

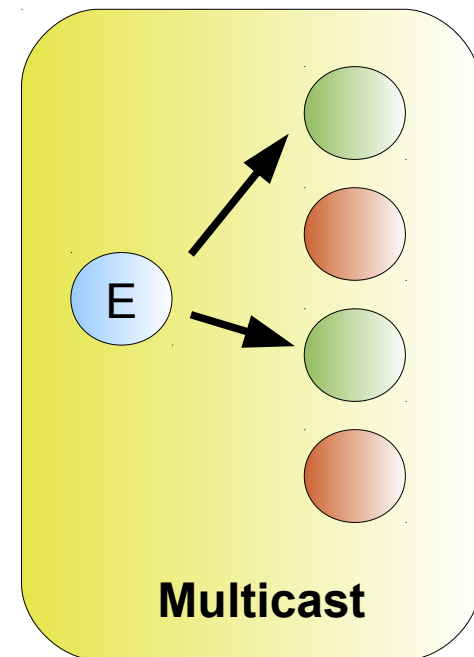
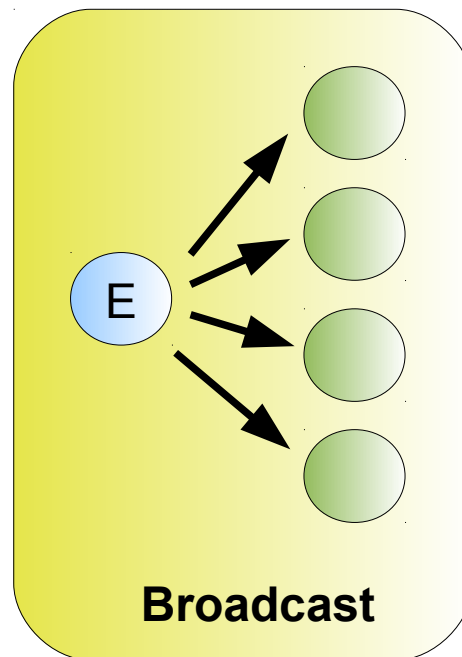
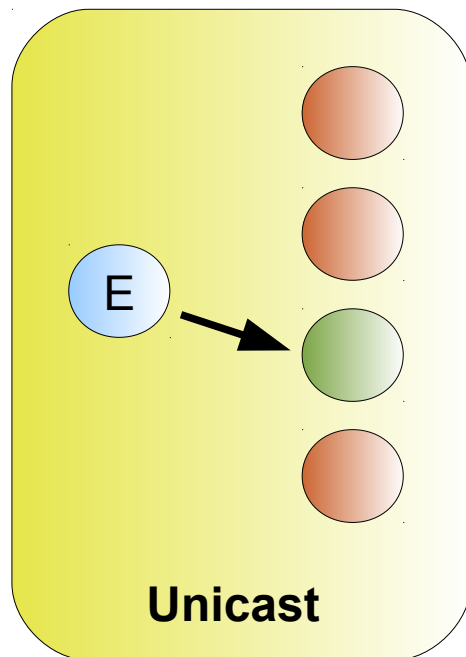
Encaminhamento Multicast

Redes e Serviços

**Licenciatura em Tecnologias e Sistemas de Informação
DETI-UA**

Multicast

- As comunicações multicast referem-se a comunicações um-para-muitos ou muitos-para-muitos.
 - Multicast ao nível da aplicação
 - Multicast ao nível da rede



Abstracção Multicast

- A informação transmitida por uma aplicação origem com uma única operação de envio é recebido por múltiplas aplicações destino em diferentes estações
- Alternativa 1: A stack protocolar TCP/IP da estação emissora estabelece ligações ponto-a-ponto com todos os destinos e envia múltiplas cópias, uma por cada destino
- Vantagens
 - ◆ Permite o uso de redes sem capacidade “multicast”
 - ◆ Permite a utilização do protocolo TCP com todas as suas vantagens
- Desvantagens
 - ◆ Exige que a aplicação emissora especifique a lista de endereços destino
 - ◆ Resulta numa utilização ineficiente dos recursos da rede



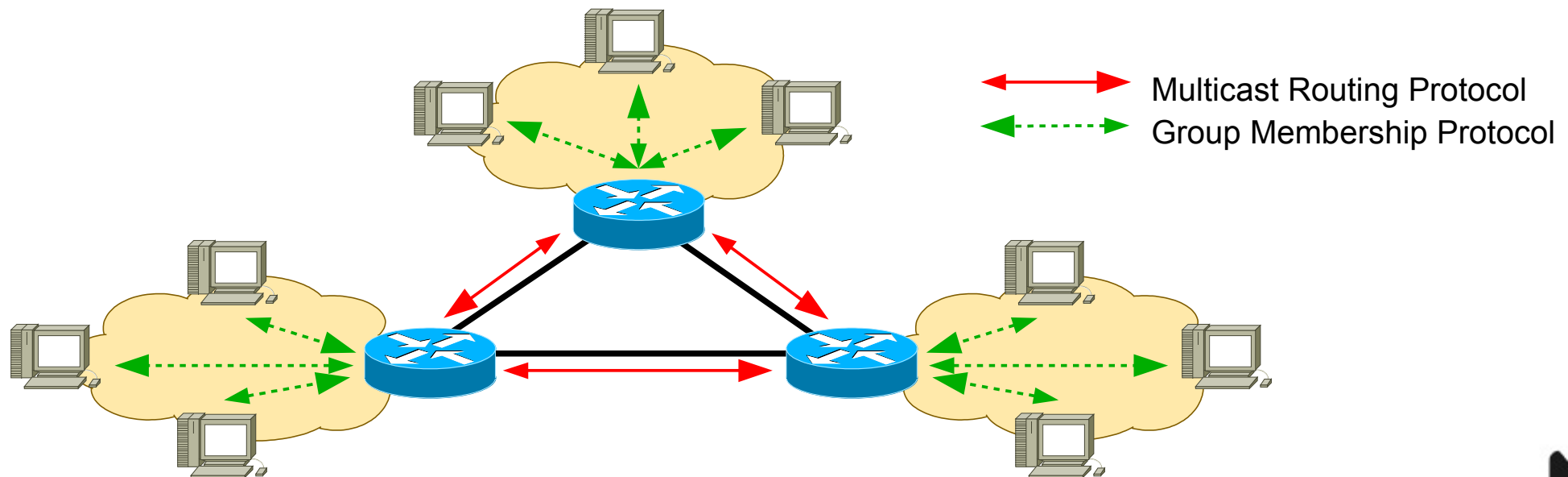
Abstracção Multicast

- Alternativa 2: A estação envia cada pacote IP uma única vez e a rede responsabiliza-se por o copiar para os múltiplos destinos
- Desvantagens
 - ◆ Exige que a rede tenha capacidade de “multicast”
 - ◆ Permite apenas a utilização do protocolo UDP com todas as suas desvantagens
- Vantagens
 - ◆ É possível ter uma utilização mais eficiente dos recursos da rede
- Pontos a discutir
 - ◆ Como é que as estações emissoras especificam as estações destino?!
 - ◆ Como é que os routers implementam capacidades “multicast”?!



Abstracção Multicast

- Redes IP multicast não são redes “connectionless” como no caso das redes “unicast”
- É necessário estabelecer percursos multicast entre os routers para saberem como encaminhar pacotes multicast
- Para isso, é necessária sinalização (entre estações e routers) e protocolos de encaminhamento (entre routers) capazes de estabelecer os percursos multicast necessários



Identificação das estações destino

- Identificação explícita é muitas vezes indesejável
- Faz-se uso de endereços IP de classe D (começam por 1110)
- As estações participantes combinam o uso de um endereço que identifica a sessão
- As estações receptoras anunciam aos routers a participação na sessão *multicast* identificada pelo endereço combinado
 - Protocolo IGMP
- Os routers encaminham os pacotes IP enviados com o endereço destino escolhido para todas as redes em que haja estações que participem na respectiva sessão
 - Protocolo de encaminhamento DVMRP, MOSPF, PIM, etc...



Endereços de Classe D

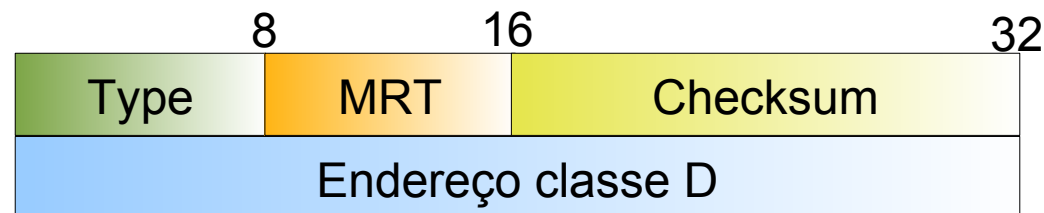


- Alguns endereços estão reservados pela Internet Assigned Numbers Authority (IANA).
- Alguns endereços têm significado próprio. Por exemplo, endereços da forma 224.0.0.X
 - 1 : All Hosts
 - 2 : All Multicast Routers
 - 4 : All DVMRP Routers
 - 5 : All OSPF routers
 - 6 : OSPF designated routers
 - 13 : All PIM routers
- A gama de 239.0.0.0 a 239.255.255.255 destina-se às redes privadas.



Internet Group Membership Protocol (IGMP)

- IGMP versão 2, RFC 2236
 - Opera entre cada estação e os routers que lhe estão directamente ligados
 - Serve para a estação anunciar ao router que quer participar numa sessão *multicast* (identificada por um endereço de classe D)



MRT - Maximum Response Time

- O IGMP corre sobre o protocolo IP (*protocol type* = 0x02)
- Os pacotes são enviados para o endereço destino 224.0.0.1 (“All Hosts”) com TTL= 1



Mensagens IGMP

- GMQ - General Membership Query
 - ◆ Enviado pelos routers para perguntar se as estações participam em alguma sessão multicast
- SMQ - Specific Membership Query
 - ◆ Enviado pelos routers para perguntar se existe alguma estação que participe numa sessão multicast específica
- MR - Membership Report
 - ◆ Enviado pelas estações para sinalizarem que participam numa sessão multicast
- LGR - Leave Group Report (Opcional)
 - ◆ Enviado pelas estações para sinalizarem que deixam de participar numa sessão multicast
- Em cada rede, o Querier Router é o router que tiver menor endereço IP na interface ligada à rede e é aquele que mantém o “diálogo” IGMP com os terminais.

