

# Routing Multicast

Area de Ingeniería Telemática  
<http://www.tlm.unavarra.es>

Grado en Ingeniería en Tecnologías de  
Telecomunicación, 3º

upna

Universidad Pública de Navarra  
Nafarroako Unibertsitate Publikoa

**Tecnologías Avanzadas de Red**  
*Área de Ingeniería Telemática*

# DVMRP

# DVMRP

- Distance Vector Multicast Routing Protocol (RFC 1075)
- El primer protocolo de enrutamiento multicast
- Distance Vector, inspirado en RIP
- Métrica número de saltos (infinito = 32) así que no sirve para la Internet, solo como IGP
- Periódicamente cada mrouter anuncia a sus vecinos las redes origen y el coste hasta ellas
- (...)

**DVMRP**

DVMRP

# DVMRP

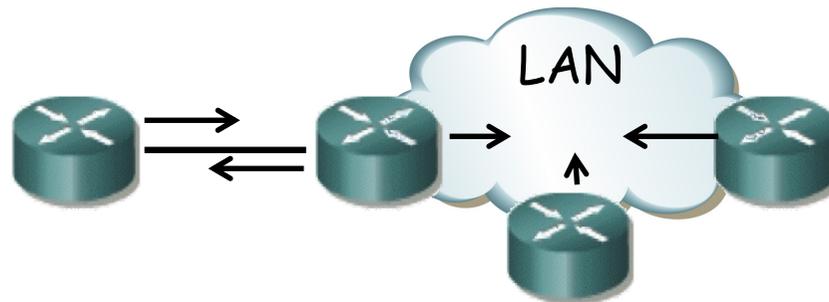
- Hace un RPF check en base a su propia tabla (otros protocolos usan la de encaminamiento unicast)
- Construye *source-based multicast trees*
- Classless
- Lenta convergencia
- Emplea un tipo de mensajes IGMP

**DVMRP**

DVMRP

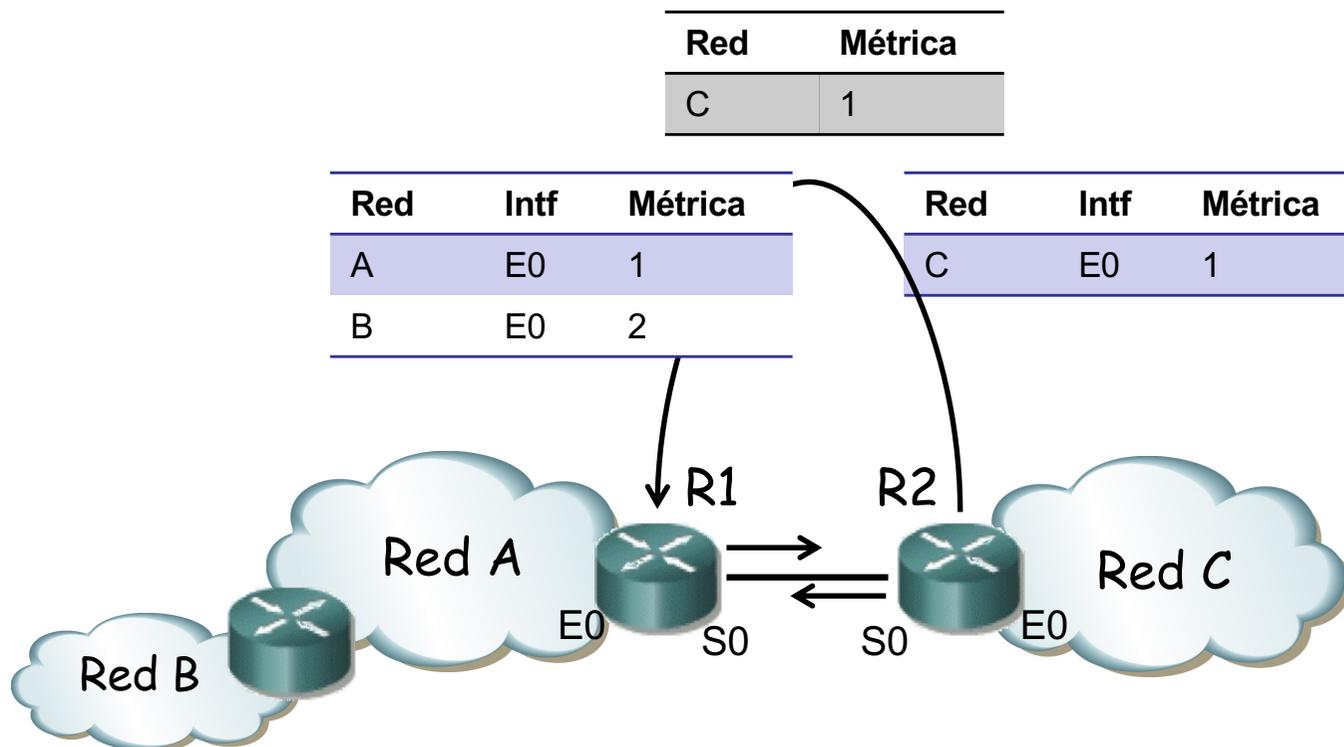
# Neighbor Discovery

- Mantiene adyacencias con routers DVMRP vecinos
- Envía periódicamente mensajes a 224.0.0.4 (*DVMRP Routers*)
- Cada uno envía en el mensaje la lista de routers que ha visto por ese interfaz
- Se verifica la bidireccionalidad del enlace viéndose listado en el anuncio del vecino



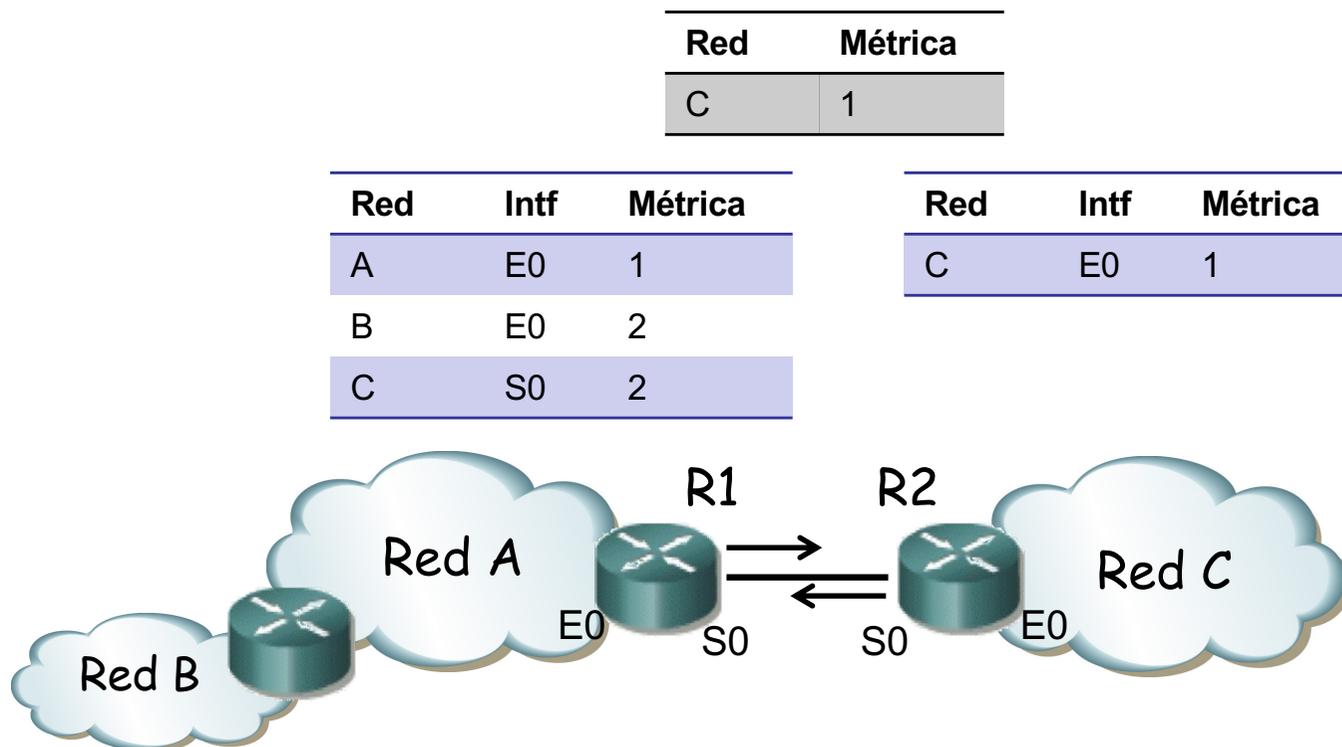
# Route Reports

- Cada mrouter tiene una lista de redes con el coste hasta ellas y el interfaz por el cual las alcanza
- Envía periódicamente a los vecinos vector de distancias a esas redes
- (... ..)



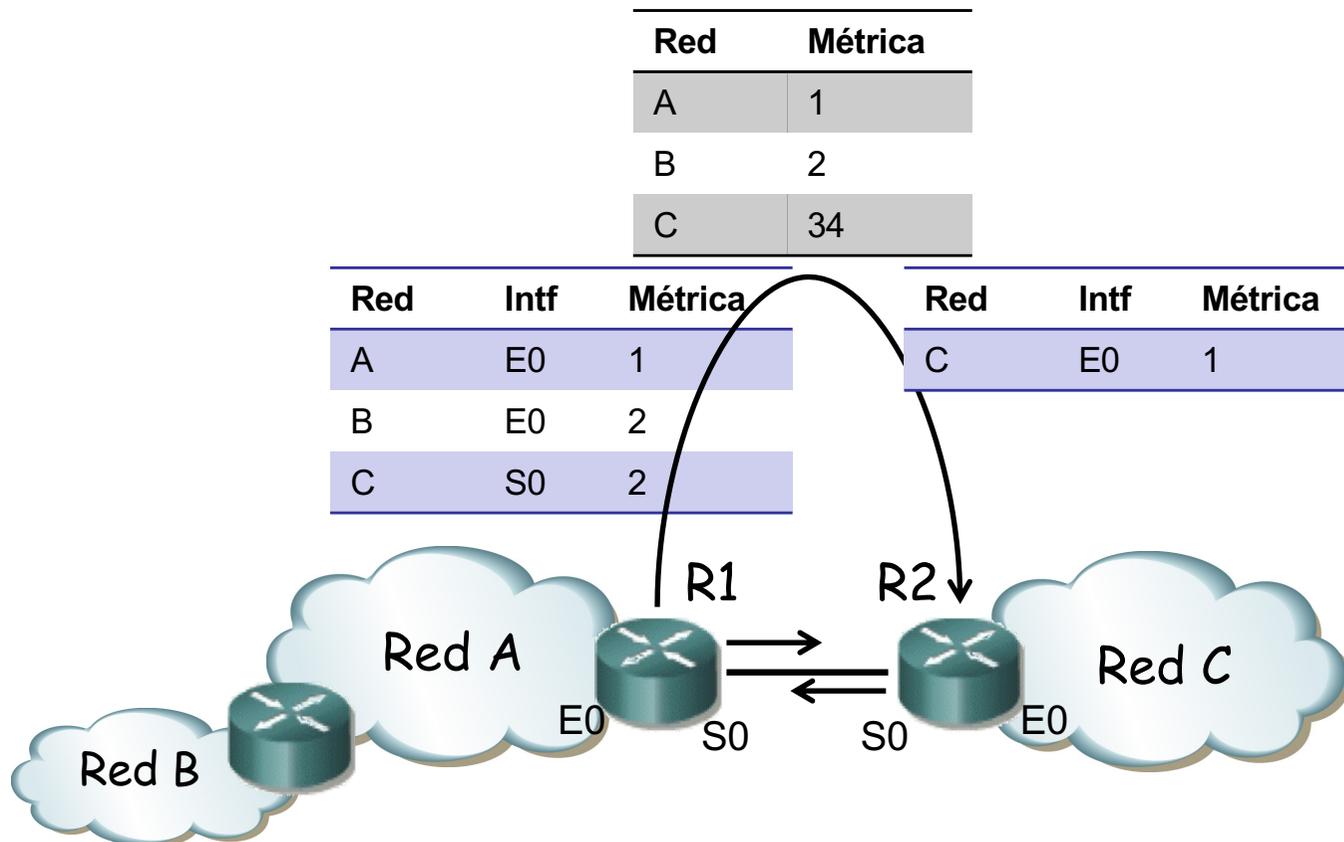
# Route Reports

- Actualiza al recibir vector de forma similar a RIP (ecuación de BF)
- Destinos a los que se llega por un interfaz los anuncia con la métrica incrementada en 32 (es una especie de *poison reverse*) (. . .)



