

upna

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Tecnologías Avanzadas de Red
Área de Ingeniería Telemática

BGP-4

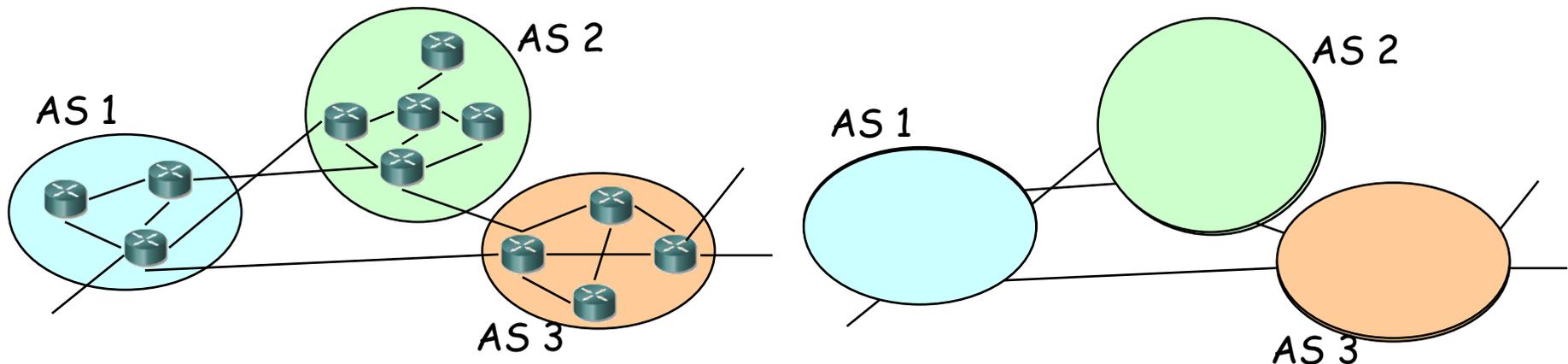
Area de Ingeniería Telemática

<http://www.tlm.unavarra.es>

Grado en Ingeniería en Tecnologías de
Telecomunicación, 3º

Enrutamiento jerárquico

- Enrutamiento jerárquico
 - IGP: Interior Gateway Protocol
 - EGP: Exterior Gateway Protocol
 - Interior/exterior respecto a “sistemas autónomos” (*Autonomous Systems*)
- Características EGP:
 - Mejor escalabilidad
 - Habilidad para agregar rutas
 - Habilidad para expresar políticas
 - Mayor carga en el router



upna

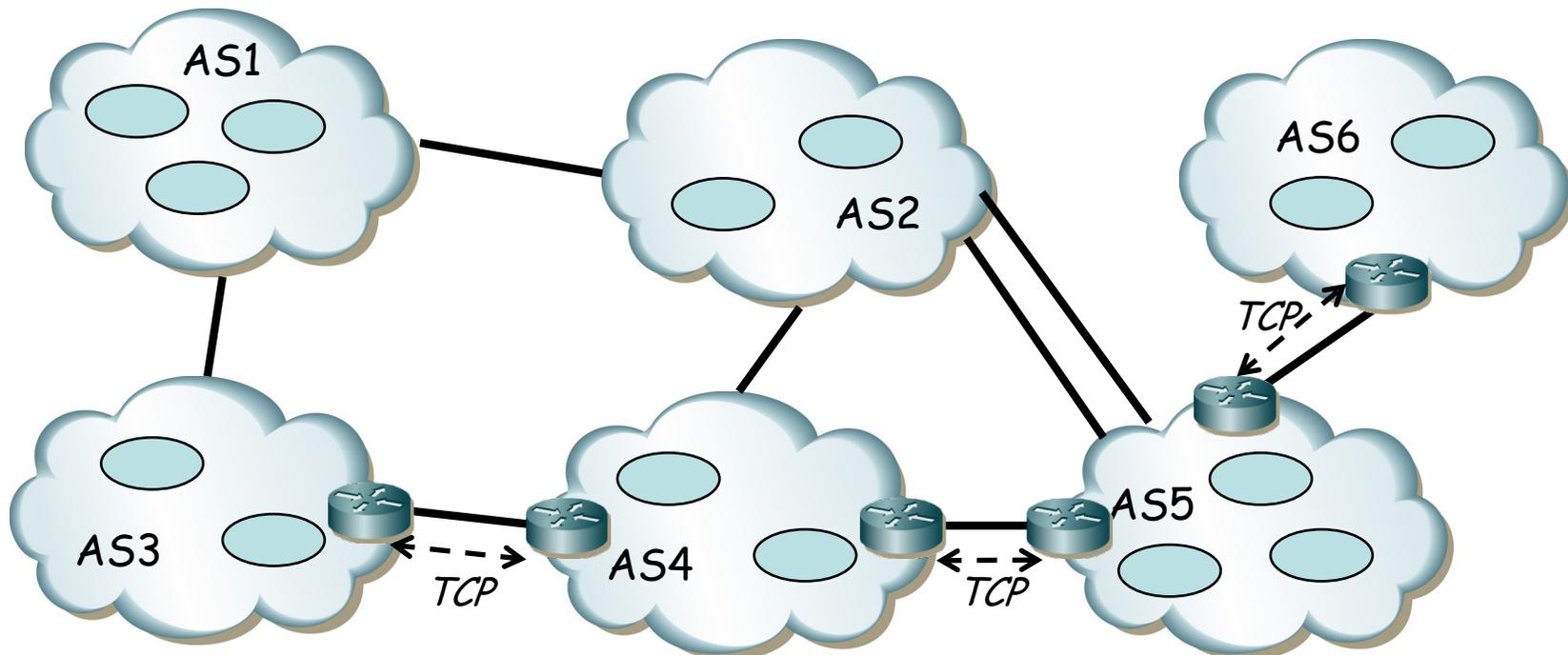
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Tecnologías Avanzadas de Red
Área de Ingeniería Telemática

