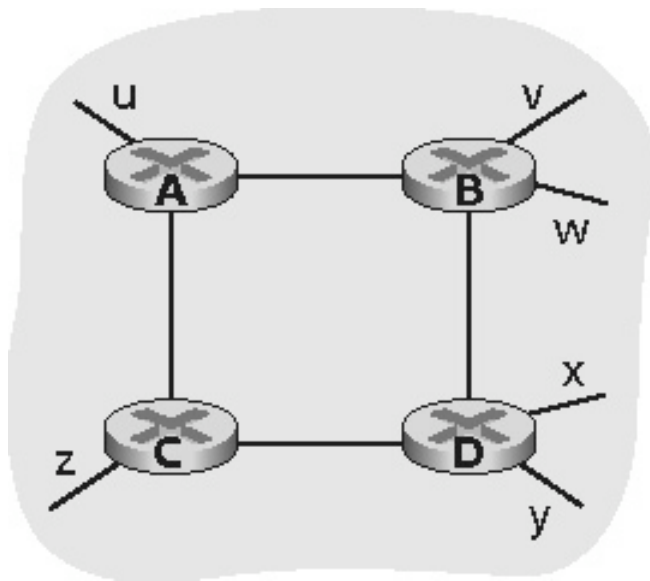


4 RIP (Routing Information Protocol)

- Algoritmo do tipo vetor distância
- Incluso na distribuição do BSD-UNIX em 1982
- Métrica de distância: # de saltos (máx. = 15 saltos)

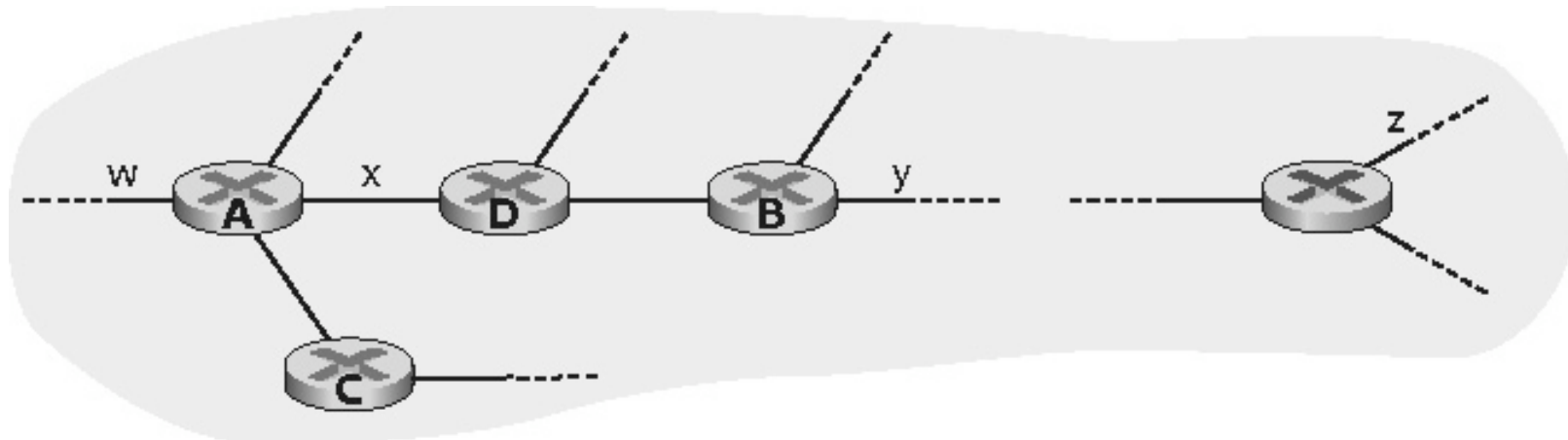


Destino	Saltos
u	1
v	2
w	2
x	3
y	3
z	2

4 Anúncio RIP

- Vetores de distância: trocados a cada 30 s via Response Message (também chamado **advertisement**, ou anúncio)
- Cada anúncio indica rotas para até 25 redes de destino

4 RIP: Exemplo



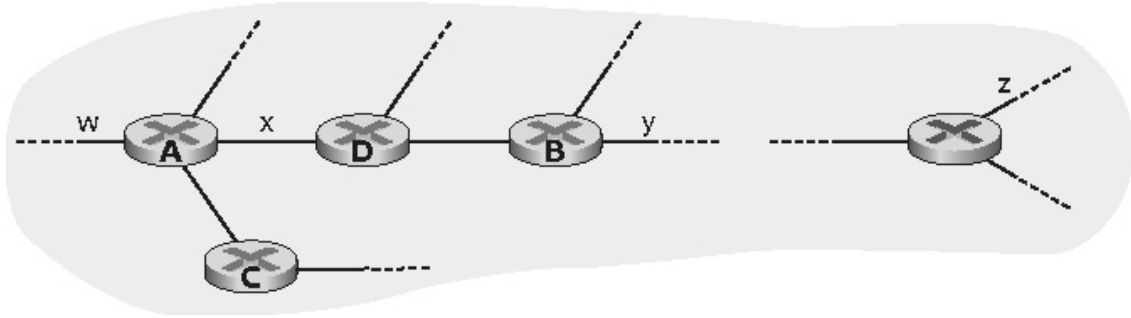
rede de destino	roteador seguinte	núm. de saltos para dest.
w	A	2
y	B	2
z	B	7
x	--	1
....

Tabela de roteamento em D

4 RIP: Exemplo

dest.	próximos saltos	
w	-	-
x	-	-
z	C	4
....

Anúncio de A para D



rede de destino	roteador seguinte	núm. de saltos até dest.
w	A	2
y	B	2
z	B A	7 5
x	--	1
....

