Capítulo 4: la ruta

4.1 Introducción
4.2 Circuito virtual y redes de datagramas
4.3 ¿Qué hay dentro de un enrutador?
4.4 IP: Protocolo de Internet
- Formato del datagrama
- Direccionamiento IPv4
- Protocolo ICMP
- Direccionamiento IPv6

4.5 Algoritmos de enrutamiento
- Estado del enlace
- Vector de distancia
- Enrutamiento jerárquico

4.6 Enrutamiento en Internet
- RIP
- OSPF
- BGP

4.7 enrutamiento de difusión y multidifusión

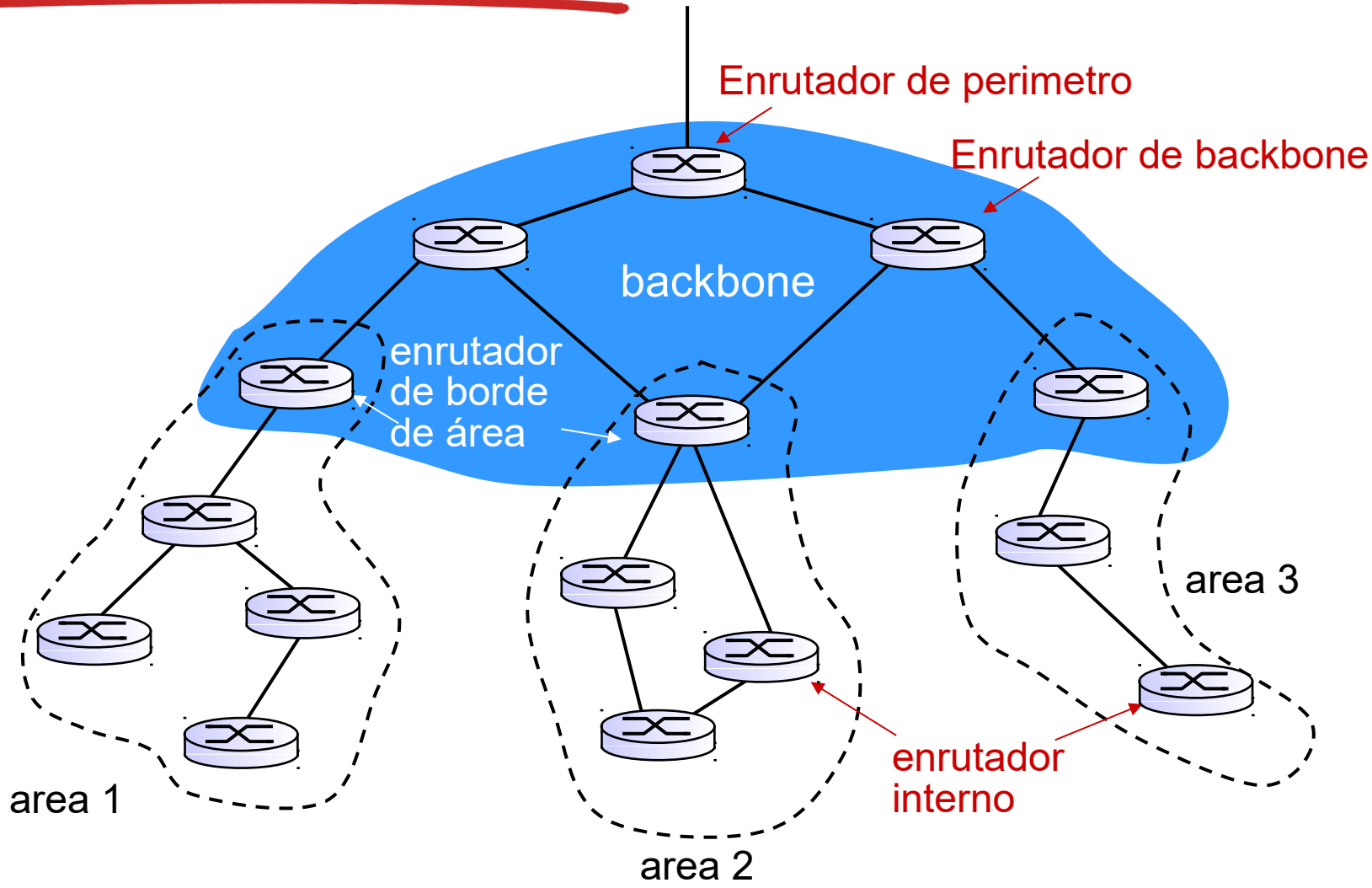
OSPF (Open Shortest Path First)

- "abierto": disponible publicamente
- usa un algoritmo de estado de enlace
  - difusión de paquetes con LS
  - mapa de topología en cada nodo
  - cálculo de ruta usando el algoritmo de Dijkstra
- los anuncios OSPF llevan una entrada por vecino
- los anuncios son inundados a *todo* el AS
  - llevado en los mensajes OSPF directamente sobre IP (en lugar de TCP o UDP)
- el protocol IS-IS de enrutamiento: casi idéntico al OSPF

OSPF "advanced" features (not in RIP)

- *seguridad:* todos los mensajes de OSPF son autenticas (para prevenir intrusiones malintencionadas)
- múltiples rutas de mismo costo permitidas (solo una ruta en RIP)
- para cada enlace, métricas de costos múltiples para diferentes TOS (por ejemplo, costo de enlace de satélite establecido "bajo" para el mejor esfuerzo ToS; alto para ToS en tiempo real)
- soporte uni y multicast integrado:
  - Multidifusión OSPF (MOSPF) utiliza la misma base de datos de topología que OSPF
- OSPF jerárquico en dominios grandes.

OSPF jerárquico



OSPF jerárquico

- *2 niveles de jerarquía:* área local, núcleo o backbone.
  - anuncios de enlace de estado (link-state) en el área local
  - cada nodo tiene la topología de detallada del área, sólo sabe la ruta más corta a redes en otras áreas.
- *Enrutadores del borde de área:* "summariza" las distancias a otras redes en su propia área, anuncia los enrutadores de borde.
- *Enrutadores de núcleo o backbone:* ejecuta enrutamiento de OSPF limitado al núcleo o backbone.
- *Enrutados de borde:* conecta a la red con otros AS.

Capítulo 4: la ruta

4.1 Introducción
4.2 Circuito virtual y redes de datagramas
4.3 ¿Qué hay dentro de un enrutador?
4.4 IP: Protocolo de Internet
- Formato del datagrama
- Direccionamiento IPv4
- Protocolo ICMP
- Direccionamiento IPv6

4.5 Algoritmos de enrutamiento
- Estado del enlace
- Vector de distancia
- Enrutamiento jerárquico

4.6 Enrutamiento en Internet
- RIP
- OSPF
- BGP

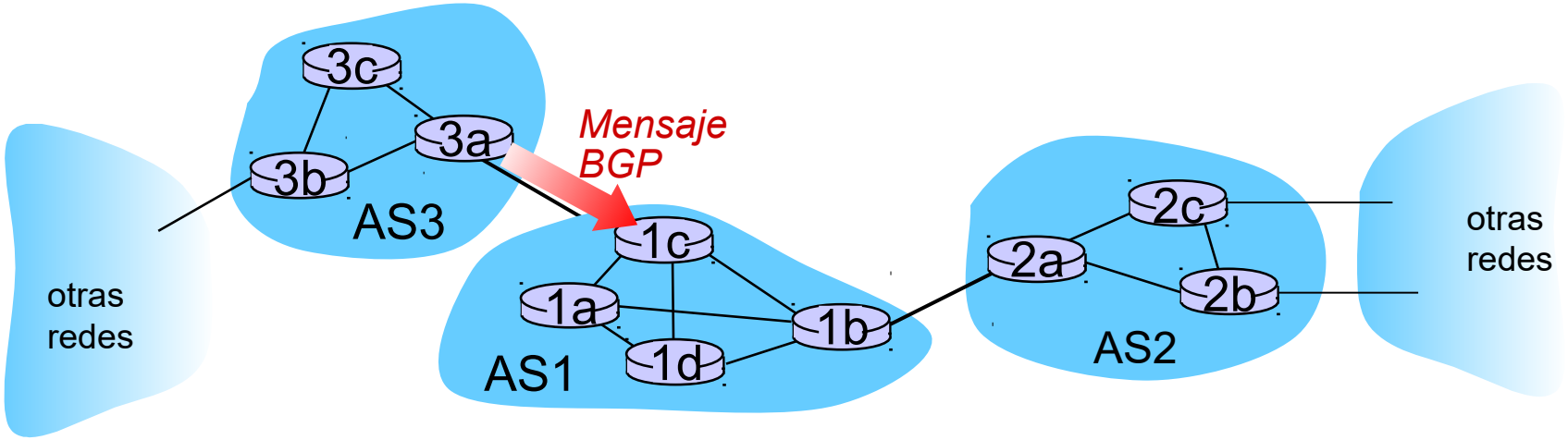
4.7 enrutamiento de difusión y multidifusión