BGP: Introducción

BGP

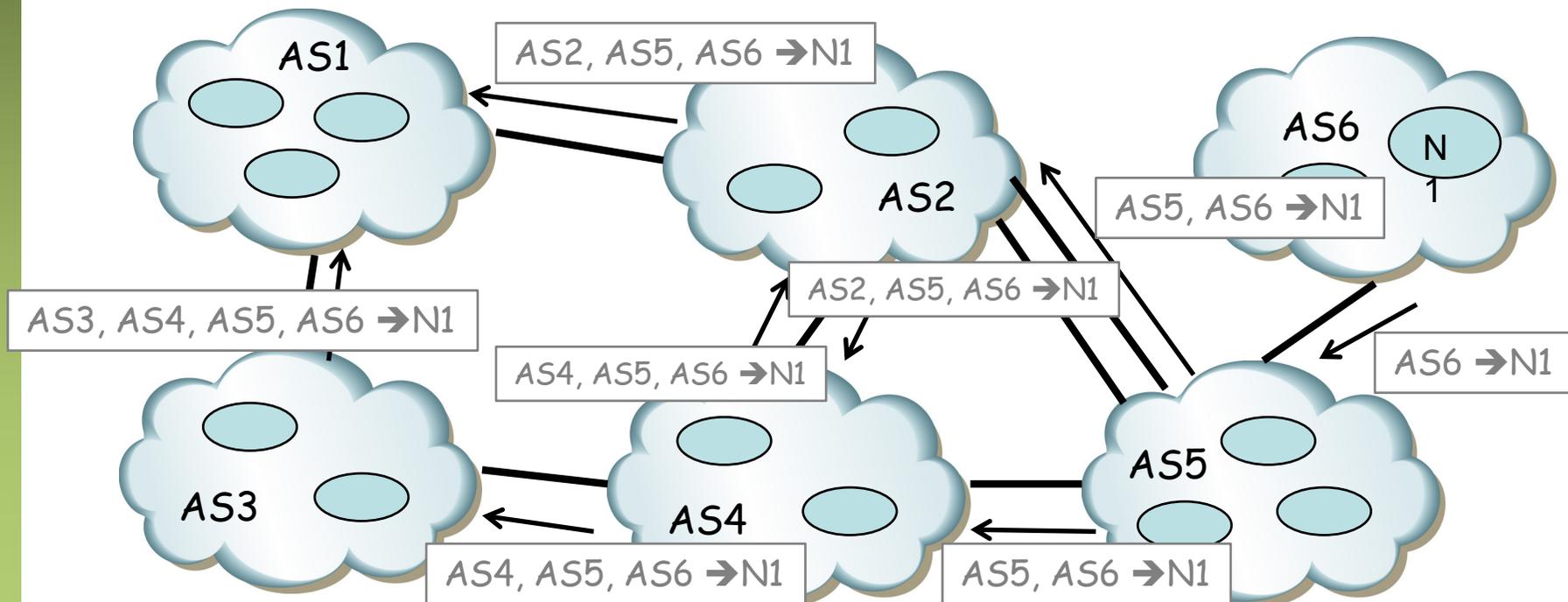
- *Border Gateway Protocol*
- BGP-4, RFC 4271
- BGP-4 primera versión classless
- Protocolo Interdomain estándar *de facto*
- Comunicación fiable mediante conexión TCP entre routers adyacentes
- Puerto 179



BGP

Path Vector

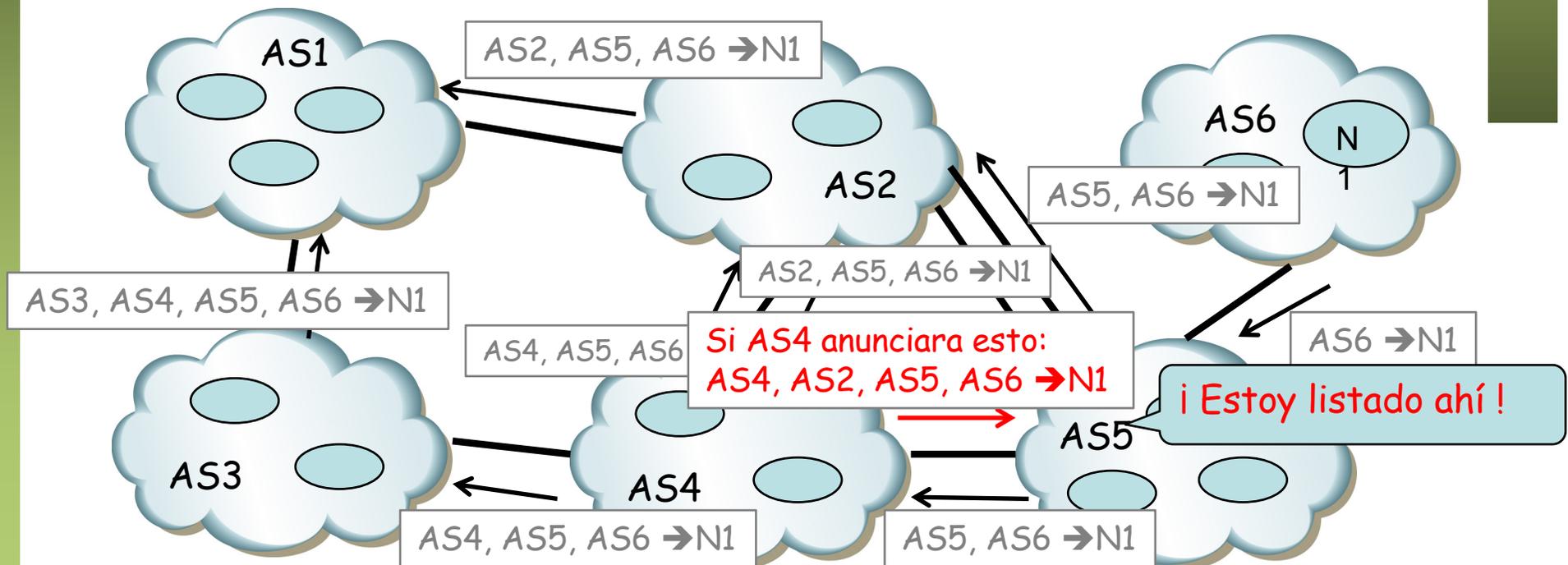
- Calcula caminos a prefijos
- Como DV recibe de vecinos, calcula sus rutas y envía a vecinos
- La adyacencia es una conexión TCP
- En vez de métrica anuncia la lista de AS en cada camino (. . .)
- Por defecto elige el camino que pasa por menor número de ASs



BGP

Path Vector

- Anunciar el camino permite evitar los ciclos
- El menor número de ASs no quiere decir que sea el menor número de saltos por routers



upna

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Tecnologías Avanzadas de Red
Área de Ingeniería Telemática

BGP: Mensajes

Mensajes

- Primero se establece la conexión TCP entre los dos *BGP speakers*
- Cuatro mensajes obligatorios

OPEN

- Tras establecerse la conexión
- Router especifica parámetros de operación: versión, identificador, AS number, *hold time*, *capabilities*, etc.
- Suele ir seguido de un intercambio de todas las rutas

KEEPALIVE

- Para comprobar periódicamente el *peering*
- Se da por rota la sesión si pasa el *hold time* sin recibirlo

NOTIFICATION

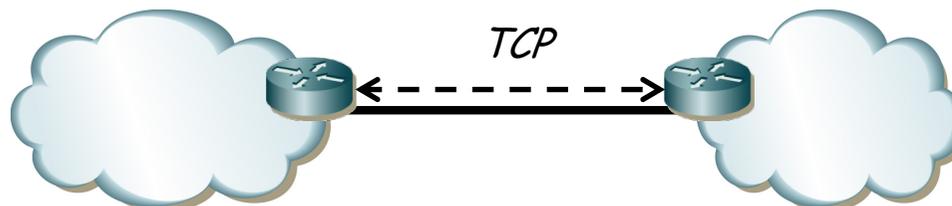
- Cuando se detecta un error
- Termina la conexión

UPDATE

- Anuncia información de enrutamiento (nuevas rutas o eliminar otras – *withdraw* –)
- Anuncia un solo camino por mensaje
- Anuncia cuando ha calculado una nueva mejor ruta al destino
- Si deja de poder alcanzarlo anuncia eso también
- Prefijo / Longitud
- **Atributos del camino:** permiten a BGP elegir el mejor

ROUTE-REFRESH (opcional)