Routing table in D

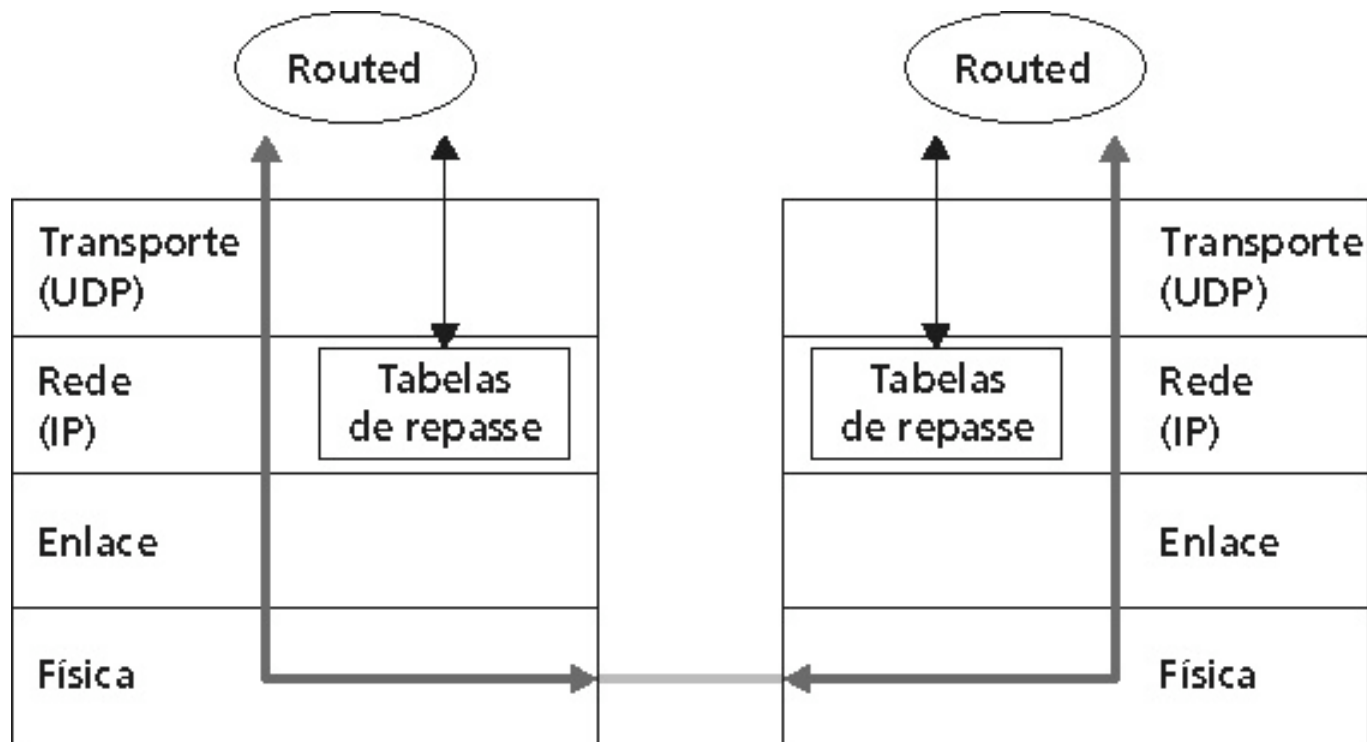
4 RIP: falha de enlaces e recuperação

Se não há um aviso depois de 180 s, --> o vizinho e o enlace são declarados mortos

- Rotas através do vizinho são anuladas
- Novos anúncios são enviados aos vizinhos
- Os vizinhos por sua vez devem enviar novos anúncios (se suas tabelas de rotas foram alteradas)
- A falha de um enlace se propaga rapidamente para a rede inteira
- Reversão envenenada é usada para prevenir loops, (distância infinita = 16 saltos)

4 RIP Processamento da tabela de rotas

- As tabelas de roteamento do RIP são manipuladas por um processo de aplicação chamado route-d (daemon)
- Anúncios são enviados em pacotes UDP com repetição periódica



4 A camada de rede

- 4.1 Introdução
- 4.2 Circuito virtual e redes de datagrama
- 4.3 O que há dentro de um roteador
- 4.4 IP: Protocolo da Internet
 - Formato do datagrama
 - Endereçamento IPv4
 - ICMP
 - IPv6
- 4.5 Algoritmos de roteamento
 - Link state
 - Distance vector
 - Roteamento hierárquico
- 4.6 Roteamento na Internet
 - RIP
 - OSPF
 - BGP
- 4.7 Roteamento de broadcast e multicast

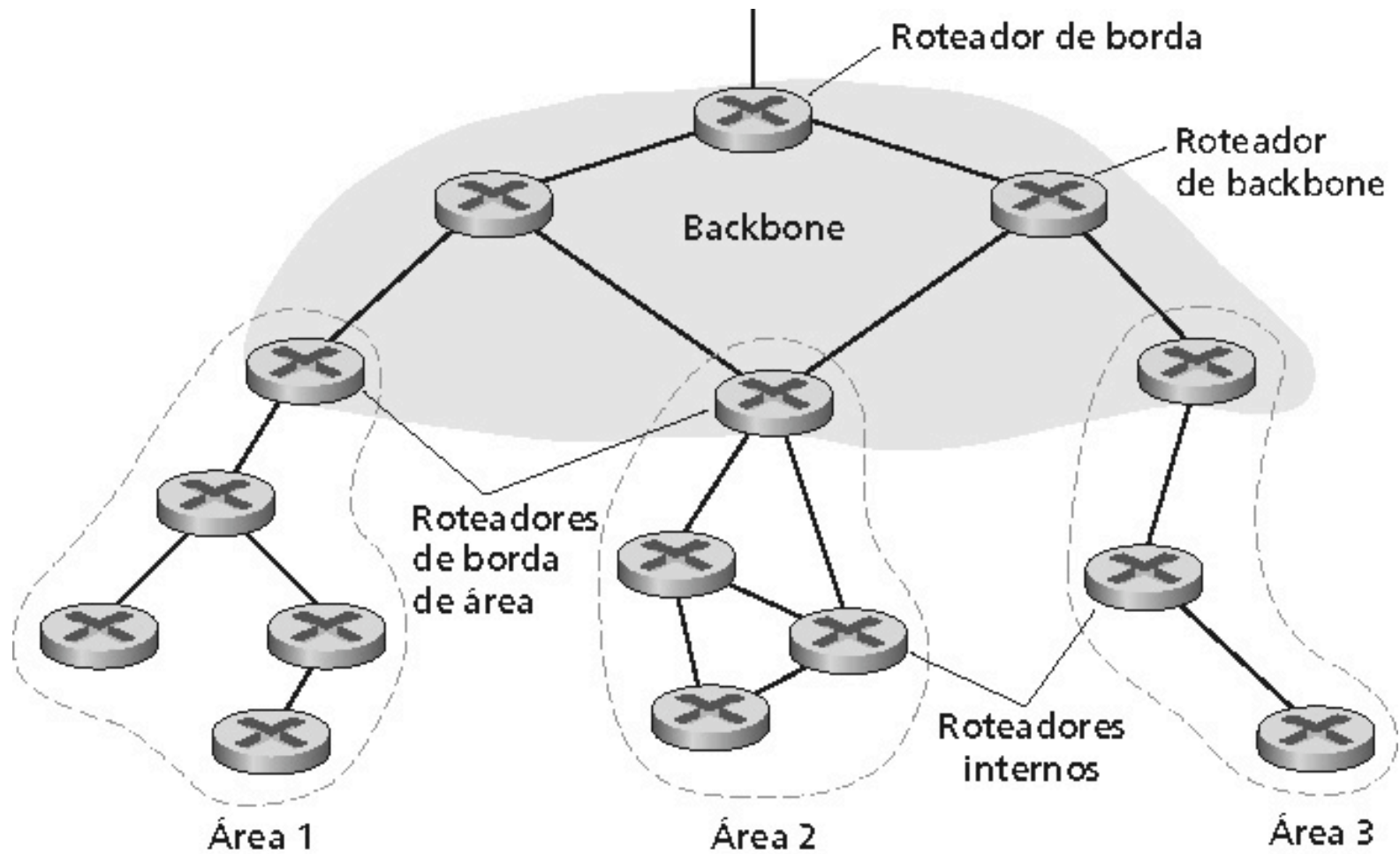
4 OSPF (Open Shortest Path First)

- “Open”: publicamente disponível
- Usa algoritmo do tipo link state
 - Disseminação de pacotes LS
 - Mapa topológico em cada nó
 - Usa algoritmo de Dijkstra para cálculo de rotas
- Anúncios do OSPF transportam um registro para cada roteador vizinho
- Anúncios são distribuídos para **todo** o AS (via flooding)
 - Transportado por mensagens OSPF diretamente sobre IP (melhor do que TCP ou UDP)

4 OSPF características “avançadas”

- **Segurança:** todas as mensagens do OSPF são autenticadas (para prevenir intrusões maliciosas)
- Múltiplos caminhos de mesmo custo são permitidos (o RIP só permite um caminho)
- Para cada link, múltiplas métricas de custo para **TOS** diferentes (ex.: custo de enlace por satélite definido baixo para tráfego de “melhor esforço” e alto para serviços de tempo real)
- Integra tráfego uni- e **multicast:**
 - Multicast OSPF (MOSPF) usa a mesma base de dados de topologia do OSPF
- **OSPF hierárquico:** OSPF para grandes domínios

4 OSPF hierárquico



4 OSPF hierárquico

- **Hierarquia de dois níveis:** área local e backbone.
 - Anúncios de link state apenas na área
 - Cada nó tem a topologia detalhada da área, mas somente direções conhecidas (caminhos mais curtos) para redes em outras áreas
- **Roteadores de borda de área:** “resumem” distâncias para redes na própria área e enviam para outros roteadores de borda de área
- **Roteadores de backbone:** executam o roteamento OSPF de forma limitada ao backbone.
- **Roteadores de borda:** conectam-se a outras AS's.

