



AMPLIACIÓN DE SISTEMAS OPERATIVOS Y REDES

Grado en Ingeniería Informática / Doble Grado

Universidad Complutense de Madrid

TEMA 1.5. Encaminamiento en Internet

PROFESORES:

Rubén Santiago Montero

Eduardo Huedo Cuesta

Rafael Rodríguez Sánchez

Introducción: El problema del encaminamiento

- En una red de conmutación de paquetes, el **encaminamiento** consiste en encontrar un camino, desde el origen al destino, a través de nodos de conmutación o encaminadores (*routers*) intermedios
- **Caminos alternativos**
 - Es necesario decidir cuál es el mejor camino posible (*camino más corto*)
 - El *camino más corto* minimiza una métrica de encaminamiento
- **Métricas de encaminamiento**
 - **Número de saltos**: tiene en cuenta el número de encaminadores y/o redes intermedias que tiene que atravesar el paquete para alcanzar el destino
 - **Distancia geográfica**: tiene en cuenta la distancia (en Km) que tiene que recorrer el paquete para alcanzar el destino
 - **Retardo promedio**: tiene en cuenta el retardo de las líneas. Dado que éste es proporcional a la distancia, esta métrica es similar a la anterior
 - **Ancho de banda**: tiene en cuenta la velocidad de transmisión de las líneas por las que tiene que circular el paquete
 - **Nivel de tráfico**: tiene en cuenta el nivel de uso de las líneas, para intentar utilizar aquellas líneas con menor nivel de saturación
 - Combinación lineal de varias métricas

Técnicas de Encaminamiento

Encaminamiento local

- No tiene en cuenta la topología de la red y usa únicamente información local
- Las técnicas más comunes son:
 - Encaminamiento aleatorio
 - Encaminamiento aislado
 - Inundación

Encaminamiento estático

- Las decisiones de encaminamiento consideran la topología de la red
- Las tablas de encaminamiento se construyen manualmente y no se adaptan a los cambios de la red

Encaminamiento dinámico

- Las tablas de encaminamiento se construyen de forma automática, mediante el intercambio periódico de información entre los encaminadores
- Permite adaptar automáticamente el encaminamiento a los cambios en la topología de la red
- Las técnicas más comunes son:
 - Encaminamiento por vector de distancias (ej. RIP)
 - Encaminamiento por estado de los enlaces (ej. OSPF)

Encaminamiento en Internet

- Internet está organizado en **Sistemas Autónomos** (*Autonomous Systems, AS*)
 - Un AS es una conjunto de redes y encaminadores gestionados y administrados por una misma autoridad
 - Cada AS se identifica mediante un número de AS (*AS Number, ASN*)
 - Hay más de 54.000 ASs
- **Encaminadores internos** del AS
 - Interconectan únicamente redes dentro del propio AS
 - Sólo conocen en detalle la organización del AS local
 - Desconocen el camino a otros ASs
 - Utilizan protocolos de encaminamiento denominados IGP (*Interior Gateway Protocol*)
- **Encaminadores externos** o frontera (*border router*) del AS
 - Interconectan varios ASs
 - Conocen el camino a otros ASs, pero no conocen la organización interna de los mismos
 - Utilizan protocolos de encaminamiento denominados EGP (*Exterior Gateway Protocol*)

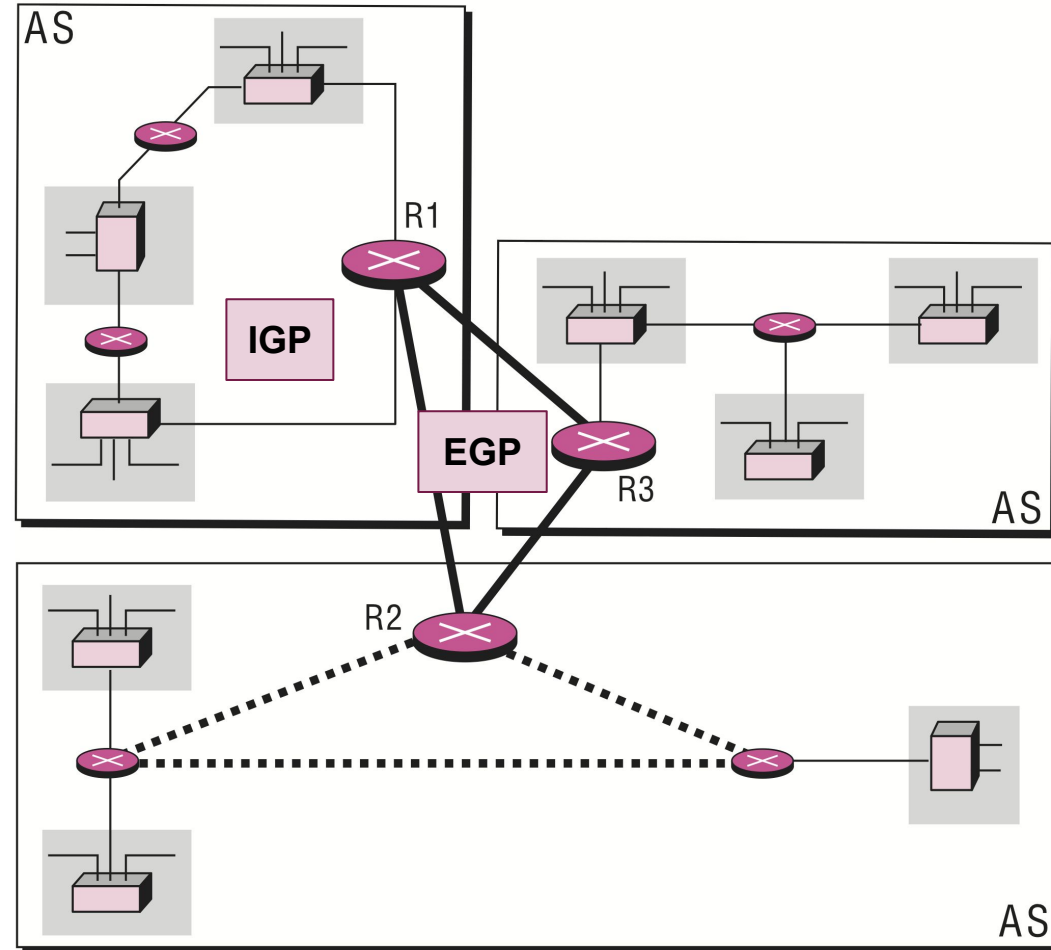
Encaminamiento en Internet

Protocolos Internos (IGP): Usados por los encaminadores internos, para el encaminamiento dentro de un AS:

- RIP: Routing Information Protocol
- OSPF: Open Shortest Path First
- EIGRP: Enhanced Interior Gateway Routing Protocol (de Cisco)
- IS-IS: Intermediate System to Intermediate System

Protocolos Externos (EGP): Usados por los encaminadores frontera, para el encaminamiento entre distintos ASs:

- EGP: Exterior Gateway Protocol (obsoleto)
- BGP: Border Gateway Protocol





AMPLIACIÓN DE SISTEMAS OPERATIVOS Y REDES

Grado en Ingeniería Informática / Doble Grado

Universidad Complutense de Madrid

Vector de Distancias. RIP

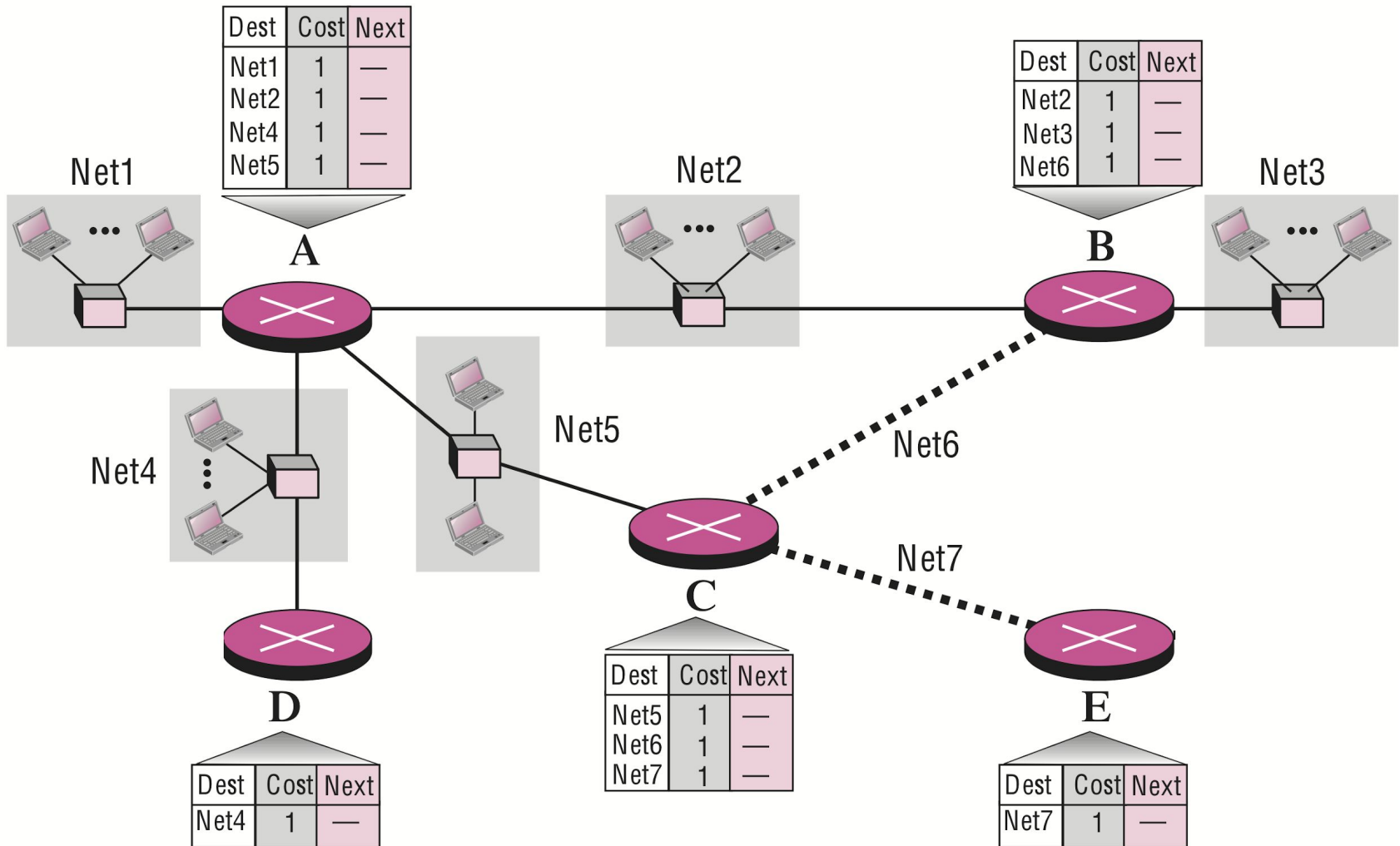
Vector de Distancias

Fundamentos

- Cada encaminador mantiene una tabla de encaminamiento con una entrada por cada posible destino en la red
- Cada entrada de la tabla contiene:
 - El destino (normalmente una red)
 - El siguiente salto (nodo o encaminador) para alcanzar dicho destino
 - La distancia o métrica para el destino, que suele ser el número de saltos
- Para construir la tabla de encaminamiento, los nodos intercambian periódicamente sus vectores de distancias (destinos y distancias) con sus vecinos
 - La distancia total al destino es la anunciada por el encaminador más la distancia al encaminador
 - Si la distancia total es menor que la actual, se sustituye la entrada
 - Si el siguiente salto de la entrada es el encaminador, se actualiza la entrada
 - La distancia total puede ser mayor debido a un cambio en la red
- El proceso iterativo de intercambio converge idealmente a los caminos óptimos
 - Este método también recibe el nombre de algoritmo de Bellman-Ford
- Ejemplo: RIP (Routing Information Protocol)

Vector de Distancias

Ejemplo: Inicialmente los encaminadores sólo conocen sus rutas directas

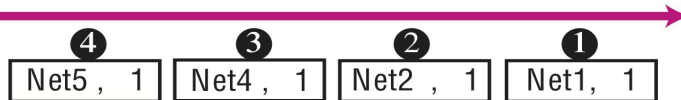


Vector de Distancias

Ejemplo: Proceso de intercambio

Dest	Cost	Next
Net1	1	—
Net2	1	—
Net4	1	—
Net5	1	—

A

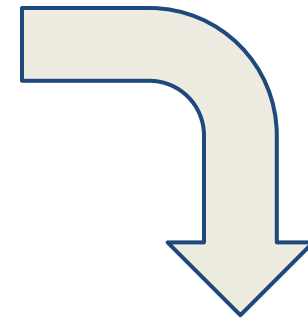


Dest	Cost	Next
Net2	1	—
Net3	1	—
Net6	1	—

B



Después del intercambio de todas las tablas



Routing Table B			Routing Table B			Routing Table B			Routing Table B		
Dest	Cost	Next	Dest	Cost	Next	Dest	Cost	Next	Dest	Cost	Next
Net1	2	A	Net1	2	A	Net1	2	A	Net1	2	A
Net2	1	—	Net2	1	—	Net2	1	—	Net2	1	—
Net3	1	—	Net3	1	—	Net3	1	—	Net3	1	—
Net6	1	—	Net6	1	—	Net4	2	A	Net4	2	A
After receiving record 1			After receiving record 2			After receiving record 3			After receiving record 4		

A

Dest	Cost	Next
Net1	1	—
Net2	1	—
Net3	2	B
Net4	1	—
Net5	1	—
Net6	2	C
Net7	2	C

B

Dest	Cost	Next
Net1	2	A
Net2	1	—
Net3	1	—
Net4	2	A
Net5	2	A
Net6	1	—
Net7	2	C

C

Dest	Cost	Next
Net1	2	A
Net2	2	A
Net3	2	B
Net4	2	A
Net5	1	—
Net6	1	—
Net7	1	—

D

Dest	Cost	Next
Net1	2	A
Net2	2	A
Net3	3	A
Net4	1	—
Net5	1	A
Net6	3	A
Net7	3	A

E

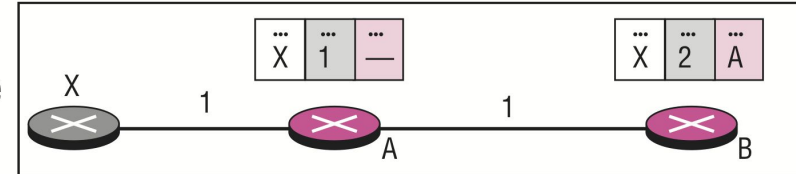
Dest	Cost	Next
Net1	3	C
Net2	3	C
Net3	3	C
Net4	3	C
Net5	2	C
Net6	2	C
Net7	1	—

Vector de Distancias

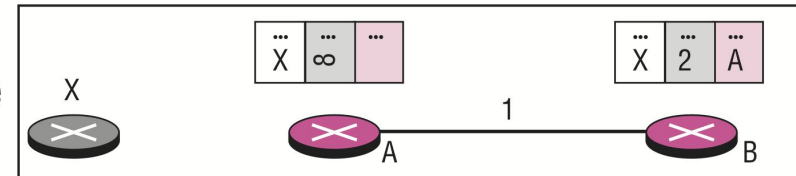
Problemas de convergencia. Cuenta a infinito

- Los cambios en la topología de la red deben propagarse a todos los encaminadores
- Cuando un enlace aumenta su distancia estos cambios tardan en propagarse
- Las actualizaciones para comunicar un enlace caído pueden no converger