- Para pedir que vuelva a anunciar los prefijos que conoce



upna

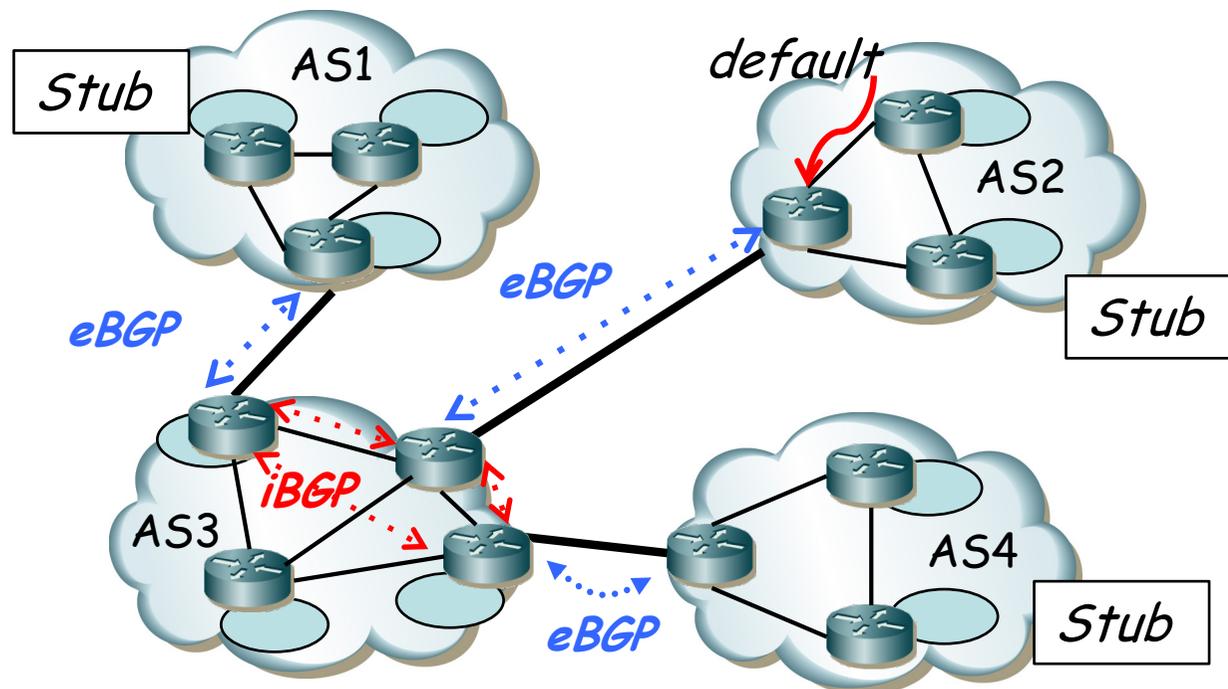
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Tecnologías Avanzadas de Red
Área de Ingeniería Telemática

eBGP vs iBGP

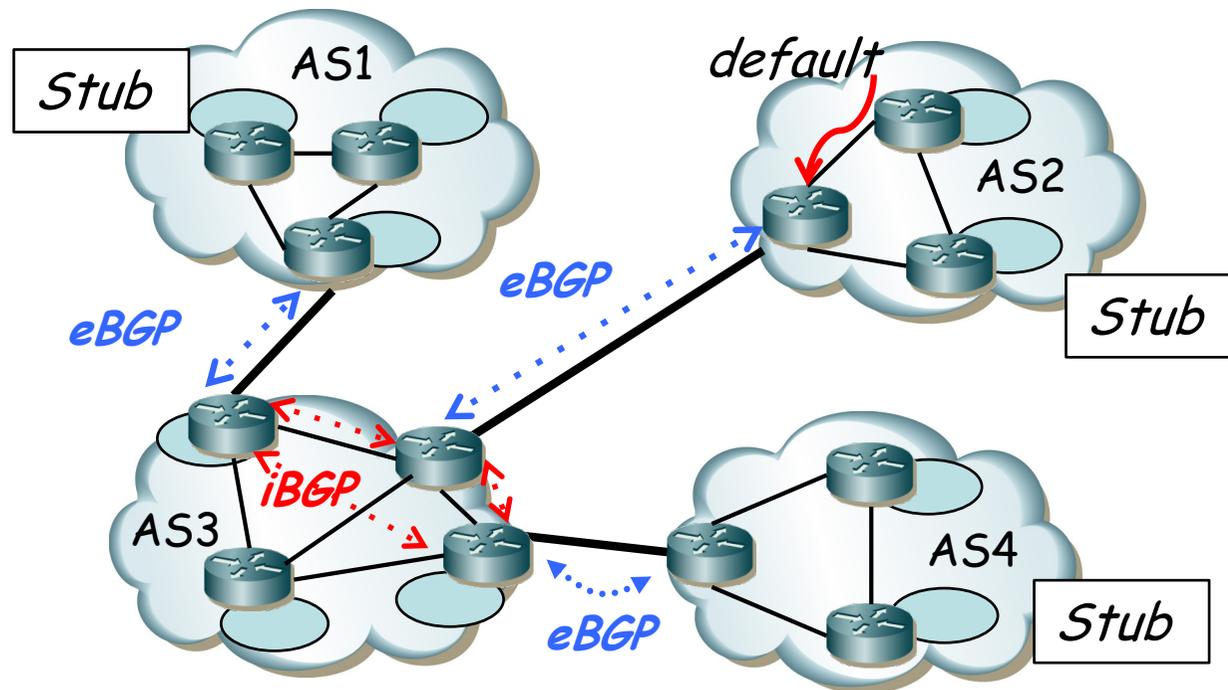
Peering en BGP

- Los *peers* de un proceso BGP pueden estar:
 - En otro AS: *external peer* ⇒ **eBGP**
 - En el mismo AS: *internal peer* ⇒ **iBGP**
- (...)



Peering en BGP

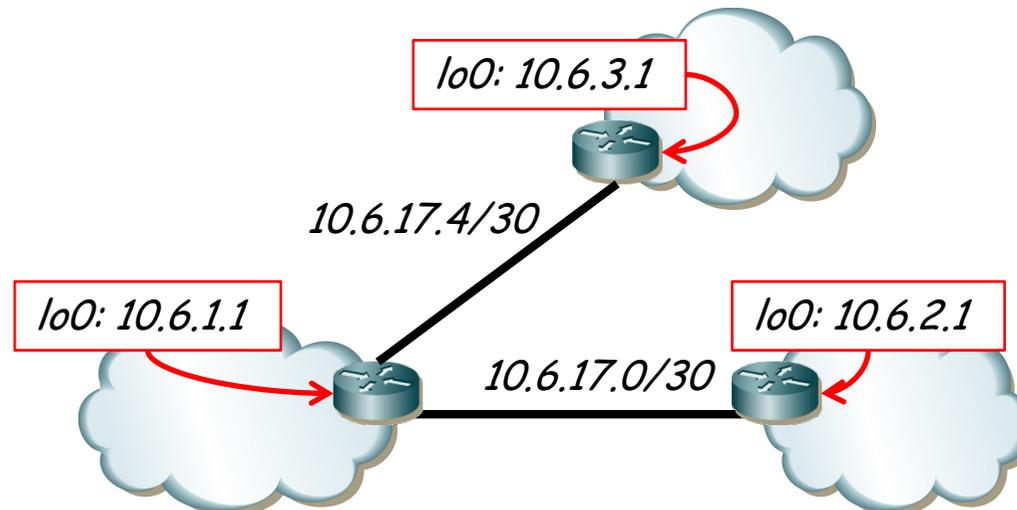
- En el mismo AS el *peering* iBGP forma una malla porque...
- No se pasan por iBGP prefijos aprendidos por iBGP
- Reconoce si es del mismo AS porque en el OPEN anuncia el ASN
- No interesa difundir todas las rutas al IGP (escalabilidad)
- iBGP permite que otros ASBRs aprendan los prefijos a anunciar
- El ASN se añade a la ruta al hacer anuncio a otro *eBGP*



