

Universidade Federal de Santa Catarina  
Departamento de Informática e de Estatística  
Curso de Ciências da Computação

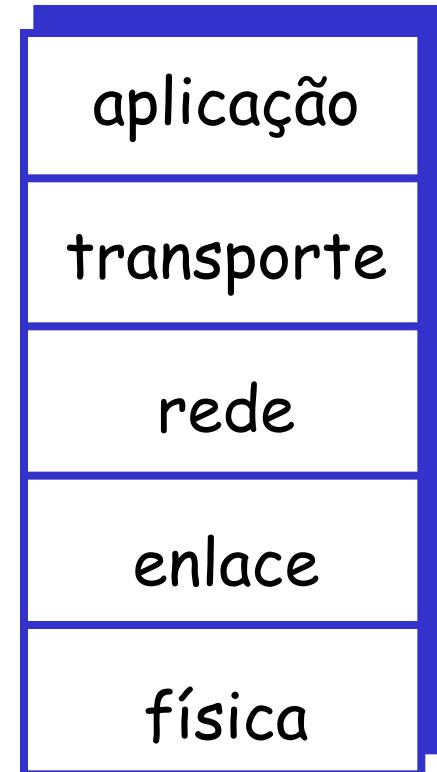
# Capítulo 6

# Camada de Rede

Prof. Roberto Willrich  
INE - UFSC  
[willrich@inf.ufsc.br](mailto:willrich@inf.ufsc.br)

# Pilha de protocolos Internet

- **Aplicação:** dá suporte às aplicações de rede
  - HTTP, FTP, SMTP, ...
- **Transporte:** transferência de dados host-a-host (processo-a-processo)
  - TCP, UDP
- **Rede:** roteamento de datagramas da origem até o destino
  - IP, protocolos de roteamento
- **Enlace:** transferência de dados entre elementos de rede vizinhos
  - Ethernet
- **Física:** bits "no fio"



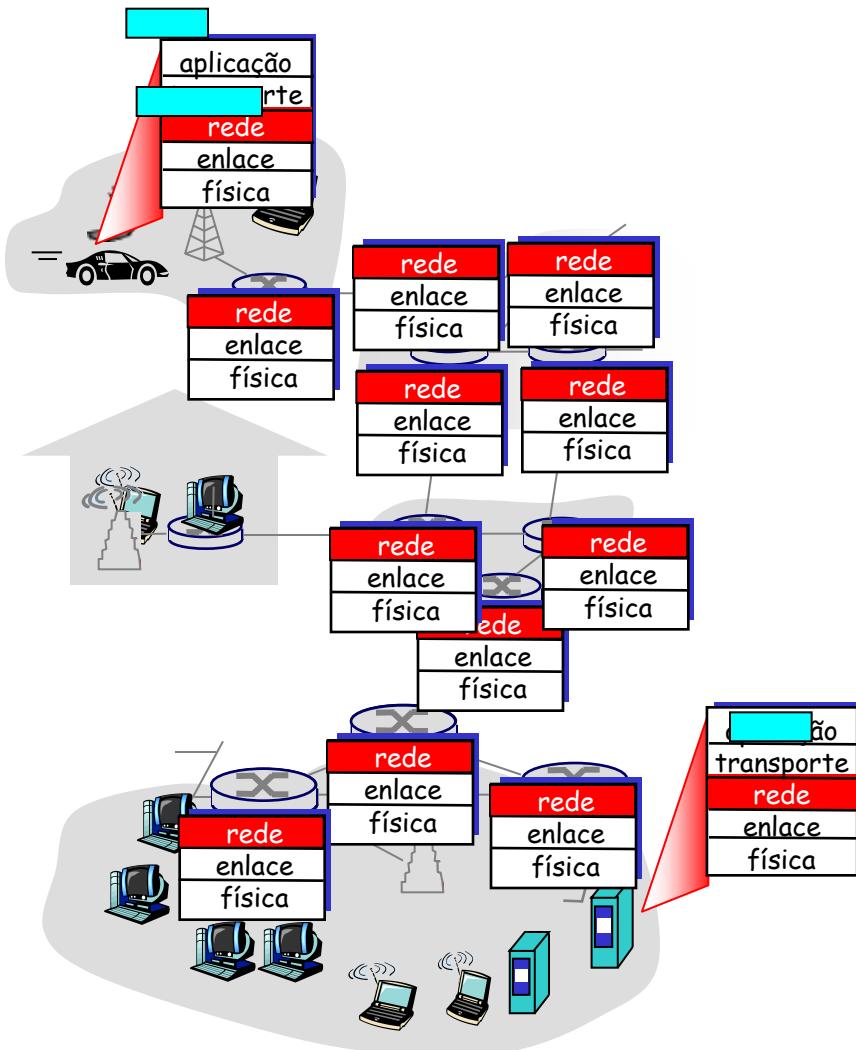
# Cap 6. qqCamada de rede

## Plano do capítulo

- Introdução: funções básicas da camada de rede
- Circuitos Virtuais e Redes de Datagramas
- Camada de rede da Internet
- Endereçamento IP
- Mapeamento de endereços: ARP e RARP
- Formato do datagrama IP
- Fragmentação e remontagem de datagramas IP
- ICMP Internet Control Message Protocol
- Protocolos de roteamento
- Arquitetura do Roteador
- IP Multicast
- IPv6

# Introdução: Camada de rede

- Encaminha segmentos de transporte do hosp. emissor ao receptor
- Lado emissor encapsula segmentos em datagramas
- Camada distribuída: no emissor, receptor e nós intermediários
- Roteador examina campos de cabeçalho em todos os datagramas IP que passam por ele



# Introdução: Duas importantes funções da camada de rede

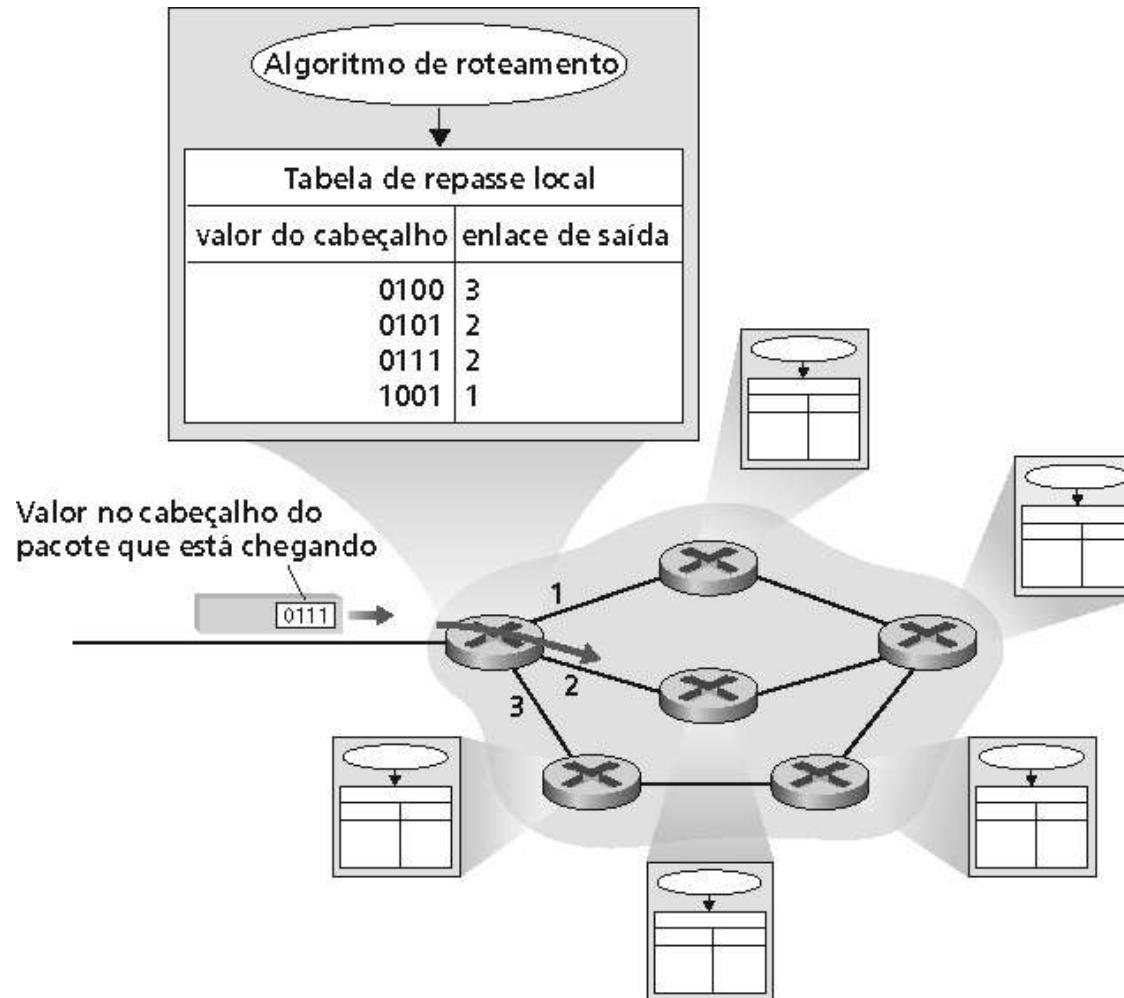
- *repasse*: mover pacotes da entrada do roteador para a saída apropriada do roteador
  
- *roteamento*: determinar rota seguida pelos pacotes da origem ao destino
  - *algoritmos de roteamento*

analogia:

- *roteamento*: processo de planejamento da viagem da origem ao destino
  
- *repasse*: processo de passar por um único cruzamento

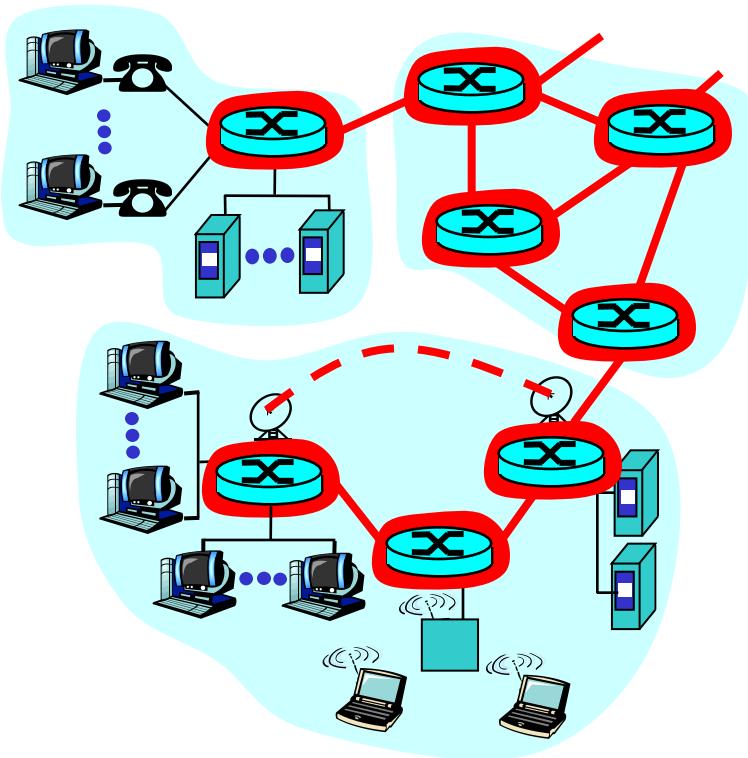
# Introdução: Camada de Rede

- Algoritmos de roteamento determinam valores em tabelas de rota



# Introdução: Camada de Rede

- Coração da rede
  - malha de roteadores interconectado
- Questão: como transferir dados pela rede?
  - Comutação de circuito: circuito dedicado por chamada: rede telefônica
  - Comutação de pacotes: dados enviados através na rede na forma de pacotes



# Introdução: Estabelecimento de Conexão

- 3a função importante em algumas arquiteturas de rede:
  - ATM, frame relay, X.25
- Antes do fluxo de datagramas, dois hospedeiros e os roteadores no caminho estabelecem uma conexão virtual
- Serviço de conexão da camada de rede e de transporte:
  - Rede: entre dois hospedeiros
  - Transporte: entre dois processos

# Introdução: Modelo de serviço de rede

P: Que *modelo de serviço* é o melhor para o “canal” que transporta datagramas do remetente ao destinatário?

exemplo de serviços para  
datagramas  
individuais:

- entrada garantida
- entrega garantida com atraso limitado

exemplo de serviços para  
fluxo de datagramas:

- entrega de datagrama na ordem
- largura de banda mínima garantida
- restrições sobre variações de atraso entre pacotes

# Introdução: Modelos de serviço da camada de rede

Arquitetura de rede	Modelo de serviço	Parâmetros garantidos				Realim. de congestionamento
		Banda	Perda	Ordem	Atraso	
Internet	melhor esforço	não	não	não	não	não (examina perdas)
ATM	CBR	taxa constante	sim	sim	sim	não há congestionamento
ATM	VBR	taxa garantida	sim	sim	sim	não há congestionamento
ATM	ABR	mínimo garantido	não	sim	não	sim
ATM	UBR	não	não	sim	não	não

- Novos serviços na Internet: Intserv, Diffserv

# Cap 6. Camada de rede

## Plano do capítulo

- Introdução: Funções básicas da camada de rede
- **Circuitos Virtuais e Redes de Datagramas**
- Camada de rede da Internet
- Endereçamento IP
- Mapeamento de endereços: ARP e RARP
- Formato do datagrama IP
- Fragmentação e remontagem de datagramas IP
- ICMP Internet Control Message Protocol
- Protocolos de roteamento
- Arquitetura do Roteador
- IP Multicast
- IPv6

# Camada de rede: serviços de conexão e sem-conexão

- Redes de datagrama
  - Fornecem serviços sem-conexão na camada de rede
- Redes de circuito virtual
  - provêem serviços de conexão na camada de rede
- Análogo aos serviços da camada de transporte, mas:
  - Serviço: hospedeiro-a-hospedeiro
  - Sem escolha: a rede provê ou um ou outro
  - Implementação: no núcleo

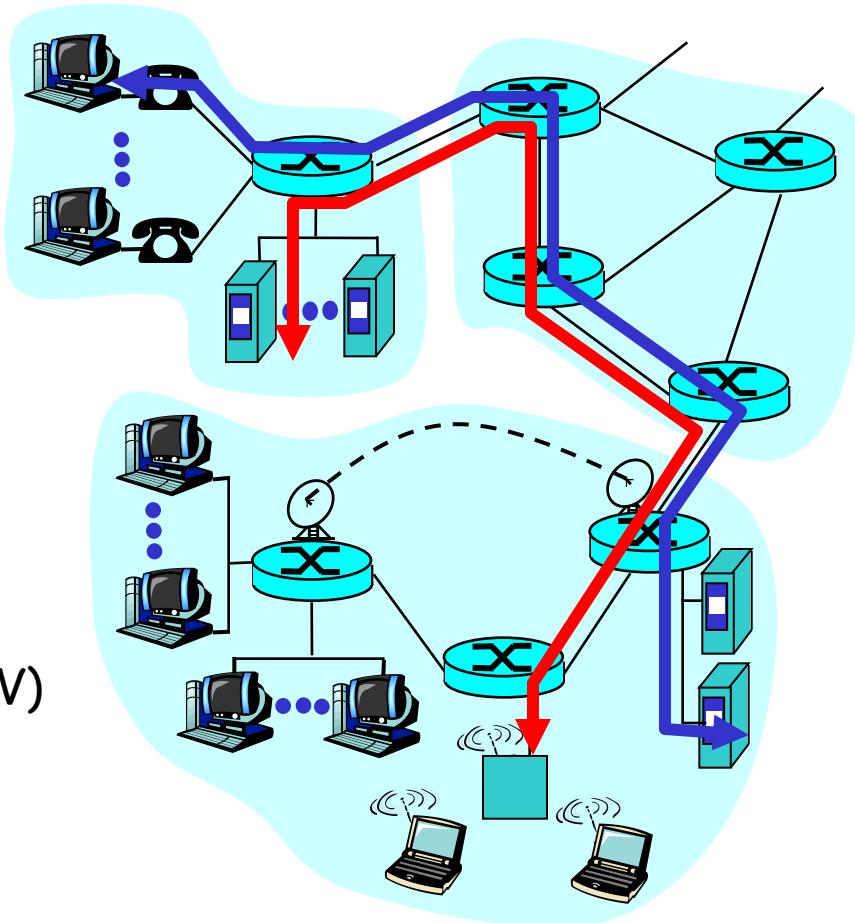
# Circuitos Virtuais (VC)

- “A ligação entre a origem e o destino emula uma ligação telefônica”
  - Orientado ao desempenho
  - A rede controla a conexão entre a origem e o destino
- Estabelecimento da conexão deve preceder o envio de dados. Liberação da conexão após os dados.
  - Cada pacote transporta um identificador do CV, não transporta o endereço completo do destino
  - Cada roteador na rota mantém informação de estado para cada conexão que passa por ele.
- O link e os recursos do roteador (banda, buffers) podem ser alocados por VC

# Comutação de Circuito

Recursos fim-a-fim reservados para a “chamada”

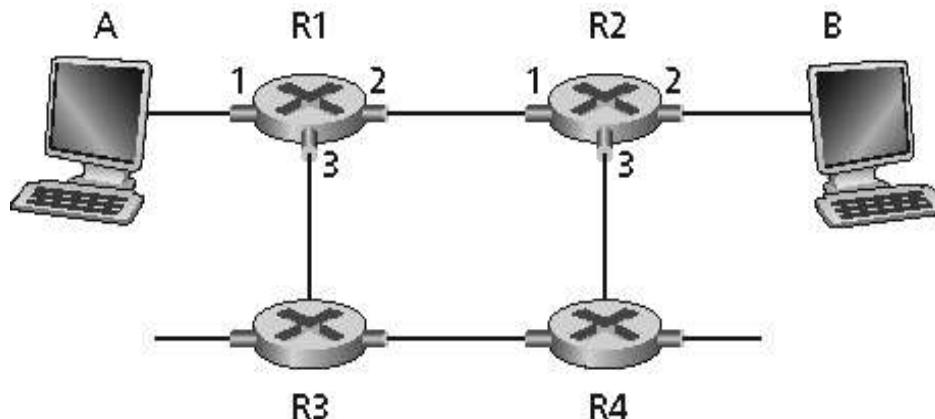
- Largura de banda do enlace, capacidade de chaveamento
- Recursos dedicados: sem compartilhamento
- Desempenho garantido
  - Taxa, atraso
- Requer uma configuração de chamada
  - Para criar um Circuito Virtual (CV)



# Implementação de VC

- Um VC consiste de:
  - Caminho da origem até o destino
  - Números de VC, um número para cada enlace ao longo do caminho
  - Entradas em tabelas de comutação em roteadores ao longo do caminho
- Pacotes pertencentes a um VC carregam um número de VC
- O número de VC deve ser trocado em cada enlace
  - Novos números de VC são obtidos da tabela de comutação

# Tabela de comutação



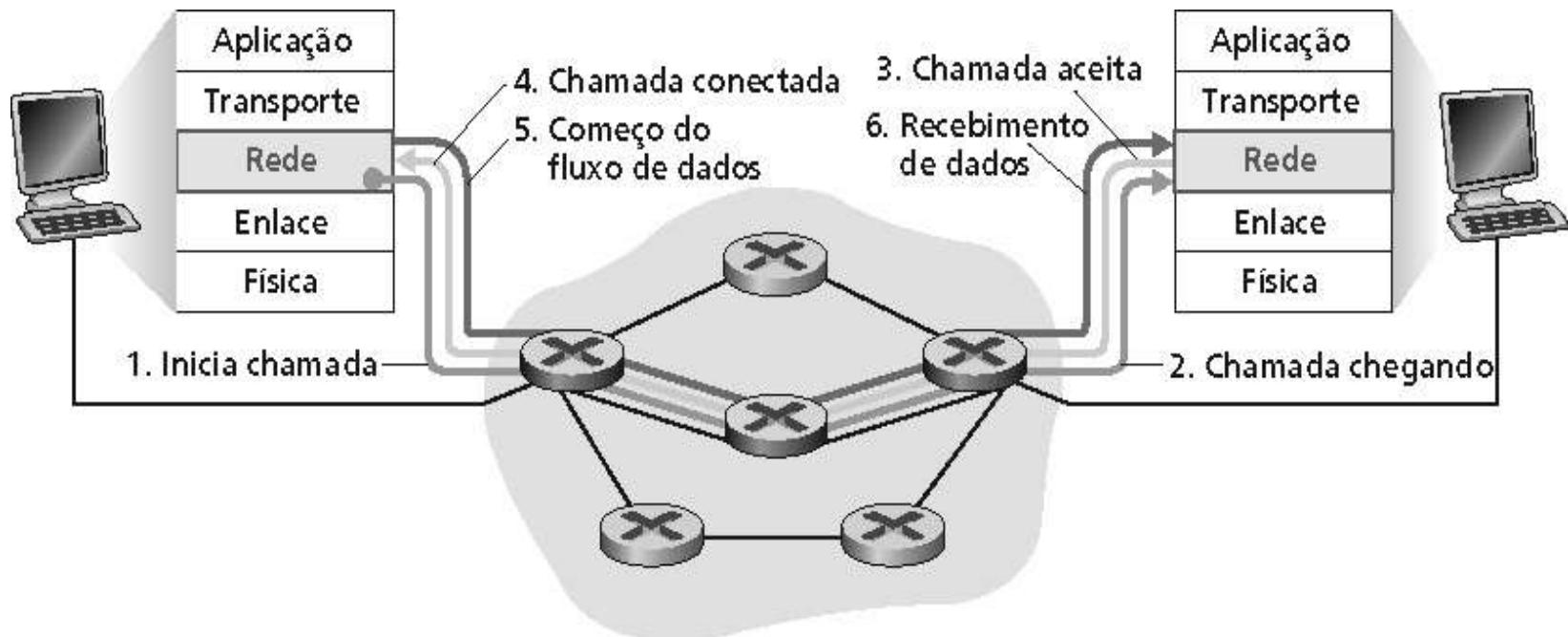
*Tabela de comutação no roteador R1:*

Interface de entrada	VC # de entrada	Interface de saída	VC # de saída
1	12	2	22
2	63	1	18
3	7	2	17
1	97	3	87
...	...	...	...

