

Proveedor de Servicio Multihoming



Talleres ISP

Proveedor de servicio Multihoming

- Ejemplos anteriores tratados con carga compartida en el tráfico entrante
 - La principal preocupación en el borde de internet
 - ¿Qué pasa con el tráfico de salida?
- Los ISPs de tránsito se esfuerzan por balancear los flujos de tráfico en ambas direcciones
 - Utilización de balance del enlace
 - Trate de mantener la mayor parte de flujos del tráfico de manera simétrica
 - Algunos ISPs de borde tratan de hacer eso también
- La original “Ingeniería de tráfico”

Proveedor de servicio Multihoming

- Balanceo de tráfico saliente requiere de información de enrutamiento entrante **La solución común es "tabla de enrutamiento completa"**
 - Raramente necesaria
 - ¿Por qué usar el "mazo de enrutamiento" para tratar de solucionar problemas de carga compartida?
 - "Mantengalo simple" es a menudo más fácil (y \$\$\$ barato) que llevar N-copias de la tabla de enrutamiento completa

¡MITOS del proveedor de servicio Multihoming!

Mitos comunes

1. **Usted necesita la tabla completa de enrutamiento para hacer multihome**
 - A la gente que vende memoria para el enrutador le gustaría que usted crea esto
 - Verdadero solamente si eres un proveedor de tránsito
 - La tabla completa de enrutamiento puede ser un obstáculo significativo para multihoming
2. **Usted necesita un GRAN enrutador para multihome**
 - El tamaño del enrutador está relacionado con las tasas de datos, no con la ejecución de BGP
 - En realidad, para hacer multihome, su enrutador necesita:
 - Tener dos interfaces,
 - Ser capaz de hablar BGP con al menos dos peers,
 - Ser capaz de manejar atributos BGP,
 - Manejar al menos un prefijo
3. **BGP es complicado**
 - En las manos equivocadas, isí puede ser! ¡Mantengalo Simple!

Proveedor de servicios Multihoming: algunas estrategias

- Tome los prefijos que necesita para ayudar a la ingeniería de tráfico
 - Mire los datos NetFlow para sitios populares
- Los prefijos originados por sus vecinos inmediatos y los vecinos de ellos van a hacer más para ayudar al balanceo de carga que los prefijos de ASNs a muchos saltos de distancia
 - Concentrarse en destinos locales
- Utilice enrutamiento por defecto tanto como sea posible
 - O use la tabla de enrutamiento completa con cuidado

Proveedor de servicio Multihoming

- Ejemplos
 - Una conexión de subida, un peer local
 - Una conexión de subida, punto de intercambio local
 - Dos conexiones de subida, un peer local
 - Tres conexiones de subida, anchos de banda de enlace desiguales
- Requiere BGP y un ASN público
- Los ejemplos asumen que la red local tiene su propio bloque de direcciones /19

Proveedor de servicio Multihoming

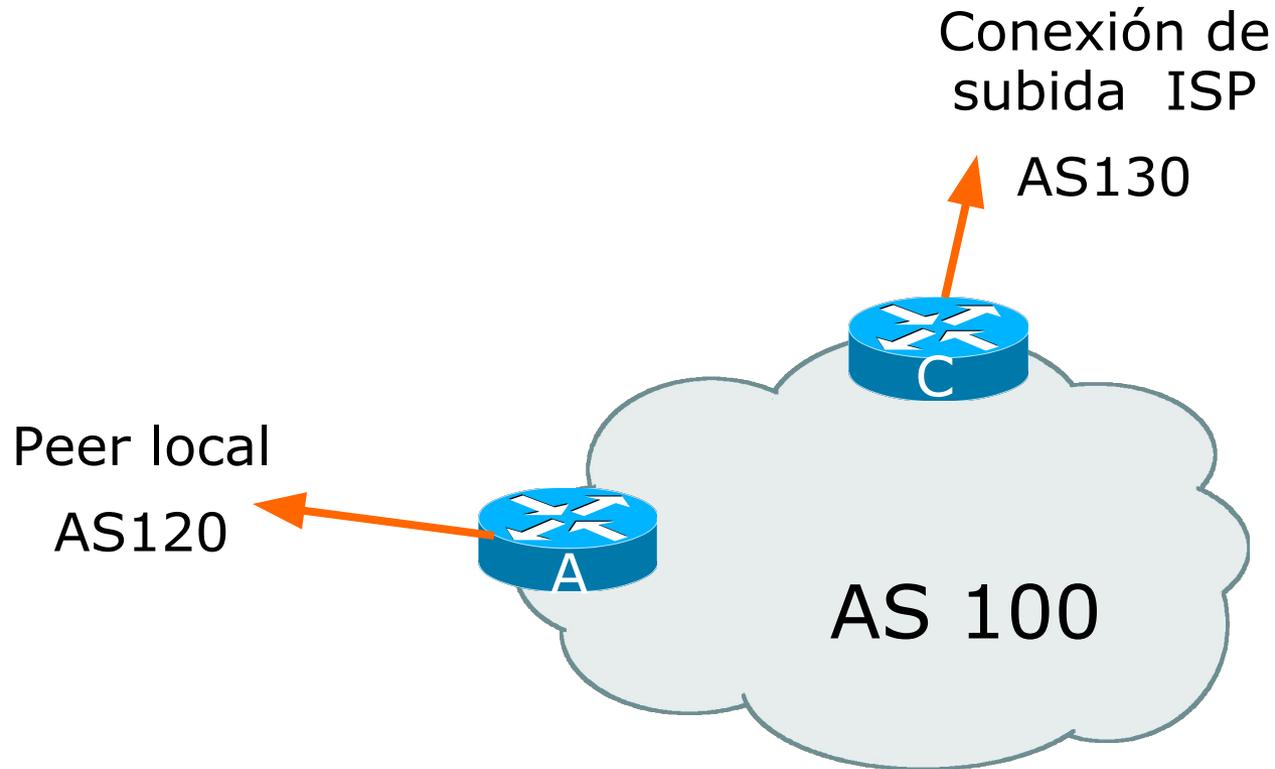


Una conexión de subida, un
peer local

Una conexión de subida, un peer local

- La situación más común en muchas regiones del internet
- Conectar al proveedor de tránsito de la conexión de subida para ver el “Internet”
- Conectarse a la competencia local de manera que el tráfico local permanece local
- Se ahorran gastos \$ sobre el tránsito de la conexión de subida para los costos del tráfico local

Una conexión de subida, Un peer local



Una conexión de subida, un peer local

- Anunciar el /19 agregado en cada enlace
- Aceptar ruta por defecto solo desde la conexión de subida
 - Ya sea 0.0.0.0/0 o una red que puede ser utilizada por defecto
- Aceptar todas las rutas originadas en el peer local

Una conexión de subida, un peer local

- Configuración Router A

Filtro de
prefijos de
entrada

```
router bgp 100
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.2 remote-as 120
  neighbor 122.102.10.2 prefix-list my-block out
  neighbor 122.102.10.2 prefix-list AS120-peer in
!
ip prefix-list AS120-peer permit 122.5.16.0/19
ip prefix-list AS120-peer permit 121.240.0.0/20
!
ip prefix-list my-block permit 121.10.0.0/19
!
ip route 121.10.0.0 255.255.224.0 null0 250
```