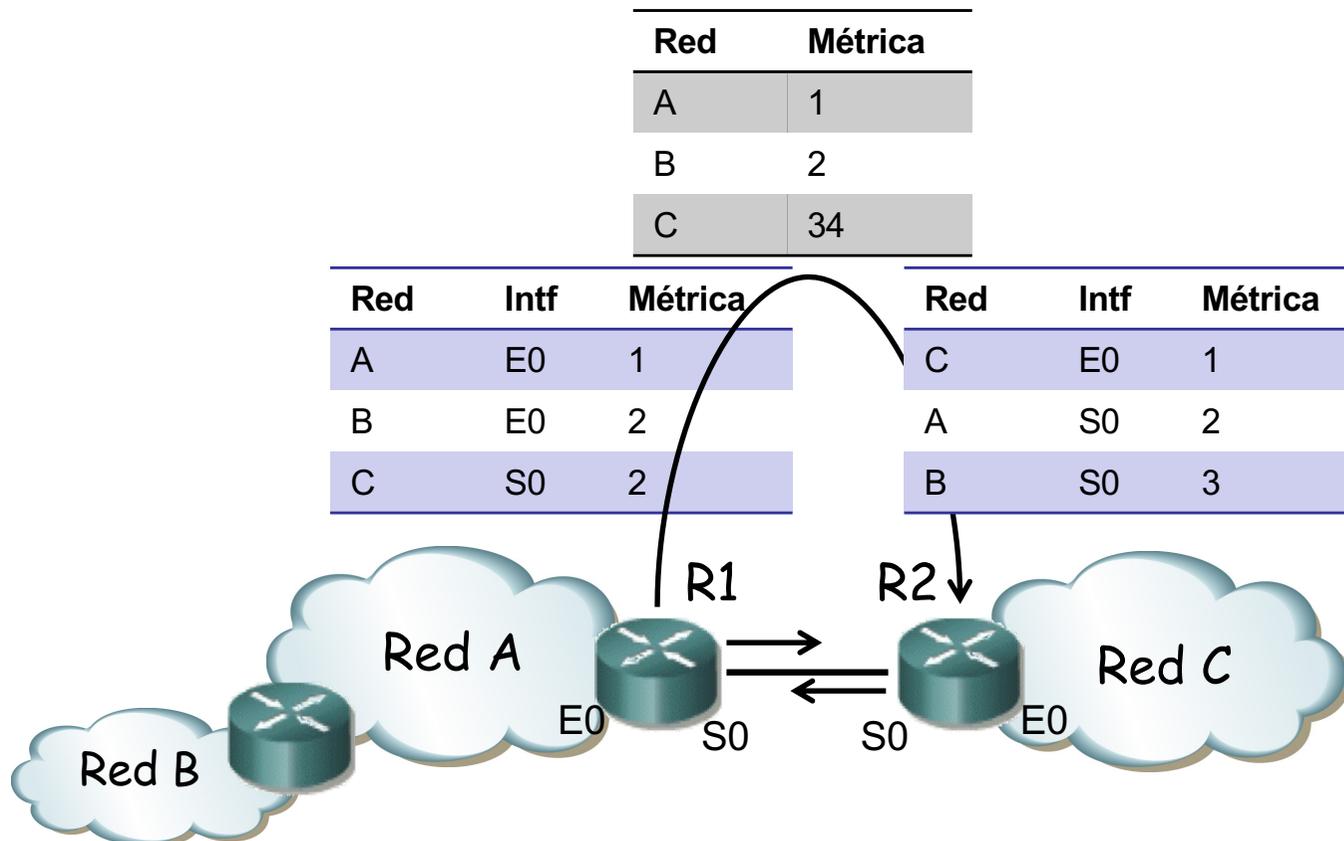
# Route Reports

- Actualiza al recibir vector de forma similar a RIP (ecuación de BF)
- Destinos a los que se llega por un interfaz los anuncia con la métrica incrementada en 32 (es una especie de *poison reverse*) (. . .)



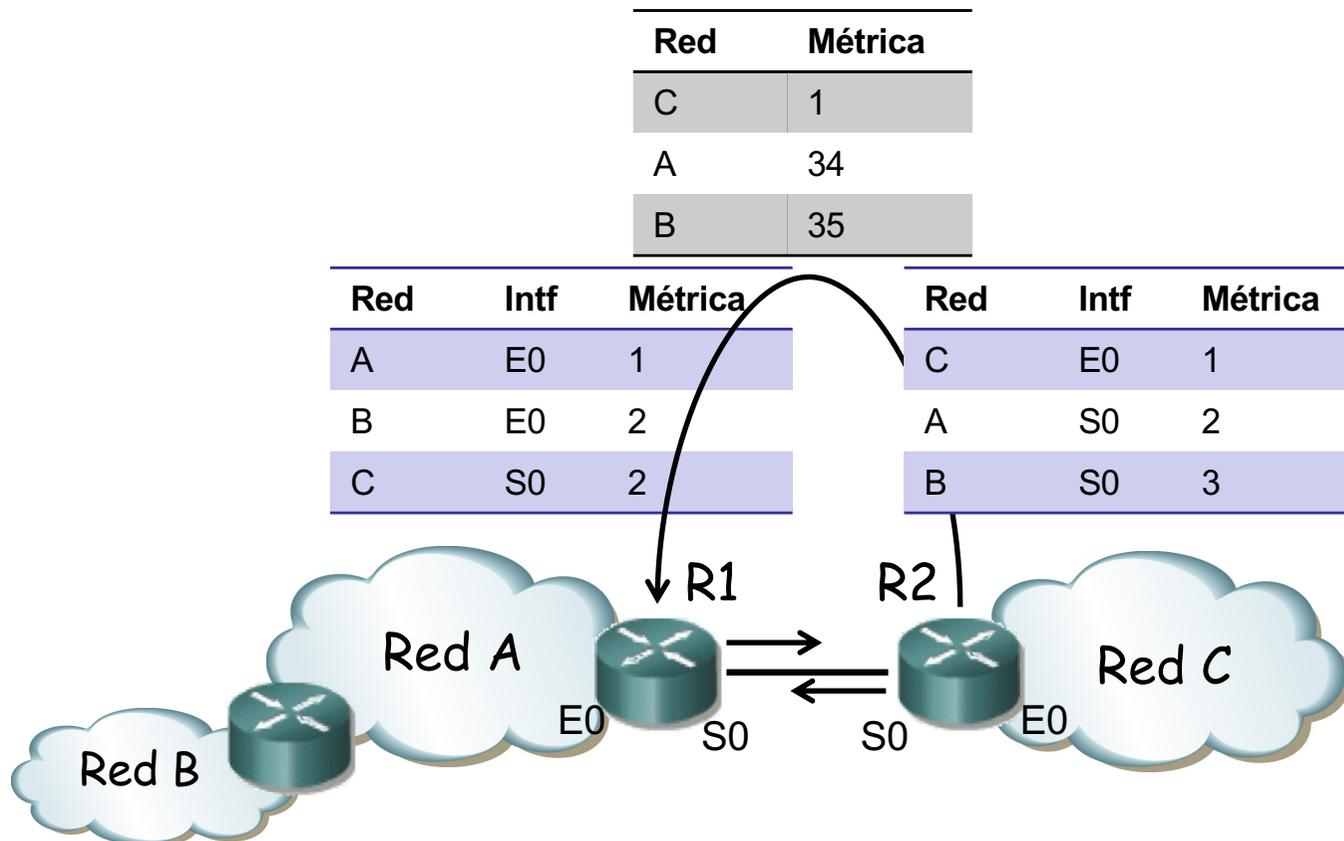
# Route Reports

- Actualiza al recibir vector de forma similar a RIP (ecuación de BF)
- Destinos a los que se llega por un interfaz los anuncia con la métrica incrementada en 32 (es una especie de *poison reverse*) (. . .)



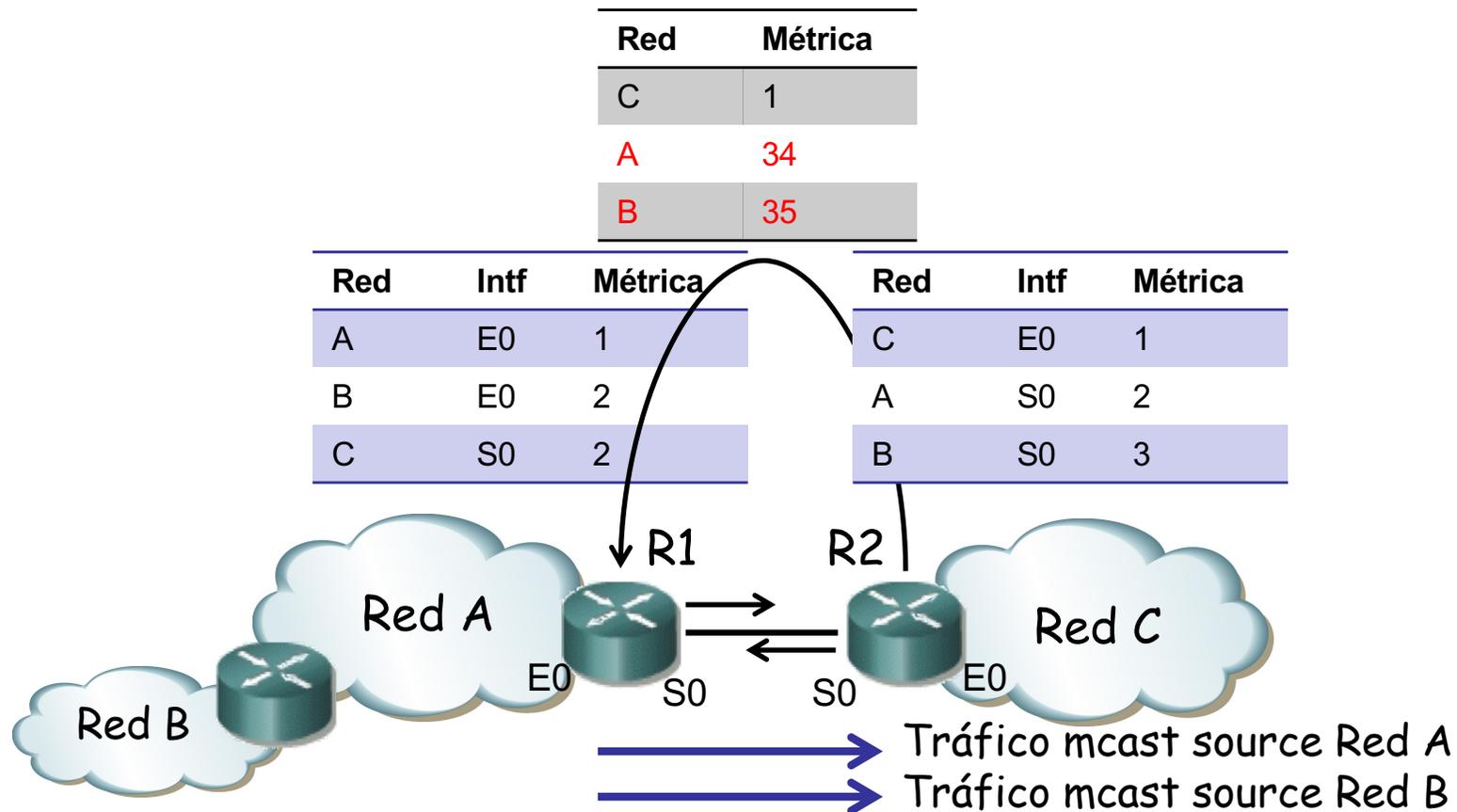
# Route Reports

- Actualiza al recibir vector de forma similar a RIP (ecuación de BF)
- Destinos a los que se llega por un interfaz los anuncia con la métrica incrementada en 32 (es una especie de *poison reverse*) (. . .)