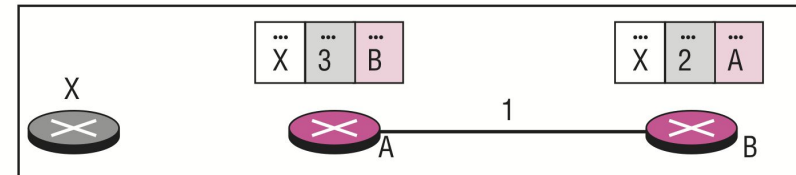
Before failure



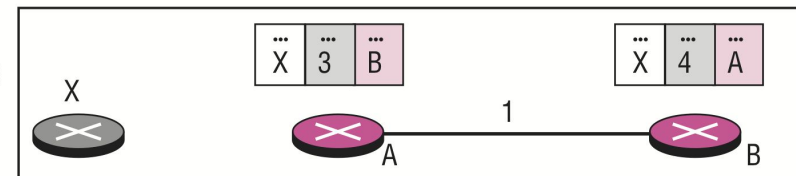
After failure



After A receives update from B

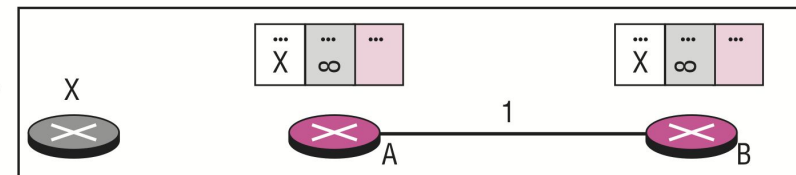


After B receives update from A



⋮

Finally



Vector de Distancias

Cuenta a infinito. Soluciones

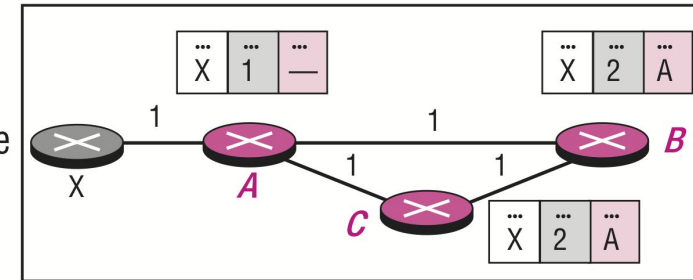
- **Infinito pequeño (*small infinity*)**
 - Por ejemplo, en RIP el infinito se establece en 16 saltos (una distancia de 16 se considera inalcanzable y, por tanto, las rutas tienen un límite de 15 saltos)
- **Horizonte dividido (*split horizon*)**
 - Los destinos aprendidos a través de un determinado enlace nunca se difunden a través de dicho enlace
 - **Ejemplo:** El nodo B no enviará al nodo A información sobre el destino X
- **Horizonte dividido con inversa envenenada (*poisoned reverse*)**
 - Los destinos aprendidos a través de un determinado enlace sí se difunden a través de dicho enlace, pero con distancia infinita
 - **Ejemplo:** El nodo B anunciará al nodo A que el destino X está a distancia infinita
- **Actualizaciones forzadas (*triggered updates*)**
 - Cuando un encaminador detecta una modificación en su tabla de rutas inmediatamente difunde esta información a sus vecinos
 - De esta forma, los cambios en la topología se propagan de forma rápida

Vector de Distancias

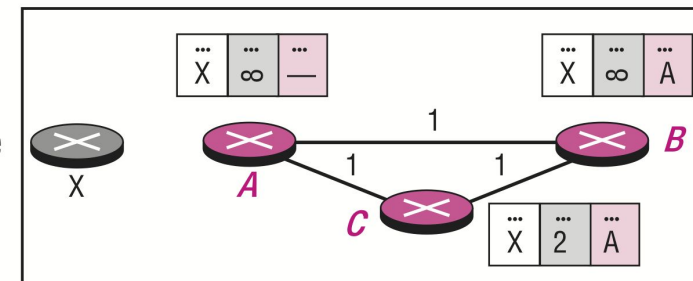
Problemas de convergencia. Bucles

- En redes con bucles el algoritmo puede no converger
- Las técnicas de horizonte dividido no evitan el problema en este caso
- Las actualizaciones forzadas aceleran la convergencia

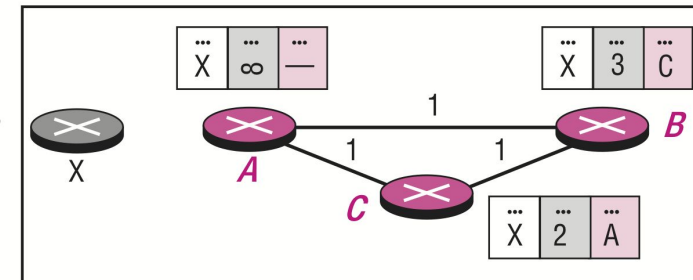
Before failure



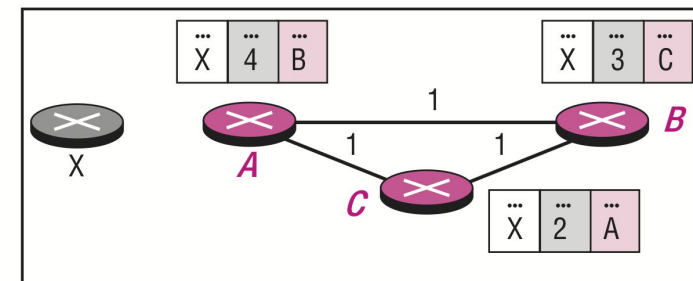
After A sends the route to B and C, but the packet to C is lost



After C sends the route to B



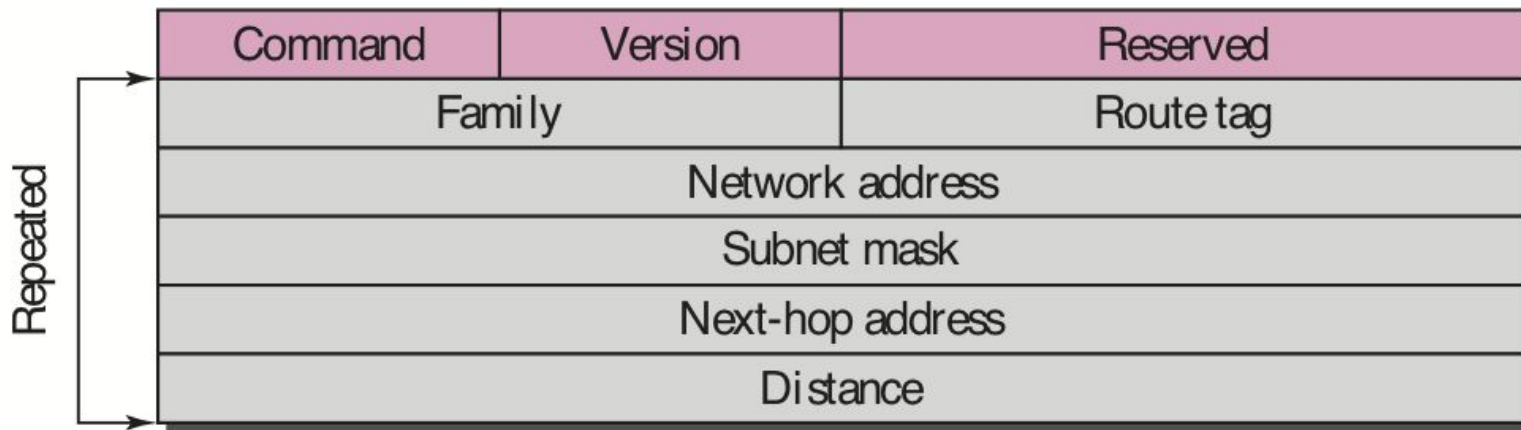
After B sends the route to A



Routing Information Protocol: RIP

- Protocolo de encaminamiento interior (IGP) por vector de distancias
- Versiones y RFCs
 - RIP-1 → RFC 1058 (1993)
 - RIP-2 → RFC 2453 (1998)
 - RIPng (para IPv6) → RFC 2080 (1997)
- El vector de distancias incluye la siguiente información de encaminamiento:
 - La lista de destinos (redes) que son alcanzables por cada encaminador
 - La distancia a la que se encuentran dichos destinos, como número de saltos
- Los mensajes se encapsulan en datagramas UDP dirigidos al puerto 520
- El infinito se establece en 16 saltos y se pueden utilizar los siguientes mecanismos
 - Horizonte dividido
 - Horizonte dividido con ruta inversa envenenada
 - Actualizaciones forzadas
- La versión 2 añade:
 - Soporte para direcciones sin clase
 - Soporte para direccionamiento *multicast* (224.0.0.9)
 - Soporte para autenticación