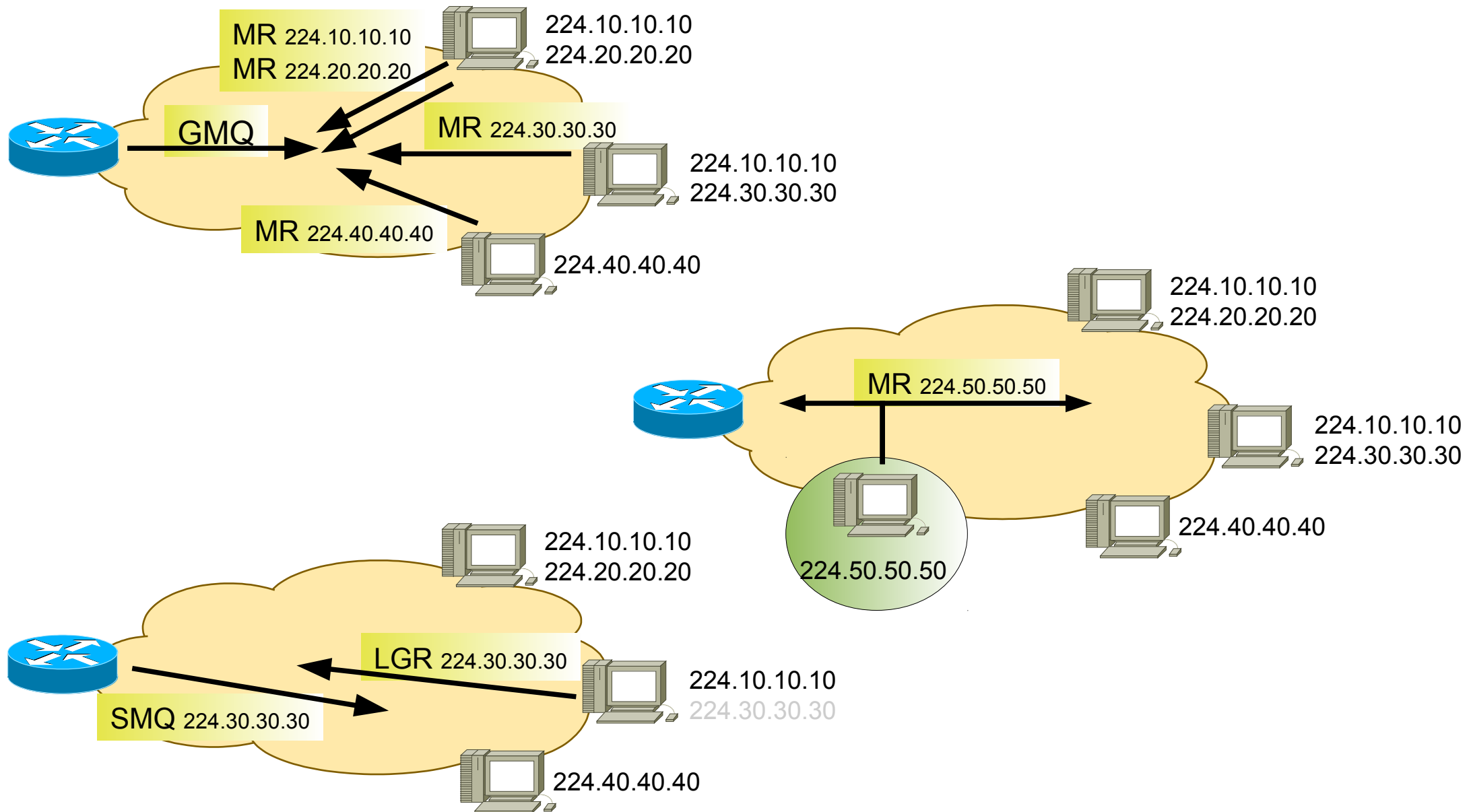


Protocolo IGMP

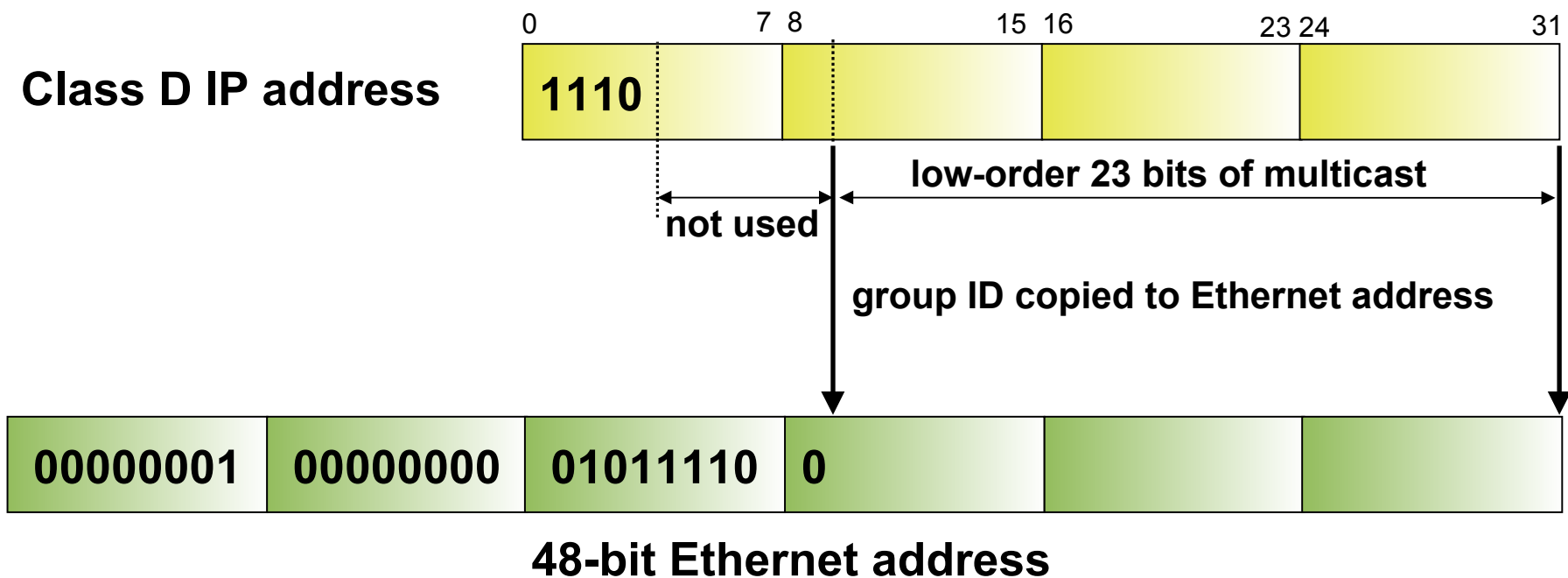
- Os routers enviam periodicamente um GMQ especificando um Maximum Response Time (MRT).
- Cada estação espera um tempo aleatório entre 0 e MRT para responder com um MR especificando um endereço multicast.
- Se entretanto, uma estação “vê” um MR para a mesma sessão, ele aborta o seu envio do MR.
- Uma estação emite um MR quando quer pertencer a uma sessão multicast.
- Opcionalmente, uma estação emite um LGR quando deixa de querer pertencer a uma sessão multicast.
- Quando um router recebe um LGR, emite um SMQ para verificar se ainda há estações pertencentes a essa sessão.



Protocolo IGMP



Conversão de endereço IPv4 de classe D em endereço IEEE 802



IGMP - conclusões finais

- Qualquer estação pode juntar-se a uma sessão multicast recebendo e enviando informação.
- A formação de sessões multicast é iniciada pelos receptores
 - ♦ Os emissores não especificam nem controlam que estações podem receber a informação
- A rede não providencia filtragem, ordenação ou privacidade dos pacotes multicast.
- O modelo de serviço multicast IP segue a mesma filosofia do modelo de serviço unicast:
 - ♦ uma camada protocolar simples e fiável em que funcionalidades adicionais são providenciadas pelas camadas protocolares superiores



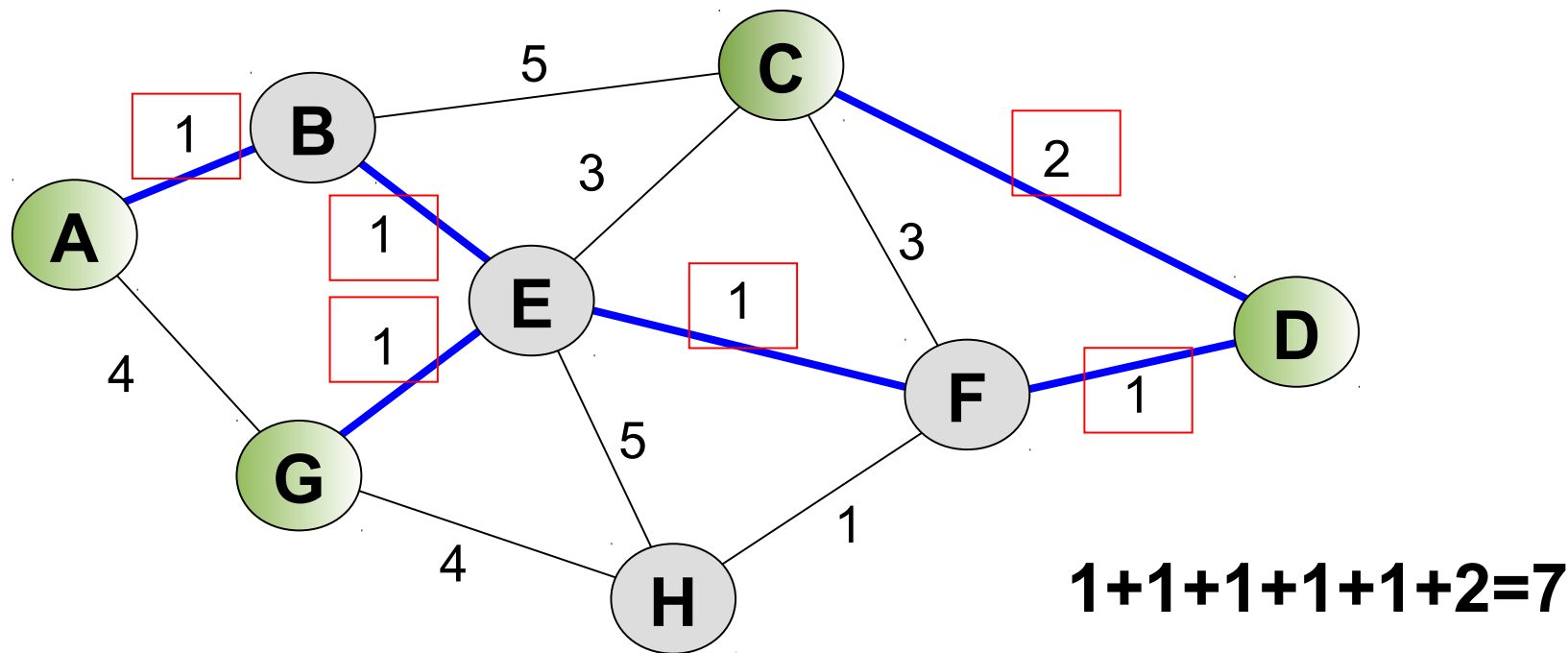
Encaminhamento Multicast

- Group-shared tree
 - Baseia-se na determinação de uma única árvore de encaminhamento por cada sessão multicast que interligue todos os routers com estações pertencentes à sessão
- Source-based tree
 - Baseia-se na determinação de uma árvore de encaminhamento por cada sessão multicast e por cada emissor



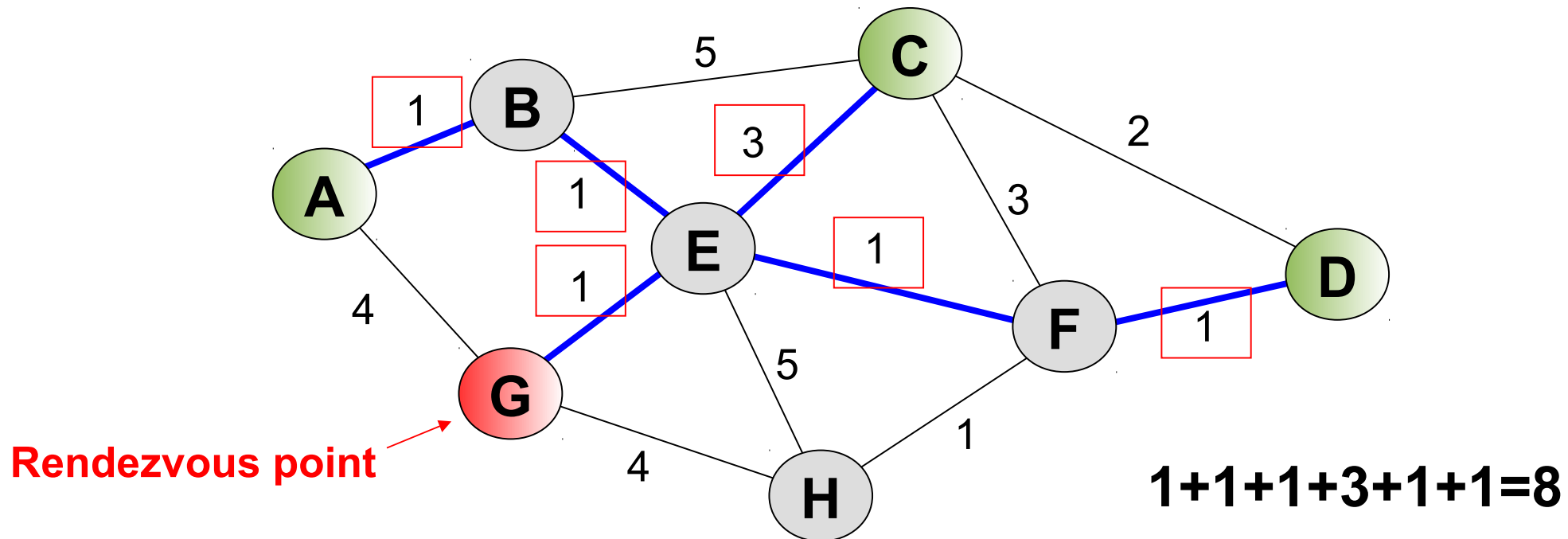
Group-shared tree

- Árvore de Steiner: determinação da árvore de custo mínimo que interliga os nós com estações de uma sessão
 - ♦ Exige um protocolo do tipo link-state
 - ♦ Algoritmo de complexidade exponencial



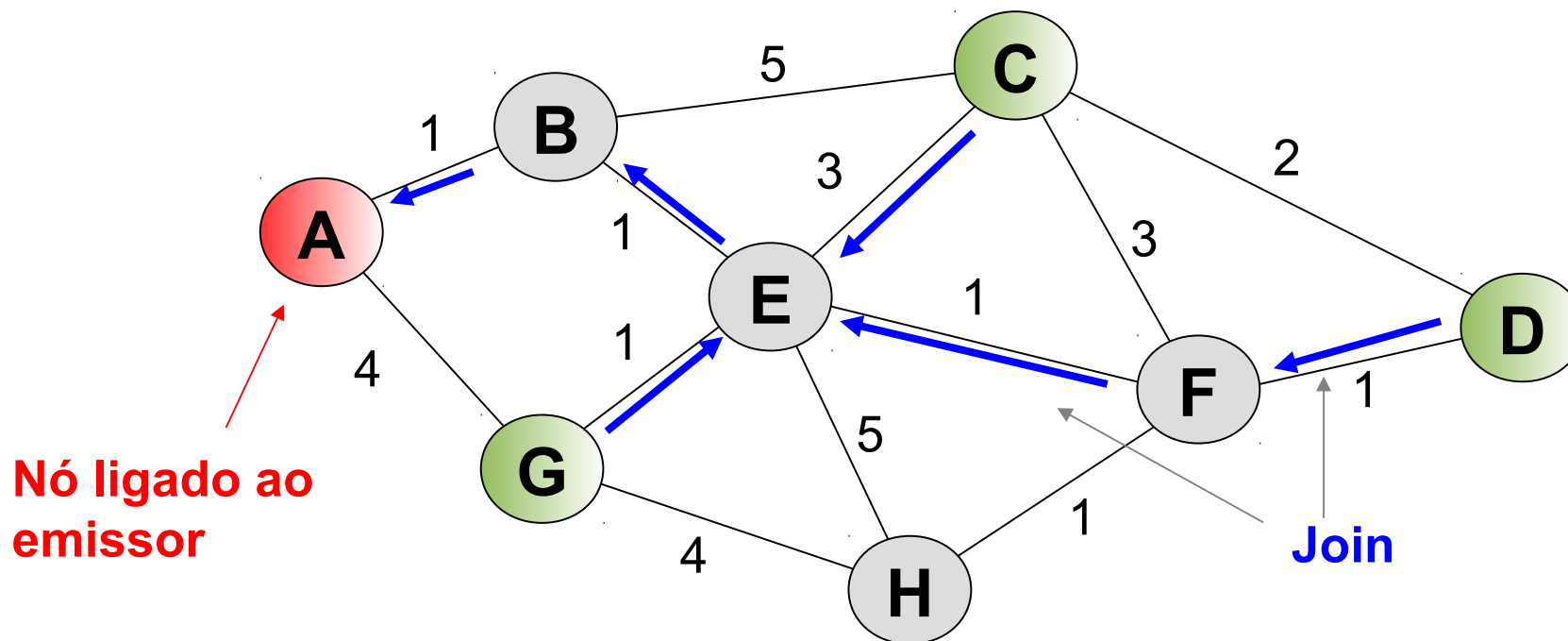
Group-shared tree

- Árvore dos percursos de custo mínimo para um nó central (“rendezvous point”)
 - O nó central é previamente escolhido (pertence necessariamente à árvore mesmo que não tenha estações terminais da sessão) e conhecido dos outros nós
 - Para se juntarem à árvore, os nós com estações terminais da sessão enviam uma mensagem “join” pelo percurso de custo mínimo entre eles e o nó central