4 A camada de rede

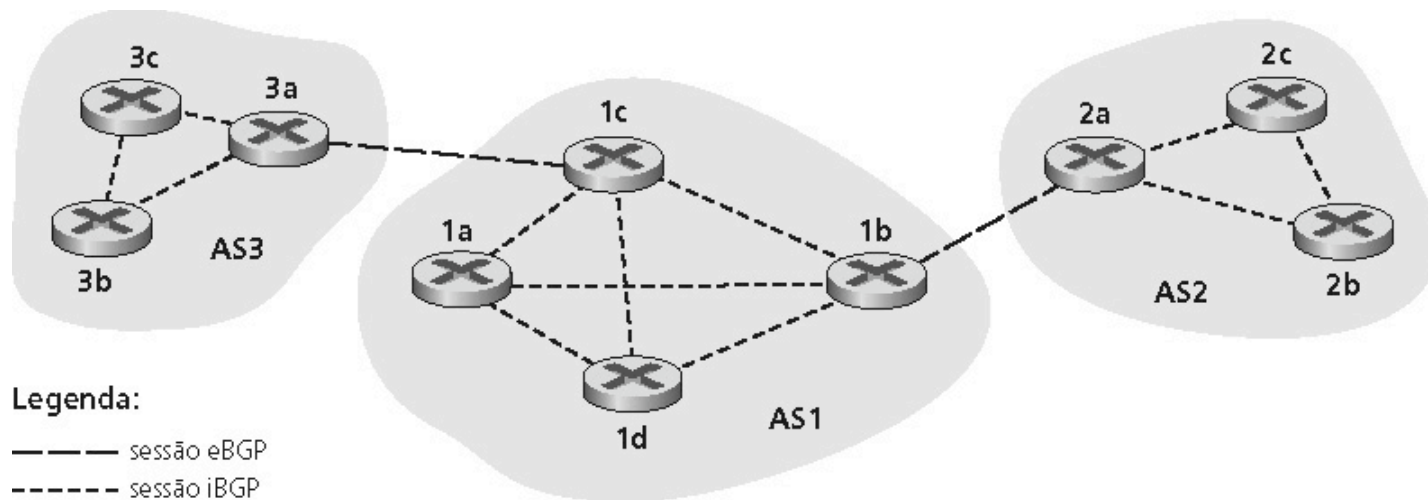
- 4.1 Introdução
- 4.2 Circuito virtual e redes de datagrama
- 4.3 O que há dentro de um roteador
- 4.4 IP: Protocolo da Internet
 - Formato do datagrama
 - Endereçamento IPv4
 - ICMP
 - IPv6
- 4.5 Algoritmos de roteamento
 - Link state
 - Distance vector
 - Roteamento hierárquico
- 4.6 Roteamento na Internet
 - RIP
 - OSPF
 - BGP
- 4.7 Roteamento de broadcast e multicast

4 Roteamento inter-AS da Internet: BGP

- **BGP (Border Gateway Protocol):** é o padrão de fato para uso na Internet
- BGP provê cada AS dos meios para:
 1. Obter informações de alcance de sub-rede dos ASs vizinhos
 2. Propagar informações de alcance para todos os roteadores internos ao AS
 3. Determinar “boas” rotas para as sub-redes baseado em informações de alcance e política
- Permite que uma subnet comunique sua existência para o resto da Internet:
“Estou aqui”

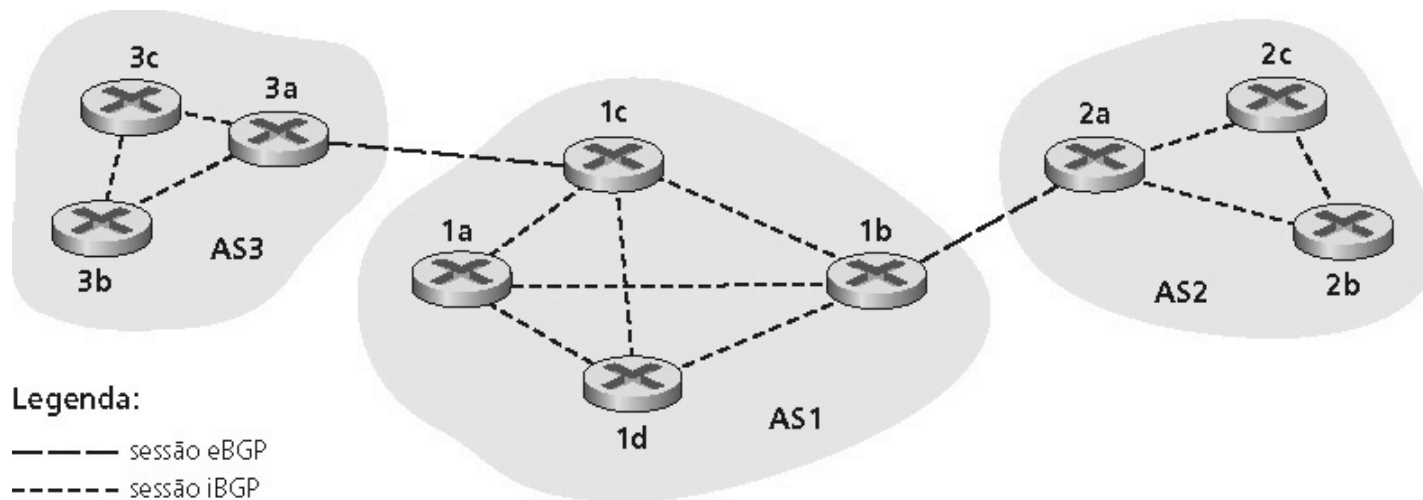
4 BGP: conceitos básicos

- Pares de roteadores (BGP peers) trocam informações de roteamento por conexões TCP semipermanentes: **sessões BGP**
- Note que as sessões BGP não correspondem aos links físicos
- Quando AS2 comunica um prefixo ao AS1, AS2 está **prometendo** que encaminhará todos os datagramas destinados a esse prefixo em direção ao prefixo
 - AS2 pode agregar prefixos em seu comunicado



4 Distribuindo informações de alcance

- Em cada sessão eBGP entre 3a e 1c, AS3 envia informações de alcance de prefixo para AS1.
- 1c pode então usar iBGP para distribuir essa nova informação de alcance de prefixo para todos os roteadores em AS1
- 1b pode recomunicar essa nova informação para AS2 por meio da sessão eBGP 1b-para-2a.
- Quando um roteador aprende um novo prefixo, ele cria uma entrada para o prefixo em sua tabela de roteamento.



4 Atributos de caminho e rotas BGP

- Quando se comunica um prefixo, o comunicado inclui os atributos do BGP.
 - Prefixo + atributos = “rota”
- Dois atributos importantes:
 - **AS-PATH**: contém os ASs pelos quais o comunicado para o prefixo passou: AS 67 AS 17
 - **NEXT-HOP**: Indica o roteador específico interno ao AS para o AS do próximo salto (next-hop). (Pode haver múltiplos links do AS atual para o AS do próximo salto.)
- Quando um roteador gateway recebe um comunicado de rota, ele usa **política de importação** para aceitar/rejeitar.