Roteadores mantêm informações de estado de conexão

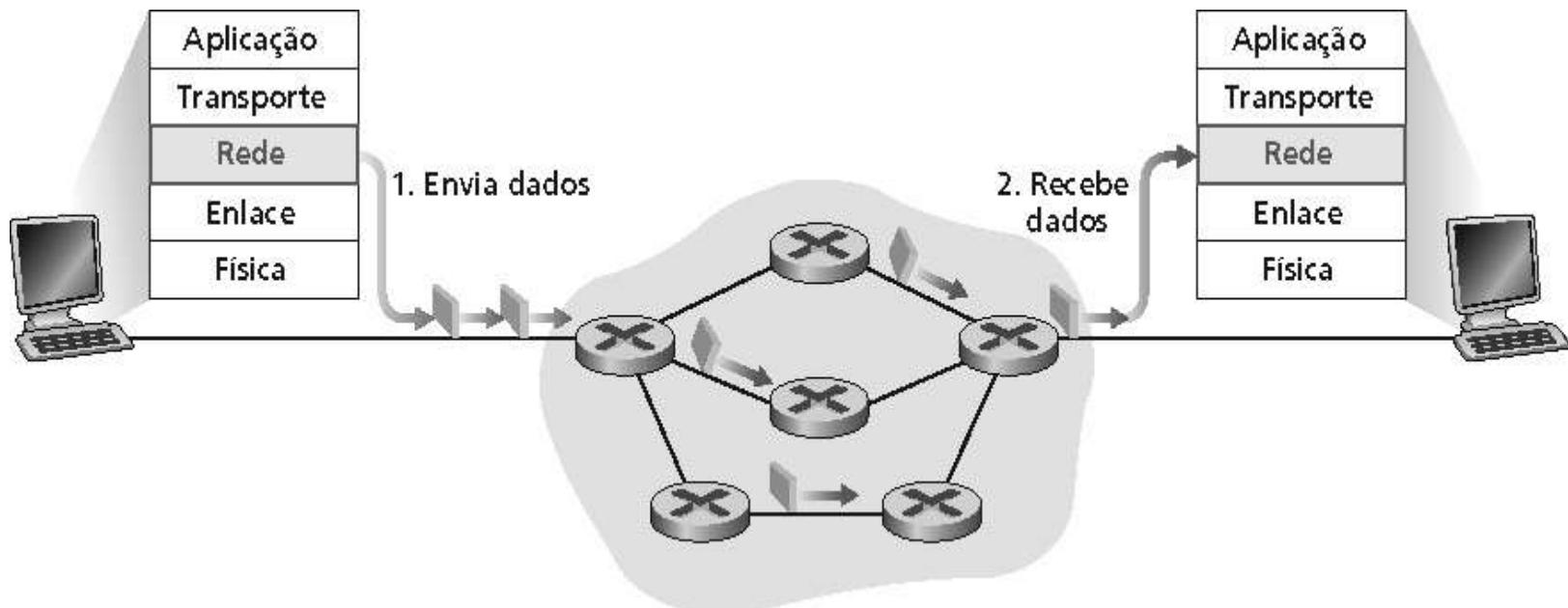
# Circuitos virtuais: protocolos de sinalização

- Usado para estabelecer, manter e encerrar circuitos virtuais
  - Usados em ATM, frame-relay e X-25
  - Não é usado na Internet atualmente



# Redes de datagrama

- Não existe estabelecimento de conexão na camada de rede
- Roteadores: não existe estado sobre conexões fim-a-fim
  - O conceito “conexão” não existe na camada de rede
- Pacotes são encaminhados pelo endereço do hospedeiro de destino
  - Pacotes para o mesmo destino podem seguir diferentes rotas



# Tabela de comutação

**4 bilhões de entradas possíveis**

<b>Faixa de Endereços de Destino</b>	<b>Interface de Enlace</b>
11001000 00010111 00010000 00000000 até	0
11001000 00010111 00010111 11111111	
11001000 00010111 00011000 00000000 até	1
11001000 00010111 00011000 11111111	
11001000 00010111 00011001 00000000 até	2
11001000 00010111 00011111 11111111	
senão	3

# Encontro de prefixos maiores

<b>Prefixo do endereço</b>	<b>Interface de Enlace</b>
11001000 00010111 00010	0
11001000 00010111 00011000	1
11001000 00010111 00011	2
senão	3

## Exemplos

DA: 11001000 00010111 00010110 10100001      Qual interface?

DA: 11001000 00010111 00011000 10101010      Qual interface?

# Datagrama versus circuito virtual

- Internet
  - Dados trocados entre computadores
    - Serviço elástico, requisitos de atraso não críticos
  - Sistemas finais inteligentes
    - Podem adaptar-se, realizar controle e recuperação de erros
    - A rede é simples; a complexidade fica nas pontas
  - Muitos tipos de enlaces
    - Características diferentes
    - Difícil obter um serviço uniforme
- ATM
  - Originário da telefonia
  - Conversação humana:
    - Tempos estritos, exigências de confiabilidade
    - Necessário para serviço garantido
  - Sistemas finais “burros”
    - Telefones
    - Complexidade dentro da rede

# Comutação de Pacotes versus de Circuitos

- Vantagem da Comutação de Circuitos
  - Existe a garantia de taxa de transmissão e atrasos constantes
    - Útil quando existe um fluxo contínuo e constante de informação
  - Não exige empacotamento de bits
    - Aumenta a eficiência na transmissão
- Problema da Comutação de Circuitos
  - utilização da rede não é muito eficiente
    - Quando o fluxo não é constante ou contínuo (a taxa variáveis)
      - Ter-se-ia que alocar canais na taxa de pico
    - Mesmo que os dois sistemas não tenham nada a transmitir, o intervalo de tempo dedicado ao circuito é mantido

# Comutação de Pacotes versus de Circuitos

## □ Vantagens da Comutação de Pacotes

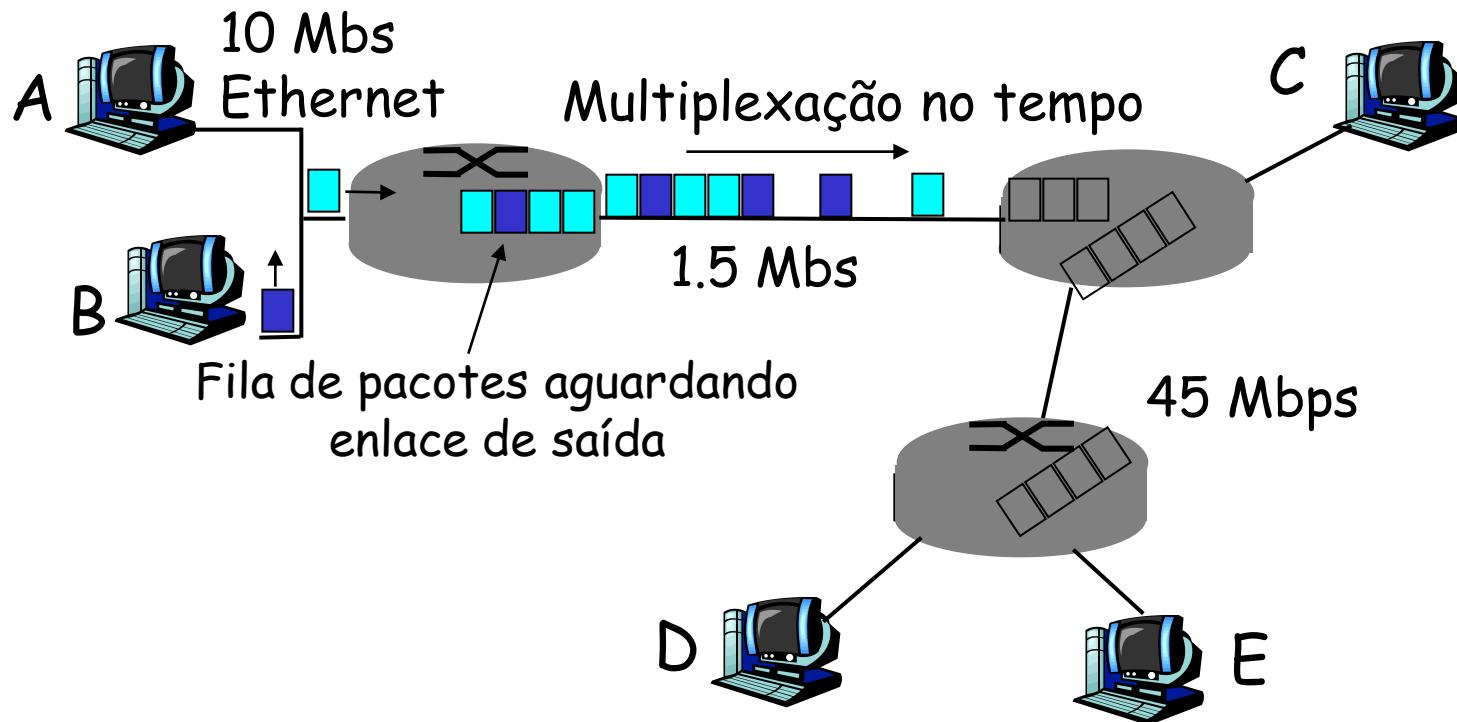
- Assegura que grandes mensagens não degradem o tempo de resposta da rede
  - grande vantagem: alta eficiência da multiplexagem da rede
- os dispositivos fonte e destino podem operar em diferentes taxas

## □ Desvantagem da Comutação de Pacotes

- relacionada ao seu comportamento estatístico
  - taxa de bits não é facilmente garantida, e atrasos de trânsito pode variar significativamente

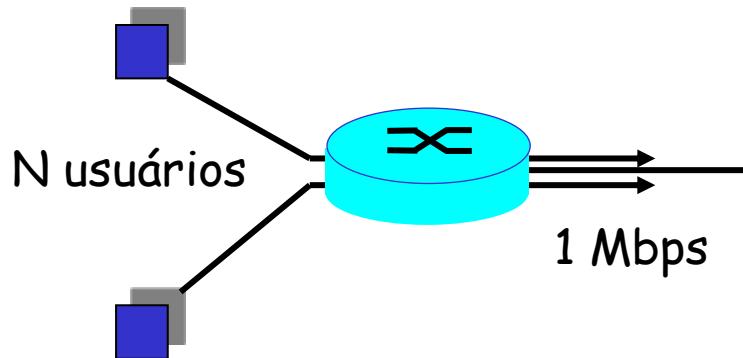
# Comutação de Pacotes

- Recursos podem não ser suficientes
  - Quantidade de recursos solicitados pode ultrapassar o montante disponível
    - Congestionamento: fila de pacotes, espera por enlaces em uso
  - Armazenar-e-transmitir:
    - Transmitido sobre um enlace
    - Espera sua vez no próximo enlace



# Comutação de Pacotes versus de Circuitos

- Comutação a pacotes permite que mais usuários utilizem a rede!
  - Enlace de 1 Mbps
    - Cada usuário usa 100Kbps quando “ativo”
    - ativo 10% do tempo
  - Comutação a circuito:
    - 10 usuários
  - Comutação a pacotes:
    - com 35 usuários, probabilidade de ter 10 usuários ativos é menor que 0.004



# Comutação de Pacote versus de Circuito

- Comutação a pacotes é melhor para tráfego em rajadas
  - Utiliza recursos quando necessário
  - Sem configuração de chamada
- Comutação a pacotes pode sofrer congestionamento: perda e atraso de pacotes
  - Protocolos de transporte de dados deve oferecer confiabilidade (TCP) e controle de congestionamento (TCP)
- Como prover um comportamento semelhante ao circuito?
  - Garantia de largura de banda necessária para aplicações de áudio e vídeo
  - Uma solução é Qualidade de Serviço

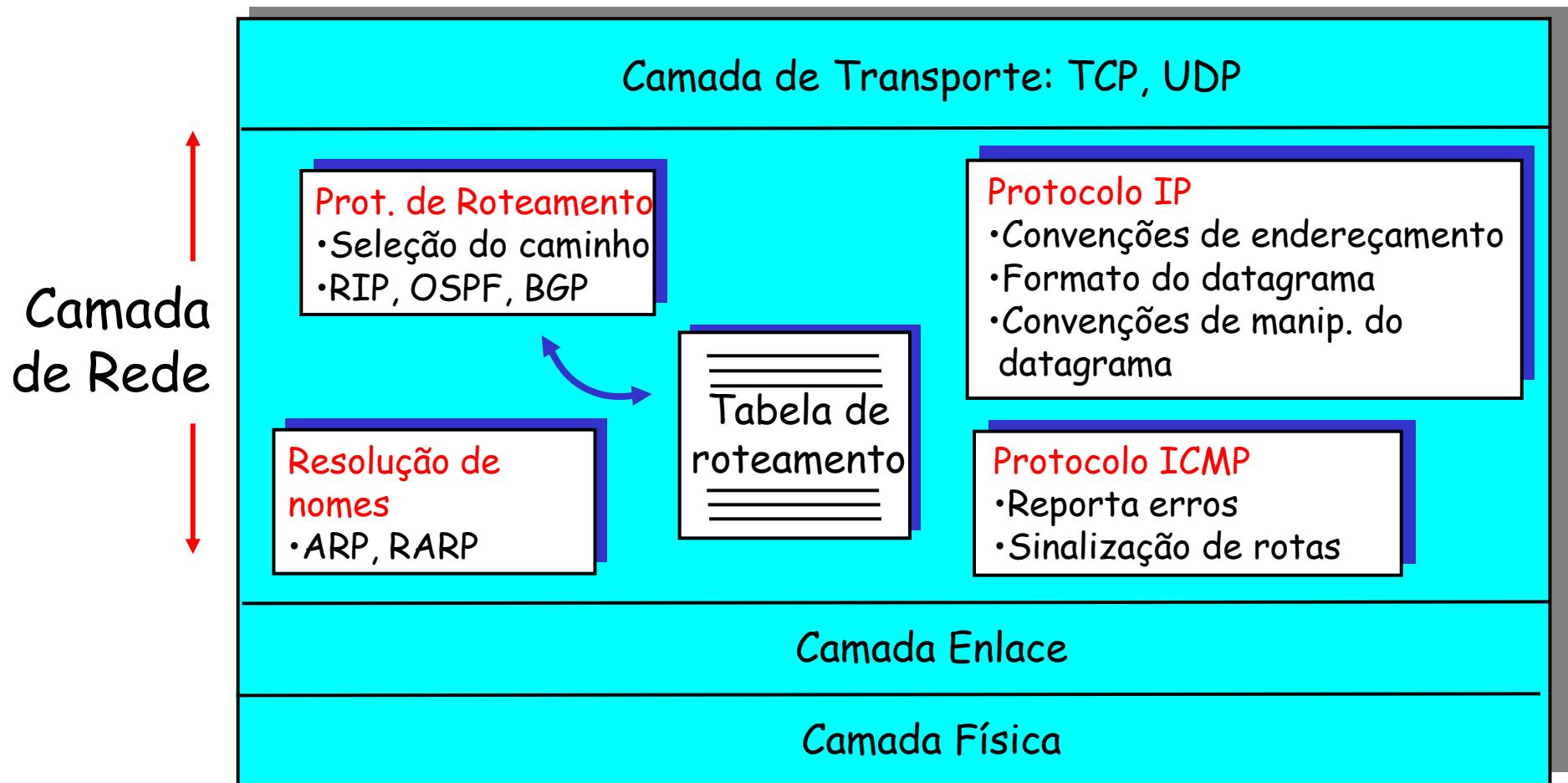
# Camada de rede

## □ Plano do capítulo

- Introdução: Funções básicas da camada de rede
- Circuitos Virtuais e Redes de Datagramas
- **Camada de rede da Internet**
- Endereçamento IP
- Mapeamento de endereços: ARP e RARP
- Formato do datagrama IP
- Fragmentação e remontagem de datagramas IP
- ICMP Internet Control Message Protocol
- Protocolos de roteamento
- Arquitetura do Roteador
- IP Multicast
- IPv6

# Camada de Rede da Internet

- Funções da camada de rede de roteadores e hosts:



# Tipos de Endereçamento

## □ Endereçamento Horizontais

- O endereço não tem relação alguma com o lugar onde estão as entidades dentro da rede
  - Exemplo: endereços globalmente administrados (IEEE 802)
    - Constituído do número de assinatura do usuário
- Dificulta o roteamento
  - Não tem informações explícitas sobre a localização da entidade
- Facilita a mobilidade
  - Não necessita uma renumeração da entidade

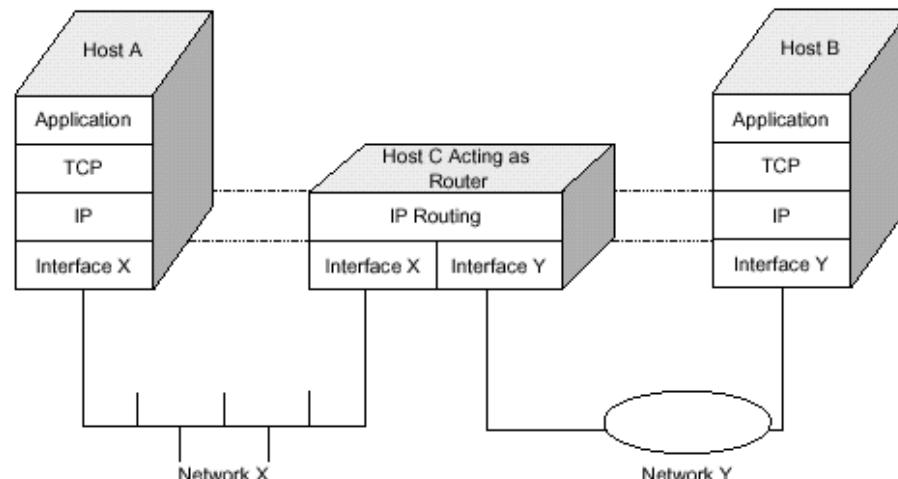
# Tipos de Endereçamento

## □ Endereçamento Hierárquico

- O endereço de uma entidade é constituído de acordo com os endereços correspondentes aos vários níveis de hierarquia de que ela faz parte
  - Exemplo: Redes Públicas de Pacote (recom. X.121)
    - Endereços são números decimais formados por três campos: código do país, código para a rede e um campo para endereçamento dentro da rede
  - Exemplo: Endereço IP (Internet Protocol)
    - Identificação do host é formado pelo endereço da rede e pelo endereço do hospedeiro

# Endereçamento IP

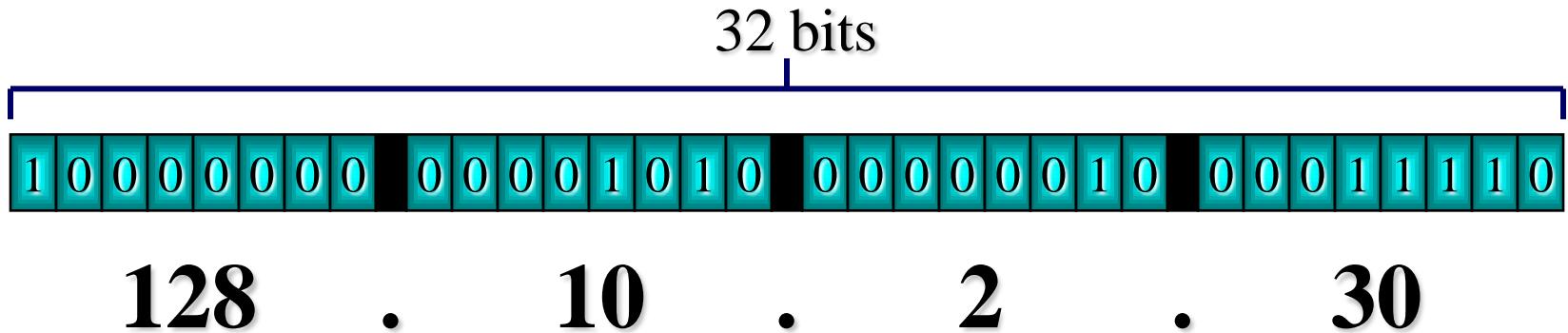
- Como Roteadores e Hosts são ligados a rede
  - Host típico tem apenas uma conexão para a rede (interface)
    - IP envia/recebe datagramas por esta interface
    - Tem um endereço IP
  - Roteador tem geralmente diversas interfaces
    - Recebe datagramas de enlaces de entrada e envia para enlaces de saída
    - Tem vários endereços IP