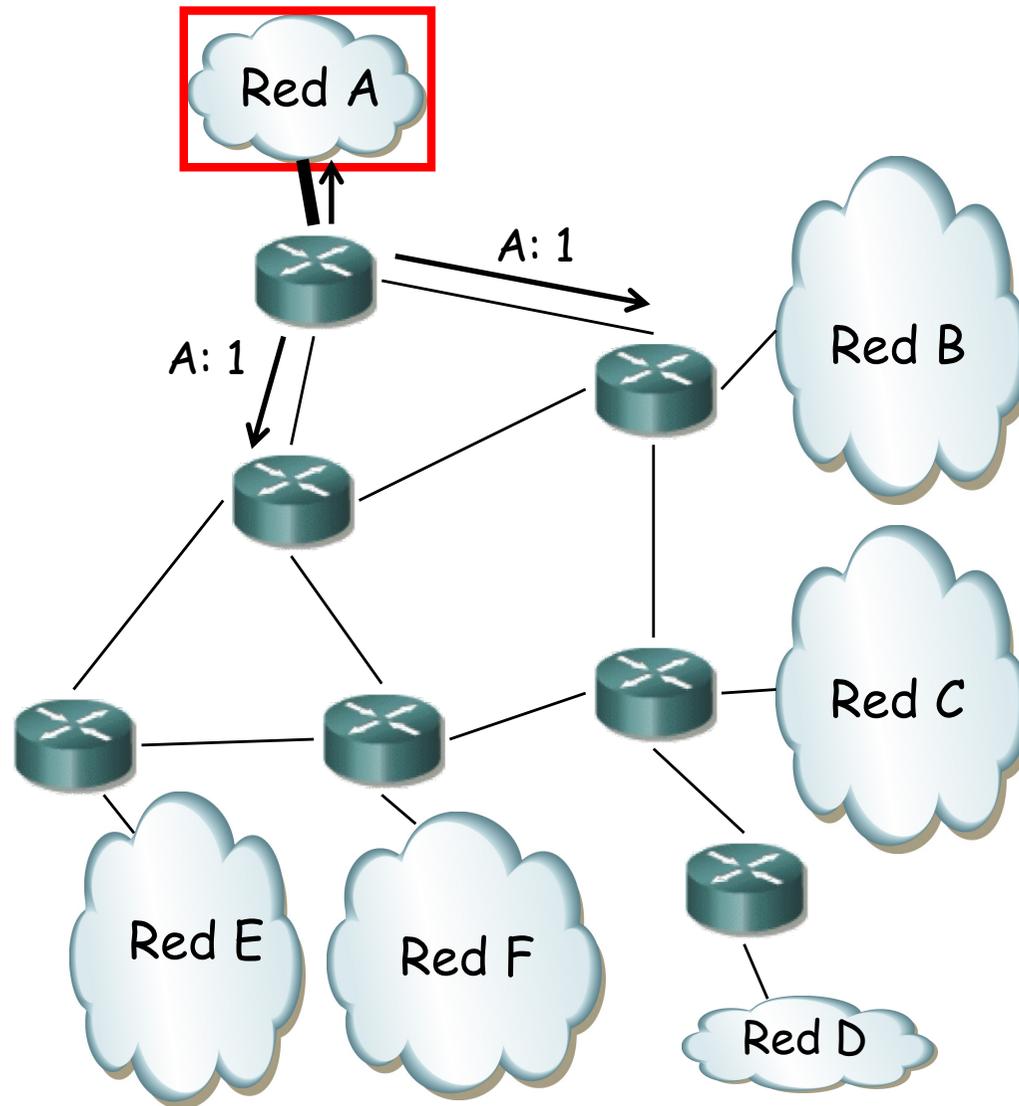


# Truncated Broadcast Trees

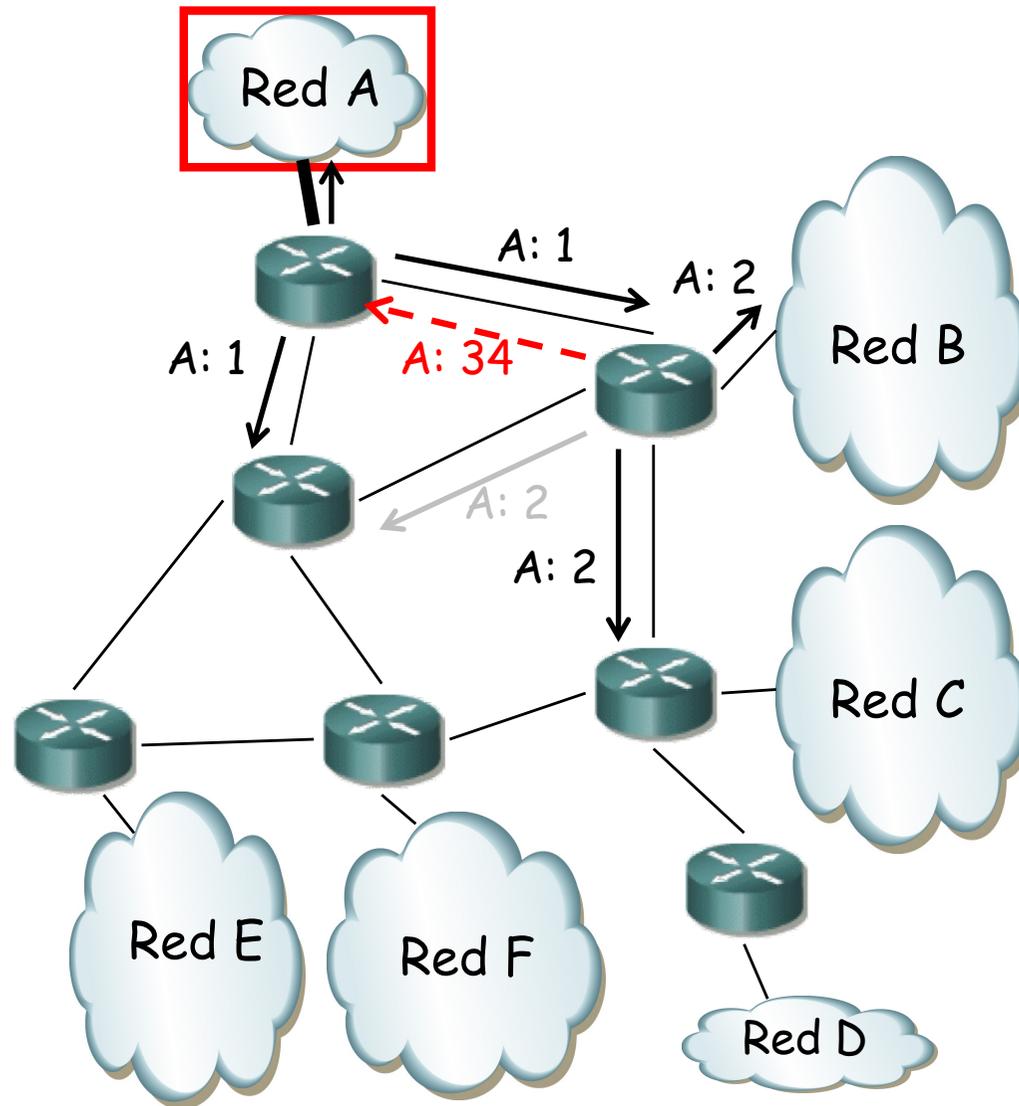
- DVMRP es de tipo denso, calculando *source trees*
- Los calcula en base a los anuncios *poison reversed*
- Un mrouter envía al router *upstream* un anuncio con las redes para las que está *downstream* con la métrica incrementada en 32
- El mrouter *upstream* aprende que debe reenviar a ese mrouter el tráfico que venga de esos orígenes



