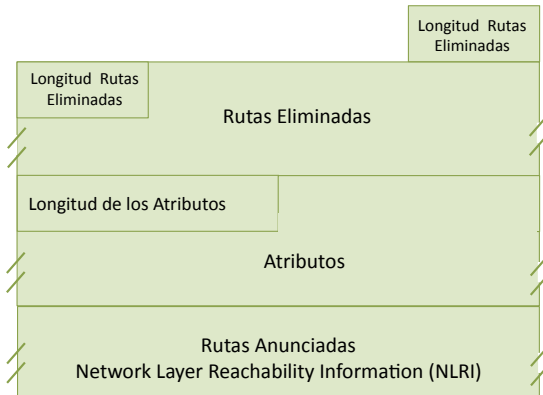
- El mensaje **OPEN** incluye información de:
  - El número del sistema autónomo al que pertenece el router que lo envía
  - El identificador BGP del router
  - El intervalo llamado **hold down** (180 segundos) que se utiliza para desconectar a un vecino si durante dicho intervalo no se han recibido mensajes **KEEPALIVE**.
  - Opciones:

Tipo de opción	Longitud del valor de la opción	Valor de la opción
----------------	---------------------------------	--------------------

Ej: se usa para saber si un router BGP soporta ASN de 32 bits.



# Mensaje UPDATE



- **Rutas eliminadas:** direcciones de subred y sus prefijos.
- **Rutas Anunciadas (NLRI):** direcciones de subred y sus prefijos.
- **Atributos:** información asociada a las rutas anunciadas.

# Actualizaciones de rutas dentro de una sesión BGP

- **Adición/eliminación de rutas**

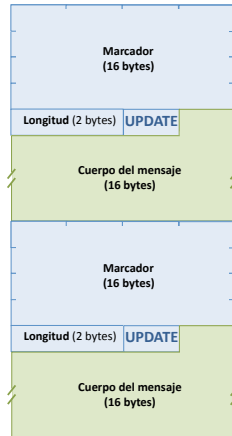
- Si se aprenden nuevos destinos o mejores rutas para llegar a destinos conocidos, se enviarán mensajes **UPDATE** para informar a los *routers* BGP vecinos.
- Si una ruta deja de estar disponible, se enviarán mensajes **UPDATE** para que no se tenga en cuenta dicha ruta de los *routers* BGP vecinos.
- Nunca se anuncia una ruta al vecino del que se aprendió dicha ruta.

- **Desconexión de router BGP vecino**

- Si no se reciben **KEEPALIVE** de un *router* en un intervalo llamado *hold down* (180 segundos, 3 periodos), o si se cierra la conexión TCP de la sesión BGP:
  - se abandona la sesión BGP con ese *router* y se eliminan todas las rutas que nos anunció.
  - si las rutas que hemos eliminado las habíamos anunciado a otros *routers*, se enviarán mensajes **UPDATE** informando de las rutas eliminadas (*withdrawn routes*) a los vecinos BGP.
- Al utilizar TCP (fiable) los mensajes **UPDATE** no se envían periódicamente.

# Concatenación de Mensajes UPDATE

- Los mensajes **UPDATE** pueden viajar concatenados: dentro de un mismo segmento TCP pueden ir varios mensajes **UPDATE** uno a continuación de otro.



- El campo **Marker** (16 bits a uno) junto con la longitud total del mensaje BGP permiten separar varios mensajes BGP dentro de un mismo segmento TCP.



