

Aula 19

Camada de Rede

Comunicação multidestinatória

Igor Monteiro Moraes

Redes de Computadores I – TCC00185

ATENÇÃO!

- Este apresentação é contém partes baseadas nos seguintes trabalhos
 - Notas de aula do Prof. Luís Henrique M. K. Costa, disponíveis em <http://www.gta.ufrj.br/ensino/CPE825/cpe825.html>
 - Notas de aula do Prof. José Augusto Suruagy Monteiro, disponíveis em <http://www.nuperc.unifacs.br/Members/jose.suruagy/cursos>
 - Material complementar do livro Computer Networking: A Top Down Approach, 5th edition, Jim Kurose and Keith Ross, Addison-Wesley, abril de 2009

Comunicação Multidestinatória

- Aplicações
 - Videoconferência
 - Ensino à distância
 - Jogos distribuídos
 - TV na Internet
 - Etc.

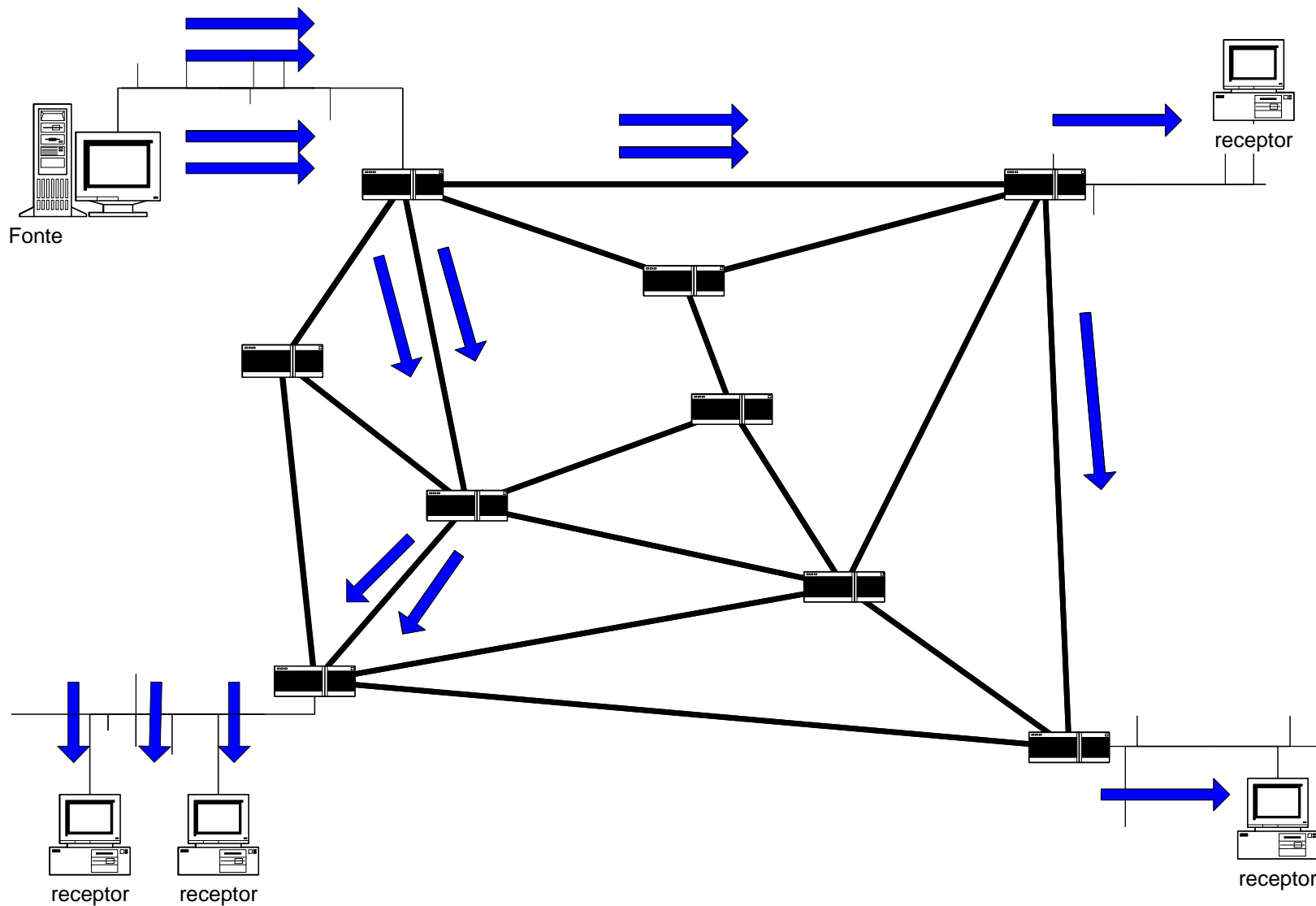
Comunicação Multidestinatória

- Aplicações
 - Videoconferência
 - Ensino à distância
 - Jogos distribuídos
 - TV na Internet
 - Etc.
- Característica comum
 - A mesma informação deve ser enviada a **múltiplos receptores**

Como enviar a N receptores?

- Usar diferentes tipos de transmissão
- Unicast
 - Transmissão ponto-a-ponto
 - 1 emissor, 1 receptor
- Broadcast
 - Envio a todos os nós da rede
- Multicast
 - Transmissão ponto-a-multiponto
 - 1 emissor, N receptores

Unicast x Multicast



Utilização do Multicast

- Vantagens
 - Produz menos pacotes
 - Utilização eficiente da banda passante da rede
 - Menor processamento em estações e roteadores

Utilização do Multicast

- Problemas
 - Como **identificar o grupo**?
 - Lista dos receptores
 - Sobrecarga de cabeçalho limita o tamanho do grupo
 - Endereço de grupo
 - Identidade e número dos receptores desconhecidos
 - Como **distribuir os pacotes**?
 - Endereçamento e roteamento (encaminhamento dos pacotes) são **diretamente** relacionados

Utilização do Multicast

- Problemas

- Como **identificar o grupo?**

- Lista dos receptores

- Sobrecarga de cabeçalho limitado

- Endereço de grupo

- Id

- C

-

- reencaminhamento dos pacotes)

- relacionados

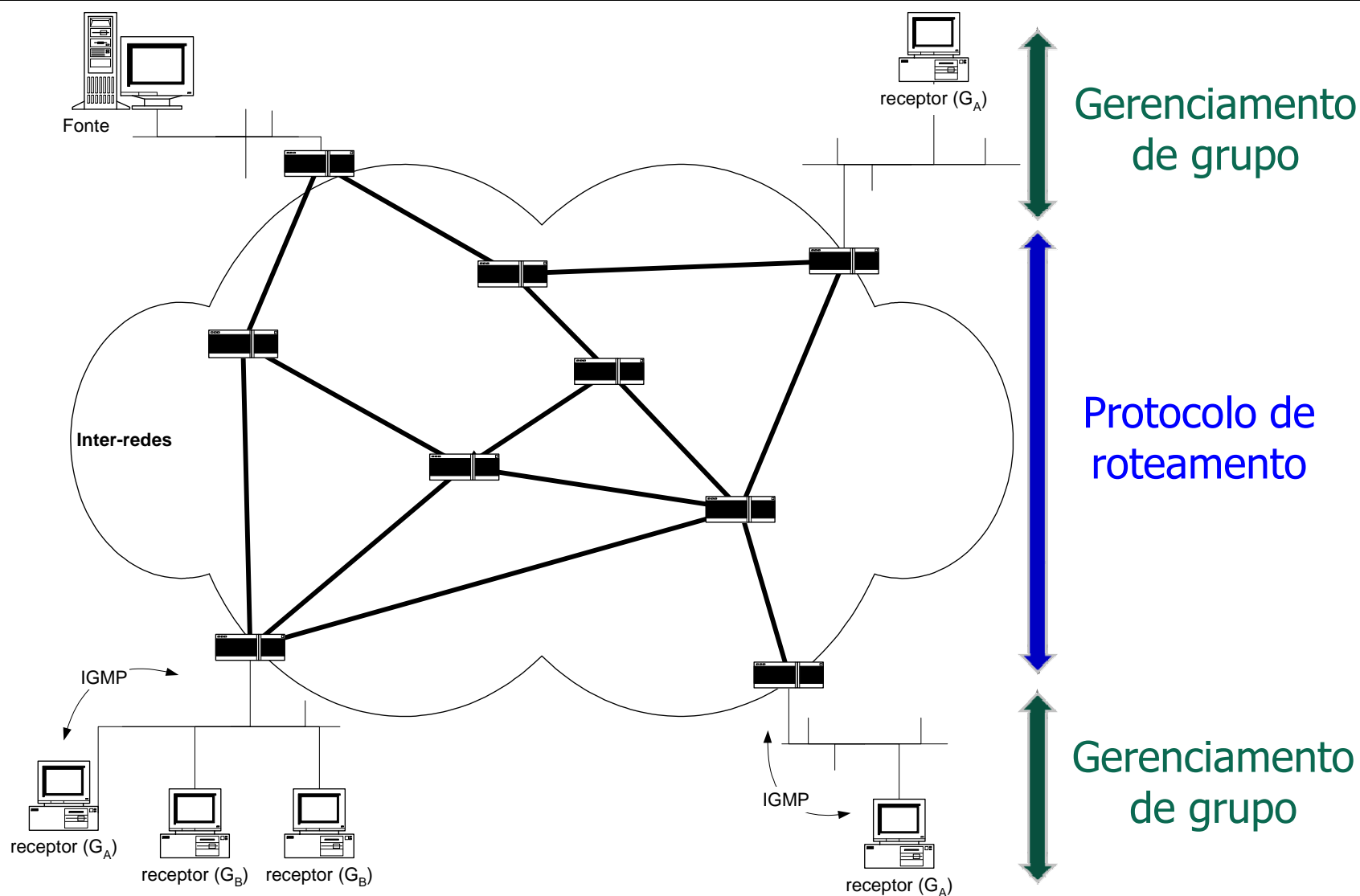
Gerenciamento de grupo

Roteamento multicast

Modelo de Serviço IP Multicast

- Proposto em 1989
- Identificação
 - Endereço de grupo
- Distribuição dos dados
 - **Gerenciamento de grupo**
 - Entrada e saída do grupo
 - “quero escutar o grupo” / “quero parar de escutar o grupo”
 - Entre a estação e seu roteador local
 - **Protocolos de roteamento**
 - Distribuição dos dados entre as redes
 - Como fazer os pacotes chegarem ao meu roteador local?