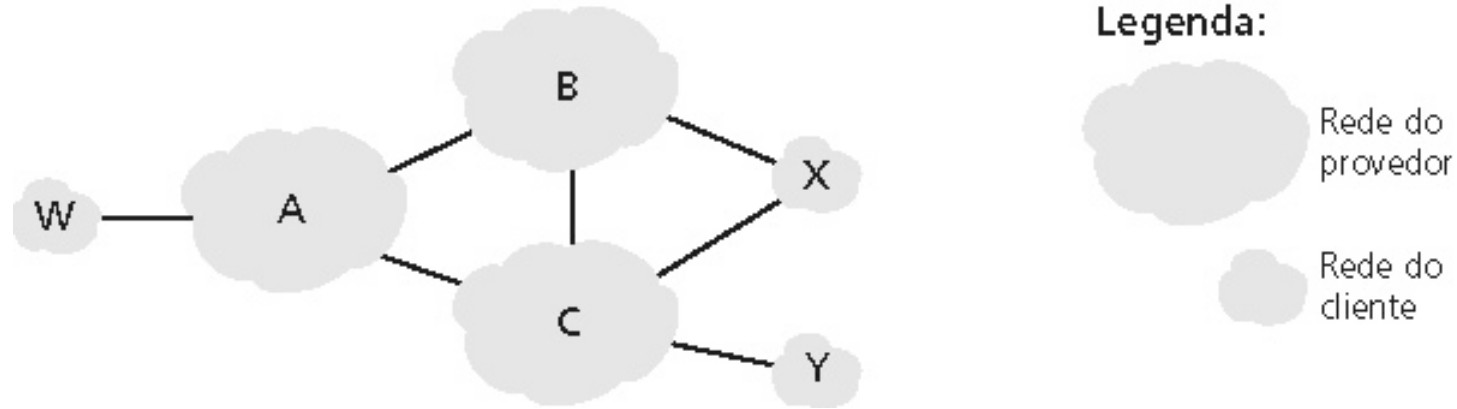
4 BGP: seleção de rota

- Um roteador pode aprender mais do que 1 rota para o mesmo prefixo. O roteador deve selecionar uma rota
- Regras de eliminação:
 - Atributo de valor de preferência local: decisão de política
 - AS-PATH (caminho) mais curto
 - Roteador do NEXT-HOP (próximo salto) mais próximo: roteamento da “batata quente”
 - Critérios adicionais

4 Mensagens BGP

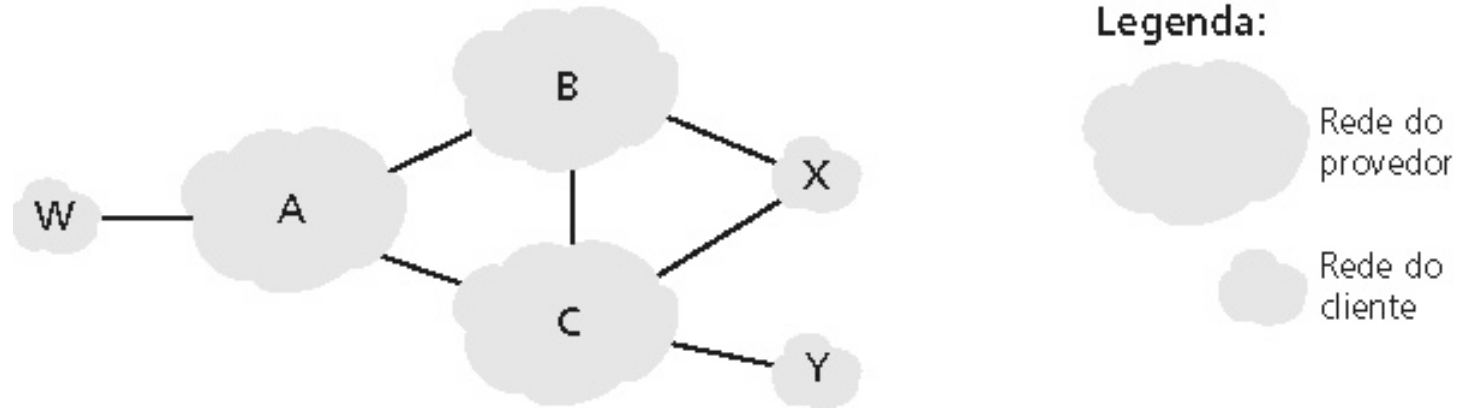
- Mensagens BGP são trocadas usando o TCP
- Mensagens BGP:
 - **OPEN**: abre conexão TCP para o peer e autentica o transmissor
 - **UPDATE**: comunica novo caminho (ou retira um antigo)
 - **KEEPALIVE** mantém a conexão ativa na ausência de atualizações (updates); também ACKs OPEN request
 - **NOTIFICATION**: reporta erros em mensagens anteriores; também usado para fechar a conexão

4 BGP: política de rotear



- A, B, C são **redes do provedor**
- X, W, Y são clientes (das redes do provedor)
- X é **dual-homed**: anexados a duas redes
 - X não quer rotear de B via X para C
 - ... então X não comunicará ao B uma rota para C

4 BGP: política de roteamento (2)



- A comunica ao B o caminho AW
- B comunica ao X o caminho BAW
- B deveria comunicar ao C o caminho BAW?
 - De jeito nenhum! B não obtém nenhum “rendimento” em rotear CBAW pois nem W nem C são seus clientes
 - B quer forçar C a rotear para W via A
 - B quer rotear **somente** de/para seus clientes!

4 Por que os protocolos intra e inter-AS são diferentes?

Políticas:

- Inter-AS: a administração quer ter controle sobre como seu tráfego é roteado e sobre quem roteia através da sua rede.
- Intra-AS: administração única, então não são necessárias políticas de decisão

Escalabilidade

- O roteamento hierárquico poupa espaço da tabela de rotas e reduz o tráfego de atualização

Desempenho:

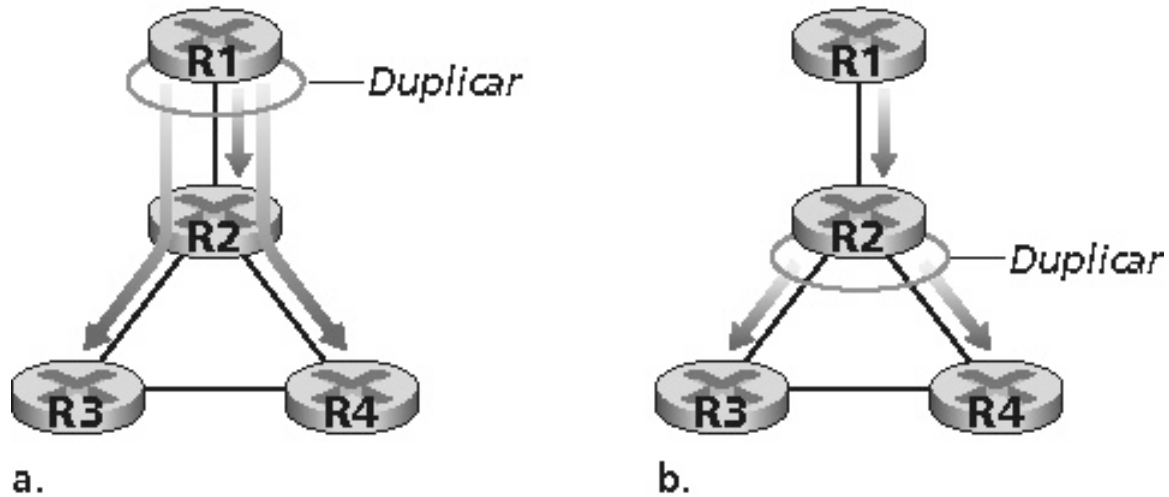
- Intra-AS: preocupação maior é desempenho
- Inter-AS: políticas podem ser dominantes em relação ao desempenho

4 A camada de rede

- 4.1 Introdução
- 4.2 Circuito virtual e redes de datagrama
- 4.3 O que há dentro de um roteador
- 4.4 IP: Protocolo da Internet
 - Formato do datagrama
 - Endereçamento IPv4
 - ICMP
 - IPv6
- 4.5 Algoritmos de roteamento
 - Link state
 - Distance vector
 - Roteamento hierárquico
- 4.6 Roteamento na Internet
 - RIP
 - OSPF
 - BGP
- 4.7 Roteamento de broadcast e multicast