Enrutamiento entre AS (inter-AS): BGP

- BGP (Border Gateway Protocol): es el protocolo inter dominio por omisión en este momento en Internet
  - "es la pega que mantiene a la Internet unidad"
- BGP provee a cada AS los mecanismos para:
  - eBGP: obtener información de accesibilidad de subred de los AS vecinos.
  - iBGP: propagar información de accesibilidad a todos los enrutadores internos del AS (intra-AS).
  - determinar rutas "buenas" a otras redes basadas en información y políticas de accesibilidad.
- permite a la subred anunciar su existencia al resto de Internet: "Estoy aquí"

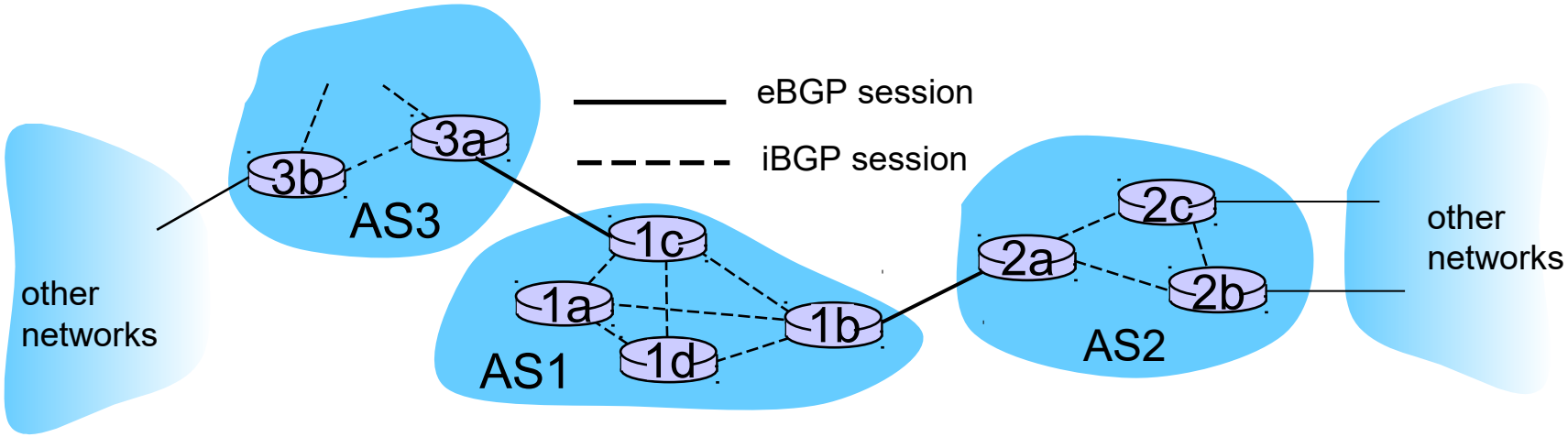
BGP: básico

- sesión BGP: los enrutadores BGP intercambian mensajes BGP entre sus "pares":
  - anuncio de rutas a diferentes prefijos de red de destino (protocolo "vector de ruta")
  - intercambio a través de conexiones TCP semi-permanentes
- cuando AS3 anuncia prefijos a AS1:
  - AS3 *le promete* que re-enviará los datagramas a través de ese prefijo
  - AS3 puede agregar prefijos en sus anuncios



BGP: distribuyendo información sobre las rutas

- usando una sesión eBGP entre 3a y 1c, AS3 envía su prefijo de accesabilidad a AS1.
  - 1c puede usar iBGP para distribuir nueva información de prefijo a todos los enrutadores en AS1
  - 1b puede volver a anunciar nueva información de accesibilidad a AS2 a través de la sesión eBGP entre 1b-2a
  - 1b puede volver a anunciar la nueva información de accesibilidad a AS2 puede volver a anunciar la nueva información de accesibilidad a AS2 durante 1b-to-2a sesión eBGP de 1b-to-2a
- cuando un enrutador aprende un nuevo prefijo, se crea un nueva entrada en la tabla de enrutamiento



Atributos de ruta y rutas BGP

- Los prefijos anunciados incluyen atributos de BGP
  - prefijos + atributos = "ruta"
- 2 atributos muy importantes:
  - AS-PATH: contiene los AS a través de los cuales ha pasado el anuncio de prefijo: ejemplo: AS 67, AS 17
  - NEXT-HOP (próximo-salto): indica el enrutador AS-interno específico al AS-del-próximo-salto (pueden ser enlaces múltiples desde AS actual hasta AS-del-próximo-salto).
- el enrutador de puerta de enlace que recibe el anuncio de ruta usa una política de importación para aceptar/rechazar
  - Ejemplo., nunca enrutes por el ASx
  - enrutamiento *basado en políticas*

Selección de la ruta BGP

- Un enrutador puede aprender más de una ruta para llegar a AS, así que selección la ruta basado en:
  1. el atributo de preferencia local: decisión basada en políticas
  2. el AS-PATH más corto
  3. el enrutador de NEXT-HOP más cercano: enrutamiento de papa caliente
  4. criterios adicionales

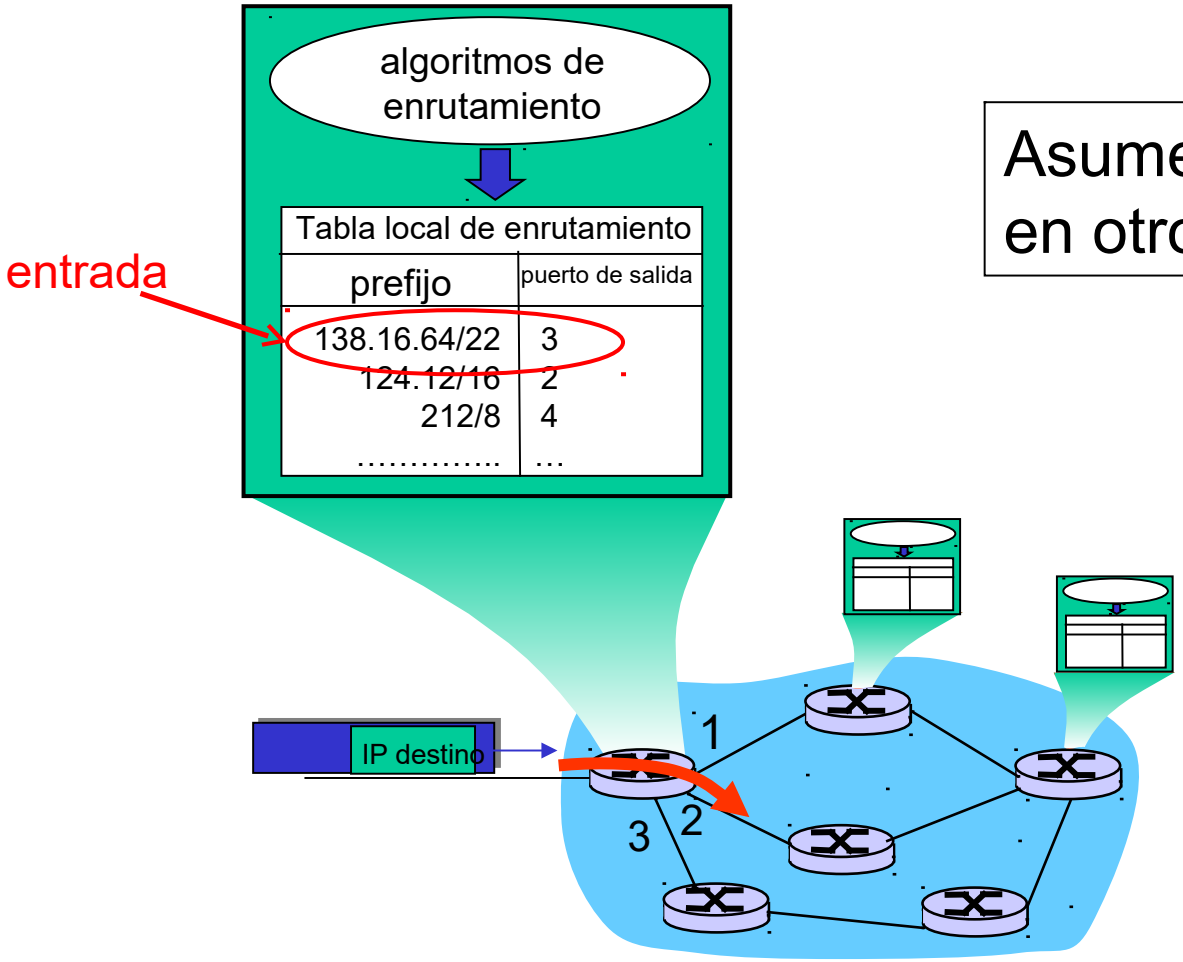
BGP: mensajes

- BGP los mensajes son intercambiados entre pares sobre una conexión TCP
- mensajes BGP:
  - OPEN: abre una conexión TCP a su par y autentica al remitente
  - UPDATE: anuncia la nueva ruta (or withdraws old)
  - KEEPALIVE: mantiene la conexión viva en ausencia de actualizaciones; también usa solicitudes ACKs y OPEN
  - NOTIFICATION: reporta errores en mensajes previos; también se usa para cerrar la conexión

Poniéndolo todo junto: ¿Cómo entra un nuevo registro en una tabla de enrutamiento?