RIP: Formato del Mensaje



- **Command** (8 bits): Request (1) o Response (2)
- **Version** (8 bits): RIP-1 (1) o RIP-2 (2)
- **Family** (16 bits): TCP/IP (2) o autenticación (0xFFFF)
- **Route tag** (16 bits): Información adicional de ruta (ej. ASN, para separar rutas internas y externas) o algoritmo de autenticación, que puede ser una contraseña (2) o un resumen del mensaje con clave (3)
- **Network address y Subnet mask** (32 bits): Dirección de red sin clase
- **Next-hop address** (32 bits): Normalmente, es 0.0.0.0 para usar la dirección del remitente del mensaje, pero podría ser un encaminador que no soporte RIP
- **Distance** (32 bits): Número de saltos al destino

RIP: Mensajes

- **Mensajes de solicitud (REQUEST)**

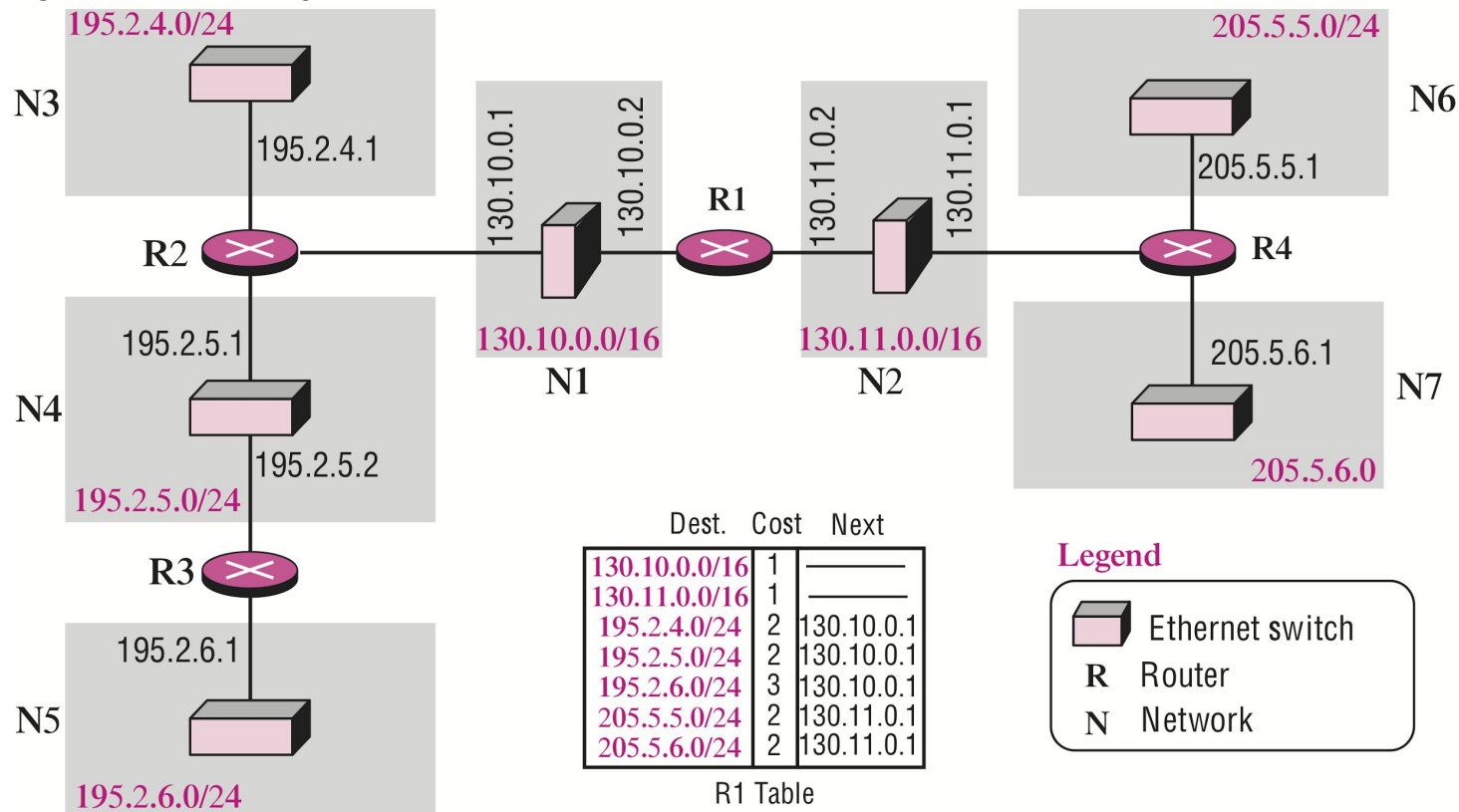
- Enviados cuando se conecta a la red → Network address = 0.0.0.0 (todas las entradas)
- Enviados cuando caduca una entrada en la tabla → Network address = entrada caducada

- **Mensajes de respuesta (RESPONSE)**

- Difundidos periódicamente (*broadcast* o *multicast*) con el vector de distancias
- Enviados en respuesta a una solicitud
- Enviados cuando cambia la distancia a la red (actualización forzada)

RIP: Ejemplo

Ejercicio: ¿Qué mensaje RIP (RESPONSE) enviará R1 a R2?



Dest.	Cost	Next
130.10.0.0/16	1	—
130.11.0.0/16	2	130.10.0.2
195.2.4.0/24	1	—
195.2.5.0/24	1	—
195.2.6.0/24	2	195.2.5.2
205.5.5.0/24	3	130.10.0.2
205.5.6.0/24	3	130.10.0.2

R2 Table

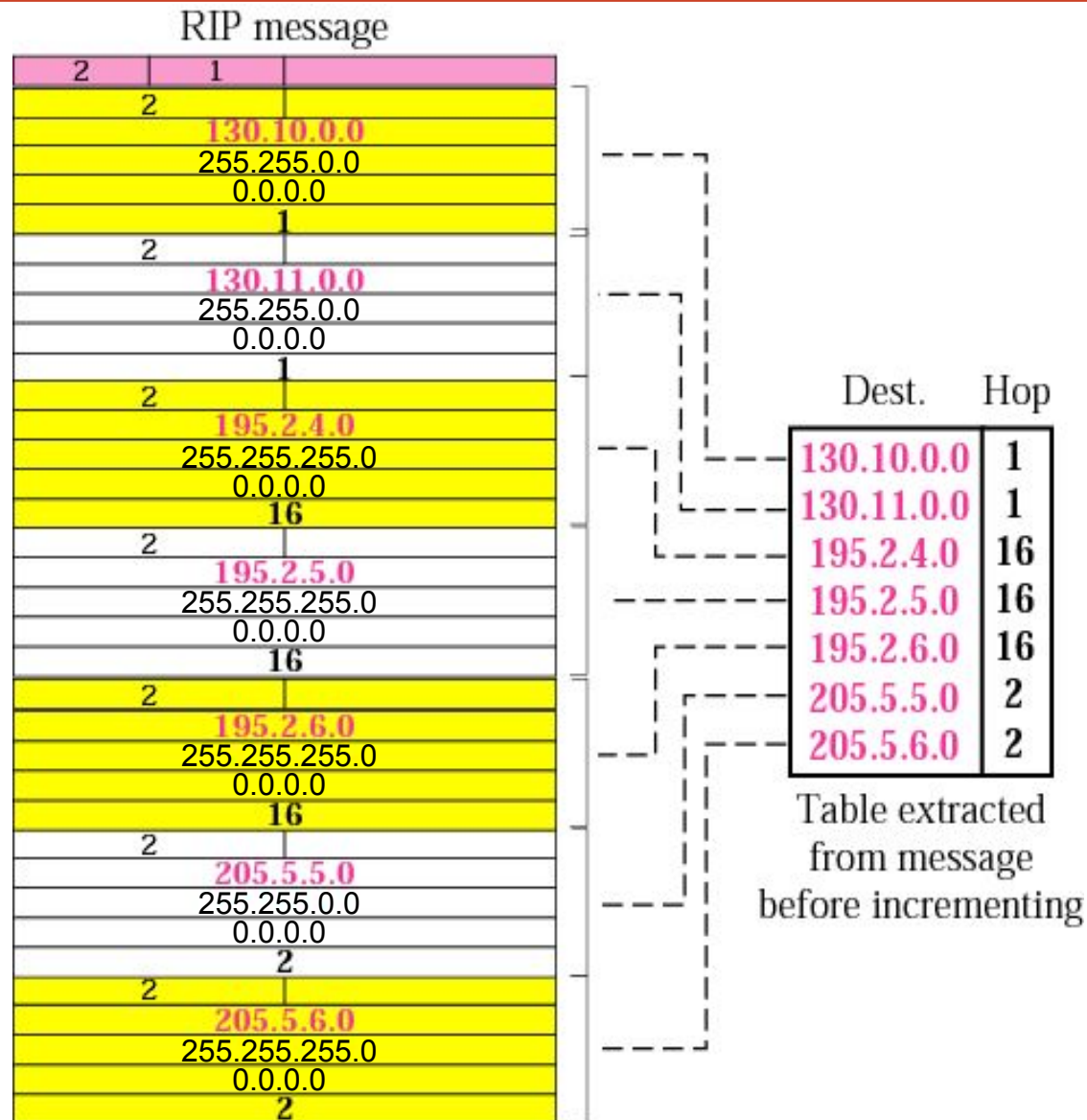
Dest.	Cost	Next
130.10.0.0/16	2	195.2.5.1
130.11.0.0/16	3	195.2.5.1
195.2.4.0/24	2	195.2.5.1
195.2.5.0/24	1	—
195.2.6.0/24	1	—
205.5.5.0/24	4	195.2.5.1
205.5.6.0/24	4	195.2.5.1