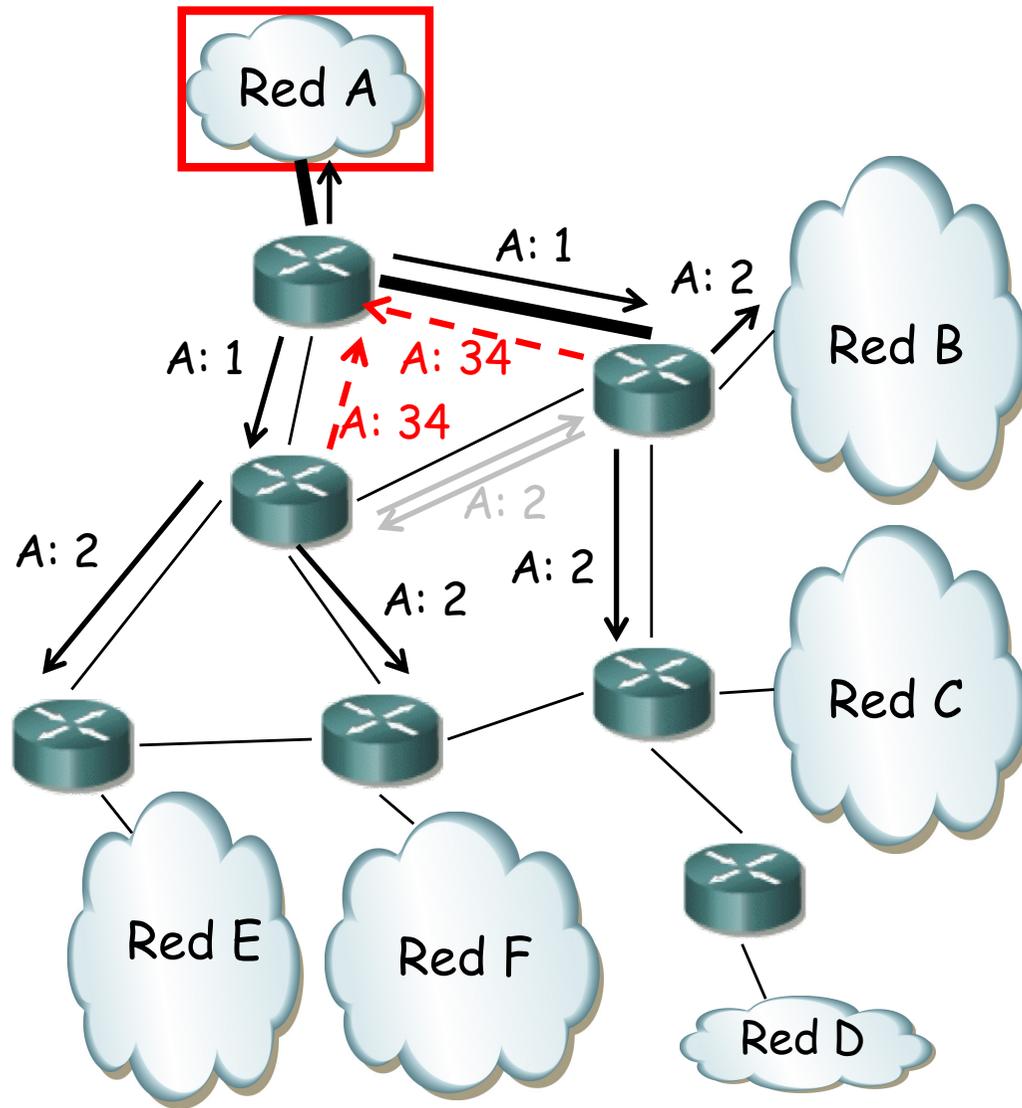
# Ejemplo



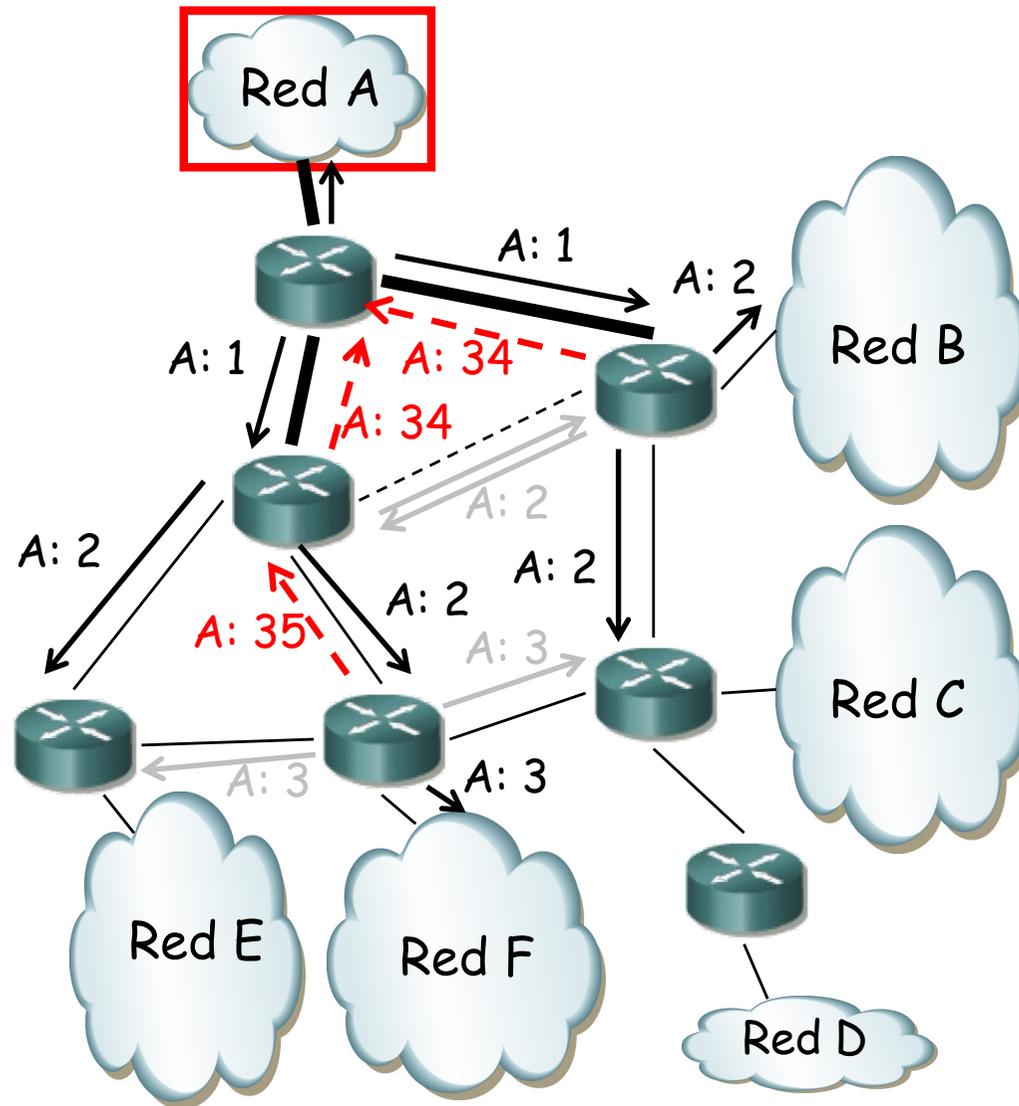
# Ejemplo



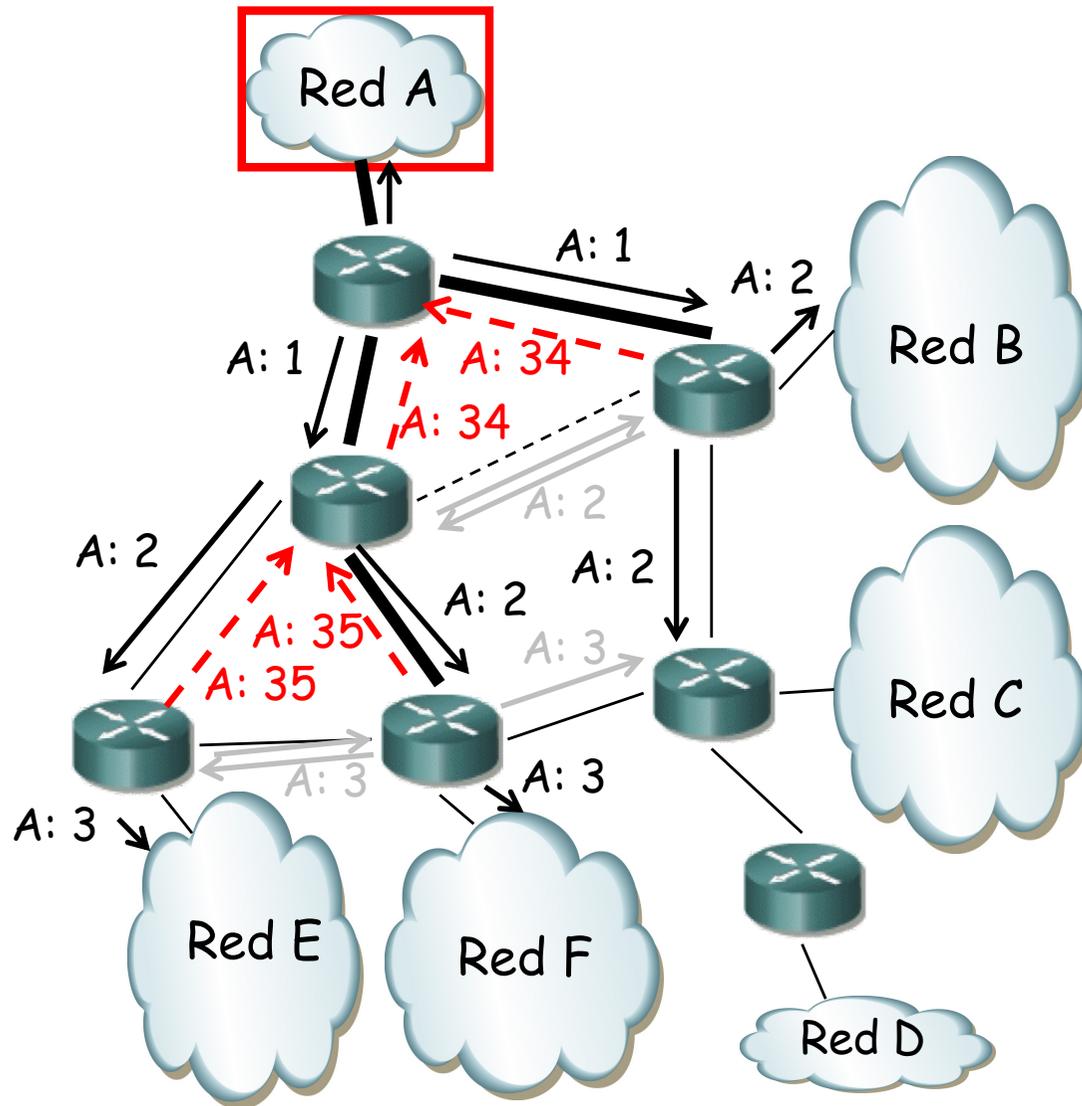
# Ejemplo



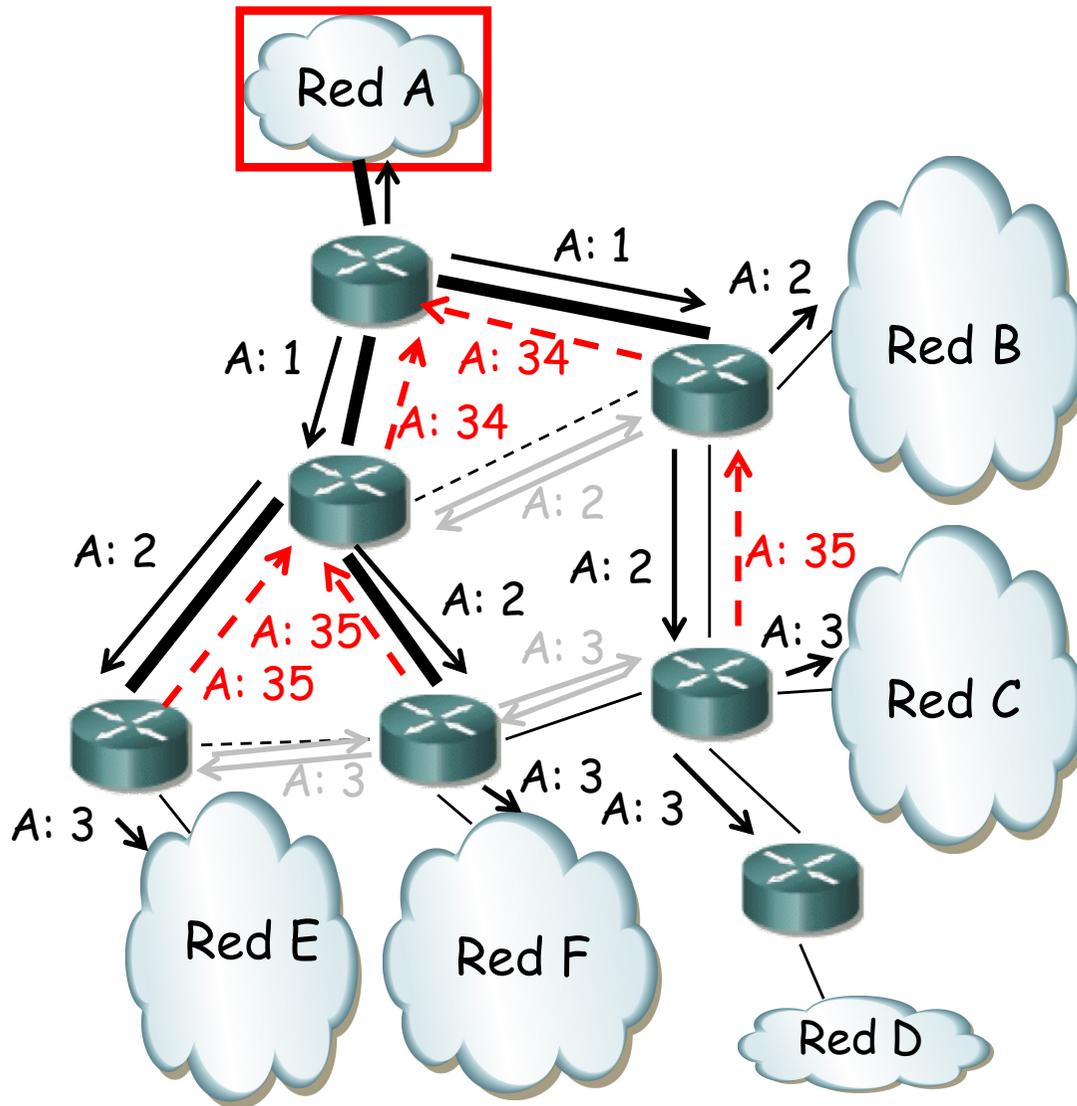
# Ejemplo



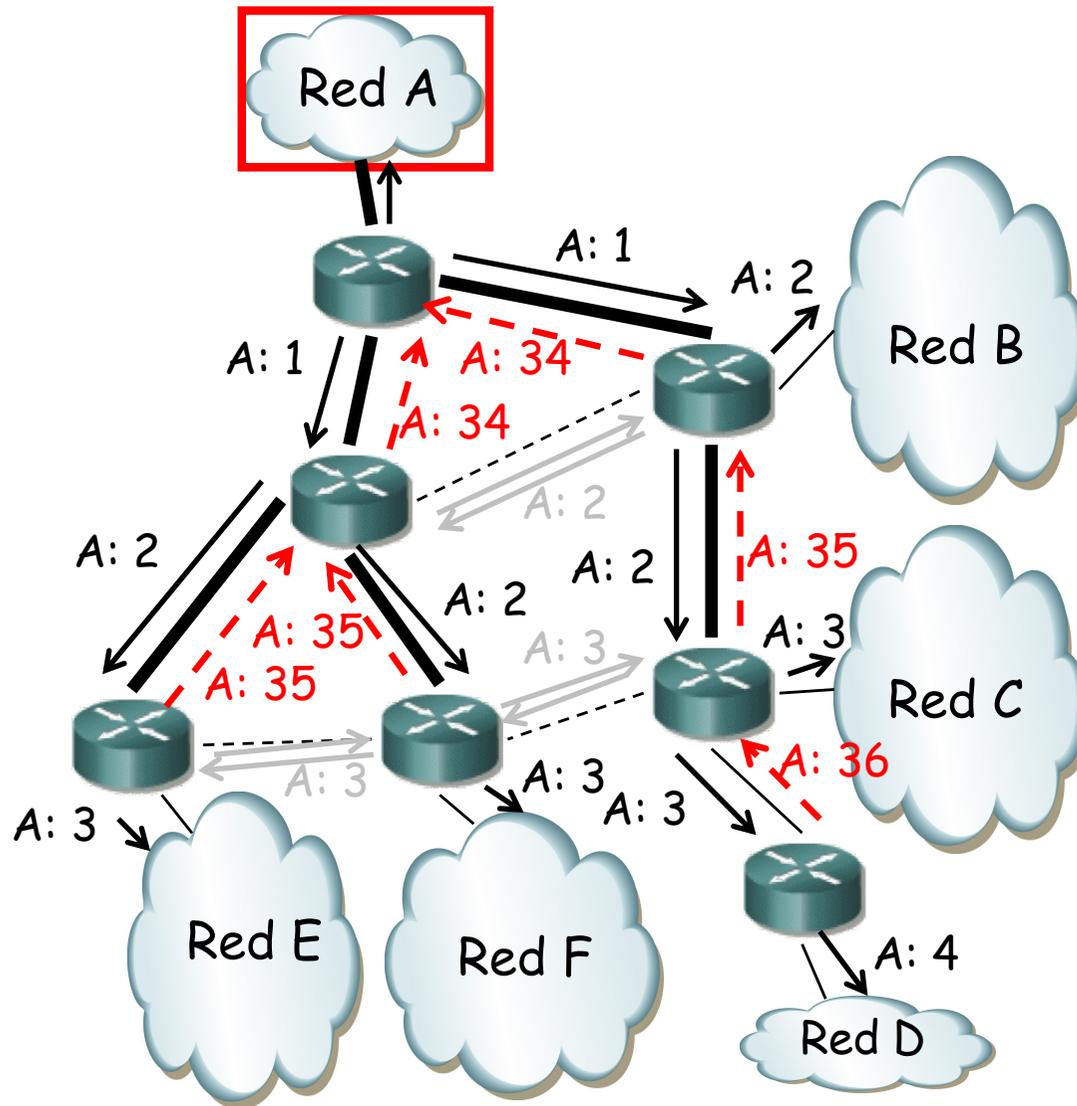
# Ejemplo



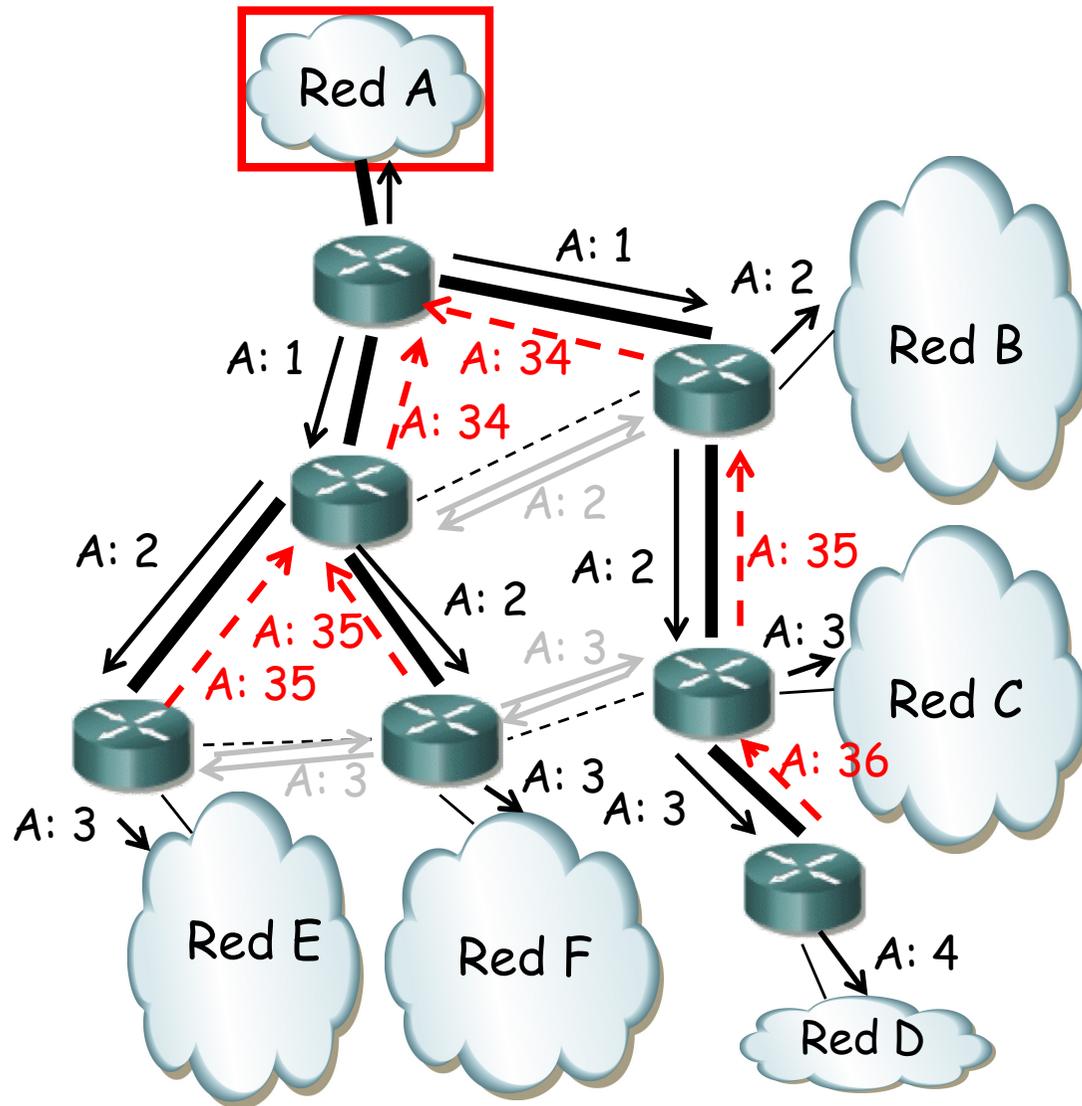
# Ejemplo



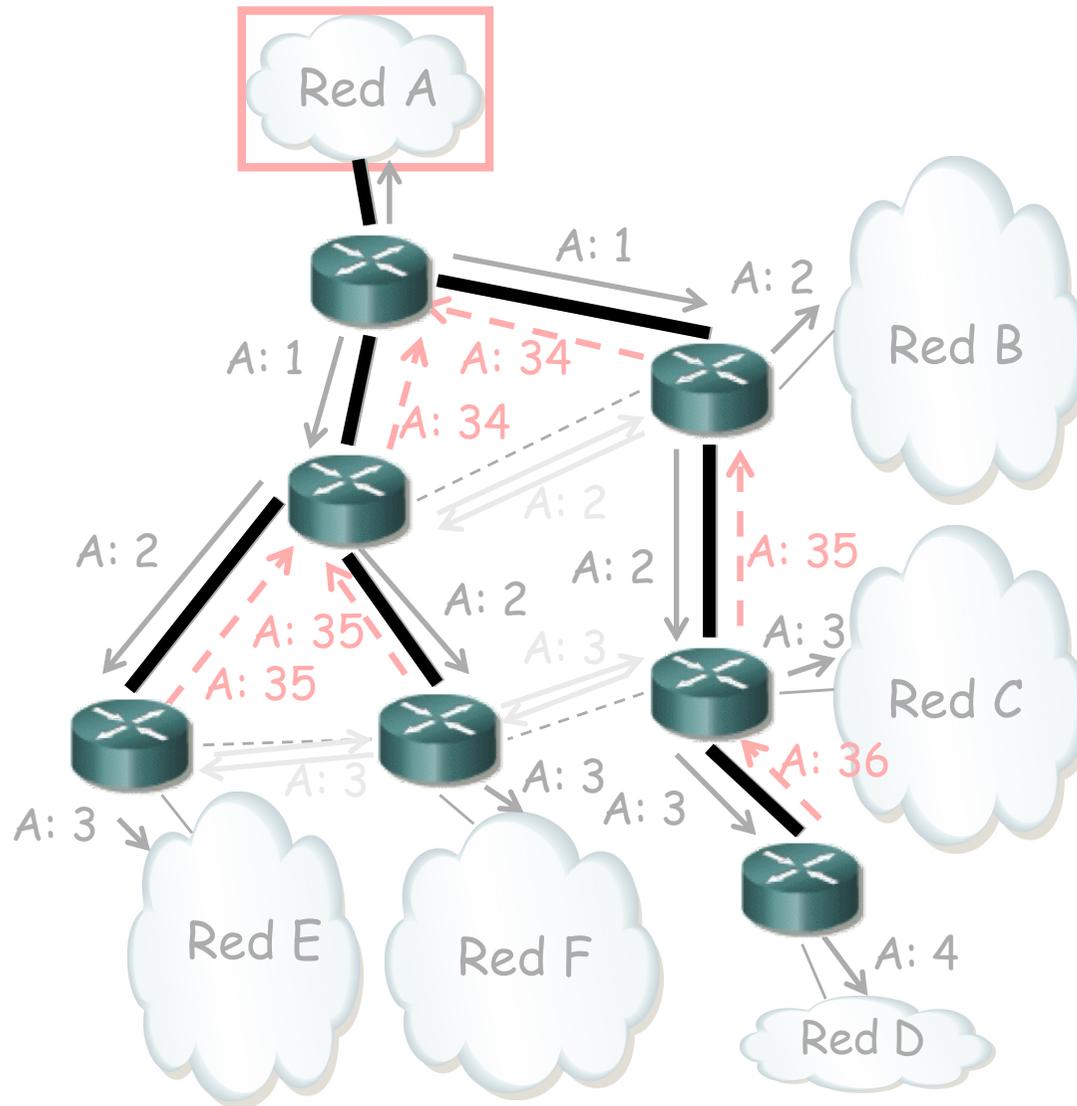
# Ejemplo



# Ejemplo

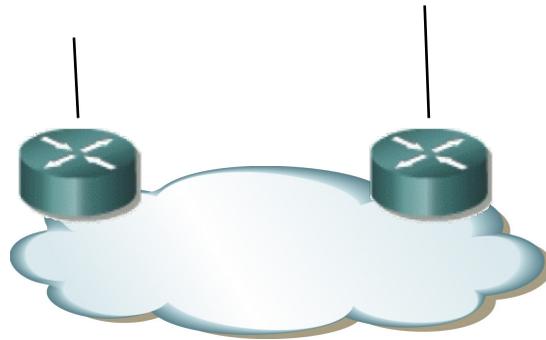


# Ejemplo



# Truncated Broadcast Trees

- Es un árbol para cada red origen
- En una LAN con varios mrouter
  - Solo uno introducirá el tráfico (*designated forwarder*)
  - El que tenga el menor coste a la fuente
  - En caso de igual coste el que tenga menor dirección IP



# *Pruning & grafting*

- ***Pruning***

- M routers extremo sin receptores envían mensaje de poda *upstream*
- Ascende por el árbol mientras los m routers no tengan clientes *downstream*
- La poda caduca y se reactivaría el *flooding* si no se retransmite la poda

- ***Grafting***

- Enviado *upstream* para reactivar *flooding* sin esperar a que caduque poda
- Los mensajes son confirmados por el nodo *upstream*

upna

Universidad Pública de Navarra  
Nafarroako Unibertsitate Publikoa

**Tecnologías Avanzadas de Red**  
*Área de Ingeniería Telemática*

# PIM: Introducción

# PIM

- *Protocol Independent Multicast* (independiente del protocolo de enrutamiento unicast que se emplee)
- Emplea la tabla de rutas del protocolo unicast para el RPF check
- Classless (si lo es el enrutamiento unicast)
- PIMv1 sobre IGMP, PIMv2 sobre IP (protocolo 103, TTL=1)

```
⊕ Frame 86 (142 bytes on wire, 142 bytes captured)
⊕ Ethernet II, Src: cc:06:07:e4:f0:01 (cc:06:07:e4:f0:01), Dst: cc:05:07:e4:f0:00 (cc:05:07:e4:f0:00)
