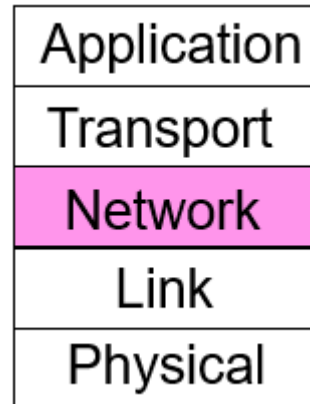


Capítulo 4

La Capa de Red – BGP



BGP

- **Situación:**
 - La naturaleza de los PPEE es muy distinta a la de los protocolos de enrutamiento de puerta de enlace interior.
 - Lo que lleva a que surjan problemas no considerados antes a resolver.
 - No hemos estudiado ningún PPEE.
 - Pero acabamos de ver algunas características de estos.
- **Problema: Construir un PPEE para internet.**
- **Solución: BGP (Border Gateway Protocol)** es el PPEE de facto que usa internet.

BGP

- Tareas que realiza BGP:
- BGP provee a cada SA un medio para:
 - Obtener *información de alcanzabilidad* de **subredes** desde SA vecinos.
 - **Propagar** la información de alcanzabilidad a todos los enrutadores dentro del SA.
 - Determinar **“buenas” rutas** a las subredes basándose en la información de alcanzabilidad y en las **políticas** del SA.
 - BGP permite a cada subred **publicar su existencia** al resto de la internet.
 - BGP se asegura que todos los SA de la internet conozcan acerca de la subred y cómo llegar allí.

BGP

- BGP permite a cada SA aprender cuáles destinos son alcanzables vía sus SA vecinos.
- En BGP ***los destinos son prefijos***, donde
 - cada prefijo representa una **subred** (según Kurose) o una **colección de subredes** (definida usando agregación de prefijos - CIDR).
- **¿Qué significa que un SA *S* avisa un prefijo a un SA *S1*?**
 - S promete que va a enviar datagramas hacia ese prefijo.

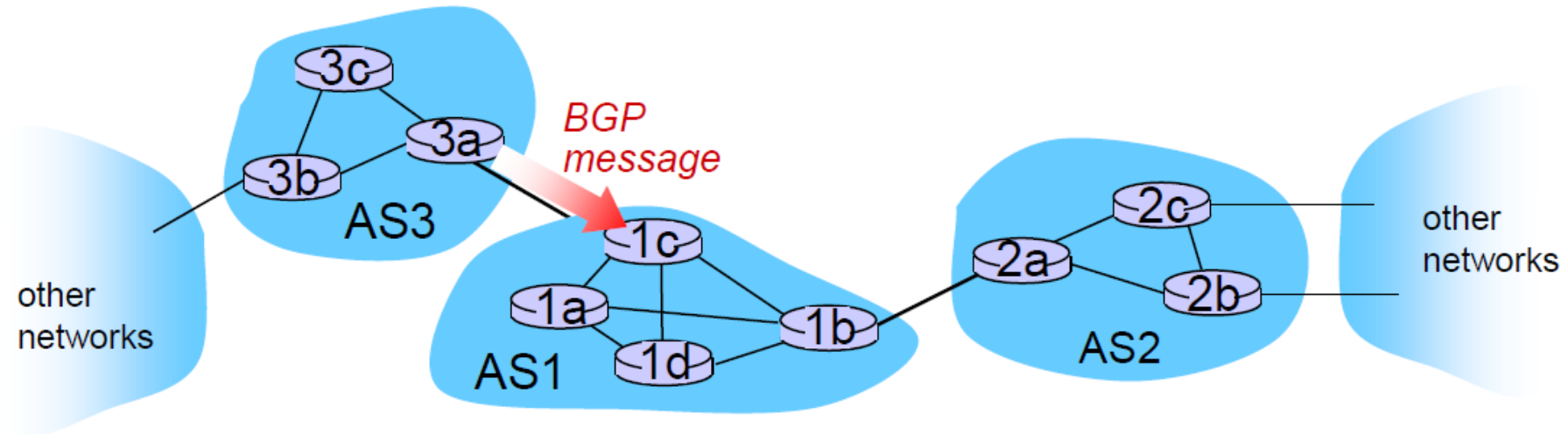
Atributos BGP

- En BGP un SA es identificado por un número globalmente único llamado **número de sistema autónomo (ASN)**.
- **Problema: ¿De cuál información se debe dar aviso cuando se avisa un prefijo?**

Atributos BGP

- **Solución:** Cuando un enrutador avisa de un prefijo a lo largo de una sesión BGP incluye con el prefijo una ruta que pasa por varios SA para llegar al prefijo.
 - Una **ruta** se compone de un prefijo más **atributos BGP**.
 - **Algunos atributos importantes:**
 - **AS-PATH:** contiene los SA por los cuales el aviso del prefijo ha pasado. Cuando un prefijo pasa por un SA, el SA agrega su ASN al atributo AS-PATH.
 - **NEXT-HOP:** es el IP de la interfaz del enrutador que comienza el AS-PATH hacia el destino.