Modelo de Serviço IP Multicast



Modelo de Serviço IP Multicast

- Grupo
 - Identificado por um **endereço de grupo**
 - Conversação **N x M aberta**
 - Qualquer estação pode participar
 - Uma estação pode pertencer a vários grupos
 - Uma fonte pode enviar dados ao grupo, tendo se inscrito neste ou não
 - **É dinâmico**
 - Uma estação pode entrar e sair a qualquer instante
 - O **número e a identidade** dos participantes do grupo são **desconhecidos**

Gerenciamento de Grupo

Criação e Destruição do Grupo

- O grupo é identificado por um endereço IP Multicast
 - Endereço IP Classe D
- Criação do grupo
 - Escolha de um **endereço multicast** e envio de dados para o grupo
- Destruição do grupo
 - Parada do envio de dados

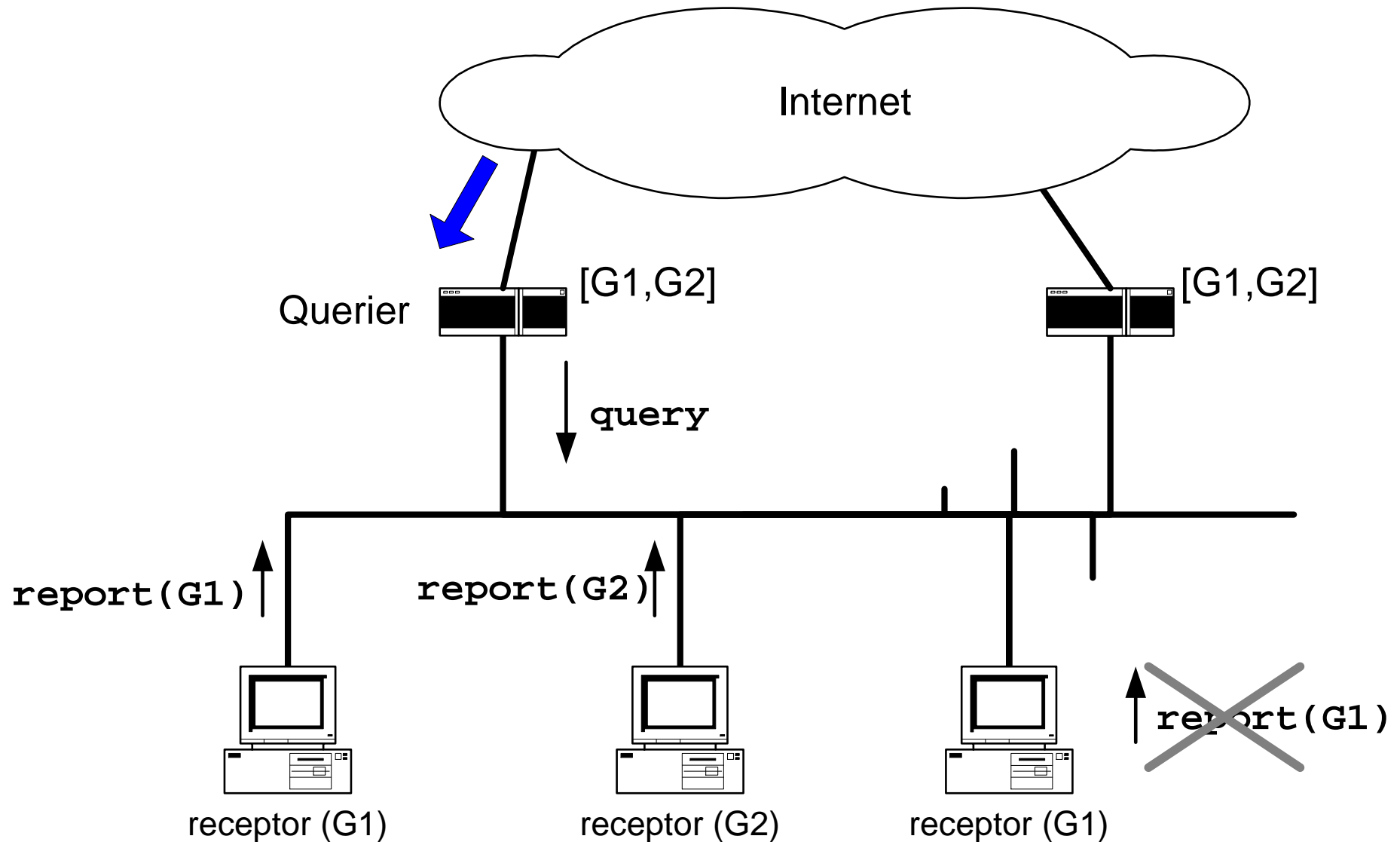
Gerenciamento de Grupo

- Quem quer ouvir que grupos?
 - “Estação de rádio”
- IGMP (*Internet Group Management Protocol*)
 - **Detecção** de estações interessadas em grupos multicast
 - Existem 4 versões do IGMP
- Escopo local
 - Diálogo entre a **estação e o primeiro roteador**
 - Criação da árvore de distribuição independente do IGMP

Funcionamento do IGMP

- Parte estação
 - Conexão ao grupo (`join(G)`)
 - Receptor envia mensagem `report(G)`
 - Envio de mensagens `report` em resposta às mensagens `query`
 - “Estes são os grupos de interesse desta estação”
- Parte roteador
 - Envio periódico de mensagens `query`
 - “Que grupos são escutados na rede?”
- Parte estação
 - Mecanismo de supressão de mensagens `report`

Funcionamento do IGMP



- Introduz o mecanismo de *fast-leave*
 - Diminuição da latência de desconexão
- Desconexão
 - Receptor envia mensagem IGMP `leave(G)`
- Regras de processamento para evitar a desconexão de outras estações
 - Ex. roteador deve enviar `query(G)` para detectar se existem ouvintes remanescentes

- Filtragem de fontes
- A estação anuncia o interesse no grupo G
 - “apenas nos dados enviados por determinadas fontes”,
ou
 - “nos dados enviados por todas, exceto determinadas fontes”
- Interface
 - `IPMulticastListen (socket, interface, mcast-address, filter mode, source-list)`
 - `filter-mode` **pode ser INCLUDE ou EXCLUDE**

Exemplo no IGMPv3

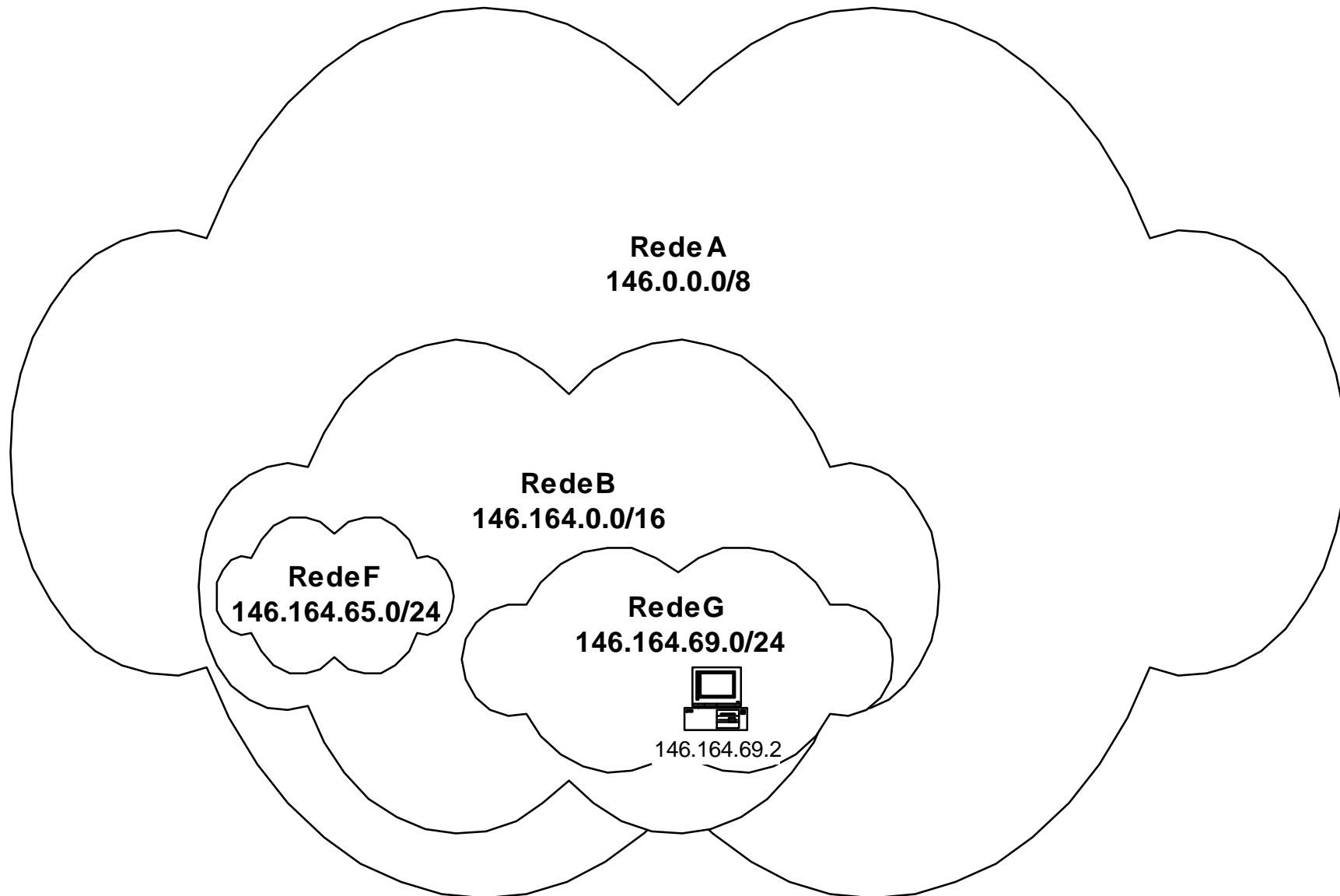
- Recepção do que apenas as fontes S1 e S2 enviam a G
 - `IPMulticastListen (sock, iface, G, INCLUDE, {S1,S2})`
- Recepção de tudo que é enviado a G, exceto por S2 e S3
 - `IPMulticastListen (sock, iface, G, EXCLUDE, {S2,S3})`
- Estado no roteador
 - `(G, EXCLUDE{S3})`

Roteamento Multicast

Endereçamento na Internet

- Endereçamento IP
 - Topológico (ou **hierárquico**: utiliza **prefixos**)
 - A posição de uma máquina determina seu endereço
 - Torna eficaz as operações de roteamento

Endereçamento Hierárquico



Problema do Multicast

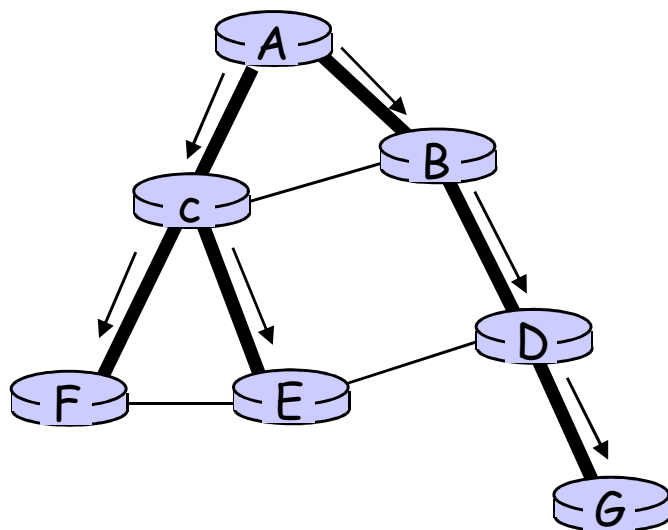
- Dado o endereçamento, como realizar a distribuição dos pacotes?
 - Endereço unicast
 - Identifica e localiza uma estação
 - Endereço de grupo
 - Hierarquia **impossível**
 - Receptores espalhados em toda a rede