⊕ Internet Protocol, Src: 192.168.0.6 (192.168.0.6), Dst: 192.168.1.254 (192.168.1.254)
⊖ Protocol Independent Multicast
    Version: 2
    Type: Register (1)
    Checksum: 0xdeff [correct]
⊖ PIM parameters
    ⊖ Flags: 0x00000000
        0... .. = Not border
        .0.. .. = Not Null-Register
⊕ Internet Protocol, Src: 192.168.20.10 (192.168.20.10), Dst: 239.1.2.3 (239.1.2.3)
⊕ Internet Control Message Protocol
```

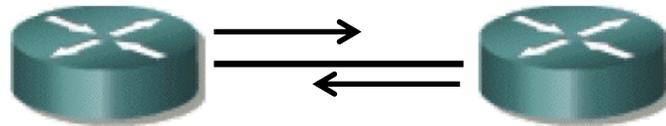
# Dos modos

- *Dense Mode (PIM-DM, RFC 3973)*
  - *Source Distribution Trees (flood-and-prune)*
  - Se supone que hay receptores por toda la red
- *Sparse Mode (PIM-SM, RFC 4601)*
  - *Source or Shared Trees*
  - *Rendezvous point*
  - Los receptores son pocos



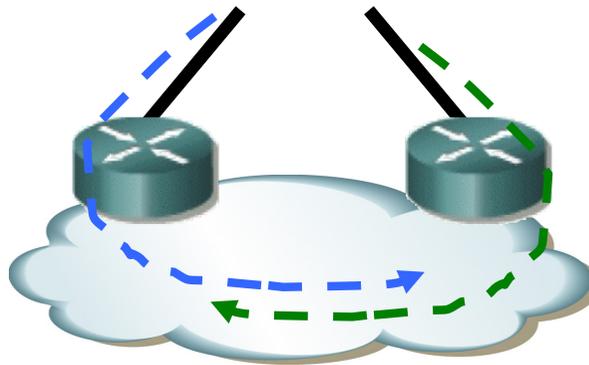
# *Neighbor discovery*

- Mensaje PIM Hello cada 30s dirigido a 224.0.0.13 (*All PIM Routers*)
- Mantiene adyacencia durante 90s si no recibe de nuevo el Hello
- En redes multiacceso sirve para seleccionar un *Designated Router* que es el de mayor prioridad (o mayor dirección IP)



# Asserts

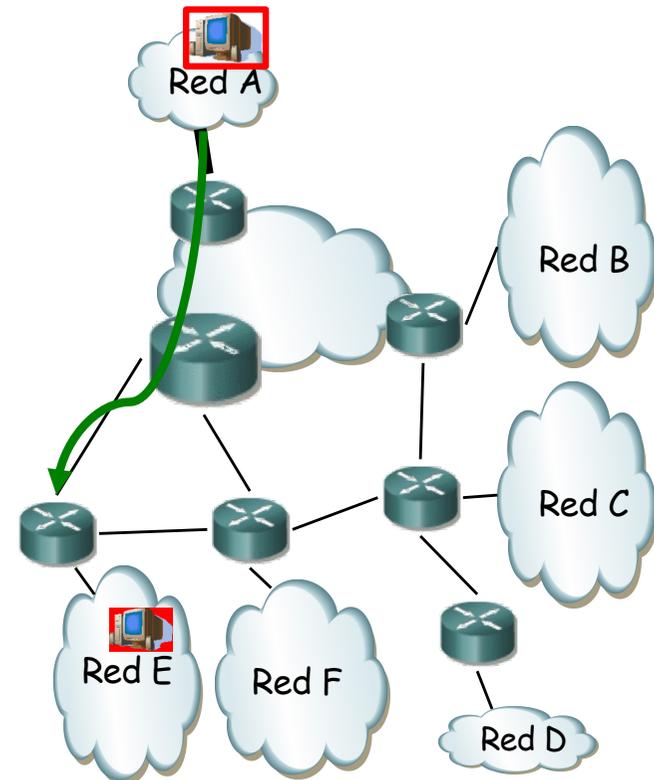
- Cuando un router recibe tráfico multicast por interfaz de salida
- Envían mensaje de *Assert* para decidir quién sigue reenviando
- Será el de menor coste al origen (o mayor IP si empatan)



# PIM-DM (Dense Mode)

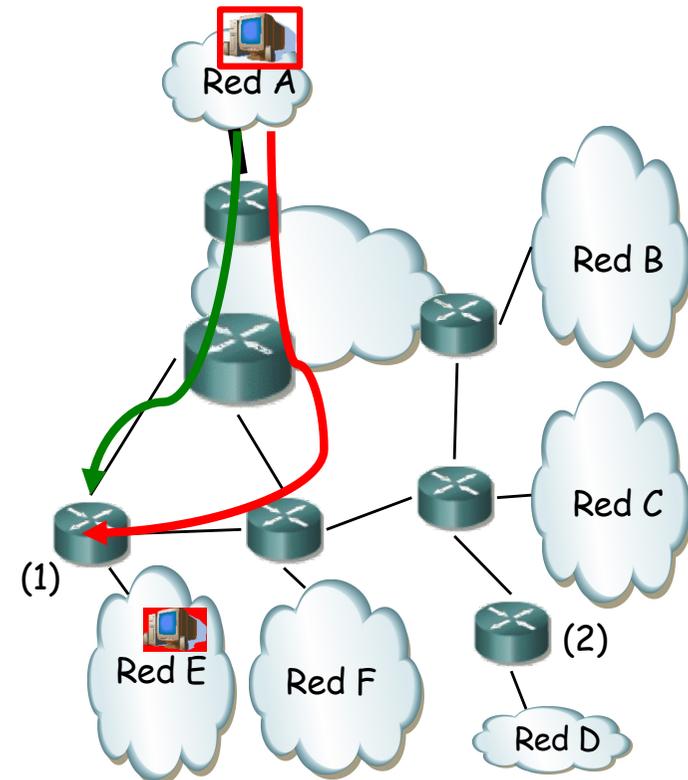
# PIM-DM: RPF

- Inicialmente inunda la red, usando RPF
- Si hay más de una entrada en la tabla de rutas para ese destino (ECMP) comprueba con la de mayor dirección de next-hop



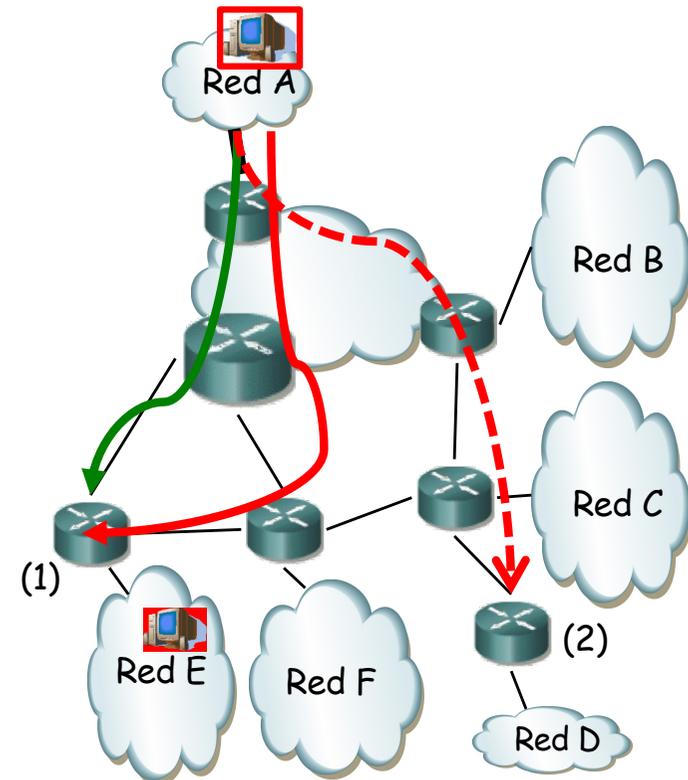
# PIM-DM: Pruning

- Para tráfico que llega por un interfaz p2p no RPF (1)
- (...)