- la respuesta es complicada!
- une todos los ingredientes: enrutamiento jerárquico (sección 4.5.3) con BGP (sección 4.6.3) y OSPF (sección 4.6.2).
- esto provee una buen recuento de BGP!

¿Cómo entra un nuevo registro en una tabla de enrutamiento?



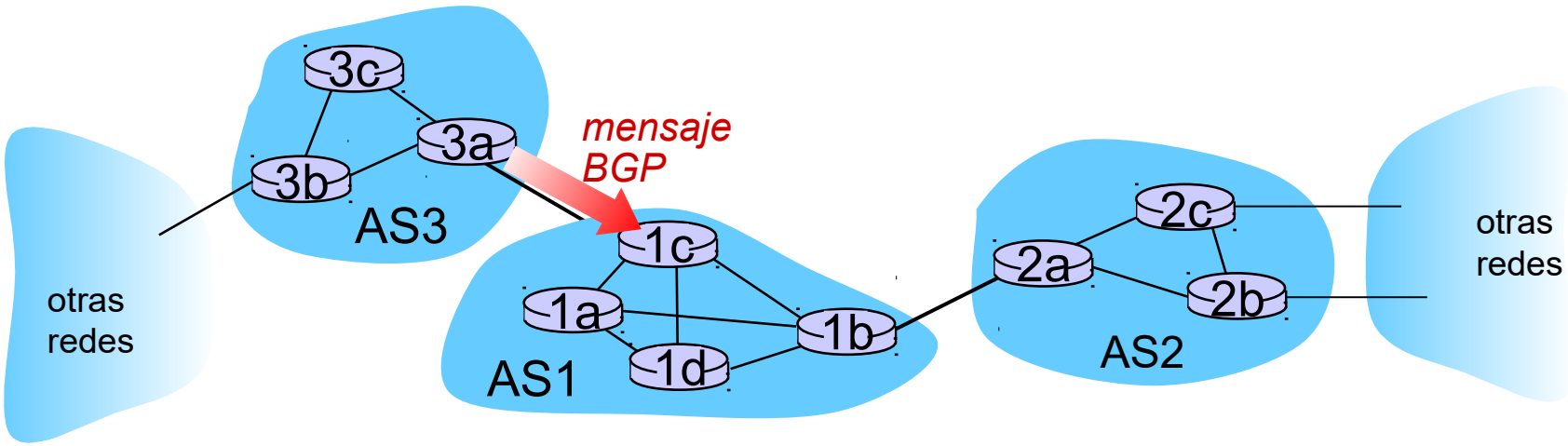
Asume el prefijo en otro AS.

¿Cómo entra un nuevo registro en una tabla de enrutamiento?

Revisión de alto-nivel

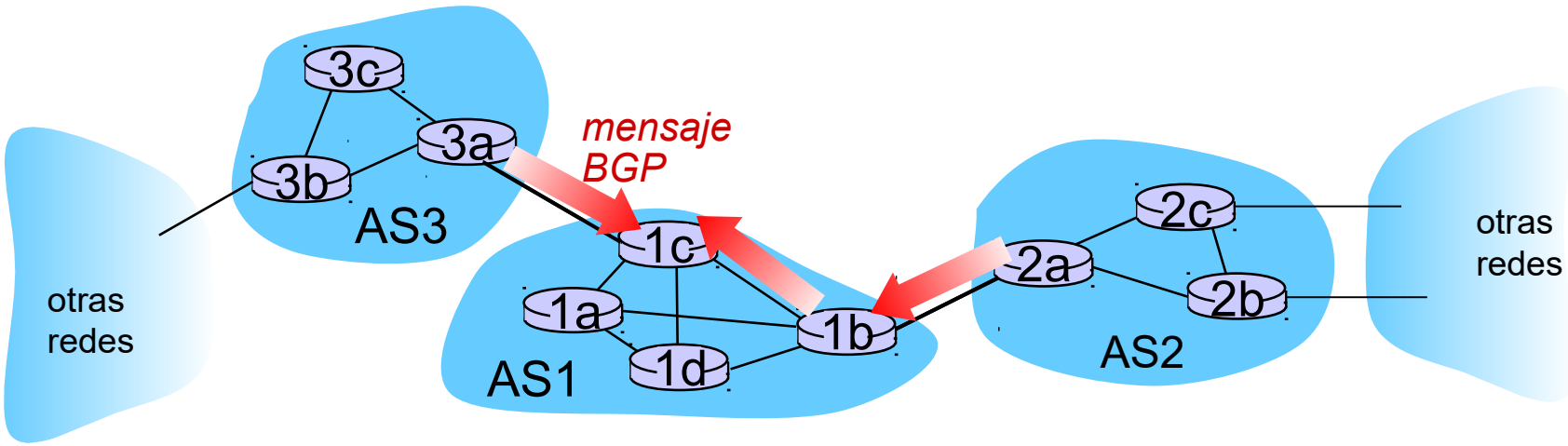
1. Enrutador se da cuenta del prefijo
2. El enrutador determina el puerto de salida del prefijo
3. El enrutador registra la ruta entra en la tabla de enrutamiento

Enrutador se da cuenta del prefijo



- El mensaje BGP contiene las "rutas"
- "ruta" es un prefijo y sus atributos: AS-PATH, NEXT-HOP,…
- ejemplo: ruta:
  - Prefijo:138.16.64/22 ; AS-PATH: AS3 AS131 ; NEXT-HOP: 201.44.13.125

Un enrutador puede recibir múltiples enrutadores



- un enrutador recibe múltiple rutas para el mismo prefijo
- tiene que seleccionar una sola ruta

Selecciona la mejor ruta para el prefijo BGP

- El enrutador selecciona la ruta basado en el atributo de la ruta más corta AS-PATH

- Ejemplo:

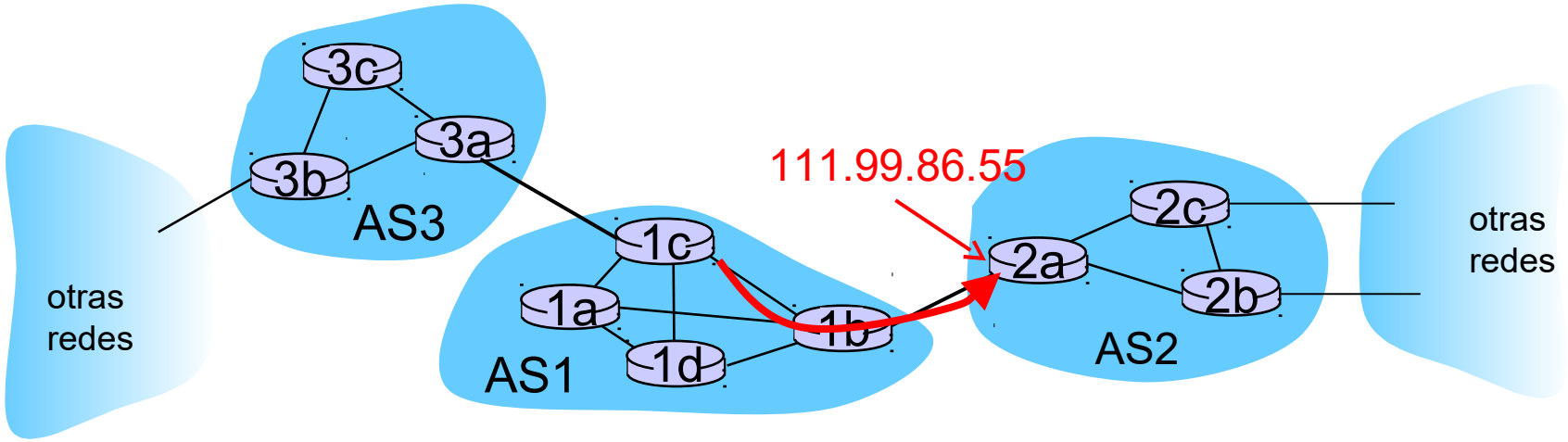
seleccionado

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

- ¿Y qué pasa si hay un empate? Ya lo vamos a ver!