Una conexión de subida, un peer local

- Configuración alternativa - Router A

```
router bgp 100
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.2 remote-as 120
  neighbor 122.102.10.2 prefix-list my-block out
  neighbor 122.102.10.2 filter-list 10 in
!
ip as-path access-list 10 permit ^(120_)+$
!
ip prefix-list my-block permit 121.10.0.0/19
!
ip route 121.10.0.0 255.255.224.0 null0
```

Filtros AS Path –
más “confiables”

Una conexión de subida, un peer local

- Configuración Router C

```
router bgp 100
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.1 remote-as 130
  neighbor 122.102.10.1 prefix-list default in
  neighbor 122.102.10.1 prefix-list my-block out
!
ip prefix-list my-block permit 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
ip route 121.10.0.0 255.255.224.0 null0
```

Una conexión de subida, un peer local

- Dos configuraciones posibles para Router A
 - Filter-lists asume que el peer sabe lo que ellos están haciendo
 - Prefix-list es de alto mantenimiento, pero más seguro
 - Algunos ISPs usan **ambos**
- El tráfico local va desde y hacia el peer local, todo lo demás va a la conexión de subida

Apartado:

Recomendaciones de configuración

- Peers privados
 - Los peering ISPs intercambian prefijos que ellos originan
 - Algunas veces intercambian prefijos de ASNs vecinos también
- Tenga en cuenta que el enrutador eBGP de peer privado debe llevar sólo los prefijos que desea recibir el peer privado
 - De otra manera podrían apuntar a una ruta por defecto hacia usted y sin intención transitar su troncal

Proveedor de servicio Multihoming

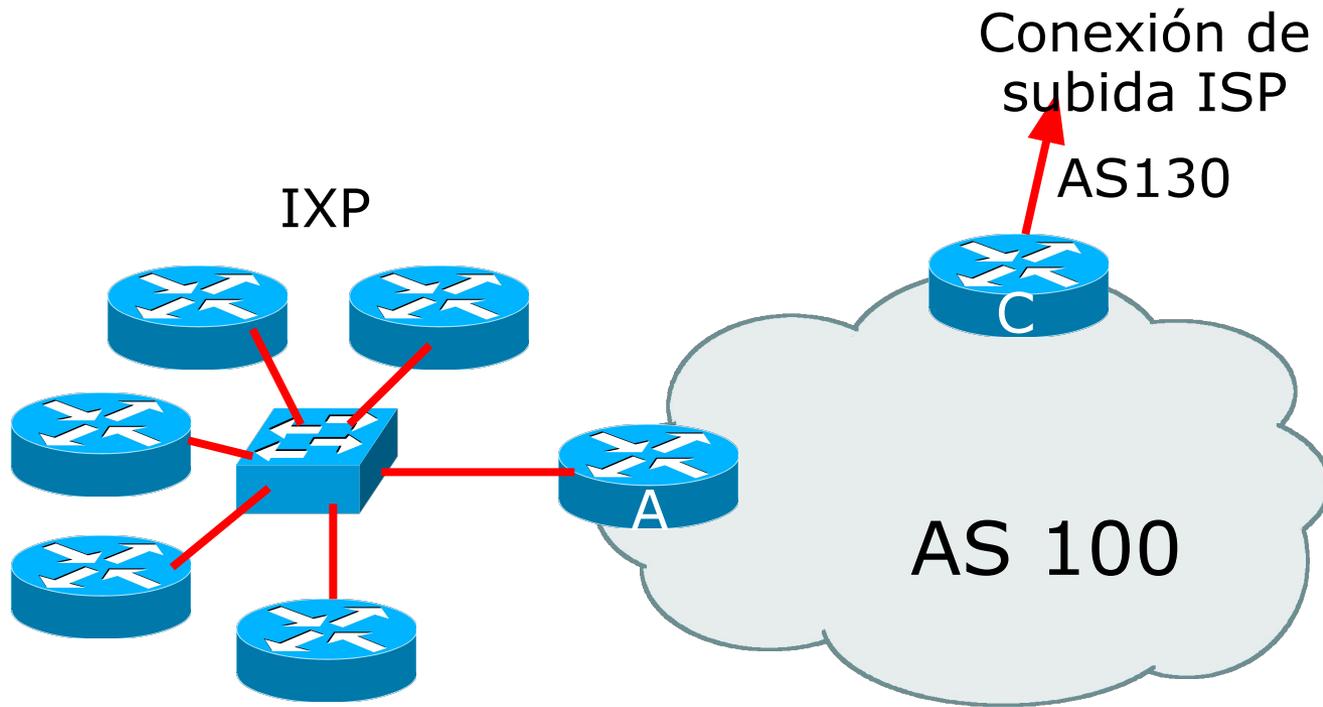


Una conexión de subida, punto
de intercambio local

Una conexión de subida, punto local de intercambio

- Situación muy común en muchas regiones del Internet
- Conectar al proveedor de tránsito de la conexión de subida para ver el “Internet”
- Conectar con el punto de intercambio de internet local de manera que el tráfico local permanezca local
 - Se ahorran gastos \$ sobre el tránsito de la conexión de subida para los costos del tráfico local
- Este ejemplo es una versión a escala del anterior

Una conexión de subida, punto local de intercambio



Una conexión de subida, punto local de intercambio

- Anuncia /19 agregado a cada vecino AS
- Acepta ruta por defecto solo de la conexión de subida
 - Ya sea 0.0.0.0/0 o una red la cual sea usada como por defecto
- Acepta todas las rutas originadas en los peers IXP

Una conexión de subida, punto local de intercambio

- Configuración Router A

```
interface fastethernet 0/0
  description Exchange Point LAN
  ip address 120.5.10.1 mask 255.255.255.224
!
router bgp 100
  neighbor ixp-peers peer-group
  neighbor ixp-peers prefix-list my-block out
  neighbor ixp-peers remove-private-AS
  neighbor ixp-peers send-community
  neighbor ixp-peers route-map set-local-pref in
!
```

...siguiente diapositiva

Una conexión de subida, punto local de intercambio

```
neighbor 120.5.10.2 remote-as 200
neighbor 120.5.10.2 peer-group ixp-peers
neighbor 120.5.10.2 prefix-list peer200 in
neighbor 120.5.10.3 remote-as 201
neighbor 120.5.10.3 peer-group ixp-peers
neighbor 120.5.10.3 prefix-list peer201 in
neighbor 120.5.10.4 remote-as 202
neighbor 120.5.10.4 peer-group ixp-peers
neighbor 120.5.10.4 prefix-list peer202 in
neighbor 120.5.10.5 remote-as 203
neighbor 120.5.10.5 peer-group ixp-peers
neighbor 120.5.10.5 prefix-list peer203 in
```

...siguiente diapositiva

Una conexión de subida, punto local de intercambio

```
!  
ip prefix-list my-block permit 121.10.0.0/19  
ip prefix-list peer200 permit 122.0.0.0/19  
ip prefix-list peer201 permit 122.30.0.0/19  
ip prefix-list peer202 permit 122.12.0.0/19  
ip prefix-list peer203 permit 122.18.128.0/19  
!  
route-map set-local-pref permit 10  
  set local-preference 150  
!
```