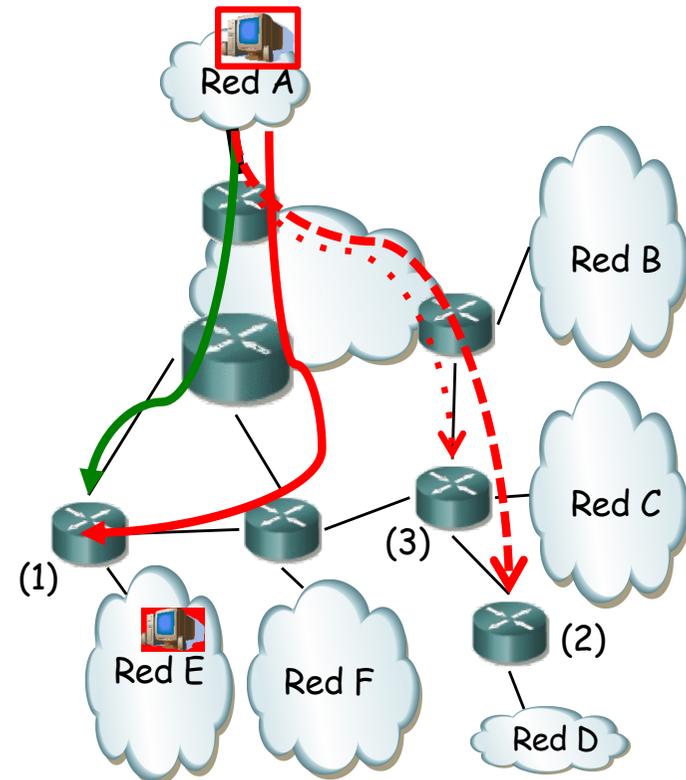
# PIM-DM: Pruning

- Para tráfico que llega por un interfaz p2p no RPF (1)
- Cuando es un router extremo sin receptores conectados (2)



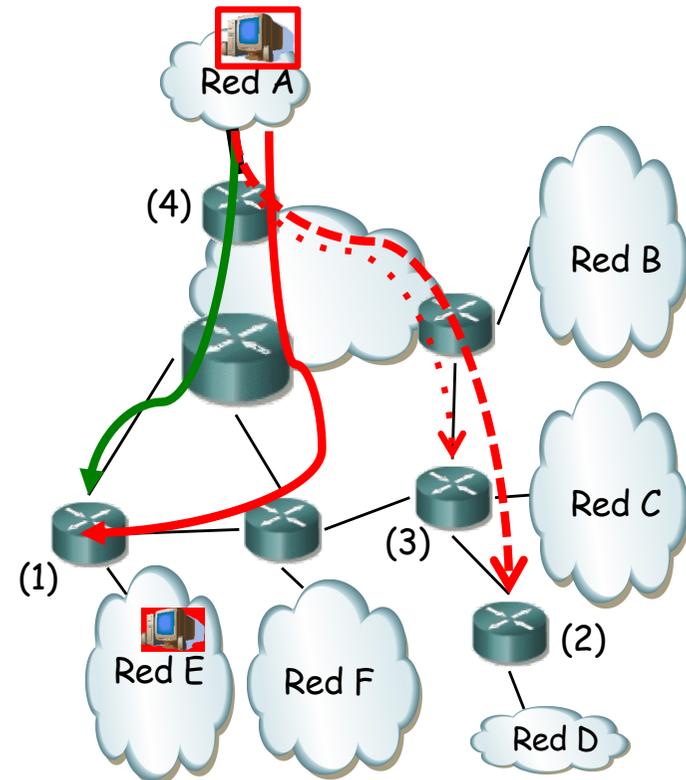
# PIM-DM: Pruning

- Para tráfico que llega por un interfaz p2p no RPF (1)
- Cuando es un router extremo sin receptores conectados (2)
- Router sin receptores conectados y que ha recibido un *Prune* de sus vecinos no RPF (3)
- (...)



# PIM-DM: Pruning

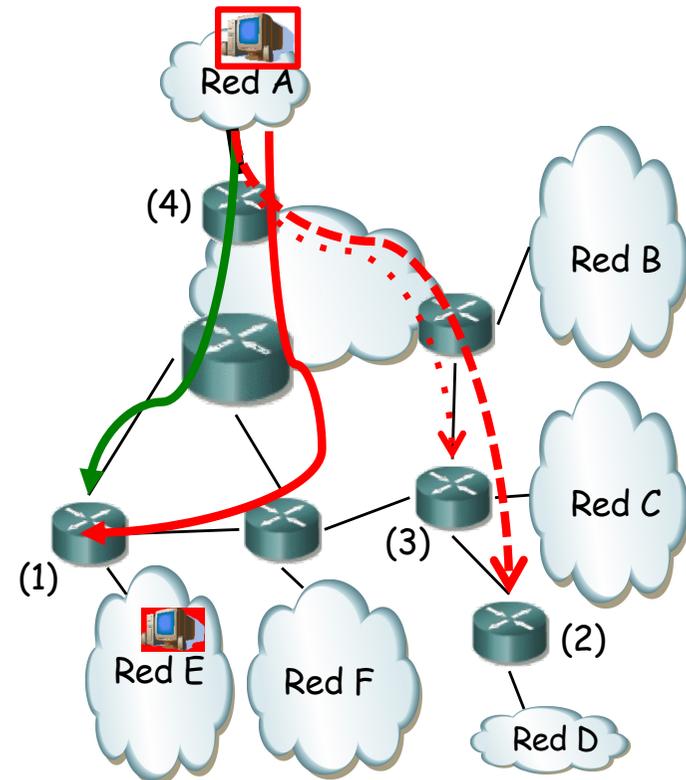
- Para tráfico que llega por un interfaz p2p no RPF (1)
- Cuando es un router extremo sin receptores conectados (2)
- Router sin receptores conectados y que ha recibido un *Prune* de sus vecinos no RPF (3)
- Router que ha recibido *Prune* de LAN, mientras otro no lo anule (*override*) con un *Join* en menos de 3s (4)
- (...)



# PIM-DM: Pruning

- El estado *pruned* caduca
- Se puede evitar con mensajes de *State Refresh* que envía el router adyacente a la fuente por el árbol refrescando las podas

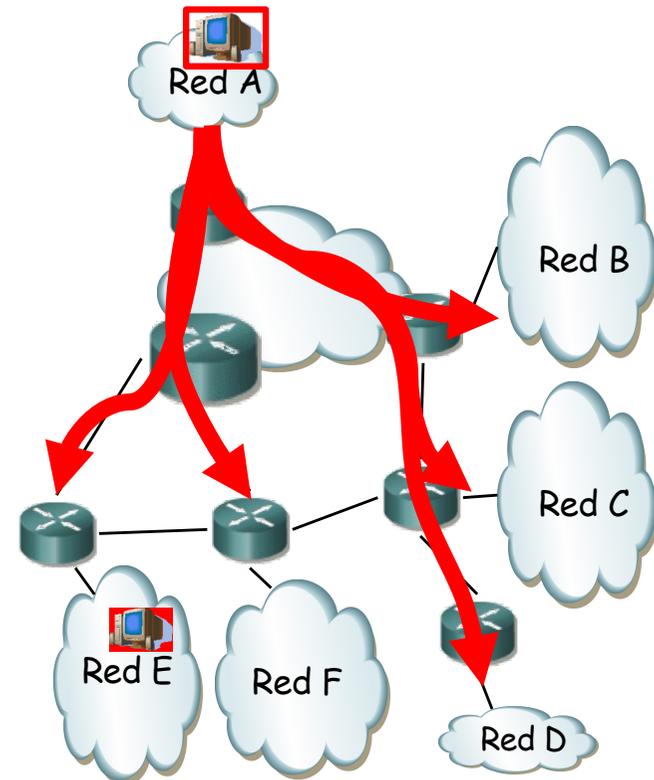
(...)





# PIM-DM

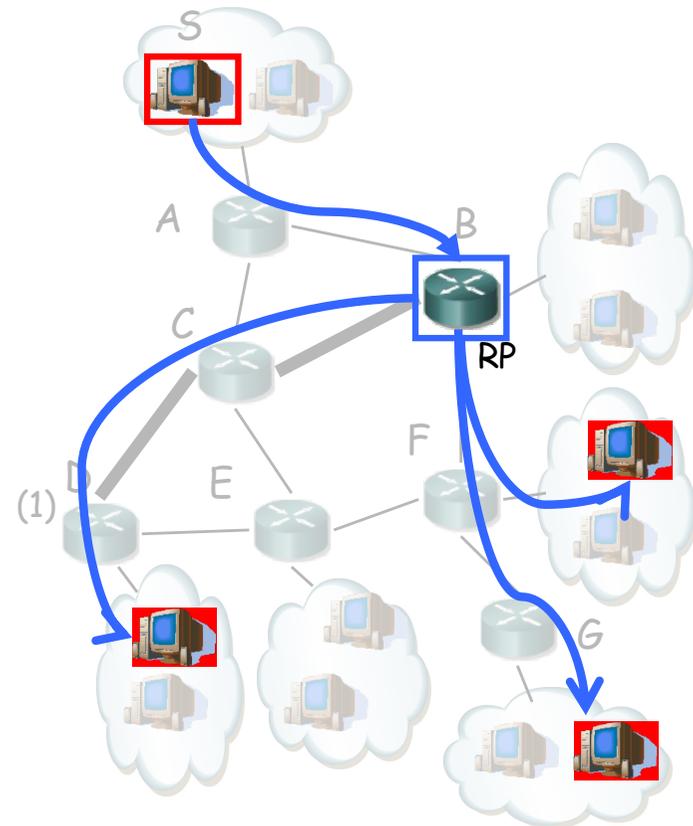
- “Inicialmente inunda la red, usando RPF”
- “El estado *pruned* caduca”
- Es decir, al tiempo vuelve a inundar
- Muchas redes no quieren tanta inundación
- Es probable que no se esté usando mucho



# PIM-SM (Sparse Mode)

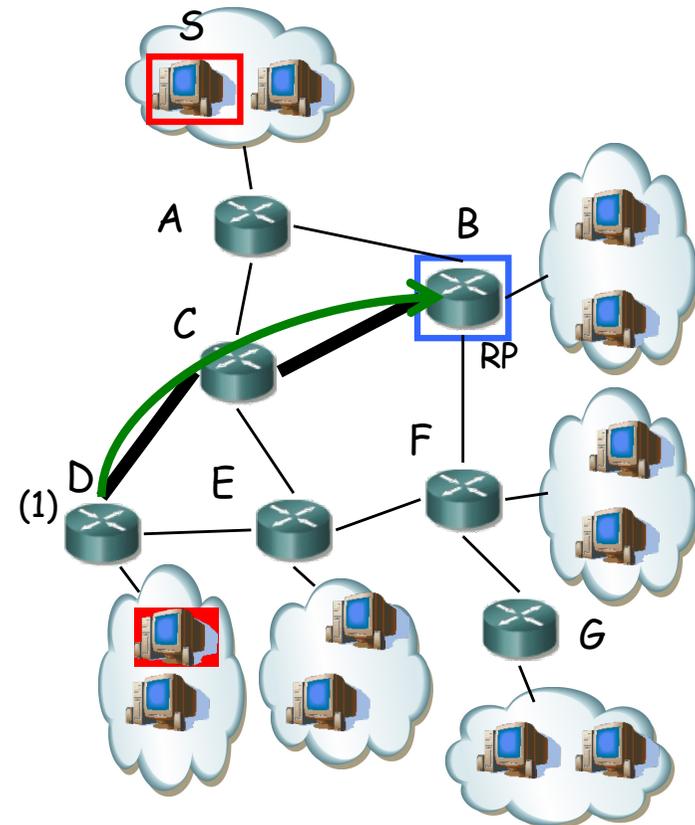
# PIM-SM: Árbol

- Fuente envía hacia un *Rendezvous Point*
- Desde el RP se envía por árbol multicast a los receptores
- Solo envía a las zonas de la red que lo han solicitado
- El *rendezvous point* (RP) es la raíz del *shared tree* y el árbol un RPT (*Rendezvous Point Tree*)



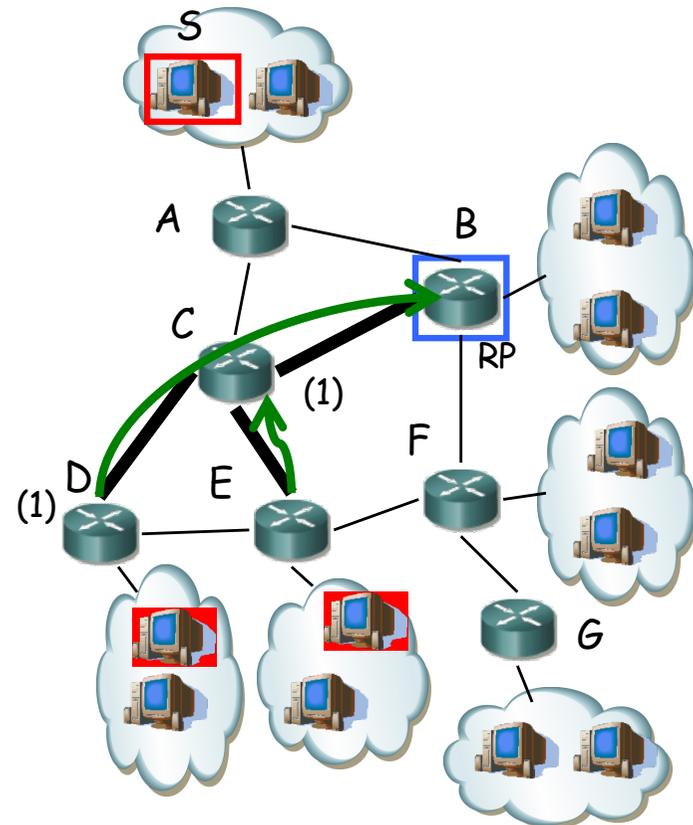
# PIM-SM: Join

- Mrouter con un receptor unido al grupo (IGMP) envía *Join* hacia el RP
- Mrouters en el camino ponen al interfaz por el que reciben el *Join* en la lista por donde reenviar para (\*,G) y envían el Join hacia el RP (1)
- (...)