Roteamento Multicast

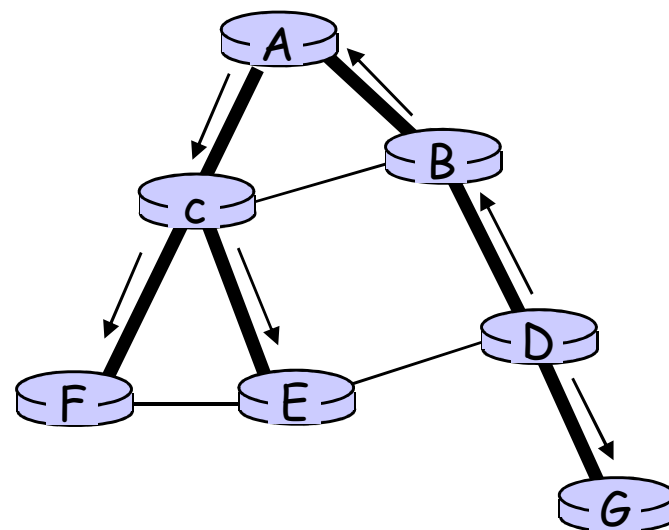
- Problema de Roteamento Multicast
- $G = (V, E)$
 - V conjunto de vértices
 - E conjunto de enlaces
- M subconjunto de V
 - Inclui fontes e receptores do grupo multicast
- **Problema:** construir uma, ou várias, topologias de interconexão, árvores, que incluem todos os nós em M
 - Árvore por fonte (*source-based tree*)
 - Árvore compartilhada (*shared tree*)

Roteamento Multicast

- **Árvore geradora**
 - Nenhum pacote redundante recebido por nenhum nó
 - Nós encaminham cópias somente ao longo da árvore geradora



(a) **Broadcast** iniciado em A



(b) **Broadcast** iniciado em D

Roteamento Multicast

- Meta
 - Achar uma árvore (ou árvores) conectando todos os roteadores com membros locais do grupo multicast
- Árvore
 - Nem todos os caminhos entre roteadores são usados
 - Baseada na fonte
 - Árvore distinta de cada fonte para receptores
 - Compartilhada
 - Mesma árvore usada por todos os membros do grupo

Algumas Soluções

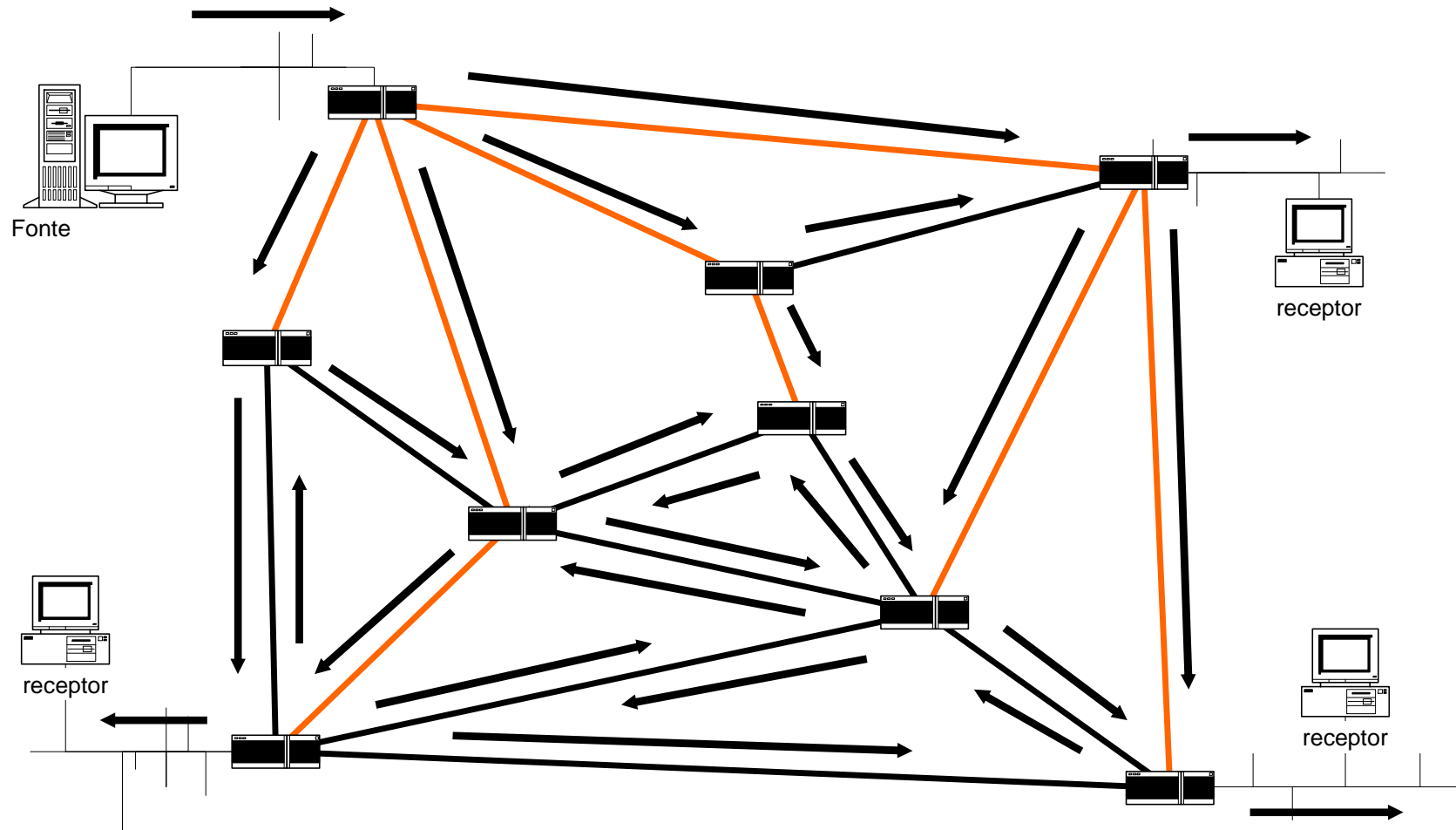
- Árvores de cobertura (*spanning trees*)
- Algoritmo de inundação
- Árvores RPF (*Reverse Path Forwarding*)
- Árvores centradas

Algumas Soluções

- Árvores de cobertura (*spanning trees*)
- Algoritmo de inundação
- **Árvores RPF (*Reverse Path Forwarding*)**
- **Árvores centradas**

- Hipótese
 - Um roteador R conhece o caminho mais curto para ir à fonte, S
 - *Reverse Path Forwarding check (RPF check)*
- *Reverse Path Broadcasting*
 - O roteador R recebe um pacote da fonte S
 - O pacote chegou pela interface utilizada por R para ir à S ? (RPF check)
 - Se sim, enviar o pacote por todas as interfaces de saída
 - Se não, descartar o pacote

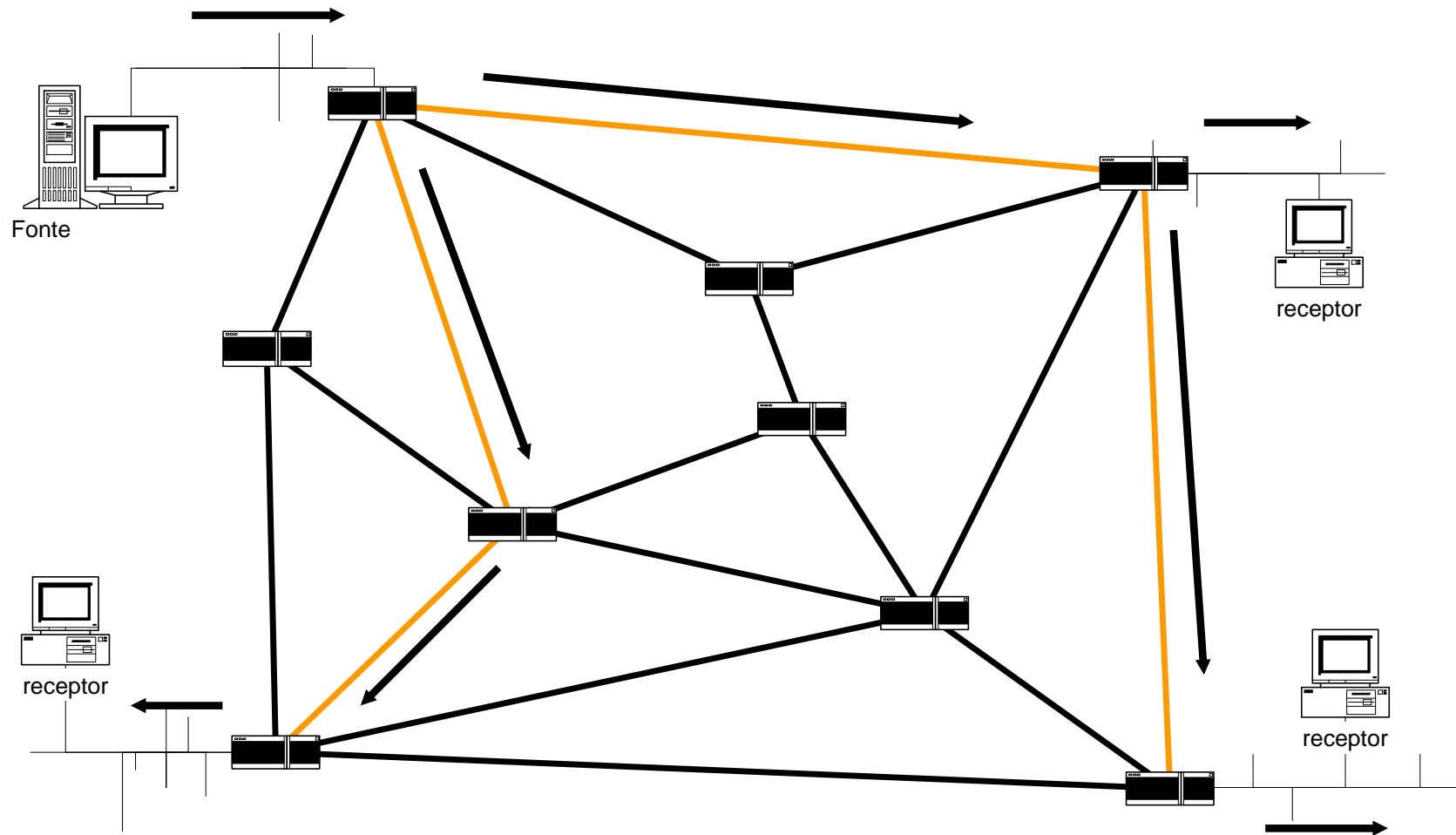
Reverse Path Broadcasting



Reverse Path Forwarding

- Nem todas as sub-árvores possuem membros do grupo
 - Ramos podem ser podados
- Hipótese
 - Um roteador R sabe se seu vizinho o utiliza como caminho para a fonte, s
- Como obter esta informação?
 - Trivial, se protocolo de estado do enlace
 - Se protocolo de vetor de distância
 - Mensagem adicional para alertar o roteador “pai”, ou
 - Mensagem de poda para eliminar a rota reversamente
- Informação por (fonte, grupo)