Atributos BGP



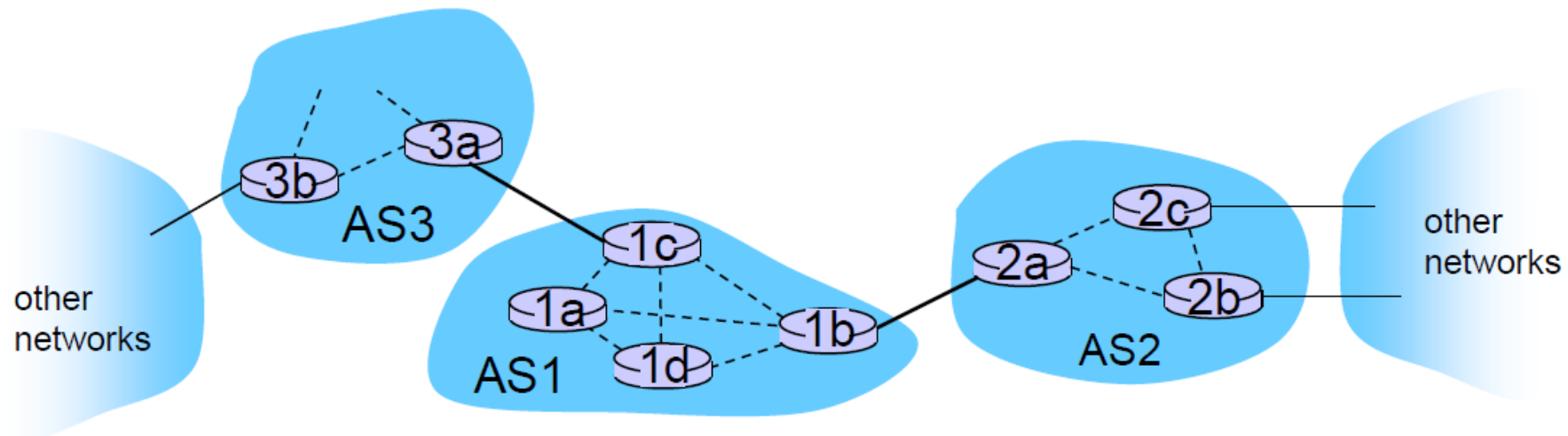
❖ Ejemplo de ruta:

- Prefix: 138.16.64/22 ; AS-PATH: AS3 AS131 ;
NEXT-HOP: 201.44.13.125

Ejemplo de uso de atributos BGP

❖ Ejemplo:

- El enrutador 3a en AS3 avisa una ruta a enrutador 1c en AS1 usando BGP,
 - la ruta incluye el prefijo avisado x y un AS-PATH al prefijo;
 - este aviso también incluye el NEXT-HOP que es la dirección IP del enrutador 3a que lleva a 1c.
- Una vez que aprende la ruta a x el enrutador 1d quiere poder enviar paquetes a x a lo largo de la ruta;
- o sea, el enrutador 1d puede querer incluir la entrada (x, I) en su tabla de reenvío donde I es el enlace que **comienza** el camino más corto de 1d hasta la puerta de enlace 1c.