Source-based tree (quando o emissor é conhecido)

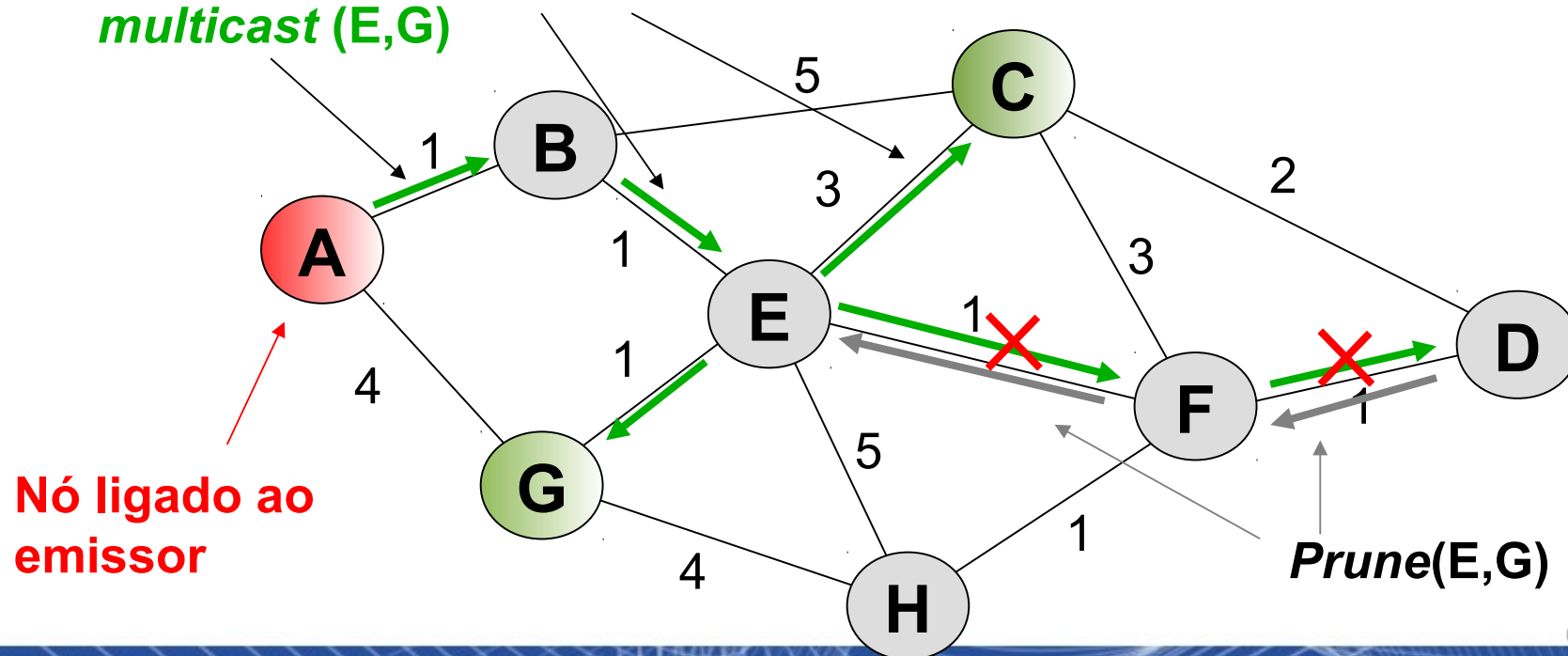
- Cada nó com receptores interessados numa sessão multicast G de uma determinado emissor com endereço E envia uma mensagem de $\text{join}(E, G)$ em direcção ao endereço E pelo percurso de custo mínimo (percurso unicast)
- No percurso da mensagem join, cada nó recebe esta mensagem por uma interface I_d , reenvia-a pela interface I_o e constrói a tabela de encaminhamento (E, G) para os pacotes multicast que entram por I_o para serem encaminhados por I_d
- O encaminhamento multicast é baseado não só no endereço destino mas também no endereço origem de cada pacote



Source-based tree (quando o emissor é conhecido)

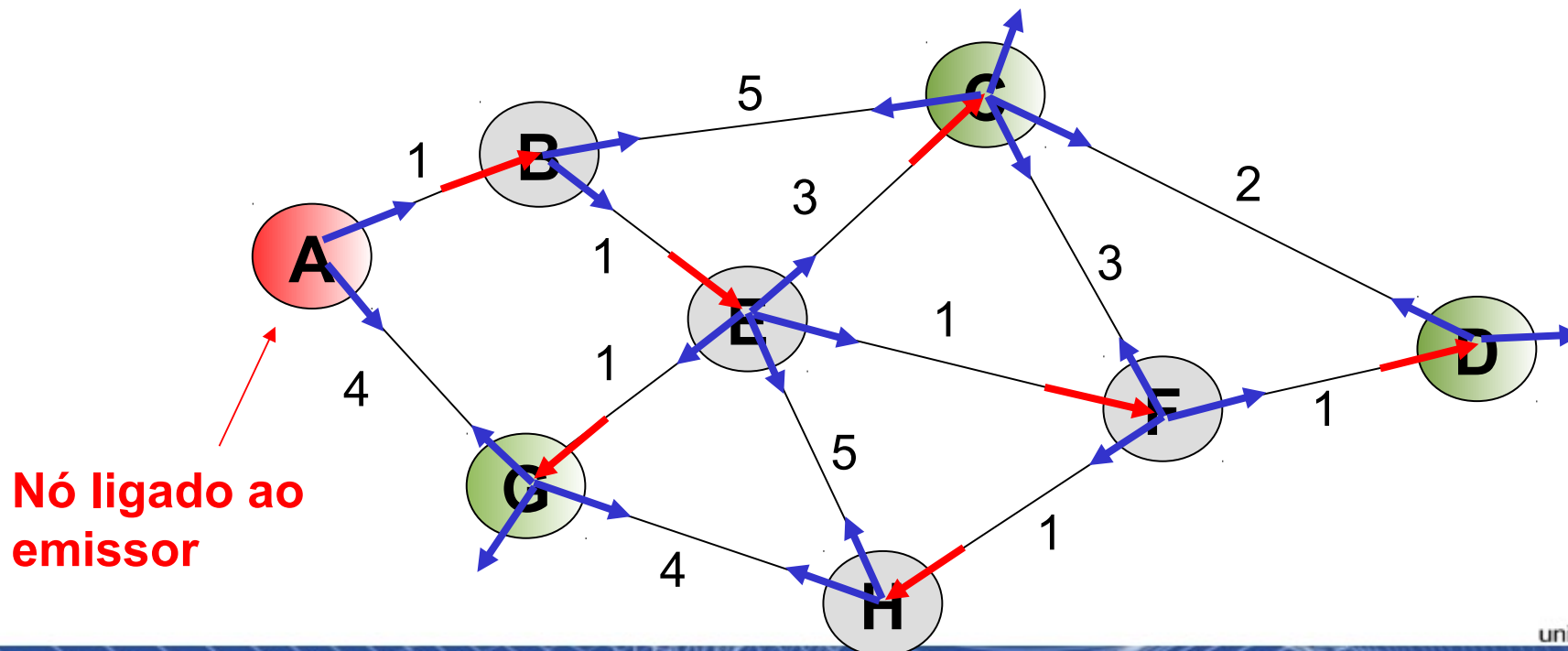
- Quando há múltiplos emissores, são estabelecidas árvores de encaminhamento multicast, uma por cada emissor
- Quando um receptor deixa de estar interessado numa sessão multicast, o nó que lhe está ligado envia uma mensagem de $\text{prune}(E,G)$ em direcção ao endereço do emissor E
- A mensagem de prune é reenviado apenas pelos nós que deixam de pertencer à árvore de encaminhamento multicast

Árvore de encaminhamento
multicast (E,G)



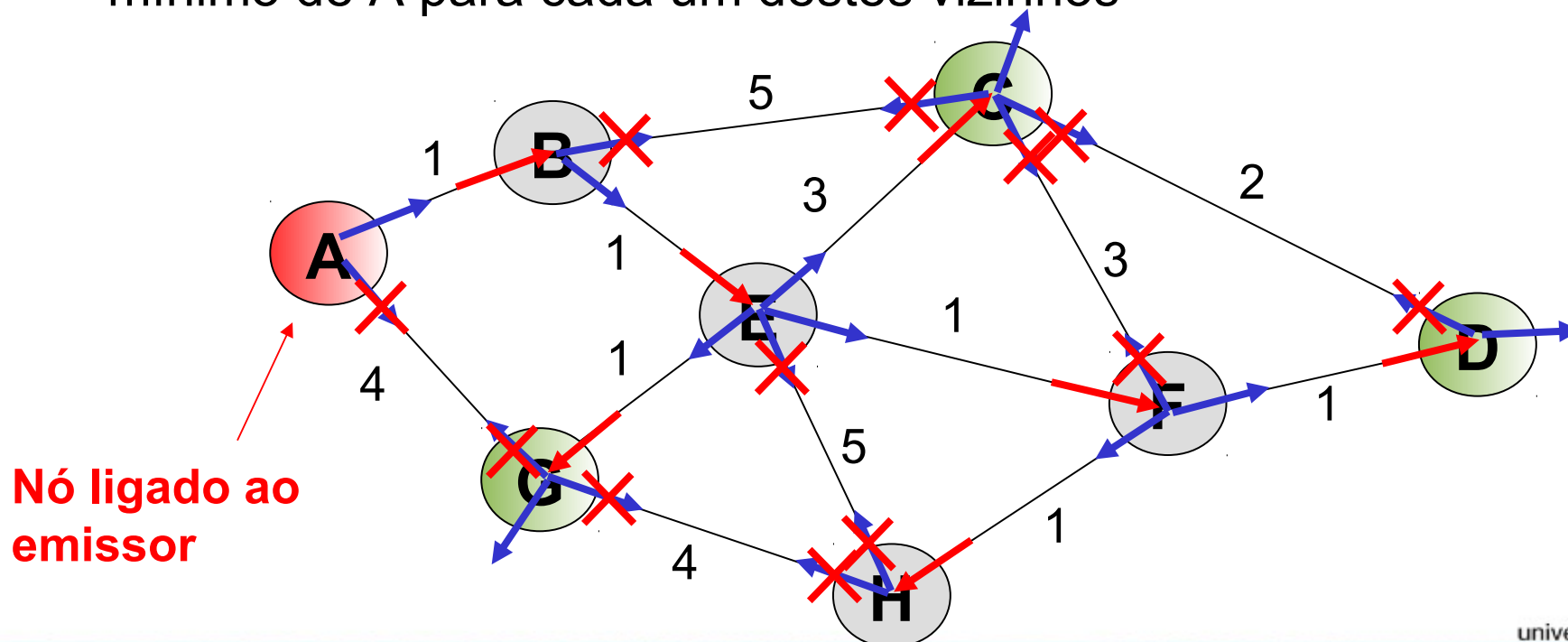
Source-based tree (quando o emissor não é conhecido)

- “Reverse Path Forwarding”: criação de uma árvore abrangente virtual
 - Cada nó N encaminha os pacotes do nó origem O recebidos do nó vizinho V para todos os outros nós vizinhos se V for o último nó no percurso de custo mínimo de O para N
 - Figura: no nó E, os pacotes com origem no nó A são encaminhados para todas as outras portas apenas se vierem do nó vizinho B porque este vizinho (o nó B) é o nó anterior a E no percurso de custo mínimo de A para E