ASBR = Autonomous System Border Router

Direccionamiento

- El enlace entre dos ASBRs empleará un direccionamiento
- Frecuentemente es parte de la asignación de uno de los dos ASs
- Un router con más de un *peer* tendría una dirección diferente para cada uno, lo cual complica la gestión
- Necesita un identificador único del router, pero si es una de sus direcciones IP, ¿qué sucede si ese interfaz falla?
- Para tener una única identificación se emplea la de un interfaz de *loopback*
- Algunas implementaciones envían los paquetes con TTL=2 si se usa *loopback* y con TTL=1 si no se usa



Atributos en BGP

Path Attributes

- Son características de una ruta BGP

Tipos según se soporten:

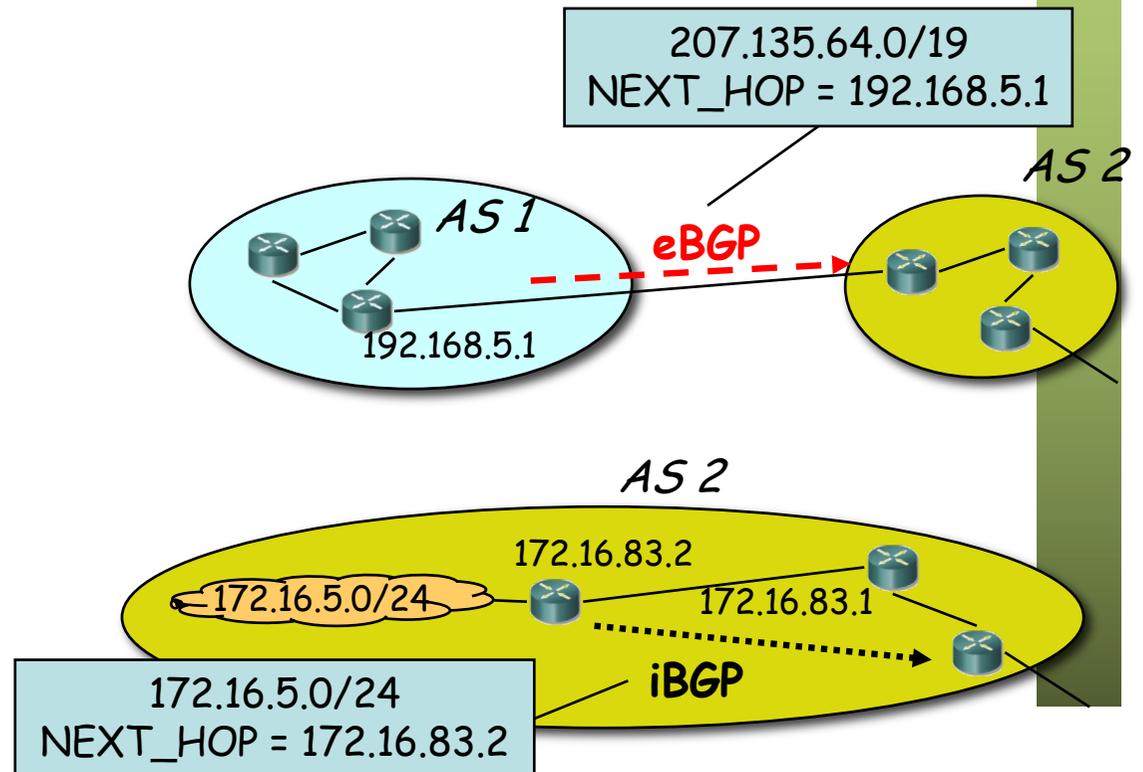
- *Well-known: mandatory* (en update) o *discretionary*
- *Optional: transitive* o *nontransitive*

ORIGIN (well-known mandatory)

- IGP, EGP o Incompleto (rutas estáticas)

NEXT_HOP (well-known mandatory)

- Si son *External Peers* es la IP del interfaz del router anunciante
- Si son *Internal Peers* y
 - Destino en el mismo AS: IP del anunciante (*recursive lookups*)
 - Destino fuera del AS: (...)



"well-known" : Debe soportarlo

"Optional" : No está obligado a soportarlo

"mandatory" : Debe aparecer en los mensajes

"discretionary" : Puede no aparecer en los mensajes

"Transitive" : Debe reenviarlo

"Nontransitive" : No debe reenviarlo

Path Attributes

- Son características de una ruta BGP

Tipos según se soporten:

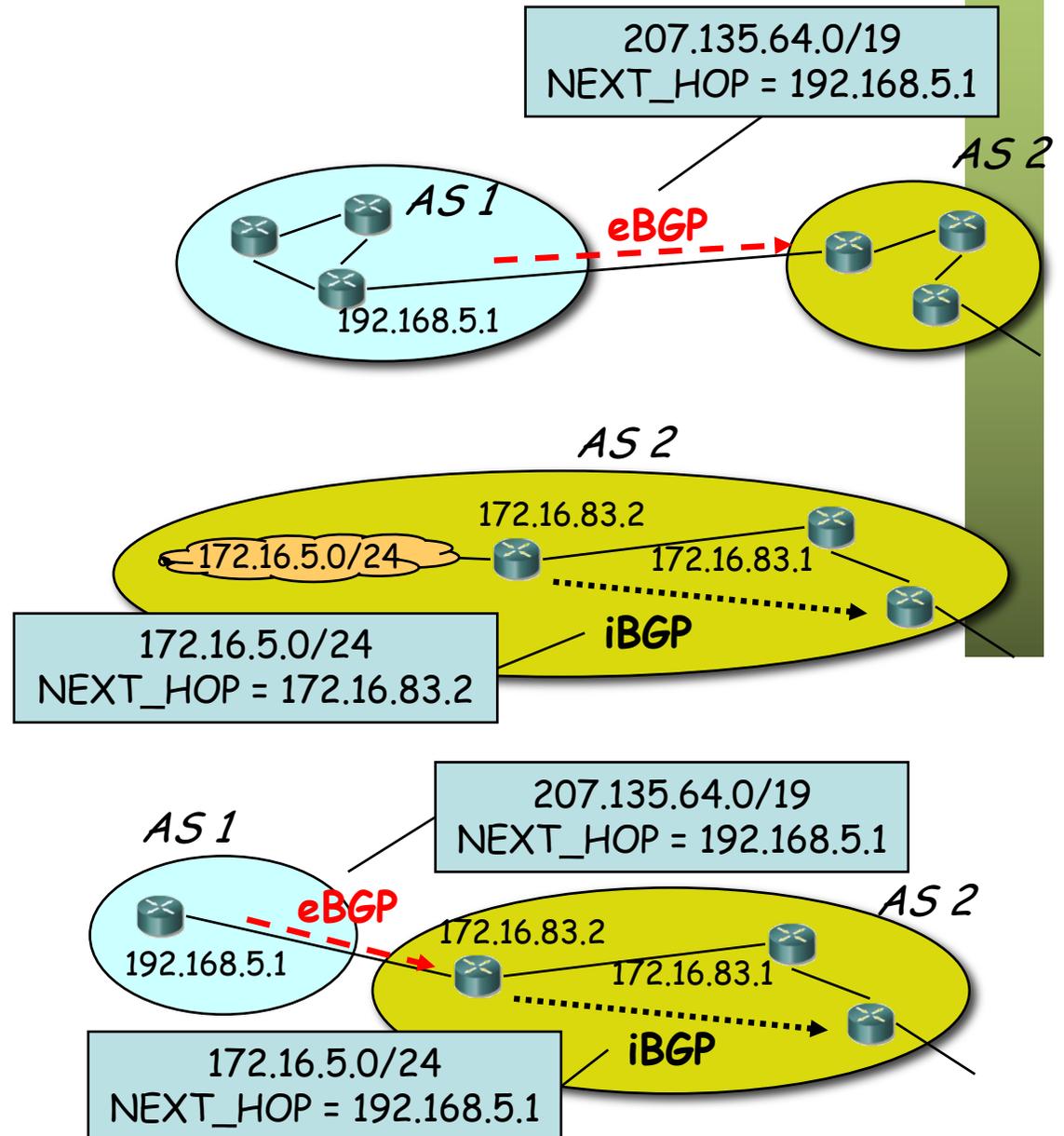
- *Well-known: mandatory* (en update) o *discretionary*
- *Optional: transitive* o *nontransitive*