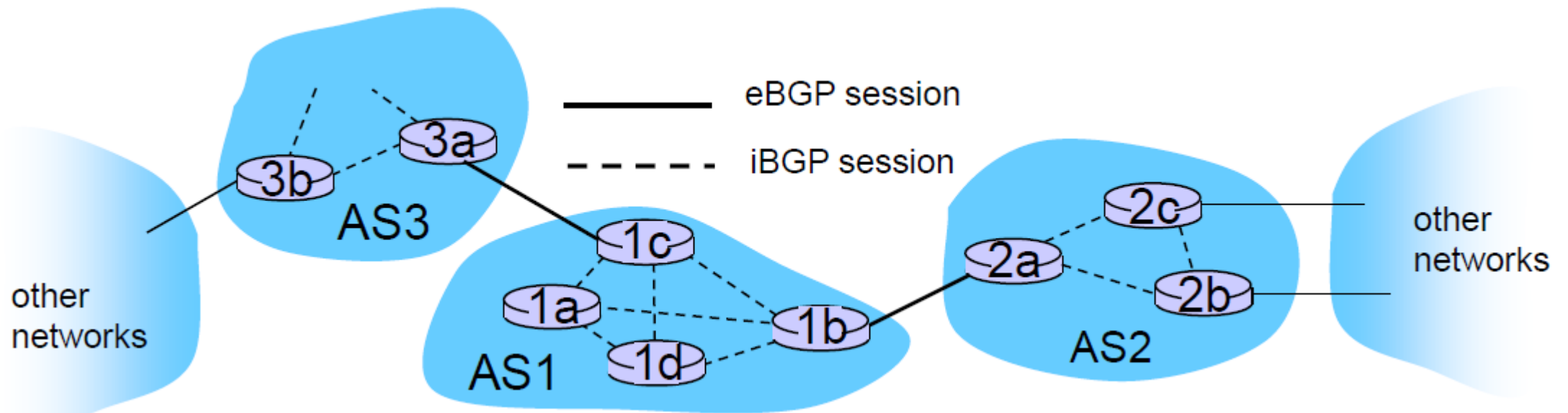
BGP

- **Problema:** ¿Cómo hacer para propagar información de rutas en BGP?
- **Solución:** En BGP pares de enrutadores *intercambian información* de rutas sobre *conexiones TCP semipermanentes* usando el puerto 179.
 - Hay típicamente una conexión BGP TCP para
 - cada enlace que conecta directamente dos **enrutadores EBSA** (o **enrutadores BGP**) en dos SA diferentes y
 - para enlaces entre enrutadores dentro del SA
 - Para cada conexión TCP, los 2 enrutadores al final de la conexión se llaman **compañeros BGP**.
 - Los compañeros BGP se *avisan rutas*.

BGP

- **Sesiones BGP:**

- La conexión TCP con todos los mensajes BGP enviados por la conexión se llama **sesión BGP**.
- Una sesión BGP entre enrutadores de dos SA se llama **sesión externa BGP (eBGP)**
- Una sesión BGP entre enrutadores en el mismo SA se llama **sesión interna BGP (iBGP)**
- Las líneas de las sesiones BGP no siempre se corresponden con los enlaces físicos.



Aviso de Rutas

- **Cuando una puerta de enlace P recibe rutas**
 - P usa las sesiones iBGP para distribuir las rutas a los otros enrutadores del SA de P .
 - Las sesiones iBGP se usan para distribuir rutas a los enrutadores dentro del SA.
- En BGP, **un vecino de un enrutador** es otro enrutador con el cual se establece una sesión BGP para intercambiar información de enrutamiento.
- El vecino se configura manualmente mediante la definición de su dirección IP y número de AS en la configuración BGP del enrutador.
- En algunos casos, especialmente con IPv6, existen mecanismos de descubrimiento automático de vecinos para simplificar la configuración.

Mensajes BGP

- Los **mensajes de BGP** son cuatro y cada uno cumple una función específica dentro del establecimiento y mantenimiento de la sesión BGP y el intercambio de información de enrutamiento:
 - **OPEN**: Se utiliza para establecer una sesión BGP entre dos enrutadores una vez que se ha establecido la conexión TCP.
 - En este mensaje se negocian **parámetros esenciales** como: versión de BGP, número de SA local, tiempo de espera para la sesión, identificador del enrutador (Router ID).
 - **UPDATE**: Se usa para anunciar nuevas rutas o retirar rutas que ya no son válidas.
 - Se envía cada vez que hay un cambio en las rutas conocidas, ya sea una nueva mejor ruta o la supresión de una existente.
 - **Contiene**: prefijo de red que se anuncia o retira, atributos BGP que ayudan a los enrutadores a decidir la mejor ruta.
 - **Estructura**: longitud de lista de rutas retiradas, lista de prefijos que se retiran, longitud de sección de atributos (para ruta anunciada), ruta anunciada (sus atributos BGP), prefijos de ruta anunciada.

Mensajes BGP

- **Mensajes BGP – Cont:**

- **KEEPALIVE:** Una vez establecida la sesión BGP, se envían periódicamente mensajes KEEPALIVE para confirmar que ambos extremos siguen activos y mantener viva la sesión.
 - Estos mensajes solo sirven para mantener la conexión activa y evitar que se cierre por inactividad.
 - El intervalo de envío se negocia en el mensaje OPEN.
- **NOTIFICATION:** Se envía cuando ocurre un error que requiere cerrar la sesión BGP.
 - Este mensaje indica la causa del error (como recepción de un mensaje mal formado, fallo en la sesión, timeout, etc.) y tras enviarlo se cierra la conexión TCP.
 - También se usa para informar condiciones inusuales o problemas en la sesión

Mensajes BGP

En resumen, el flujo típico es:

- Se establece la conexión TCP.
- Se intercambian mensajes OPEN para negociar parámetros.
- Se envían KEEPALIVE periódicos para mantener la sesión.
- Se intercambian UPDATE para anunciar o retirar rutas.
- Si hay errores, se envía NOTIFICATION y se cierra la sesión.

Tablas y estructuras internas de BGP

- Las **tablas y estructuras internas** que usa BGP para gestionar rutas son:
 - **Adj-RIB-In:** Es la tabla donde el enrutador almacena todas las rutas recibidas de cada vecino BGP.
 - Actúa como la "bandeja de entrada" de rutas, almacenando la información recibida sin modificar antes de aplicar políticas o selección.
 - Hay una Adj-RIB-In por vecino, y dentro de cada una se organizan las rutas por prefijo, junto con sus atributos BGP.
 - Se actualiza dinámicamente con cada mensaje UPDATE recibido: Esto permite reflejar cambios en la topología o políticas de los vecinos.
 - Si llega una ruta por primera vez para un prefijo de un vecino: se la agrega.
 - Si llega una ruta actualizada para un prefijo de un vecino (p.ej. cambió el AS_PATH o el NEXT_HOP): la ruta actualizada reemplaza la anterior para ese vecino.
 - Si llega un retiro, la ruta correspondiente se elimina de la tabla.