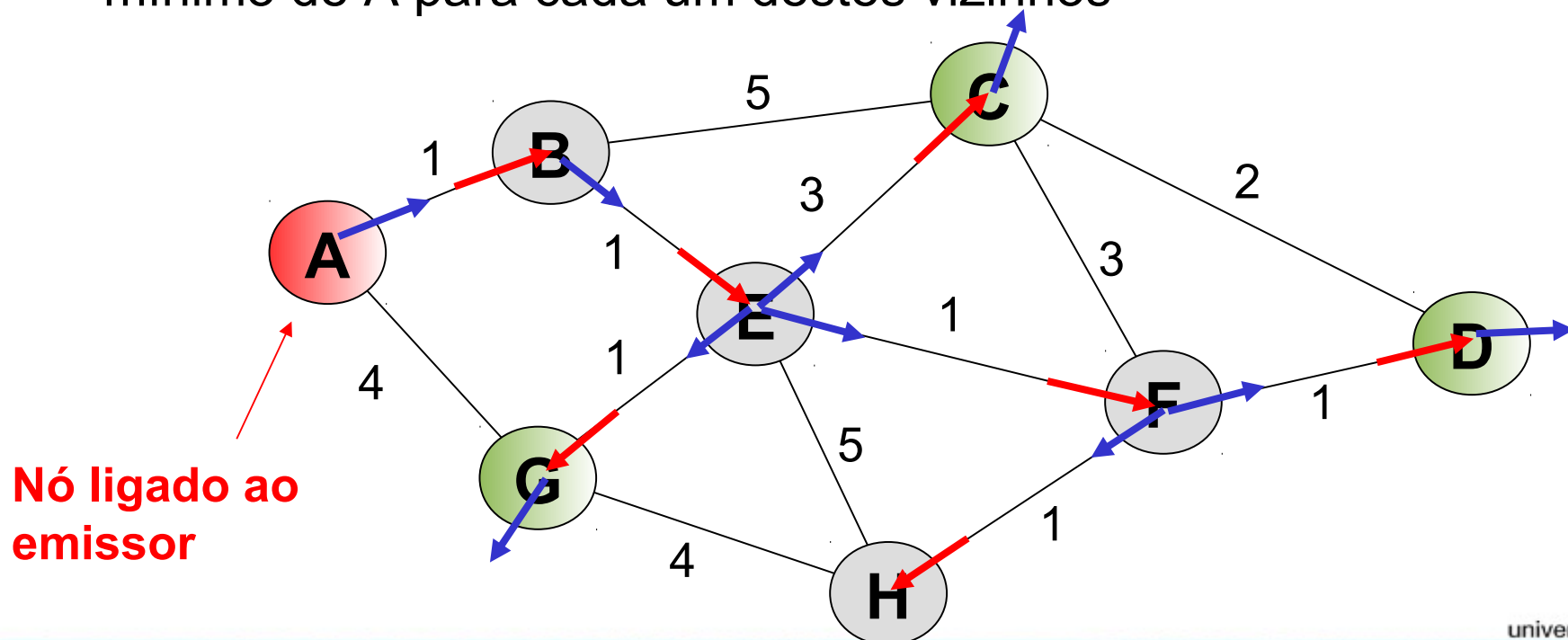
Source-based tree (quando o emissor não é conhecido)

- “Reverse Path Forwarding com Pruning” (I)
 - Para cada nó origem O, o nó N sabe para que nós vizinhos V ele é o último nó no percurso de custo mínimo de O para V
 - O encaminhamento é feito só para estes vizinhos
 - Figura: o nó E só encaminha os pacotes do nó origem A para os vizinho C, F e G porque o nó E é o último nó nos percursos de custo mínimo de A para cada um destes vizinhos



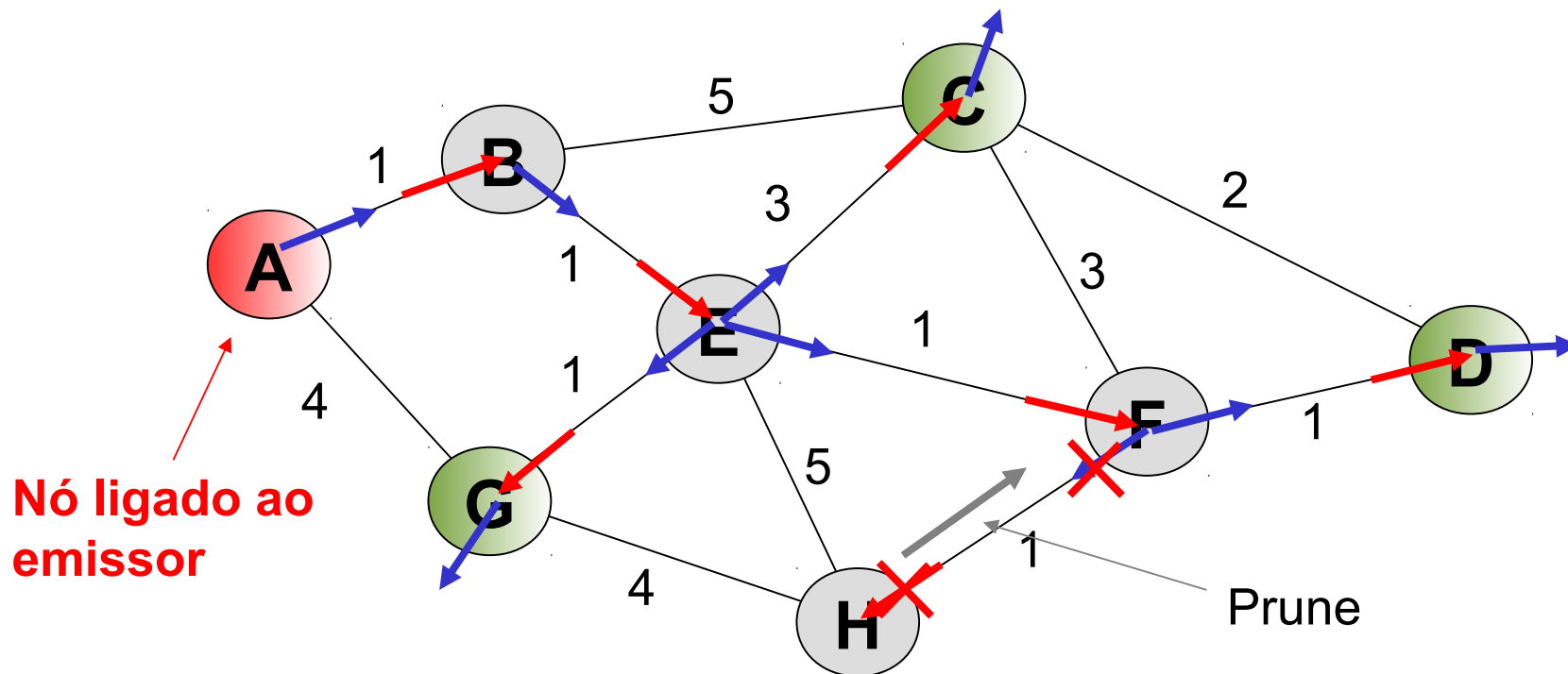
Source-based tree (quando o emissor não é conhecido)

- “Reverse Path Forwarding com Pruning” (I)
 - Para cada nó origem O, o nó N sabe para que nós vizinhos V ele é o último nó no percurso de custo mínimo de O para V
 - O encaminhamento é feito só para estes vizinhos
 - Figura: o nó E só encaminha os pacotes do nó origem A para os vizinho C, F e G porque o nó E é o último nó nos percursos de custo mínimo de A para cada um destes vizinhos



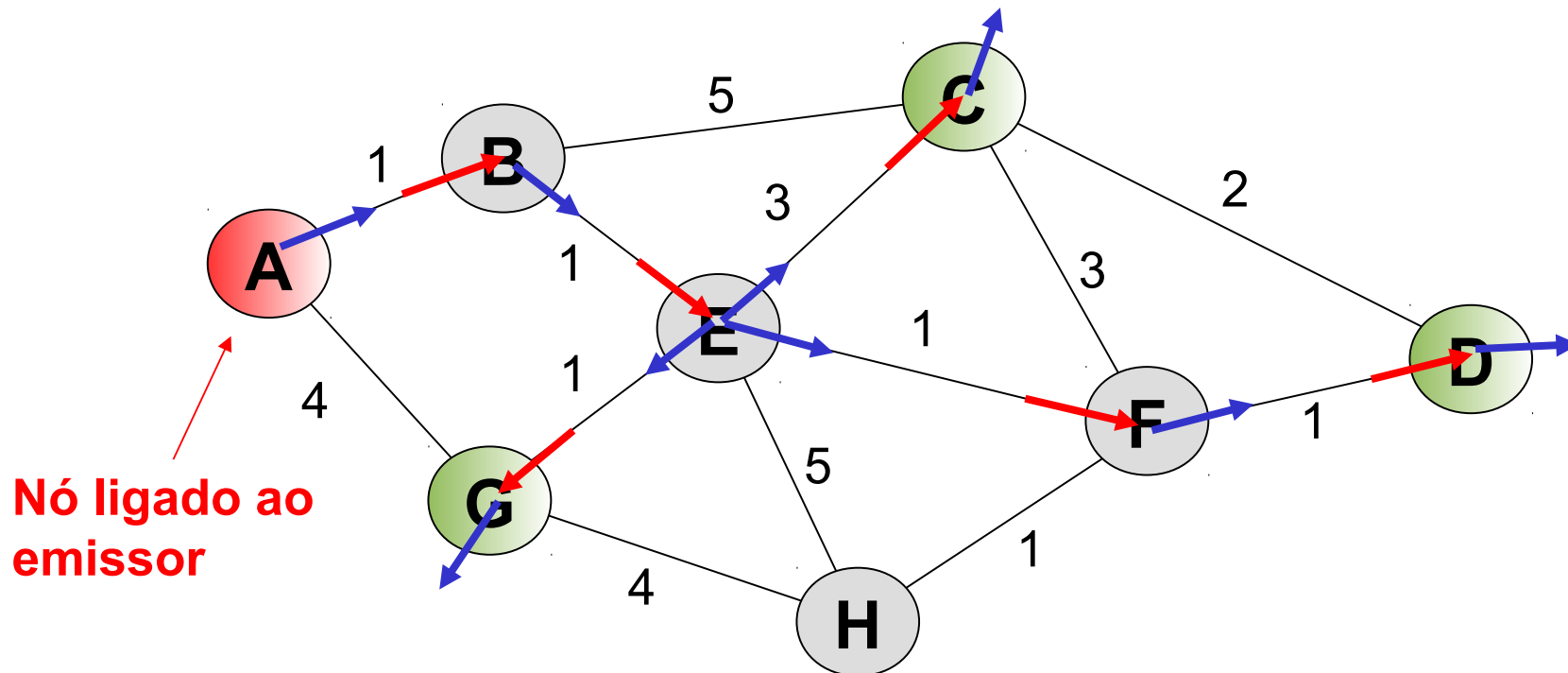
Source-based tree
(quando o emissor não é conhecido)

- “Reverse Path Forwarding com Pruning” (II)
 - Um nó sem estações terminais interessados na sessão e sem nós vizinhos para reencaminhar pacotes multicast, envia uma mensagem prune para o vizinho por quem recebe os pacotes multicast
 - Figura: o nó H envia uma mensagem prune para o vizinho F e deixa, assim, de receber pacotes multicast



Source-based tree (quando o emissor não é conhecido)

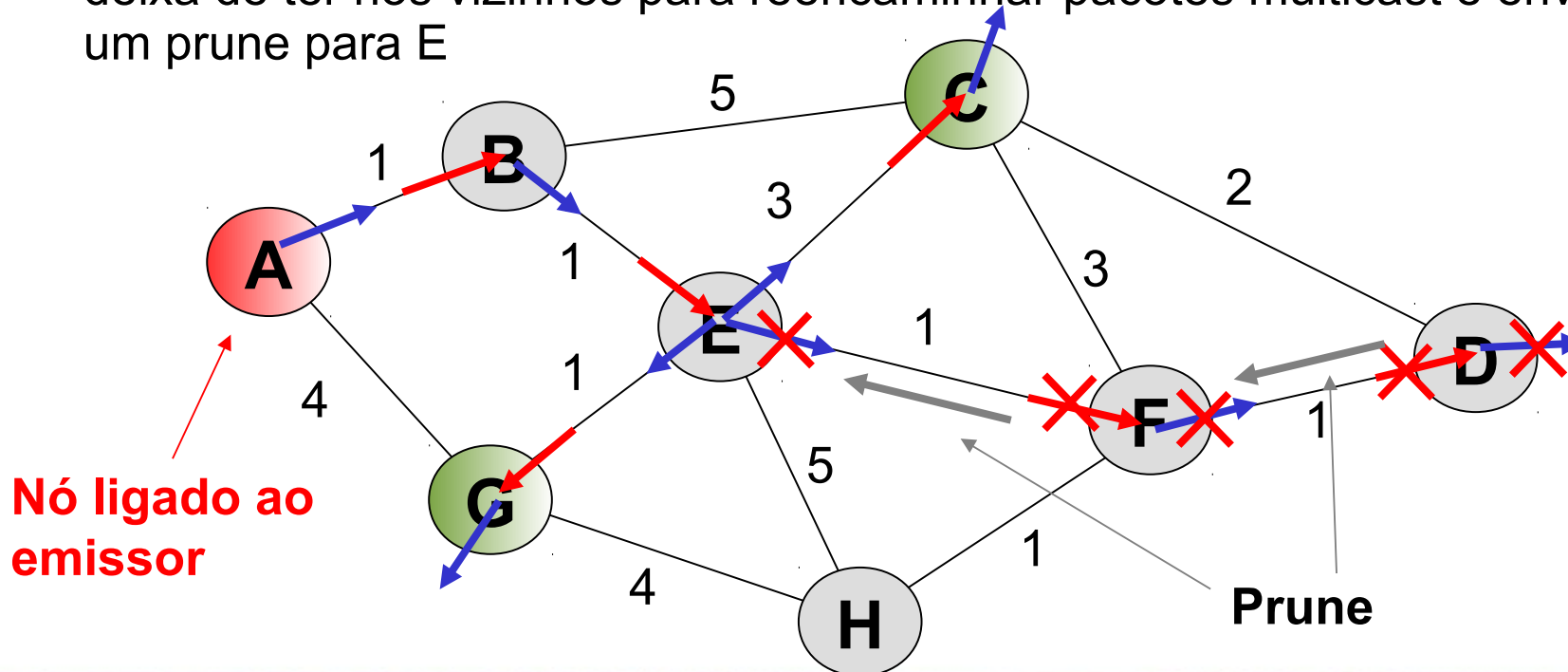
- “Reverse Path Forwarding com Pruning” (II)
 - Um nó sem estações terminais interessados na sessão e sem nós vizinhos para reencaminhar pacotes multicast, envia uma mensagem prune para o vizinho por quem recebe os pacotes multicast
 - Figura: o nó H envia uma mensagem prune para o vizinho F e deixa, assim, de receber pacotes multicast



Source-based tree (quando o emissor não é conhecido)

- “Reverse Path Forwarding com Pruning” (II)

