

Tutorial Introducción a las Tecnologías Fundamentales de Internet

Introducción a BGP

LACNIC 25 – 2 de Mayo 2016

Algunos conceptos importantes

Internet es una gran red que interconecta redes. Las redes se agrupan en **sistemas autónomos (AS)**

AS (Sistema Autónomo): Grupo de redes IP que **poseen una política de ruteo propia e independiente**. Cada AS hace su propia gestión del tráfico que fluye entre él y los restantes ASs de Internet. Se identifican con un número de 16 o 32 bits (mas reciente)



FIGURA 1: ROUTERS INTERNOS DE CADA SISTEMA AUTÓNOMO Y ROUTERS DE BORDE

Una característica importante de BGP, es la a interacción de información de ruteo entre sí. Hay desde el punto de vista técnico limitaciones de intercambio de rutas entre un sistema autónomo y otro. Cualquier organización puede interconectarse con cualquier otro sistema autónomo, mas allá del tamaño o complejidad de su red. Hay razones para que un sistema autónomo deba tener políticas de intercambio de rutas. Mas adelante veremos mas detalles de como pueden influir en estas cuestiones.

Algunos conceptos importantes (Cont)

Interconexión de ASs:

corresponden. Hubo casos en que teníamos sesiones tanto en IPv4 como IPv6, mientras que es posible que sólo tengamos configurada una sesión de acuerdo a la topología de la red.

4.5_

Peerings y tránsito

Si bien anteriormente mencionamos que el routing no conlleva una jerarquía en Internet de diseño, en la práctica existen distintos tipos de diseño según la función que cumplen. Es así que podemos hablar de peering entre dos sistemas autónomos cuando acuerdan para intercambiar las rutas de cada uno. Podemos hablar de sistemas autónomos de tránsito cuando un sistema autónomo permite a otros utilizar sus recursos de Internet (o a otros sistemas autónomos).

Algunos conceptos importantes (Cont)

IGP: Interior Gateway Protocol, realiza el proceso de routing DENTRO del AS.

Ej: RIP, EIGRP, OSPF, IS-IS.

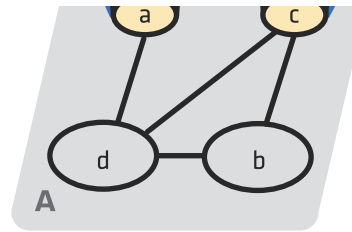


FIGURA 1: ROUTERS INTERNOS DE CADA SISTEMA AUTÓNOMO Y ROUTERS DE BORDE

Una característica importante de BGP, es la ausencia de intercambio de información de ruteo entre sistemas autónomos. Hay desde el punto de vista técnico limitaciones en las posibilidades de intercambio de rutas entre un sistema autónomo y otro. Cualquier organización puede interconectarse con otras, mas allá del tamaño o complejidad de su red interna. Hay razones para que un sistema autónomo deba depender de otros. Mas adelante veremos mas razones que pueden influir en estas cuestiones.

Algunos conceptos importantes (Cont)

EGP: Exterior Gateway Protocol, realiza el proceso de routing ENTRE los AS.

Ej: **BGP**

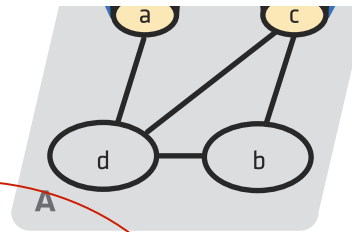


FIGURA 1: ROUTERS INTERNOS DE CADA SISTEMA AUTÓNOMO Y ROUTERS DE BORDE

Una característica importante de BGP, es la ausencia de intercambio de información de ruteo entre sistemas autónomos. Hay desde el punto de vista técnico limitaciones en las capacidades de intercambio de rutas entre un sistema autónomo y otro. Cualquier organización puede interconectarse con otros sistemas autónomos, sin importar el tamaño o complejidad de su red interna. Sin embargo, hay razones para que un sistema autónomo deba depender de otros sistemas autónomos. Veremos más adelante que puede haber razones técnicas y políticas que pueden influir en estas cuestiones.

Algunos conceptos importantes (Cont)

- **Tabla de ruteo:** Tabla electrónica que almacena las rutas a los diferentes destinos. Cada router consulta la tabla de enrutamiento con el fin de encontrar la mejor ruta para la transferencia de datos.
 - Existe una tabla por cada protocolo que el router maneja: tabla de isis, tabla de rip, tabla de BGP.
 - Los protocolos “compiten” para que sus rutas sean las que finalmente se utilicen (Mediante una propiedad que se llama “Distancia Administrativa”)
 - La ruta “elegida” es la que pasa a formar parte de la tabla de ruteo
 - Importante: consideramos varias tablas de protocolos (una por cada uno de ellos) pero solo **1 Tabla de Ruteo**

Algunos conceptos importantes (Cont)

- **Next-Hop (NH):** Me indica, en la tabla de rutas, a través de qué IP alcanzaré el destino. (Sólo me dice el siguiente salto, no la ruta completa)
- No necesariamente es la IP del router directamente conectado.

BGP

- **BGP (Border Gateway Protocol):** protocolo de ruteo externo (EGP) utilizado en Internet para el intercambio de rutas de diferentes Ass
- Usa TCP como protocolo de transporte (port 179)
- Originalmente lanzado en 1989 a través del RFC 1105. La versión más actualizada es BGP4 y ha sido definida en 1995 a través de la RFC 4271.
- BGP4 es el primer protocolo de enrutamiento exterior que soporta CIDR y sumarización de rutas.

BGP (Cont)

- Toma decisiones basándose en los atributos de una ruta.
- Existen dos casos
 - **iBGP**: dentro del sistema autónomo (Los peers no necesitan estar directamente conectados)
 - **eBGP**: entre sistemas autónomos

No pasa la complejidad interna a INTERNET.

Nota: **Cuidado**: iBGP no es un IGP

Cómo trabaja BGP?

- **Mediante el “aprendizaje” y “anuncio” de rutas:** para que un router sepa hacia donde encaminar un paquete, alguien tiene que decirle por dónde hacerlo (**aprende**). De la misma manera, si queremos que alguien llegue a nuestras redes tenemos que “comunicar” que las tenemos (**anuncia**).
- Para BGP, el ruteo interno es una caja negra.

Como es el proceso?

- 1_ Se inicia una sesion BGP entre dos routers
- 2_ Cada router comienza a *anunciar* las rutas
- 3_ Cada router comienza a *aprender* rutas que almacena en la tabla de BGP

Aprender y Anunciar

- **Aprender una ruta:** es incorporar en la tabla de bgp alguna ruta que me está enseñando otro router.
 - **Anunciar una ruta:** comunicarle a otro router la ruta que tengo almacenada en la tabla de BGP, para llegar a determinado destino.
- Importante al anunciar:*** No basta con tener la ruta en la tabla de BGP, la debo tener tambien en la tabla de ruteo.

Cómo es el proceso? (Cont)

1_ Se inicia una sesión BGP entre dos routers

- Los routers que conforman una sesión BGP se denominan “Neighbors”: punto remoto con el cual enseñar y/o aprender rutas.
- Para que la sesión BGP se establezca, el potencial neighbor debe ser alcanzable por IP.
- La sesión establecida puede ser eBGP o iBGP

Cómo es el proceso? (Cont)

2_ Cada router comienza a anunciar las rutas

- El router podrá anunciar las rutas que:
 - Aprenda de otras sesiones BGP
 - El propio router inserte en la tabla de BGP (Ya veremos cómo)
- Cuidado: la ruta debe estar en la tabla de ruteo
- Pertenezcan a otros protocolos (en caso de utilizar redistribute... Cuidado!)

Cómo es el proceso? (Cont)

3_ Cada router comienza a aprender rutas que almacena en la tabla de BGP

- Según la política de anuncio del otro neighbor

Tener en cuenta...

- Qué tráfico afecta las rutas que aprendo?
 - Saliente
- Qué tráfico afecta las rutas que enseño?
 - Entrante
- Qué pasa si no aprendo nada?
 - Depende de lo que tenga configurado localmente en mi tabla de ruteo.

Resumamos lo que vimos...

- La interconexión de redes en Internet se realiza entre AS
- Los AS utilizan BGP para comunicarse
- Los routers que *hablan* BGP con otros AS son routers de borde.
- Los routers saben que ruta tomar segun una tabla de ruteo
- Cada protocolo de ruteo posee su propia tabla de rutas
- Existe eBGP e iBGP
- BGP trabaja aprendiendo y anunciando rutas
- El NH me indica el próximo salto

Configuración Básica

- **router bgp <ASN>**

- crea el proceso BGP dentro del router

- **neighbor <ip> remote-as <AS-remoto>**

- vecinos internos (mismo AS)
- vecinos externos (distinto AS)
- generalmente los externos comparten una subred, son adyacentes. Los internos (iBGP) pueden estar en cualquier parte del Sistema Autónomo.