Tablas y estructuras internas de BGP

- **Tablas y estructuras internas de BGP – cont:**

- **Loc-RIB:** Contiene las mejores rutas seleccionadas para cada prefijo tras aplicar el **algoritmo de selección BGP y políticas de entrada** sobre las rutas almacenadas en las Adj-RIB-In.
 - La Loc-RIB representa la **visión local del router sobre las mejores rutas BGP disponibles** y es la base para anuncios y para la instalación en la tabla RIB.
 - **El proceso BGP** compara las múltiples rutas recibidas para un prefijo desde las Adj-RIB-In. Se divide en:
 - Aplica filtros y modifica atributos según políticas de entrada.
 - A continuación se selecciona la mejor ruta según diversos criterios.
 - Finalmente la ruta seleccionada se almacena en la Loc-RIB.

Tablas y estructuras internas de BGP

- **Tablas y estructuras internas de BGP – cont:**
 - **Adj-RIB-Out:** Contiene las rutas que el enrutador va a anunciar a cada vecino BGP.
 - Esta tabla controla qué rutas se envían a cada vecino mediante mensajes UPDATE.
 - **Preparación de rutas para anuncio:**
 - Se construye a partir de la Loc-RIB aplicando políticas de salida. Las mismas pueden modificar atributos de la ruta, filtrar rutas para que no se anuncien a ciertos vecinos, controlar qué rutas se anuncian a qué vecinos.
 - Puede haber diferencias en las rutas anunciadas a distintos vecinos, incluso para el mismo prefijo.
 - **RIB (Routing Information Base):** Contiene las rutas activas usadas para el reenvío de paquetes.
 - Solo las rutas seleccionadas en la Loc-RIB que son mejores que las rutas existentes en la RIB se instalan en la RIB para ser usadas en el reenvío de paquetes.
 - Las rutas de Loc-RIB se comparan con otras rutas en la RIB para decidir cuál ruta se instala para el encaminamiento.

Atributos BGP

- **Otros atributos BGP** influyen en las políticas y la selección de rutas a un prefijo. Estos son:
 - **LOCAL_PREF**: es un valor usado para controlar la preferencia de la ruta dentro del sistema autónomo.
 - **MED**: Influye en la preferencia entre múltiples puntos de entrada a un AS.
 - MED influye en cómo los sistemas autónomos (AS) vecinos envían tráfico hacia tu AS cuando existen múltiples puntos de entrada/salida.
 - Indica a los AS vecinos cuál es el mejor punto de entrada para alcanzar un prefijo dentro de tu AS. Minimiza la latencia o costos al preferir enlaces más cercanos o de mayor capacidad.
 - La idea es que se va a elegir la ruta con el MED más bajo.
 - Solo se compara entre rutas del mismo AS vecino.

Atributos BGP

- **Otros atributos BGP – cont:**

- **COMMUNITY (comunidad):** Es un atributo opcional que actúa como una etiqueta para agrupar rutas y aplicar políticas dinámicas.

- **Propósito:**

- Agrupar prefijos con características comunes (ejemplo: rutas de un cliente específico).
- Automatizar políticas: Aplicar acciones como filtrado, modificación de atributos o control de propagación basado en etiquetas.

- **Formato:** valor de 32 b representado como SA:valor (ejemplo: 64500:100)

- **Ejemplos:**

- NO_EXPORT (0xFFFFF01): No anunciar la ruta fuera del AS local.
- NO_ADVERTISE (0xFFFFF02): No anunciar a ningún vecino.
- INTERNET (0xFFFFF00): Anunciar la ruta públicamente.

- **Casos de uso:**

- Filtrado selectivo: Un proveedor puede usar comunidades para que los clientes controlen qué rutas se anuncian a otros sistemas autónomos.
- Priorización: asignar LOCAL_REF alto a rutas con una comunidad específica.
- Marcar rutas para preferir enlaces de bajo costo.