R3 Table

Dest.	Cost	Next
130.10.0.0/16	2	130.11.0.2
130.11.0.0/16	1	—
195.2.4.0/24	3	130.11.0.2
195.2.5.0/24	3	130.11.0.2
195.2.6.0/24	4	130.11.0.2
205.5.5.0/24	1	—
205.5.6.0/24	1	—

R4 Table

RIP: Ejemplo



RIP: Temporizadores

Temporizador periódico (25-35 s)

- Intervalo de envío de mensajes RESPONSE para anunciar el vector de distancias
- El protocolo RIP establece un valor de 30 s para este temporizador. En la práctica, se usa un valor aleatorio entre 25 y 35 s

Temporizador de expiración (180 s)

- Controla el periodo de validez de una entrada de la tabla de encaminamiento
- Si no se recibe actualización de la entrada durante un intervalo de 180 s, la entrada deja de considerarse válida

Temporizador de “recolección de basura” (120 s)

- Cuando una entrada de la tabla de rutas expira, el encaminador no la elimina inmediatamente de la tabla de encaminamiento
- La entrada se sigue anunciando con métrica 16 (destino inalcanzable) durante un periodo adicional de 120 s

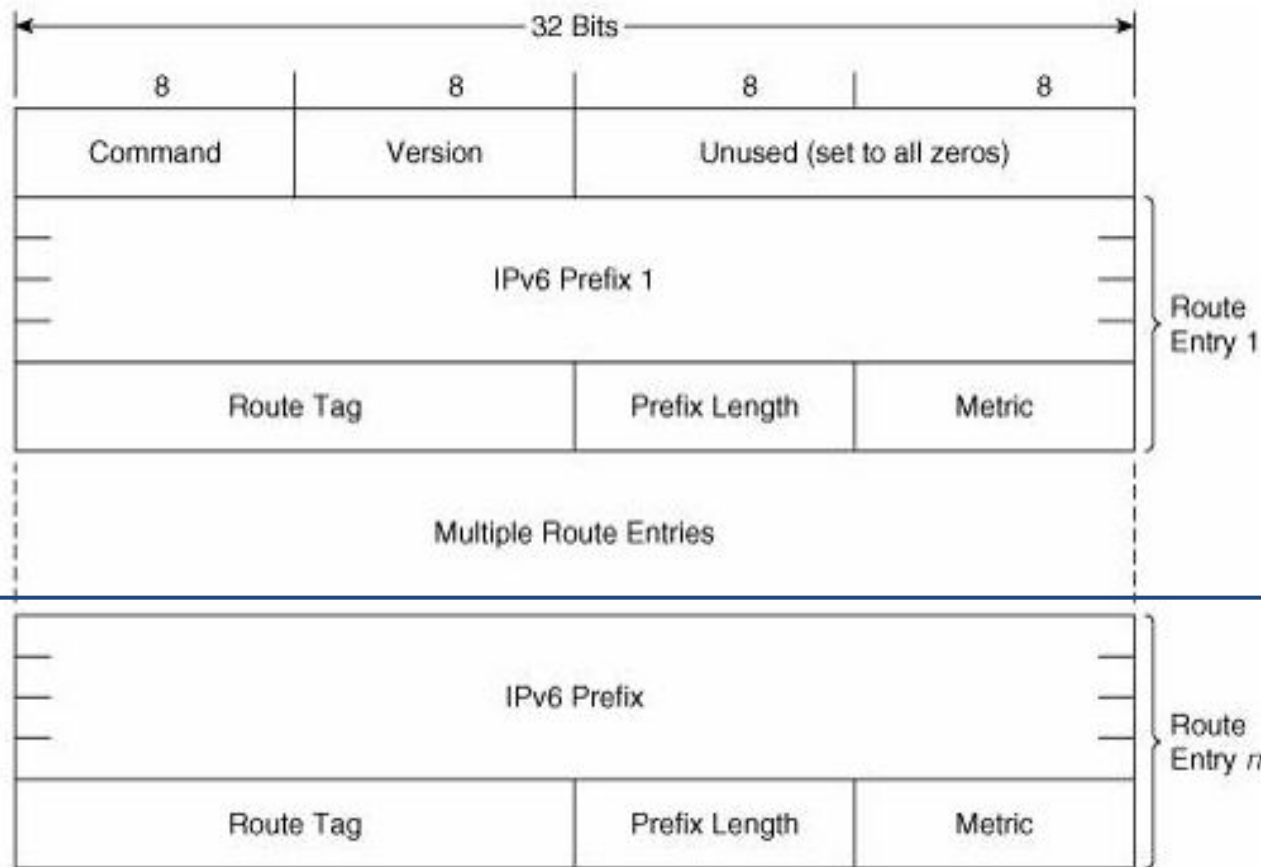
RIP: Limitaciones

- Puede generar **gran cantidad de tráfico *broadcast* o *multicast***, debido a la difusión periódica de los vectores de distancias (mensajes RESPONSE)
- **No admite métricas alternativas** al número de saltos
- Una vez calculadas las tablas, **no se admiten caminos alternativos** para equilibrar la carga de la red
- Cuando la red crece, los cambios pueden tardar bastante **tiempo en propagarse** hasta todos los puntos de la red
- El **infinito** se establece en 16 saltos. Redes grandes pueden necesitar más saltos

RIPng: RIP para IPv6

- RIPng (RIP *next generation*) es la adaptación del protocolo RIP-2 para IPv6
- Diferencias con RIP-2:
 - Los mensajes RIPng se encapsulan en datagramas UDP dirigidos al puerto 521 y se difunden a la dirección IPv6 multicast FF02::9
 - El vector de distancias contenido en los mensajes de tipo RESPONSE, en lugar de direcciones de red IPv4, anuncia prefijos de red IPv6
 - La información de ruta contenida en un vector de distancias no incluye el campo Next Hop (casi duplicaría el tamaño de cada entrada)
 - En su lugar, se puede incluir una entrada especial (con 0xFF en el campo Metric) que afecta a las entradas siguientes
 - No implementa autenticación, sino que utiliza los mecanismos de cifrado y autenticación disponibles en IPv6

RIPng: Formato del Mensaje



Entrada de la tabla de rutas:

- **IPv6 Prefix** (128 bits): prefijo de red IPv6 de la red destino anunciada
- **Prefix Length** (8 bits): longitud del prefijo de red anunciado
- **Route Tag** (16 bits) y **Metric** (8 bits): igual que en RIP-2