• **Nota:** Cisco IOS permite ejecutar sólo un proceso BGP a la vez, por lo tanto, **un router no puede pertenecer a más de un Sistema Autónomo.**

Configuración Básica

- **network <red> [mask <máscara>]**

- Se da de alta una red en la tabla de BGP y la marca como local al AS
- Sintaxis similar pero diferente significado con el comando network de un IGP (en RIP, quedan determinadas las interfaces que están sobre esa red para enviar y recibir actualizaciones, así como que redes conectadas directamente a esas interfaces se anunciarán)
- La máscara permite especificar supernetting o subnetting

Configuración Básica

- Ejemplo:

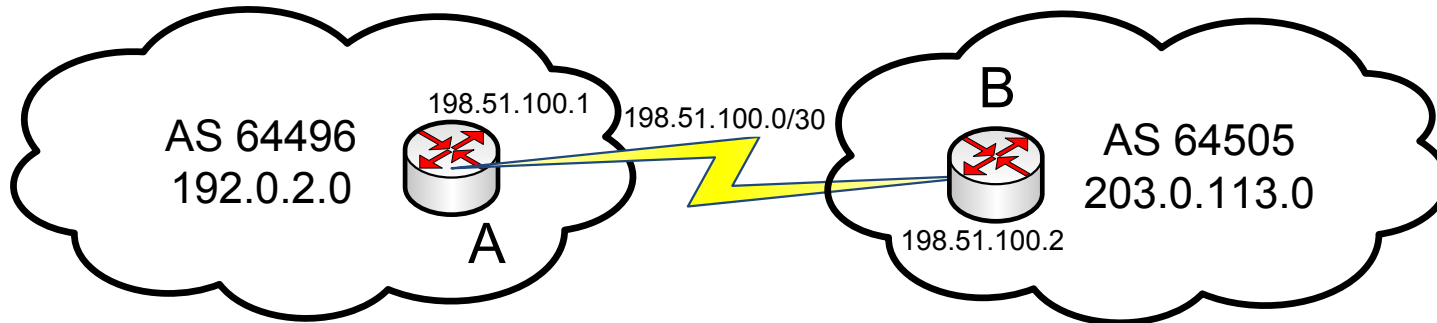
```
router bgp 64496
```

```
network 203.0.113.0 mask 255.255.255.0
```

```
neighbor 192.0.2.2 remote-as 64500
```

```
neighbor 192.0.2.22 remote-as 64505
```

Configuración Básica



Configuración Router A

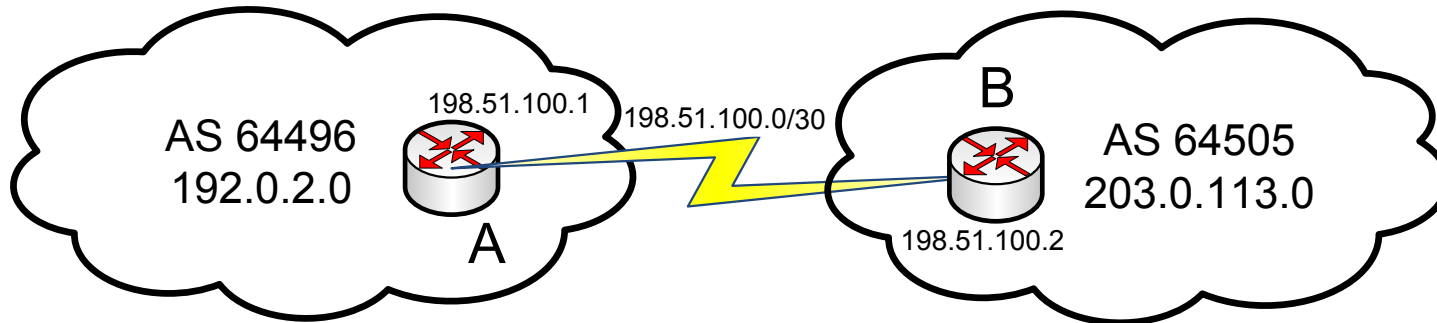
```
router bgp 64496
network 192.0.2.0
neighbor 198.51.100.2 remote-as 64505
```

Configuración Router B

```
router bgp 64505
network 203.0.113.0
neighbor 198.51.100.1 remote-as 64496
```

- Esas rutas deben existir en la tabla de ruteo del router local o no serán enviadas en las actualizaciones.
- También podría usarse cada comando `network` con el parámetro `mask 255.255.255.0`
- Las rutas aprendidas son propagadas por defecto (pueden ser filtradas por una política de ruteo).

Configuración Básica



Configuración Router A

```
router bgp 64496
network 192.0.2.0
neighbor 198.51.100.2 remote-as 64505
```

Configuración Router B

```
router bgp 64505
network 203.0.113.0
neighbor 198.51.100.1 remote-as 64496
```

- Para establecer las relaciones con otros router BGP están los comandos neighbor.
- Ese comando sirve para identificar un vecino con el cual el router local establece la sesión.
- El argumento ASN determina si el router vecino es EBGP o IBGP.

Configuración Básica

- Comandos “show” básicos

- `show ip bgp`
- `show ip bgp summary`
- `show ip bgp <prefix>`
- `show ip bgp neighbors <ip> advertised-routes`
- `show ip bgp neighbors <ip> routes`
- `show ip bgp regexp <regexp>`

show ip bgp

route-server.phx1>sh ip bgp

BGP table version is 1318967429, local router ID is 67.17.81.28

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,

r RIB-failure, S Stale, m multipath, b backup-path, x best-external

Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i1.0.0.0/24	67.16.148.37	50	200	0	15169 i
*>i1.0.4.0/24	67.16.148.37	50	200	0	3257 4826 56203 i
*>i1.0.7.0/24	67.16.148.37	50	200	0	3257 4826 56203 56203 56203 i
*>i1.0.204.0/22	67.16.148.37	100	201	0	3356 4651 9737 23969 i
*>i1.1.1.0/24	67.16.148.37	50	200	0	15169 i
*>i1.1.8.0/24	67.16.148.37	100	200	0	4134 i
*>i1.1.64.0/19	67.16.148.37	0	300	0	10026 2519 i
*>i1.11.0.0/21	67.16.148.37	100	201	0	3356 15412 15412 15412 15412 9848 38091
38091 38091 38091 38091 38091 18313 i					
*>i2.0.0.0/16	67.16.148.37	50	199	0	5511 3215 i
*>i2.18.250.0/23	67.16.148.37	50	199	0	5511 34164 34164 i
*>i5.63.160.0/24	67.16.148.37	50	200	0	1299 44395 44395 201921 e
*>i5.102.160.0/21	67.16.148.37	0	202	0	286 8881 8881 8881 8881 8881 8881 8881
199284 i					
*>i5.202.254.0/23	67.16.148.37	50	200	0	1299 20771 48159 48159 48159 198357 ?
* i6.64.166.0/24	67.16.148.37	50	200	0	209 721 27064 1503 i
*>i6.64.166.0/24	67.16.148.37	100	200	0	209 721 27064 1468 i
*>i6.64.181.0/24	67.16.148.37	50	200	0	209 721 i
*>i17.253.52.0/23	67.16.148.37	100	201	0	3356 714 6185 i
* i17.253.66.0/24	67.16.148.37	50	200	0	4637 714 6185 e
* i17.253.66.0/24	67.16.148.37	150	200	0	2516 714 6185 e
*>i17.253.66.0/24	67.16.148.37	150	200	0	2516 714 6185 e
*>i17.253.72.0/23	67.16.148.37	50	200	0	1299 4766 714 i
*>i17.253.80.0/24	67.16.148.37	50	200	0	4637 714 6185 e
*>i20.143.208.0/20	67.16.148.37	100	300	0	7473 7474 17647 i
*>i23.2.36.0/22	67.16.148.37	50	200	0	2914 20940 20940 16625 i



show ip bgp

- Las primeras tres columnas listan el status de cada ruta.
- *, en la primer columna: la ruta tiene un next-hop válido. Otras opciones:
 - "s" suppressed: BGP conoce la ruta, pero no está siendo anunciada, usualmente porque es parte de una ruta sumariada.
 - "d" dampened: BGP detiene el anuncio de una ruta que produce un efecto denominado flapping (se levanta y se baja)
 - "h" history: BGP conoce la red, pero no hay una ruta válida hacia ella.
 - "r" RIB* failure: la ruta es anunciada a BGP pero no es instalada en la tabla de ruteo.
 - "S" stale: indica que la ruta está “detenida” y requiere ser refrescada cuando se reestablezca la conexión con su vecino.