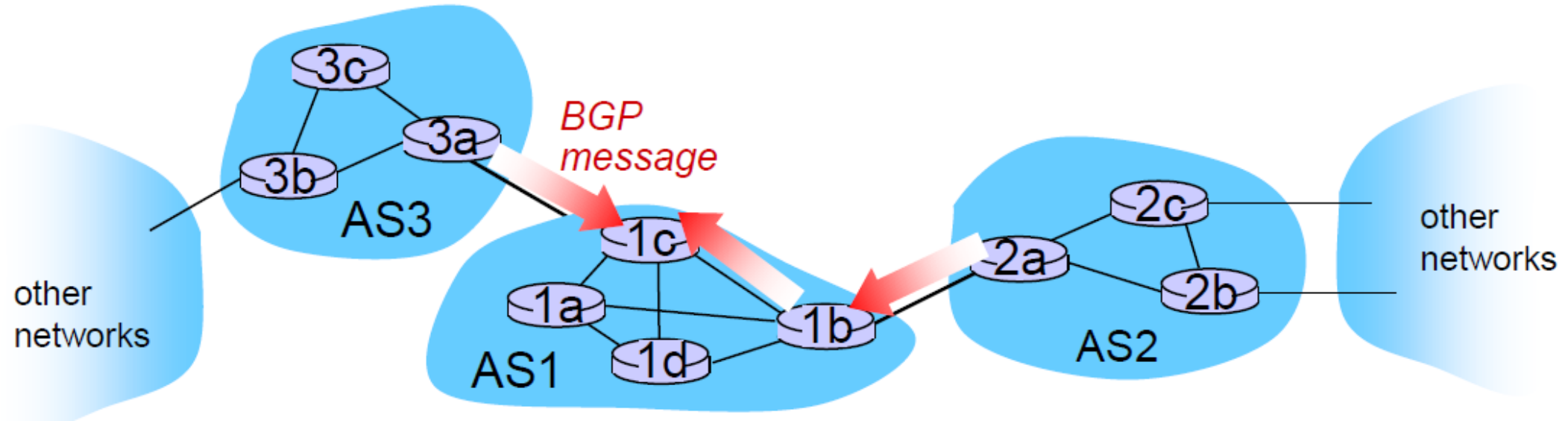
Políticas BGP

- Las **políticas en BGP** son reglas que controlan cómo se aceptan, modifican y anuncian las rutas entre vecinos.
- En conjunto, estas políticas permiten controlar el flujo de información de enrutamiento, optimizar rutas, evitar bucles, cumplir acuerdos comerciales y mantener la estabilidad y seguridad de la red.

Políticas BGP

- **Clasificación de las políticas BGP:**
 - **Políticas de entrada:** Se aplican a las rutas recibidas de un vecino BGP antes de almacenarlas en la tabla local (Loc-RIB). Su función es:
 - **Filtrar rutas no deseadas** (bloquear prefijos o rutas con ciertos atributos – por ejemplo, rechazar prefijos con AS-PATH que incluye AS no confiables).
 - **Modificar atributos** para influir en la selección de ruta (por ejemplo, incrementar el atributo LOCAL_PREF para dar preferencia a ciertas rutas recibidas de un vecino sobre otras).
 - **Clasificar rutas mediante comunidades** para aplicar reglas internas.
 - **Políticas de salida:** Se aplican a las rutas que el enrutador va a anunciar a sus vecinos. Sirven para:
 - **Filtrar rutas que no se desean enviar a ciertos vecinos.**
 - **Modificar atributos** como AS_PATH: Añadir el ASN local al AS_PATH antes de anunciar rutas a un vecino externo para evitar bucles y cumplir con las reglas BGP.
 - **Controlar anuncios mediante comunidades** para segmentar el comportamiento según acuerdos o topología.

Elección entre las rutas a un prefijo



- ❖ **Situación:** Un enrutador puede recibir múltiples rutas al mismo prefijo.
- ❖ La **mejor ruta** a un prefijo debe guardarse en la BIE.
- ❖ **Problema:** ¿Cómo escoge el enrutador una de esas rutas al mismo prefijo?

Elección entre las rutas a un prefijo

- **Solución:** Se usa el siguiente algoritmo para la selección de la mejor ruta BGP entre las rutas que están en Adj-RIB-In:
 1. **Verificar que el NEXT_HOP es alcanzable:** el NEXT_HOP debe poder resolverse en la tabla de enrutamiento local. Rutas con NEXT_HOP inalcanzable se descartan.
 2. **LOCAL_PREF:** Las rutas con el *mayor valor de preferencia local* son elegidas.
 - A las rutas se les asigna un **valor de preferencia local** que puede haber sido fijado por el enrutador o aprendido de otro enrutador en el mismo SA (esto lo define el administrador del SA).
 - LOCAL_PREF es un atributo propagado dentro del AS que influye en la selección interna.
 3. **Longitud del AS_PATH:** De las rutas restantes, la ruta con el *camino AS-PATH más corto* es elegida (la métrica es la cantidad de saltos SA).
 4. **MED:** Se prefiere la ruta con el MED más bajo. Por defecto, se compara solo entre rutas del mismo AS vecino.

Elección entre las rutas a un prefijo

- **Solución – cont:**

5. Se prefiere la ruta aprendida vía eBGP frente a iBGP.
6. **Costo IGP al NEXT_HOP:** De las rutas restantes la ruta con el *enrutador NEXT-HOP más cercano es elegida*;
 - o sea, se considera el enrutador NEXT-HOP con el camino más corto determinado por el algoritmo de enrutamiento intra-SA (a esto se lo llama **hot potato routing**).
7. **Ruta más antigua:** para evitar fluctuaciones, se prefiere la ruta aprendida primero.
8. **Router ID del vecino:** Se prefiere la ruta aprendida del vecino con el Router ID más bajo.
9. **Dirección IP del vecino:** se prefiere la ruta aprendida del vecino con la dirección IP más baja.