Una conexión de subida, intercambio local

- Tenga en cuenta que el Router A no genera el agregado para AS100
 - Si el router A se desconecta del troncal, a continuación el agregado ya no se anuncia al IX
 - La conmutación por error de BGP funciona como se esperaba
- Tenga en cuenta que la entrada route-map, establece la preferencia local más alta que el valor por defecto
 - Este es un recordatorio visual que la mejor trayectoria BGP para el tráfico local será a través del IXP
 - (y permite para casos futuros donde el operador puede necesitar tomar rutas BGP desde su(s) conexión de subida

Una conexión de subida, punto local de intercambio

- Configuración Router C

```
router bgp 100
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.1 remote-as 130
  neighbor 122.102.10.1 prefix-list default in
  neighbor 122.102.10.1 prefix-list my-block out
!
ip prefix-list my-block permit 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
ip route 121.10.0.0 255.255.224.0 null0
```

Una conexión de subida, punto local de intercambio

- Nota de la configuración del Router A
 - Prefix-list es de alto mantenimiento, pero seguro
 - Ninguna generación de agregado AS100
- El tráfico IXP va y viene del IXP local, todo lo demás va a la conexión de subida

Apartado:

Recomendaciones de configuración IXP

- peers IXP
 - El peering en ISPs en el IXP intercambia prefijos que ellos originan
 - Algunas veces ellos intercambian prefijos de vecinos ASN también
- Tenga en cuenta que el enrutador de borde IXP debe llevar solamente los prefijos que quiere que los peers IXP reciban y los destinos que quiere que ellos estén en capacidad de alcanzar
 - De otra manera ellos podrían apuntar a una ruta por defecto hacia usted y sin intención transitar su troncal
- Si el enrutador IXP está en IX, y distante de su troncal
 - No origine su bloque de direcciones en su enrutador

Proveedor de servicio Multihoming

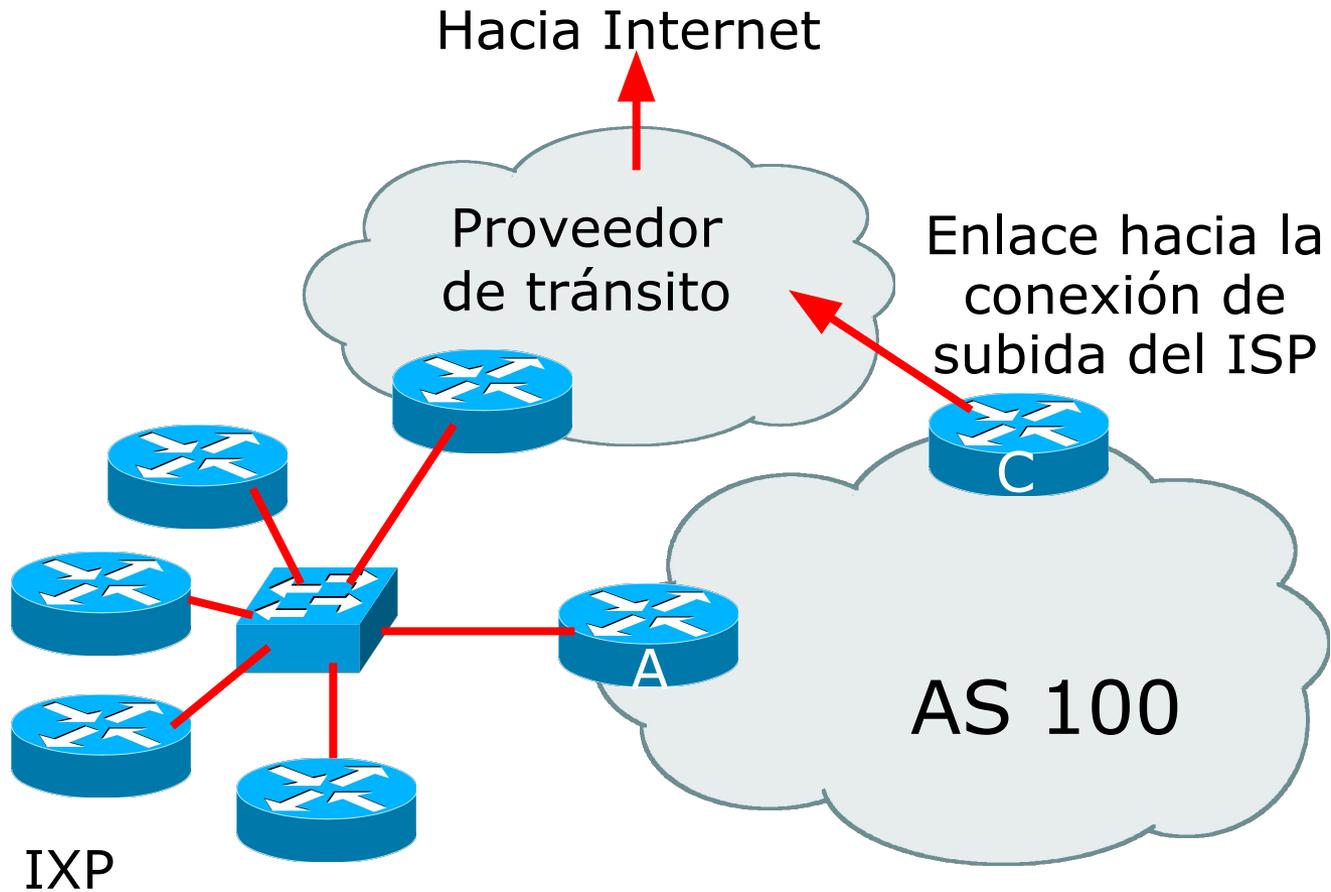


Punto de intercambio local, con
conexión de subida también
siendo peer

Punto local de intercambio, con conexión de subida siendo también un peer

- Situación bastante común
 - Varios ISPs locales proveen acceso al mercado local
 - Uno o dos proveedores de tránsito internacional con licencia
 - Tránsitos con licencia también hacen peer con el IXP
- Cómo asegurar eso:
 - Tráfico de tránsito va al enlace de tránsito
 - Tráfico de peering va al enlace de peering