# Endereçamento IPv4

## □ Notação Decimal Pontuada

### ○ Exemplo:

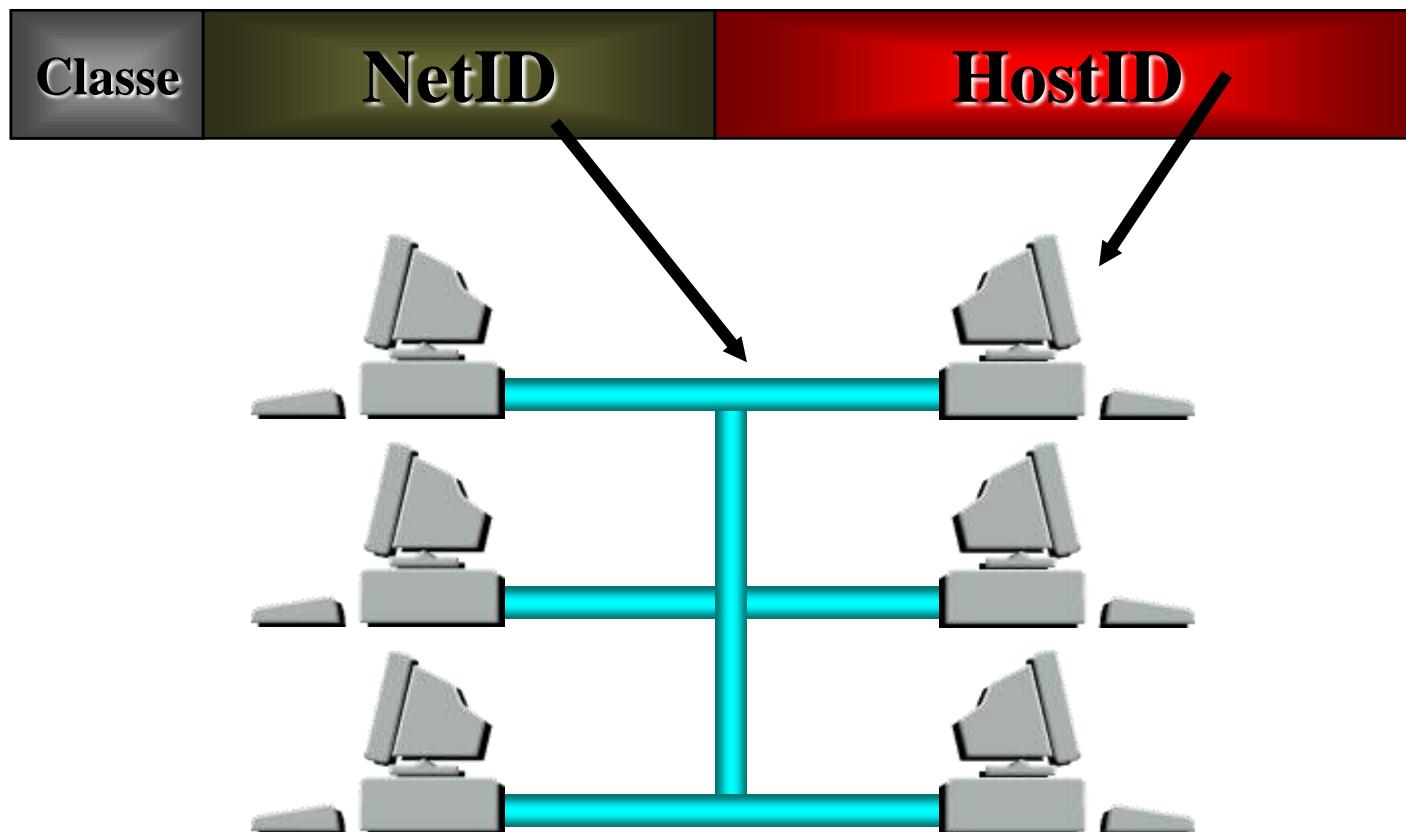


# Endereçamento IP

- Capacidade de Endereçamento
  - Palavra de 32 bits (4 bytes)
    - Por exemplo: 150.162.60.200
  - Permite endereçar mais de 4 bilhões de máquinas
    - $2^{32} = 4.294.967.296$
    - Mas diversos endereços são reservados

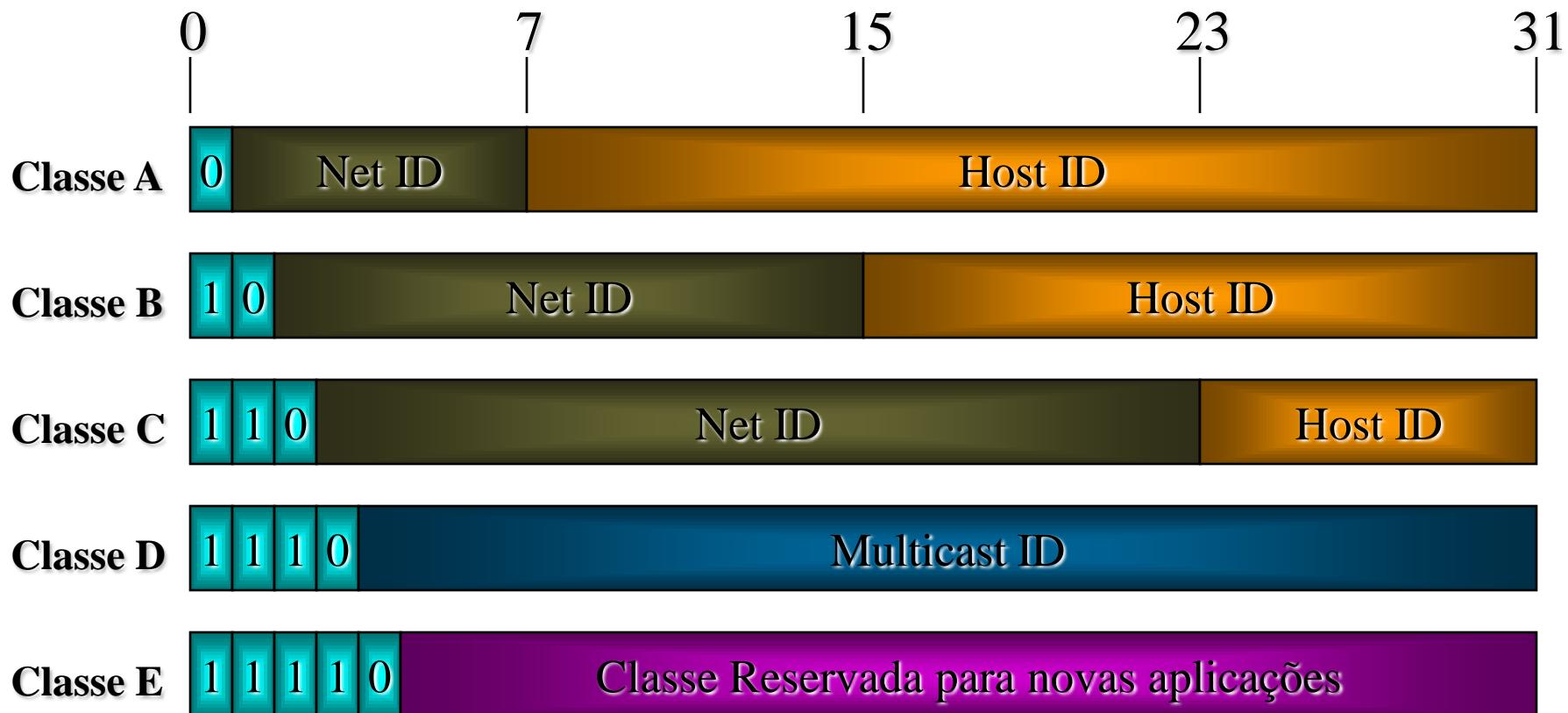
# Endereçamento IP

- Definição de classes de endereçamento
  - Redes de classe A, B, C, D e E
    - Pelo fato que as redes da Internet variarem muito de tamanho



# Endereçamento IP

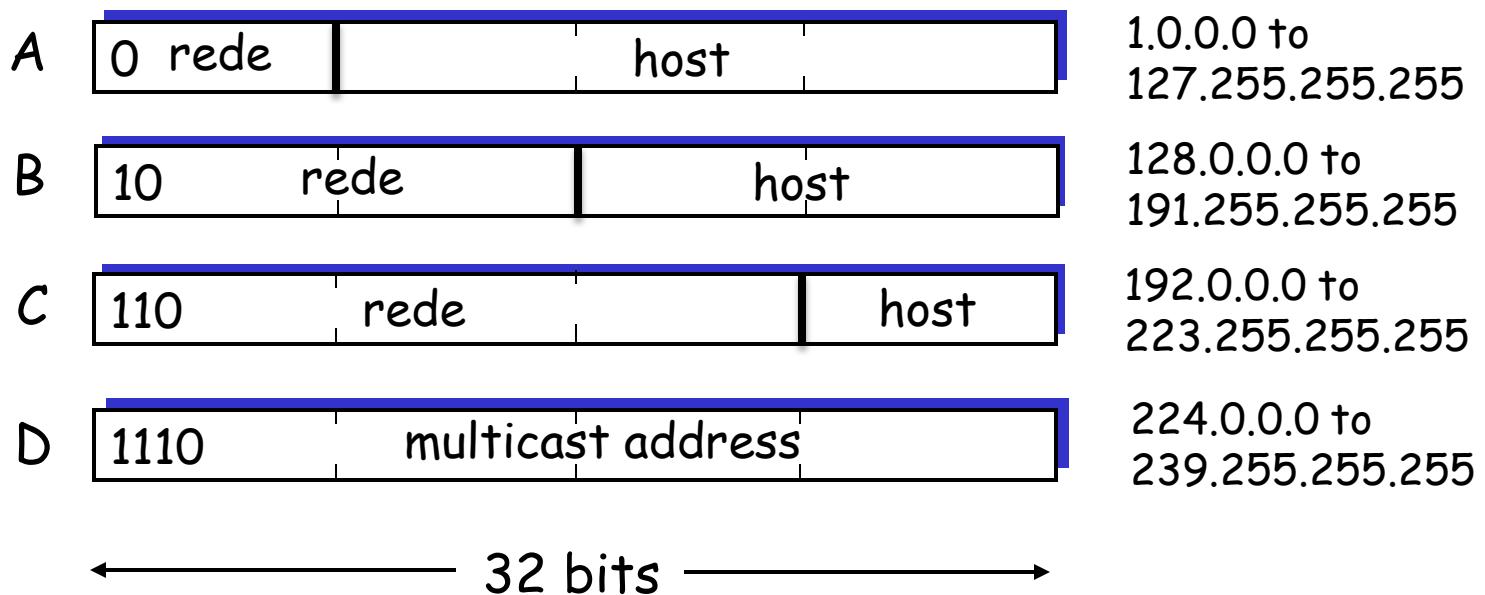
## □ Classes de Endereçamento



# Endereços IP

- Endereçamento “class-full”:

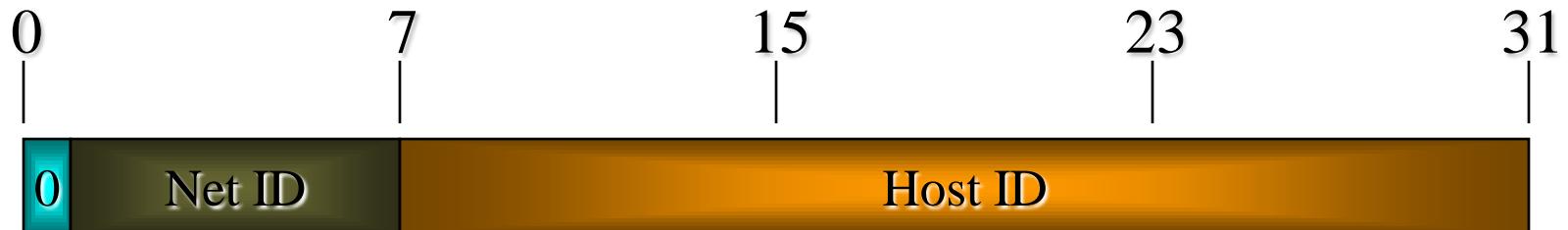
class



# Endereços Especiais

Prefixo	Sufixo	Tipo de endereço	Propósito
Todos 0's	Todos 0's	Este computador	Usado durante o boot
Rede	Todos 0's	Rede	Identifica a rede
Rede	Todos 1's	Broadcast direto	Broadcast para uma rede específica
Todos 1's	Todos 1's	Broadcast limitado	Broadcast na rede local
127	Qualquer	Loopback	Teste

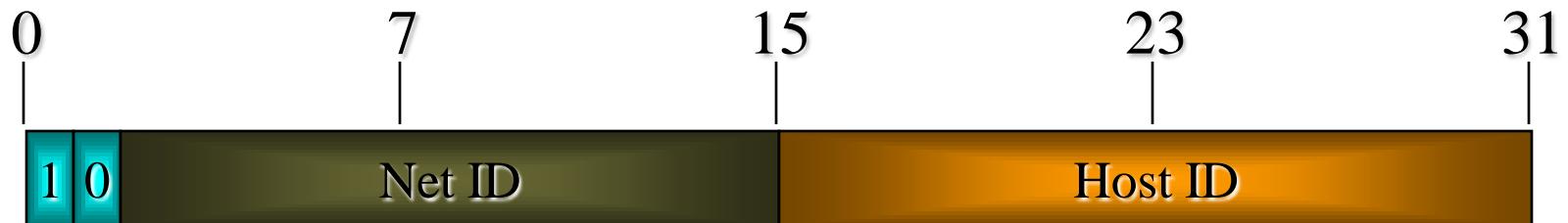
# Endereçamento IP



## Classe A

- Usada em redes de grande porte
  - Endereços de rede variam de 1 a 126
  - Cada rede pode ter 16 milhões de hosts
- Exemplo: rede Arpanet

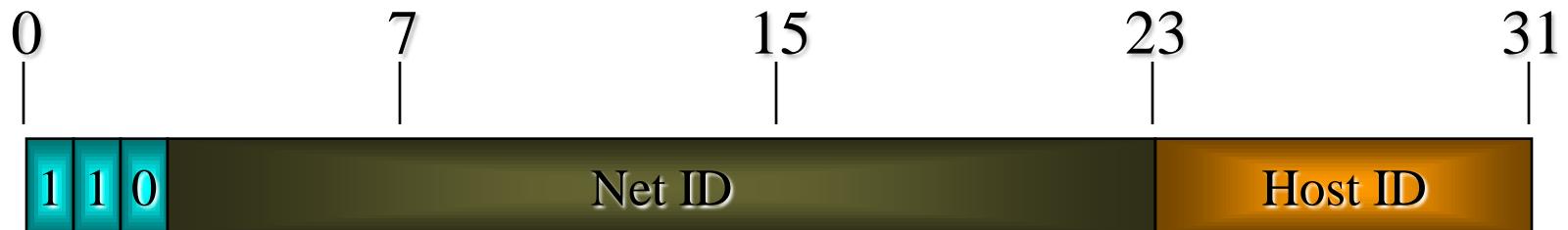
# Endereçamento IP



## Classe B

- Endereços de rede variando de 128.1 até 191.255
  - Cada rede pode ter 65 mil hosts

# Endereçamento IP

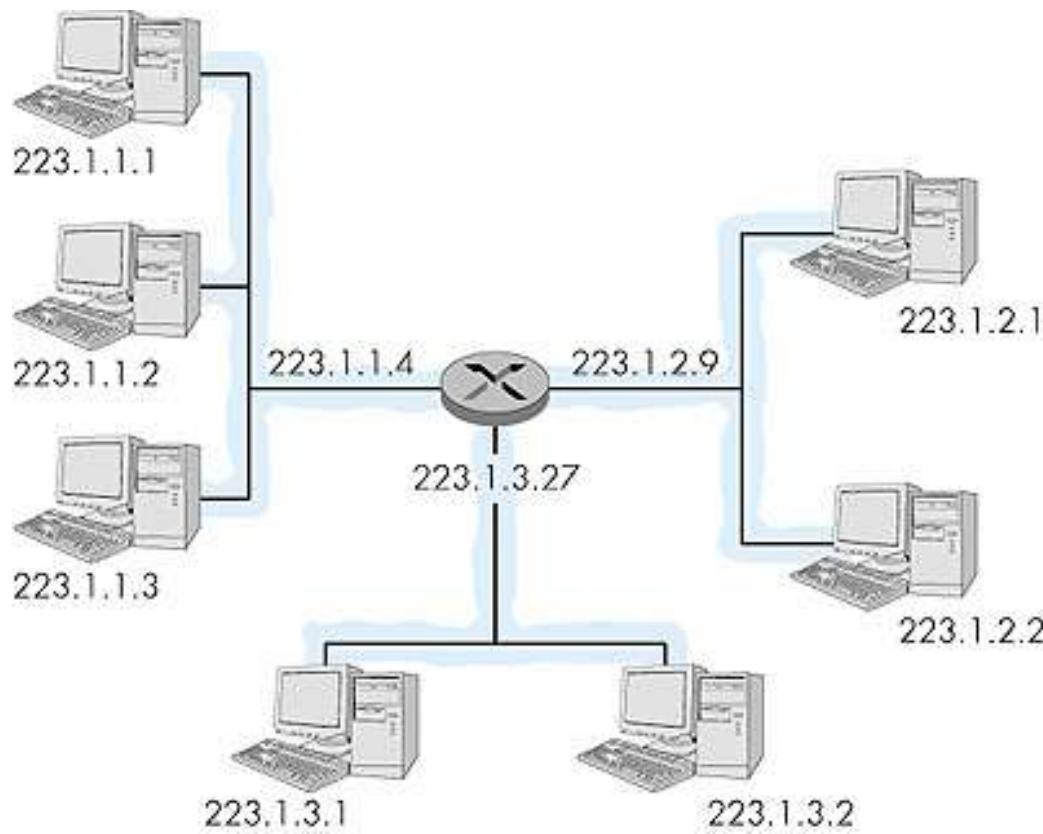


## Classe C

- Endereços de rede variando de 192.1.1 até 223.254.254
- Cada rede pode ter 254 hosts

# Endereçamento IP

- Como Roteadores e Hosts são ligados a rede
  - Endereços IP são associados a interfaces
  - Endereços IP podem também identificar uma rede
    - Campo de host com todos bits iguais a 0



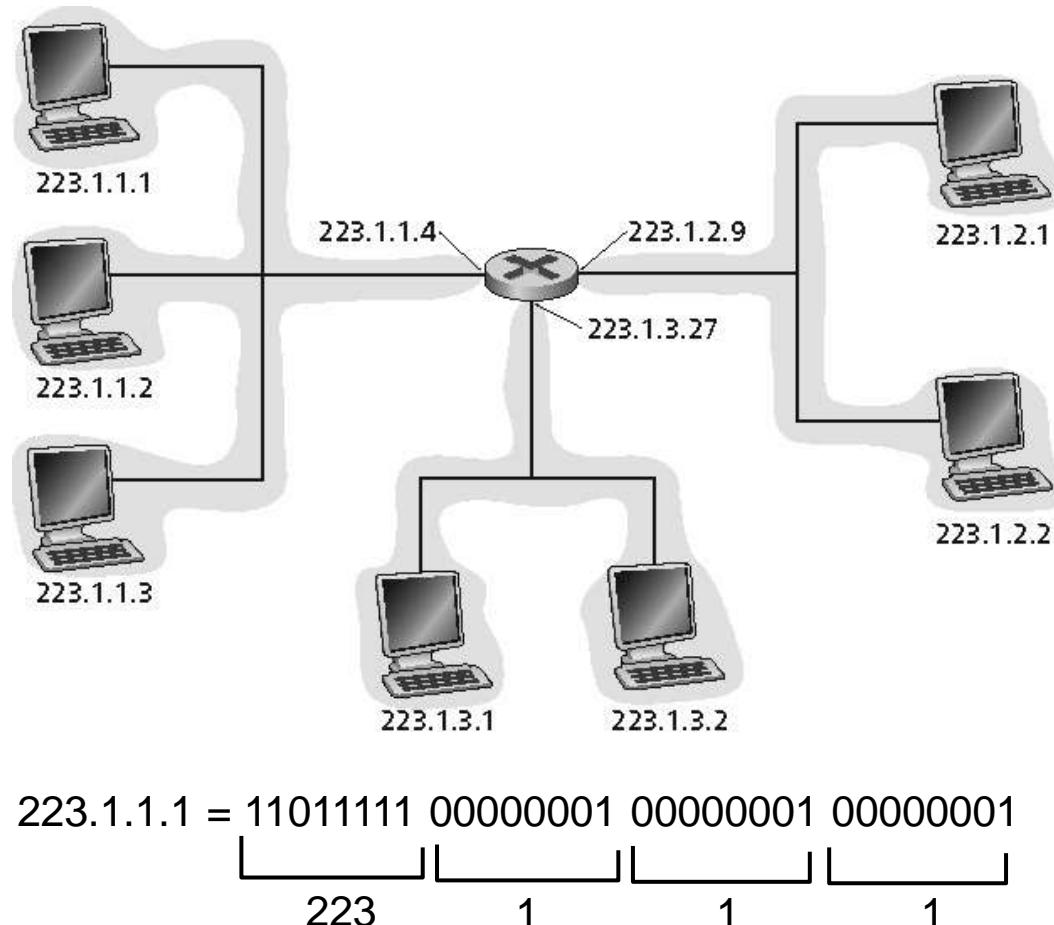
# Endereçamento IP

- Endereço IP:

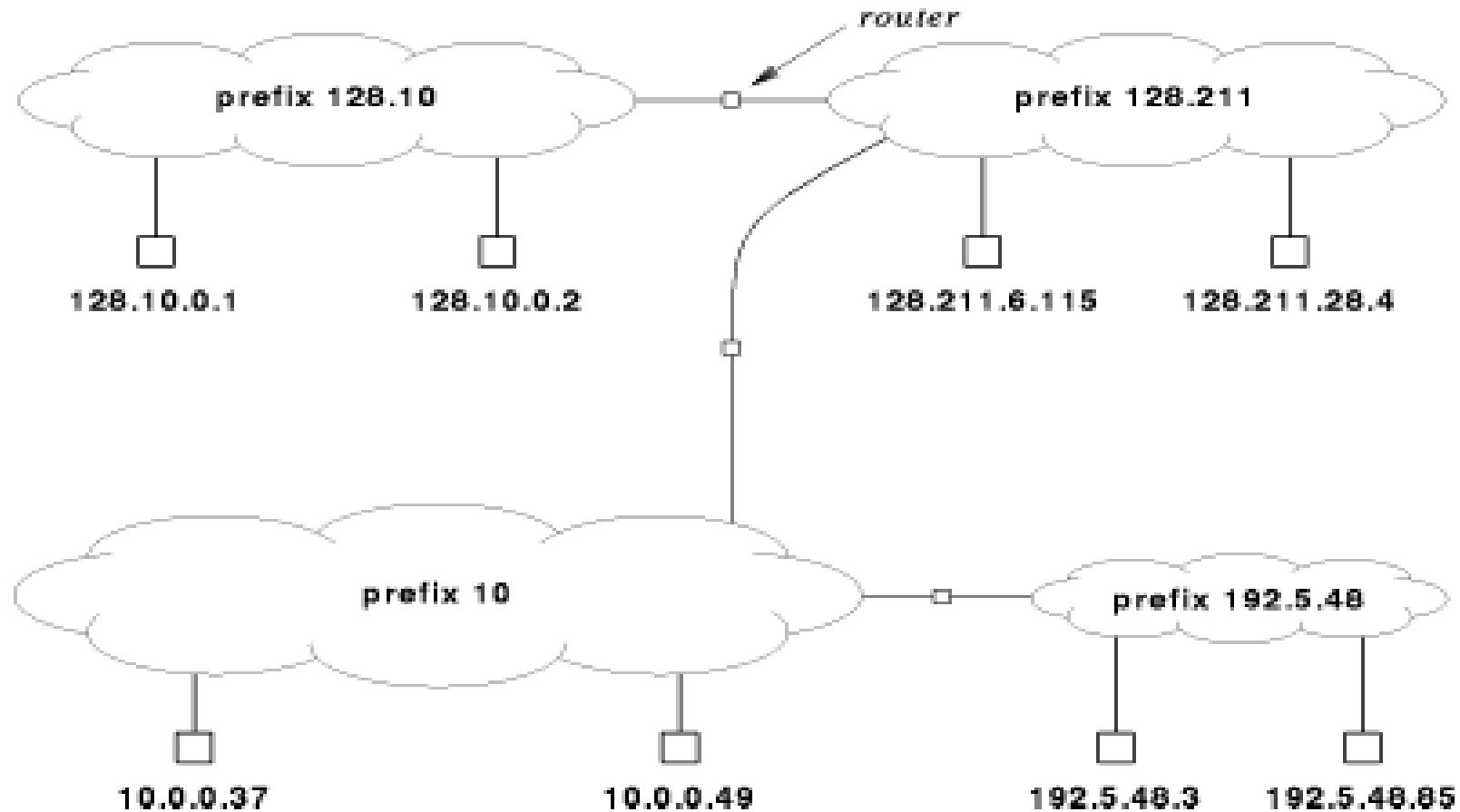
- Parte rede (bits mais significativos)
  - Parte host (bits menos significativos)

- O que é uma rede? (da perspectiva de endereço IP)

- Dispositivos com interface com o mesma parte rede do endereço IP
  - Pode alcançar fisicamente outro dispositivo sem intervenção do roteador

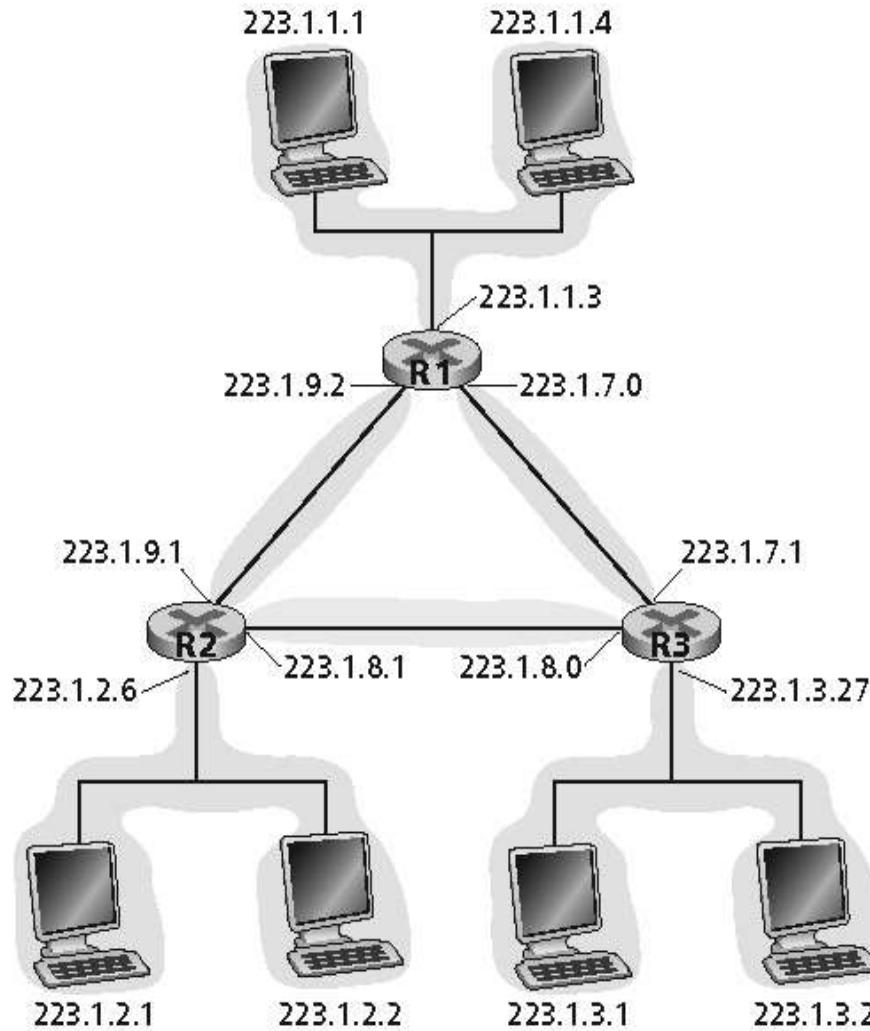


# Exemplo de Endereçamento



# Exemplo de Endereçamento

Quantas redes?



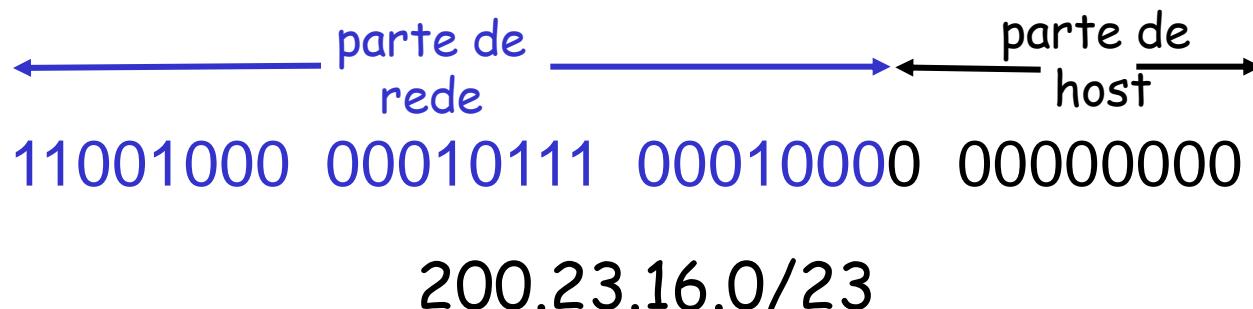
# Endereçamento IP

## □ Sub-redes

- Com o crescimento de uma empresa o números de hosts possíveis de uma classe pode ser insuficiente
  - Se uma empresa tiver mais de 254 hosts e tiver um endereço classe C?
  - Se uma rede de Classe B aloca endereços para 65K hosts, mesmo se só existem 2000 hosts naquela rede?
- Solução: Sub-redes
  - Permitir que uma rede seja dividida em diversas partes para uso interno
    - Mas externamente é vista como uma única rede

# Endereçamento IP: CIDR

- Endereçamento “Classful”:
  - Uso ineficiente do espaço de endereçamento, exaustão do espaço de endereços
- **CIDR: classless interdomain routing**
  - A porção de endereço de rede tem tamanho arbitrário
  - Formato do endereço: **A.B.C.D/x**, onde **x** é o número de bits na parte de rede do endereço

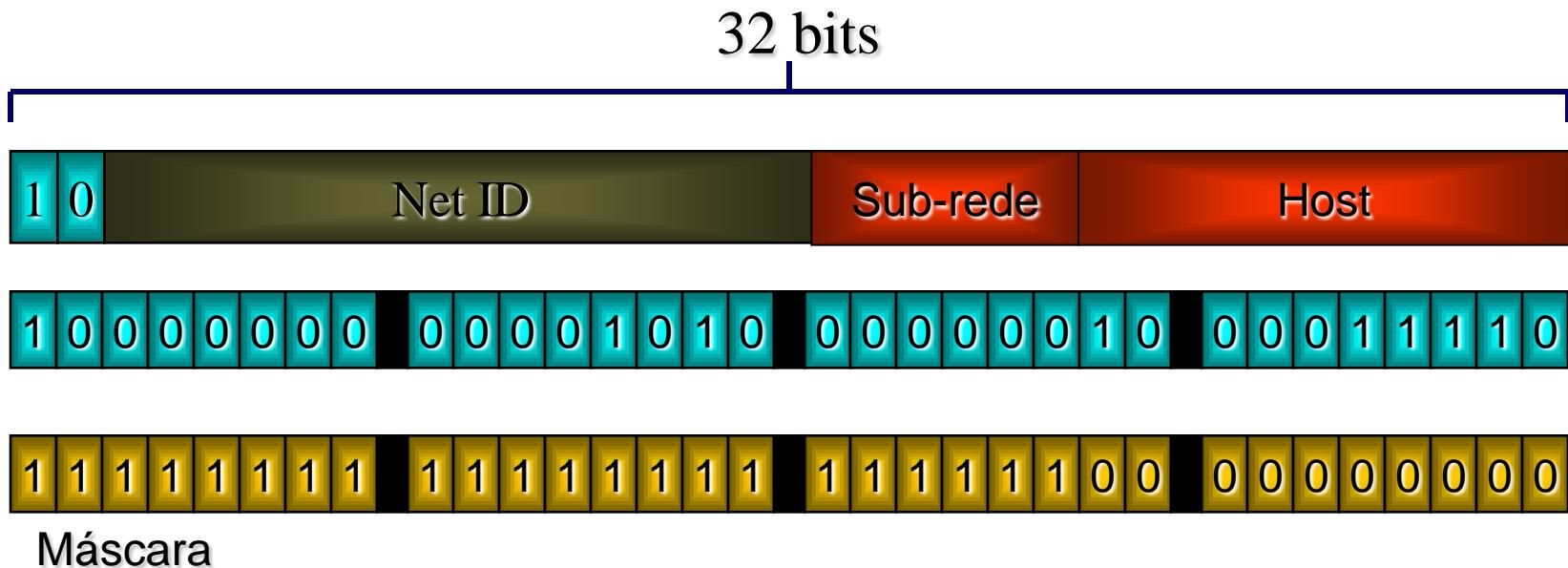


## Endereçamento IP

## Sub-redes

## ○ Considerando Classe B

- Máscara abaixo permite criar até 62 LANs ( $2^6$ ) com 1022 ( $2^{10}-2$ ) hosts cada
  - Ex.: 128.10.2.30/22

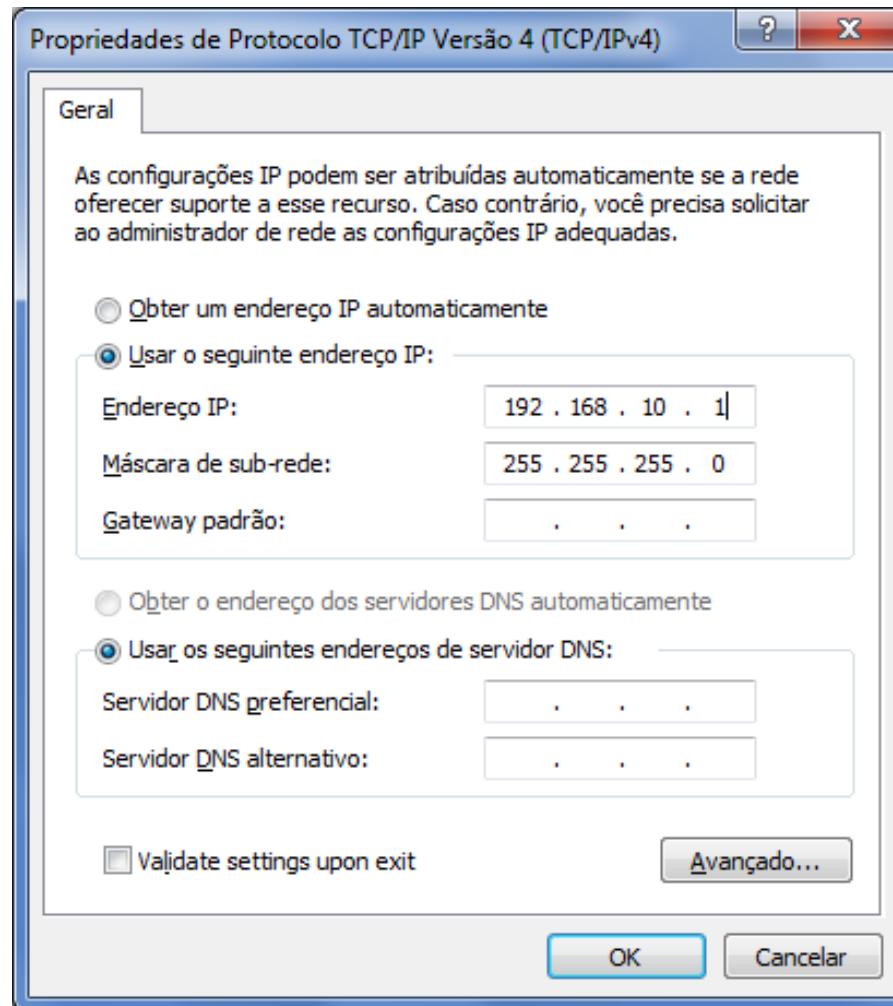


# Atribuindo endereços

- Como um host obtém seu endereço IP?
  - Endereço de rede é fixo para uma rede
  - Existem duas formas para atribuir um endereço de host
- Configuração Manual
  - O endereço IP é configurado no computador pelo administrador do sistema
- Uso do DHCP
  - Dynamic Host Configuration Protocol (DHCP)
    - Um servidor DHCP na rede recebe pedidos DHCP de um cliente e aloca um endereço IP para o cliente

# Atribuindo endereços

## □ Configuração Manual



# DHCP

- DHCP: Dynamic Host Configuration Protocol (RFC 2131)
  - é um protocolo que oferece configuração dinâmica de terminais
    - com concessão de endereços IP de host e outros parâmetros de configuração para clientes de rede.
    - é o sucessor do BOOTP.
  - DHCP utiliza UDP e a porta 67
- Baseia-se no modelo cliente-servidor:
  - Um cliente solicita informações de configuração (endereço IP, máscara de rede, gateway, servidores DNS,...)
  - Servidor DHCP mantém o gerenciamento centralizado dos endereços IP usados na rede
    - Mais de um por rede (aumento da confiabilidade)

# DHCP

- O servidor DHCP possui um pool de endereços (ex: 192.168.1.20 até 192.168.1.254) disponíveis para “alugar” por um determinado período de tempo (lease time) para os clientes
- O tempo do “alocação” dos endereços é configurável
  - Se o tempo de alocação do endereço for muito longo...
    - ... faz com que o endereço que foi alocado para o cliente esteja impossibilitado de ser usado por outro host nesse período de tempo
  - Se o tempo de alocação do endereço for muito curto...
    - ... faz com que o cliente tenha que solicitar com frequência a “renovação” da alocação
- Para configurar o lease time, deve ser levado em consideração a quantidade de hosts que a rede possui e o quanto dinâmico é o ambiente

# DHCP

- Ilustração de informações obtidas por um cliente DHCP
  - Comando ipconfig /all (no command do windows)

```
C:\ Prompt de Comando - more lixo

Adaptador de Rede sem Fio Conexão de rede sem fio:
  Sufixo DNS específico de conexão. . . . . : Intel(R) PRO/Wireless 3945ABG N
  Descrição . . . . . : network Connection
  Endereço Físico . . . . . : 00-1C-BF-1C-53-7B
  DHCP Habilitado . . . . . : Sim
  Configuração Automática Habilitada. . . . . : Sim
  Endereço IPv4. . . . . : 192.168.0.195<Preferencial>
  Máscara de Sub-rede . . . . . : 255.255.255.0
  Concessão Obtida. . . . . : quarta-feira, 21 de outubro de
  2009 07:42:34
  Concessão Expira. . . . . : quinta-feira, 22 de outubro de
  2009 07:42:33
  Gateway Padrão. . . . . : 192.168.0.1
  Servidor DHCP . . . . . : 192.168.0.1
  Servidores DNS. . . . . : 192.168.0.1
  NetBIOS em Tcpip. . . . . : Habilitado

Adaptador Ethernet Conexão local:
  Estado da mídia. . . . . : mídia desconectada
  Sufixo DNS específico de conexão. . . . . : inf.ufsc.br
  Descrição . . . . . : Broadcom NetLink (TM) Gigabit E
  thernet
  Endereço Físico . . . . . : 00-1B-24-95-70-BF
  DHCP Habilitado . . . . . : Sim
-- Mais <10% --
```

# DHCP

- Lease Life Cycle (ciclo de vida de alocação)
  - Alocação (allocation):
    - O cliente não possui um endereço e então faz a requisição por alocação
  - Re-alocação (reallocation):
    - Se o cliente já possui um endereço alocado, quando ele ligar ou reiniciar o host, ele irá contatar o servidor para confirmar a alocação
  - Renovação (renewal):
    - Depois que um certo período do lease time tiver passado, o cliente contata o servidor para renovar a locação
  - Liberação (release):
    - O cliente pode decidir a qualquer momento que não deseja mais utilizar o endereço que lhe foi alocado, e pode encerrar a locação