AMPLIACIÓN DE SISTEMAS OPERATIVOS Y REDES

Grado en Ingeniería Informática / Doble Grado

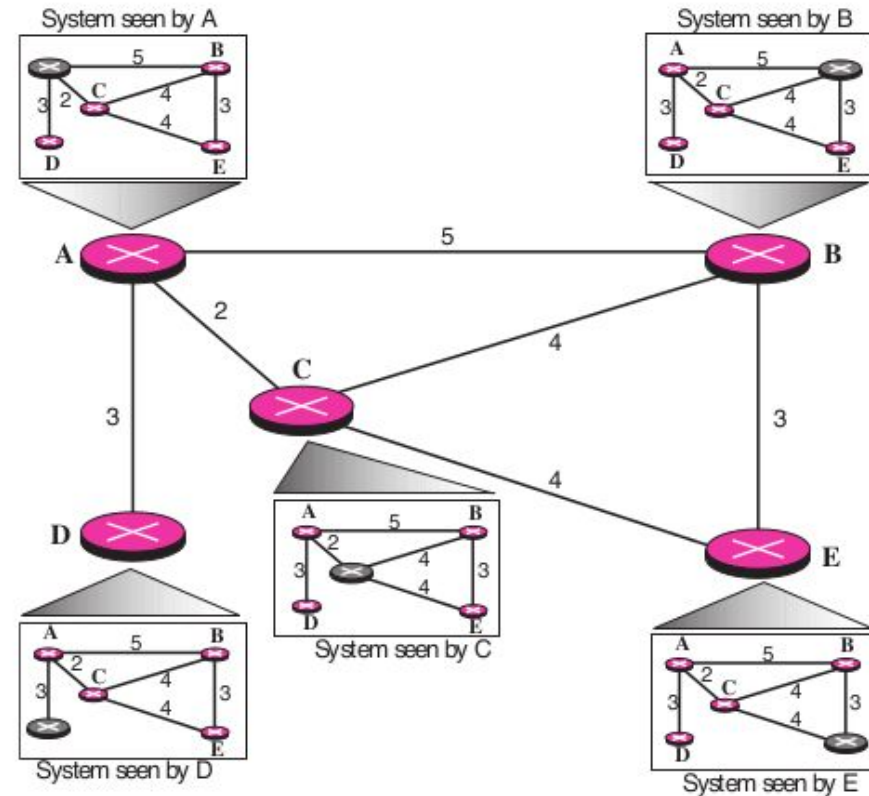
Universidad Complutense de Madrid

Estado de Enlaces. OSPF

Estado de Enlaces

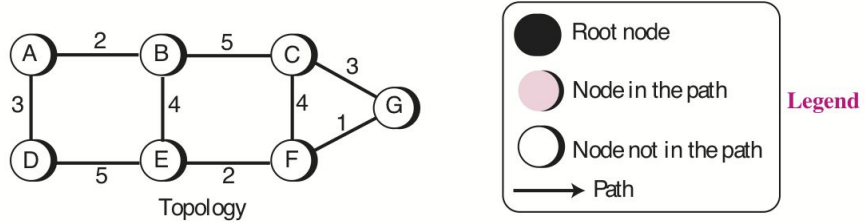
Fundamentos

- Cada encaminador mantiene una base de datos (*link state database*) con la información sobre la topología exacta de la red
- Para construir esta base de datos, cada encaminador:
 - Identifica sus nodos vecinos y su distancia (estado de enlace)
 - Distribuye esta información a **todos** los nodos de la red (inundación)
- Usando la información completa de la red (grafo), cada nodo construye un mapa de rutas (árbol) con él mismo como origen (raíz) usando el algoritmo de Dijkstra
- Ejemplo: OSPF (Open Shortest Path First)

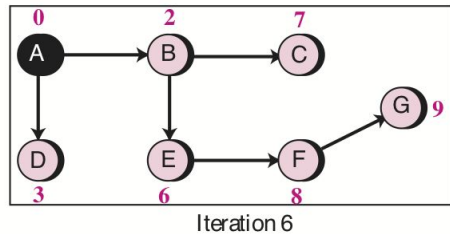
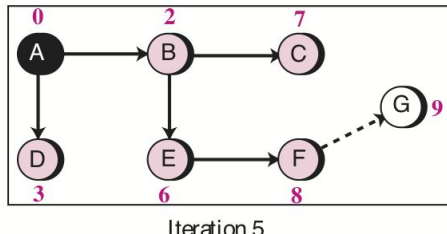
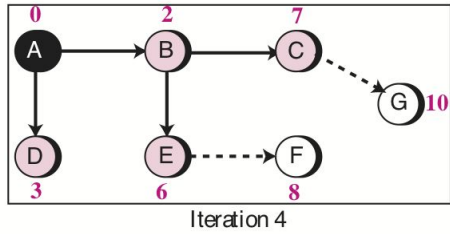
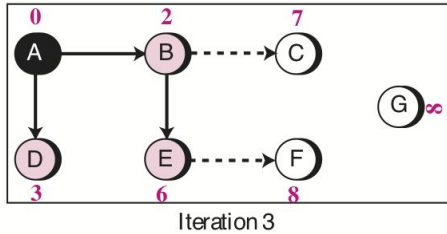
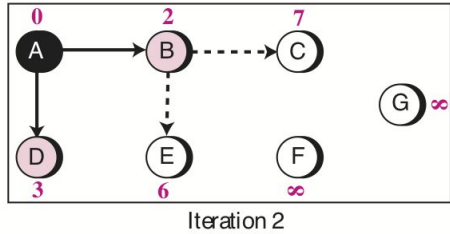
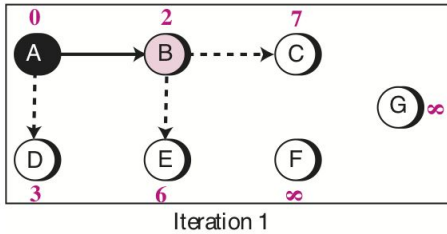
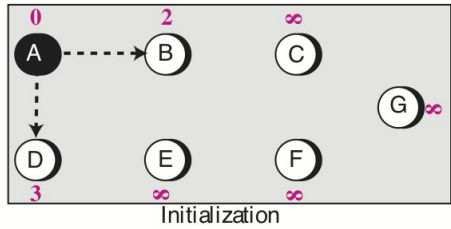


Estado de Enlaces

Ejemplo: Calcular las rutas desde el nodo A. Derivar la tabla de encaminamiento resultante



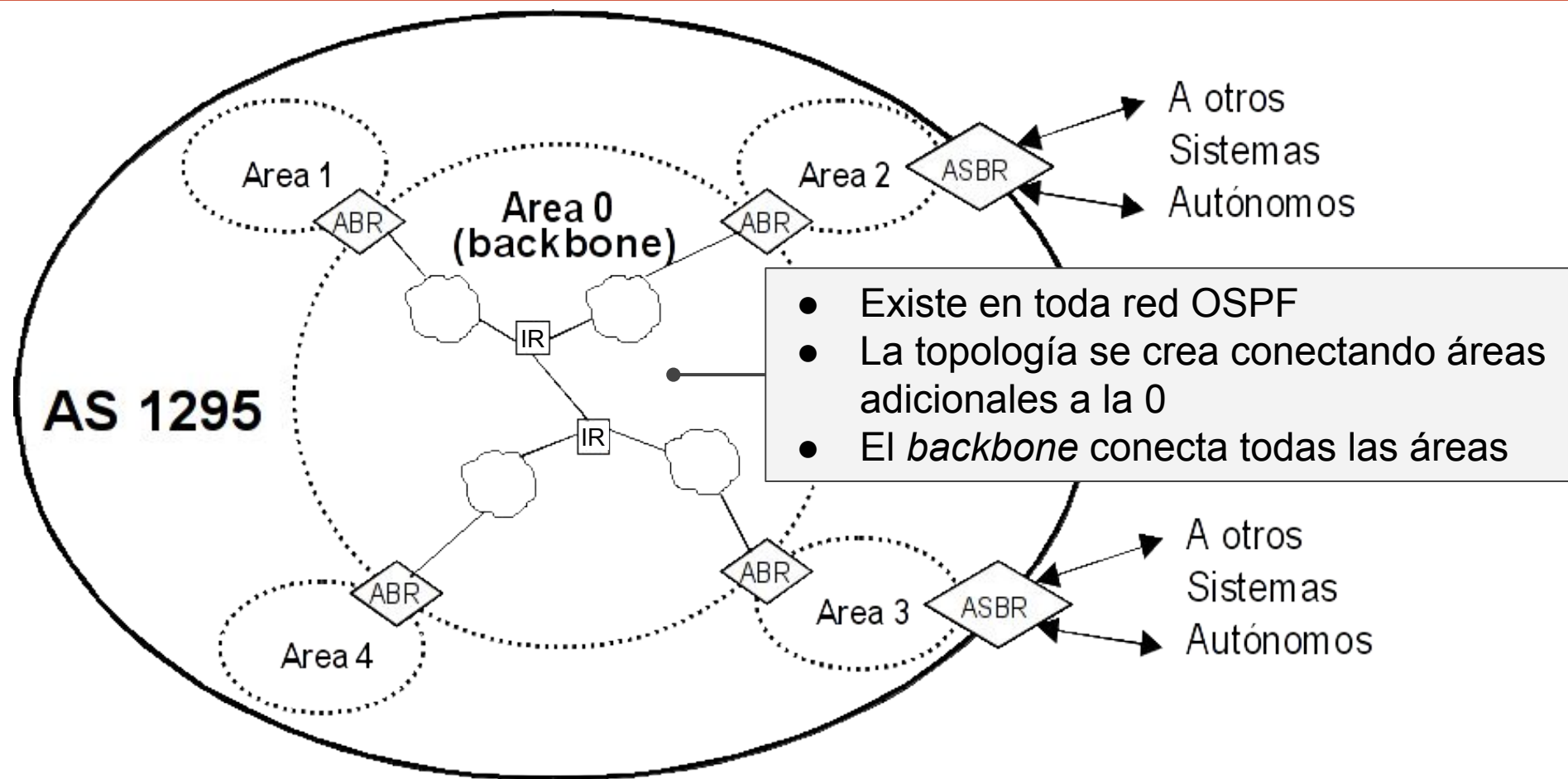
Destination	Cost	Next Hop
A	0	-
B	2	-
C	7	B
D	3	-
E	6	B
F	8	B
G	9	B