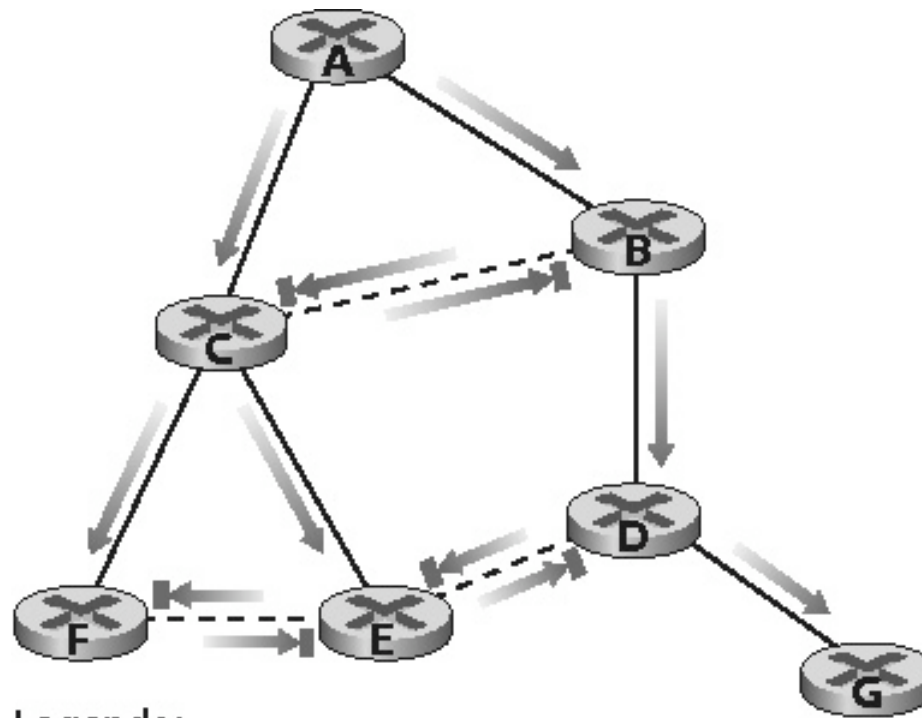
4 Duplicação na origem *versus* duplicação na rede

Criação/transmissão de duplicatas



(a) duplicação na origem, (b) duplicação na rede

4 Repasse pelo caminho reverso

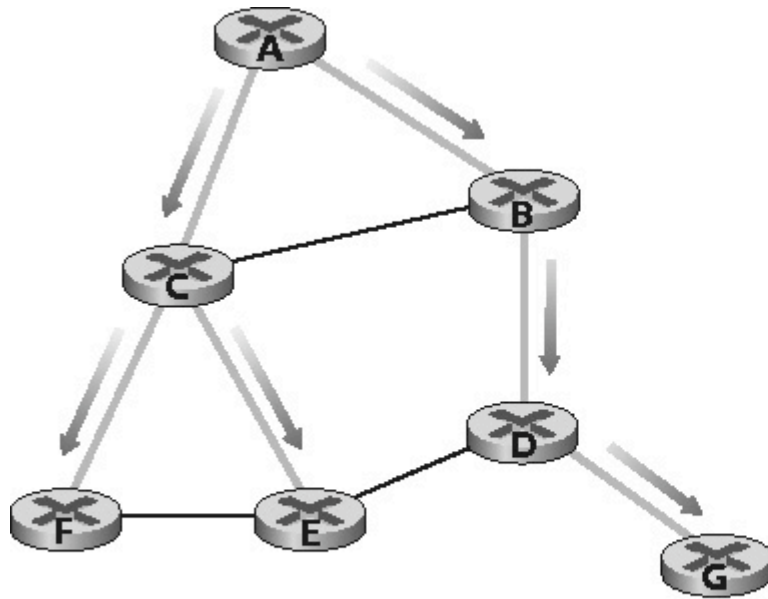


Legenda:

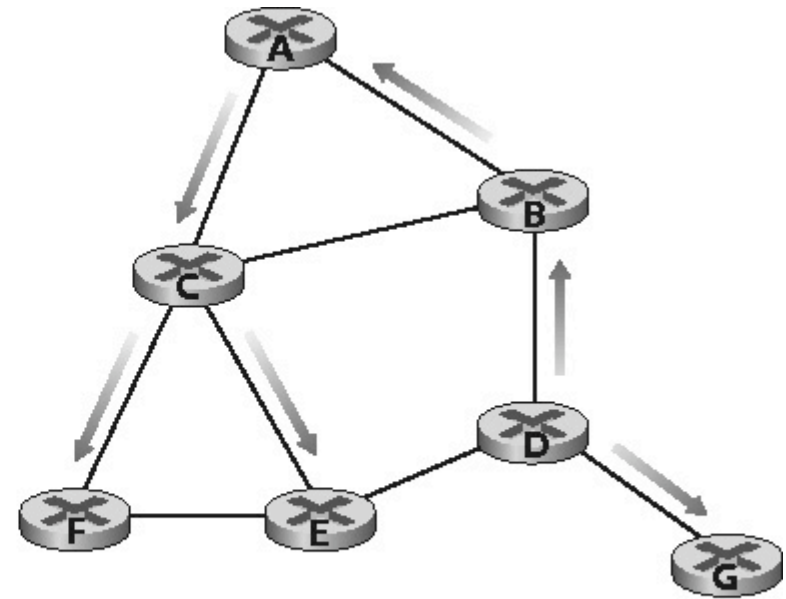
→ pacote (pkt) será repassado

→ pacote (pkt) não será repassado além do roteador receptor

4 Broadcast ao longo de uma spanning tree



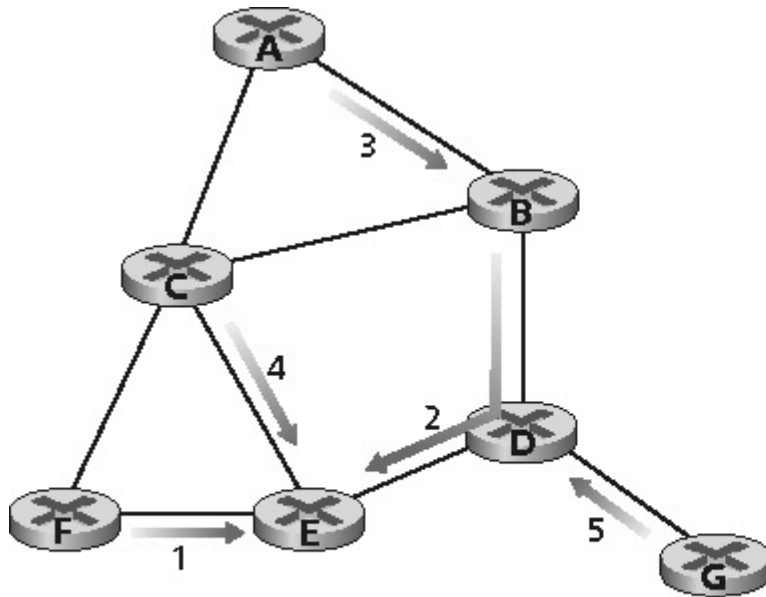
a. Broadcast iniciado em A



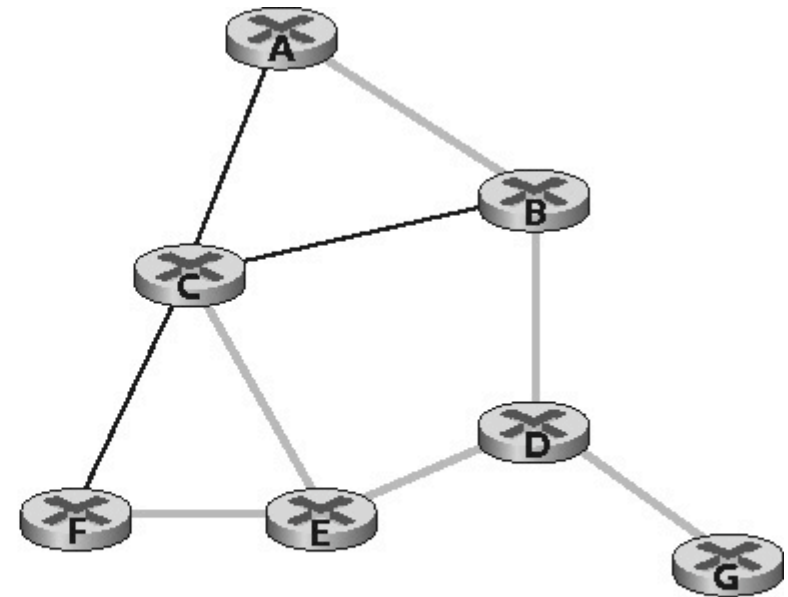
b. Broadcast iniciado em D

4

Construção centro-baseada de uma spanning tree



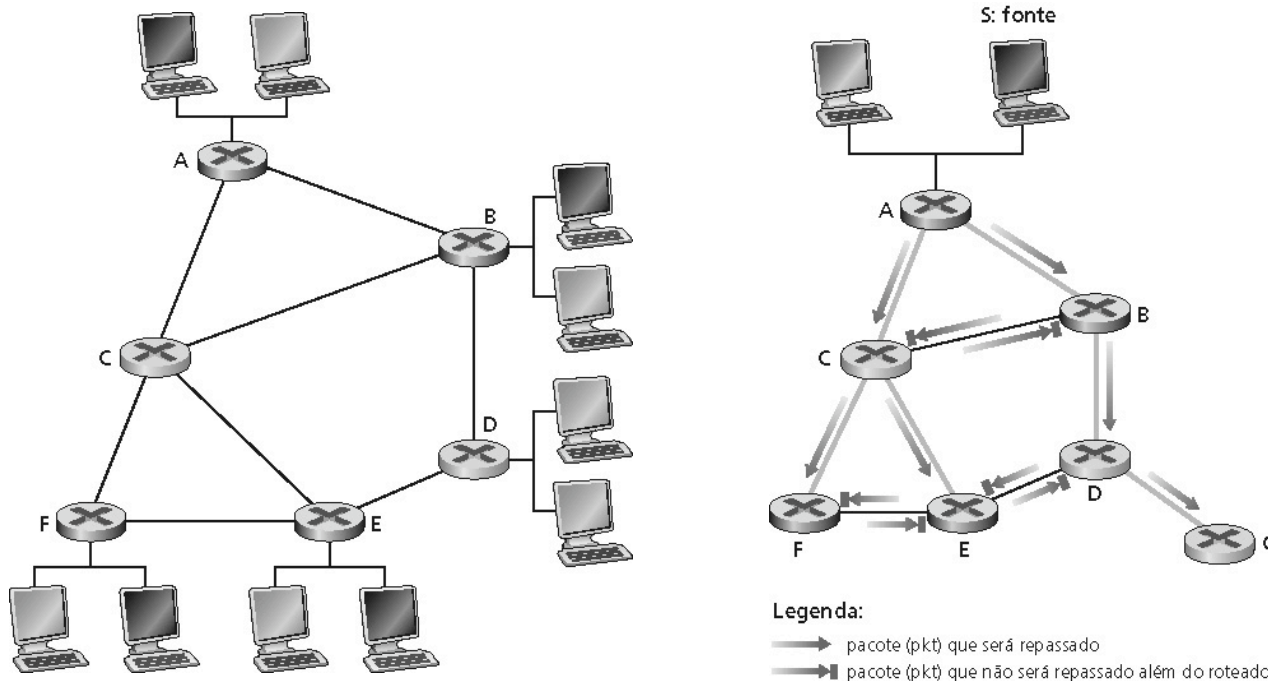
a. Construção da spanning tree passo a passo



b. Spanning tree construída

4 Roteamento multicast: indicação do problema

- **Objetivo:** encontrar uma árvore (ou árvores) conectando roteadores que possuem membros de grupo multicast local
 - **Árvore:** não são todos os caminhos entre os roteadores usados
 - **Baseada na fonte:** uma árvore diferente de cada transmissor para os receptores
 - **Árvore compartilhada:** a mesma árvore é usada por todos o membros do grupo



4 Métodos para construir multicast trees

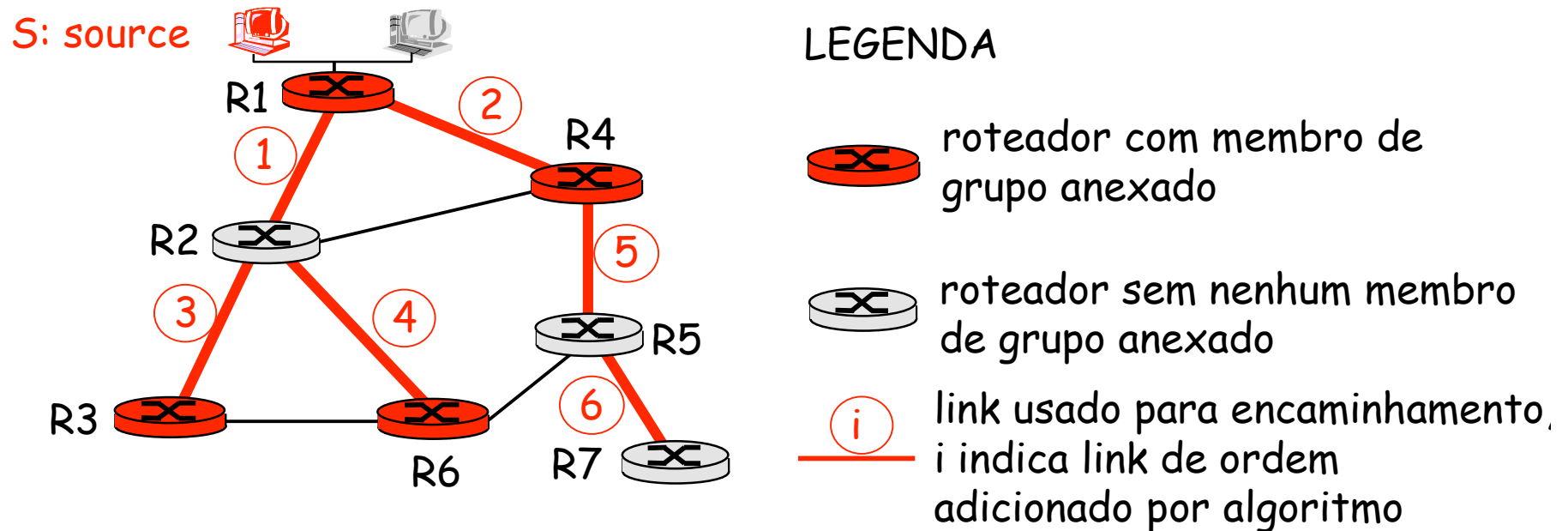
Métodos:

- **Árvore baseada na fonte:** uma árvore por origem
 - Shortest path trees
 - Repasse pelo caminho reverso
- **Árvore compartilhada pelo grupo:** grupo usa uma árvore
 - Minimal spanning (Steiner)
 - Center-based trees

... primeiro veremos métodos básicos e, então, os protocolos específicos que adotam estes métodos

4 Shortest Path Tree

- mcast forwarding tree: árvore de rotas de caminho mais curto da origem para todos os receptores
 - Algoritmo de Dijkstra