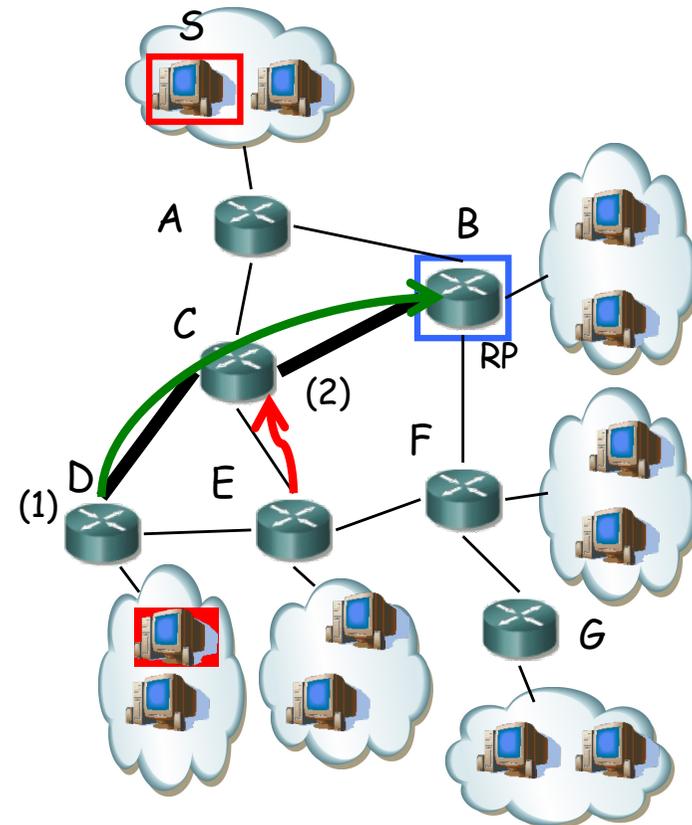
# PIM-SM: Join

- Mrouter con un receptor unido al grupo (IGMP) envía *Join* hacia el RP
- Mrouters en el camino ponen al interfaz por el que reciben el *Join* en la lista por donde reenviar para (\*,G) y envían el *Join* hacia el RP (1)
- Si mrouter ya tiene estado para (\*,G) no envía el *Join* hacia el RP
- (...)



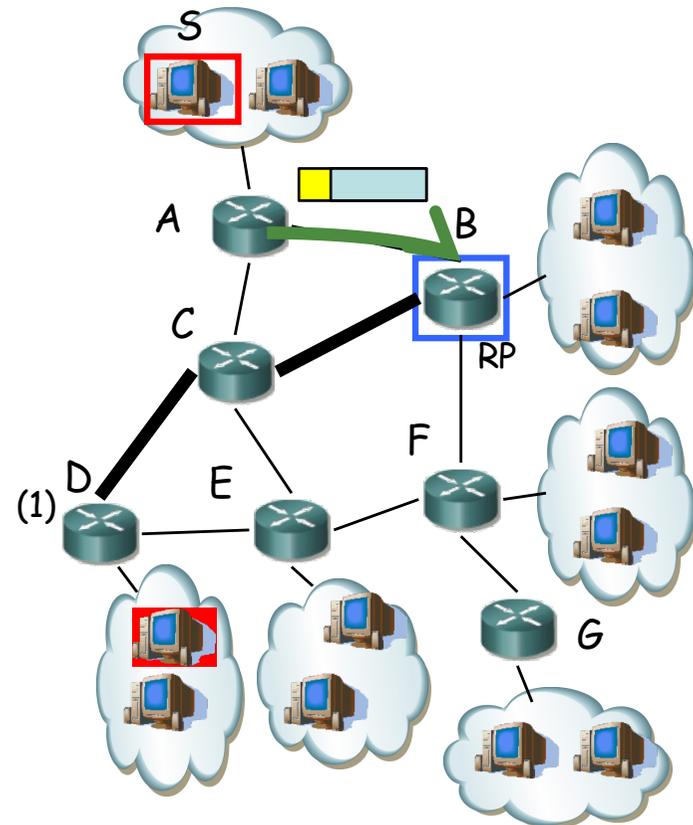
# PIM-SM: Prune

- Si un router queda sin receptores envía *Prune*
- El *Join/Prune* se reenvía periódicamente
- En LANs emplea el mismo mecanismo con *override* que PIM-DM



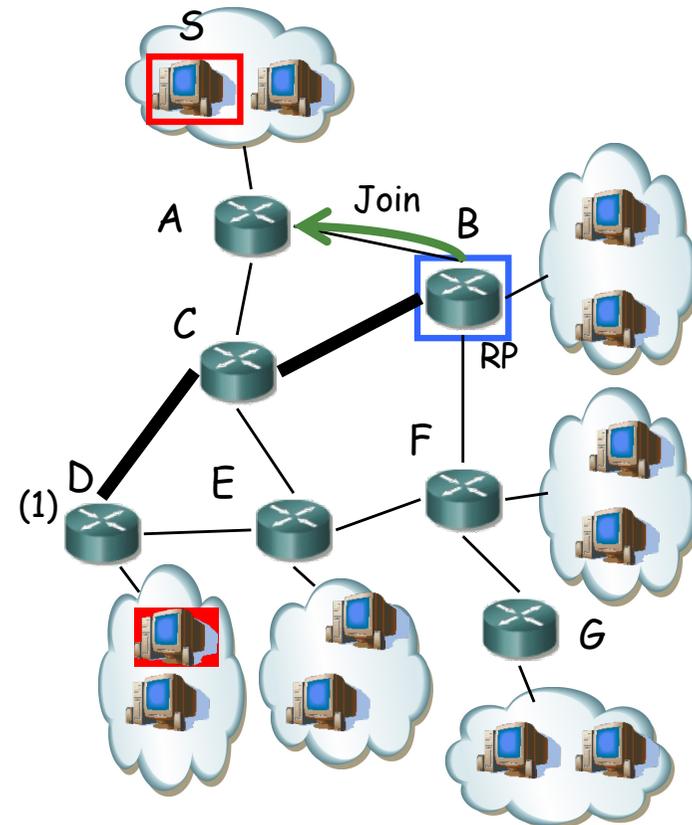
# PIM-SM: RPT

- La fuente envía los paquetes al RP encapsulados en paquetes PIM unicast a él
- El RP distribuye desde él el paquete nativo multicast por el RPT
- (...)



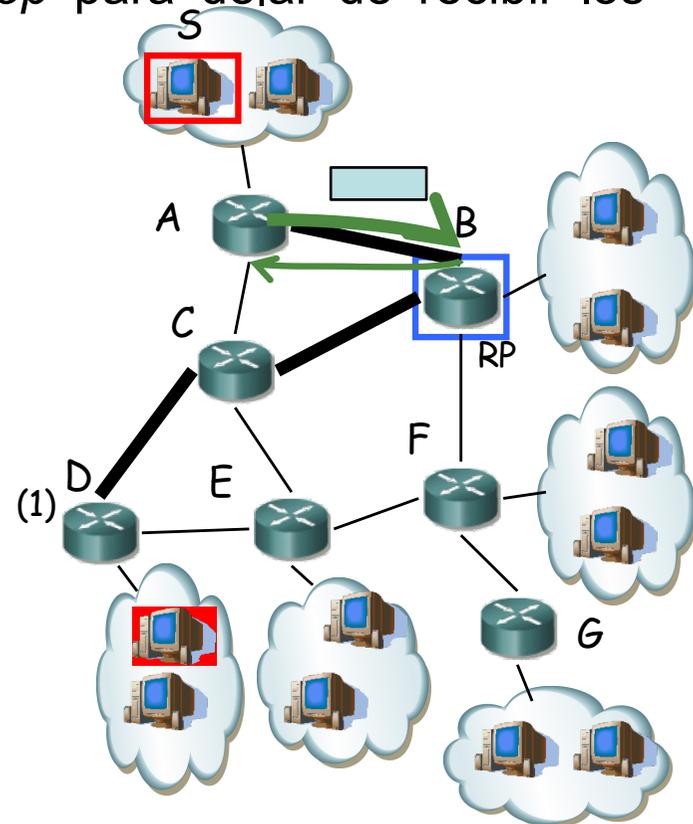
# PIM-SM: RPT

- La fuente envía los paquetes al RP encapsulados en paquetes PIM unicast a él
- El RP distribuye desde él el paquete nativo multicast por el RPT
- RP envía *Join* a (S,G) upstream
- (...)



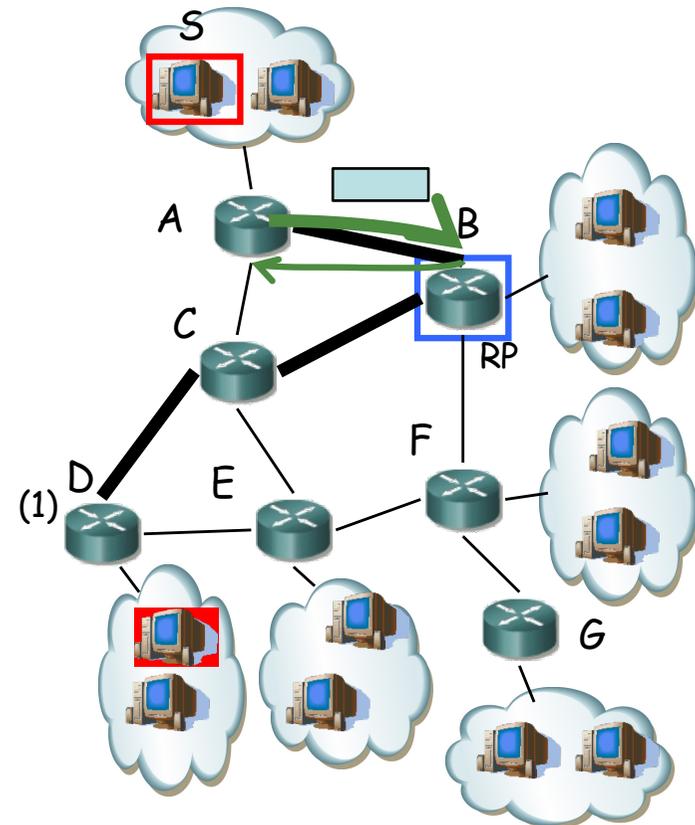
# PIM-SM: RPT

- La fuente envía los paquetes al RP encapsulados en paquetes PIM unicast a él
- El RP distribuye desde él el paquete nativo multicast por el RPT
- RP envía *Join* a (S,G) upstream
- Fuente empieza a enviar los paquetes en multicast por árbol (S,G) hasta el RP
- Al recibirlos, el RP envía un *Register-Stop* para dejar de recibir los encapsulados unicast
- Desde S al RP fluye el tráfico por un (S,G) y desde el RP fluye por un (\*,G)



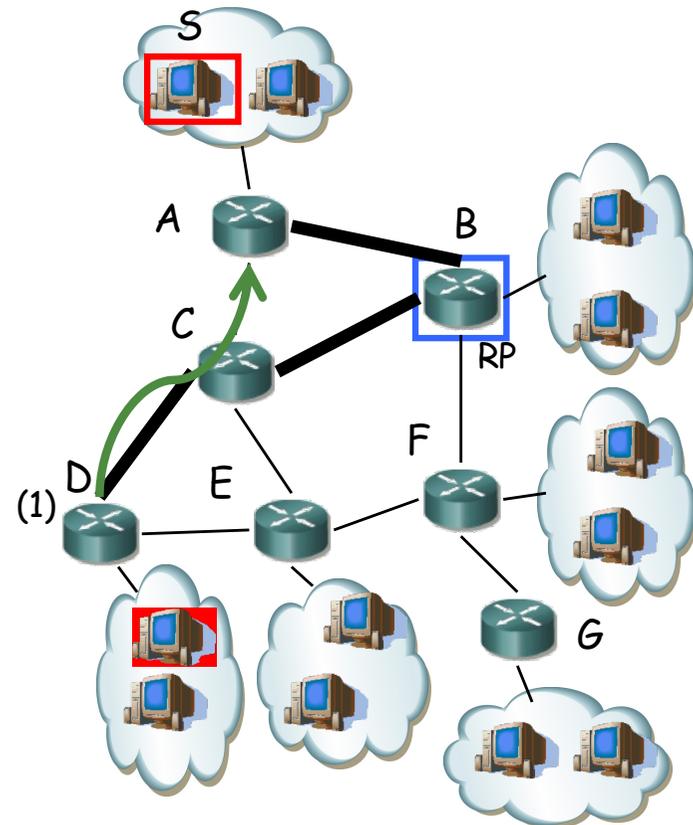
# PIM-SM: RP discovery

- El RP para un grupo puede estar configurado estáticamente en el resto de m routers
- O se puede descubrir automáticamente
  - Bootstrap Router (BSR) Protocolo, RFC 5059 (actualiza PIMv2)
  - Cisco AutoRP



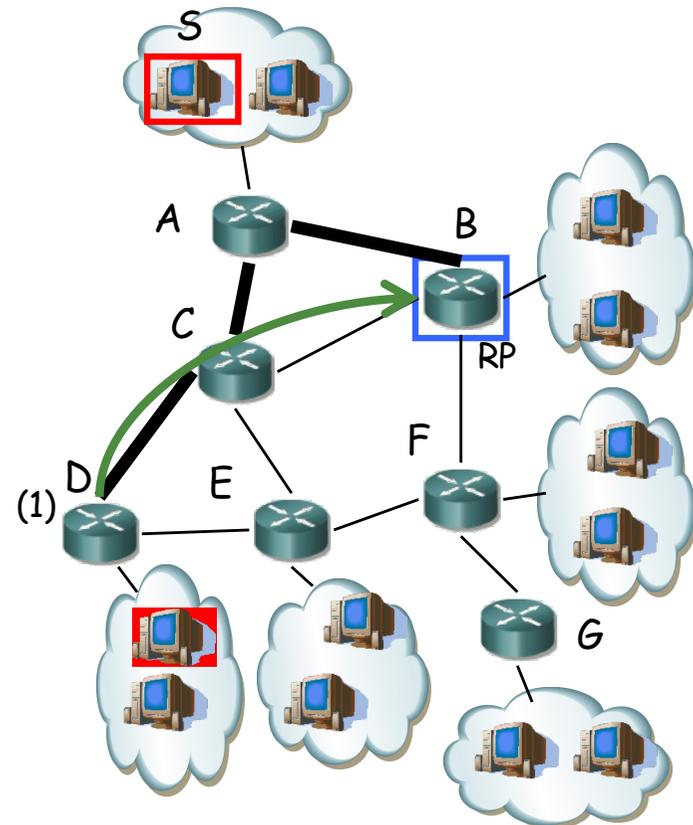
# PIM-SM: SPT

- *Shortest Path Tree* (SPT)
- El camino por el *shared tree* puede no ser rentable
- Mrouter junto a cliente puede enviar un *Join* a (S,G)
- Eso implica un árbol desde la fuente
- (...)