# Contenidos

- 1 CIDR
- 2 Motivación
- 3 Esquema de Internet: ISPs
- 4 Los anuncios de subredes provocan tráfico en sentido opuesto
- 5 Relaciones entre Sistemas Autónomos
- 6 Exportación de subredes teniendo en cuenta las relaciones entre ASs
- 7 BGP**
  - Mensajes BGP
  - Selección de la mejor ruta**
  - Atributos: ORIGIN, AS\_PATH, NEXT\_HOP, LOCAL\_PREF
  - iBGP y eBGP
- 8 Referencias

# Selección de la mejor ruta

- Por defecto, todas las subredes recibidas en anuncios BGP se incorporan a la tabla BGP (BGP RIB). Por tanto, en esta tabla se muestran todas las posibilidades que conoce un router para alcanzar cada una de las subredes.
- Si existe más de una posibilidad de ruta para alcanzar una determinada subred, es necesario seleccionar cuál será la ruta preferida atendiendo a los siguientes criterios:
  - ❶ Si el siguiente salto que debe utilizarse para alcanzar una subred es inaccesible, se descarta la ruta.
  - ❷ Se prefiere ruta con mayor valor de Weight (parámetro configurable en Cisco y quagga).
  - ❸ Se prefiere ruta con mayor valor de atributo LOCAL\_PREF.
  - ❹ Se prefiere ruta generada localmente en el fichero de configuración de BGP.
  - ❺ Se prefiere ruta con el atributo AS\_PATH más corto.
  - ❻ Se prefiere ruta en función de atributo ORIGIN: IGP mejor que EGP y mejor que INCOMPLETE.
  - ❼ Se prefiere ruta con menor atributo Multi-Exit Discriminator (MED).
  - ❽ ...
- Una vez seleccionadas las rutas preferidas:
  - Las rutas preferidas se incorporan a la tabla de encaminamiento de la máquina.
  - Las rutas preferidas se anuncian (si es conveniente) a los ASs vecinos.

# Contenidos

- 1 CIDR
- 2 Motivación
- 3 Esquema de Internet: ISPs
- 4 Los anuncios de subredes provocan tráfico en sentido opuesto
- 5 Relaciones entre Sistemas Autónomos
- 6 Exportación de subredes teniendo en cuenta las relaciones entre ASs
- 7 BGP**
  - Mensajes BGP
  - Selección de la mejor ruta
  - **Atributos: ORIGIN, AS\_PATH, NEXT\_HOP, LOCAL\_PREF**
  - iBGP y eBGP
- 8 Referencias

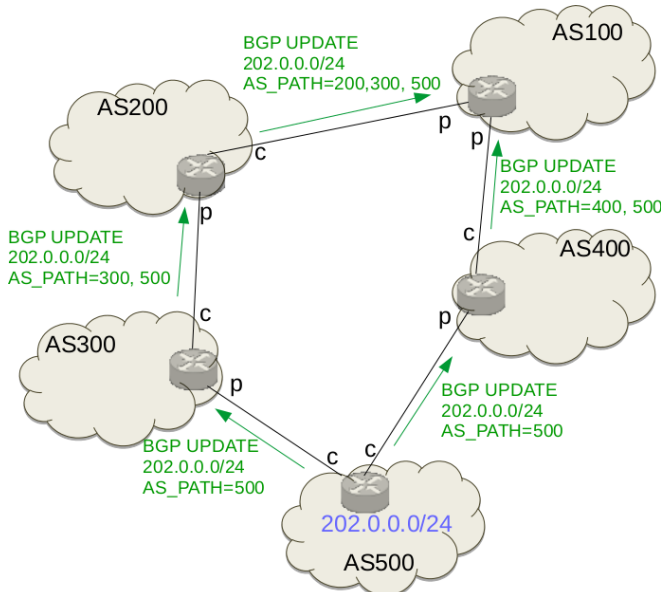
# Atributo ORIGIN

- Atributo obligatorio que define el origen de la información, lo genera el router BGP que crea el anuncio de una determinada subred, los routers que aprenden y reenvían el anuncio de esa misma subred no modifican el valor de **ORIGIN**:
  - IGP (0): Las subredes anunciadas se han generado en un proceso BGP de algún router.
  - EGP (1): Las subredes anunciadas se han aprendido a través de un protocolo de encaminamiento exterior, distinto de BGP.
  - INCOMPLETE (2): Las subredes anunciadas se han aprendido por otras vías (por ejemplo, la redistribución de rutas desde otro protocolo de encaminamiento, rutas estáticas).