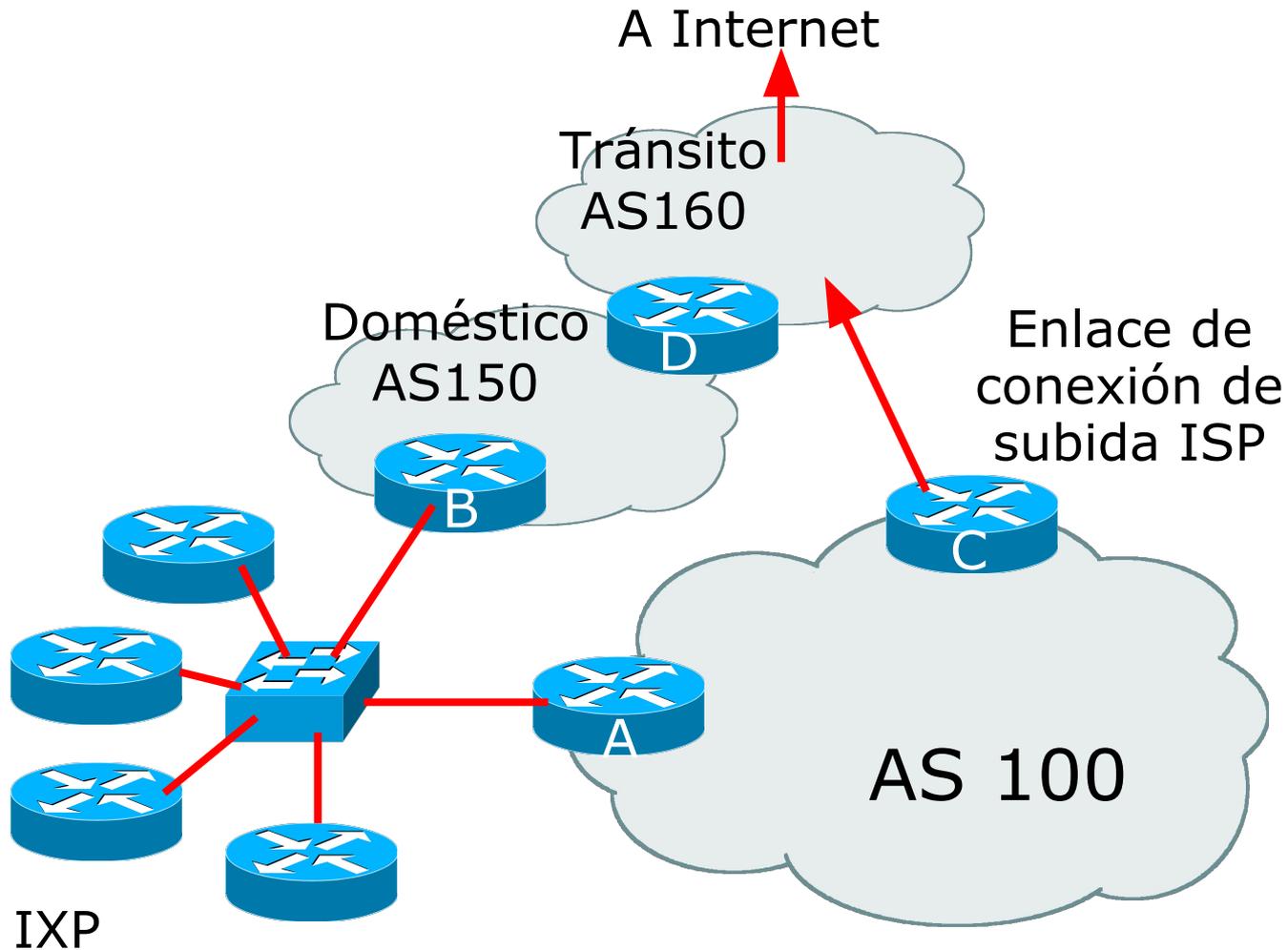
Punto local de intercambio, con conexión de subida siendo también un peer



Punto local de intercambio, con conexión de subida siendo también un peer

- Tráfico de salida desde AS100:
 - Conexión de subida envía tabla completa BGP al AS100
 - Conexión de subida envía rutas domésticas a los peers IXP
 - AS100 utiliza IXP para tráfico doméstico
 - AS100 utiliza el enlace de la conexión de subida para el tráfico internacional
- Tráfico entrante hacia AS100:
 - AS100 envía bloque de direcciones a los peers IXP
 - AS100 envía bloque de direcciones a la conexión de subida
 - La mejor trayectoria desde la conexión de subida al AS100 es preferida vía IXP (ver escenario previo)
- **Problema: ¿Cómo separar el tráfico entrante doméstico e internacional?**

Solución: Separación de AS



Solución: Separación de AS

- El proveedor de tránsito necesita separar su red en doméstica (AS150: rutas locales) y tránsito (AS160: rutas internacionales)
- Clientes de tránsito conectan al AS de tránsito (AS160)
- Peers de AS domésticos (AS150) en el IX, pasan rutas domésticas solamente:
- Tráfico entrante al AS100 ahora:
 - AS100 envía bloque de direcciones al peer IXP (incluyendo AS150)
 - AS100 envía bloque de direcciones a la conexión de subida (AS160)
 - **Importante: Router D no pasa prefijos aprendidos de peers IX hacia AS160**
 - Mejor trayectoria para la conexión de subida al AS100 prefiere la vía del enlace de tránsito

Punto local de intercambio, con conexión de subida siendo también un peer

- Proveedores de tránsito, quienes hacen peer con los clientes en un IX para rutas locales necesitan dividir sus ASNs en dos:
 - Un AS para rutas negocios/domésticas
 - Un AS para rutas de tránsito internacional
- Dos ASNs son justificables para los RIRs porque los dos ASNs tienen políticas de enrutamiento completamente diferentes
 - Peers AS domésticos en IXP
 - AS de tránsito conecta a clientes de tránsito y conexiones de subida
- Esta solución es mucho más fácil de implementar que otras soluciones tales como política de enrutamiento de direcciones fuente complejas