BGP table version is 134358, local router ID is 198.51.100.1

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, S Stale

Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	192.0.2.0/26	198.51.100.1		0		64496 65511 i
*>	192.0.2.128/26	198.51.100.1		0		64496 64506 i
*>i		198.51.100.22	0	0		64506 e
*	203.0.113.128/25	198.51.100.1		0		64496 64501 i
*>		198.51.100.114	0	0		64501 i
*>	203.0.113.0/25	198.51.100.1		0		64496 ?

show ip bgp

- La cuarta columna lista las redes. La que no tiene una máscara de subred, usa la máscara classfull (A /8, B /16, C /24). Cuando el router aprende la misma red desde múltiples fuentes, sólo lista la red una vez.

```
BGP table version is 134358, local router ID is 198.51.100.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, S
Stale
```

```
Origin codes: i - IGP, e - EGP, ? - incomplete
```

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	192.0.2.0/26	198.51.100.1			0	64496 65511 i
*>	192.0.2.128/26	198.51.100.1			0	64496 64506 i
*>i		198.51.100.22	0	0		64506 e
*	203.0.113.128/25	198.51.100.1			0	64496 64501 i
*>		198.51.100.114	0	0		64501 i
*>	203.0.113.0/25	198.51.100.1			0	64496 ?

show ip bgp summary

```
BGP router identifier 192.0.2.51, local AS number 65537
BGP table version is 48347, main routing table version 48347
2558 network entries and 3869 paths using 389968 bytes of memory
527 BGP path attribute entries using 28600 bytes of memory
250 BGP AS-PATH entries using 6304 bytes of memory
1 BGP community entries using 24 bytes of memory
191 BGP route-map cache entries using 3056 bytes of memory
1028 BGP filter-list cache entries using 12336 bytes of memory
BGP activity 14929/71963 prefixes, 42175/38306 paths
```

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
192.0.2.114	4	64496	38357	31873	48347	0	0	6d08h	1342
192.0.2.122	4	64498	31728	34362	48347	0	0	6d09h	97
192.0.2.182	4	65551	32203	34388	48347	0	0	2w1d	70
192.0.2.190	4	65537	31146	62851	48347	0	0	2w1d	4
192.0.2.238	4	65537	0	0	0	0	0	never	Active
203.0.113.1	4	65537	38217	52708	48347	0	0	6d08h	2329
203.0.113.11	4	65537	31057	57875	48347	0	0	6d08h	3
203.0.113.23	4	65537	31066	57722	48347	0	0	2w1d	1
203.0.113.114	4	65537	244830	57881	48347	0	0	6d07h	0
198.51.100.2	4	65537	33504	31208	48347	0	0	6d07h	18
198.51.100.117	4	65537	31048	31242	48347	0	0	2d14h	2

show ip bgp <prefix>

```
route-server.phx1>sh ip bgp 170.210.0.0
BGP routing table entry for 170.210.0.0/16, version 6559918
Paths: (5 available, best #5, table default)
  Not advertised to any peer
  6762 7303 4270
    67.16.148.37 from 4.69.244.197 (4.69.244.197)
      Origin IGP, metric 100, localpref 200, valid, internal
      Community: 3549:2406 3549:30840
      Originator: 67.17.82.183, Cluster list: 0.0.6.8
  6762 7303 4270
    67.16.148.37 from 4.69.244.133 (4.69.244.133)
      Origin IGP, metric 100, localpref 200, valid, internal
      Community: 3549:2101 3549:30840
      Originator: 67.17.82.114, Cluster list: 0.0.6.17
  6762 7303 4270
    67.16.148.37 from 4.69.244.196 (4.69.244.196)
      Origin IGP, metric 100, localpref 200, valid, internal
      Community: 3549:2406 3549:30840
      Originator: 67.17.82.183, Cluster list: 0.0.6.8
  6762 7303 4270
    67.16.148.37 from 4.69.244.132 (4.69.244.132)
      Origin IGP, metric 100, localpref 200, valid, internal
      Community: 3549:2101 3549:30840
      Originator: 67.17.82.114, Cluster list: 0.0.6.17
  6762 7303 4270
    67.16.148.37 from 4.69.243.136 (4.69.243.136)
      Origin IGP, metric 100, localpref 200, valid, internal, best
      Community: 3549:2352 3549:30840
      Originator: 67.17.81.117, Cluster list: 0.0.6.7
```

show ip bgp neighbors <ip> advertised-routes

```
show ip bgp neighbors 192.0.2.51 advertised-routes
```

BGP table version is 48402, local router ID is 192.0.2.65

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal

Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	192.0.2.64/26	0.0.0.0		0	32768	i
*>	192.0.2.128/27	198.51.100.3		0		64496 64506 i
*>	192.0.2.160/27	198.51.100.22	0	0		64506 e
*>	192.0.2.192/27	198.51.100.35		100		65550 i
*>	192.0.2.224/27	198.51.100.2		0		65549 i

show ip bgp neighbors <ip> routes

```
show ip bgp neighbors 192.0.2.51 routes
```

BGP table version is 48402, local router ID is 192.0.2.51

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal

Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i203.0.113.0/26	198.51.100.1	0	600	0	65549 i
*>i203.0.113.64/26	198.51.100.65	0	600	0	65547 65548 i
*>i203.0.113.128/27	198.51.100.129	0	600	0	65547 i
*>i203.0.113.160/27	198.51.100.161	0	600	0	65548 i
*>i203.0.113.192/27	198.51.100.193	0	600	0	65496 e
*>i203.0.113.224/27	198.51.100.225	0	600	0	65500 ?

Configuración Básica

- Comandos “clear”: a veces necesario para “refrescar/limpiar” una sesión BGP
 - Formato: clear ip bgp <neighbor-address>
 - Las nuevas versiones soportan “soft in” y “soft out” (lo que aprende y lo que enseña).
 - Implica menos carga para el router.
 - Los clear indiscriminados pueden afectar al router (CPU, inestabilidades).
 - (OJO: clear ip bgp *)

Atributos de rutas

Parámetros que utiliza el BGP para determinar cuál es la ruta que deberá “competir” para ser parte de la Tabla de Ruteo.

Veremos en este tutorial:

- Origin
- Next-hop
- AS_PATH

Y También:

- Multi_EXIT_DISC (Multi-Exit Discriminator) (MED)
- LOCAL_PREF (Local Preference)
- WEIGHT

Atributos de rutas: origen

- Informa a todos los sistemas autónomos como fue introducido el prefijo de red

- 3 valores: IGP, EGP, incomplete

- i originada en un IGP, anunciada con “network”

- e originada en un EGP (BGP a BGP)

- ? origen desconocido, normalmente producto de una

redistribución incompleta desde otro protocolo de ruteo

BGP table version is 134358, local router ID is 198.51.100.1

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, S Stale

Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	192.0.2.0/26	198.51.100.1			0	64496 65511 i
*>	192.0.2.128/26	198.51.100.1			0	64496 64506 i
*>i		198.51.100.22	0	0		64506 e
*	203.0.113.128/25	198.51.100.1			0	64496 64501 i
*>		198.51.100.114	0	0		64501 i
*>	203.0.113.0/25	198.51.100.1			0	64496 ?

Atributos de rutas: next-hop

- No necesariamente es la IP del router directamente conectado.
- En un IGP, el next-hop (NH) es la IP del router que anunció la ruta.
- En BGP:
 - En las sesiones eBGP, el NH es la IP del neighbor que anunció la ruta.
 - En las sesiones iBGP:
 - Rutas originadas dentro del AS: NH es la IP del router que la origina.
 - Rutas incorporadas por eBGP: se transporta inalterado el NH aprendido por eBGP.
- Un next-hop 0.0.0.0 indica que el router local originó la ruta.

Atributos de rutas: next-hop

BGP table version is 134358, local router ID is 198.51.100.1

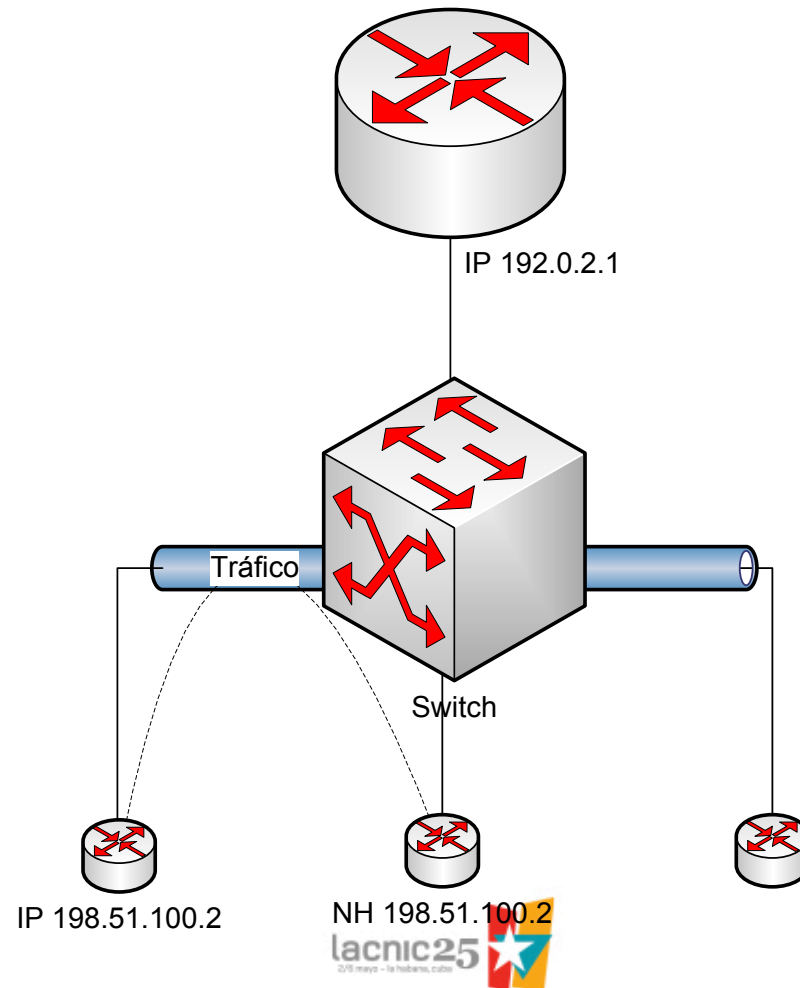
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
S Stale