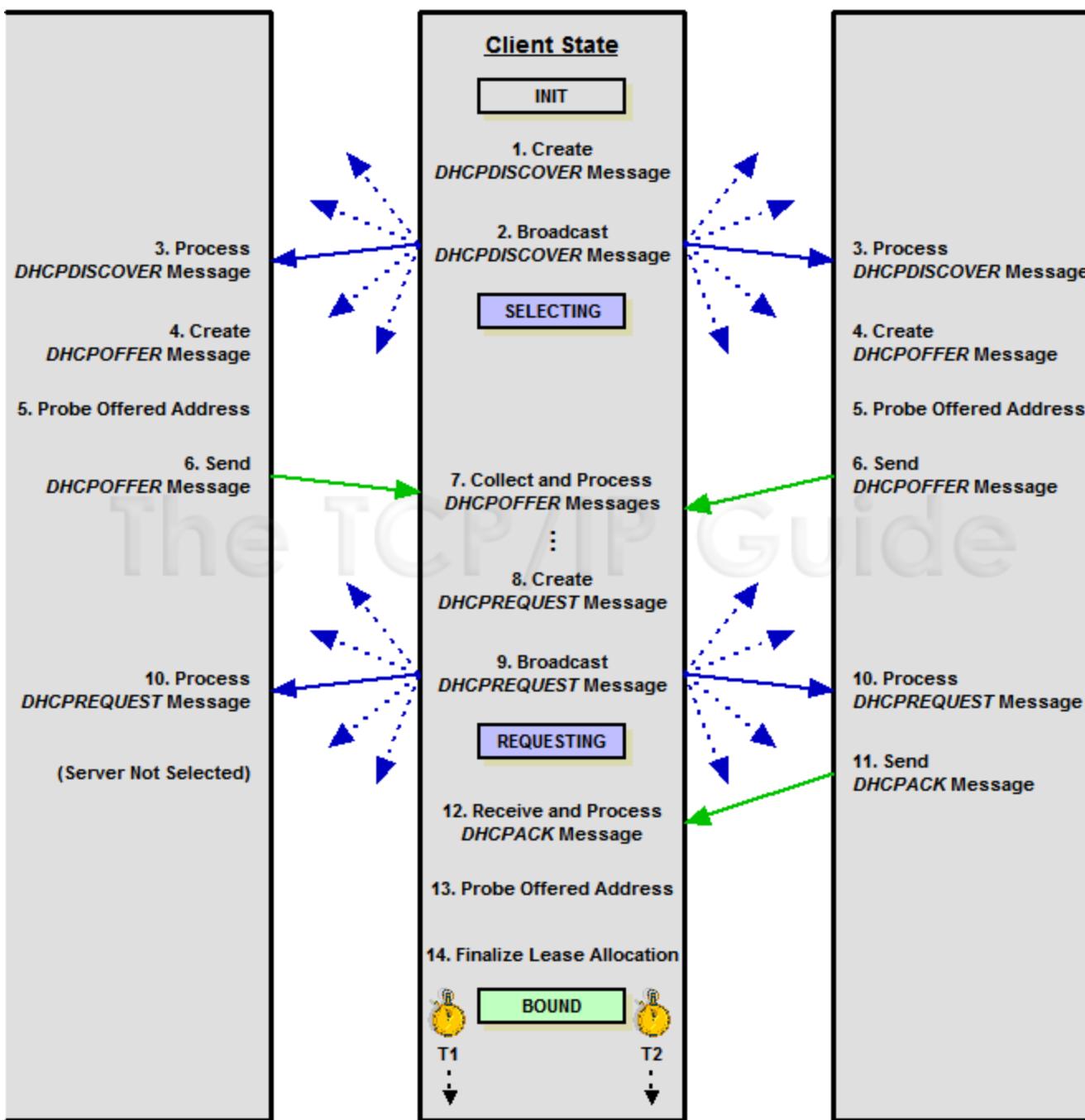
# DHCP

- Alocação (allocation) - figura no próximo slide
  - Cliente envia uma mensagem broadcast "DHCP discover" para encontrar servidores DHCP
  - Caso haja um servidor DHCP na rede (ou até mais de um) o mesmo responde com uma mensagem "DHCP offer", ofertando alocação de um endereço para o cliente
  - Entre a(s) oferta(s) recebidas(s) - de acordo com o número de servidores que responderam - o cliente solicita uma das ofertas (normalmente a primeira recebida - depende da implementação do sistema operacional) com uma mensagem "DHCP request"
  - O servidor que recebeu a requisição pela oferta então confirma a alocação através da mensagem "DHCP ack"

## Server #1

## Client

## Server #2



# DHCP

## □ Re-alocação (reallocation)

- Se um cliente é inicializado e ele já possui o alocação de um endereço, ele não precisa passar por todo o processo de alocação
  - O cliente envia uma mensagem para encontrar o servidor que possui as informações sobre sua alocação "DHCP request"
  - O servidor responde ao cliente confirmando que o alocação ainda é válido "DHCP ack"

# DHCP

## □ Renovação (renewal)

- Cada cliente tem associado com seu endereço um renewal timer (T1), normalmente setado para 50% do tempo do alocação
- Quando T1 esgotar, o cliente tentará renovar seu alocação com seu servidor DHCP antes que o mesmo expire
- Caso não consiga efetuar a renovação até que o tempo do alocação expire, o cliente entrará no processo de rebinding
  - Aqui, irá disparar um broadcast para ver se algum servidor será capaz de renovar seu endereço atual
- Por fim, caso nenhum servidor faça a renovação, o cliente terá que refazer o processo de alocação

# DHCP

## □ Liberação (release)

- Por alguma razão, um usuário pode decidir terminar com a alocação de um determinado IP
- O cliente envia então uma mensagem DHCPRelease ao servidor DHCP que mantém seu contrato
- O servidor libera o endereço IP para que o mesmo possa então ser utilizado por outro cliente

# DHCP

## □ Configuração Manual x Configuração Automática

### Configuração Manual

Endereço IP atribuído manualmente em cada computador

Possibilidade de atribuir endereços inválidos ou conflitantes

Sobrecarga de trabalho aos administradores da rede

Pouco amigável com usuários móveis

### Configuração Automática

Endereço IP atribuído dinamicamente para cada computador

Cientes sempre terão configurações de endereçamento corretas

Diminui a sobrecarga de trabalho dos administradores de rede / Administração centralizada

Mais amigável com usuários móveis

# Mapeamento de Endereços

- Mapeamento do endereço de um host em um endereço de sub-rede (muitas vezes o endereço de enlace)
  - Problema a ser resolvido pelo nível de rede
- Existem duas técnicas
  - Mapeamento direto
    - Estação sabe como computar o endereço da sub-rede
  - Resolução através da vinculação dinâmica
    - Através de um protocolo de resolução (p.e. ARP)

# Mapeamento de Endereços

- ARP (Address Resolution Protocol)
  - Endereços IP: virtuais (software)
  - Hardware não consegue localizar host utilizando o endereço IP
  - ARP
    - Resolução de Endereço: mapear endereço físico para endereço lógico

# Mapeamento de Endereços

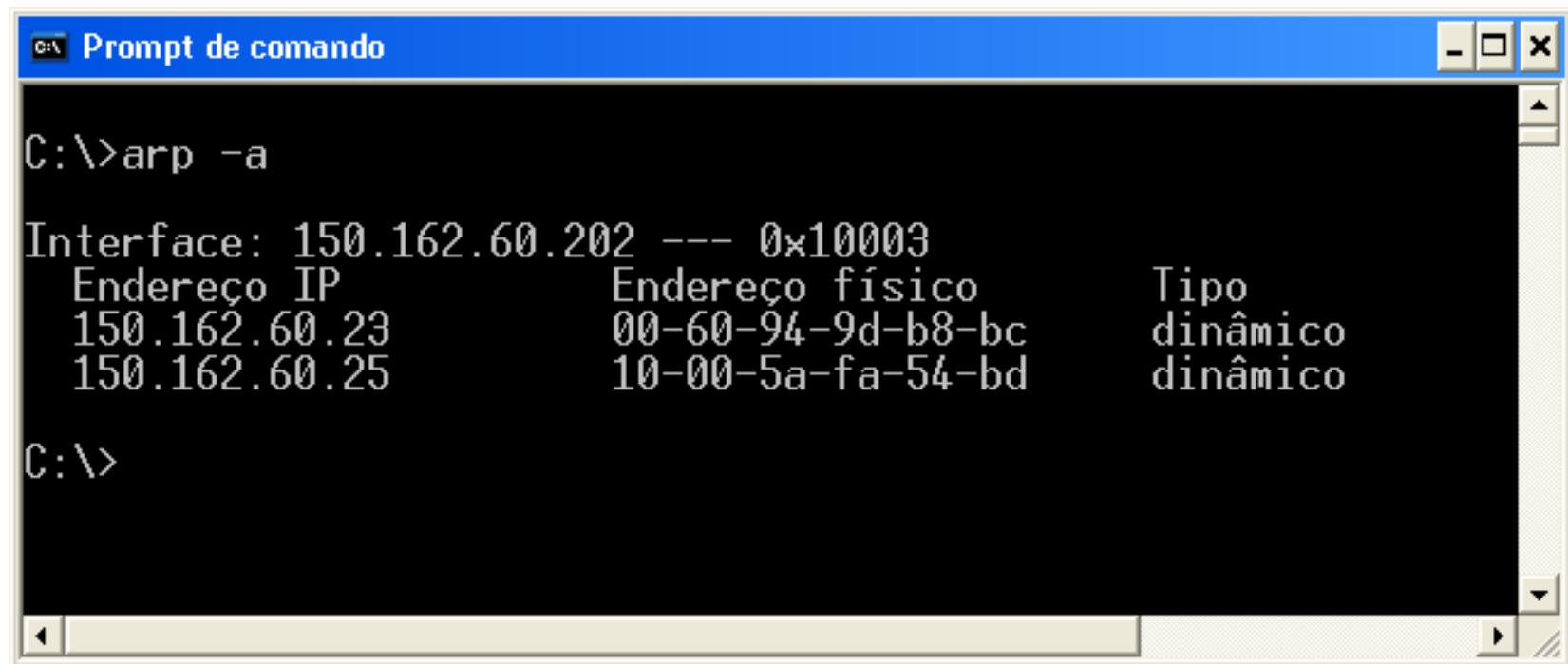
- ARP define a tradução de endereços IP em endereços físicos ETHERNET:
  - Cada adaptador ETHERNET possui endereço único na LAN;
  - ARP utiliza uma tabela de tradução:

Endereço IP	Endereço Ethernet
150.162.60.60	08:00:20:11:FD:25
150.162.60.12	08:00:20:11:FD:25
150.162.60.50	08:00:20:11:FE:F9

- Se o endereço procurado não está na tabela, então difunde um pacote, para toda a rede física, perguntando pelo endereço;
- Se alguém responde, então o endereço é inserido na tabela.

# Mapeamento de Endereços

## 1) Pesquisa em tabelas.



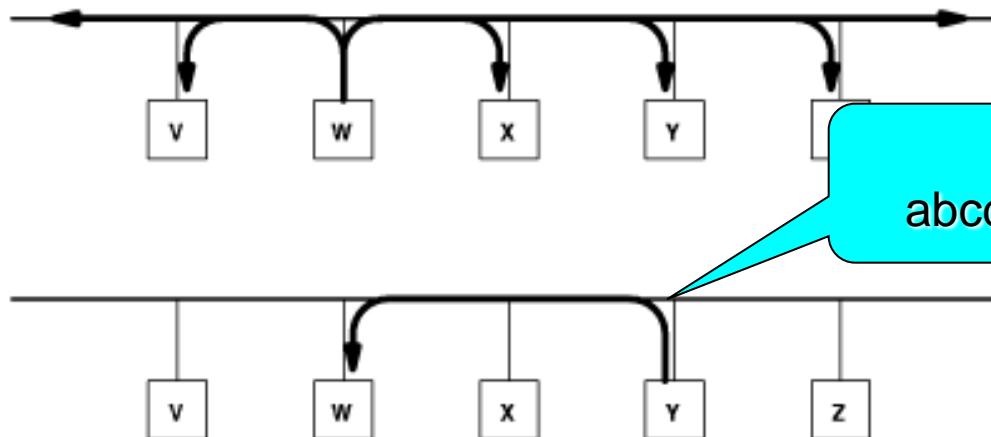
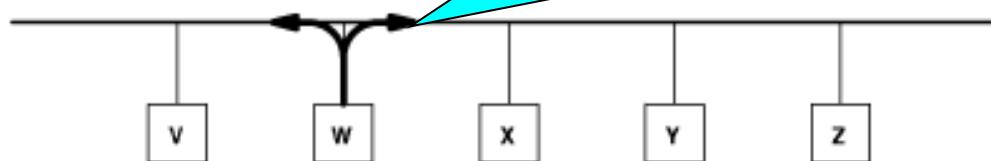
A screenshot of a Windows Command Prompt window titled "Prompt de comando". The window shows the output of the command "arp -a". The output is a table with three columns: IP Address, Physical Address, and Type. The table has a header row and two data rows. The header row contains "Endereço IP", "Endereço físico", and "Tipo". The first data row contains "150.162.60.23" and "00-60-94-9d-b8-bc", with "dinâmico" in the Type column. The second data row contains "150.162.60.25" and "10-00-5a-fa-54-bd", also with "dinâmico" in the Type column. The command prompt prompt "C:\>" is visible at the bottom.

Endereço IP	Endereço físico	Tipo
150.162.60.23	00-60-94-9d-b8-bc	dinâmico
150.162.60.25	10-00-5a-fa-54-bd	dinâmico

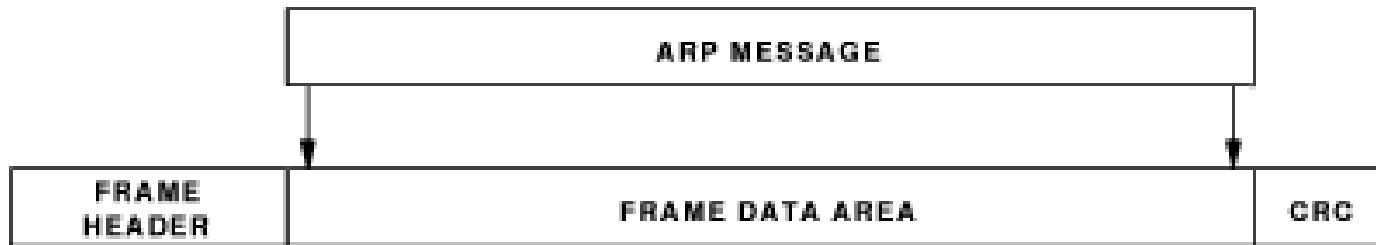
# Como fazer ?

## 2) Troca de Mensagem

Qual é o endereço  
Ethernet do Host  
x.y.w.z?



# Enviando uma mensagem ARP



Mensagem ARP encapsulada dentro de um quadro Ethernet

# Mapeamento de Endereços

- RARP (Reverse Address Resolution Protocol)
  - Usado para descobrir o endereço IP a partir do endereço de sub-rede (endereço Ethernet)

# Atribuindo endereços

- Obtendo um endereço de rede
  - O administrador da rede da organização deve contactar a provedora de serviços internet (ISP)
  - Que poderia fornecer endereços a partir de um grande bloco de endereços que teria sido alocado para a ISP
  - Se ISP tem alocado o bloco de endereço 200.23.16.0/20
  - ISP poderia dividir o bloco de endereço em 8 blocos menores de endereços para cada uma das organizações suportadas pela ISP
    - Bloco da ISP 11001000 00010111 00010000 00000000  
200.23.16.0/20
    - Organização 0 11001000 00010111 00010000 00000000  
200.23.16.0/23
    - Organização 1 11001000 00010111 00010010 00000000  
200.23.18.0/23
    - Organização 2 11001000 00010111 00010100 00000000  
200.23.20.0/23
    - ...
    - Organização 7 11001000 00010111 00011110 00000000  
200.23.30.0/23

# Atribuindo endereços

- Como uma ISP obtém seu bloco de endereços?
  - Endereços IP são gerenciados pela *Internet Corporation for Assigned Names and Numbers (ICANN)*
    - Aloca não apenas endereços IP, mas também gerenciam servidores raiz DNS
  - Atualmente endereços são gerenciados por registradores Internet regionais
    - *American Registry for Internet Number (ARIN, América do norte e do sul e parte da África)*
    - *Reseaux IP Europeans (RIPE, Europa e vizinhanças)*
    - *Asia Pacific Network Information Center (APNIC)*.

# NAT (Network Address Translator)

- Tecnologia que permite ligar uma rede com endereços IP privados (não utilizáveis na Internet) à Internet através de um servidor de NAT
  - Contornar o problema de falta de números IP
  - Traduz endereço válido em privado e vice-versa
- Endereços Privados
  - Faixas de endereços privados são definidas na RFC 1597:
    - 10.0.0.0 → 10.255.255.255
    - 172.16.0.0 → 172.31.255.255
    - 192.168.0.0 → 192.168.255.255
  - Qualquer empresa pode utilizar estas faixas
- Exemplo de uso
  - Rede com 100 computadores com esquema de endereçamento 10.10.0.0/255.255.0.0
    - Pode ter acesso à Internet usando um único endereço IP válido: o endereço IP da interface externa do NAT.
    - Uma grande economia de endereços IP!

# NAT (Network Address Translator)

- Duas formas de tradução
  - **NAT Básico**: a tradução simples de endereço IP global para endereço IP privado (sem mapeamento de portas)
  - **NAPT/PAT (Port Address Translation)**: Envolve a tradução de endereços IP e de número de portas

# NAT (Network Address Translator)

## □ Servidor NAT

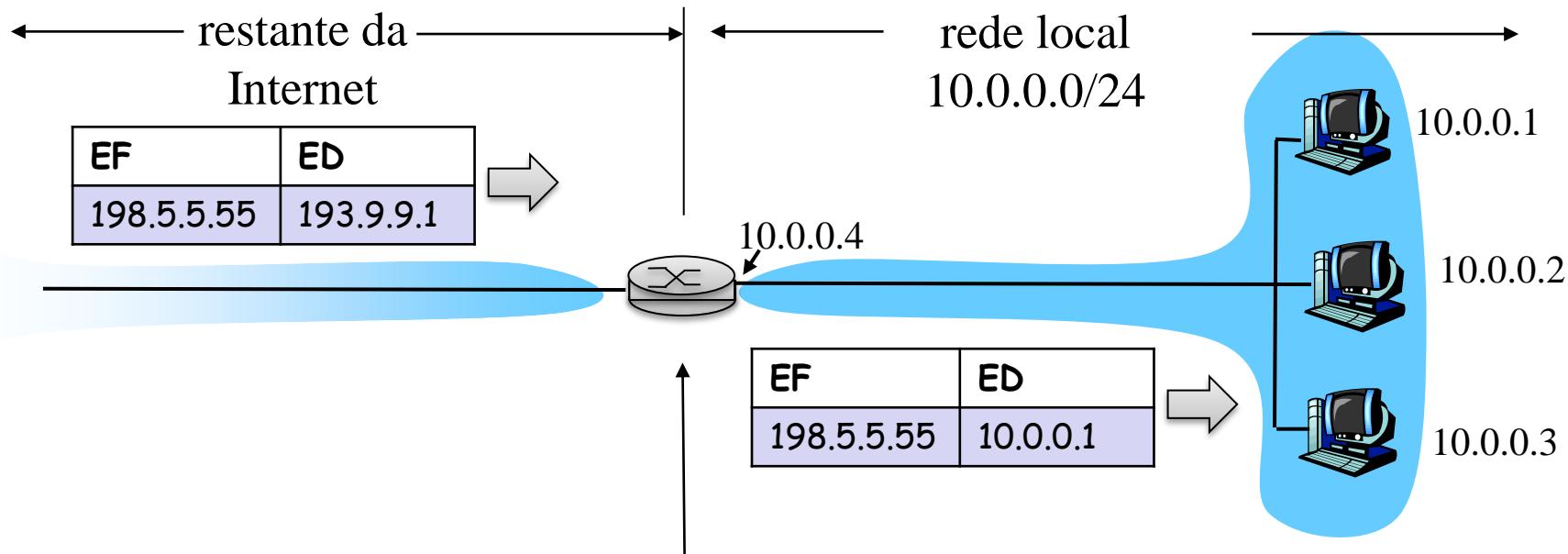
- Pode-se usar um equipamento específico (router com capacidade de NAT), um servidor Linux, ou uma máquina Windows com ICS (Internet Connection Sharing) ou outro software como Wingate ou Winroute

## □ Configuração

- IP da interface interna: 10.10.0.1 ou 192.168.0.1 ou outro reservado
- IP da interface externa: Um ou mais endereços válidos na Internet
  - obtidos a partir da conexão com o provedor de Internet

# NAT (Network Address Translator)

## □ NAT Básico



Index	Local IP (lado LAN)	Global IP (lado WAN)
1	10.0.0.1	193.99.99.1
1	10.0.0.2	193.99.99.2
1	10.0.0.3	193.99.99.3