Elección entre las rutas a un prefijo

- Elegir ruta basada en el AS-PATH más corto

❖ Ejemplo:

- ❖ AS2 AS17 to 138.16.64/22
- ❖ AS3 AS131 AS201 to 138.16.64/22

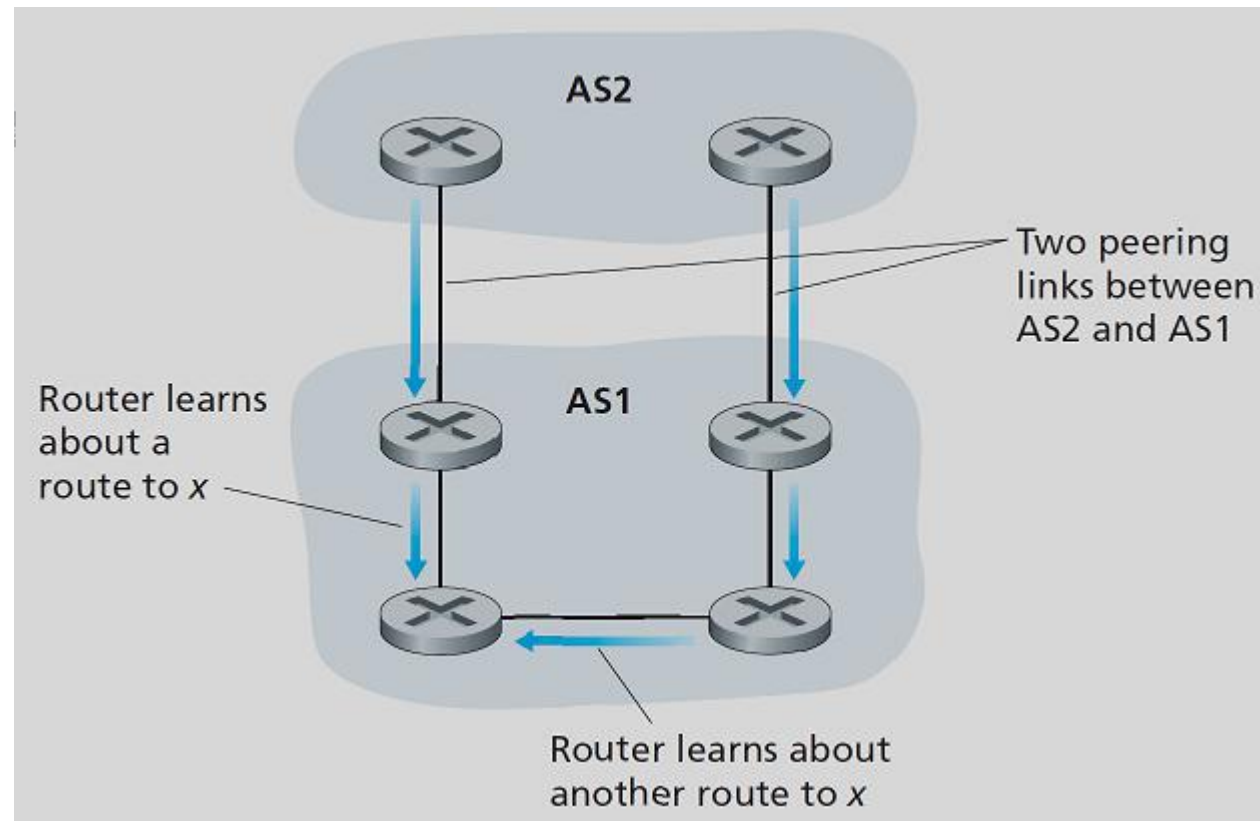
select



Elección entre las rutas a un prefijo

Key:

← Route advertisements message for destination x



Ejercicio:

- Un enrutador en AS1 puede aprender acerca de dos rutas al mismo prefijo x .
- Estas dos rutas pueden tener el mismo AS-PATH a x, pero pueden tener diferentes valores de NEXT-HOP correspondientes a los distintos enlaces de compañerismo ente AS1 y AS2. **¿Cómo elegir entre las rutas a usar?**
- **Solución:** Usando los valores de NEXT HOP y el algoritmo de enrutamiento intra-SA el enrutador puede determinar el costo del camino a cada enlace de compañerismo,
 - y luego entre ellos elegir el camino de menor costo para determinar el enlace de compañerismo apropiado.