Origin codes: i - IGP, e - EGP, ? - incomplete

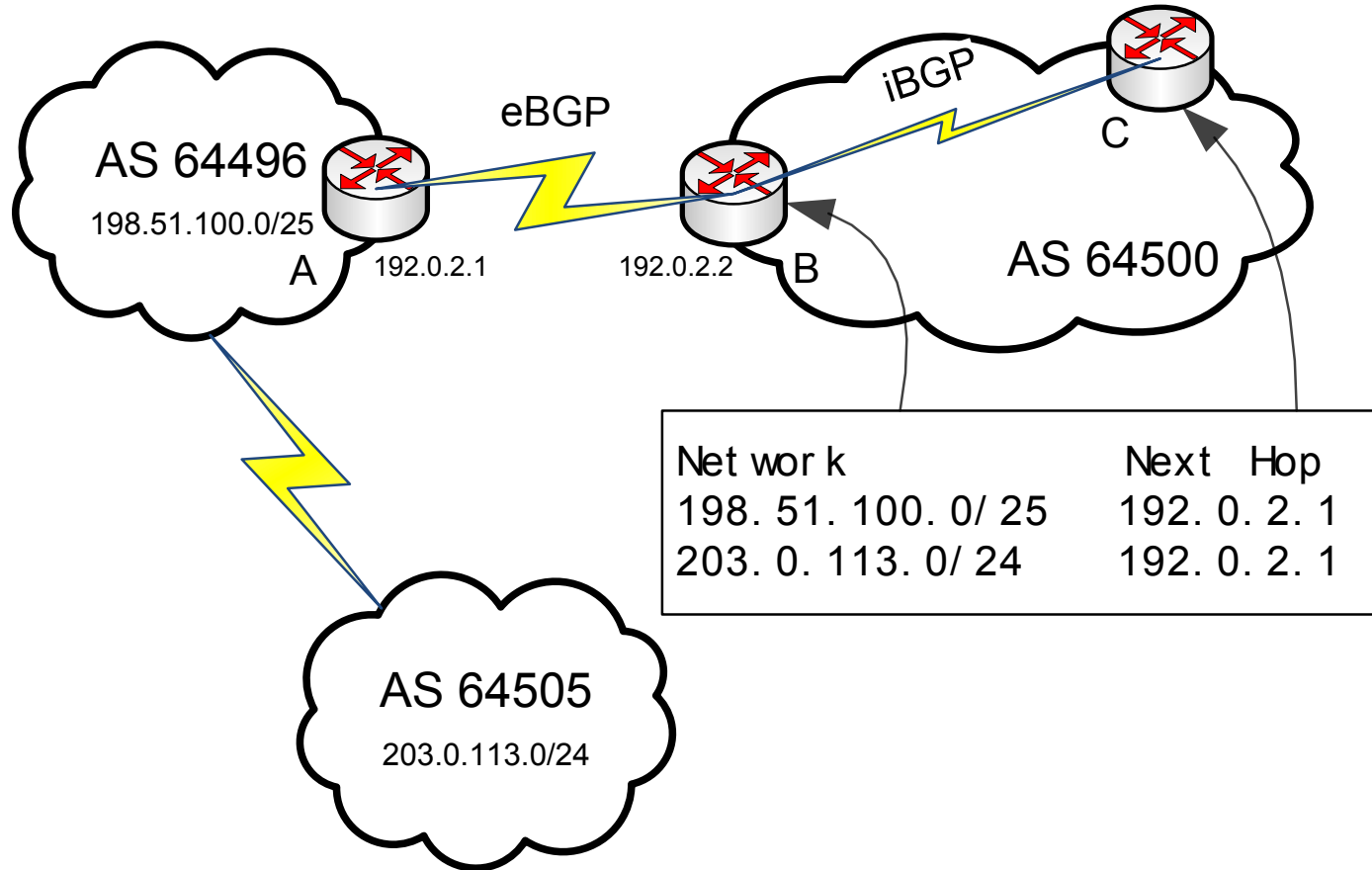
Network	Next Hop	Metric	LocPrf	Weight	Path
*> 192.0.2.0/26	198.51.100.1			0	64496 65511 i
*> 192.0.2.128/26	198.51.100.1			0	64496 64506 i
*>i	198.51.100.22	0	0		64506 e
* 203.0.113.128/25	198.51.100.1			0	64496 64501 i
*>	198.51.100.114	0	0		64501 i
*> 203.0.113.0/25	198.51.100.1			0	64496 ?

Atributos de rutas: next-hop

- En medios multiacceso (Ethernet, FR, ATM), el NH es la IP de la interfaz del router que originó la ruta
 - Ejemplo: IXP



Atributos de rutas: next-hop



Atributos de rutas: AS-Path

- Secuencia de ASN que se deben atravesar para llegar al AS destino.
- Un AS Path en blanco, significa que la ruta fue originada en el AS local.

BGP table version is 134358, local router ID is 198.51.100.1

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, S Stale

Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 192.0.2.0/26	198.51.100.1			0	64496 65511 i
*> 192.0.2.128/26	198.51.100.1			0	64496 64506 i
*>i	198.51.100.22	0	0		64506 e
* 203.0.113.128/25	198.51.100.1			0	64496 64501 i
*>	198.51.100.114	0	0		64501 i
*> 203.0.113.0/25	198.51.100.1			0	64496 ?

Atributos de rutas: AS-Path

- Puede ser utilizado para detectar loops.

BGP table version is 134358, local router ID is 198.51.100.1

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, S Stale

Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	192.0.2.0/26	198.51.100.1	95		0	64496 65511 i
*>	192.0.2.128/26	198.51.100.1	95		0	64496 64506 i
*>i		198.51.100.22	95		0	64506 e
*	203.0.113.128/25	198.51.100.1	95		0	64496 64501 i
*>		198.51.100.114		100	0	64501 i
*>	203.0.113.0/25	198.51.100.1	0	100	0	64496 64497 64498 64499 64500 64503 64540 64540 65540 65540 655440 65540 65540 65540 65540 i

- El primer AS (de izquierda a derecha) muestra el AS adyacente desde donde se aprendió la ruta (el AS local no se muestra). El resto, representa la secuencia de ASN que se debe atravesar para llegar a destino. El último ASN es el AS origen de la ruta.

Insertando prefijos en BGP

Comando “Network”

- Ejemplo:

```
router bgp 64496  
  network 203.0.113.128 mask 255.255.255.128
```

- Importante: Una ruta coincidente debe existir en la tabla de ruteo para que la red puede ser anunciada (aprendida por otro protocolo de ruteo, ruta estática, directamente conectada, etc)

Insertando redes en BGP

- Ruta “pull up”, ejemplo

```
ip route 203.0.113.128 255.255.255.128 null0
router bgp 64496
  network 203.0.113.128 mask 255.255.255.128
```

- Se inserta la ruta en la tabla de ruteo
- La ruta estática a “null0” es llamada ruta “pull up”.
 - sólo se usará si no hay una ruta más específica en la tabla de ruteo.
- Forma práctica de generar una ruta agregada.

Insertando prefijos en BGP

- ...mucho cuidado con redistribute!
 - `redistribute <routing-protocol>`
 - Significa que todas las rutas del protocolo *<routing-protocol>* serán transferidas dentro del protocolo de ruteo vigente.
 - Es altamente recomendable evitarlo.
- Se pierde el control de lo anunciado.

Atributos de rutas: MED

- MED: Multi-Exit Discriminator
- MED solo es enviado a vecinos eBGP, para preferir puertas de entrada a nuestro AS
- Los path con el valor MED más bajo son los más preferidos.
- Se utiliza en conexiones multihomed.

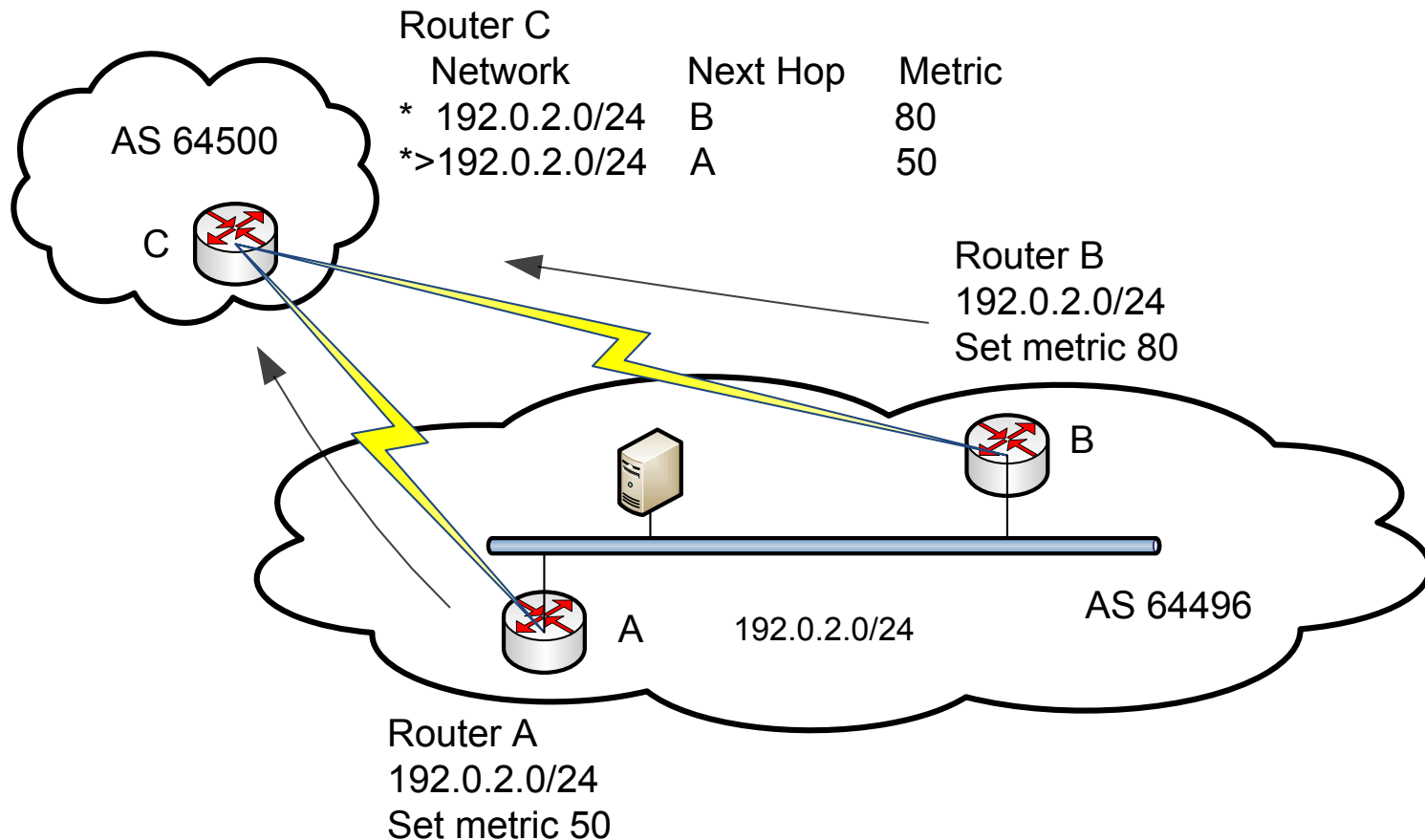
BGP table version is 134358, local router ID is 198.51.100.1