# PIM-SM: SPT

- *Shortest Path Tree* (SPT)
- El camino por el *shared tree* puede no ser rentable
- Mrouter junto a cliente puede enviar un *Join* a (S,G)
- Eso implica un árbol desde la fuente
- Cuando le llega el tráfico por los dos árboles envía un *Prune* del (\*,G)
- Si el RP está *pruned* por el resto de interfaces puede enviar *Prune* hacia S



upna

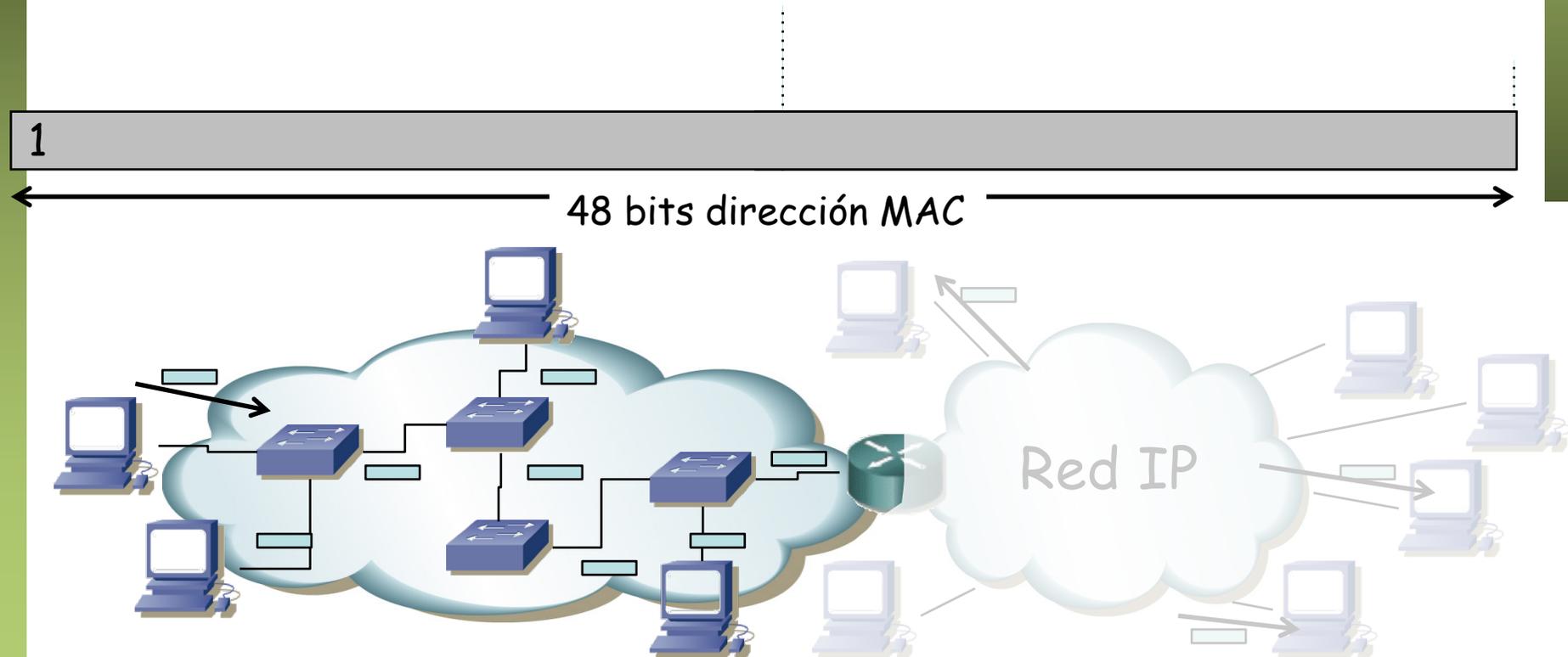
Universidad Pública de Navarra  
Nafarroako Unibertsitate Publikoa

**Tecnologías Avanzadas de Red**  
*Área de Ingeniería Telemática*

# Multicast en LAN Ethernet

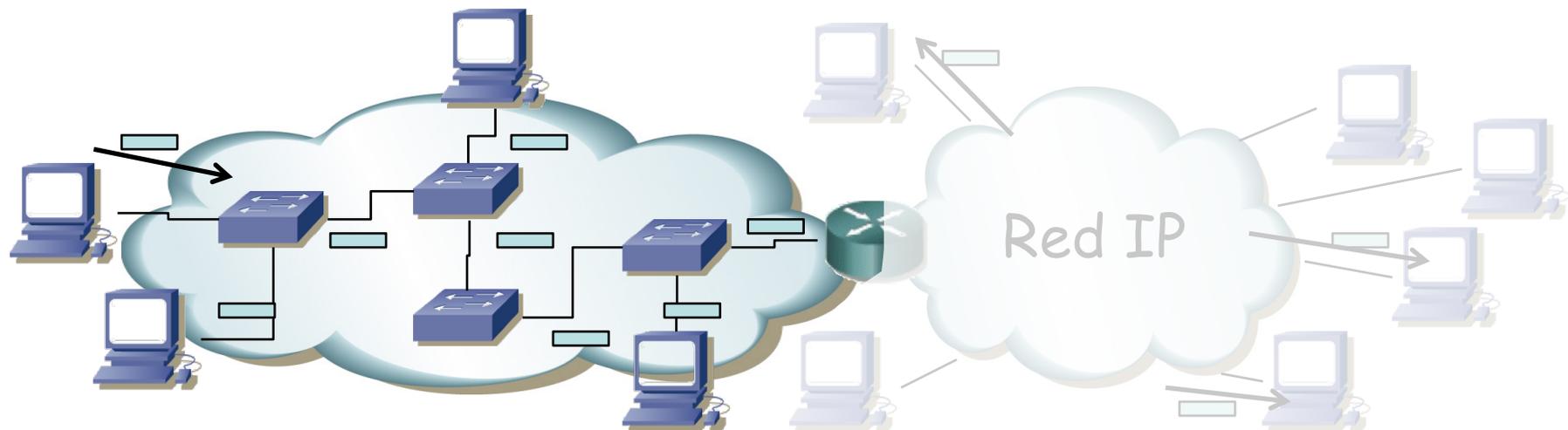
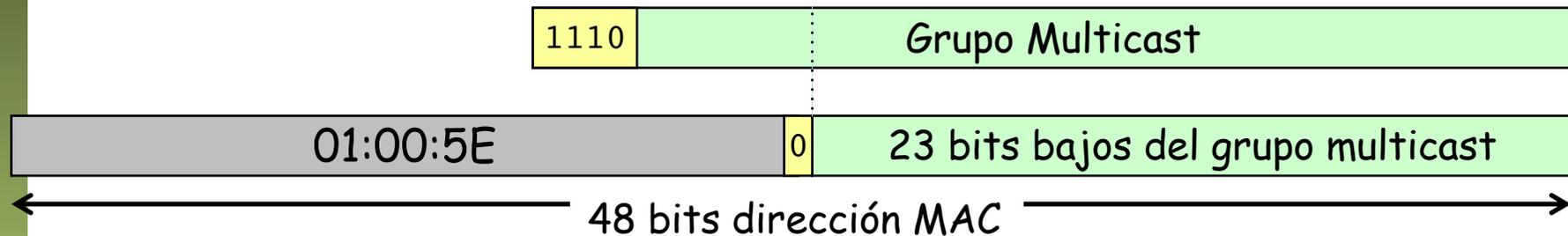
# Direccionamiento en LAN

- Se puede enviar un paquete a un conjunto de hosts de la LAN
- Direcciones MAC mcast en Ethernet tienen el primer bit (en transmisión) a 1
- Es el bit menos significativo del primer byte de la dirección MAC luego el primer byte es un número impar
- Los conmutadores tendrán que hacer inundación
- Llega a los interesados pero también al resto de hosts de la LAN



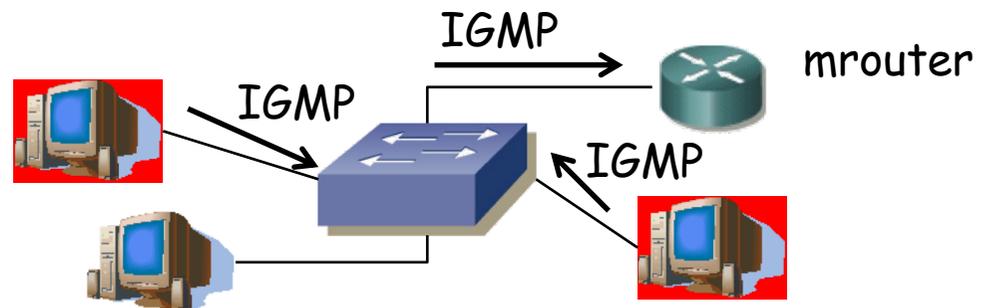
# Direccionamiento en LAN

- OUI 0x00005E reservado para IANA (John Postel)
- De ahí la mitad se reservan para este mapeo (siguiente bit a 0)
- Paquete IP multicast se mapea en trama a MAC multicast (32 a 1)
- Al indicar una aplicación que le interesa un grupo multicast la NIC entregará al nivel IP todas las tramas que correspondan a esa MAC



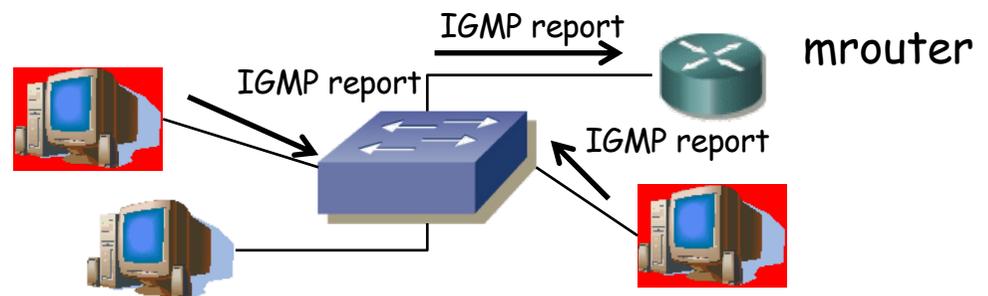
# IGMP snooping

- Switches por defecto hacen inundación del tráfico multicast
- Un switch que hace *IGMP snooping* procesa los mensajes IGMP para decidir por dónde no hace falta que reenvíe el tráfico multicast
- Procesa los *IGMP Report* de los hosts para saber por dónde reenviar
- En la CAM mantendrá información indicando por qué puertos reenviar los paquetes que van a una MAC multicast
- Siempre *flooding* para paquetes no IGMP a 224.0.0.0/24 porque para esos no se usa IGMP (locales al enlace)



# IGMP snooping

- Reenvía los *IGMP Report* solo por los puertos donde haya mrouter
  - Evita así que otros hosts del grupo supriman sus *Reports*
  - Si los suprimen no sabrá si siguen perteneciendo al grupo
  - El mrouter debe anunciarse como tal, por ejemplo con *IGMP Queries*
- RFC 4541 “*Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches*”
- RFC sobre implementaciones y recomendaciones



# Otros protocolos de encaminamiento multicast

# Otros protocolos

## MOSPF

- Multicast Extensions to OSFP (RFC 1584)
- LSAs de pertenencia a un grupo
- Crea *source trees*
- Dijkstra para cada (S,G) y cada vez que se une un nuevo receptor

## MBGP

- Multiprotocol Extensios for BGP-4 (RFC 4760)
- Para transportar información para más protocolos que IPv4 (IPv6, IPX...), unicast y multicast
- Distribuye la información que se puede emplear para el RPF de forma independiente para el tráfico multicast que para el unicast
- No calcula árboles, sigue haciendo falta otro protocolo