4 Reverse Path Forwarding

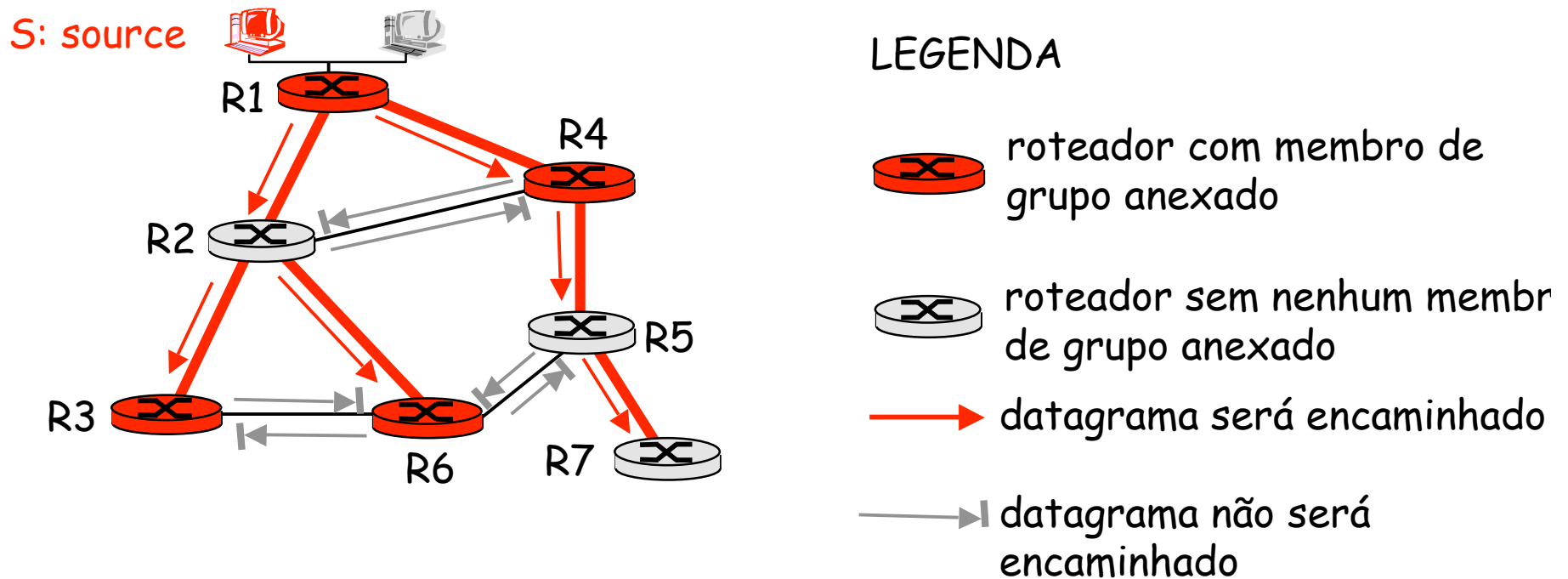
- Baseia-se no conhecimento dos roteadores sobre caminhos de unicast mais curtos dele até o transmissor
- Cada roteador possui comportamento de encaminhamento simples:

if (datagrama mcast recebido no link de entrada no menor caminho
retorna ao centro)

then dispara datagramas para todos os links de saída

else ignora datagrama

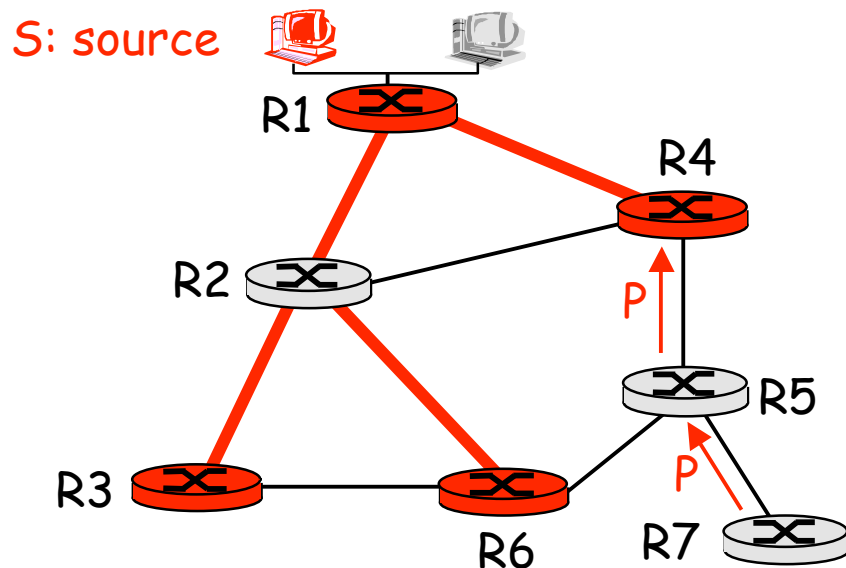
4 Reverse Path Forwarding: exemplo







- Resultado é um reverse SPT de origem específica.
- Pode ser uma má escolha com links assimétricos

4 Reverse Path Forwarding: pruning

- Árvores de encaminhamento contêm subárvores com membros de grupo sem multicast
 - Não necessita encaminhar datagramas por subárvores abaixo
 - Mensagens “prune” são enviadas por upstream pelo roteador com membros de grupo sem nenhum downstream



LEGENDA

-  roteador com membro de grupo anexado
-  roteador sem nenhum membro de grupo anexado
-  mensagem prune
-  links com encaminhamento multicast

4 Shared-Tree Steiner Tree

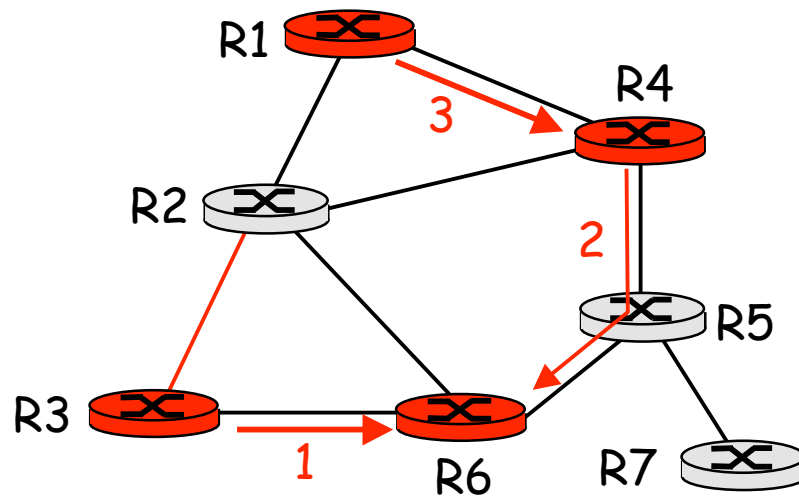
- **Steiner Tree:** árvore de custo mínimo conectando todos os roteadores com membros de grupo anexados
- Problema é NP-completo
- Existe uma heurística excelente
- Não é usado na prática:
 - Complexidade computacional
 - Informação sobre toda a rede é necessária
 - Monolítica: reexecuta sempre que um roteador precisa se juntar/deixar.