Proveedor de servicio Multihoming

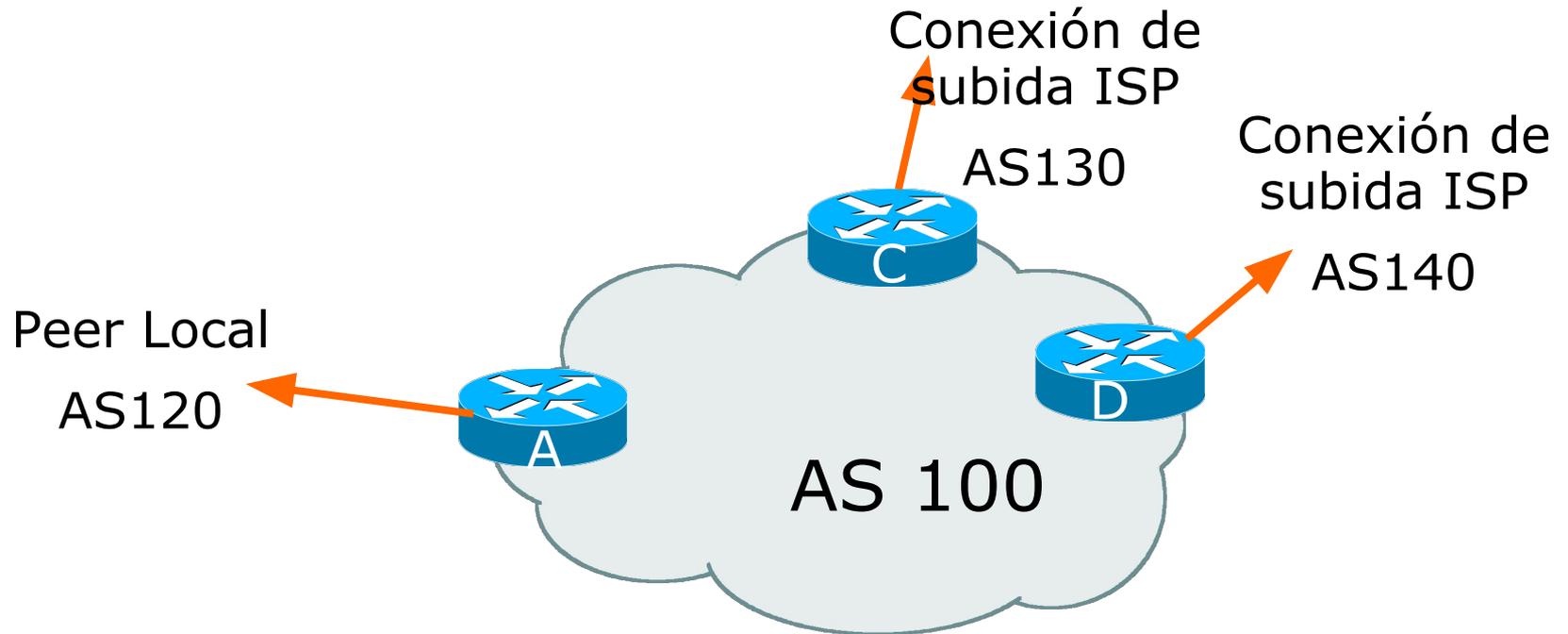


Dos conexiones de subida, un
peer local

Dos conexiones de subida, un peer local

- Conecta a ambos proveedores de tránsito a conexiones de subida para ver “Internet”
 - Provee redundancia externa y diversidad - la razón de multihome
- Conecta al peer local de manera que el tráfico local permanece local
 - Se ahorran gastos \$ sobre el tránsito de la conexión de subida para los costos del tráfico local

Dos conexiones de subida, un peer local



Dos conexiones de subida, un peer local

- Anuncia /19 agregado en cada enlace
- Acepta ruta por defecto solamente para conexiones de subida
 - Ya sea 0.0.0.0/0 o una red que pueda ser usada como por defecto
- Acepta todas las rutas originadas por el peer local
- Note la separación de Router C y D
 - Router de borde individual significa que no hay redundancia
- Router A
 - Misma configuración de enrutamiento como en el ejemplo de una conexión de subida y un peer local

Dos conexiones de subida, un peer local

- Configuración Router C

```
router bgp 100
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.1 remote-as 130
  neighbor 122.102.10.1 prefix-list default in
  neighbor 122.102.10.1 prefix-list my-block out
!
ip prefix-list my-block permit 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
ip route 121.10.0.0 255.255.224.0 null0
```

Dos conexiones de subida, un peer local

- Configuración Router D

```
router bgp 100
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.5 remote-as 140
  neighbor 122.102.10.5 prefix-list default in
  neighbor 122.102.10.5 prefix-list my-block out
!
ip prefix-list my-block permit 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
ip route 121.10.0.0 255.255.224.0 null0
```

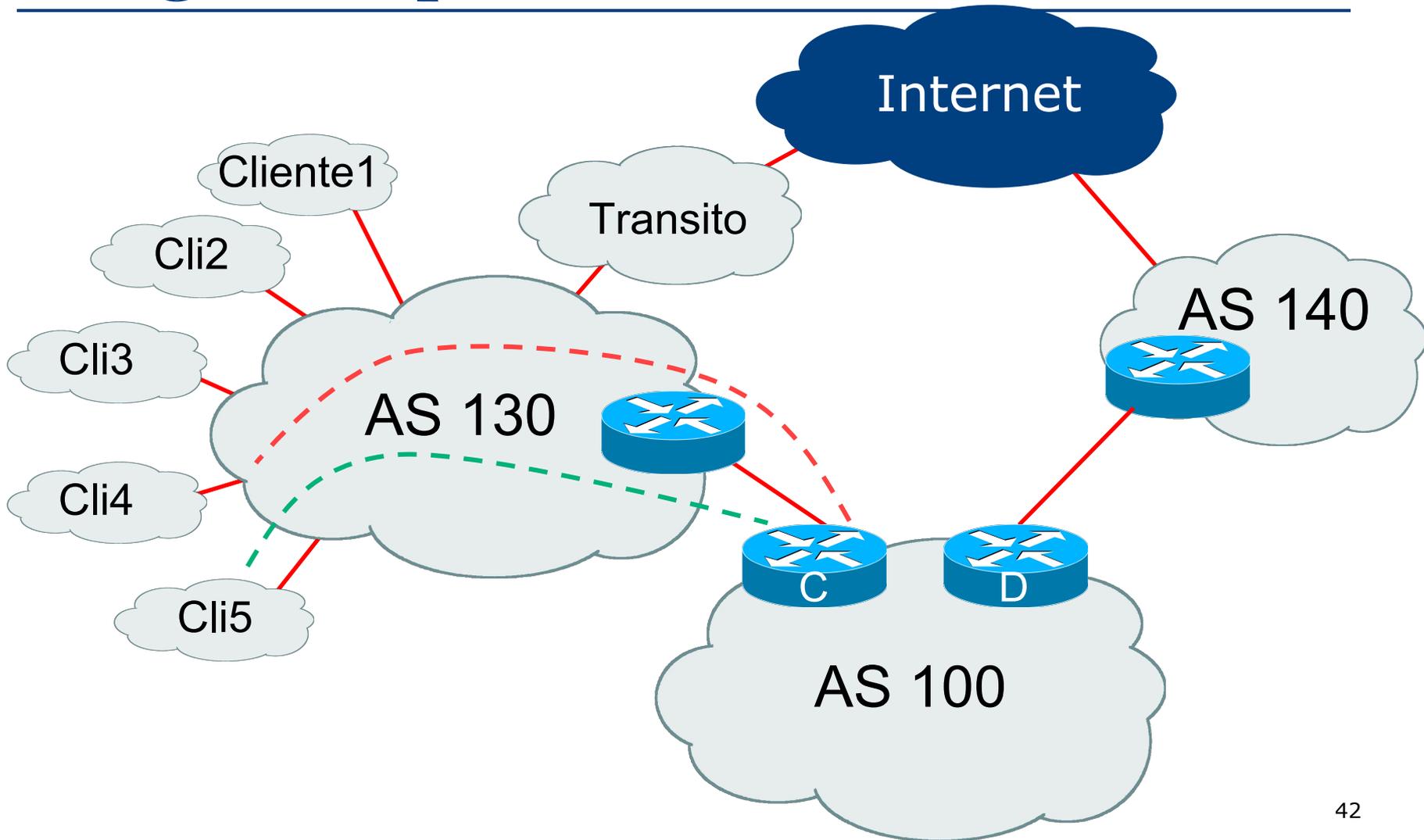
Dos conexiones de subida, un peer local

- Esta es la configuración simple para el Router C y D
- El tráfico de salida hacia las dos conexiones de subida tomará la salida más cercana
 - Requiere enrutadores económicos
 - Esto no es utilizado en la práctica, especialmente en enlaces internacionales
 - Necesita carga compartida para ser mejor