# Atributo AS\_PATH (I)

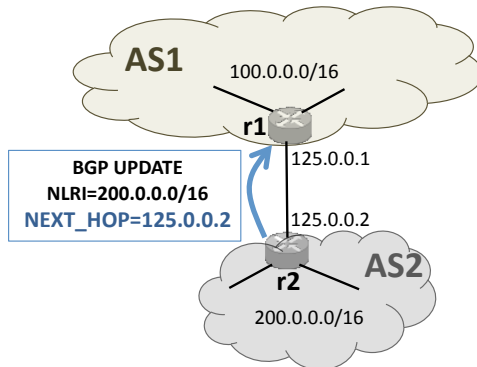
- Es un vector asociado a cada ruta anunciada, con la lista de los identificadores de los AS que ha recorrido el anuncio.
- El **AS\_PATH** permite evitar bucles en las rutas: Cuando un encaminador BGP recibe el anuncio de una ruta, comprueba si su identificador ya está en el **AS\_PATH**
  - si está, es que hay un bucle, y descarta el anuncio
  - si no está, añade su identificador **al principio de la lista de AS\_PATH** y propaga el anuncio (siempre que su política de exportación lo permita) a sus vecinos BGP.

# Atributo AS\_PATH (II)



# Atributo NEXT\_HOP

- Un encaminador BGP pone su dirección IP en el atributo **NEXT\_HOP** de cada ruta que anuncia
- Ésta es la información que permite encaminar luego, ya que los datagramas IP fluyen en el sentido contrario a los anuncios



# Atributo LOCAL\_PREF

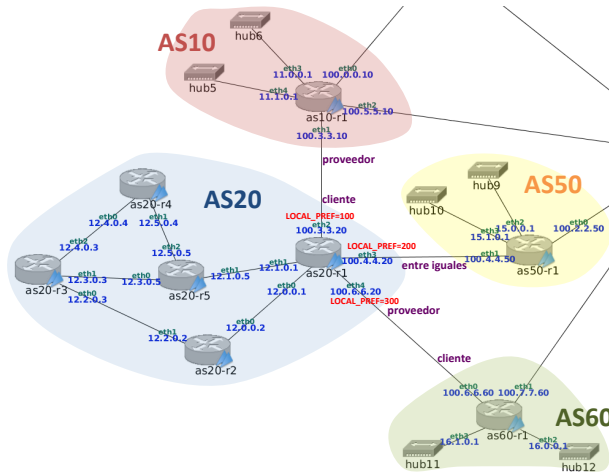
- El atributo **LOCAL\_PREF** es un **valor numérico** que un encaminador puede asignar localmente a un anuncio de ruta recibido de forma que con él pueda fijar las preferencias de elección para una ruta según el origen del anuncio:

clientes > iguales > proveedores.

- El atributo **LOCAL\_PREF** se configura localmente dentro de un AS y no se propaga fuera de dicho AS.
- Cuando un *router* recibe varias rutas para llegar a una misma subred destino, las incluye todas en su tabla BGP. Ahora debe decidir cuál será la ruta preferida que se incluirá en la tabla de encaminamiento de la máquina y se anunciará a los vecinos. El atributo **LOCAL\_PREF** es uno de los parámetros importantes en el proceso de decisión de la ruta preferida.



# Atributo LOCAL\_PREF: Ejemplo



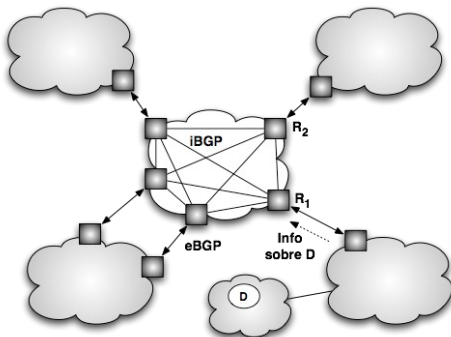
- `as20-r1` marca los anuncios que le llegan por `eth4` con **LOCAL\_PREF=300** por venir de un **cliente**.
- `as20-r1` marca los anuncios que le llegan por `eth3` con **LOCAL\_PREF=200** por venir de un **igual**.
- `as20-r1` marca los anuncios que le llegan por `eth2` con **LOCAL\_PREF=100** por venir de un **proveedor**.