ORIGIN (well-known mandatory)

- IGP, EGP o Incompleto (rutas estáticas)

NEXT_HOP (well-known mandatory)

- Si son *External Peers* es la IP del interfaz del router anunciante
- Si son *Internal Peers* y
 - Destino en el mismo AS: IP del anunciante (*recursive lookups*)
 - Destino fuera del AS: IP del *peer* externo



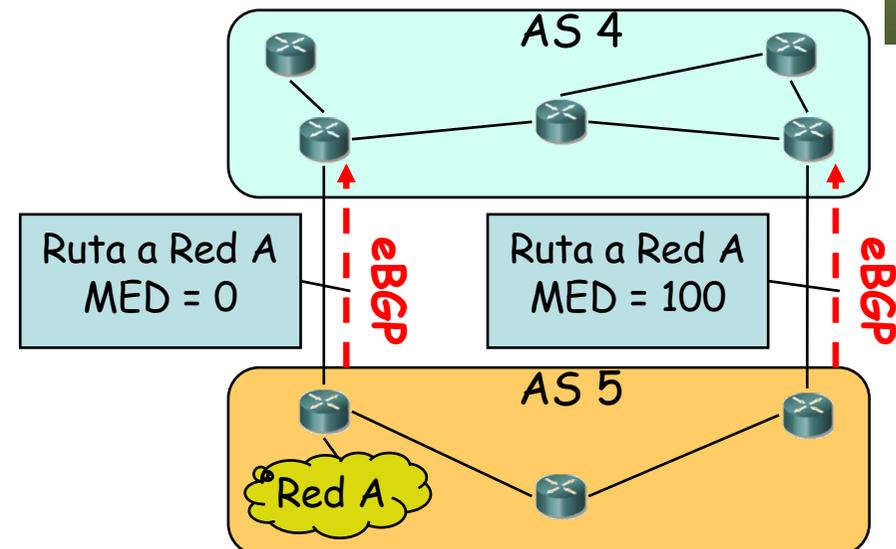
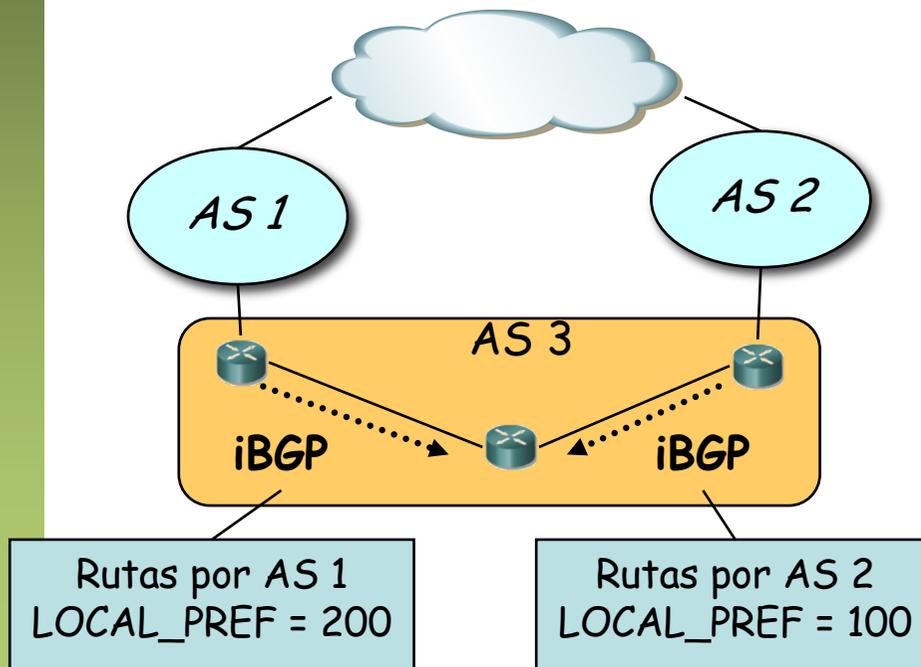
Path Attributes

LOCAL_PREF (well-known discretionary, nontransitive)

- Solo en iBGP
- Comunica el grado de preferencia por una ruta
- La ruta de mayor valor es seleccionada

MED (optional, nontransitive)

- Multi-Exit-Discriminator
- Cuando hay múltiples links a un AS
- Anuncia el *ingress point* preferido
- Es una métrica y se selecciona el de menor MED
- No se propaga a más ASs (debe borrarlo al pasar la ruta a otro AS)



Un criterio de selección

1. Ruta con el mayor **LOCAL_PREF**
2. Si iguales, la ruta de **AS_PATH** más corto
3. Si iguales, la ruta de origen menor (**ORIGIN** IGP < EGP < Incomplete)
4. Si iguales y van al mismo AS, la de menor **MED**
5. Si igual, la de menor **métrica** del IGP hasta el NEXT_HOP
6. Si iguales y van al mismo AS, se puede instalar todas las rutas o escoger la de menor identificador de router