Dos conexiones de subida, un peer local

- Mejores opciones de configuración:
 - Acepta todo el enrutamiento de ambas conexiones de subida
 - **Caro e innecesario**
 - Acepta ruta por defecto de una conexión de subida y algunas rutas de la otra conexión de subida
 - **¡El camino a seguir!**

Carga compartida con diferentes ISPs



Dos conexiones de subida, un peer local con rutas completas

- Configuración Router C

Permite todos los prefijos, menos los bloques RFC6890



```
router bgp 100
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.1 remote-as 130
  neighbor 122.102.10.1 prefix-list rfc6890-deny in
  neighbor 122.102.10.1 prefix-list my-block out
  neighbor 122.102.10.1 route-map AS130-loadshare in
!
ip prefix-list my-block permit 121.10.0.0/19
!
! See http://tools.ietf.org/html/rfc6890
...siguiente diapositiva
```

Dos conexiones de subida, un peer local con rutas completas

```
ip route 121.10.0.0 255.255.224.0 null0
!
ip as-path access-list 10 permit ^(130_)+$
ip as-path access-list 10 permit ^(130_)+_[0-9]+$
!
route-map AS130-loadshare permit 10
  match ip as-path 10
  set local-preference 120
!
route-map AS130-loadshare permit 20
  set local-preference 80
!
```

Dos conexiones de subida, un peer local con rutas completas

- Configuración Router D

```
router bgp 100
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.5 remote-as 140
  neighbor 122.102.10.5 prefix-list rfc6890-deny in
  neighbor 122.102.10.5 prefix-list my-block out
!
ip prefix-list my-block permit 121.10.0.0/19
!
```

! See <http://tools.ietf.org/html/rfc6890>

Permite todos los prefijos, menos los bloques RFC6890

Dos conexiones de subida, un peer local con rutas completas

- Router C configuration:
 - Acepta todas las rutas de AS130
 - Etiqueta prefijos originados por AS130 y los vecinos AS de AS130 con la preferencia local 120
 - El tráfico a esos AS sobrepasará el enlace AS130
 - Prefijos remanentes son etiquetados con la preferencia de 80
 - El tráfico a todos los otros AS sobrepasará el enlace AS140
- La configuración del Router D es la misma que el Router C sin el route-map

Dos conexiones de subida, un peer local con rutas completas

- Rutas completas desde la conexión de subida
 - Resumen de las rutas recibidas:

ASN	Rutas completas	rutas parciales
AS140	500000 @Ip 100	
AS130	30000 @Ip 120 470000 @Ip 80	
Total	1000000	

Dos conexiones de subida, un peer local con rutas completas

- Rutas completas desde la conexión de subida
 - Caro – necesita mucha memoria y CPU
 - Need to play preference games
 - El ejemplo anterior es solamente un ejemplo - ien la vida real necesitará mejorar un buena parametrización (fine-tunning)!
 - El ejemplo previo no considera el tráfico entrante - ver antes en la presentación de ejemplos

Dos conexiones de subida, un peer local de rutas parciales: estrategia

- Preguntar a una conexión de subida por la ruta por defecto
 - Fácil de originarse por defecto hacia un vecino BGP
- Preguntar a otra conexión de subida por tabla de enrutamiento completa
 - entonces filtre esta tabla de enrutamiento basada sobre el vecino ASN
 - Ej: quiere que el tráfico de sus vecinos vayan a través del enlace de ese ASN
 - La mayor parte de lo que se envía por la conexión de subida se tira
 - Más fácil que preguntar a la conexión de subida para establecer filtros personalizados para usted ⁴⁹

Dos conexiones de subida, un peer local de rutas parciales

- Configuración Router C

```
router bgp 100
 network 121.10.0.0 mask 255.255.224.0
 neighbor 122.102.10.1 remote-as 130
 neighbor 122.102.10.1 prefix-list rfc6890-deny in
 neighbor 122.102.10.1 prefix-list my-block out
 neighbor 122.102.10.1 filter-list 10 in
 neighbor 122.102.10.1 route-map tag-default-low in
!
```

Permite todos los prefijos, menos los bloques RFC6890

Lista de filtros AS filtra prefijos basada en el origen ASN

Dos conexiones de subida, un peer local de rutas parciales

```
ip prefix-list my-block permit 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
ip route 121.10.0.0 255.255.224.0 null0
!
ip as-path access-list 10 permit ^(130_)+$
ip as-path access-list 10 permit ^(130_)+_[0-9]+$
!
route-map tag-default-low permit 10
  match ip address prefix-list default
  set local-preference 80
!
route-map tag-default-low permit 20
!
```

Dos conexiones de subida, un peer local de rutas parciales

- Configuración Router D

```
router bgp 100
  network 121.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.5 remote-as 140
  neighbor 122.102.10.5 prefix-list default in
  neighbor 122.102.10.5 prefix-list my-block out
!
ip prefix-list my-block permit 121.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
!
ip route 121.10.0.0 255.255.224.0 null0
```

Dos conexiones de subida, un peer local de rutas parciales

- Configuración Router C:
 - Acepta todas las rutas de AS130
 - (o conseguir que se envíen menos)
 - Filtrar ASNs por lo que solo AS130 y vecinos de AS130 son aceptados
 - Permitir default, y establecer preferencia local a 80
 - El tráfico a esos ASs pasará a través del enlace AS130
 - El tráfico a todos los otros ASs pasará a través del enlace AS140
 - Si el enlace AS140 falla, hay respaldo de conexión por AS130 - y viceversa

Dos conexiones de subida, un peer local de rutas parciales

- Rutas parciales de conexiones de subida
 - Resumen de rutas recibidas:

ASN	Rutas completas		Rutas parciales	
AS140	500000	@lp 100	1	@lp 100
AS130	30000	@lp 120	30000	@lp 100
	470000	@lp 80	1	@lp 80
Total	1000000		30002	

Distribuyendo ruta por defecto con el IGP

- Configuración de Router C IGP

```
router ospf 100
default-information originate metric 30
!
```

- Configuración Router D IGP

```
router ospf 100
default-information originate metric 10
!
```

- Trayectoria primaria es vía Router D, con backup vía Router C
 - preferible a llevar ruta por defecto en iBGP

Dos conexiones de subida, un peer local de rutas parciales

- Rutas parciales de conexiones de subida
 - No caro - solo se llevan las rutas necesarias para la carga compartida
 - Necesita filtrar en AS paths
 - El ejemplo previo es solamente un ejemplo - en la vida real necesita una buena parametrización mejorada!
 - El ejemplo previo no considera el tráfico entrante - ver antes en la presentación de ejemplos

Apartado:

Recomendación de configuración

- Cuando se distribuye default interno por iBGP u OSPF/ISIS
 - Asegúrese que los enrutadores conectan a los peers privados o a los IXPs **NO** lleven la ruta por defecto
 - De otra manera ellos apuntarán una ruta por defecto hacia usted y sin intención transitar su troncal
 - Solución simple para enrutadores Peer/IXP:
`ip route 0.0.0.0 0.0.0.0 null0`