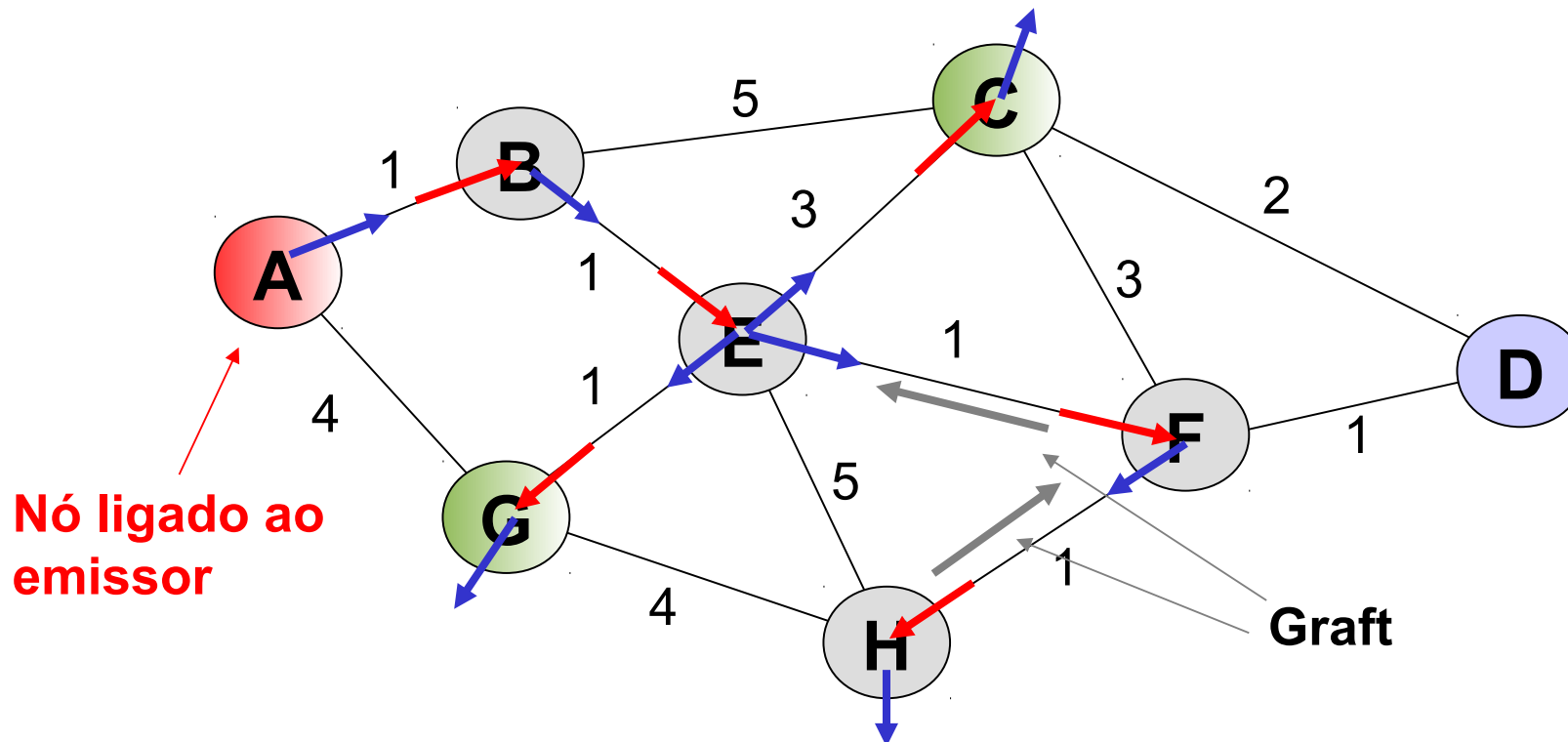
- Um nó sem estações terminais interessados na sessão e sem nós vizinhos para reencaminhar pacotes multicast, envia uma mensagem prune para o vizinho por quem recebe os pacotes multicast
- Figura: o nó H envia uma mensagem prune para o vizinho F e deixa, assim, de receber pacotes multicast
- Figura: o nó D deixa de ter receptores e envia um prune para F que, por sua vez, deixa de ter nós vizinhos para reencaminhar pacotes multicast e envia também um prune para E



Source-based tree

(quando o emissor não é conhecido)

- “Reverse Path Forwarding com Pruning” (III): aparecimento de nós membro de uma sessão multicast
 - Primeira estratégia: usar uma mensagem graft para anular um prune
 - Segunda estratégia: associar um tempo de vida ao prune ao fim do qual os pacotes voltam a ser encaminhados



Group-shared tree vs. Source-based tree

- Group-shared tree
 - Minimiza a informação de estado que tem de ser mantida em cada router
 - Minimiza o número de ligações que usa para suportar tráfego *multicast*
 - Concentra os problemas de congestionamento de tráfego num número reduzido de ligações
- Source-based tree
 - Penaliza a informação de estado em cada router pois é baseada em cada uma das possíveis origens
 - Distribui o tráfego *multicast* por um número maior de ligações
 - Provoca menos problemas de congestionamento de tráfego



Distance Vector Multicast Routing Protocol (DVMRP)

- Algoritmo do tipo “*source-based tree*”
 - É do tipo “Distance Vector” (RFC 1075) parecido com o RIP
- Usa uma estratégia de RPF com “*pruning*”
 - Tal como o RIP, as distâncias são dadas em número de saltos (“hops”)
 - Os vectores distância representam a distância de cada possível origem
 - Por cada possível origem, cada router anuncia também aos seus vizinhos quando eles são o último salto no percurso desde a origem
- Mensagens de “*prune*” (poda)
 - enviadas com um tempo de vida
- As mensagens de “*graft*” (enxerto)
 - servem para eliminar uma mensagem de “*prune*”



Distance Vector Multicast Routing Protocol (DVMRP)

- Algoritmo do tipo “*source-based tree*”
 - É do tipo “Distance Vector” (RFC 1075) parecido com o RIP
- Usa uma estratégia de RPF com “*pruning*”
 - Tal como o RIP, as distâncias são dadas em número de saltos (“hops”)
 - Os vectores distância representam a distância de cada possível origem
 - Por cada possível origem, cada router anuncia também aos seus vizinhos quando eles são o último salto no percurso desde a origem
- Mensagens de “*prune*” (poda)
 - enviadas com um tempo de vida
- As mensagens de “*graft*” (enxerto)
 - servem para eliminar uma mensagem de “*prune*”



Multicast Open Shortest Path First (MOSPF)

- Algoritmo do tipo “source-based tree”
 - É uma extensão do protocolo OSPF (RFC 1584)
 - Implementa uma estratégia “RPF com *pruning*”
- Faz uso do facto de cada router conhecer a topologia completa da rede
 - Desta forma, cada router pode processar localmente a árvore dos percursos de custo mínimo de cada sessão *multicast*
- Não suporta túneis
 - Nas mensagens OSPF, existe uma flag que quando está a Null significa que o router não suporta multicast
 - Estes routers não podem pertencer à árvore dos percursos de custo mínimo
 - Este protocolo exige que quaisquer dois routers *multicast* tenham de ter entre si pelo menos um percurso em que todos os routers intermédios suportem *multicast*



Protocol-Independent Multicast (PIM)

- O protocolo PIM (RFC 2362) endereça dois cenários extremos de utilização
- PIM, dense mode
 - A maioria das redes tem estações que pretendem usar encaminhamento *multicast*
 - Consequentemente, a maioria dos routers da rede necessita de encaminhar pacotes *multicast*
- PIM, sparse mode
 - As estações que pretendem usar encaminhamento *multicast* concentram-se num número reduzido de redes
 - Consequentemente, o número de routers que necessitam de encaminhar pacotes *multicast* é pequeno comparativamente ao número total de routers

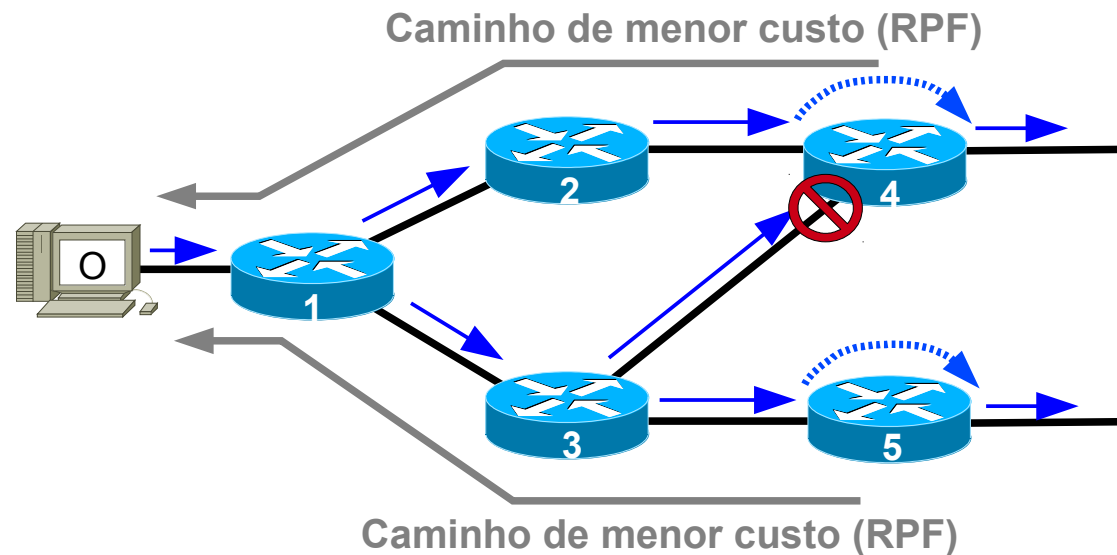


PIM Dense Mode

- Implementa uma estratégia “RPF com *pruning*”
- Exige que todos os routers tenham o protocolo activo
- É mais simples porque:
 - Não calcula tabelas de encaminhamento
 - Em vez disso, usa as tabelas de encaminhamento construídas por um qualquer protocolo unicast
 - ➔ Desta forma, é independente do protocolo de encaminhamento em uso
 - Assume que as rotas ponto-a-ponto são simétricas
- Em situações de múltiplos percursos de custo mínimo, aceita apenas os da interface com maior endereço IP
- As tabelas de encaminhamento *unicast* não permitem determinar para que vizinhos os pacotes *multicast* devem ser reencaminhados. Assim:
 - os routers reencaminham por omissão para todos os vizinhos que não tenham enviado mensagens de *prune*
 - usam as mensagens de *prune* para sinalizar os vizinhos que não devem reencaminhar os pacotes para eles