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, S Stale

Origin codes: i - IGP, e - EGP, ? - incomplete

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	192.0.2.0/26	198.51.100.1	95		0	64496 65511 i
*>	192.0.2.128/26	198.51.100.1	95		0	64496 64506 i
*>i		198.51.100.22	95		0	64506 e
*	203.0.113.128/25	198.51.100.1	95		0	64496 64501 i
*>		198.51.100.114		100	0	64501 i
*>	203.0.113.0/25	198.51.100.1	0	100	0	64496 64497 64498 i

Atributos de rutas: MED



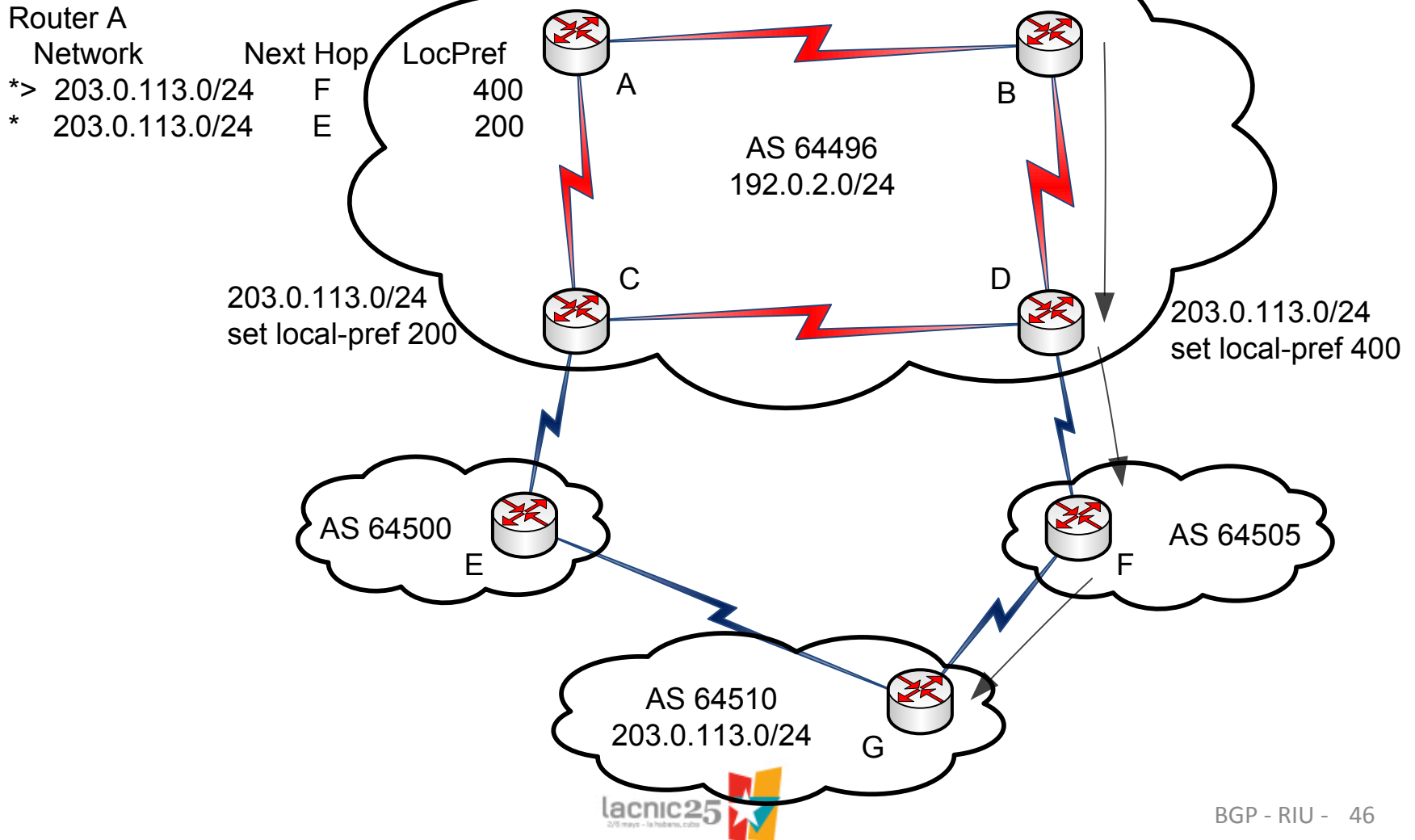
- Los routers A y B anuncian el prefijo 192.0.2.0/24 con MED 50 y 80 a un eBGP vecino (AS 64500).
- Se elegirá el camino por el router A, pues tiene menor MED.

Atributos de rutas: LOCAL_PREF

- Indica un grado de preferencia respecto a otras rutas al mismo destino.
- Es un atributo **local al sistema autónomo** (se propaga por iBGP pero no por eBGP).
- Mayor local preference indica mejor preferencia:
- Se setea mediante route-maps.

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	192.0.2.0/26	198.51.100.1	95		0	64496 65511 i
*>	192.0.2.128/26	198.51.100.1	95		0	64496 64506 i
*>i		198.51.100.22	95		0	64506 e
*	203.0.113.128/25	198.51.100.1	95		0	64496 64501 i
*>		198.51.100.114		100	0	64501 i
*>	203.0.113.0/25	198.51.100.1	0	100	0	64496 64497
	64498 i					