# NAT (Network Address Translator)

## □ Tarefas e comportamentos

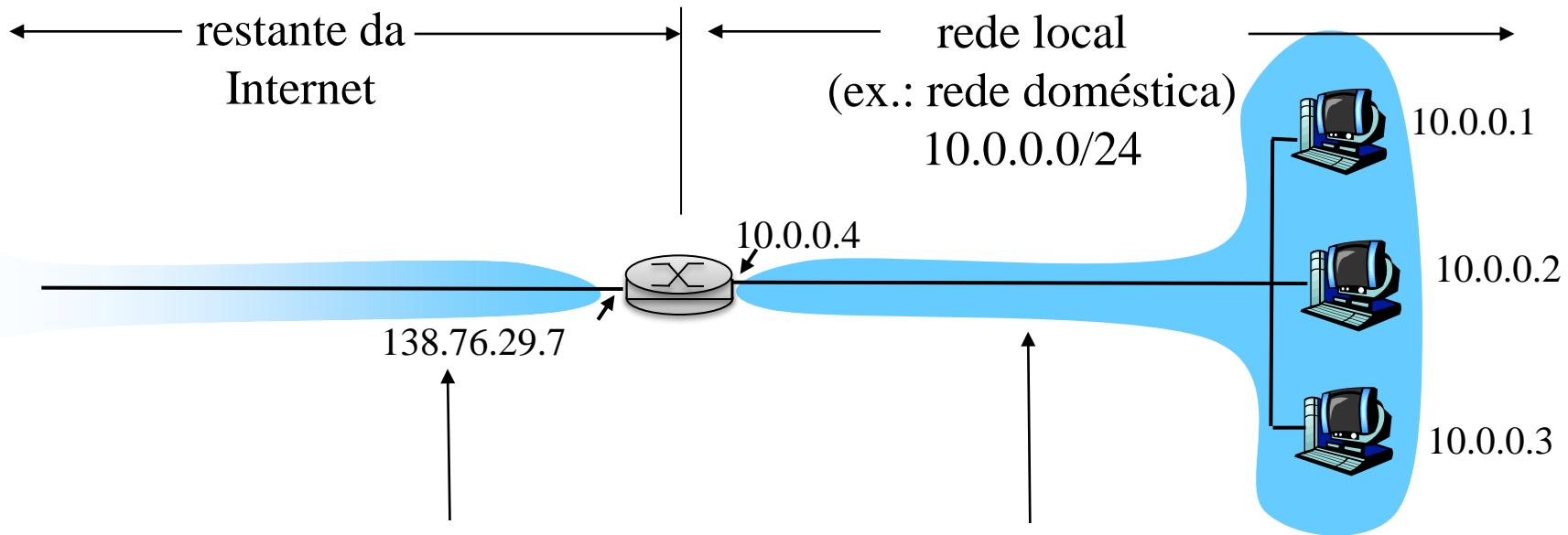
- Modificar o endereço IP de acordo com a tabela NAT
- Modificar o checksum do IP e TCP
- Modificar pacotes ICMPs
- Modificações nos campos do FTP, NetBIOS sobre TCP/IP, SNMP, DNS, Kerberos, X-Windows, SIP, H.323, Ipsec, IKE...)
- Pacotes emitidos e recebidos deveriam não ter ciência da existência do NAT

# NAT (Network Address Translator)

## □ Duas formas de tradução

- **Forma estática**: onde se estabelece uma relação entre endereços locais e endereços da Internet
  - Mapeamento estático um-para-um
- **Forma dinâmica**: onde o mapeamento de endereços locais e endereços da Internet é feito conforme a necessidade de uso
  - Um pool de endereços IP globais
  - Conexões iniciadas por hosts privados alocam um endereço IP global e a tradução é feita durante a existência da conexão
  - Estado deve ser mantido e conexões podem falhar se não existirem end. IP disponíveis

# NAPT: Exemplo



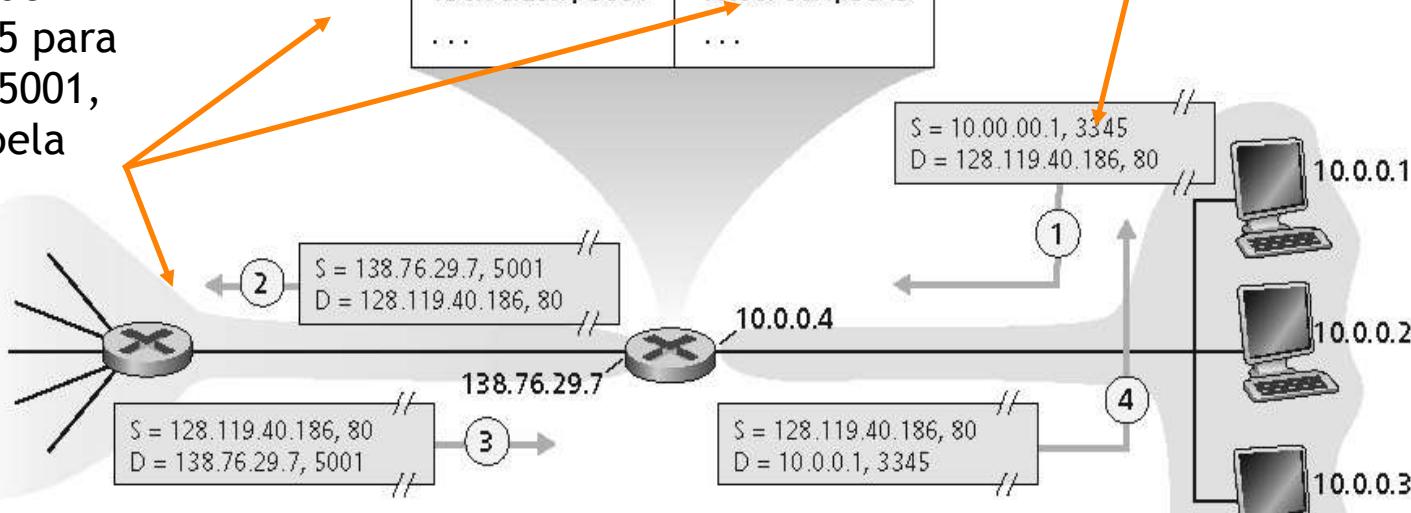
**todos os datagramas que saem da rede local possuem o mesmo e único endereço IP do NAT de origem:**  
138.76.29.7,  
números diferentes de portas de origem

datagramas com origem ou destino nesta rede possuem endereço 10.0.0.0/24 para origem, destino (usualmente)

# NAAPT

NAT translation table	
WAN side	LAN side
138.76.29.7, 5001	10.0.0.1, 3345
...	...

2: roteador NAT substitui end. origem do datagrama de 10.0.0.1, 3345 para 138.76.29.7, 5001, atualiza a tabela



# NAFT

- Quando um cliente acessa a Internet
  - Pacote contém o endereço IP da rede interna.
    - Ex.: 10.10.0.10.
    - Pacote não pode ser enviado à Internet pois endereço é privado
- Servidor NAT:
  - Datagramas que saem:
    - Substitue (endereço IP de origem, porta #) de cada datagrama para (endereço IP do NAT, nova porta #)
    - Clientes/servidores remotos responderão usando (endereço IP do NAT, nova porta #) como endereço de destino.
  - Mantem a tabela de tradução do NAT
    - Cada (endereço IP de origem, porta #) para o par de tradução (endereço IP do NAT, nova porta #).
  - Datagramas que chegam
    - substitue (endereço IP do NAT, nova porta #) nos campos de destino de cada datagrama pelos correspondentes (endereço IP de origem, porta #) armazenados da tabela NAT

# NAPT

- Campo número de porta com 16 bits:
  - 60.000 conexões simultâneas com um único endereço de LAN
- NAT é controverso:
  - Roteadores deveriam processar somente até a camada 3
    - Violação do argumento fim-a-fim
  - A possibilidade de NAT deve ser levada em conta pelos desenvolvedores de aplicações, ex., aplicações P2P
  - A escassez de endereços deveria ser resolvida pelo IPv6

# NAT (Network Address Translator)

## □ Traduções estáticas

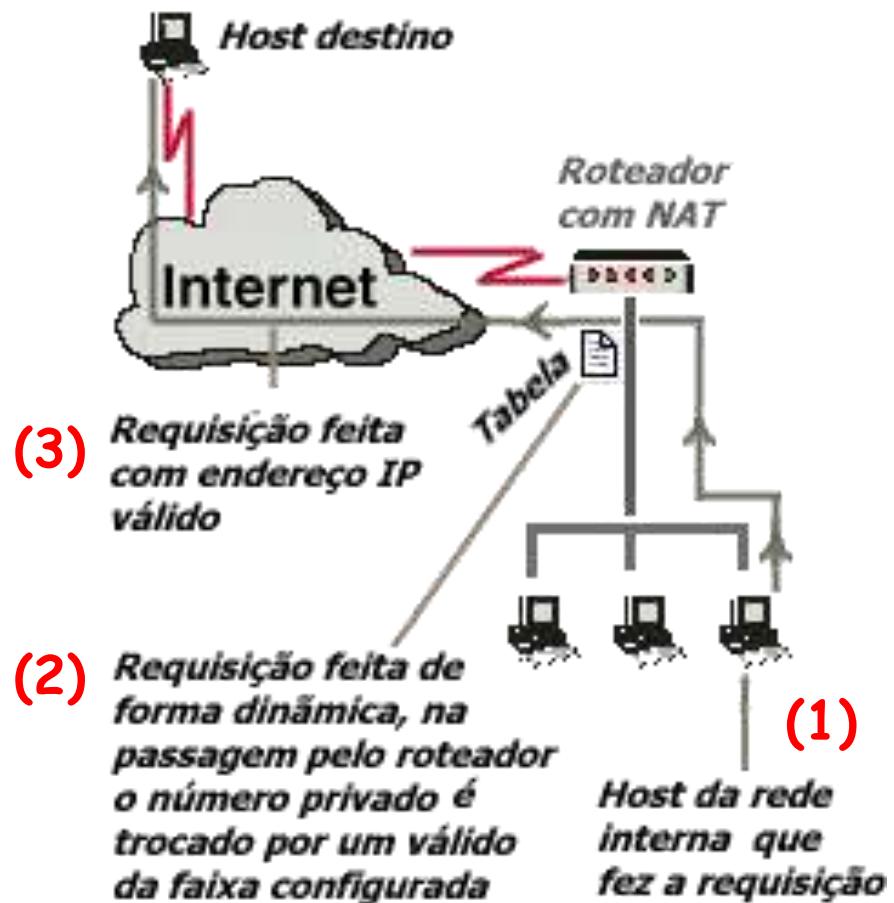
- Úteis quando disponibilizamos serviços na rede interna, como exemplo, um site Web
- quando o pedido de conexão chega ao roteador, o NAT consulta a tabela de endereços e transcreve para o IP interno correspondente, permitindo assim, que seja possível fazer uma conexão no sentido da Internet para a rede interna.



# NAT (Network Address Translator)

## □ Traduções Dinâmicas

- úteis quando se pretende dar acesso aos computadores no sentido da rede corporativa para Internet
- Endereço válido pode ser único ou uma faixa de endereços
- No retorno do pedido: NAT consulta a tabela de traduções e responde a máquina que fez a requisição.



# NAT (Network Address Translator)

## □ Outras aplicações

### ○ Balanceamento de carga

- Uso de um IP global representando um "servidor virtual"
- Na forma de várias máquinas com Ips locais
- Servidor NAT traduz endereço global em local de acordo com um algoritmo de distribuição

### ○ Alta disponibilidade

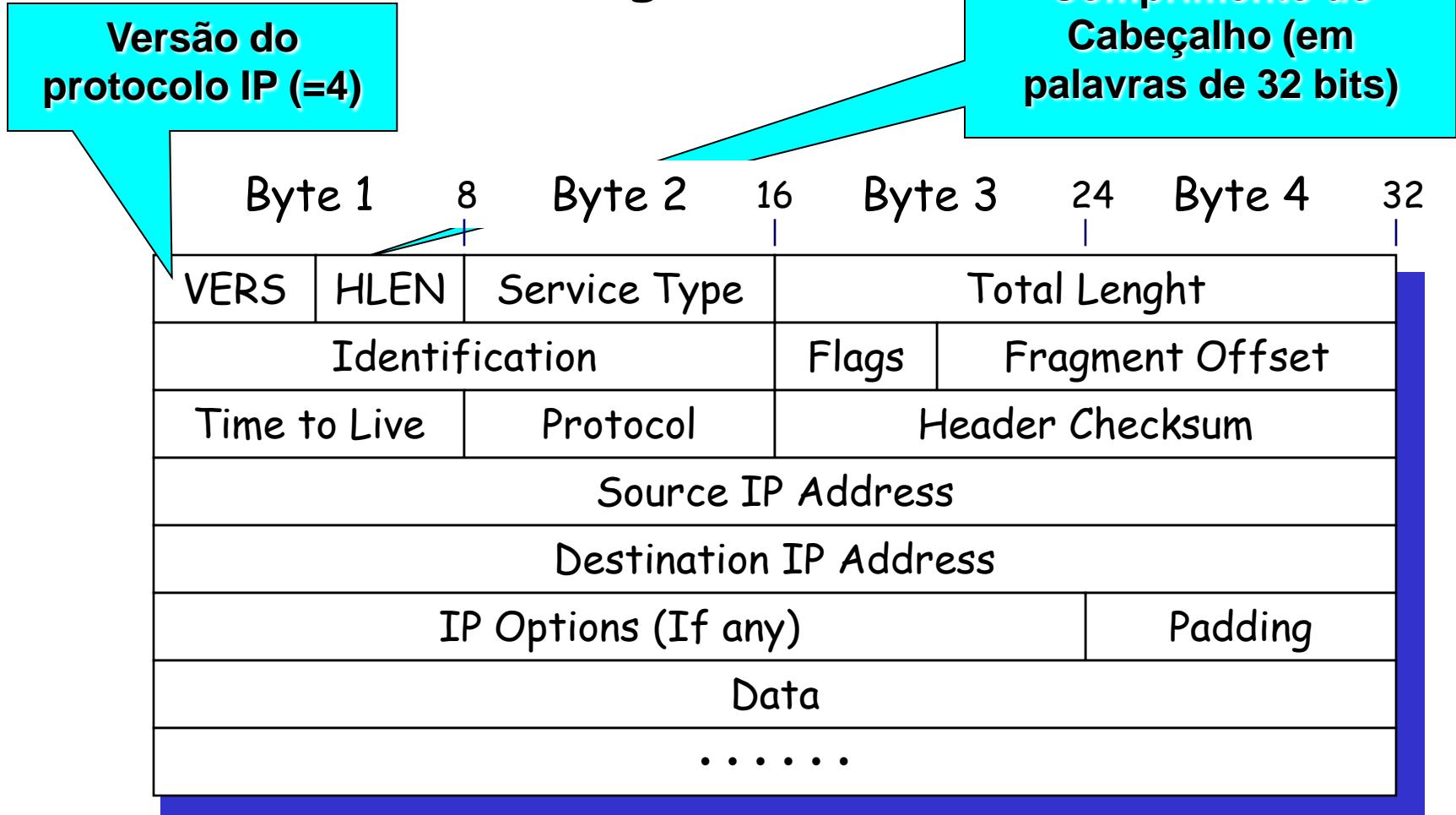
- Se uma máquina falhar, o servidor virtual ainda continua operando (com os outros servidores)

# Protocolo IP

- Comunicação não é confiável
  - Não é usado reconhecimentos
  - Não existe controle de erro
    - Exceto um checksum do cabeçalho
      - Garante que as informações usadas pelos roteadores estão corretas
- Não existe controle de fluxo e de congestionamento
- Fornece um serviço de segmentação e remontagem de datagramas longos
  - Para que eles possam ser transferidos através de redes onde o tamanho máximo (MSS) permitido para os pacotes é pequeno

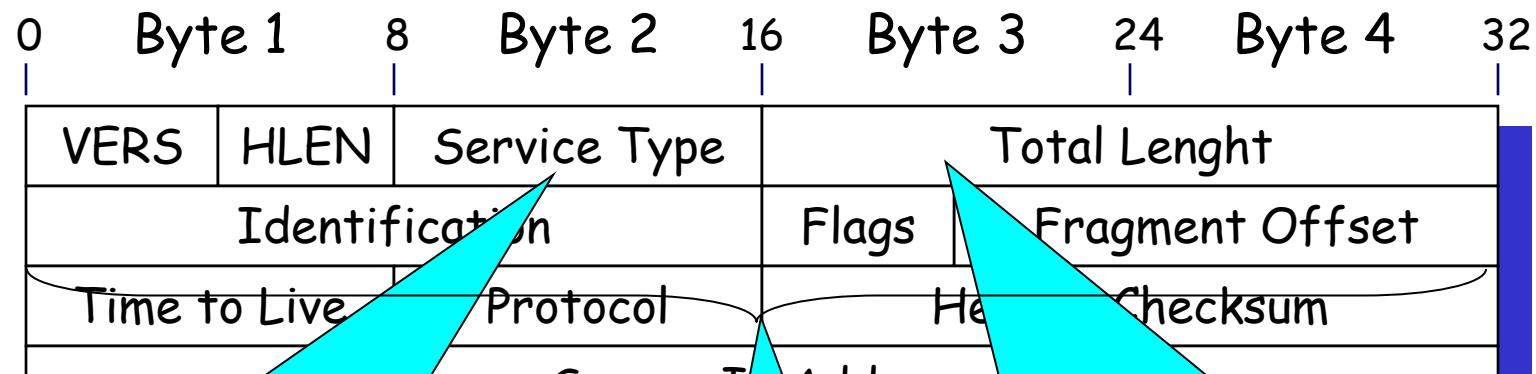
# Formato do Datagrama IP

## □ Formato do Datagrama IP



# Formato do Datagrama IP

## □ Formato do Datagrama IP



Identifica a Qualidade de Serviço a ser fornecida para o pacote

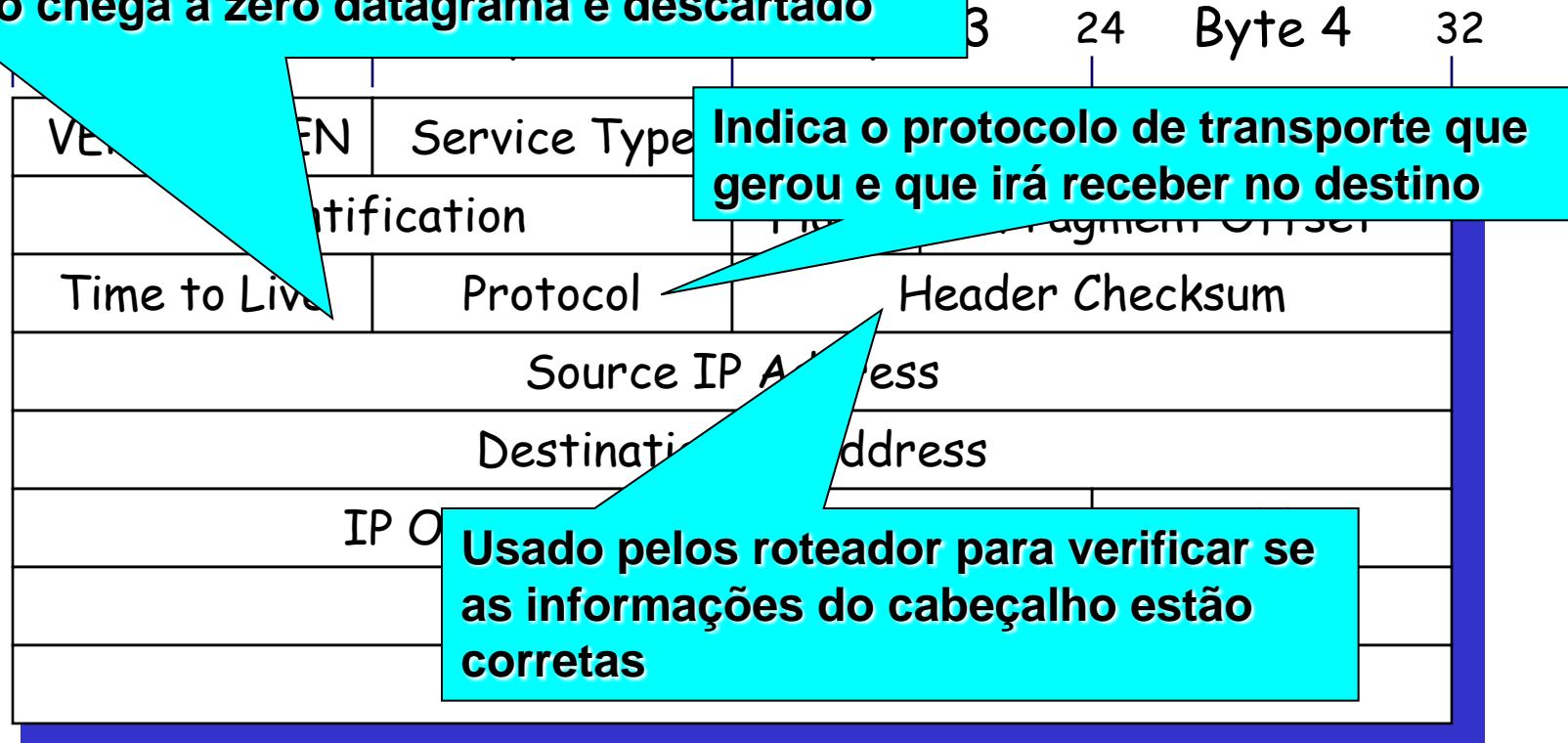
Armazena o tamanho do datagrama medido em bytes  
Tam. Max.  $2^{16} = 65536$  bytes