Proveedor de servicio Multihoming



Tres upstreams, anchos de
banda desiguales

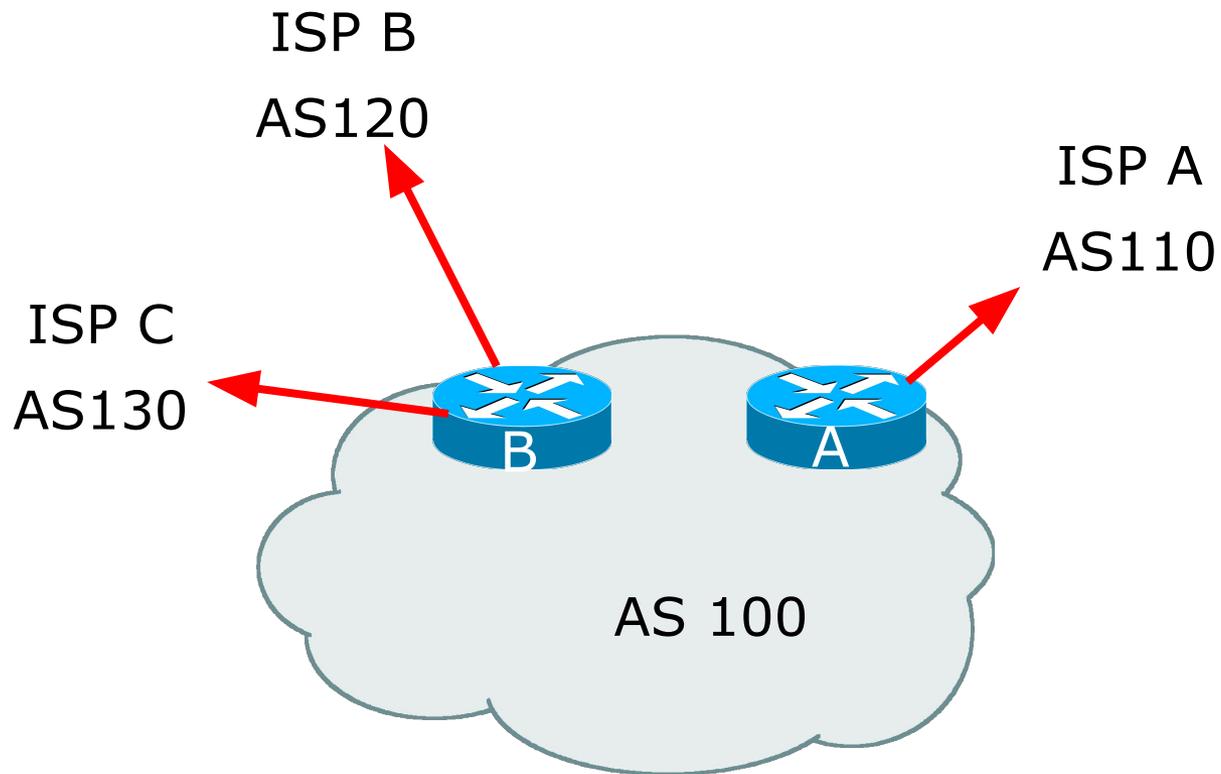
Tres upstreams, anchos de banda desiguales

- El Sistema Autónomo tiene tres conexiones de subida
 - 16Mbps a ISP A
 - 8Mbps a ISP B
 - 4Mbps a ISP C
- ¿Cuál es la estrategia aquí?
 - Una opción es tabla completa por cada uno
 - 3x 533k prefijos \Rightarrow 1600k paths
 - Otra opción es tabla de enrutamiento parcial y rutas por defecto por cada una
 - ¿Cómo?

Estrategia

- Dos enrutadores externos (da la redundancia al enrutador)
 - **NO** necesita tres enrutadores para esto
- Conecte el ancho de banda más grande a un enrutador
 - La mayoría del tráfico entrante y saliente irá por ahí
- Conecte los otros dos enlaces al segundo enrutador
 - Provee máxima capacidad de respaldo si el enlace primario falla
- Utilice el enlace más grande por defecto
 - La mayoría del tráfico entrante y saliente irá por ahí
- Haga la ingeniería de tráfico sobre los dos enlaces más pequeños
 - Enfóquese en las necesidades del tráfico local

Diagrama



- Router A tiene 16Mbps conecta a ISP A
- Router B tiene 8Mbps y 4Mbps que conectan a los ISPs B y C

Estrategia de balanceo de carga de salida

- Disponible BGP se alimenta de los proveedores de Tránsito:
 - Tabla completa
 - Prefijos de clientes y default
 - Ruta por defecto
- Estas son las opciones comunes en Internet hoy en día
 - Muy raro para cualquier proveedor que ofrezca algo diferente
 - Muy raro para cualquier proveedor que personalice la alimentación de BGP para un cliente

Estrategia de balanceo de carga de salida

- Acepte solo una ruta por defecto del proveedor con la mayor conectividad, ISP A
 - Porque la mayoría del tráfico utilizará este enlace
- Si ISP A no proporciona un default:
 - Continúe corriendo BGP con ellos, pero descarte todos los prefijos
 - Apunte la ruta estática por defecto al enlace de subida
 - Distribuya el default en IGP
- Solicite la tabla completa a los ISP B y C
 - Mucho de esto será tirado
 - (“Default plus customers” no es suficiente)

Estrategia de balanceo de carga de salida

- ¿Cómo decidir qué mantener y qué descartar de los ISPs B y C?
 - La mayor parte del tráfico utilizará el enlace ISP A - así que tenemos que encontrar un subconjunto bueno/útil
- Descartar los prefijos que viajan por los ISPs de tránsito global
 - Los ISPs de tránsito global generalmente aparecen en la mayoría como non-local o AS-paths regionales
- Descartar prefijos con el ASN del ISP A en la trayectoria
 - tiene más sentido para el tráfico a esos destinos ir vía el enlace de ISP A

Estrategia de balanceo de carga de salida

- ISPs de Tránsito Global (Tier-1) incluyen:

209	CenturyLink	(Qwest)
701	VerizonBusiness	(UUNET)
1229	TeliaSonera	(Telia)
1239	Softbank	(Sprint)
1668	AOL TDN	
2914	NTT America	(NTT/Verio)
3549	Level 3	(GlobalCrossing)
3356	Level 3	
3561	CenturyLink	(Savvis, ex C&W)
7018	AT&T	

Peering ISP B de entrada, filtro AS-PATH

```
ip as-path access-list 1 deny _209_
ip as-path access-list 1 deny _701_
ip as-path access-list 1 deny _1239_
ip as-path access-list 1 deny _3356_
ip as-path access-list 1 deny _3549_
ip as-path access-list 1 deny _3561_
ip as-path access-list 1 deny _2914_
ip as-path access-list 1 deny _7018_
!
ip as-path access-list 1 deny _ISPA_
ip as-path access-list 1 deny _ISPC_
!
ip as-path access-list 1 permit _ISPB$
ip as-path access-list 1 permit _ISPB_[0-9]+$
ip as-path access-list 1 permit _ISPB_[0-9]+_[0-9]+$
ip as-path access-list 1 permit _ISPB_[0-9]+_[0-9]+_[0-9]+$
ip as-path access-list 1 deny .*
```