Atributos de rutas: LOCAL_PREF

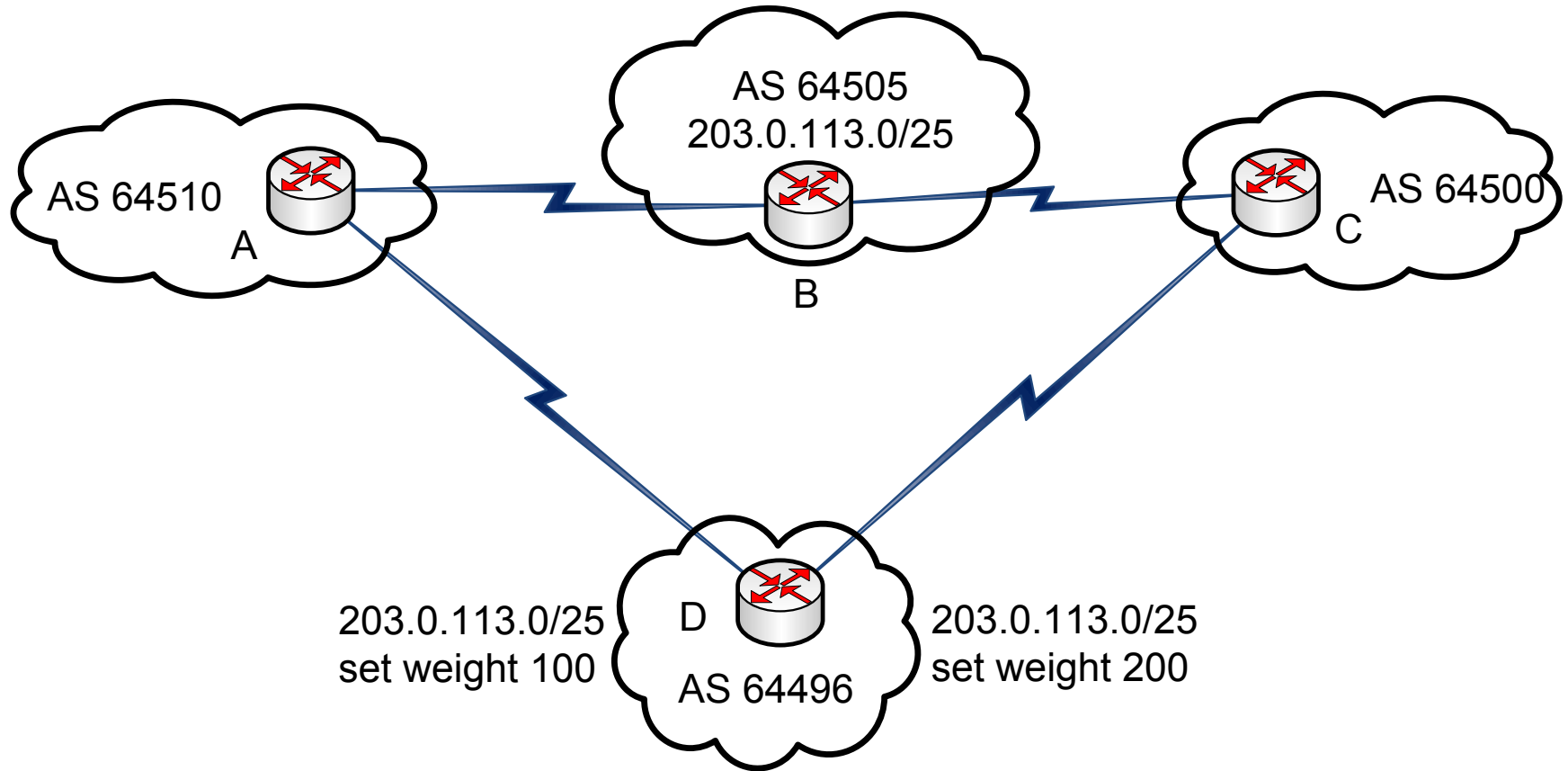


Atributos de rutas: Weight

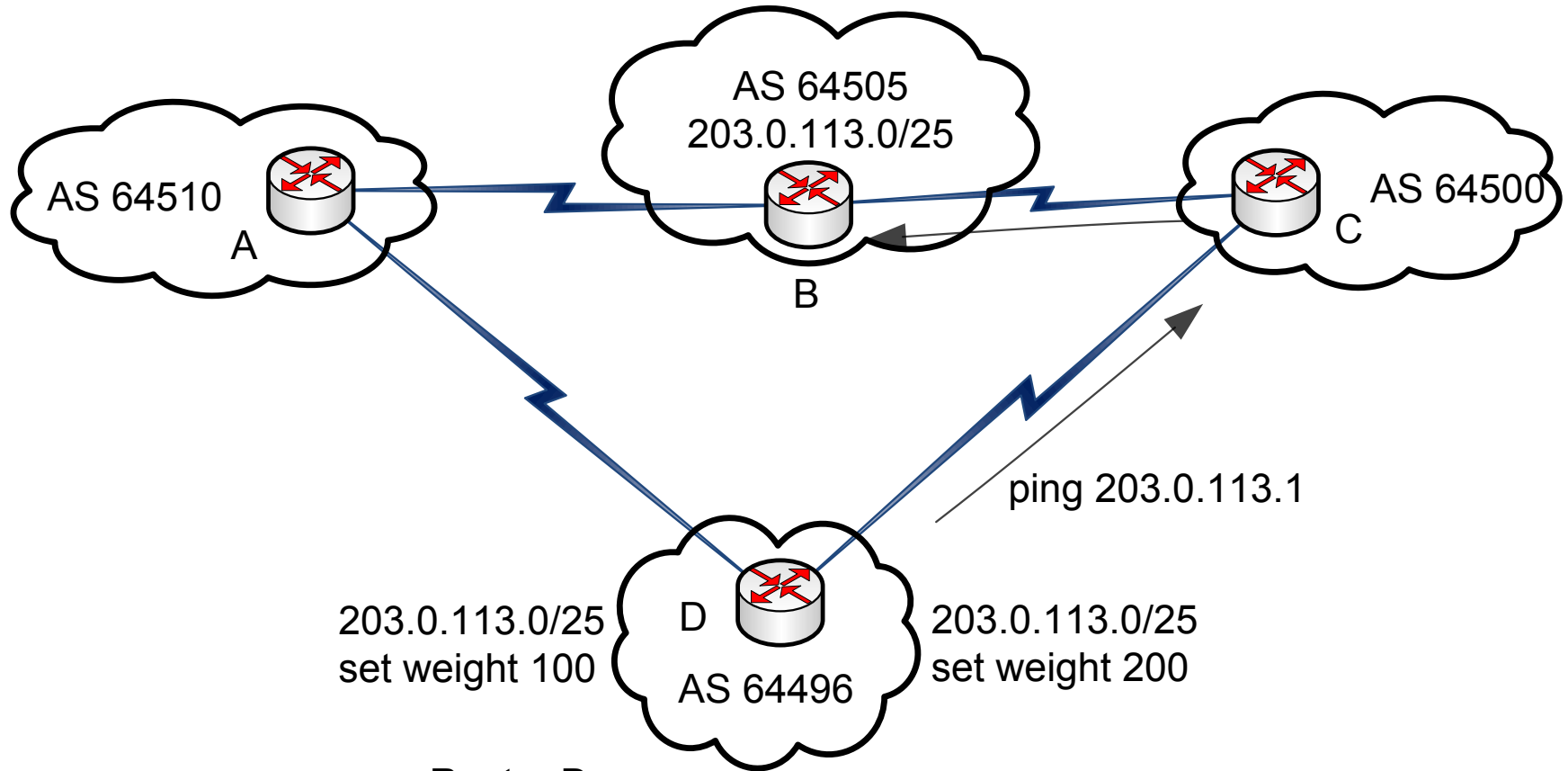
- Es similar a local_pref, sólo que tiene efecto local en el router que se define.
- Tiene precedencia más alta que cualquier otro atributo (algoritmo de selección de camino).
- Sirve para discriminar rutas entre proveedores conectados al mismo router.
- Se setea mediante route-maps.
- Propietario de CISCO.

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	192.0.2.0/26	198.51.100.1	95		300	64496 65511 i
*>	192.0.2.128/26	198.51.100.1	95		0	64496 64506 i
*>i		198.51.100.22	95		100	64506 e
*	203.0.113.128/25	198.51.100.1	95		0	64496 64501 i
*>		198.51.100.114		100	0	64501 i
*>	203.0.113.0/25	198.51.100.1	0	100	0	64496 64497 64498 i