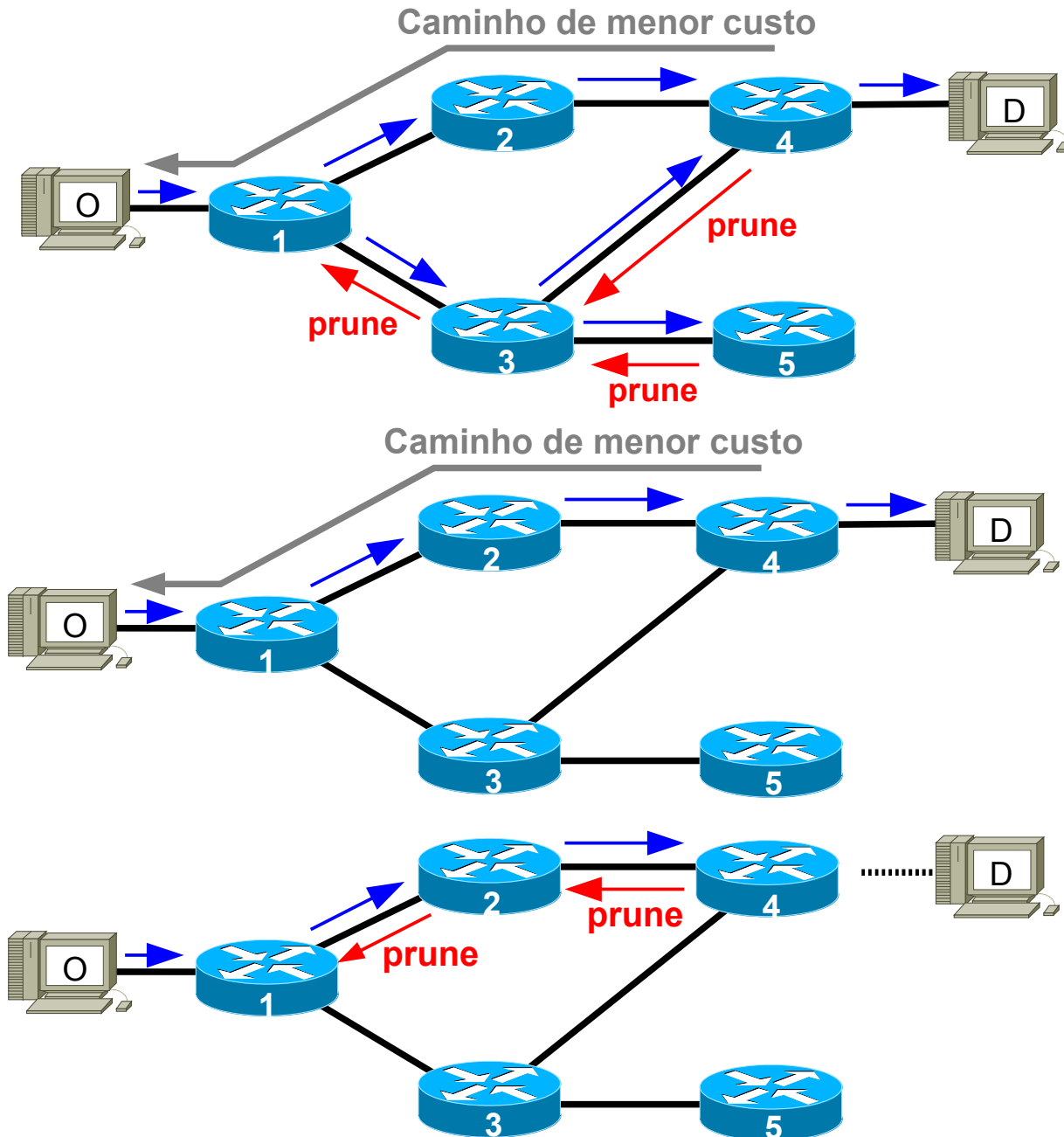


PIM Dense Mode – Flooding Inicial



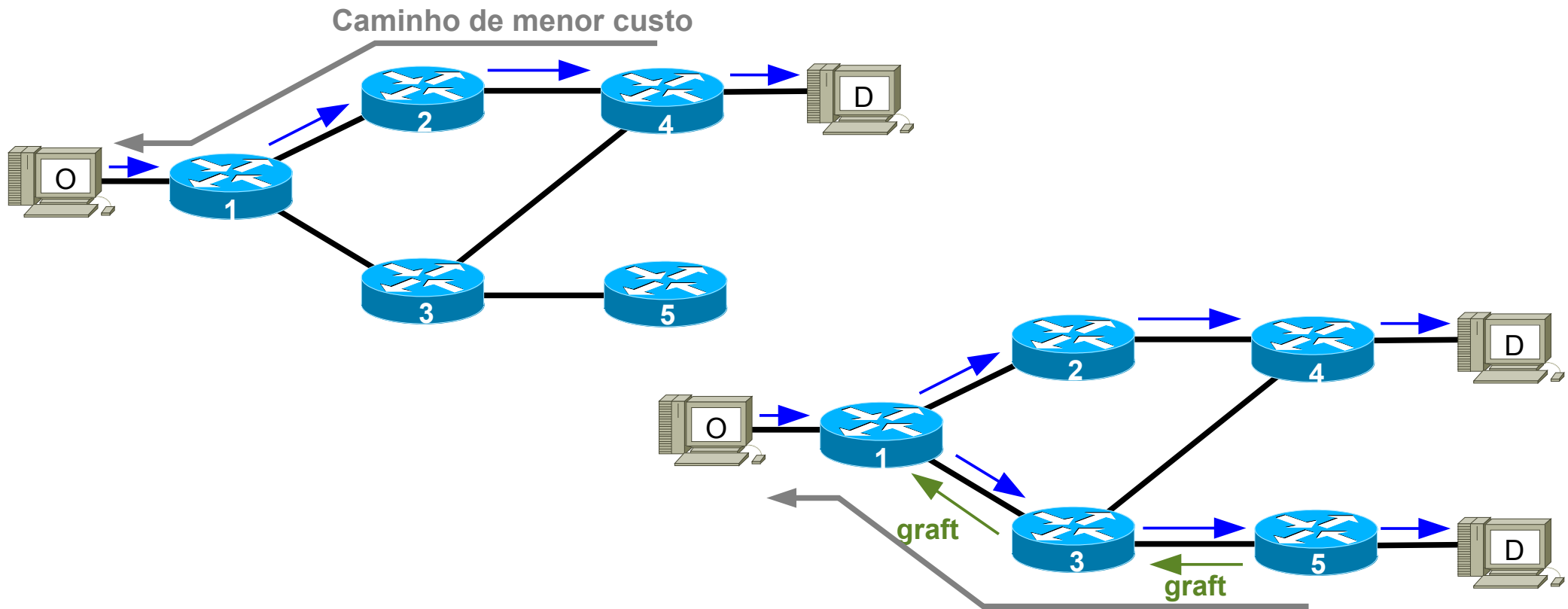
- Quando um router recebe o tráfego multicast pelo interface que fornece o percurso de custo mínimo para a fonte (RPF interface), reenvia o tráfego para todos os seus vizinhos (PIM-DM)
- Quando um router recebe tráfego multicast num interface que não o que fornece o percurso de custo mínimo para a fonte (not the RPF interface), descarta os pacotes multicast

PIM DM - Mensagem *Prune*



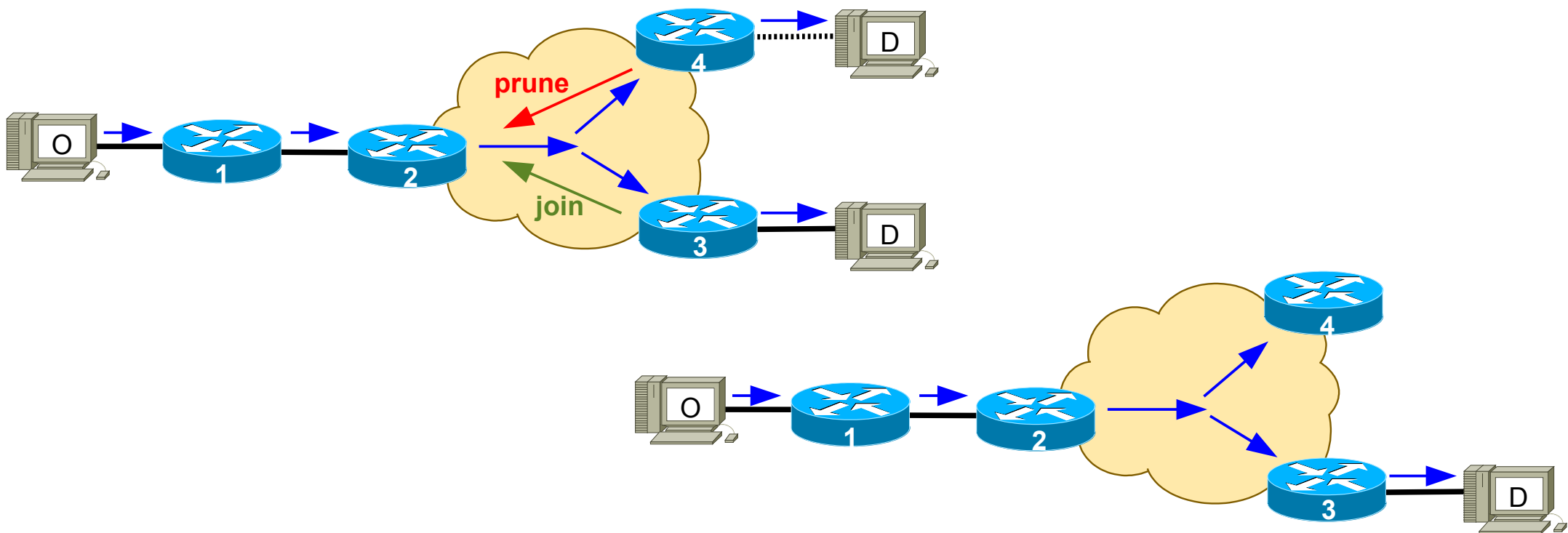
- Os routers encaminham o tráfego multicast para todos os vizinhos que não tenham enviado mensagens de *Prune*
- Uma mensagem *Prune* é enviada por:
 - Um router que não tenha clientes interessados na sessão multicast em causa (Ex: Router 5)
 - Um router que tenha recebido o tráfego multicast em mais que um interface (Ex: Router 4)
 - Envia Prune para todos os interfaces onde recebeu o tráfego excepto no que proporciona o percurso de menor custo até à fonte
 - Um router que tenha recebido prunes em todos os interfaces para onde encaminhou tráfego multicast (Ex: Router 3)
- Quando um Router deixa ter clientes interessados numa sessão multicast envia uma mensagem *Prune* pelo interface onde está a receber o respectivo tráfego

PIM DM - Mensagem *Graft*



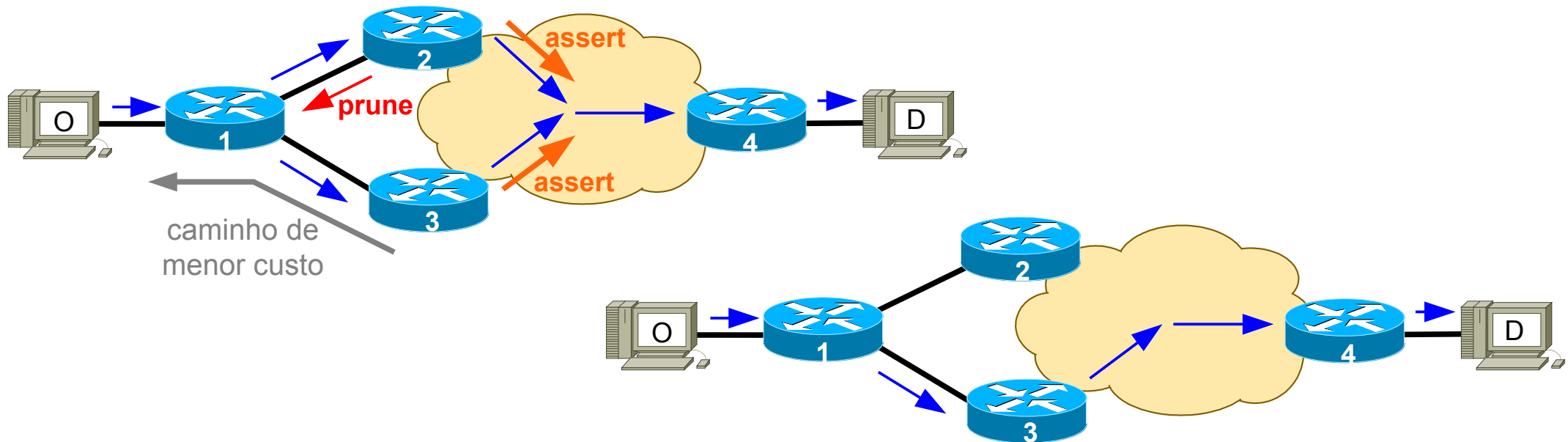
- Para um router recomençar a receber tráfego multicast envia uma mensagem *Graft* para anular o *Prune* enviado anteriormente por si

PIM DM – Mensagem *Join*



- Quando um Router deixa de ter clientes de uma determinada sessão multicast envia uma mensagem *Prune*
- Mas se for enviada para um meio partilhado caso hajam mais routers interessados na respectiva sessão multicast deverão enviar uma mensagem *Join* para anular a mensagem *Prune* enviada

PIM DM – Mensagem Assert



- Caso exista mais que um Router a reenviar tráfego de uma determinada sessão multicast para um meio partilhado estes devem decidir qual é o responsável por esse envio
- Todos enviam uma mensagem *Assert* com o respectivo custo até à fonte
- É escolhido o Router
 - ◆ Que proporciona o percurso de custo mais baixo até à fonte
 - ◆ Em caso de igualdade é escolhido o Router com o **maior endereço IP** (ligado ao meio partilhado)

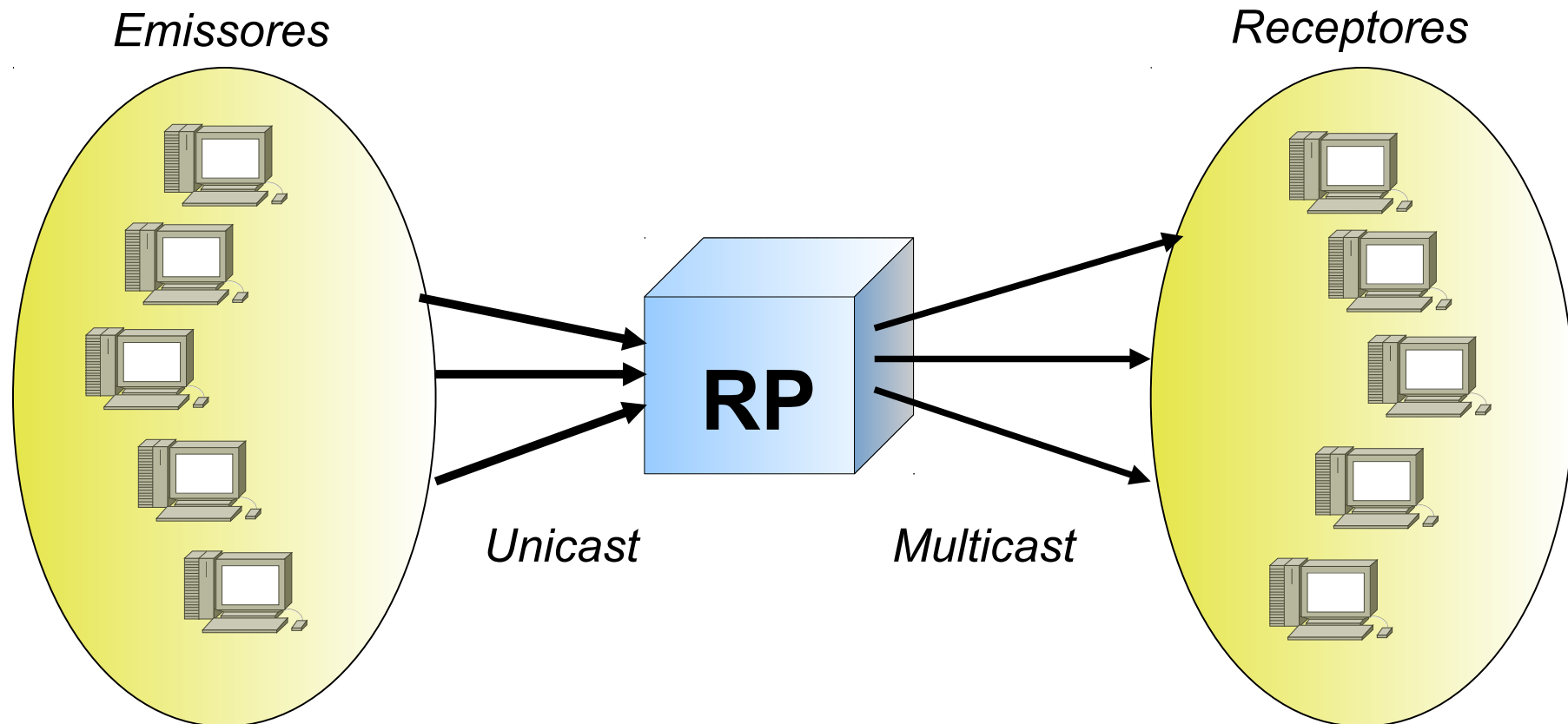
PIM Sparse Mode

- O PIM, Dense Mode é um protocolo data-driven
 - ♦ Exige que os routers que não têm clientes multicast estejam periodicamente a enviar mensagens prune para evitar que lhes enviem pacotes multicast
- O PIM, Sparse Mode é um protocolo receiver-driven
 - ♦ Cada router anuncia explicitamente (com mensagens *Join*) que quer suportar determinadas sessões multicast
- O PIM, Sparse Mode usa inicialmente uma estratégia de “Group-Shared tree” baseado num router Rendezvous Point (RP)
 - ♦ O RP pode ser configurado administrativamente
 - ♦ Existem também métodos automáticos tais como CISCO RP Discovery que usa o endereço multicast 224.0.1.40
 - ♦ Podem existir diferentes RPs para diferentes sessões multicast