No necesita prefijos
ISPA e ISPC vía ISPB

Estrategia de balanceo de carga saliente: configuración peering

- Parte 1: tirar prefijos de ISPs de tránsito global
 - Esto puede ser bien parametrizado si el volumen del tráfico no es suficiente
 - (Más prefijos de entrada = más tráfico de salida)
- Parte 2: tirar los prefijos que transitan en la red del ISP A y C
- Parte 3: Permitir prefijos desde el ISP B, sus vecinos BGP, y sus vecinos, y sus vecinos
 - Más AS_PATH permite cláusulas, los prefijos más permitidos de entrada, los que más generan tráfico de salida
 - Muchos prefijos de entrada significaría más tráfico saliente del que el enlace del ISP B puede manejar

Estrategia de balanceo de carga saliente

- Similar al filtro AS-PATH puede ser construido para el peering BGP del ISP C
- Si los mismos prefijos son escuchados por ambos ISPs (B y C), entonces establecen proximidad de su ASN origen a ISP B o C
 - ej: ISP B podría estar en Japón, con su ASN vecino en Europa, todavía ISP C podría estar en Europa might be in Japan, with the neighbouring ASN in Europe, yet ISP C might be in Europe
 - el tránsito al ASN vía ISP C tiene más sentido en este caso

Estrategia de balanceo de carga entrante

- El enlace saliente más grande debería anunciar sólo el agregado
- Los otros enlaces solo deben anunciar:
 - a) El agregado con AS-PATH se antepone
 - b) Subprefijos del agregado, se elijen de acuerdo al volumen de tráfico de esos subprefijos, y de acuerdo a los servicios de esos subprefijos
- Ejemplo:
 - El enlace a ISP B podría ser usado solo para Banda ancha/clientes de marcado (dial) -- por lo que el número de tales clientes están fuera de un prefijo contiguo
 - El enlace a ISP C podría ser usado para clientes de línea comercial de alquiler - por lo que el número de tales clientes está fuera de un subprefijo contiguo

Router A: ejemplo de configuración eBGP

```
router bgp 100
  network 100.10.0.0 mask 255.255.224.0
  neighbor 122.102.10.1 remote 110
  neighbor 122.102.10.1 prefix-list default in
  neighbor 122.102.10.1 prefix-list aggregate out
!
ip prefix-list default permit 0.0.0.0/0
ip prefix-list aggregate permit 100.10.0.0/19
!
```

Router B: ejemplo de configuración eBGP

```
router bgp 100
  network 100.10.0.0 mask 255.255.224.0
  neighbor 120.103.1.1 remote 120
  neighbor 120.103.1.1 filter-list 1 in
  neighbor 120.103.1.1 prefix-list ISP-B out
  neighbor 120.103.1.1 route-map to-ISP-B out
  neighbor 121.105.2.1 remote 130
  neighbor 121.105.2.1 filter-list 2 in
  neighbor 121.105.2.1 prefix-list ISP-C out
  neighbor 121.105.2.1 route-map to-ISP-C out
!
ip prefix-list aggregate permit 100.10.0.0/19
!
...siguiente diapositiva
```

Router B: ejemplo de configuración eBGP

```
ip prefix-list ISP-B permit 100.10.0.0/19
ip prefix-list ISP-B permit 100.10.0.0/21
!
ip prefix-list ISP-C permit 100.10.0.0/19
ip prefix-list ISP-C permit 100.10.28.0/22
!
route-map to-ISP-B permit 10
  match ip address prefix-list aggregate
  set as-path prepend 100
!
route-map to-ISP-B permit 20
!
route-map to-ISP-C permit 10
  match ip address prefix-list aggregate
  set as-path prepend 100 100
!
route-map to-ISP-C permit 20
```

← /21 a ISP B
"clientes de
mercado"

← /22 a ISP C
"biz customers"

← ej: solo anteponer
sobre enlace ISP B

← ej: Anteposición
doble sobre enlace
ISP C

¿Qué hay sobre el backup saliente?

- Tenemos:
 - Ruta por defecto de ISP A por eBGP
 - Tabla completa mayormente descartada de ISPs B y C
- Estrategia:
 - Originar ruta por defecto por OSPF sobre router A (con métrica 10) - enlace a ISP A
 - Originar ruta por defecto por OSPF sobre Router B (con métrica 30) - enlaces a ISP B y C
 - Agregar sobre Router B:
 - Ruta estática por defecto al ISP B con distancia 240
 - Ruta estática por defecto al ISP C con distancia 245
 - Cuando el enlace se cae, la ruta estática es retirada

Backup saliente: estado estable

- Estado estable (todos los enlaces arriba y activos):
 - La ruta por defecto es hacia el Router A - métrica 10 para OSPF
 - (Porque el default aprendido por eBGP → default está en RIB → OSPF originará default)
 - El respaldo por defecto es hacia Router B - métrica 20 para OSPF
 - Prefijos eBGP aprendidos para conexiones de subida son distribuidos por el iBGP a lo largo del troncal
 - (Default puede ser filtrado en iBGP para evitar "un error de falla RIB")

Backup saliente: ejemplos de falla

- El enlace a ISPA caído, a IPSs B y C levantados:
 - La ruta por defecto es hacia Router B - métrica 20 para OSPF
 - (eBGP default se ha ido de RIB, por lo que OSPF en Router A retira el default)
- Lo anterior es cierto si el enlace hacia B o C están caídos
- Enlace hacia ISPs B y C están caídos, el enlace hacia ISP está levantado:
 - Ruta por defecto es hacia Router A - métrica 10 para OSPF
 - (defaults estáticos sobre Router B removidos de RIB, por lo que OSPF sobre el Router B retira el default)

Otras consideraciones

- Ruta por defecto no debería ser propagada a los dispositivos terminales de peers de no-tránsito y clientes
- Rara vez hay necesidad de llevar default en iBGP
 - Mejor filtrar default en peering mallados iBGP
- Aún así llevar otros prefijos eBGP a través de la malla iBGP
 - De otra manera los enrutadores seguirán las reglas de la ruta por defecto, resultando en flujo de tráfico subóptimo
 - no es un gran problema porque no lleva la tabla completa

Router A: ejemplo de configuración iBGP

```
router bgp 100
  network 100.10.0.0 mask 255.255.224.0
  neighbor ibgp-peers peer-group
  neighbor ibgp-peers remote-as 100
  neighbor ibgp-peers prefix-list ibgp-filter out
  neighbor 100.10.0.2 peer-group ibgp-peers
  neighbor 100.10.0.3 peer-group ibgp-peers
!
ip prefix-list ibgp-filter deny 0.0.0.0/0
ip prefix-list ibgp-filter permit 0.0.0.0/0 le 32
!
```

Resumen de tres upstreams, ancho de banda desigual

- Ejemplo basado en muchas implementaciones de trabajo en topologías multihoming/balanceo de carga
- Muchas variaciones posibles - ésta es una:
 - Fácil de parametrizar
 - Liviano en los recursos del enrutador de borde
 - Liviano en la infraestructura del enrutador troncal
 - Tabla de BGP limitada \Rightarrow más rápida convergencia

Proveedor de servicio Multihoming



Talleres ISP FIN