upna

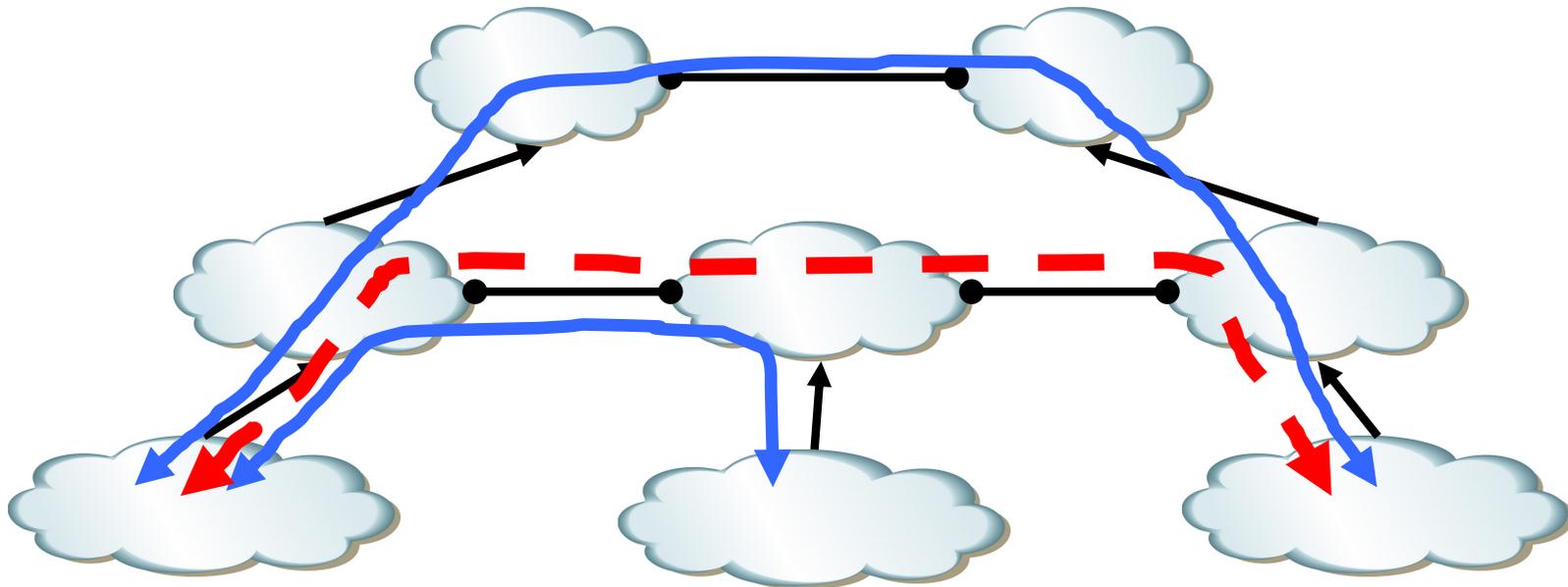
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Tecnologías Avanzadas de Red
Área de Ingeniería Telemática

BGP e Internet

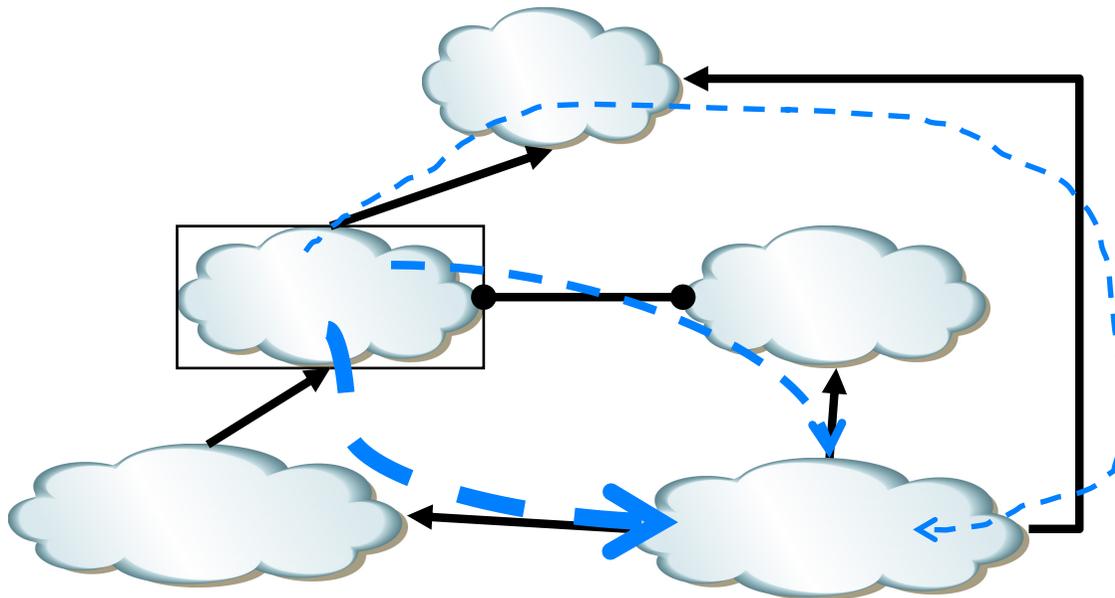
Jerarquía y economía

- En la Internet tenemos enlaces
 - Cliente-Proveedor (de pago)
 - Entre iguales (normalmente no se pagan)
- Por un enlace entre pares no se hace tránsito (...)
- Preferencia habitual:
 - (...)



Jerarquía y economía

- En la Internet tenemos enlaces
 - Cliente-Proveedor (de pago)
 - Entre iguales (normalmente no se pagan)
- Por un enlace entre pares no se hace tránsito (...)
- Preferencia habitual:
 1. Por cliente
 2. Por *peer*
 3. Por proveedor



Políticas

- Anunciar una ruta implica que se está dispuesto a encaminar el tráfico a ese destino

Los administradores pueden implementar diferentes políticas:

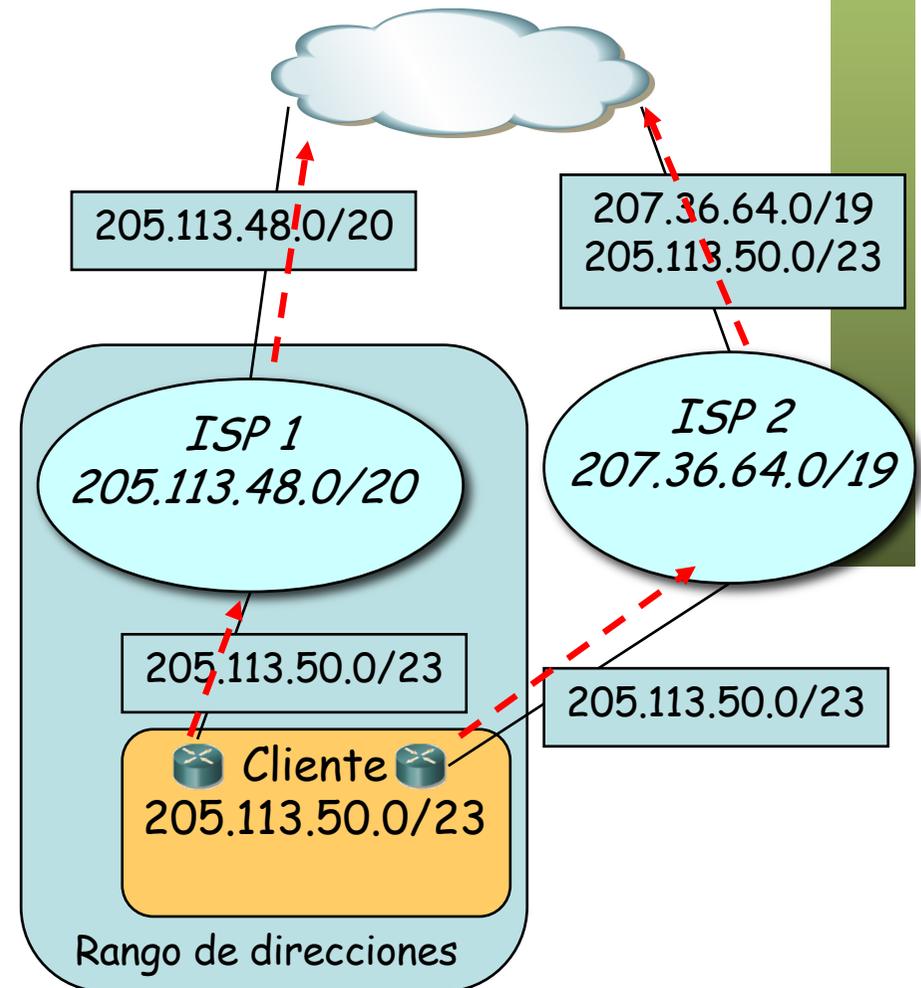
- No anunciar un destino a un vecino
- No usar caminos que pasen por cierto AS
- Ignorar el MED y usar *shortest-paths (hot potato routing)*
- Añadir varias veces su ASN
- Etc.