4 Center-based trees




- Única árvore de entrega compartilhada por todos
- Um roteador é identificado como “centro” da árvore para se juntar:
 - Roteador de borda envia uma **join-msg** unicast endereçada ao roteador de centro
 - **join-msg** “processada” pelos roteadores intermediários e encaminhada rumo ao centro
 - **join-msg** ou encontra um ramo da árvore para seu centro, ou chega até o centro
 - O caminho tomado pela **join-msg** torna-se um novo ramo da árvore para esse roteador

4 Center-based trees: um exemplo

Suponha que R6 escolheu como centro:



LEGENDA

-  roteador com membro de grupo anexado
-  roteador sem nenhum membro de grupo anexado
-  ordem de caminho onde são geradas mensagens join

4 Roteamento de multicast da Internet: DVMRP

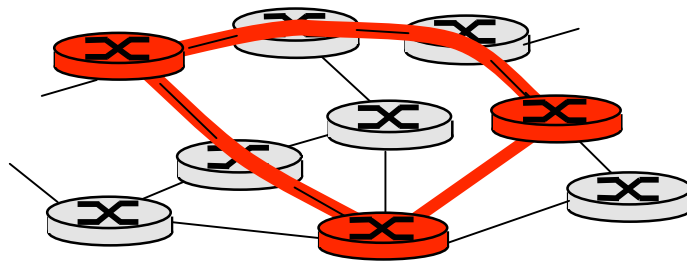
- **DVMRP**: distance vector multicast routing protocol, RFC1075
- **flood and prune**: reverse path forwarding, source-based tree
 - Árvore RPF baseada em tabelas de roteamento do DVMRP construídas pelos roteadores de comunicação do DVMRP
 - Sem suposições sobre unicast subjacente
 - Datagrama inicial para grupo de multicast disparado para todos os lados via RPF
 - Roteadores não querem grupo: enviam mensagens “prune” de upstream

4 DVMRP

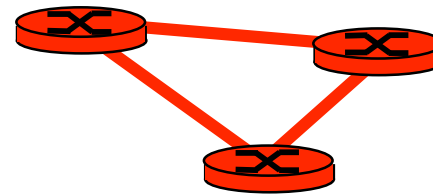
- **soft state:** o roteador DVMRP periodicamente (1 min) “esquece” os ramos que estão podados:
 - Dados mcast fluem novamente por ramos não podados
 - Roteador de downstream: poda novamente ou senão continua a receber dados
- Roteadores podem rapidamente se reintegrar à árvore
 - IGMP seguinte se junta na folha
- Probabilidades e extremidades
 - Comumente implementado em roteadores comerciais
 - Roteamento Mbone é feito usando DVMRP

4 Tunelamento

P.: Como conectar “ilhas” de roteadores multicast num “mar” de roteadores unicast?



topologia física



topologia lógica

- Datagrama mcast encapsulado dentro de um datagrama “normal” (sem endereço mcast)
- O datagrama IP normal é enviado pelo “túnel” via unicast IP regular para o roteador mcast receptor
- O roteador mcast receptor desencapsula para obter o datagrama mcast

4 PIM: Protocol Independent Multicast

- Não depende de nenhum algoritmo de roteamento unicast específico (trabalha com todos)
- Dois cenários diferentes de distribuição de multicast:

Denso

- Membros de grupo densamente empacotados, em proximidade.
- Largura de banda mais abundante

Esparso

- # redes com membros de grupo pequeno wrt # redes interconectadas
- Membros “extensamente dispersados”
- Largura de banda não abundante

4 Conseqüência de dicotomia esparsa-densa

Densa

- Sociedade de grupo de roteadores **assumida** até que os roteadores sejam explicitamente podados
- Construção dirigida ao dado (**data-driven**) na árvore multicast (ex., RPF)
- Largura de banda e processamento de roteador sem grupo **deliberados**

Esparsa

- Não há sociedade até que os roteadores se juntem explicitamente
- Construção dirigida ao receptor (**receiver-driven**) da árvore multicast (ex., center-based)
- Largura de banda e processamento de roteador sem grupo **conservativos**

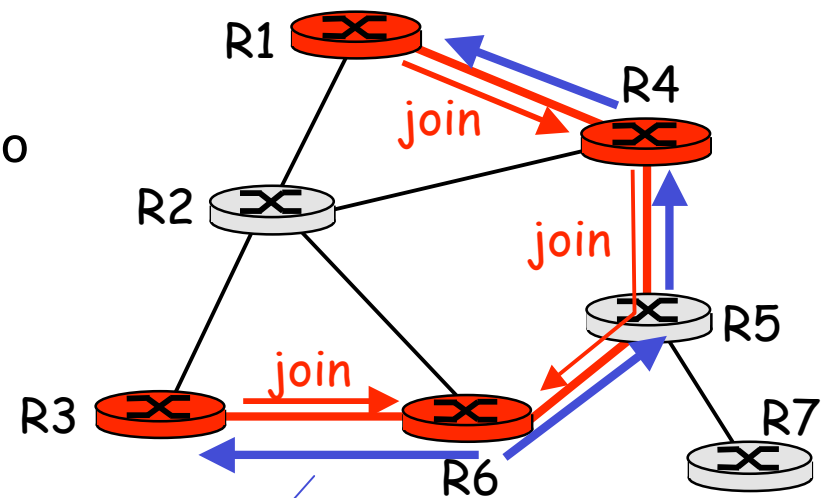
4 PIM-Modo denso

Flood and prune RPF, similar ao DVMRP mas

- Protocolo unicast em questão provê informações RPF para os datagramas que chegam
- Flood de downstream menos complicado (menos eficiente) do que o DVMRP reduz a confiança no algoritmo de roteamento em questão
- Possui mecanismo de protocolo para o roteador detectar que ele é um roteador **leaf-node**

4 PIM-Modo esparso

- Método **center-based**
- Roteador envia mensagens *join* para o ponto de encontro RP (**rendezvous point**)
 - Roteadores intermediários atualizam o estado e encaminham **join**
- Após juntar via RP, o roteador pode mudar para uma árvore de origem específica
 - Aumento de desempenho: menos concentração, caminhos menores



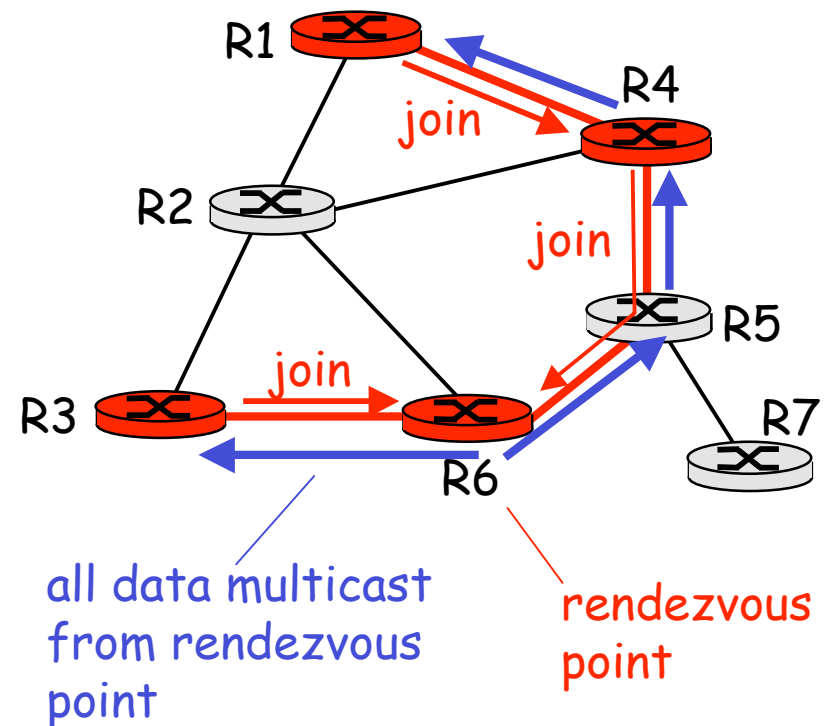
o multicast de todos os dados é feito a partir do ponto de encontro

ponto de encontro

4 PIM-Modo esparso

Transmissor(es):

- Unicast de dados para o RP, que distribui pela árvore **RP-roteada**
- RP pode estender o upstream da árvore multicast para a origem
- RP pode enviar a mensagem **stop** se não houver receptores anexados
 - “Ninguém está escutando!”



4 Camada de rede: resumo

O que foi coberto:

- Serviços da camada de rede
- Princípios de roteamento: link state e distance vector
- Roteamento hierárquico
- IP
- Protocolos de roteamento da Internet: RIP, OSPF, BGP
- O que há dentro de um roteador?
- IPv6

Próxima etapa: a camada de enlace de dados!