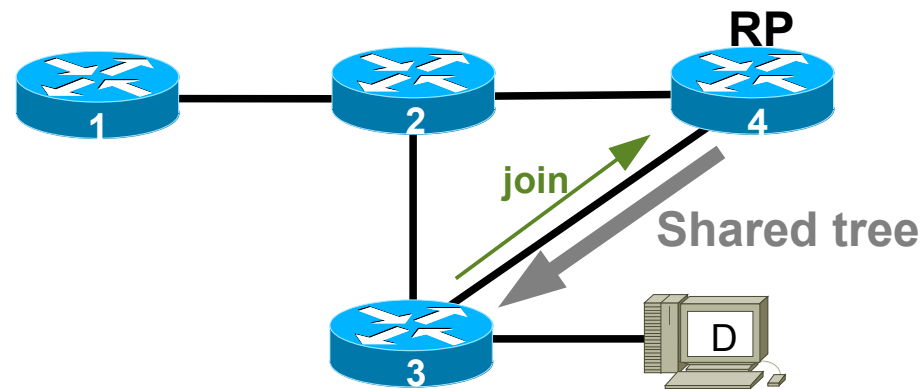
PIM Sparse Mode

- Inicialmente, os pacotes de uma sessão são enviados para o RP por unicast (através de túneis)
- Os pacotes são distribuídos pelo RP para os receptores por multicast

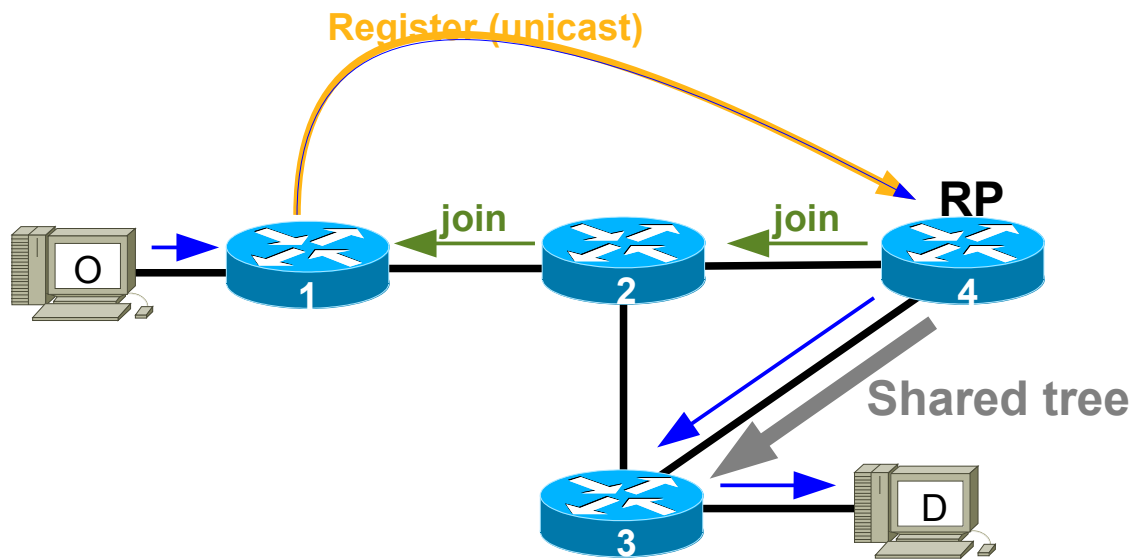


PIM SM – Adesão à *Group-Shared Tree*

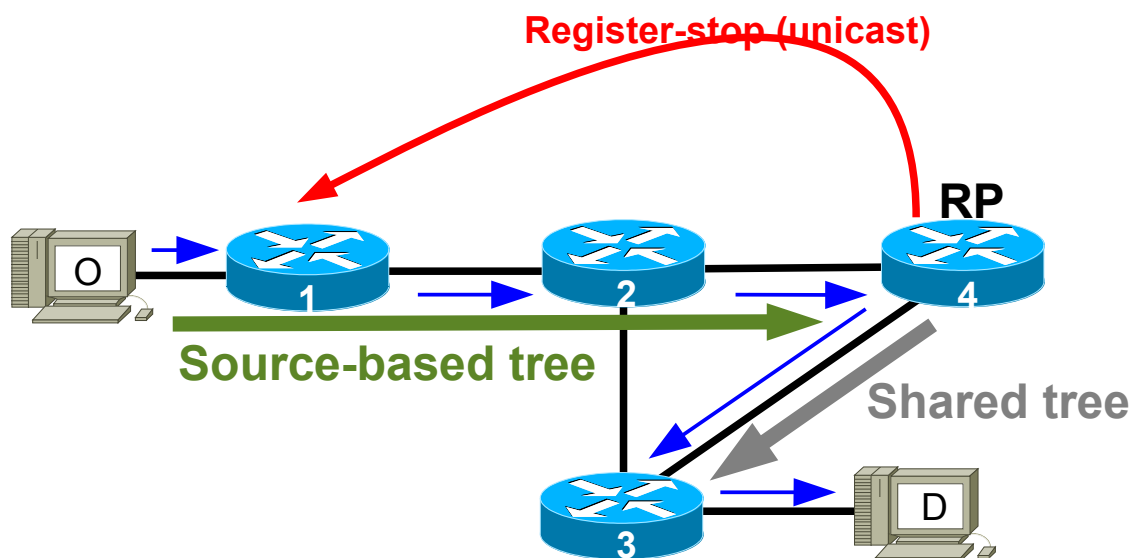
- Os routers receptores enviam uma mensagem *Join* em direcção ao RP de modo a se juntarem à group-shared tree que tem a sua raiz no Rendezvous Point



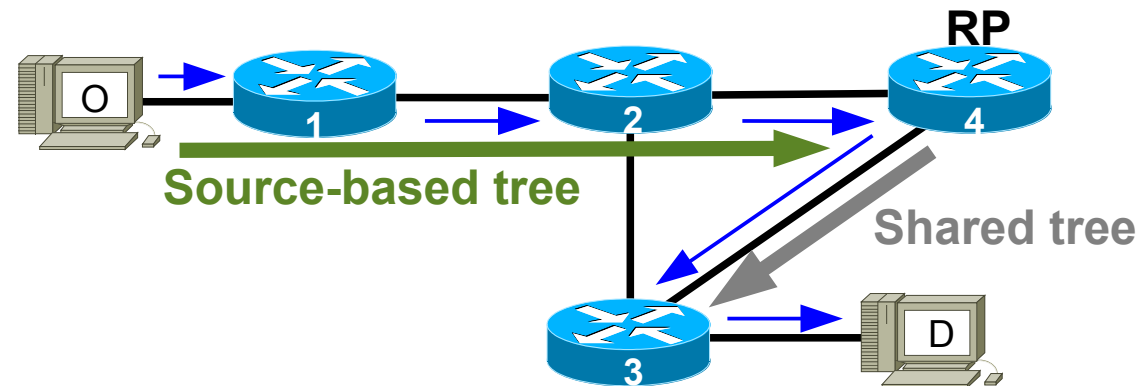
PIM SM – Nova fonte de tráfego



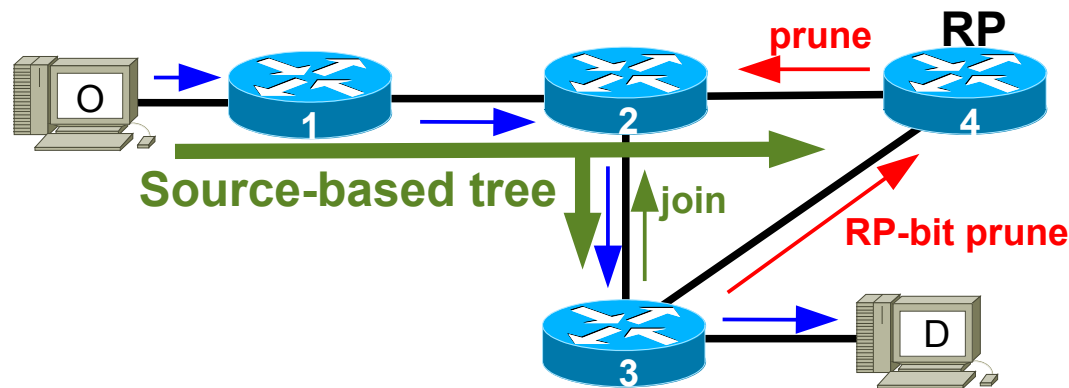
- Um router emissor envia os pacotes multicast encapsulados em pacotes unicast (túnel IP) para o RP através de mensagens *Register*
- O RP recupera os pacotes multicast e, simultaneamente:
 - ◆ Reenvia o pacote pela árvore estabelecida para todos os receptores
 - ◆ No primeiro pacote recebido, envia uma mensagem *Join* para o router emissor para estabelecer uma source-based tree do router emissor
- Assim que o router emissor começa a enviar os pacotes multicast pela source-based tree, o RP envia um *Register-stop* para deixar de receber os pacotes encapsulados



PIM SM – Adesão à *Source-based Tree*



- Quando o bit rate agregado (de todas as fontes) excede um dado limiar um router pode optar por aderir à Source-based Tree.
 - Envia um *Join* em direcção à fonte do fluxo multicast, que passará por todos os routers até que encontra um router que já pertença à Source-based Tree.



- Quando um pacote for recebido através desta árvore, é enviado uma mensagem *RP-bit Prune* em direcção ao RP

Source-Specific Multicast para IP

- O SSM é uma extensão ao modelo de encaminhamento Multicast
 - ♦ Tem por base um encaminhamento puramente baseado em *Source-based trees*.
 - ♦ A implementação SSM do PIM (PIM-SSM) tem por base um subconjunto dos mecanismos do PIM-SM.
- Os receptores especificam a sessão multicast e a fonte
 - ♦ O router receptor estabelece uma árvore do tipo *Source-based tree* enviando um “join” diretamente para a fonte.
- Endereços reservados para SSM
 - ♦ IPv4 da gama 232.0.0.0/8.
 - ♦ IPv4 da gama FF3x::/32.
- Necessita do IGMPv3
 - ♦ Para além da sessão multicast, o receptor especifica de que fonte quer receber os pacotes multicast.