Problemas

- Hay políticas que no convergen
- Hay políticas que pueden converger dependiendo del orden de los mensajes
- Hay políticas que convergen pero dejan de hacerlo si un enlace se cae
- Dadas las políticas y la topología, decidir si convergerá es NP-completo

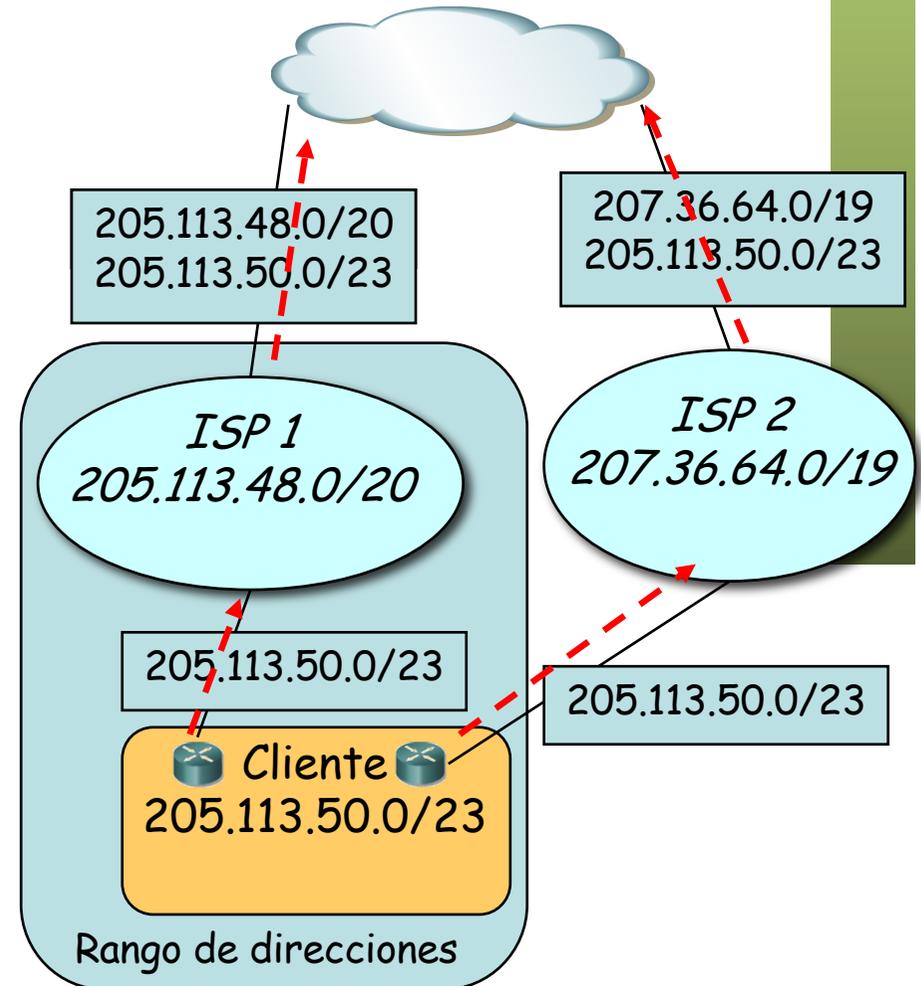
Multihoming

- Para ofrecer redundancia
- El rango de direcciones pertenece al ISP 1
- Habrá que anunciarlo también al ISP 2
- Ahora la ruta por ISP 2 es más específica
- *Address leaking*: ISP 1 debería anunciar también la ruta específica (...)



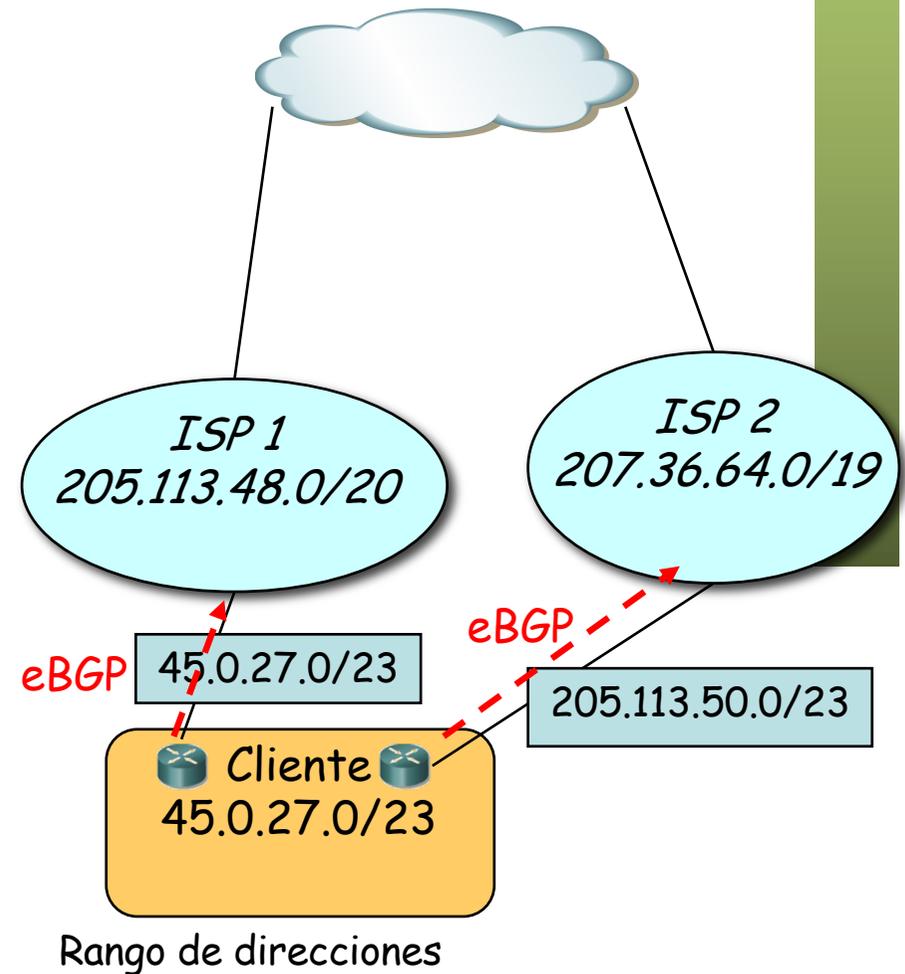
Multihoming

- Para ofrecer redundancia
- El rango de direcciones pertenece al ISP 1
- Habrá que anunciarlo también al ISP 2
- Ahora la ruta por ISP 2 es más específica
- *Address leaking*: ISP 1 debería anunciar también la ruta específica
- (...)