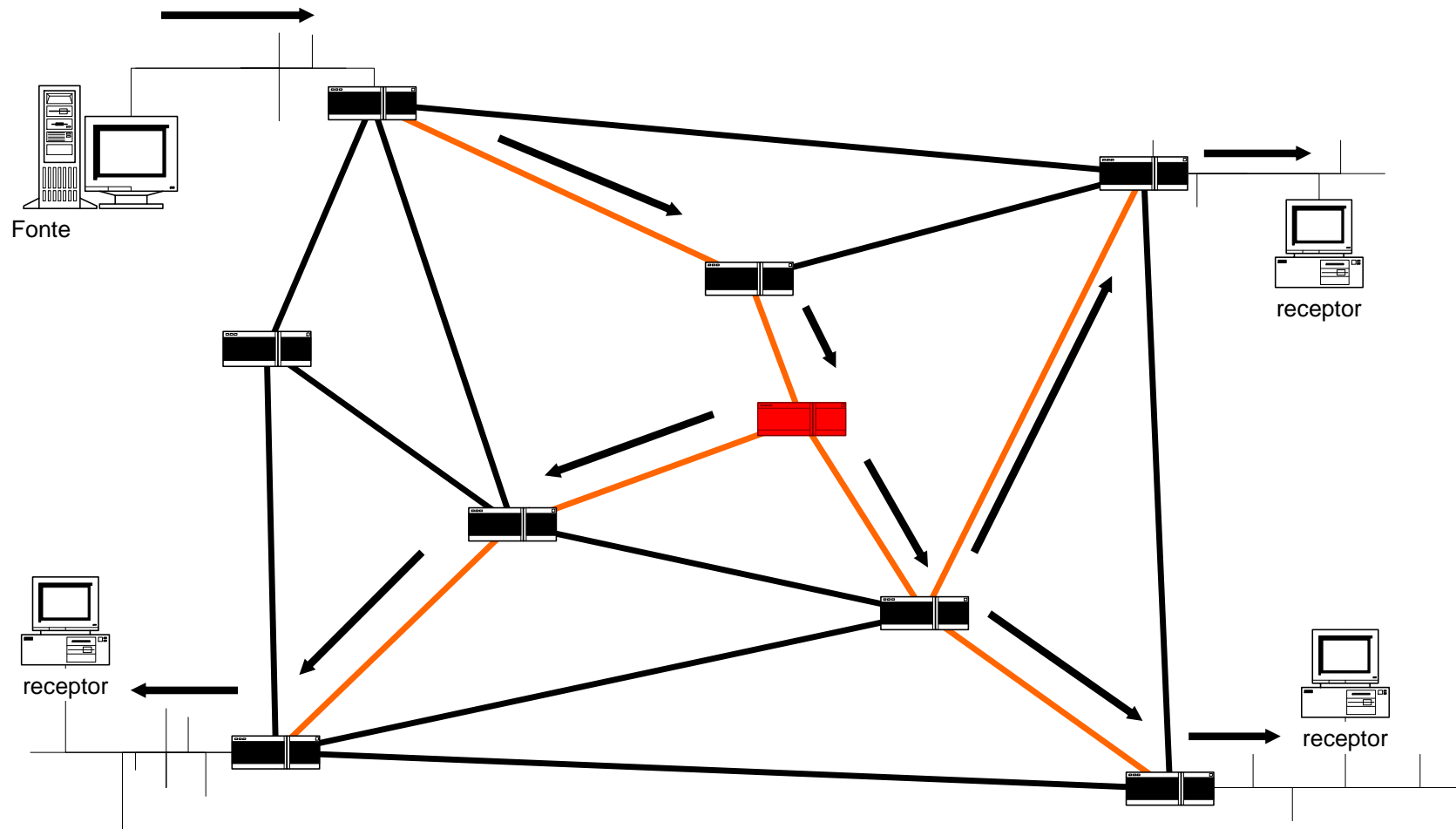
Árvore RPF



Árvores Centradas

- Construída a partir de um nó central (*core*)
- **Compartilhada** por diversas fontes
 - Diversas fontes utilizam o mesmo nó central
 - “Pedidos de conexão” são enviados ao nó central

Árvores Centradas

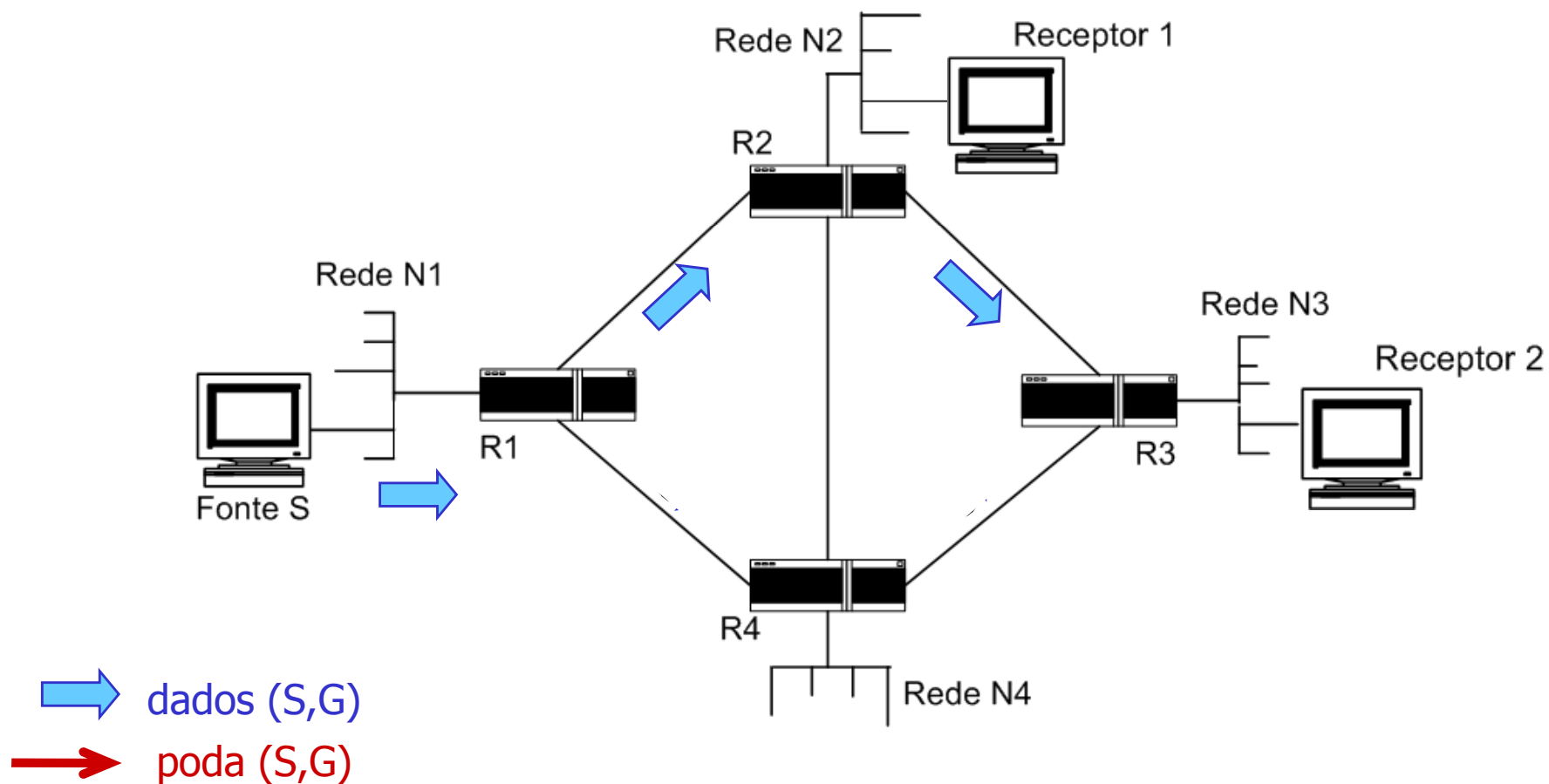


Protocolos de Roteamento

- Árvore de distribuição *multicast*
 - Por Fonte
 - Compartilhadas
 - Nó central
- Mecanismos de construção
 - Inundação-e-poda
 - Inscrição explícita