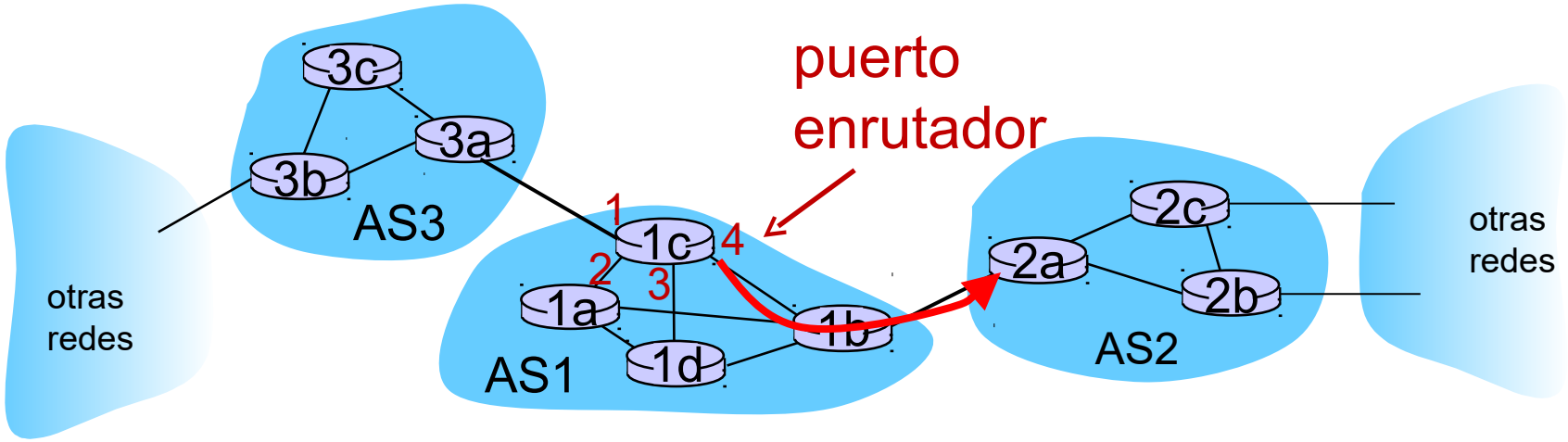
Buscando la mejor ruta interna BGP

- Usa la ruta seleccionada por el atributo NEXT-HOP
  - El atributo NEXT-HOP de la ruta es la dirección IP de la interfaz del enrutador que comienza AS PATH
- ejemplo:
  - AS-PATH: AS2 AS17; NEXT-HOP: 111.99.86.55
- El enrutador usa OSPF para buscar el camino más corto de 1c a 111.99.86.55



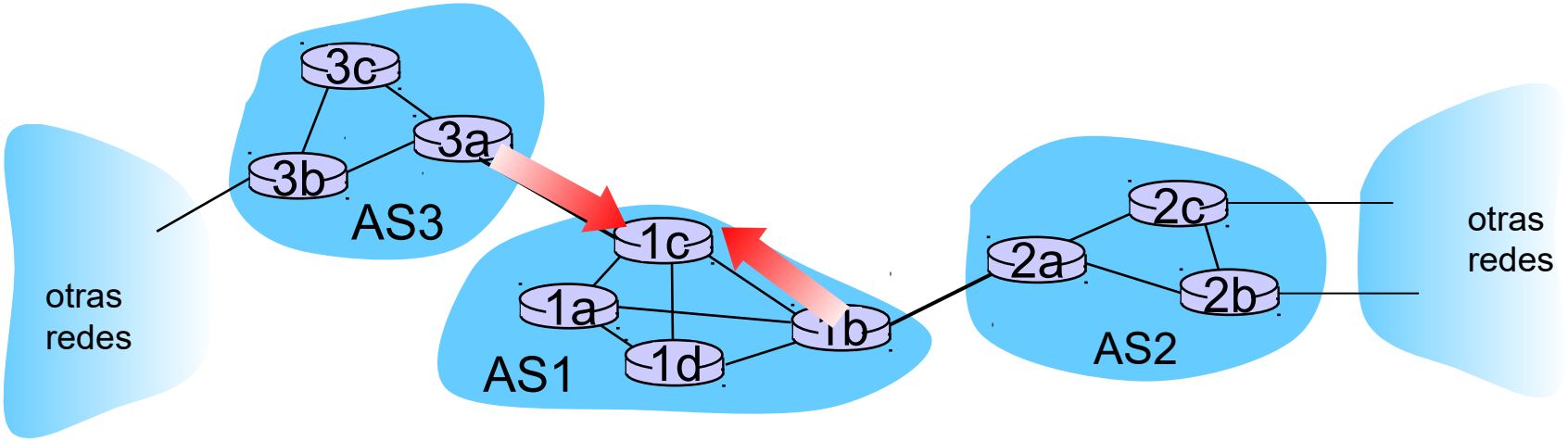
El enrutador identifica el puerto para la ruta

- Identifica el puerto OSPF con el camino más corto
- Agrega el puerto-prefijo a la tabla de enrutamiento:
  - (138.16.64/22 , puerto 4)



Enrutamiento de la papa caliente

- Supongamos que hay 2 o más, mejores rutas Inter-AS
- Luego escojemos la ruta con el NEXT-HOP más cercano
  - usa OSPF para determinar cual es el gateway más cercano
  - P: desde 1c, escoge AS3 AS131 ó AS2 AS17?
  - R: enrutador AS3 AS201 ya que es el más cercano



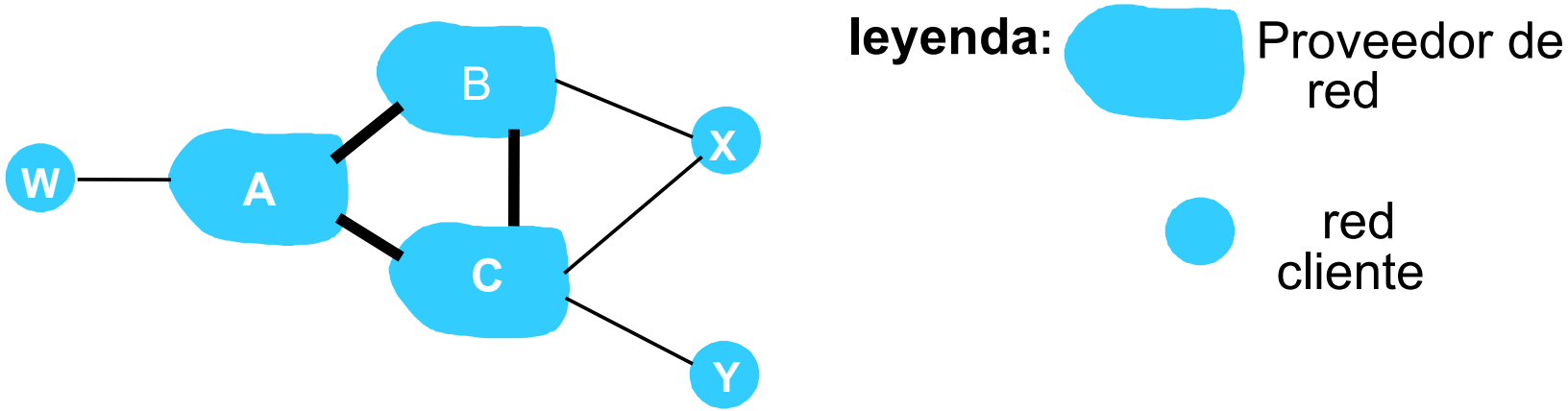
¿Cómo entra un nuevo registro en una tabla de enrutamiento?