Multihoming

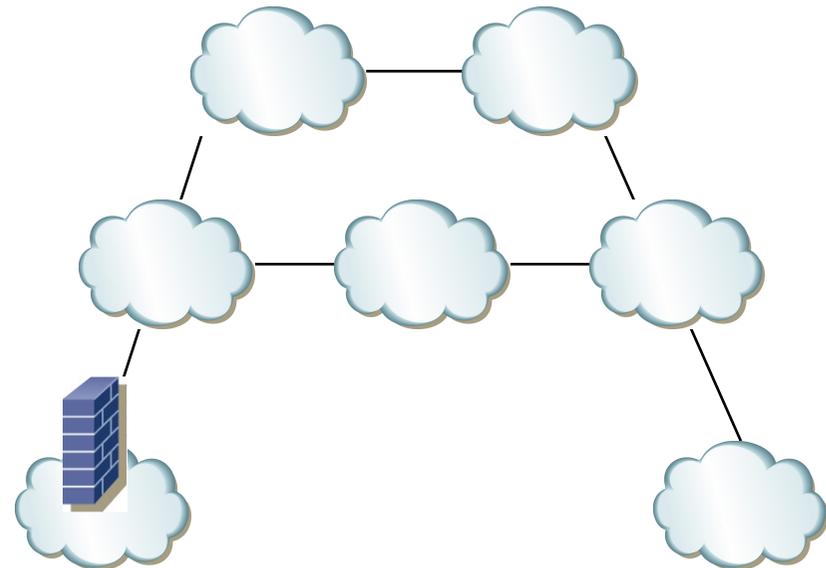
- Para ofrecer redundancia
- El rango de direcciones pertenece al ISP 1
- Habrá que anunciarlo también al ISP 2
- Ahora la ruta por ISP 2 es más específica
- *Address leaking*: ISP 1 debería anunciar también la ruta específica
- Más habitual tener un espacio de direcciones propio
- Ser un AS y correr BGP



Precauciones

Martians

- Algunos prefijos no se deben anunciar ni enrutar paquetes de ellos
- Ruta por defecto (0.0.0.0/0)
- Direccionamiento privado
 - 10.0.0.0/8
 - 172.16.0.0/12
 - 192.168.0.0/16
- *Link-local* (169.254.0.0/16)
- TEST_NET (192.0.2.0/24, etc.)
- Clases D y E (224.0.0.0/3)
- Reservados para IANA



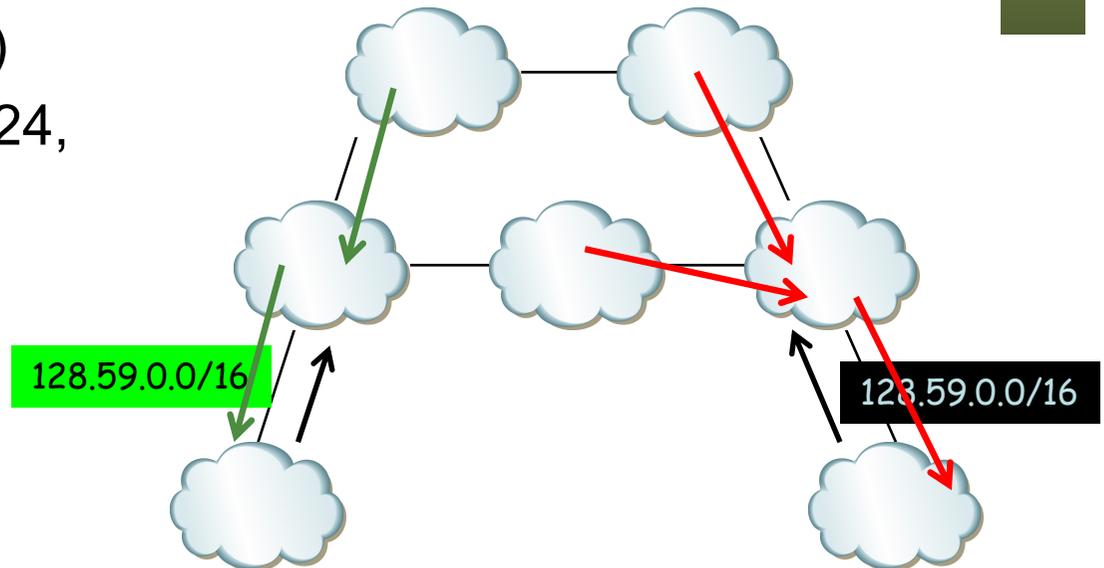
Precauciones

Martians

- Algunos prefijos no se deben anunciar ni enrutar paquetes de ellos
- Ruta por defecto (0.0.0.0/0)
- Direccionamiento privado
 - 10.0.0.0/8
 - 172.16.0.0/12
 - 192.168.0.0/16
- *Link-local* (169.254.0.0/16)
- TEST_NET (192.0.2.0/24, etc.)
- Clases D y E (224.0.0.0/3)
- Reservados para IANA

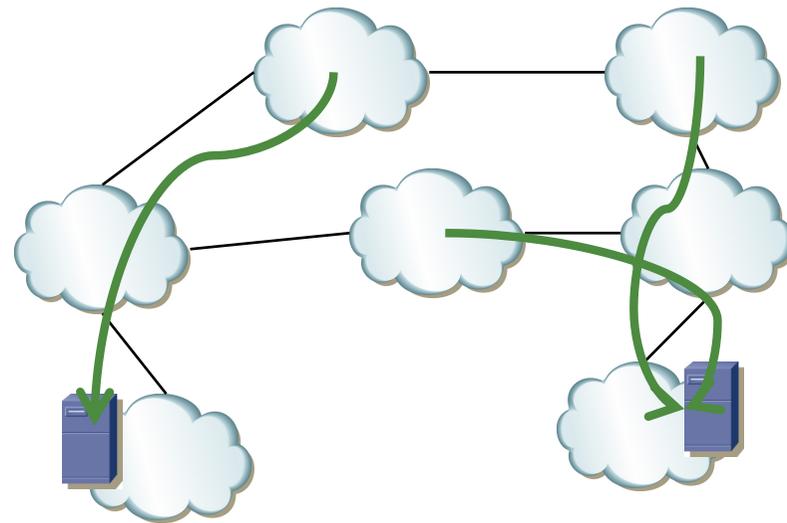
Black holes

- Si un AS anuncia un prefijo al que no está conectado
- El real puede dejar de ser accesible desde ciertas redes
- O puede hacer pasar tráfico por él



Anycast

- Servidores con misma dirección IP (contenido replicado o no)
- Todos en la misma red física o en diferentes
- Anuncios por ejemplo por diferentes proveedores
- Clientes acceden a servidor según proximidad
- Permite distribución de contenidos
- También se puede hacer en el IGP
- Ejemplo: F-root name server



Otras características

- Agregación de rutas
 - Gracias a CIDR
 - Combinar prefijos de dos o más ASs y anunciar el combinado
- *Route Reflectors*
 - Mejorar escalabilidad de iBGP (que crea un *full-mesh*)
 - Un router reflector actúa como un concentrador
- *Confederations*
 - Mejora escalabilidad de iBGP
 - Dividir AS en varios de forma que entre ellos sea eBGP
 - La confederación tiene un ASN y cada sub-AS puede tenerlo o usar uno privado
- *Route Flap Dampening*
 - Para evitar rápidas oscilaciones en una ruta
 - Aumenta el tiempo de convergencia