Open Shortest Path First: OSPF

- Protocolo de encaminamiento interno (IGP) por estado de los enlaces
- Se desarrolló como alternativa a RIP para aliviar sus limitaciones:
 - Equilibrado de carga entre caminos equivalentes
 - Particionado lógico de la red para reducir la cantidad de información anunciada
 - Convergencia más rápida, por propagar inmediatamente los cambios en las rutas
 - Soporte para máscaras de longitud variable (VLSM) y CIDR
 - Soporte para autenticación de cualquier nodo que anuncie rutas
- Utiliza un protocolo propio de encapsulado (89) y direcciones *multicast*:
 - 224.0.0.5 o FF02::5 - todos los encaminadores OSPF de una red
 - 224.0.0.6 o FF02::6 - los encaminadores designados OSPF de una red
- Versiones y RFCs
 - OSPF version 2 → RFC 2328 (1998)
 - OSPF version 3 (para IPv6) → RFC 5340 (2008)

OSPF: Áreas



Área: agrupación lógica de encaminadores y redes, con un identificador único (Area ID) de 32 bits

- Los encaminadores mantienen únicamente información de su área
- Limitan el número de intercambios de información de los enlaces

OSPF: Encaminadores y Redes

Encaminadores

- Cada encaminador tiene un identificador único (Router ID) de 32 bits en la red OSPF
- La información que se almacena y se intercambia depende del tipo:
 - **Internal Router (IR)**
 - Localizado en un área (todos sus interfaces están en el área)
 - Mantiene sólo información de la topología de su área
 - **Area Border Router (ABR)**
 - Conectado a dos o más áreas, una de ellas la 0
 - Mantiene una DB para cada una de las áreas a las que está conectado
 - **Autonomous System Boundary Router (ASBR)**
 - Situado en la frontera del AS, transmite rutas externas a la red OSPF
 - Puede inyectar rutas aprendidas mediante otro protocolo, como RIP

Redes

- Definen la frecuencia y el tipo de comunicaciones entre los encaminadores
- OSPF define las siguientes redes: punto-a-punto, multi-acceso con *broadcast*, multi-acceso sin *broadcast* y punto-multipunto

OSPF: Vecindades y Adyacencias

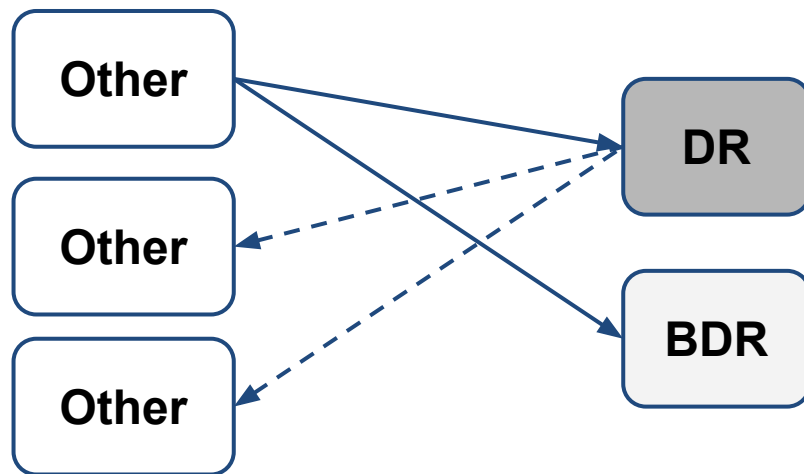
Relación de vecindad

- Se establece entre encaminadores que comparten un enlace común, pertenecen a la misma área OSPF y usan el mismo mecanismo de autenticación
- Configuración de los temporizadores del protocolo OSPF Hello
- Configuración del área

Relación de adyacencia

- Se establece entre encaminadores vecinos
- Dos encaminadores adyacentes intercambian información de estado de los enlaces
- Permite limitar la información intercambiada entre los encaminadores (no se comunican todos los vecinos)
- Las relaciones de adyacencia se desarrollan según el tipo de red
 - Punto-a-punto: entre los dos dispositivos vecinos
 - Multi-acceso: el encaminador designado (DR) y el encaminador designado de respaldo (BDR) son adyacentes al resto de los encaminadores de la red

OSPF: Vecindades y Adyacencias



- Los encaminadores envían su información de estado de enlaces al DR y BDR (224.0.0.6 o FF02::6)
- - → El DR envía la información agregada al resto (224.0.0.5 o FF02::5)

- El proceso de distribución de la información de enlaces es una optimización de la estrategia de inundación
- En caso de fallo del DR, el BDR asume sus funciones
- Los anuncios del DR (ej. después de una actualización) no son inmediatos, para solapar el envío de múltiples actualizaciones
- Los mensajes de actualización (Link State Update) se confirman (Link State Ack) para asegurar la fiabilidad

OSPF: Funcionamiento

Descubrimiento de vecinos: El protocolo OSPF Hello

- El protocolo se utiliza para establecer y mantener la relación entre vecinos
 - Cada encaminador envía periódicamente mensajes Hello a través de todos sus interfaces
 - Los mensajes Hello incluyen una lista con los Router IDs de todos los encaminadores en la red de los que se han recibido mensajes Hello recientemente
 - Cuando un encaminador recibe un mensaje Hello que contiene su Router ID en la lista, establece una relación de vecindad con el encaminador que lo envió
- En redes multiacceso, también se selecciona el DR usando información de prioridad contenida en los mensajes Hello
 - Los dos encaminadores con mayor prioridad en una red se convierten en DR y BDR respectivamente
 - En caso de empate en la prioridad, se elige el encaminador con mayor ID

OSPF: Funcionamiento

Sincronización de la Base de Datos de Estado de los Enlaces

- Los encaminadores adyacentes deben sincronizar sus bases de datos para establecer la adyacencia completa
- **Intercambio de Base de Datos**
 - El intercambio sigue un patrón maestro-esclavo
 - El maestro (con mayor Router ID) inicia el intercambio de mensajes (Database Description), que resumen el contenido de la base de datos
 - Cada mensaje es confirmado con otro mensaje que incluye la información del esclavo
- **Carga de la Base de Datos**
 - Durante la fase anterior, los encaminadores van detectando información obsoleta o no presente en su base de datos
 - En esta fase, se solicita al encaminador adyacente una copia de esa información (Link State Request)
 - La información proporcionada (Link State Update) debe ser confirmada (Link State Ack), siguiendo un mecanismo similar al de actualización

OSPF: Funcionamiento

Construcción de las tablas de encaminamiento

- Cuando el encaminador dispone de la información del estado de los enlaces, construye su árbol de rutas
- El árbol de rutas incluye tanto encaminadores (Router ID) como redes (dirección IP) y la métrica asociada
- A partir del árbol de rutas se construye la tabla de encaminamiento que contiene el destino (red/host), el siguiente salto y la métrica total asociada