Comparación de la ruta en la Loc-RIB con la ruta existente en la RIB

- Las **rutas iBGP** son las rutas intercambiadas entre enrutadores que pertenecen al mismo sistema autónomo.
- Las **rutas eBGP** son las rutas intercambiadas entre enrutadores que pertenecen a diferentes sistemas autónomos.
- Las **rutas estáticas** son rutas configuradas manualmente por el administrador de red en cada enrutador.
 - Indican explícitamente el camino que deben seguir los paquetes hacia una red destino, especificando el siguiente salto o la interfaz de salida.
 - No cambian automáticamente; si hay un cambio en la topología, el administrador debe actualizar las rutas manualmente.
- Las **rutas IGP** son rutas aprendidas automáticamente mediante protocolos de enrutamiento internos al sistema autónomo (p.ej: RIP, OSPF, EIGRP.)

Comparación de la ruta en la Loc-RIB con la ruta existente en la RIB

- El criterio usado para comparar una ruta en la Loc_RIB con la ruta existente en la RIB es la **distancia administrativa** y la **métrica del protocolo**.
- La **distancia administrativa** es un valor que indica la preferencia relativa entre rutas aprendidas por diferentes protocolos (BGP, OSPF, etc.)
 - Por ejemplo, las rutas eBGP tienen una distancia administrativa por defecto de 20, las rutas iBGP de 200, y las rutas estáticas o IGP's tienen otros valores.
- La ruta con menor distancia administrativa se prefiere para instalarse en la RIB y usarse para el reenvío.

Procesamiento y flujo de rutas en BGP

- El procesamiento y flujo de rutas en BGP sigue un ciclo estructurado que involucra varias tablas internas y operaciones clave para recibir, seleccionar, almacenar y anunciar rutas.
- **Etapas principales:**
 1. **Recepción de rutas y almacenamiento en Adj-RIB-In:** Cuando un enrutador BGP recibe un mensaje UPDATE de un vecino, las rutas anunciadas se almacenan en la tabla Adj-RIB-In correspondiente a ese vecino.
 - Cada vecino tiene su propia Adj-RIB-In, que contiene todas las rutas recibidas de ese vecino, organizadas por prefijo y con sus atributos BGP.
 - Esta tabla funciona como una "bandeja de entrada" donde se guarda la información recibida sin modificar antes de aplicar políticas o selección

Procesamiento y flujo de rutas en BGP

- **Etapas principales – cont:**

- 2. **Aplicación de políticas de entrada y selección de la mejor ruta:**

- Las rutas en Adj-RIB-In pasan por las políticas de entrada, que pueden filtrar rutas no deseadas o modificar atributos.
 - Luego, el proceso BGP ejecuta el algoritmo de selección de la mejor ruta, que compara todas las rutas recibidas para un mismo prefijo desde todos los vecinos.
 - Solo la ruta mejor seleccionada para cada prefijo se mantiene para uso posterior.

- 3. **Almacenamiento en Loc-RIB:**

- La ruta seleccionada se almacena en la Loc-RIB, que contiene las mejores rutas BGP conocidas por el router.
 - La Loc-RIB representa la visión local del router sobre las rutas óptimas, y es la base para anuncios y para la instalación en la tabla de enrutamiento del sistema (RIB)

Procesamiento y flujo de rutas en BGP

- **Etapas principales – cont:**

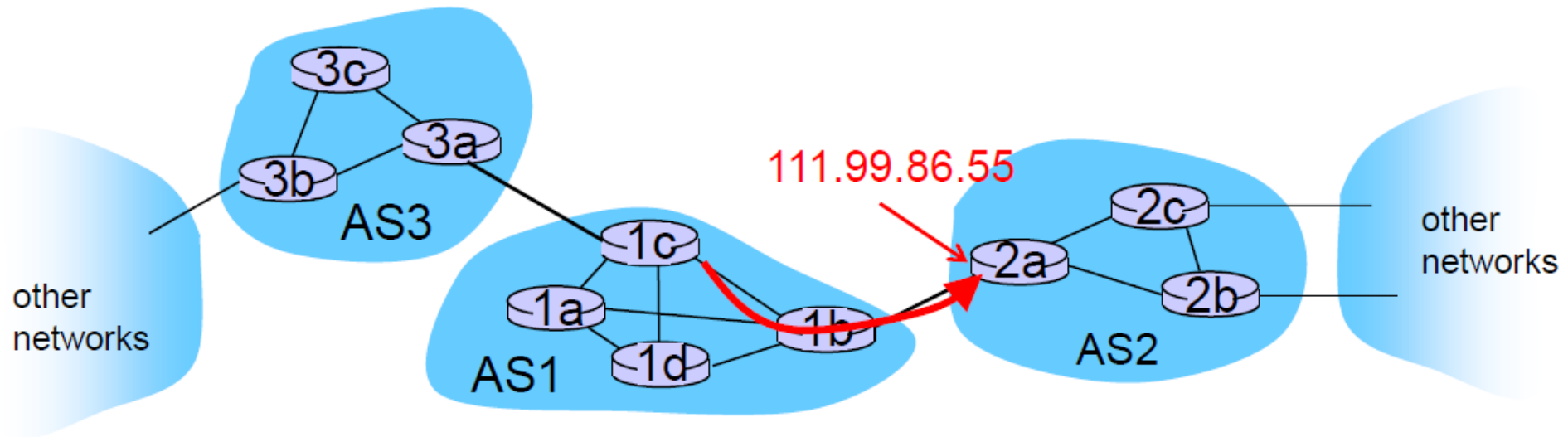
- 4. Aplicación de políticas de salida y preparación de Adj-RIB-Out:**

- A partir de las rutas en Loc-RIB, el enrutador aplica las políticas de salida para cada vecino, que pueden modificar atributos o filtrar rutas que no se desean anunciar.
 - Las rutas resultantes se almacenan en la Adj-RIB-Out para cada vecino, que contiene las rutas que efectivamente se van a anunciar a ese vecino.