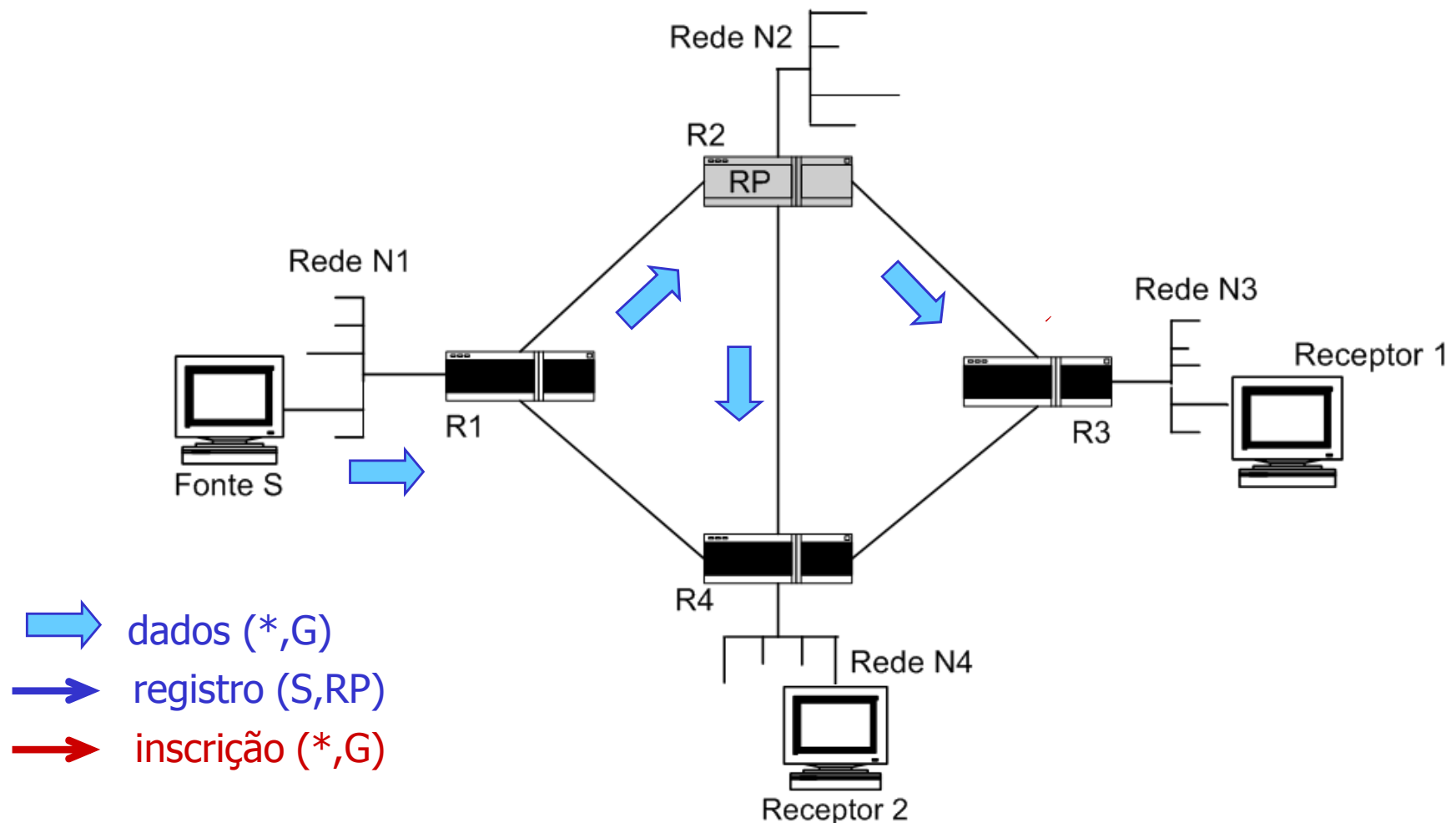
Protocolos de Roteamento

- Inundação-e-poda



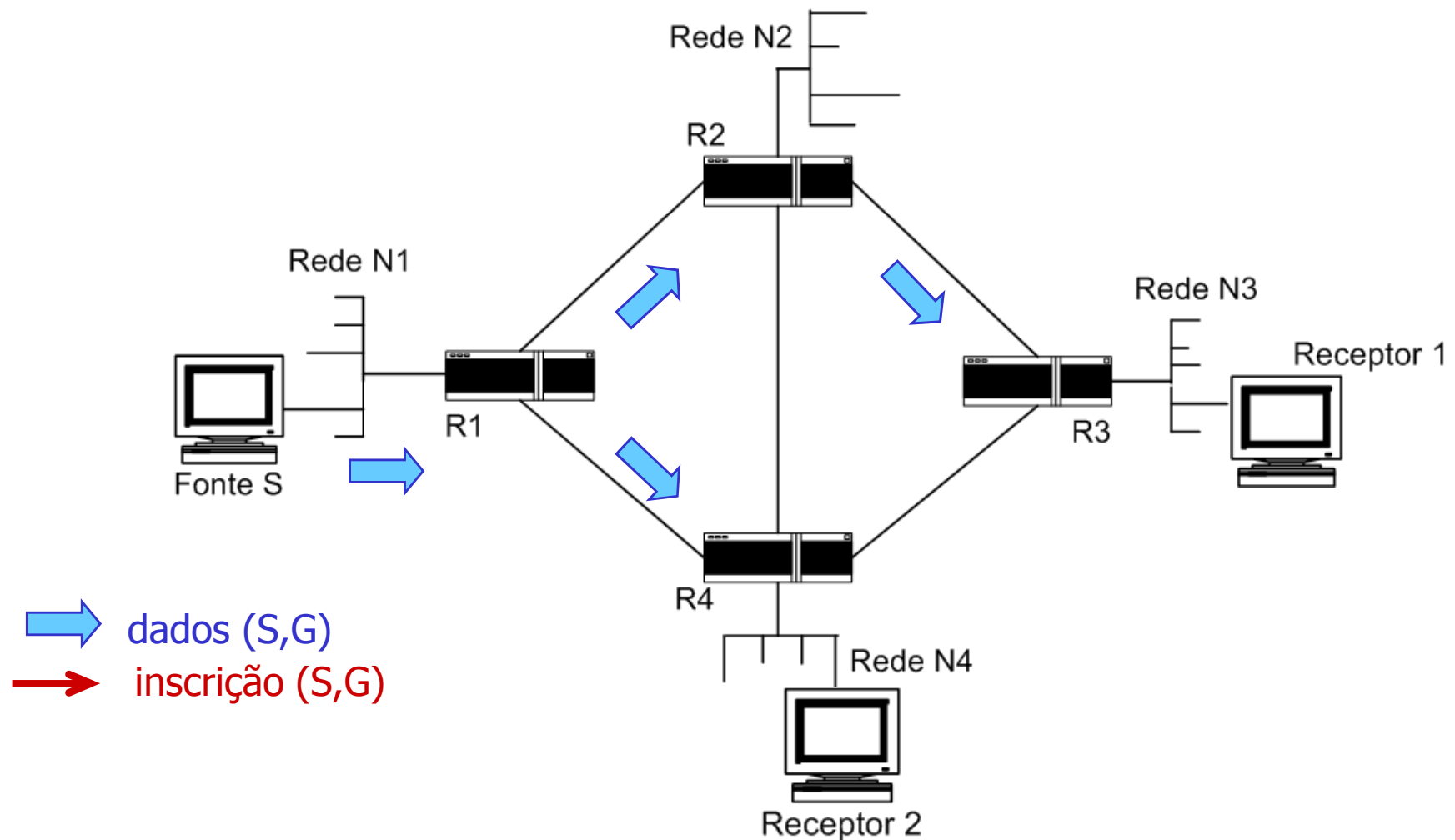
Protocolos de Roteamento

- Inscrição explícita – ponto central



Protocolos de Roteamento

- Inscrição explícita – fonte

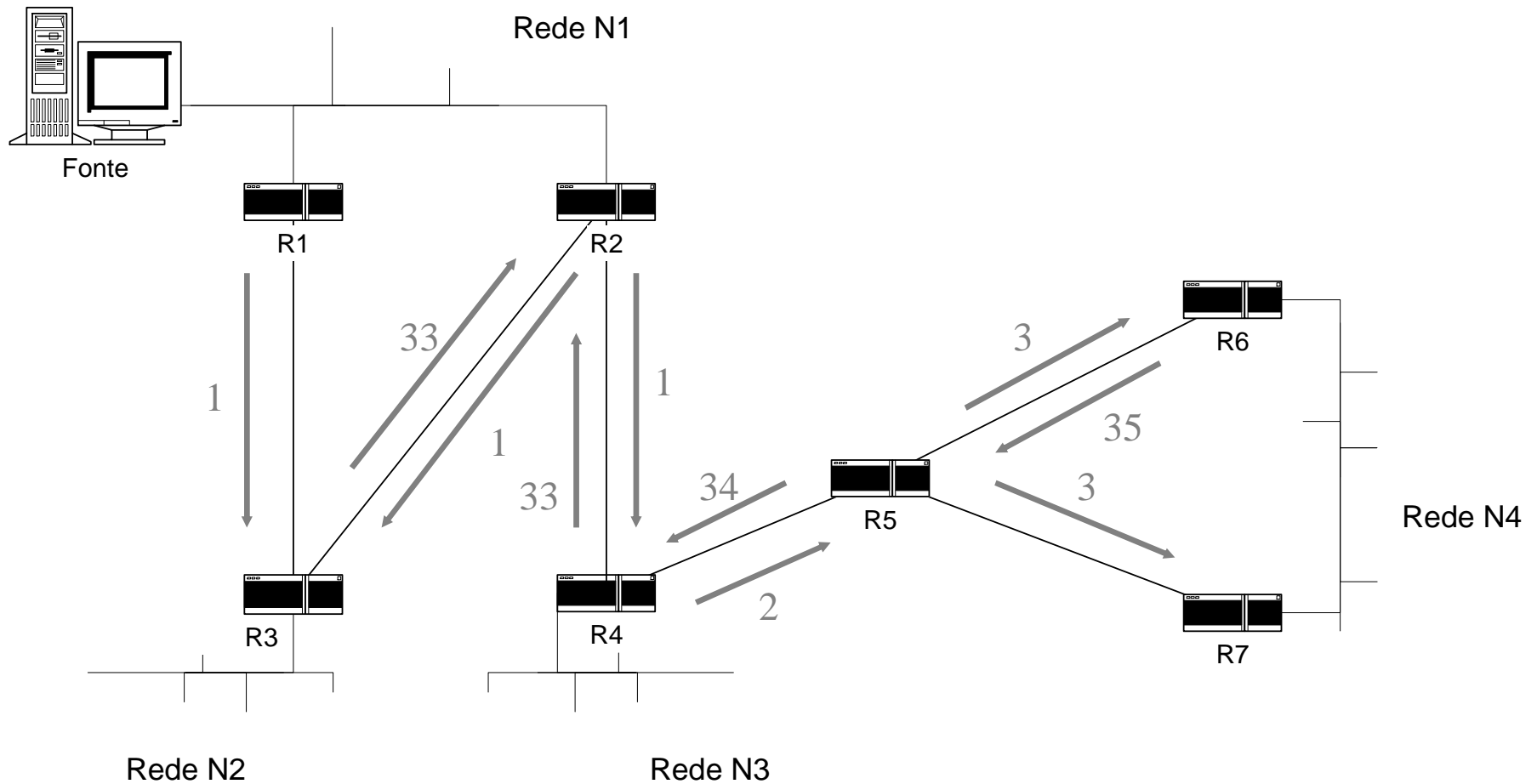


Roteamento Multicast Intradomínio

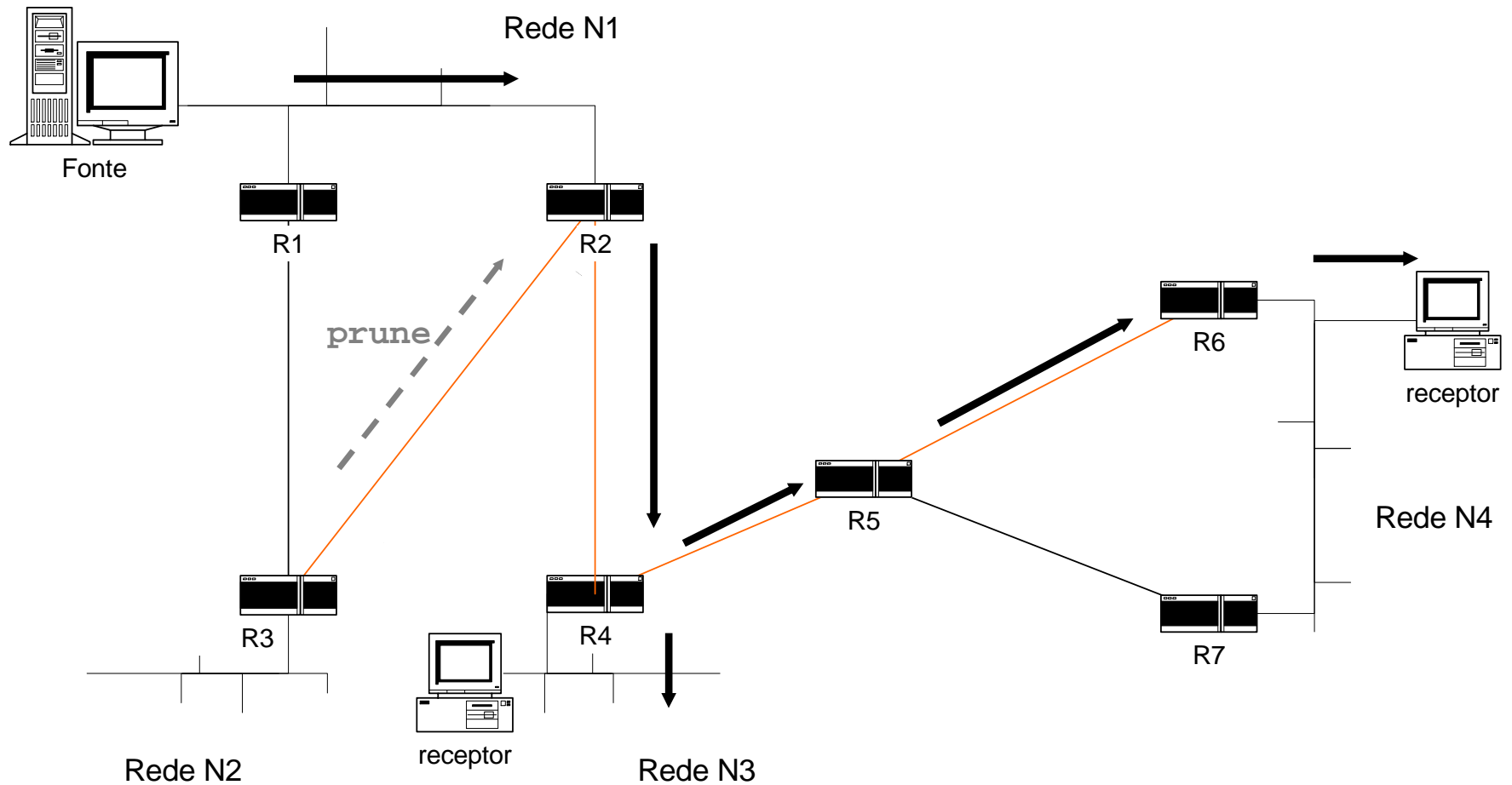
- DVMRP (*Distance Vector Multicast Routing Protocol*)
 - Primeiro protocolo utilizado no MBone
 - Backbone experimental para tráfego multicast
 - Início dos anos 90
- MOSPF (*Multicast Open Shortest Path First*)
- CBT (*Core Based Trees*)
- PIM (*Protocol Independent Multicast*)
 - PIM-DM (*PIM Dense-Mode*)
 - PIM-SM (*PIM Sparse Mode*)