Atributos de rutas: Weight



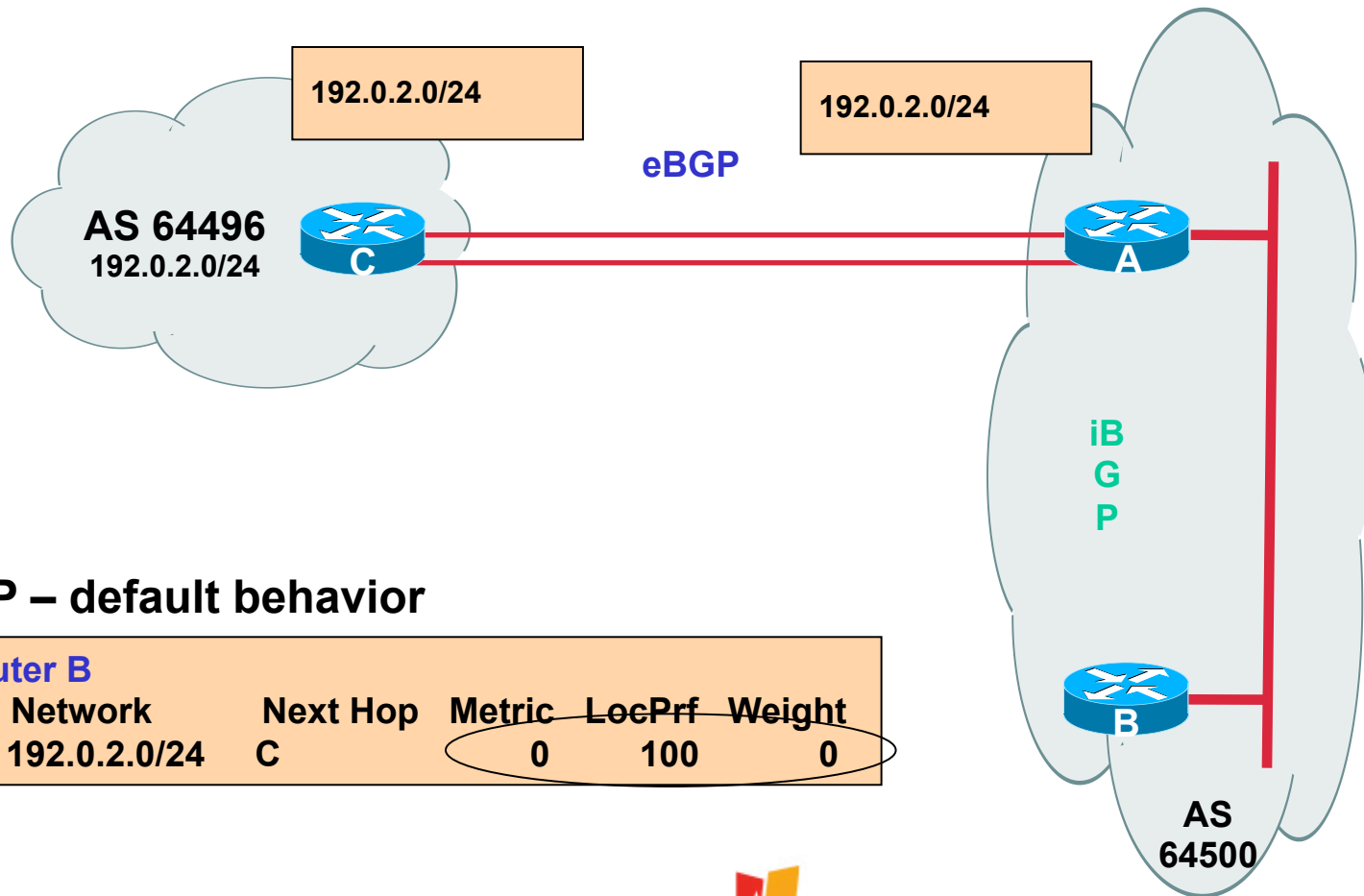
Atributos de rutas: Weight



Router D

	Network	Next Hop	Weight
*	203.0.113.0/25	A	100
*>	203.0.113.0/25	C	200

Local-pref, MED y Weight

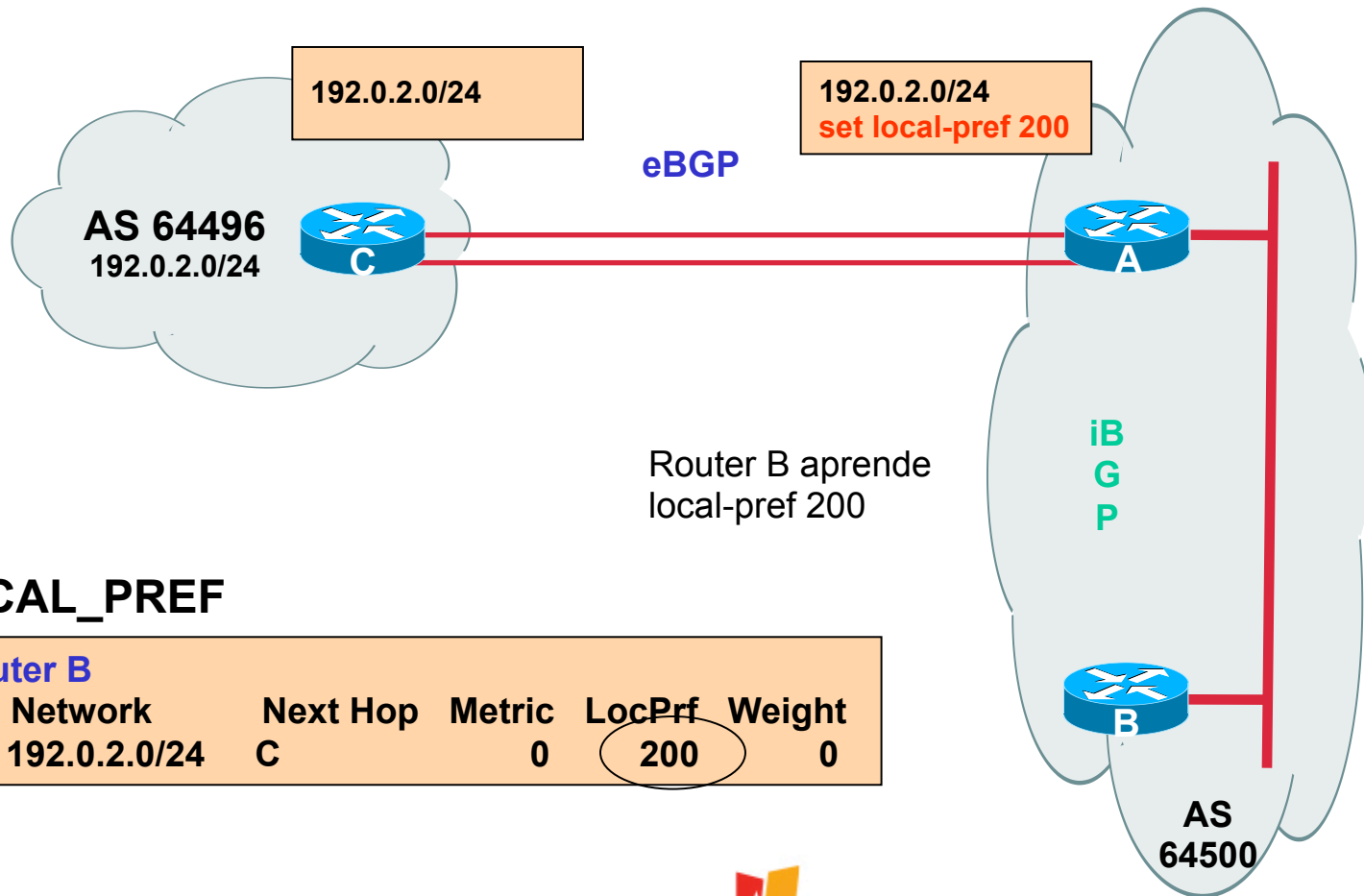


BGP – default behavior

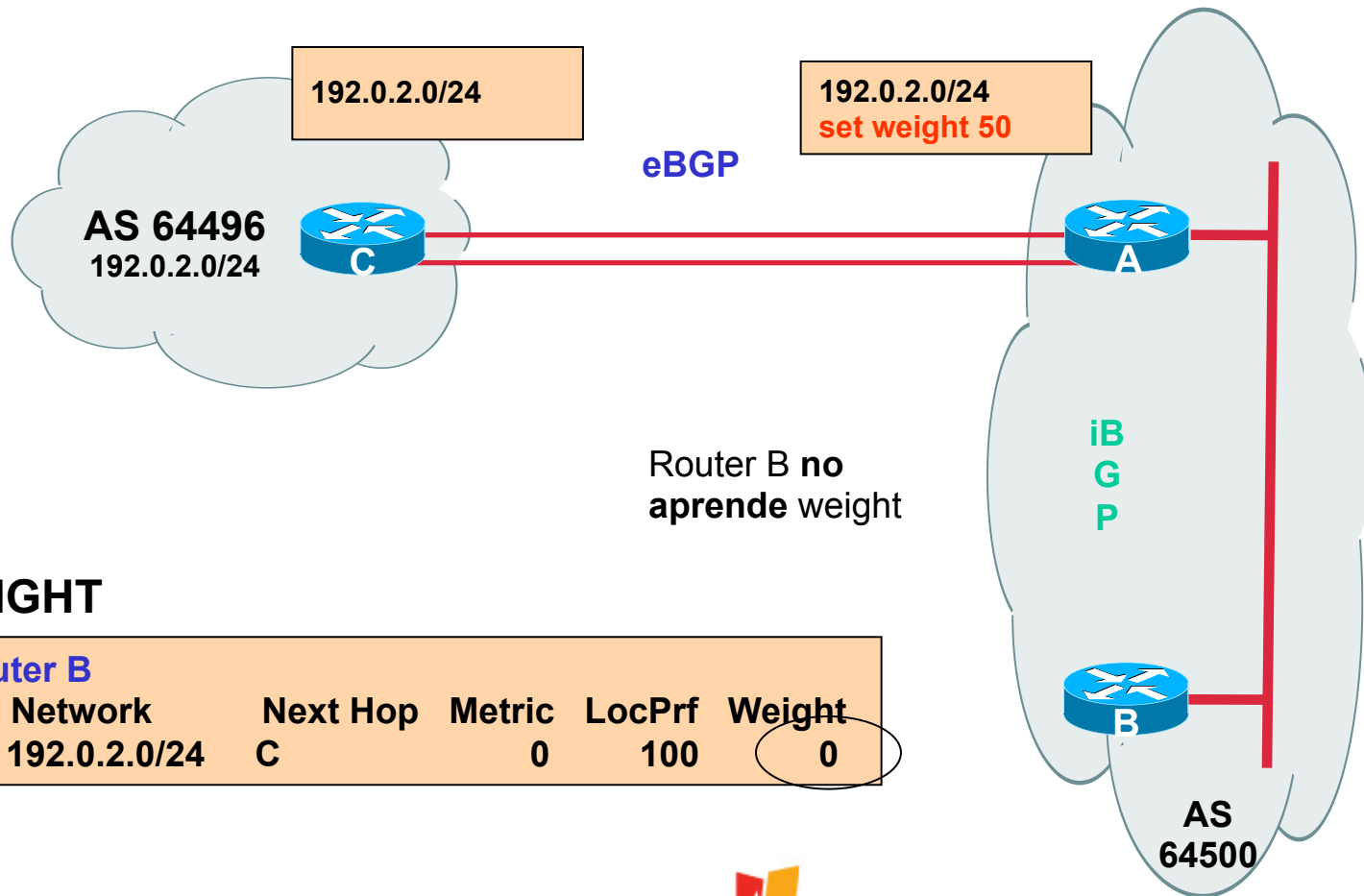
Router B

	Network	Next Hop	Metric	LocPrf	Weight
*>i	192.0.2.0/24	C	0	100	0

Local-pref, MED y Weight



Local-pref, MED y Weight

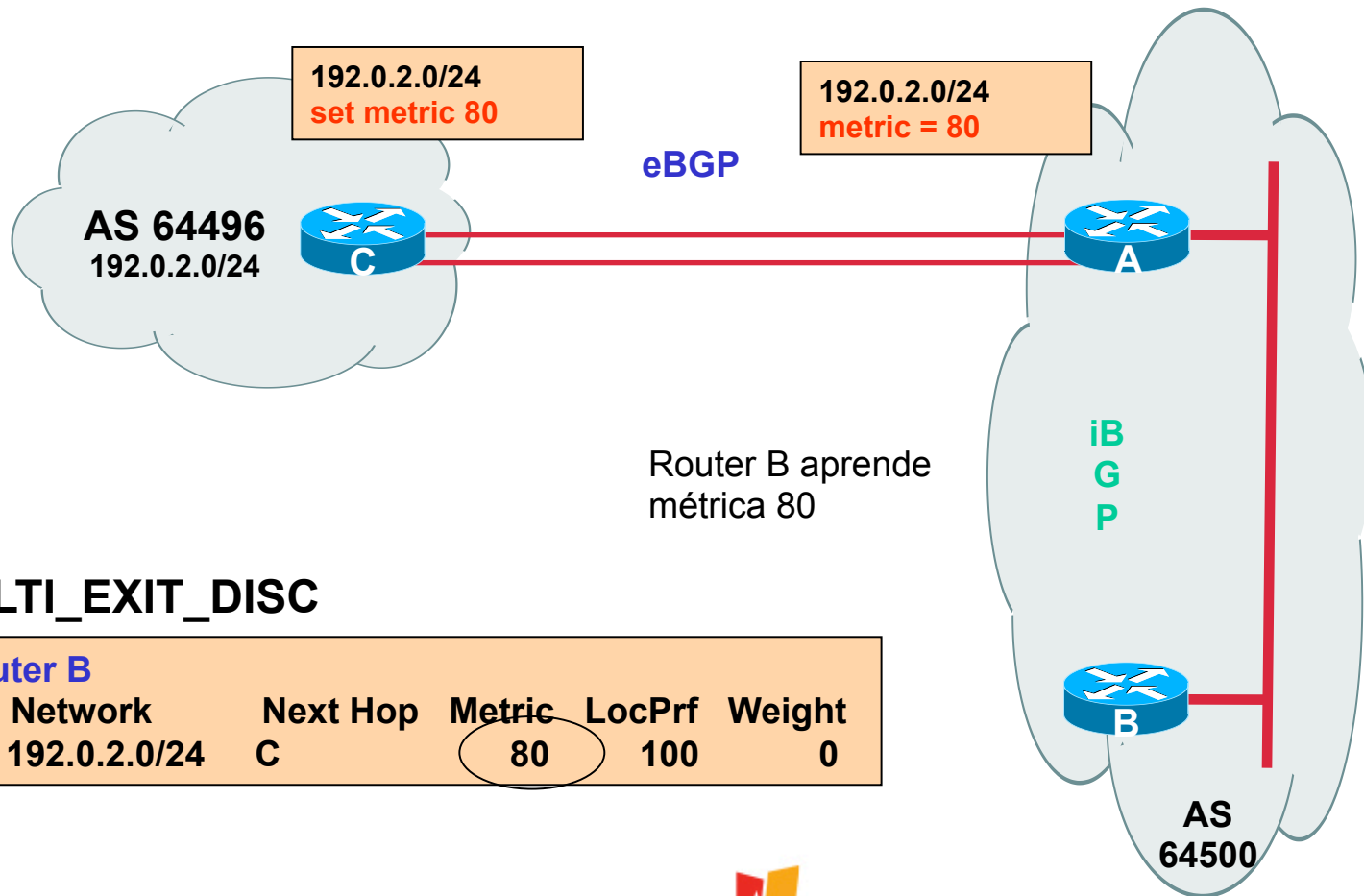


WEIGHT

Router B

	Network	Next Hop	Metric	LocPrf	Weight
*>i	192.0.2.0/24	C	0	100	0

Local-pref, MED y Weight



MULTI_EXIT_DISC

Router B

	Network	Next Hop	Metric	LocPrf	Weight
*>i	192.0.2.0/24	C	80	100	0

Selección del mejor camino

- 1.Si el next-hop es inaccesible, descartar la ruta.
- 2.Si el camino es interno (iBGP), synchronization está habilitado y la ruta no está en el IGP, descartarla.
- 3.Se prefiere el camino con mayor peso “weight” (propietario Cisco).
- 4.Luego, se prefiere la ruta de mayor “local preference”.

Selección del mejor camino (Cont)

5. En caso del mismo local-pref, se prefiere una ruta que es originada por el router (comando network o redistribución).
6. Si la ruta no fue originada por el router y local-pref es igual, se prefiere la ruta con el path de sistemas autónomos más corto (shortest as-path).
7. Si todo es igual, se prefiere el menor código de “origen” (IGP<EGP<Incomplete).

Selección del mejor camino (Cont)

8. A igualdad de “origen”, se prefiere el menor valor de MED. Esta comparación se realiza sólo si los neighbors de los que se aprendieron la ruta pertenecen todos a un mismo AS (a menos que se especifique “bgp always-compare-med”).
9. Se prefieren rutas aprendidas por eBGP que por iBGP.
10. Se prefiere la ruta cuyo next-hop tiene menor métrica en el IGP.
11. Si hasta aquí no hay decisión, se prefiere la ruta correspondiente al neighbor de menor router-id.

Puesta a punto de BGP

BGP vs IGP's

Protocolos de ruteo interno

- Ejemplos OSPF, RIP, etc..
- Usados para transportar las direcciones de la **infraestructura**.
- No usados para transportar los prefijos de Internet o clientes.
- Diseñados de tal forma que se minimice el número de prefijos.

Puesta a punto de BGP (Cont)

BGP vs IGP's

- iBGP: usado para transportar:
 - Los prefijos de Internet a través del Backbone
 - Los prefijos de los clientes
 - Utilizar /32 para las interfaces loopbacks de enrutamiento y levantar las sesiones de iBGP con ellas
 - Usar Peer Groups (No es contenido de este tutorial)
 - Usar Passwords en la sesiones de iBGP (robos de prefijos)

Puesta a punto de BGP (Cont)

BGP vs IGP's

- IGP
 - El IGP debe contener a las redes de la DMZ
- eBGP usados para
 - Intercambiar prefijos con otros AS
 - Implementar políticas de ruteo

Puesta a punto de BGP (Cont)

BGP

- Nunca hacer
 - Redistribuir prefijos BGP en un IGP
 - Redistribuir rutas IGP dentro de BGP