Resumen

1. Enrutador se da cuenta del prefijo
   - Vía anuncios de BGP el enrutador se da a conocer a otros enrutadores
2. El enrutador determina el puerto de salida del prefijo
   - Usa BGP para seleccionar encuentra la mejora ruta inter-AS
   - Usa OSPF para encontrar la mejor ruta intra-AS que a su vez lleva la mejor ruta inter-AS
   - El enrutador identifica el puerto para esa mejor ruta
3. El enrutador registra la ruta en la tabla de enrutamiento

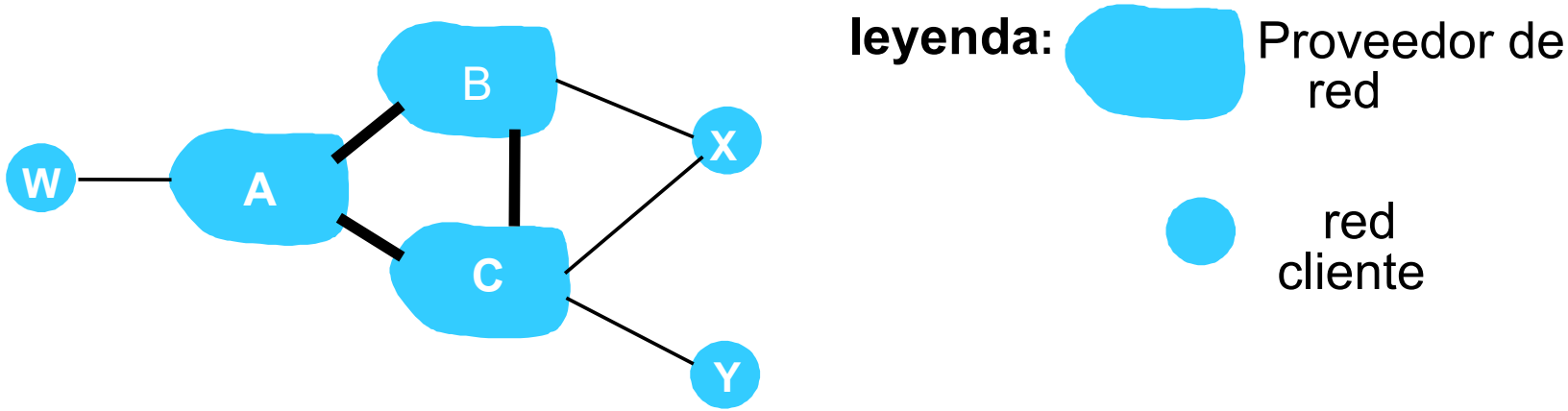
Kurose and Ross, Computer Networking: A Top Down Approach, 6th edition, Addison-Wesley, 2012

Políticas de enrutamiento BGP



- A,B,C son *redes de proveedor*
- X,W,Y son clientes (de esas redes de proveedores)
- X tiene *doble-proveedor:* porque tiene 2 redes para acceder a internet
  - X no quiere enrutar a B vía X para llegar a C
  - .. por lo que X no anunciará la ruta a B de cómo llegar a C

Políticas de enrutamiento BGP (2)



- A anunca el camino AW para B
- B anunca el camino BAW para X
- podría B anunciar el camino BAW para C?
  - De ninguna manera! B no tendría ninguna ganancia por enrutar CBAW ya que ni W ni C son clientes de B
  - B quiere forzar a C enrutar w vía A
  - B quiere enrutar *sólo* desde/hacia sus clientes!

¿Porqué un enrutamiento diferente para Intra-AS e Inter-AS?

política:

- inter-AS: el administrador quiere control sobre cómo enrutar el tráfico que pasa a través de sus equipos, decide quien pasa por su red.
- intra-AS: una sola administración, así que no hay decisión que maneja en términos de políticas

escala:

- El enrutamiento jerárquico ahorra en tamaño de la tabla, reduce el tráfico de actualización

rendimiento:

- intra-AS: puede enfocarse en rendimiento
- inter-AS: políticas pueden prevalecer sobre el rendimiento

Capítulo 4: la ruta

4.1 Introducción
4.2 Circuito virtual y redes de datagramas
4.3 ¿Qué hay dentro de un enrutador?
4.4 IP: Protocolo de Internet
- Formato del datagrama
- Direccionamiento IPv4
- Protocolo ICMP
- Direccionamiento IPv6

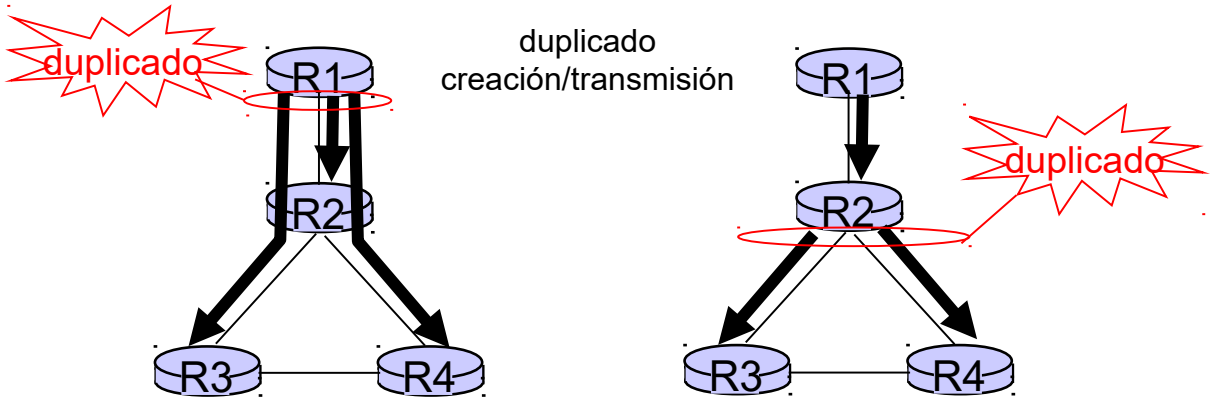
4.5 Algoritmos de enrutamiento
- Estado del enlace
- Vector de distancia
- Enrutamiento jerárquico

4.6 Enrutamiento en Internet
- RIP
- OSPF
- BGP

4.7 enrutamiento de difusión y multidifusión

Enrutamiento de difusión

- entrega paquetes desde la fuente a todos los otros nodos
- duplicación en la fuente es ineficiente:



duplicación en la fuente

duplication en la red

- duplicación en la fuente: ¿cómo hace la fuente para determinar la dirección del receptor?

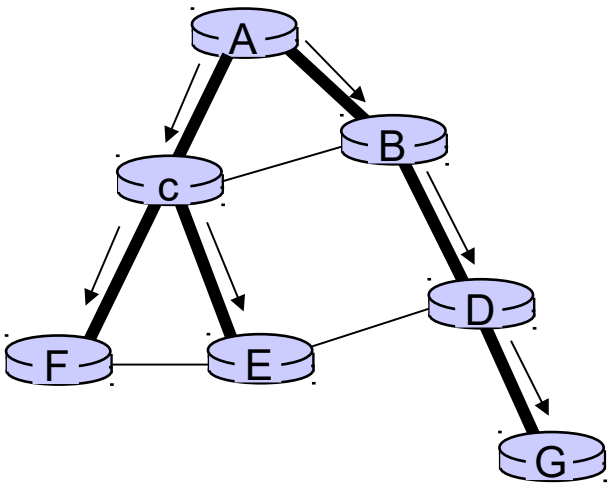
Duplicación en la red

- *inundación:* cuando el nodo recibe el paquete de difusión, envía copia a todos los vecinos
  - problemas: ciclos y tormenta de difusión
- *Inundación controlada:* nodo solo transmite el paquete si no ha transmitido el mismo paquete antes
  - nodo realiza un seguimiento de los identificadores de paquetes ya emitidos
  - o reenvío de ruta inversa (RPF): solo reenvía el paquete si llegó en la ruta más corta entre el nodo y la fuente
- *árbol de expansión:*
  - no hay paquetes redundantes recibidos por ningún nodo

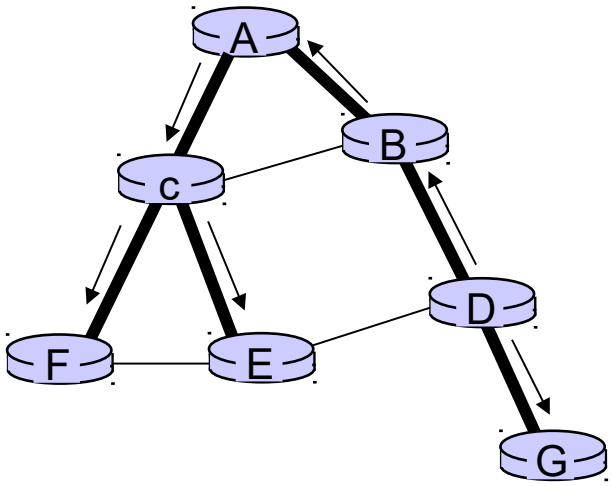
Kurose and Ross, Computer Networking: A Top Down Approach, 6th edition, Addison-Wesley, 2012

Árbol de expansión

- primero se construye un árbol de expansión
- Luego los nodos reenvian/hacen copias sólo a lo largo del árbol de expansión