- Utiliza vetores de distância
 - Semelhante ao RIP (*Route Information Protocol*)
 - Constrói rotas unicast para cada fonte multicast
 - *Poison-reverse* especial utilizado para marcar interfaces filhas
- Distribuição de dados
 - Inundação e poda (*flood-and-prune*)
 - Teste RPF baseado em sua tabela de roteamento unicast
- A inundação é periódica
 - Descoberta de fontes ativas

Funcionamento do DVMRP



Envio de Dados no DVMRP

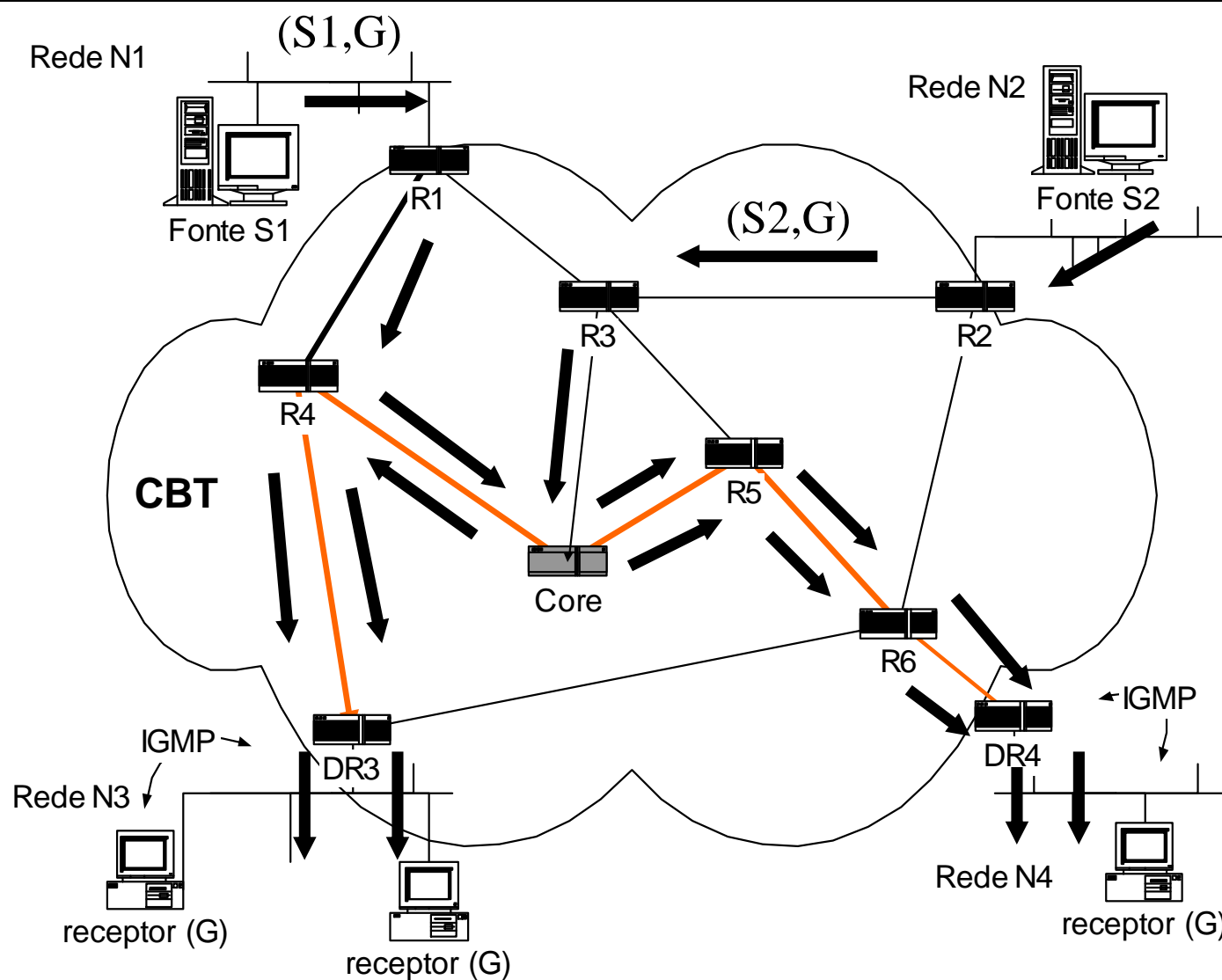


- Algoritmo simples
- Protocolo de roteamento unicast *próprio*
- Inundação periódica da rede com *dados*
- Vetores de distância
 - Convergência lenta, como no RIP

- Utiliza árvores centradas
 - Compartilhadas e bidirecionais
- Roteador central – *core*

- Construção da árvore
 - Mensagens *join*
 - Enviadas pelos receptores na direção do *core*

Envio de Dados no CBT



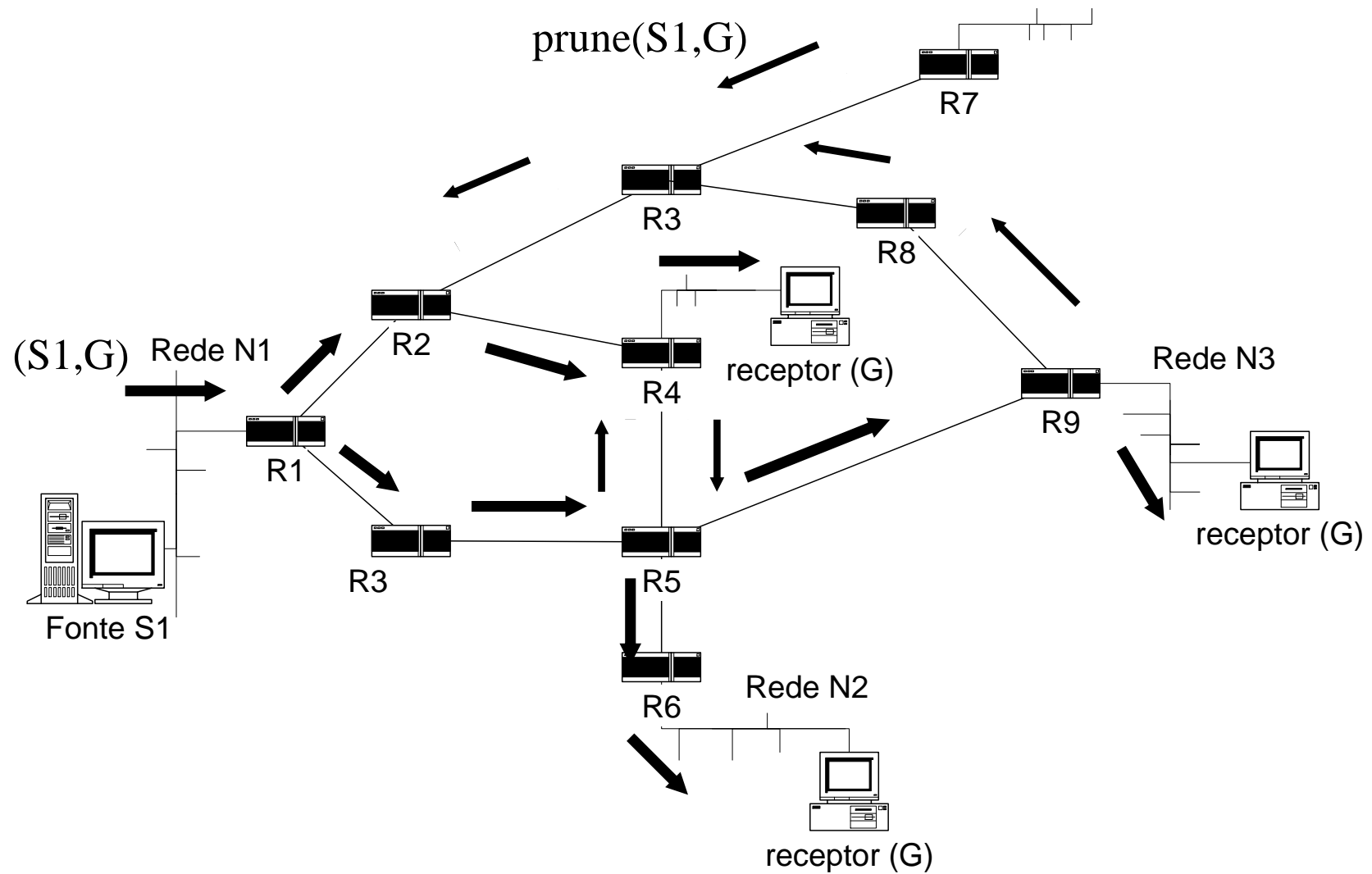
- Escalabilidade
 - Estado apenas nos roteadores na árvore de distribuição
 - Ao contrário de DVMRP e MOSPF
 - Estado por (grupo), em vez de por (fonte, grupo)
- Desvantagens
 - Concentração de tráfego próximo ao *core*
 - Rotas sub-ótimas entre a fonte e o receptor
 - Maiores atrasos
- Localização do *core* é crítica

- *Protocol Independent Multicast* (PIM)
 - Independente do protocolo de roteamento unicast
- *Dense-Mode* (PIM-DM)
 - Receptores densamente distribuídos
 - Árvores por fonte
 - Inundação-e-poda (semelhante ao DVMRP)
- *Sparse-Mode* (PIM-SM)
 - Receptores esparsamente distribuídos na rede
 - Árvores compartilhadas (como o CBT)
 - Unidirecionais

- *Reverse Path Multicast*
 - Utiliza o teste RPF
 - Mas não constrói lista de interfaces filhas como o DVMRP
 - Tráfego enviado em todas as interfaces de saída
 - Duplicação de pacotes, todos os enlaces da rede são utilizados, mas
 - Independência do roteamento unicast
 - Evita base de dados com pais/filhos

- *Reverse Path Multicast*
 - Após a inundação inicial, mensagens de poda são enviadas
 - Por roteadores que não possuem receptores do grupo
 - Por roteadores que não possuem vizinhos interessados no grupo
 - Por roteadores que receberam tráfego por uma interface incorreta (RPF)