AMPLIACIÓN DE SISTEMAS OPERATIVOS Y REDES

Grado en Ingeniería Informática / Doble Grado

Universidad Complutense de Madrid

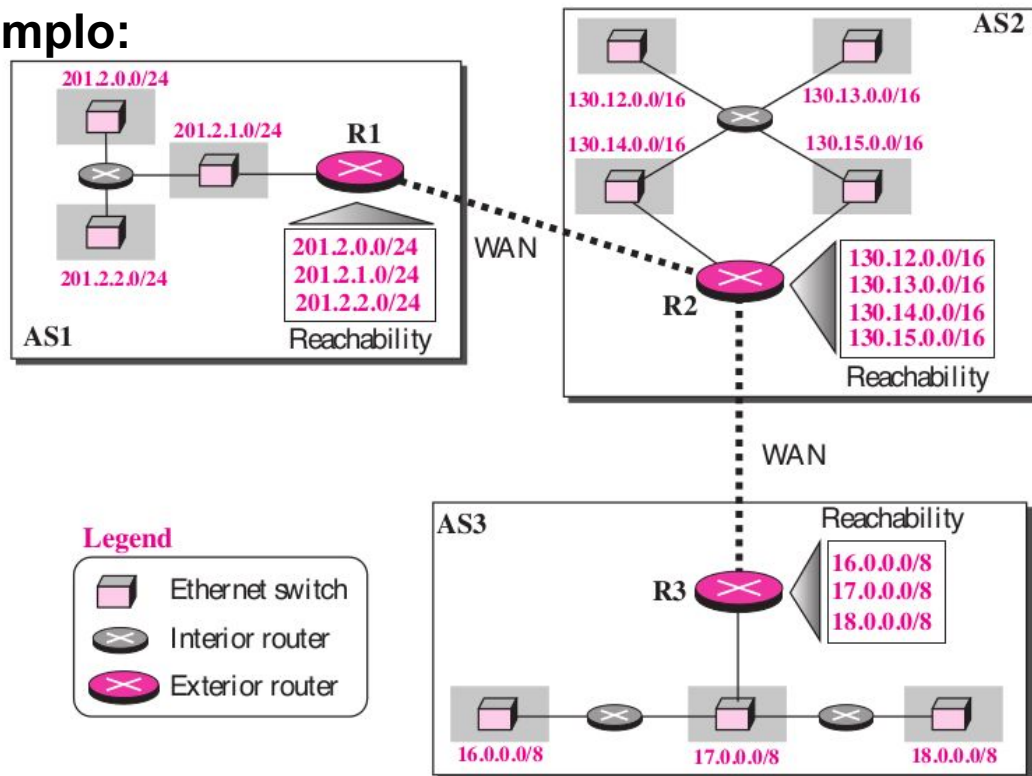
Vector de Rutas. BGP

Vector de Rutas

- Las técnicas anteriores no pueden aplicarse para el encaminamiento inter-AS
 - El encaminamiento por *vector de distancias* presenta inestabilidades con pocos saltos entre las redes y problemas de convergencia
 - El encaminamiento por *estado de enlaces* converge rápidamente pero requiere el intercambio de gran cantidad de información
- El encaminamiento por vector de rutas (Path Vector Routing) se basa en el encaminamiento por vector de distancias, e intenta resolver los problemas de convergencia para el encaminamiento inter-AS
 - A partir de la información sobre los destinos alcanzables en el AS, mediante un proceso de intercambio, cada encaminador obtiene:
 - La lista de destinos (redes) alcanzables
 - La *ruta* completa al destino, como lista de ASs que han de atravesarse
 - Detección de bucles sencilla, descartando las rutas en las que el propio AS ya es parte del camino
 - Permite implementar políticas comprobando si un determinado AS es parte de la ruta
 - Uso de CIDR para agregar direcciones de red en las tablas de rutas

Vector de Rutas

Ejemplo:



Después del intercambio de todas las tablas

R1

Network	Path
201.2.0.0/22	AS1 (ThisAS)
130.12.0.0/14	AS1, AS2
16.0.0.0/6	AS1, AS2, AS3

Path-Vector Routing Table

R2

Network	Path
201.2.0.0/22	AS2, AS1
130.12.0.0/14	AS2 (ThisAS)
16.0.0.0/6	AS2, AS3

Path-Vector Routing Table

R3

Network	Path
201.2.0.0/22	AS3, AS2, AS1
130.12.0.0/14	AS3, AS2
16.0.0.0/6	AS3 (ThisAS)

Path-Vector Routing Table

Border Gateway Protocol (BGP)

- Protocolo de encaminamiento exterior (Inter-AS) por vector de rutas
- La función principal de un sistema BGP es intercambiar información sobre las redes alcanzables con otros sistemas BGP
 - La información incluye la lista de ASs atravesados por la propia información
 - Esta información es suficiente para construir un grafo de conectividad de ASs para las redes alcanzables, libre de bucles
 - Cada AS puede aplicar ciertas políticas para aceptar y anunciar rutas
- La comunicación entre encaminadores se realiza mediante TCP, puerto 179
- La versión actual, BGP-4, soporta CIDR y agregación de rutas

BGP: Sistemas Autónomos (AS)

- Tipos de AS:
 - **Stub**: Conectado únicamente a otro AS, es destino u origen del tráfico
 - **Tránsito**: Conectado a varios AS, permite el tráfico de *tránsito* (tráfico que viene de un AS y se dirige a otro AS)
 - **Multihomed**: Conectado a varios AS por redundancia y balanceo de carga, no permite el tráfico de *tránsito*, siendo destino y origen del tráfico
- Cada AS puede aplicar **políticas** para limitar el flujo de tráfico de datos en la red
 - Las políticas no forman parte de BGP y reflejan acuerdos de tránsito (con un coste que depende del proveedor) o de intercambio (normalmente gratuitos)
- En función de la ruta, cada AS puede, por ejemplo:
 - Configurarse como AS *multihomed* (no de tránsito), anunciando solo sus redes directamente alcanzables
 - No actuar como AS de tránsito para determinados AS
 - Evitar un determinado AS como tránsito
 - Escoger la ruta en función de la relación con otros AS (clientes, socios, proveedores...)

BGP: Funcionamiento

- Los encaminadores intercambian sus tablas de rutas cuando establecen la conexión inicial y envían actualizaciones incrementales si las tablas cambian
- Mensajes:
 - **OPEN:** Establecimiento de la sesión BGP (semipermanente)
 - Identificador de AS y de encaminador
 - Parámetros de configuración (tiempo *hold* y autenticación)
 - **UPDATE:** Actualización incremental de la información de encaminamiento
 - Cada mensaje puede incluir una red alcanzable en CIDR con sus atributos, incluida la ruta, y una lista de redes retiradas (*withdrawn*)
 - **NOTIFICATION:** Se envía a los vecinos cuando se detecta un error
 - Implica un cierre de la sesión y las rutas asociadas serán inválidas
 - Ejemplos: tiempo *hold* excedido, error en los mensajes, falta de atributos...
 - **KEEPALIVE:** Para asegurar que la sesión permanezca activa
 - En respuesta a un mensaje OPEN y periódicamente para informar de la presencia del encaminador (no usa *keepalive* de TCP)
 - Si pasado un tiempo (*hold*) no se recibe información, se cierra la sesión

BGP: Atributos

- Los mensajes UPDATE incluyen las redes alcanzables y atributos de cada ruta
 - Los atributos permiten evaluar caminos alternativos al mismo destino
 - Son generados por cada encaminador, que puede modificar los recibidos
- Tipos de atributos:
 - **Bien conocidos (*well-known*):** Deben ser admitidos por todas las implementaciones BGP
 - Pueden ser obligatorios (*mandatory*) o discrecionales (*discretionary*)
 - Los atributos obligatorios se deben incluir en cada actualización
 - **Opcionales:** Son específicos de cada implementación
 - Pueden ser transitivos (*transitive*) o no
 - Los atributos transitivos se debe incluir en las actualizaciones aunque no sean implementados por el encaminador
- Ejemplos de atributos bien conocidos y obligatorios (*well-known mandatory*):
 - **ORIGIN:** Origen de la información de ruta (IGP, EGP o INCOMPLETE). No debe modificarse por otro encaminador BGP
 - **AS_PATH:** La ruta como secuencia de ASs
 - **NEXT_HOP:** Dirección IP del siguiente salto para alcanzar el destino

Arquitectura de Internet