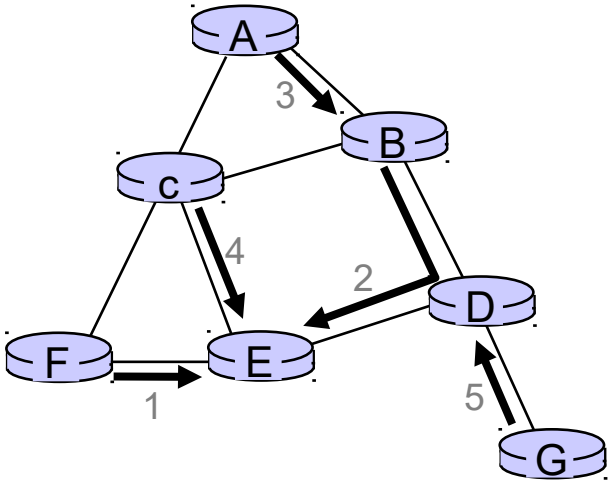
(a) difusión comenzada por A



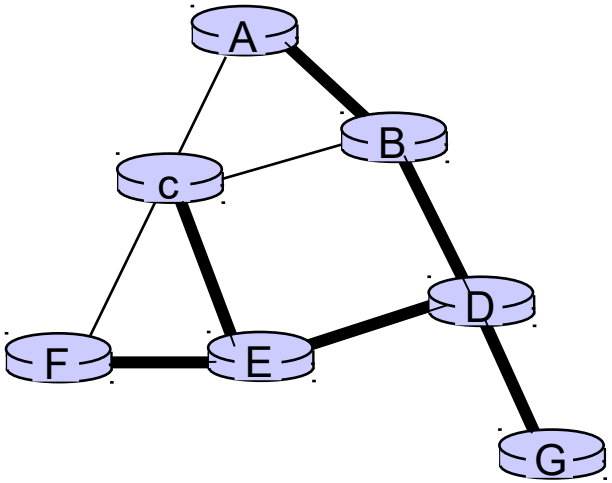
(b) difusión comenzada por D

Árbol de expansión: creación

- Nodo del centro
- cada nodo envía un mensaje de unión único al nodo central
  - mensaje reenviado hasta que llegue a un nodo que ya pertenece al árbol de expansión



(a) construcción paso a paso del árbol de expansión (centro: E)

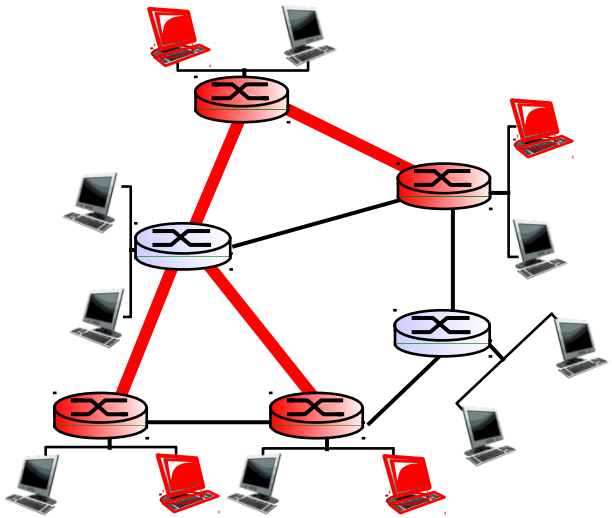


(b) árbol de expansión construido

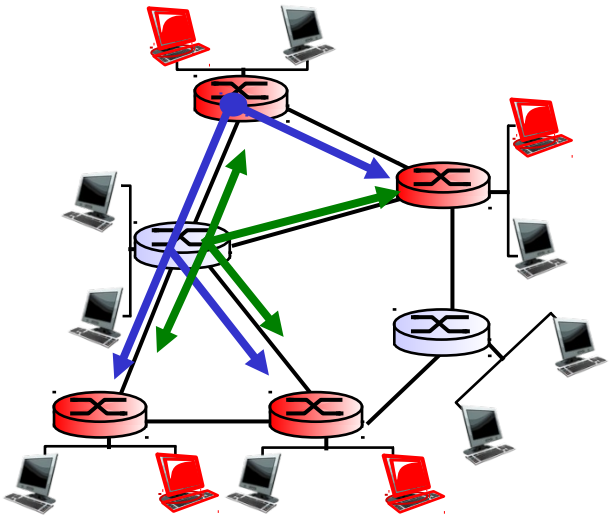
Enrutamiento de multidifusión

objetivo: encontrar un árbol (o árboles) conectando enrutadores que tienen mcast para miembros del grupo

- *árbol:* no se usan todos los caminos entre los enrutadores
- *árbol-compartido:* mismo árbol usado por todos los miembros de grupo
- *basado en fuentes:* árbol diferente desde cada emisor has el receptor



árbol compartido



árbol basado en fuentes

Enfoques para la construcción de árboles de mcast

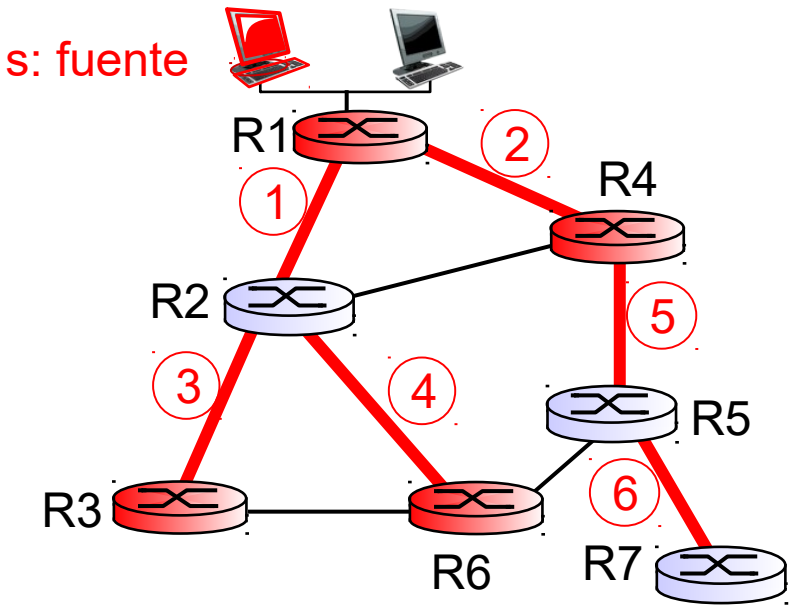
enfoques:

- *árbol basado en fuente:* un árbol por fuente
  - árboles de ruta más corta
  - reenvío de ruta inversa
- *árbol compartido de grupo:* el grupo usa un solo árbol
  - mínima expansión (Steiner)
  - árboles basados en su centro

…primero vemos los enfoques básicos, luego los protocolos específicos que adoptan estos enfoques

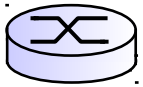
Árbol del camino más corto

- árbol de reenvío mcast: árbol de rutas de ruta más cortas desde la fuente hasta todos los receptores
  - Algoritmo de Dijkstra



LEYENDA

Enrutador con un miembro de grupo



Enrutador sin un miembro de grupo

enlace utilizado para el reenvío, i indica el ordenen el que es agregado por algoritmo

Reenvío de ruta inversa

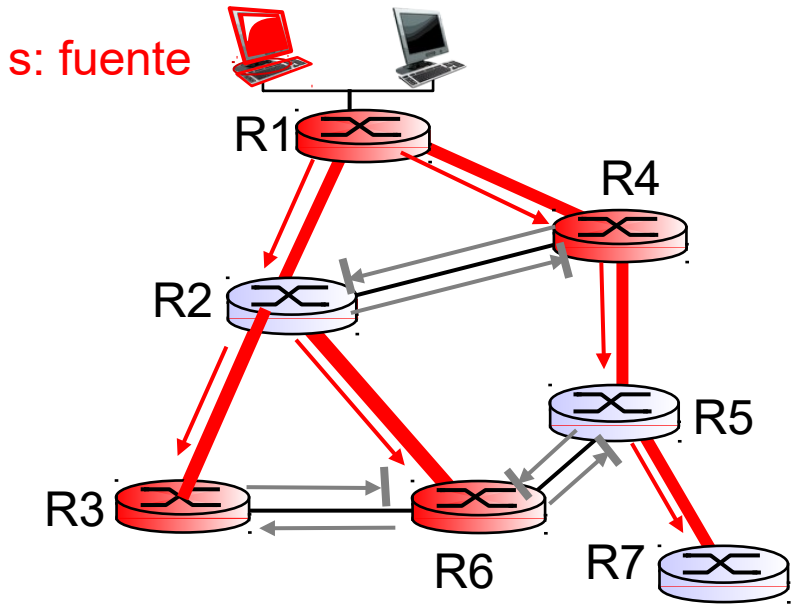
- Confía en el conocimiento del enrutador de la ruta más corta de unidifusión desde él hasta el remitente.
- cada enrutador tiene un comportamiento de reenvío simple:

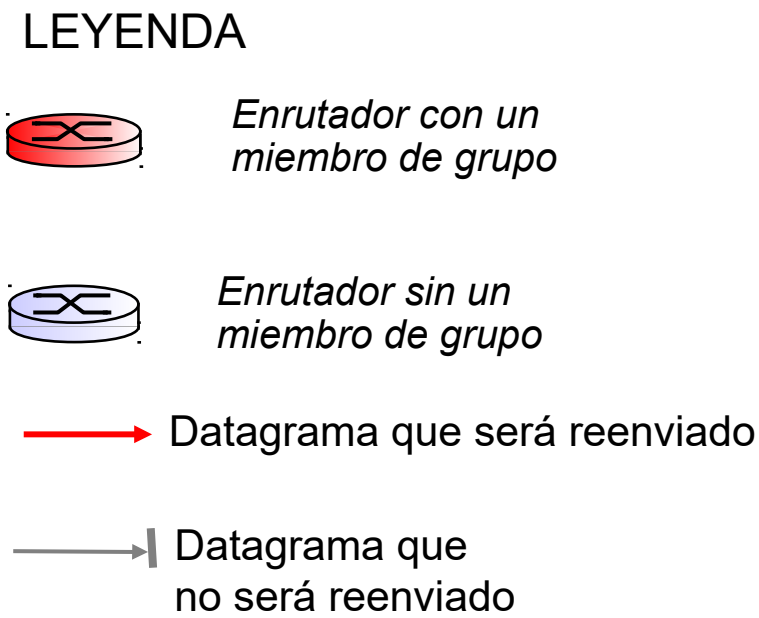
if (el datagrama recibido de mcast en el enlace entrante es el camino más corto de regreso al centro)

then inundación en todos los enlaces salientes con el datagrama

else ignora el datagrama

Reenvío de ruta inversa: ejemplo

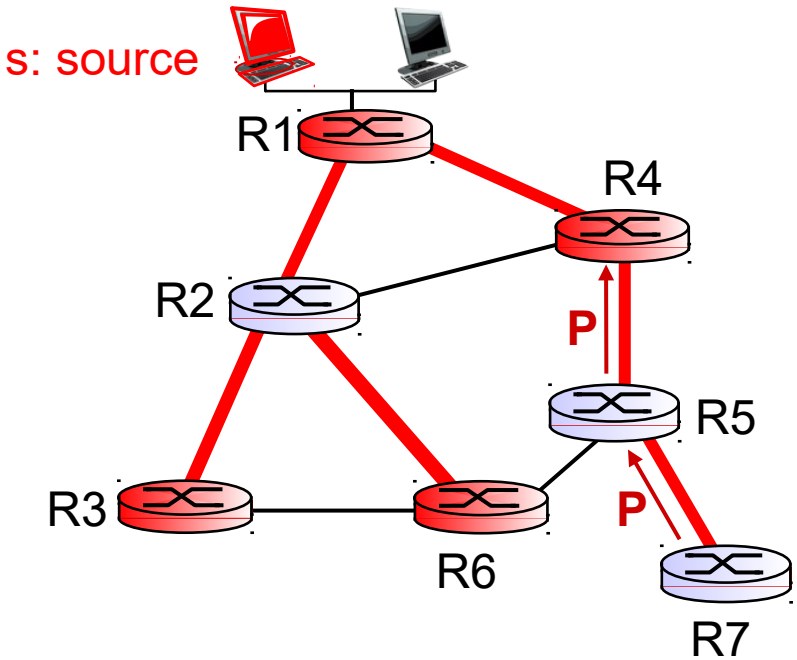


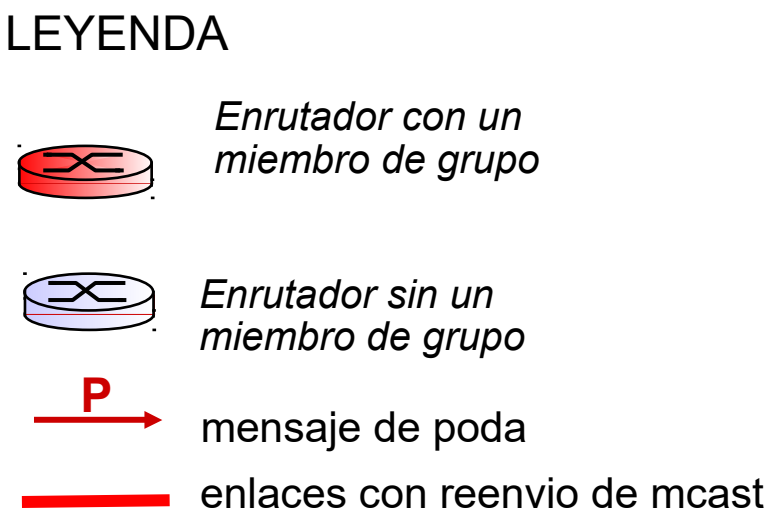


- El resultado es un SPT inverso específico de la fuente
  - puede ser una mala elección con enlaces asimétricos

Reenvío de ruta inversa: poda

- el árbol de reenvío contiene subárboles sin miembros de grupo de mcast
  - no es necesario reenviar datagramas hacia abajo del subárbol
- mensaje de "podar" es enviado en sentido ascendente por enrutador sin miembros del grupo aguas abajo





Arbol compartido: arbol de steiner

- *arbol de steiner:* Árbol de costo mínimo que conecta todos los enrutadores con miembros del grupo adjuntos
- el problema es NP-completo
- existe una heurística excelente
- no es utilizado en la práctica:
  - complejidad computacional
  - es necesaria información sobre toda la red
  - monolítico: vuelva a ejecutar cada vez que un enrutador necesite unirse/salirse

Kurose and Ross, Computer Networking: A Top Down Approach, 6th edition, Addison-Wesley, 2012

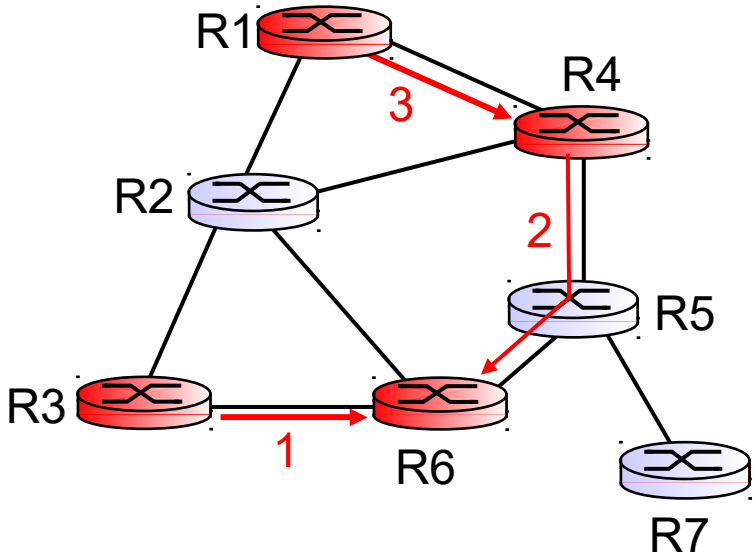
Arboles basados en su centro

- único árbol de entrega compartido por todos
- un enrutador identificado como "centro" del árbol
- para unirse:
  - el enrutador de borde envía un mensaje de unidifusión join-msg dirigido al enrutador central
  - join-msg "procesado" por enrutadores intermedios y reenviado hacia el centro
  - join-msg llega a la rama de árbol existente de este centro o llega al centro
  - ruta tomada por join-msg se convierte en una nueva rama de árbol para este enrutador

Kurose and Ross, Computer Networking: A Top Down Approach, 6th edition, Addison-Wesley, 2012

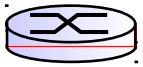
Arboles basados en su centro: ejemplo

Supongamos que R6 es elegido como centro:



LEYENDA

Enrutador con un miembro de grupo



Enrutador sin un miembro de grupo

orden de ruta en la que se unen los mensajes generados

Enrutamiento Multicasting de Internet: DVMRP

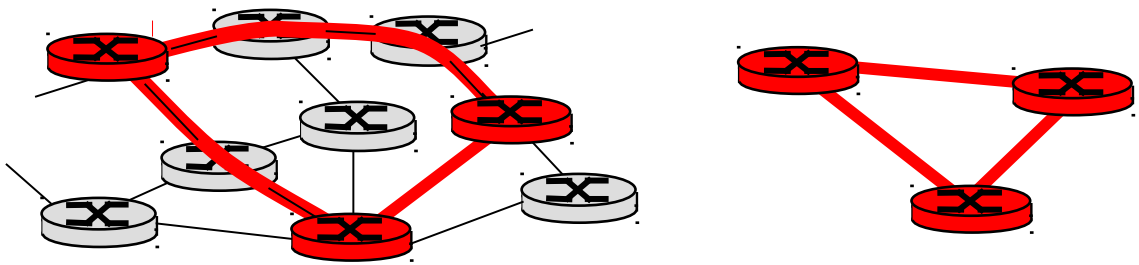
- DVMRP: protocolo de enrutamiento de multidifusión de vector de distancia, RFC1075
- inundación y poda: reenvío de ruta inversa, árbol basado en fuente
  - Árbol RPF basado en las tablas de enrutamiento propias de DVMRP construidas por enrutadores DVMRP que se están comunicando
  - sin suposiciones sobre unicast subyacente
  - datagrama inicial para el grupo de difusión inundado en todas partes a través de RPF
  - enrutadores que no quieren grupo: enviar mensajes de poda hacia arriba

DVMRP

- *estado suave: el enrutador DVMRP periódicamente (1 minuto) "se olvida" de las ramas se podan:*
  - *los datos de difusión vuelven a fluir hacia abajo en la rama sin podar*
  - *enrutador aguas abajo: vuelva a iniciar o continúa recibiendo datos*
- *los enrutadores pueden volver a injertarse rápidamente en un árbol*
  - *después de IGMP se una a una hoja at leaf*
- *retazos*
  - *comúnmente implementado en enrutadores comerciales*

Túneles

P: ¿cómo conectar "islas" de enrutadores de multidifusión en un "mar" de enrutadores de unidifusión?



topología física topología lógica

- datagrama mcast encapsulado dentro de datagrama "normal" (no dirigido por multidifusión)
- datagrama de IP normal enviado a través del "túnel" a través de unidifusión de IP común para recibir el enrutador mcast (recuperación de IPv6 dentro del túnel IPv4)
- recibir enrutador mcast se desencapsula para obtener un datagrama mcast

PIM: Protocol Independent Multicast

- no depende de ningún algoritmo de enrutamiento unicast subyacente específico (funciona con todos)
- dos escenarios de distribución de multidifusión diferentes:

denso:

- miembros del grupo densamente empaquetados, en proximidad "cercana".
- ancho de banda más abundante

escazo:

- # redes con miembros del grupo pequeñas wrt # redes interconectadas
- miembros del grupo "ampliamente dispersos"
- ancho de banda no abundante

Consecuencias de la dicotomía dispersa-densa:

denso

- la pertenencia a un grupo de enrutadores es asumida hasta que los enrutadores explícitamente sean podados
- construcción basada en datos en el árbol de difusión (por ejemplo, RPF)
- ancho de banda derrochador

escazo

- sin membresía hasta que los enrutadores se unan explícitamente
- construcción impulsada por el receptor de un árbol difuso (por ejemplo, basado en el centro)
- ancho de banda y procesamiento de enrutamiento no grupal y conservador

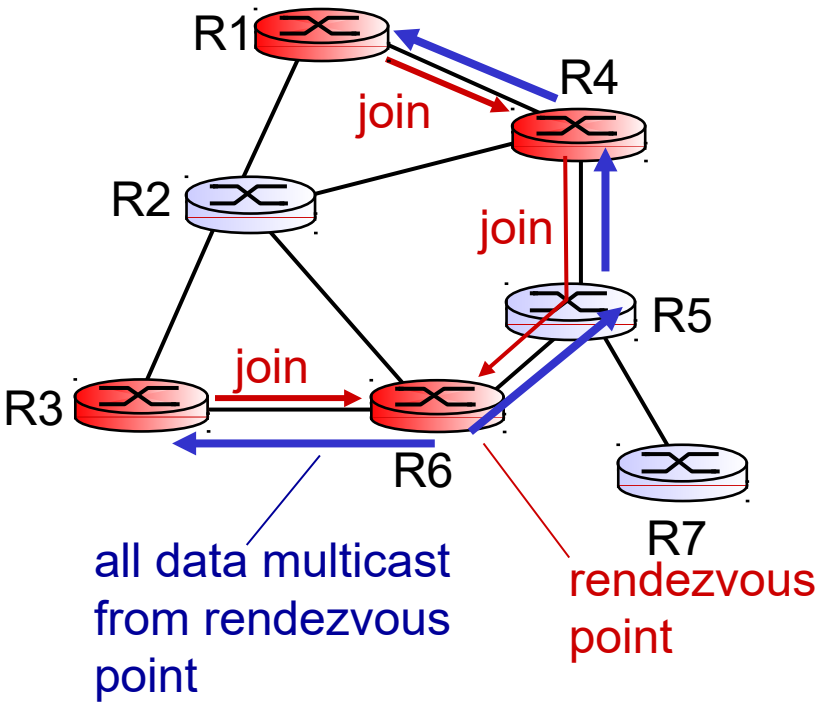
PIM- modo denso

flood-and-prune RPF: similar a DVMRP pero ...

- el protocolo unicast subyacente proporciona información RPF para el datagrama entrante
- Inundación aguas abajo menos complicada (menos eficiente) que DVMRP reduce la dependencia del algoritmo de enrutamiento subyacente
- tiene un mecanismo de protocolo para que el enrutador detecte que es un enrutador de nodo de hoja

PIM – modo escazo

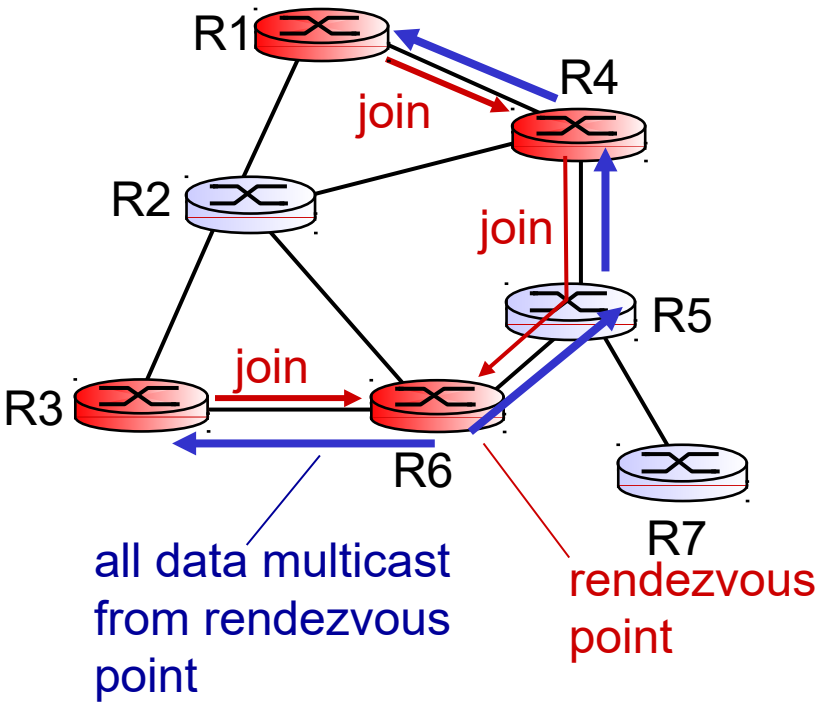
- enfoque basado en el centro
- el enrutador envía el mensaje de unión al punto de encuentro (RP)
  - los enrutadores intermedios actualizan el estado y la unión directa
- después de unirse mediante RP, el enrutador puede cambiar al árbol específico de la fuente
  - mayor rendimiento: menos concentración, trayectorias más cortas



PIM – modo escazo

remitente(s):

- datos de unidifusión a RP, que distribuye el árbol con raíz de RP
- RP puede extender el árbol de transmisión río arriba a la fuente
- RP puede enviar stop msg si no hay receptores conectados
  - "¡Nadie está escuchando!"



Listo el capítulo 4!

4.1 Introducción
4.2 Circuito virtual y redes de datagramas
4.3 ¿Qué hay dentro de un enrutador?
4.4 IP: Protocolo de Internet
- Formato del datagrama
- Direccionamiento IPv4
- Protocolo ICMP
- Direccionamiento IPv6

4.5 Algoritmos de enrutamiento
- Estado del enlace
- Vector de distancia
- Enrutamiento jerárquico

4.6 Enrutamiento en Internet
- RIP
- OSPF
- BGP

4.7 enrutamiento de difusión y multidifusión

- comprendimos los principios detrás de los servicios de la capa de red
- conocimos los modelos de servicio de capa de red, reenvío versus enrutamiento
- cómo funciona un enrutador, enrutamiento (selección de ruta), difusión, multidifusión
- instanciación, implementación en Internet