- 5. Anuncio de rutas a vecinos:** Finalmente, el contenido de la Adj-RIB-Out se usa para enviar mensajes UPDATE a cada vecino, anunciando las rutas que el enrutador ha decidido compartir.

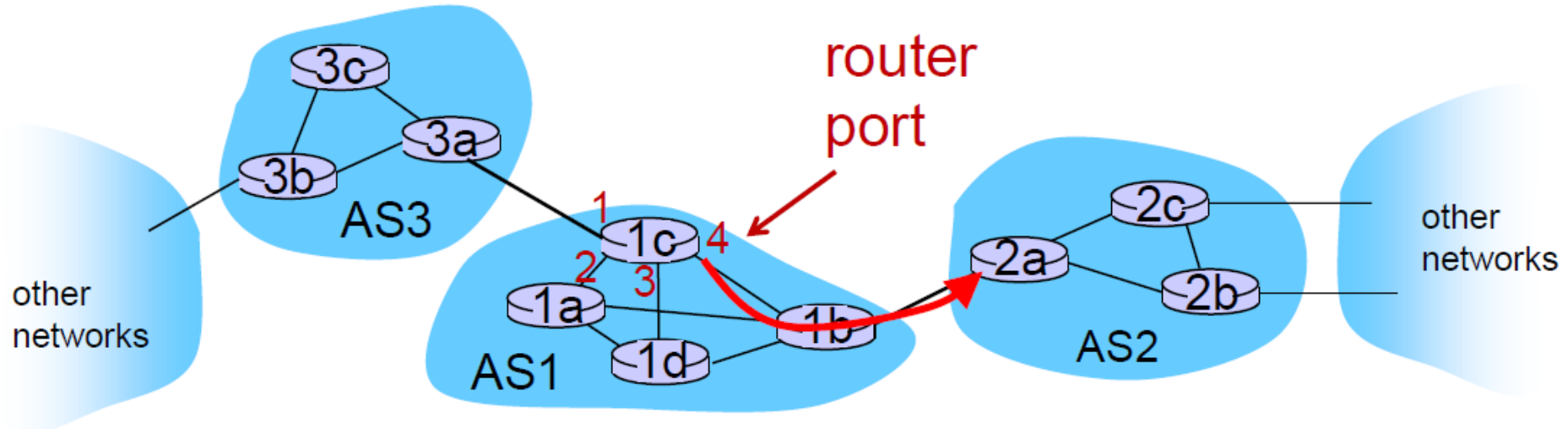
Encontrar la mejor intra-ruta a una ruta BGP

- **Ejercicio:** Dado el aviso de ruta al sistema autónomo AS1
 - Prefijo 138.16.64/22, AS-PATH: AS2 AS17 ; NEXT-HOP: 111.99.86.55
 - Supongamos que es el único aviso de ruta a ese prefijo y que respeta las políticas de AS1.
 - ¿Cómo hace el enrutador 1c para poner entrada para el prefijo anterior en su tabla de reenvío?



Encontrar la mejor intra-ruta a una ruta BGP

- ❖ **Solución:** 1c provee la dirección IP del atributo NEXT-HOP al algoritmo OSPF.
- ❖ 1c usa OSPF para encontrar el camino más corto de 1c a la subred para el enlace entre 1b y 2a (que tiene **111.99.86.55**).
- ❖ Supongamos que el puerto de 1c a lo largo de ese camino más corto es el puerto 4.
- Entonces **1c** agrega el puerto 4 de entrada para el prefijo de la red de destino a su tabla de reenvío:
 - (138.16.64/22 , port 4)



Cómo se hace una entrada en tabla de reenvío

Problema: ¿cómo un enrutador hace una entrada en la tabla de reenvío de un prefijo x perteneciente a otro SA?

Solución:

1. La tabla RIB contiene la ruta óptima R al prefijo x a considerar para la tabla de reenvío.
2. Determinar el puerto de salida del enrutador para el prefijo x :
 - Usar OSPF para encontrar la mejor ruta intra-SA que lleva a NEXT_HOP de R .
 - El enrutador identifica el puerto de salida del enrutador para esa mejor ruta.
3. Ingresar el puerto del prefijo en la tabla de reenvío
 - La tabla de reenvío se **sincroniza continuamente** con la RIB para reflejar cambios en la topología o en las rutas.