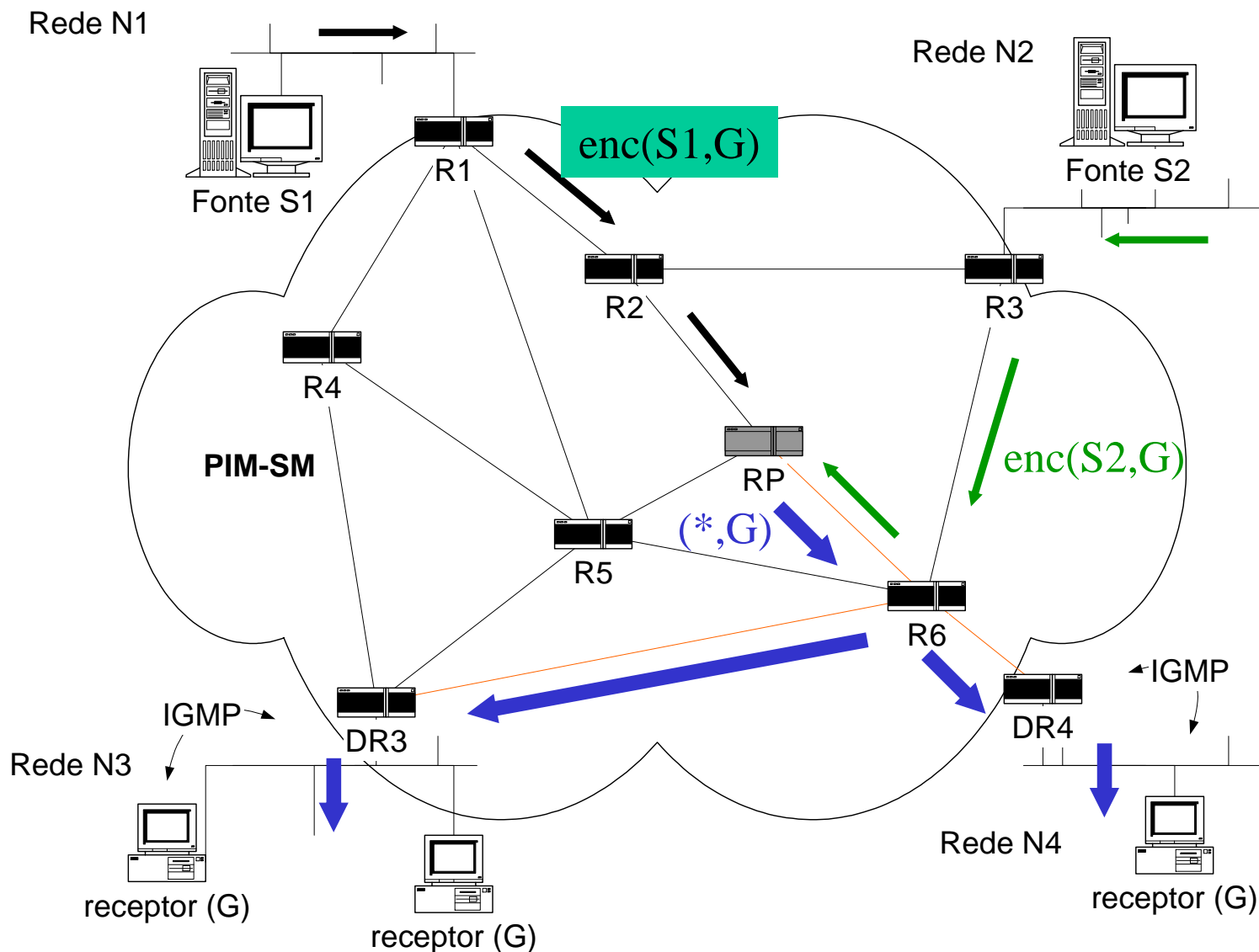
PIM-DM



- Árvore SPT reversa (RSPT)
 - União dos caminhos mais curtos dos receptores até a fonte
- Todos os roteadores da rede armazenam estado (fonte, grupo) para todas as fontes/grupos ativos
- Inundação periódica é necessária
 - Descoberta de novos membros do grupo

- Árvores de distribuição centradas ($(*, G)$, como o CBT)
 - Nó central – roteador RP (*rendez-vous point*)
 - Unidirecional
- Construção da árvore
 - Mensagens *join*
- Mecanismo de mapeamento entre grupos e RPs
- Fontes se “registram” com o RP
 - Dados são enviados ao RP (encapsulados em mensagens **PIM-register**)

Árvore Compartilhada no PIM-SM



- Árvores por fonte (S, G)
- Troca entre tipos de árvore realizada por configuração
 - Taxa de envio de dados
- Roteador local envia mensagens $join(S, G)$
 - Mas não pára o envio de $join(*, G)$
 - Tráfego de outras fontes deve continuar
 - Envia mensagem de poda especial ($RP-bit-prune(S, G)$)
 - Evita a recepção de dados de s em duplicata

- RP também pode enviar `join(S,G)`
- Possibilidade de árvores por fonte
 - Diminui a importância da localização do RP
 - Reduz o atraso fonte-receptores

Problemas

Outros Problemas do Modelo de Serviço

- Como limitar o alcance (ou escopo) do tráfego multicast?
 - Até onde vai o tráfego enviado por uma fonte?
 - Receptores **não** são conhecidos
- Como evitar a colisão de endereços?
 - Duas aplicações escolhem o mesmo endereço multicast

Alcance do Tráfego Multicast

- É necessário definir **escopos**
- Escopos
 - Por endereço
 - Utilizando o campo TTL
 - Administrativos

Roteamento Interdomínio

- Nem todos os roteadores são multicast
- Diferentes protocolos nos diferentes domínios
- Problemas com o PIM-SM
 - Mecanismo escalável de mapeamento entre RPs e grupos
 - Interdependência entre provedores de serviço introduzida pelos RPs

Interdomínio: Estado Atual

- *Border Gateway Multicast Protocol* (BGMP) [RFC 3913]
- Projeto semelhante ao BGP
 - “Anuncio as rotas que me interessam anunciar”
 - “Sou a raiz dos grupos que me pertencem”

Border Gateway Multicast Protocol

- Árvores compartilhadas bidirecionais
 - Podem ser construídos ramos por fonte
- A raiz da árvore é um Sistema Autônomo (AS)
 - Maior estabilidade e tolerância a falhas
 - ASs devem ser associados a endereços de grupo multicast
- A raiz da árvore do grupo G é o AS ao qual G está associado
 - Probabilidade de este AS possuir receptores de G

Novas Propostas

- Modelo de Serviço IP Multicast
 - Endereço IP class-D = grupo de estações
 - Qualquer estação pode se inscrever no grupo
 - Qualquer estação pode enviar dados para o grupo
 - Alocação de endereços multicast é problemática
 - Protocolos: IGMP + protocolos de roteamento
- IP Multicast não foi implantado na Internet

Source Specific Multicast

- SSM (*Source-Specific Multicast*)
 - Conversação 1 x N
 - *Subscribe channel* $\langle S, G \rangle$
 - Fornece base para o controle de acesso
 - Apenas S pode enviar para (S, G) , outras fontes são bloqueadas
 - Alocação de endereços multicast (G)
 - Problema local à fonte
 - Roteadores RP e o protocolo MSDP não são necessários

Componentes do Serviço SSM

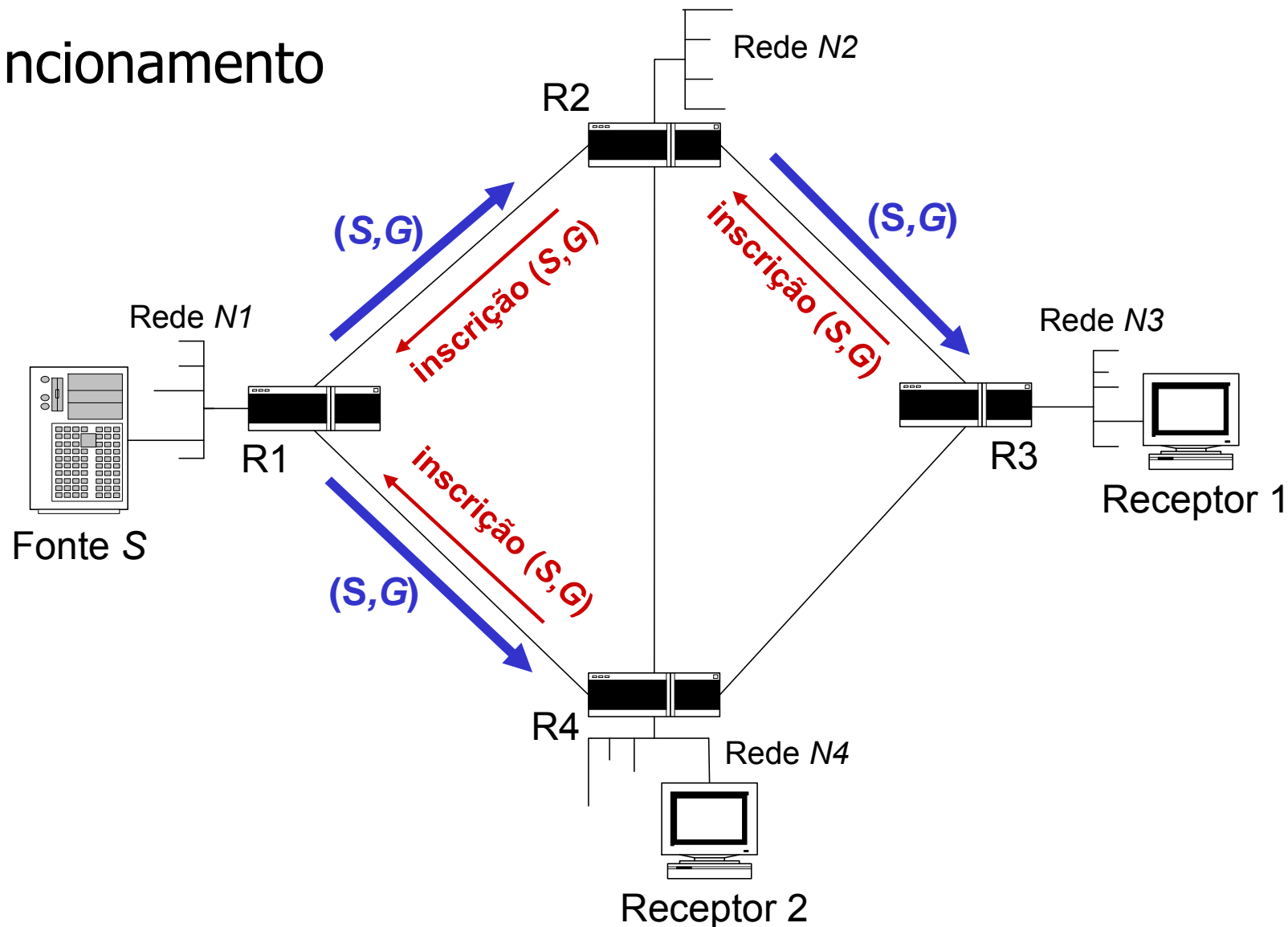
- Faixa de endereços exclusiva - 232/8 (IANA)
- Roteamento: PIM-SSM
 - Versão modificada do PIM-SM
 - Pode implementar ambos os serviços (SM & SSM)
- IGMPv3 (MLDv2 no IPv6)
 - Suporta a filtragem de fontes
 - (INCLUDE, EXCLUDE)

Funcionamento do PIM-SSM

- Regras do PIM-SSM
 - somente $join(S,G)$ é permitido na faixa 232/8
 - $join(*,G)$ e $join(S,G)$ permitidos na faixa restante
 - roteadores de borda (DR no PIM)
 - implementam $join(S,G)$ imediato
 - roteadores de núcleo
 - devem evitar as árvores compartilhadas em 232/8