# Contenidos

- 1 CIDR
- 2 Motivación
- 3 Esquema de Internet: ISPs
- 4 Los anuncios de subredes provocan tráfico en sentido opuesto
- 5 Relaciones entre Sistemas Autónomos
- 6 Exportación de subredes teniendo en cuenta las relaciones entre ASs
- 7 BGP**
  - Mensajes BGP
  - Selección de la mejor ruta
  - Atributos: ORIGIN, AS\_PATH, NEXT\_HOP, LOCAL\_PREF
  - iBGP y eBGP
- 8 Referencias

# Sesiones iBGP, eBGP

- **eBGP**: sesiones BGP entre encaminadores BGP **vecinos** de AS distintos (funcionamiento descrito hasta ahora)
- **iBGP**: sesiones BGP entre encaminadores BGP no necesariamente vecinos del mismo AS: necesario IGP para encaminar segmentos TCP



# iBGP

- Un AS puede tener más de un *router* BGP que mantenga sesiones eBGP con AS vecinos
- Cada uno de ellos obtendrá un conjunto de rutas de sus vecinos de otros AS.
- Utilizando iBGP esos *routers* intercambian dentro del AS las rutas aprendidas por cada uno para que todos acaben teniendo las rutas aprendidas por todos los demás
  - Ejemplo: en la figura, R2 tiene que aprender prefijos anunciados para la ruta hacia D aunque R2 no mantenga una sesión eBGP con el AS vecino
- Una vez que todos los encaminadores BGP de un AS tienen todas las rutas y han decidido consistentemente cuáles importar, se inyectan en el IGP (y, en su caso, se anuncian a otros AS vecinos).

# iBGP

- Todos los encaminadores deben usar las mismas reglas para importar rutas, para lo cual se usan las sesiones iBGP
- Topología de las sesiones iBGP: **todos con todos**
  - Es decir, cada *router* BGP del AS mantiene una sesión iBGP con cada uno de los otros *routers* BGP del AS.
- iBGP no es un IGP
- ¿Por qué no utilizar un IGP en lugar de iBGP?
  - En general, los *routers* BGP de un AS NO son vecinos, con lo que los mensajes que envía uno a otro es mucho más probable que se pierdan que si lo fueran
  - Los protocolos IGP no van sobre TCP, se apoyan en retransmisiones periódicas: pero si se pierde un mensaje de BGP transportado por el IGP no se puede recuperar, porque BGP no anuncia periódicamente
  - No se podrían expresar atributos: un IGP busca la métrica más corta, no utiliza atributos (como LOCAL\_PREF).

# Contenidos

- 1 CIDR
- 2 Motivación
- 3 Esquema de Internet: ISPs
- 4 Los anuncios de subredes provocan tráfico en sentido opuesto
- 5 Relaciones entre Sistemas Autónomos
- 6 Exportación de subredes teniendo en cuenta las relaciones entre ASs
- 7 BGP
- 8 Referencias**

# Referencias

- Hari Balakrishnan, **Encaminamiento unicast en redes de área extensa en Internet**. Artículo disponible en la página web de la asignatura.
- Charles M. Kozierok, **TCP/IP GUIDE. A Comprehensive, Illustrated Internet Protocols Reference**, No Starch Press, 2005: capítulo 39  
([http://www.tcpipguide.com/free/t\\_OpenShortestPathFirstOSPF.htm](http://www.tcpipguide.com/free/t_OpenShortestPathFirstOSPF.htm))
- Christian Huitema, **Routing in the Internet, 2nd edition**, Ed. Prentice Hall, 2000.
- RFC 4271, **BGP version 4**:  
<http://www.faqs.org/rfcs/rfc4271.html>