 - NO Usar IGP para transportar los prefijos de los clientes u otras redes externas

Puesta a punto de BGP (Cont)

BGP

- Auto-Sumarización

- Si este comando se activa en BGP, cambia el comportamiento del comando network: el comando network omite la máscara.

```
router bgp 64496  
no auto-summary
```

Puesta a punto de BGP (Cont)

BGP

- Que anuncios no debería recibir
 - No recibir los prefijos definidos en el RFC1918
 - No aceptar mis propios prefijos
 - No aceptar el default (a menos que se requiera)
 - No aceptar prefijos mayores de /24

Puesta a punto de BGP (Cont)

BGP - Sesiones estandar sin loopback

Consideremos un router que tiene 2 caminos para conectarse con un vecino BGP. Sus interfaces con numeración IP son:

- Router1: Serial0, 192.0.2.1 /24, Serial2, 198.51.100.1/24, loopback0, 203.0.113.1 /32
- Router2: Serial0, 192.0.2.2/24, Serial2 198.51.100.2/24, loopback0, 203.0113.2/32

El router 1 puede configurarse así:

Puesta a punto de BGP (Cont)

BGP - Sesiones estandar sin loopback

```
router bgp 64496  
  neighbor 192.0.2.2 remote-as 64505
```

- En este caso, BGP usará automáticamente 192.0.2.1 como la IP fuente para realizar la conexión TCP. Sin embargo, si la interfaz serial0 es puesta en shutdown (administrativamente o no), la conexión BGP también, aunque el Router 2 esté funcionando.

Puesta a punto de BGP (Cont)

BGP - Sesiones con loopback

- Los neighbor no necesariamente deben estar directamente conectados y pueden utilizar las interfaces loopback, en lugar de la interfaz física, la configuración sería:

Router1:

```
router bgp 64496
neighbor 192.0.2.2 remote-as 64505
neighbor 192.0.2.2 update-source loopback0
```

Router2:

```
router bgp 64505
neighbor 192.0.2.1 remote-as 64496
neighbor 192.0.2.1 update-source loopback0
```

Puesta a punto de BGP (Cont)

BGP - Sesiones con loopback

Router1:

```
router bgp 64496
neighbor 192.0.2.2 remote-as 64505
neighbor 192.0.2.2 update-source loopback0
```

Router2:

```
router bgp 64505
neighbor 192.0.2.1 remote-as 64496
neighbor 192.0.2.1 update-source loopback0
```

En este caso, perdiendo una de las conexiones físicas no implica que el BGP se caerá. Mientras cada router tenga una ruta válida para alcanzar la otra interfaz loopback, la conexión BGP estará funcionando. Y aun mejor, no tendremos un solo punto de falla.

Resumen..

- BGP es un protocolo de ruteo **externo**
- Trabaja mediante el anuncio y aprendizaje de rutas
- Establece “sesiones BGP” con “neighbors” del mismo AS (iBGP) o de otros ASs (eBGP)
- Como otros protocolos de ruteo, tiene su propia tabla de rutas (Tabla BGP)
- Mediante un algoritmo de decision, la “mejor” ruta es la *candidata* a pasar a la tabla de ruteo
- El Next Hop me indica el próximo salto para alcanzar un destino.
- Los atributos de las rutas son los que determinan la elección de la mejor ruta
- Existen recomendaciones prácticas para lograr una administración del ruteo mas controlada y eficiente.

Preguntas? Questions?