PIM-SSM

- Funcionamento



Benefícios do SSM

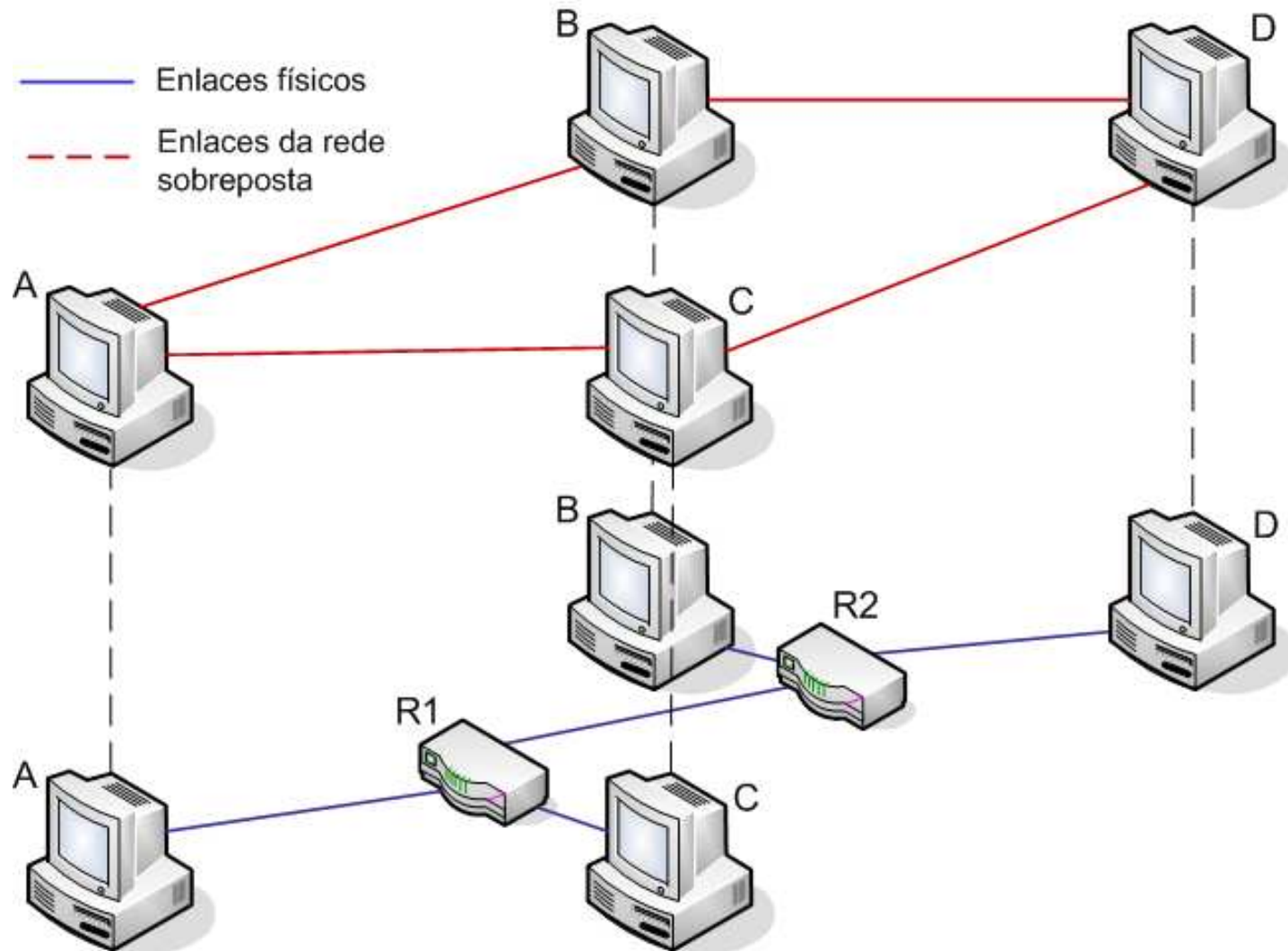
- Evita a interferência entre aplicações
 - Controle local
- Controle de acesso
 - Um receptor só recebe dados de uma fonte
 - Apenas uma fonte transmite para um canal
- Autenticação e contagem dos participantes
 - Pedidos de assinatura enviados à fonte
- Árvores por fonte
 - Distribuição de vídeo
 - Menor atraso entre a fonte e os assinantes

- IETF
 - ASM (Any Source Multicast)
 - conversação M x N
 - *Join **group** G*
 - SSM (Source-Specific Multicast)
 - conversação 1 x N
 - *Subscribe **channel** <S,G>*
- Protocolo PIM-SSM [RFC 3569]
 - Modificação do PIM-SM

Multicast Aplicativo

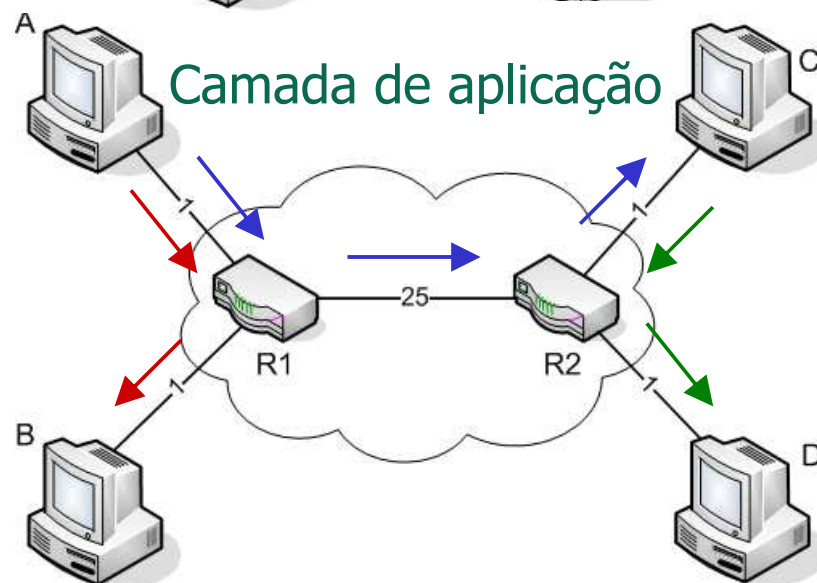
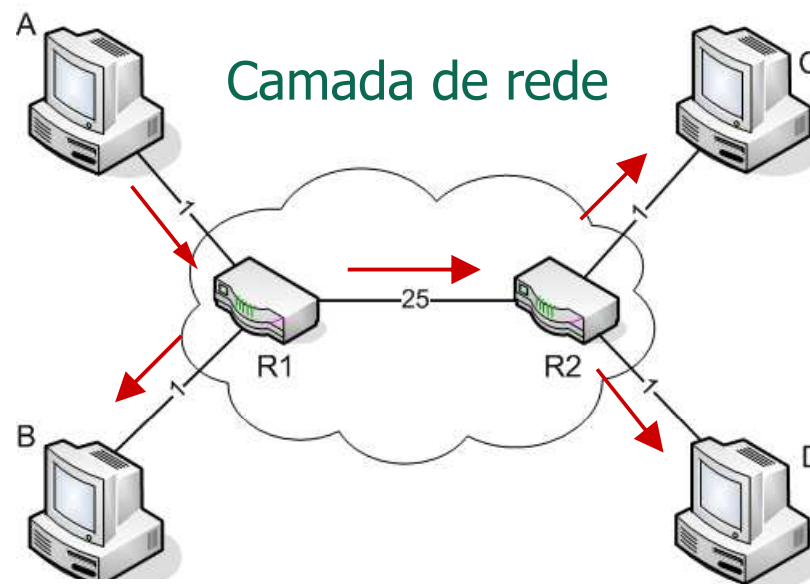
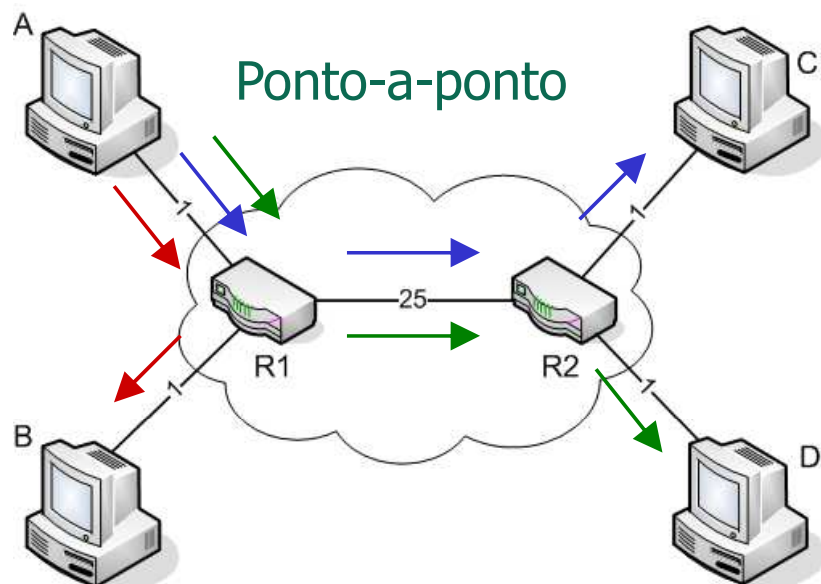
- Comunicação multidestinatória na camada de aplicação
- Os participantes são responsáveis pelo
 - Roteamento
 - Gerenciamento de participantes
 - Replicação da informação
- Existência apenas da comunicação ponto-a-ponto
- Construção de uma rede sobreposta (*overlay*)
 - Enlaces são túneis ponto-a-ponto
 - Não é necessário modificar a infra-estrutura de rede

Multicast Aplicativo



- Duas formas para construir a rede sobreposta
 - Conjunto de servidores e *proxies*
 - Nós âncoras
 - Sistemas par-a-par
 - Não há nós dedicados
 - Compartilhamento de recursos x utilização do serviço
- Desvantagem
 - Desempenho em relação a comunicação multidestinatória na camada de rede
 - Redundância nos enlaces físicos
 - Aumento da latência

Desempenho



Futuro: Multicast no IPv6

- Todos os nós devem suportar o multicast
 - Implementações não precisam suportar túneis multicast
- Modelo de serviço idêntico ao IPv4
- Escopo
 - Definido explicitamente no endereço multicast
 - Implementação
 - IPv4: endereço unicast é utilizado na identificação da interface
 - Inadequado no IPv6, uma interface pode ter vários endereços

Protocolos Multicast no IPv6

- Roteamento
 - PIM-SM, PIM-SSM
 - Poucas modificações, já existem implementações
- Gerenciamento de grupo
 - MLD (Multicast Listener Discovery)
 - MLDv1 implementa IGMPv2 [RFC2710]
 - MLDv2 implementa IGMPv3 [RFC 3810]
- MSDP não *deve* existir
- BGMP deve ser implantado

Observações Finais

- Arquitetura IP Multicast
 - Continua complexa
 - Ainda possui problemas de escalabilidade
 - Estado armazenado nos roteadores
- Faltam ferramentas de gerenciamento
- Modelo de tarifação em discussão
- Conclusão: ainda há muito trabalho a fazer

Aula 19

Camada de Rede

Comunicação multidestinatória

Igor Monteiro Moraes

Redes de Computadores I – TCC00185