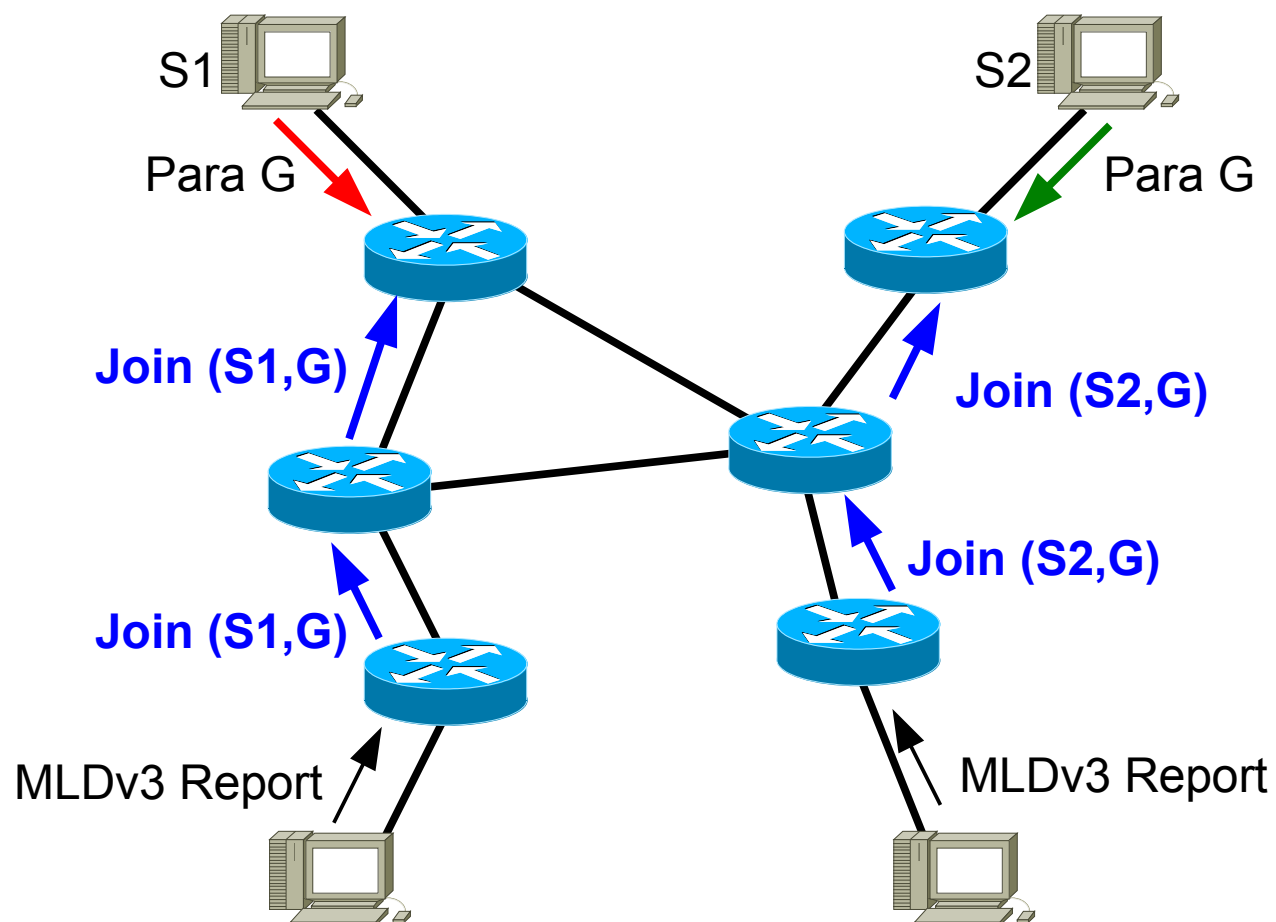
IGMP versão 3

- A versão 3 do IGMP adiciona o suporte para "source filtering" ou "filtragem de fontes".
 - ◆ Permite que um terminal reporte o interesse em receber pacotes enviados para um determinado endereço multicast de:
 - APENAS de uma fonte específica, ou
 - TODOS com a exceção de fontes específicas.
 - ◆ Permite a adesão simultânea a múltiplos grupos multicast.
 - ◆ Implica a definição de um novo pacote do tipo "Report".
 - Version 3 Membership Report.
- Define grupos de adesão
 - ◆ Em modo INCLUDE – Adere a um grupo multicast, das fontes listadas.
 - Se a lista for vazia, anula a adesão.
 - ◆ Em modo EXCLUDE – Adere a um grupo multicast, das fontes não listadas.
 - Se a lista for vazia, faz uma adesão para todas as fontes.
- Permite interoperabilidade com o IGMPv1 e IGMPv2
 - ◆ Suporta os pacotes Version 1 Membership Report, Version 2 Membership Report e Version 2 Leave Group.



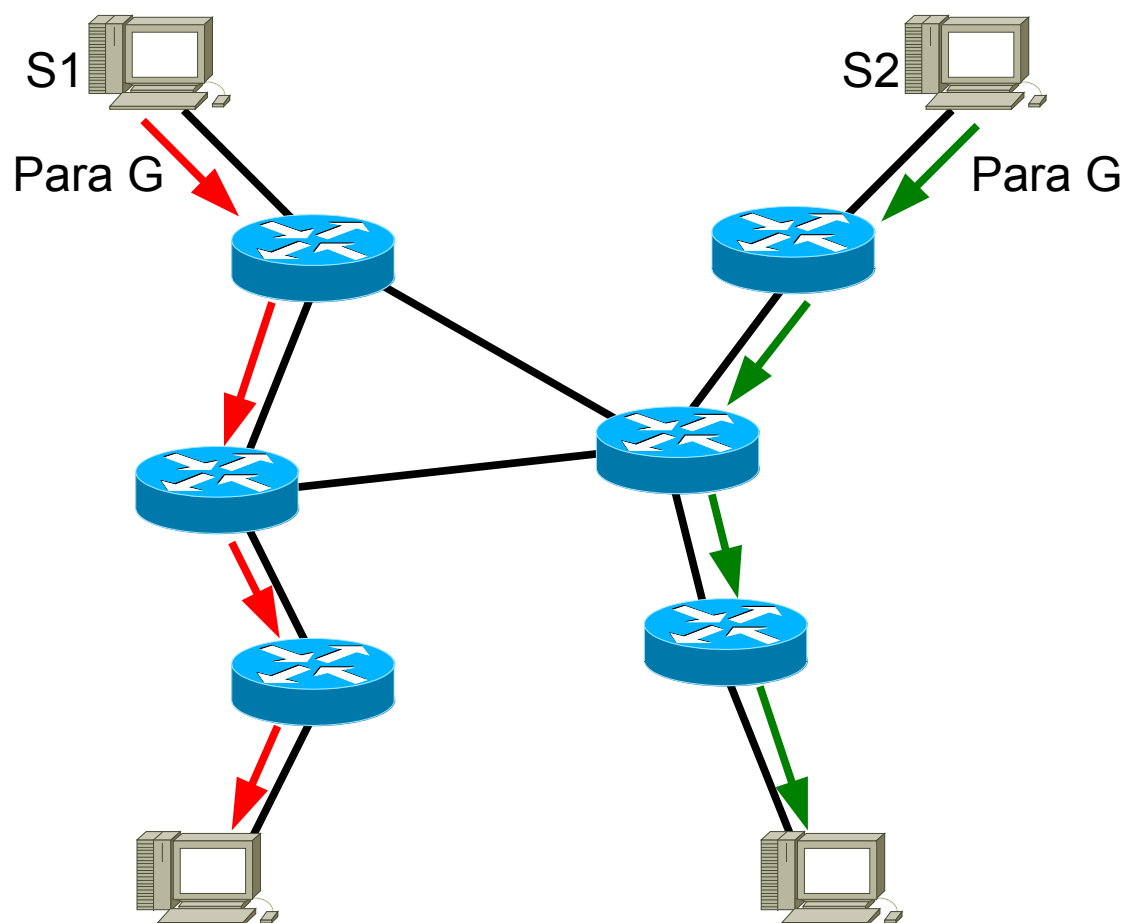
Exemplo PIM-SSM (1)

- MLD Reports são enviados para aderir ao grupo multicast G da fonte S1 e S2.
- OS routers enviam pacotes PIM Join para construir a Source-based tree para o grupo G da fonte S1 e S2.



Exemplo PIM-SSM (2)

- Os routers após a criação da *Source-based tree* reencaminham o tráfego do grupo multicast G até aos terminais.

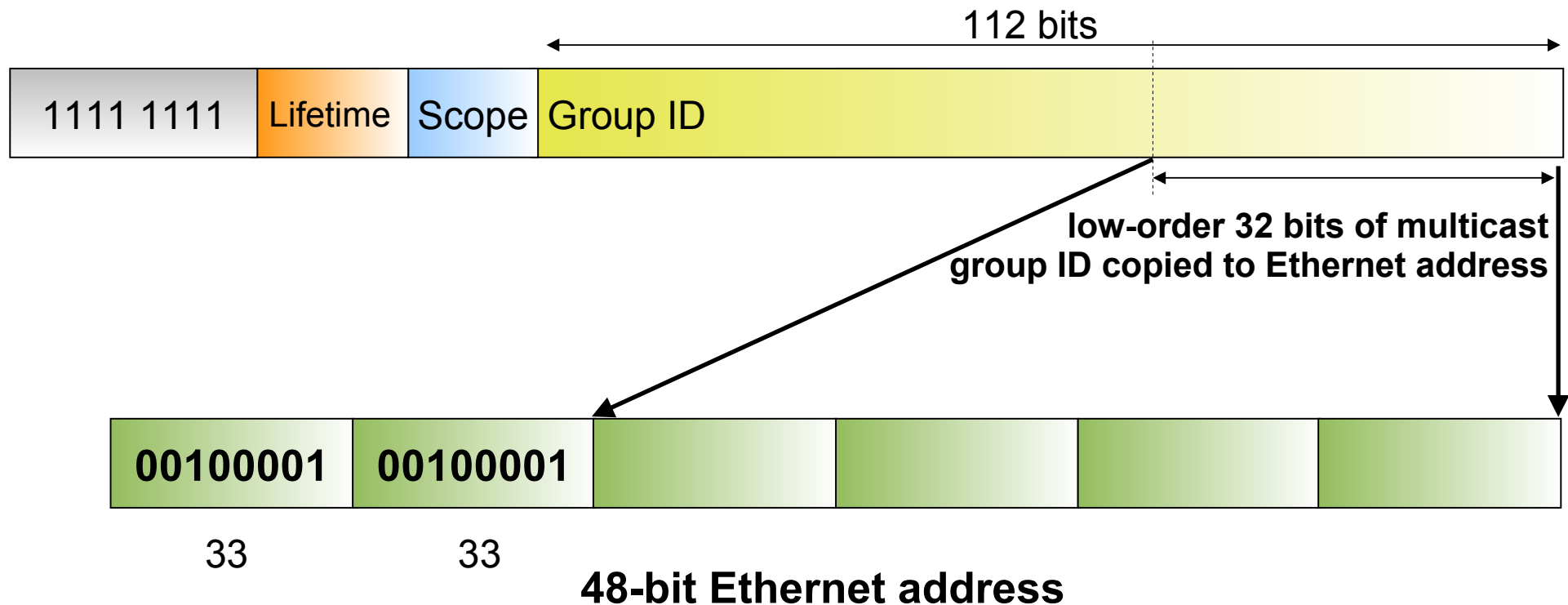


Multicast em IPv6

- O protocolo IGMP é substituído pelo protocolo Multicast Listener Discovery (MLD).
 - O MLD é equivalente ao IGMP do IPv4,
 - O MLD é uma adaptação do IGMP à semântica do IPv6.
 - ➔ MLDv1 \Leftrightarrow IGMPv2.
 - ➔ MLDv2 \Leftrightarrow IGMPv3.
 - As mensagens MLD são transportadas sobre ICMPv6.
 - ➔ Multicast Listener Query, Multicast Listener Report e Multicast Listener Done.
 - O MLD usa endereços link local como origem.
- As árvores de encaminhamento podem ser construídas em modo *Sparse ou Source-specific*.
 - O modo *dense* não é usado!
 - O modo Source-specific requer o MLDv2.
 - O protocolo PIM continua a ser usado em IPv6.
 - ➔ PIM-SM e PIM-SSM.



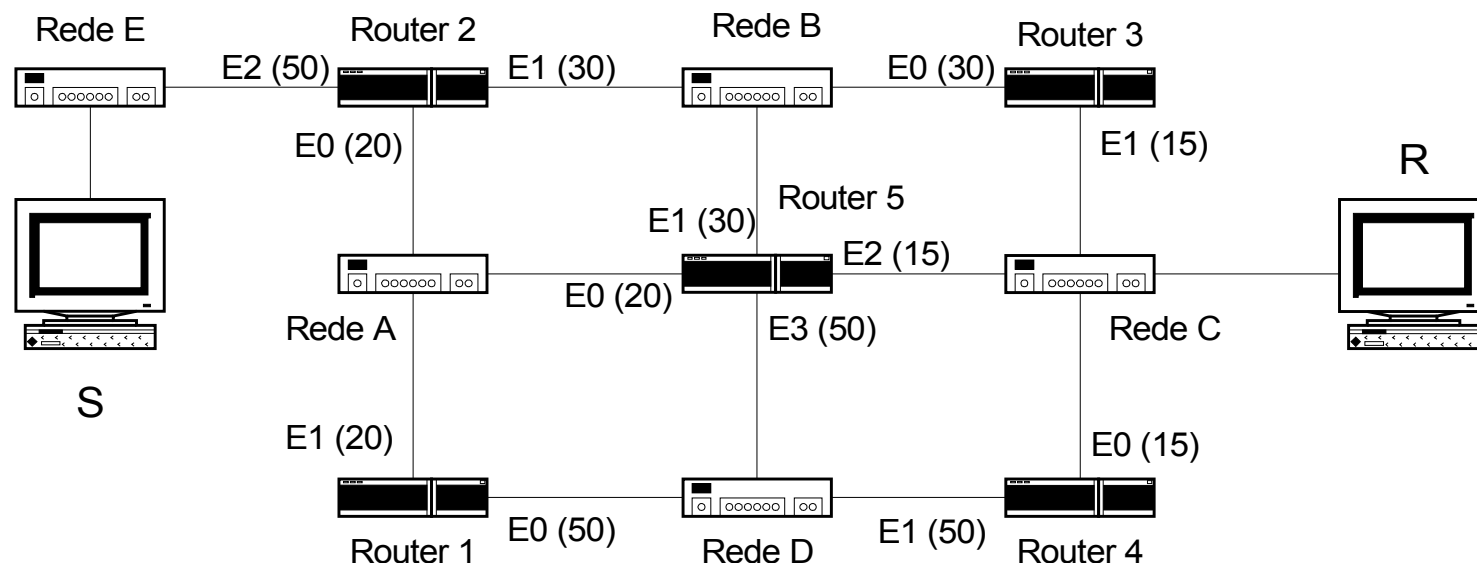
Conversão de endereço IPv6 Multicast em endereço IEEE 802



Exemplo 1:

Na rede da figura, considere que os *routers* estão configurados com o OSPF (custos das portas indicados na figura) e com o PIM *dense-mode*. Assuma que o terminal S envia periodicamente pacotes para o endereço 230.20.1.1 e que o terminal R já aderiu à sessão *multicast* 230.20.1.1.

- (1) Descreva como é que o primeiro pacote enviado por S se propaga pela rede.
- (2) Apresente a tabela de encaminhamento *multicast* de cada um dos *routers*; para cada router use a notação: $O - G - E - S_1, S_2, \dots$ (O – endereço origem; G – grupo multicast; E – interface de entrada; S_1, S_2, \dots – lista de interfaces de saída)



Exemplo 2:

Na rede da figura, considere que os *routers* estão configurados com o OSPF (custos das portas indicados na figura) e com o PIM *sparse-mode* em que o endereço do *Rendezvous Point* é o endereço da interface E1 do *Router 1*. Indique que pacotes são trocados e qual o seu percurso:

- (1) no início, quando o terminal R adere à sessão 225.3.3.3;
- (2) depois, quando o terminal S envia um pacote para o destino 225.3.3.3.