Usados para fragmentação e remontagem

# Formato do Datagrama IP

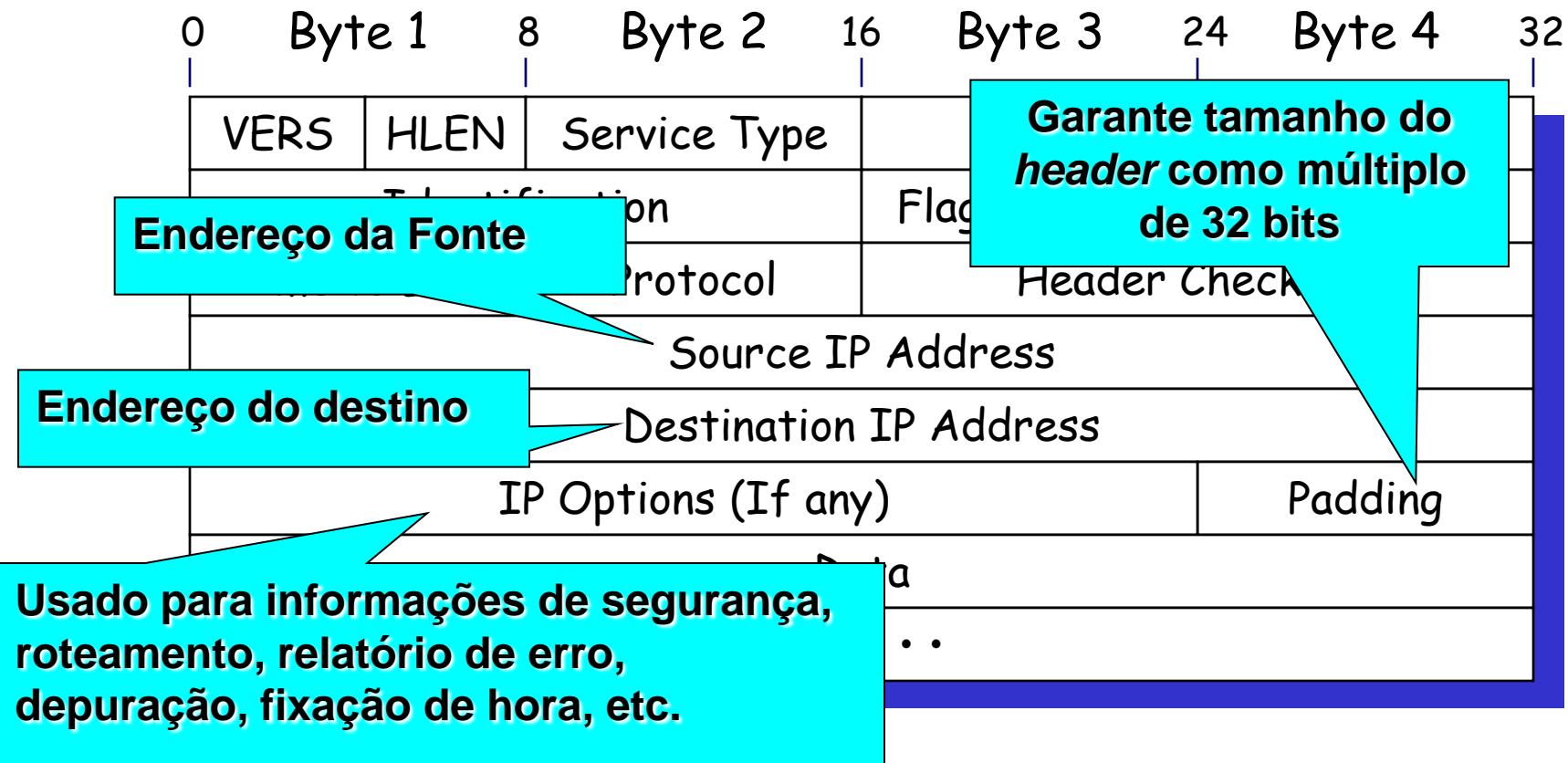
Usado para limitar o “tempo” de transmissão dos datagramas:

- Recebe um valor inicial
- Decrementado quando passa por um roteador
- Quando chega a zero datagrama é descartado

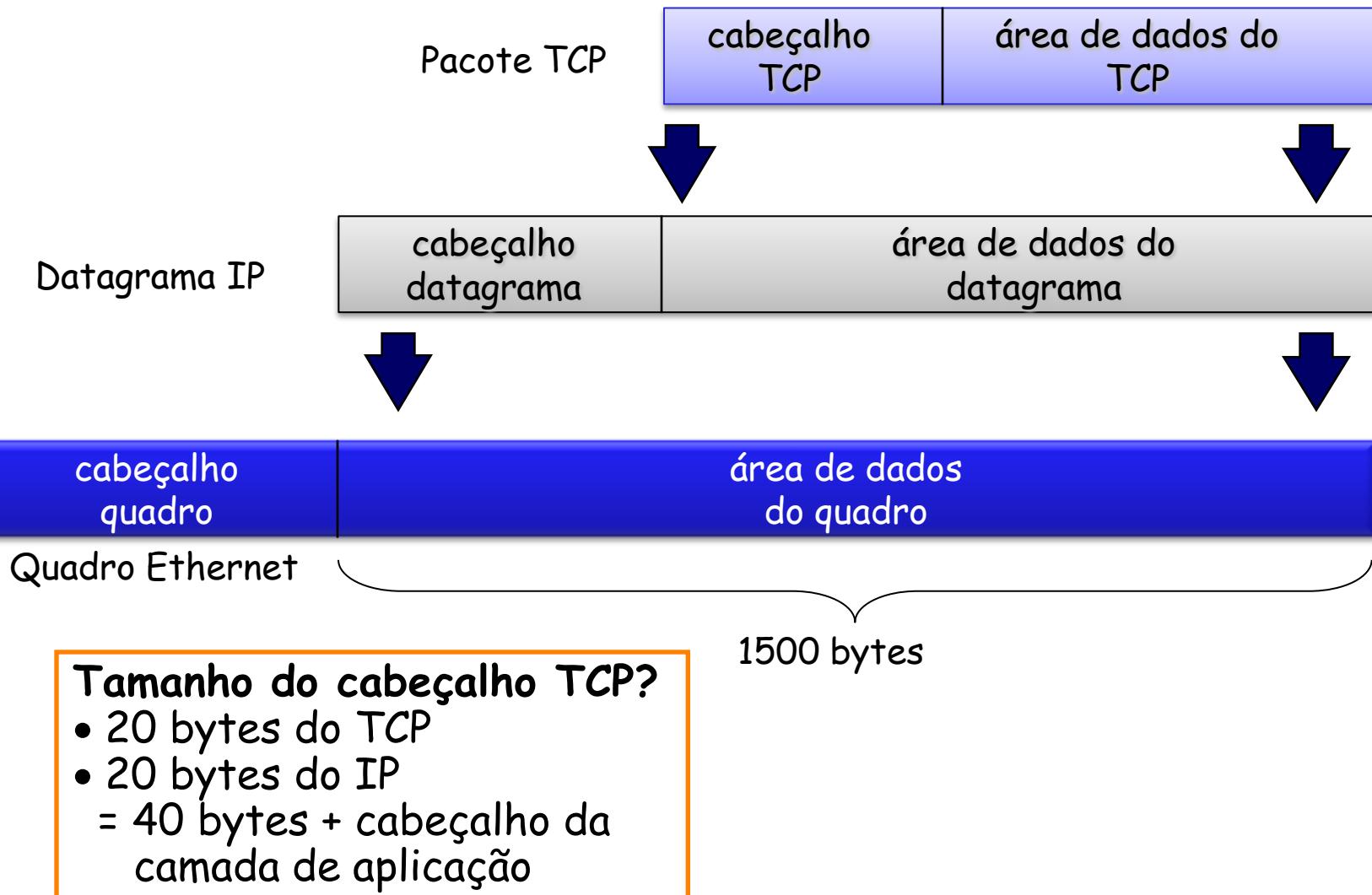


# Formato do Datagrama IP

## □ Formato do Datagrama IP

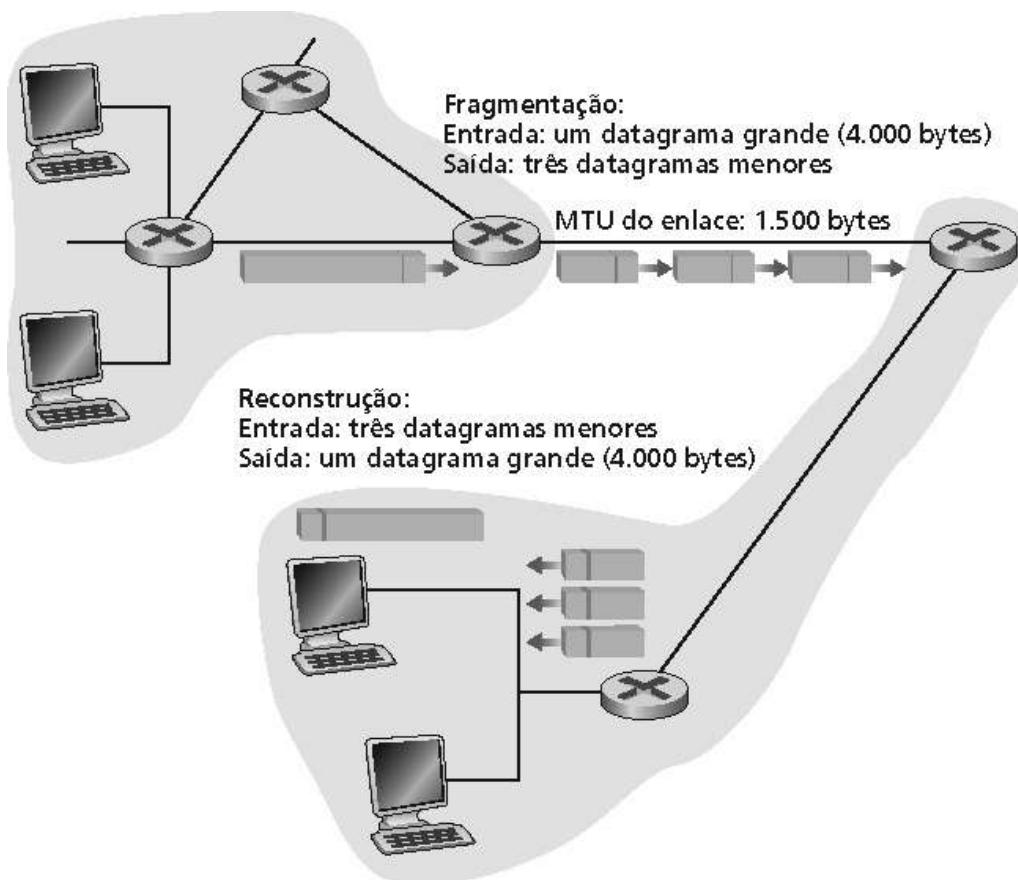


# Encapsulamento de Datagramas

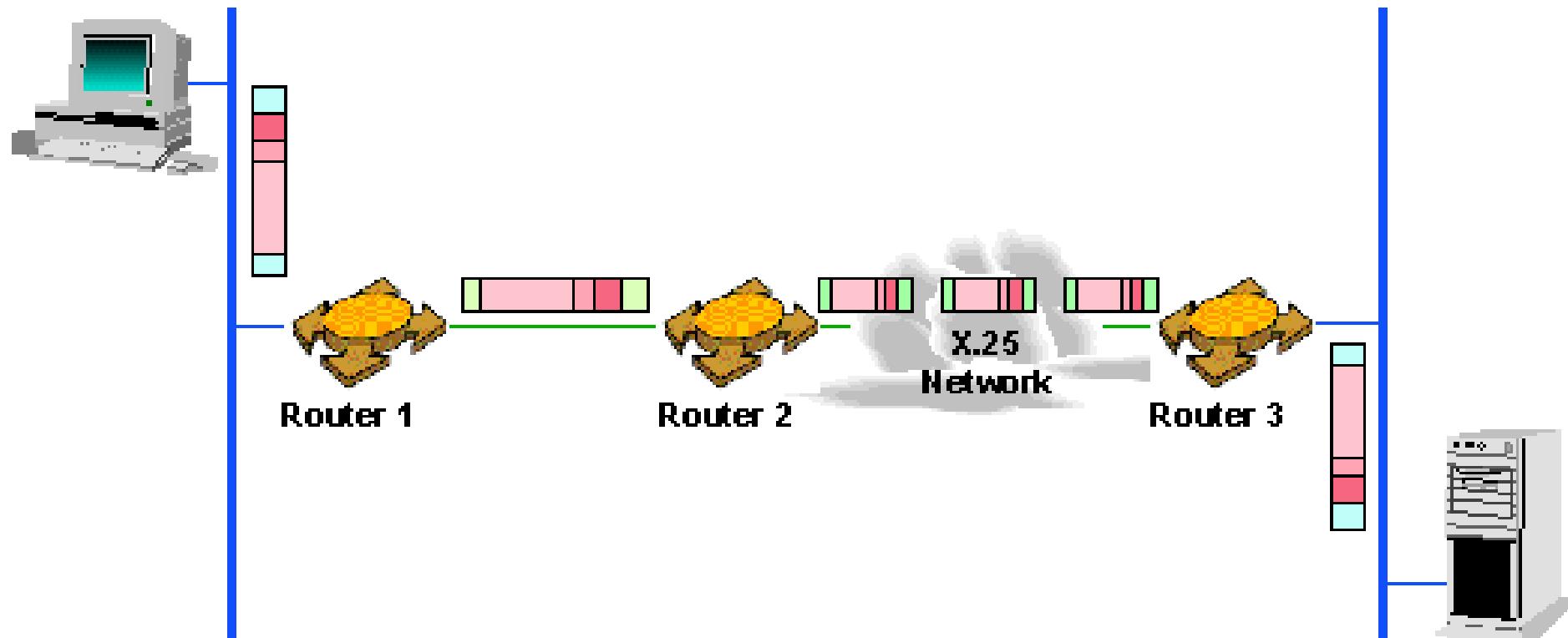


# IP Fragmentação e Remontagem

- Enlaces de rede têm MTU (max.transfer size) - corresponde ao maior frame que pode ser transportado pela camada de enlace.
  - tipos de enlaces diferentes possuem MTU diferentes (ethernet: 1500 bytes)
- Datagramas IP grandes devem ser divididos dentro da rede (fragmentados)
  - um datagrama dá origem a vários datagramas
  - "remontagem" ocorre apenas no destino final
  - O cabeçalho IP é usado para identificar e ordenar datagramas relacionados



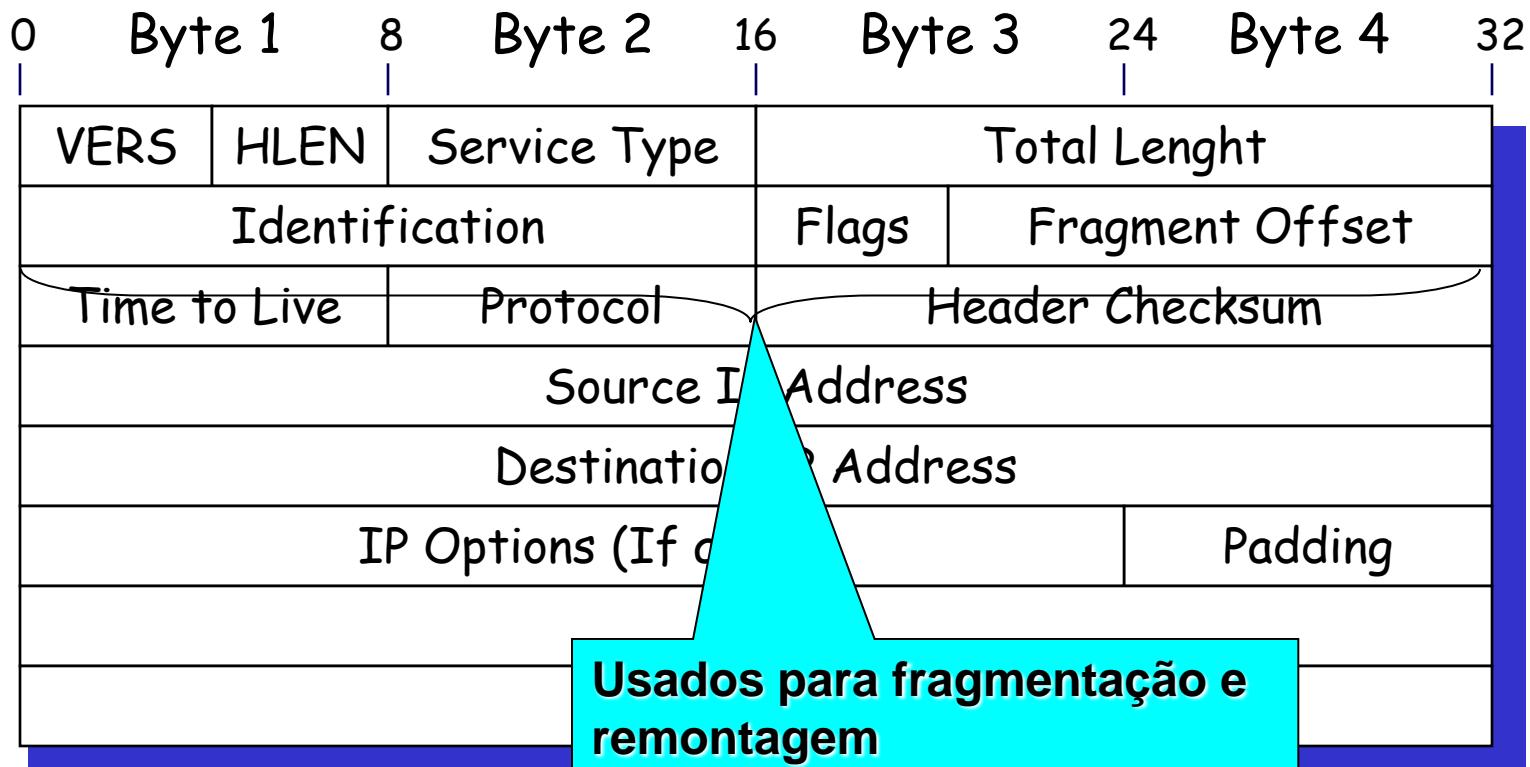
# Fragmentação de Datagramas



MTU = 620

# Formato do Datagrama IP

## □ Formato do Datagrama IP

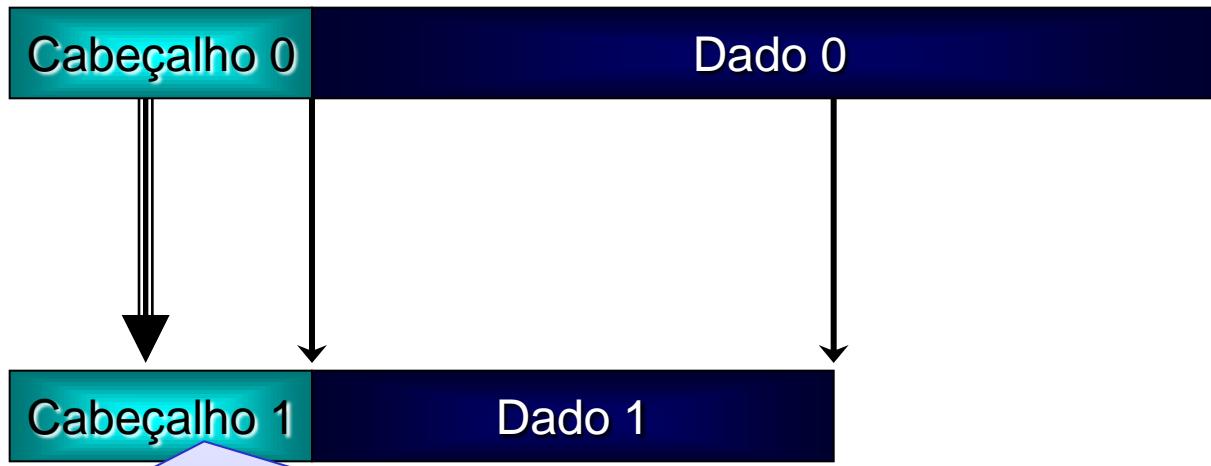


# Fragmentação e Remontagem

- Campo flags do cabeçalho IP (3 bits)
  - 1 Bit reservado
  - 1 Bit DF (don't fragment)
    - 1 indica que o pacote não pode ser fragmentado
      - Pois receptor não pode remontar
      - Roteadores tentam encontrar caminho sem fragmentação
        - » Se não conseguir o pacote é descartado
  - 1 Bit MF (more fragments)
    - 1 indica que existem mais fragmentos

# Fragmentação e Remontagem

- Exemplo: fragmentando em dois quadros

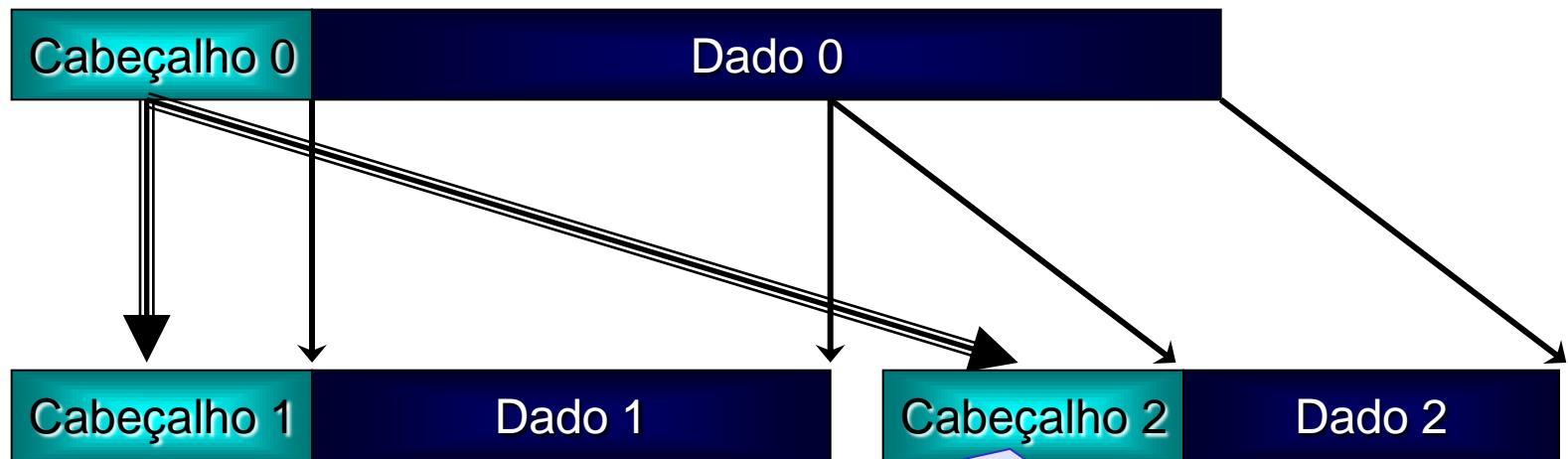


Cabeçalho:

- Copia completa do cabeçalho
  - *identification* permanece inalterado: serve para identificar todos os fragmentos de um datagrama IP
  - Fragment-offset<sub>0</sub> = 0 (se não já for fragmentado)
- Seta flag MF=1
- Campo total length é atualizado

# Fragmentação e Remontagem

- Exemplo: fragmentando em dois quadros



Cabeçalho:

- Copia completa do cabeçalho
  - *identification* permanece inalterado: serve para identificar todos os fragmentos de um datagrama IP
  - $\text{Fragment-offset}_1 = \text{fragmet-offset}_0 + \text{qtdDados}_0$
- Seta flag MF=0
- Campo total length é atualizado

# IP Fragmentação e Remontagem

	tamanho =4000	ID =x	fragflag =0	offset =0	
--	------------------	----------	----------------	--------------	--

(4000-20=3980 bytes de dados)

Um grande datagrama se torna vários datagramas menores

	tamanho =1500	ID =x	fragflag =1	offset =0	
--	------------------	----------	----------------	--------------	--

(1500-20=1480 bytes de dados)

	tamanho =1500	ID =x	fragflag =1	offset =1480	
--	------------------	----------	----------------	-----------------	--

(1500-20=1480 bytes de dados)

	tamanho =1040	ID =x	fragflag =0	offset =2960	
--	------------------	----------	----------------	-----------------	--

(1040-20=1020 bytes de dados)

# Fragmentação e Remontagem

## □ Remontagem

- Módulo IP destinatário combina os datagramas IP que possuem o mesmo valor para os campos *identification*, *protocol*, *source address* e *destination address*
  - Feita através da colocação da porção de dados de cada fragmento na posição relativa indicada pelo valor de *fragment-offset*
    - Primeiro fragmento possui o *fragment-offset* igual a 0
    - Último fragmento tem um *flag more fragments* igual a 0

# ICMP: Internet Control Message Protocol

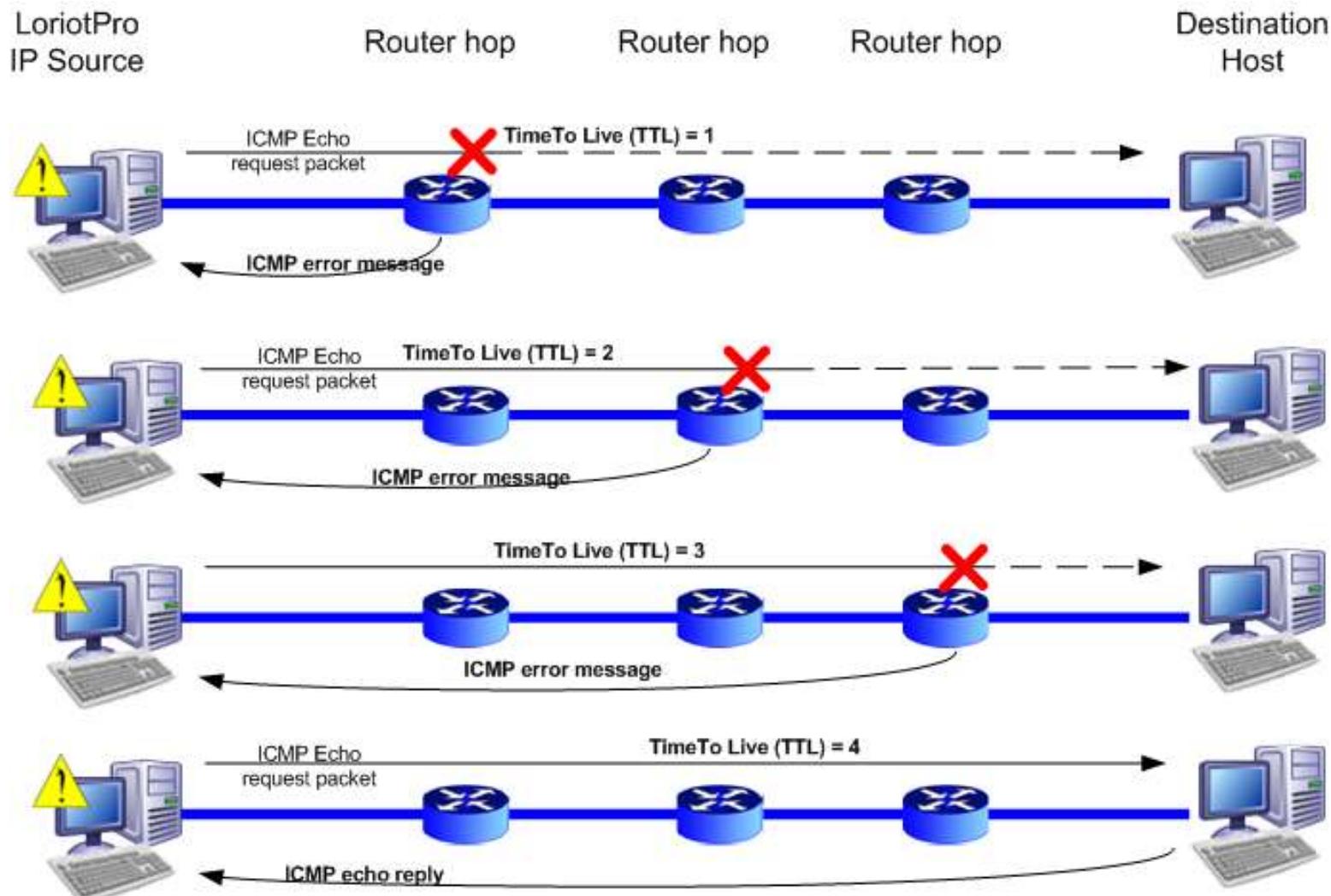
- Usado por hosts, routers, para comunicações a nível de rede
  - Relato de erros: protocolo, porta, rede, host não encontrado
  - echo request/reply (usado pelo ping)
- Msgs ICMP
  - são transportados em pacotes IP
- Mensagens ICMP: type, code mais primeiros 8 bytes do datagrama IP causando erro

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

# Traceroute e ICMP

- O transmissor envia uma série de segmentos UDP para o destino (ICMP Echo Request)
  - 1o possui TTL = 1
  - 2o possui TTL = 2
  - ... Incrementa TTL em cada tentativa até encontrar o destino
- Quando o enésimo datagrama chega ao enésimo roteador:
  - O roteador descarta o datagrama
  - E envia à origem uma mensagem ICMP (type 11, code 0)
  - A mensagem inclui o nome do roteador e o endereço IP
- Quando a mensagem ICMP chega, a origem calcula o RTT
- O traceroute faz isso três vezes
- Critério de interrupção
  - O segmento UDP finalmente chega ao hospedeiro de destino
  - O destino retorna o pacote ICMP "Echo Reply" (type 0, code 0)

# Traceroute e ICMP



# Traceroute e ICMP

```
Prompt de Comando
C:\Users\willrich>tracert www.ietf.org
Rastreando a rota para www.ietf.org [12.22.58.30]
com no máximo 30 saltos:

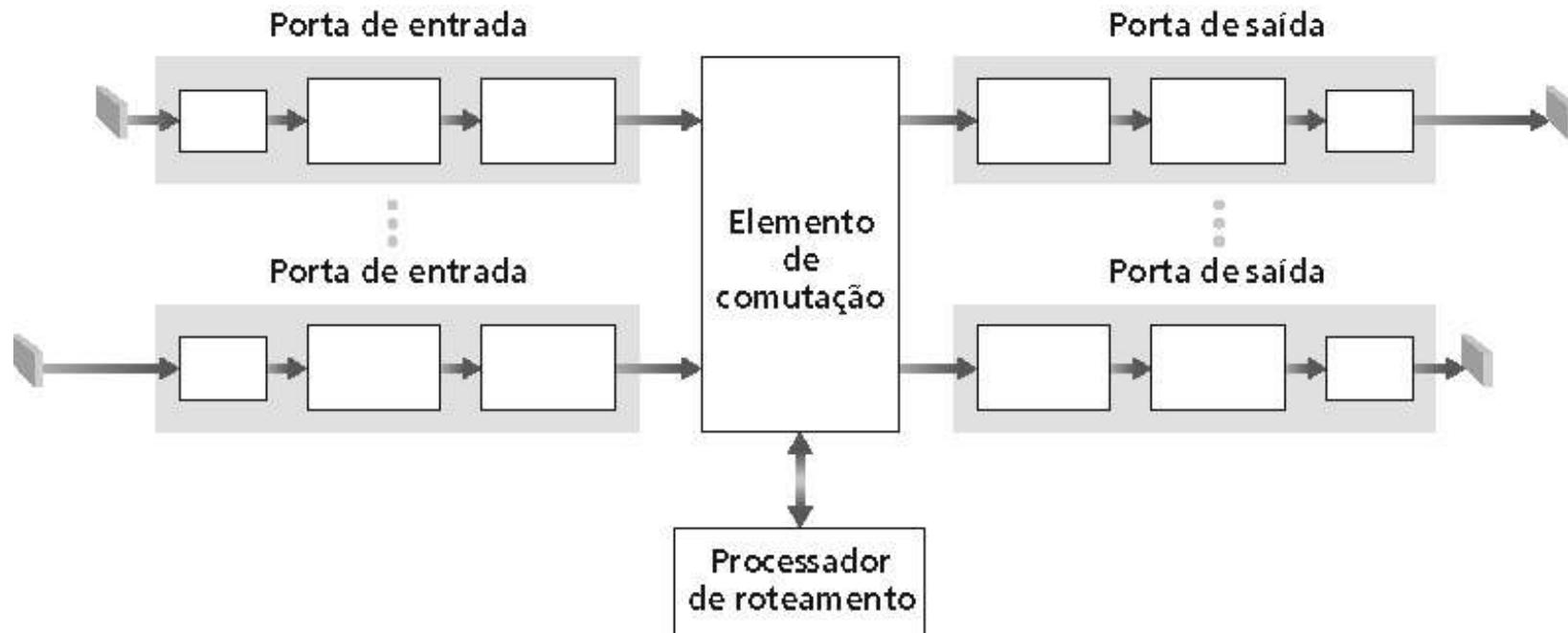
      1   16 ms    6 ms     *   150.162.239.253
      2   29 ms   25 ms   16 ms  popsc-10g-ufsc-te-1-2-rt1.bb.pop-sc.rnp.br [200.
237.194.45]
      3   30 ms   71 ms    9 ms  xe-2-0-0-2910-r0-sc.bkb.rnp.br [200.143.254.161]
      4   374 ms  183 ms  117 ms  xe-3-1-1-3000-r0-sp.bkb.rnp.br [200.143.252.65]
      5   158 ms  173 ms  179 ms  200.143.254.234
      6   133 ms  134 ms  177 ms  66.165.175.25
      7   *       168 ms    *     t0-0-0-5.br2.mia.terremark.net [66.165.161.93]
      8   *       213 ms  161 ms  12.88.168.13
      9   210 ms    *     *     cr1.ormfl.ip.att.net [12.122.143.50]
     10   247 ms    *     218 ms  cr2.hs1tx.ip.att.net [12.122.1.5]
     11   218 ms    *     *     cr1.dlstx.ip.att.net [12.122.28.157]
     12   203 ms  206 ms  218 ms  cr2.dlstx.ip.att.net [12.122.1.210]
     13   249 ms  223 ms  215 ms  cr2.la2ca.ip.att.net [12.122.28.178]
     14   205 ms  206 ms  225 ms  cr1.la2ca.ip.att.net [12.122.2.165]
     15   305 ms    *     230 ms  cr82.sj2ca.ip.att.net [12.122.1.146]
     16   243 ms  256 ms    *     gar14.sffca.ip.att.net [12.122.110.9]
     17   *       205 ms  214 ms  12.94.198.6
     18   205 ms  232 ms  220 ms  mail.ietf.org [12.22.58.30]

Rastreamento concluído.

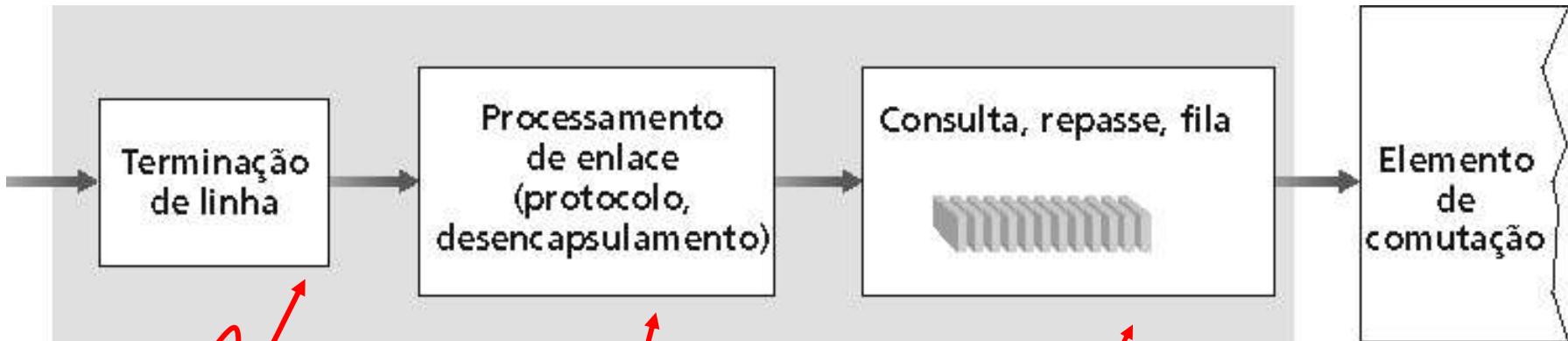
C:\Users\willrich>
```

# Arquitetura de um Roteador

- Duas funções chaves do roteador:
  - Executar algoritmos/protocolos de roteamento (RIP, OSPF, BGP)
  - Comutar datagramas de enlaces de entrada para enlaces de saída



# Funções da porta de entrada



**Camada Física:**

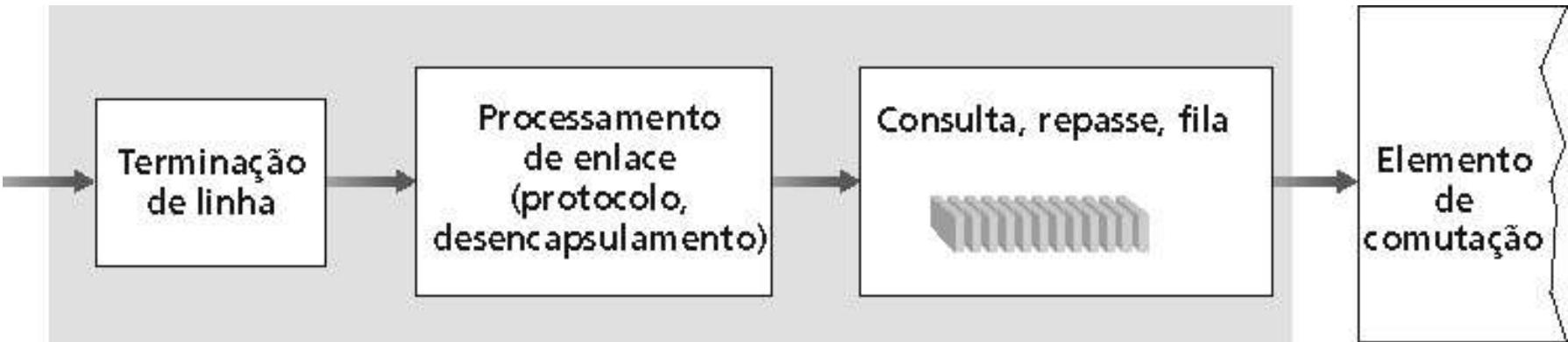
Recepção a nível de bit

**Camada de Enlace:**  
(ex.:Ethernet)

**Comutação descentralizada:**

- Dado o destino do datagrama, procura a porta de saída usando a tabela de roteamento na memória da porta de entrada
- Objetivo: completar o processamento da porta de entrada na 'velocidade da linha'
- Fila: se os datagramas chegam mais rápido do que a taxa de comutação para o switch

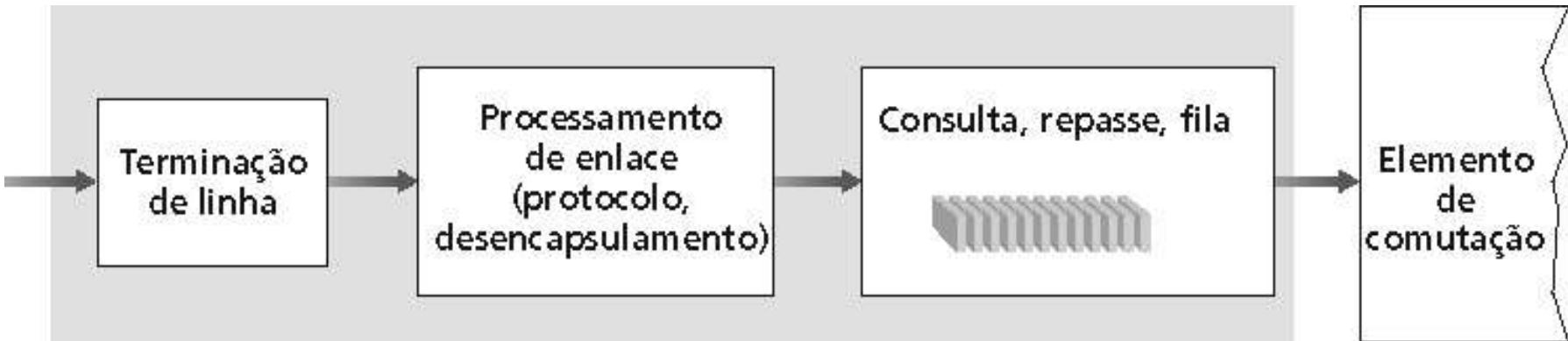
# Funções da porta de entrada



## □ Funções da fila de entrada

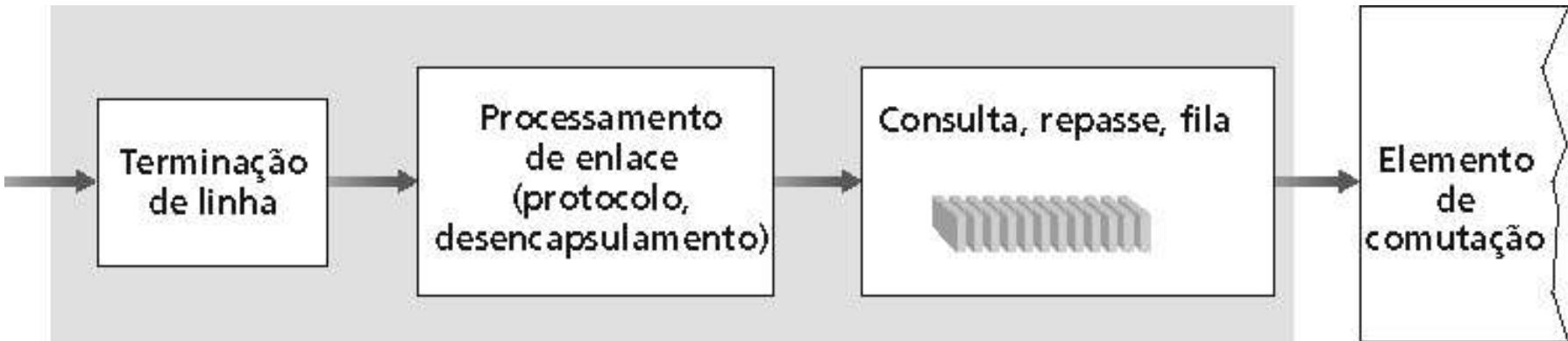
- Consulta tabela de roteamento e encaminha pacote para porta de saída
  - Escolha da porta de saída é feita usando a tabela de roteamento
  - Tabela de roteamento é computada pelo processador de roteamento e cada porta de entrada mantém uma cópia
  - Com cópias locais a decisão de comutação pode ser feita localmente sem invocar o processador centralizado
    - Tal comutação descentralizada evita a ocorrência de um gargalo de processamento em um único ponto

# Funções da porta de entrada



- Em processadores com capacidades de processamento limitadas
  - A porta de entrada pode simplesmente encaminhar o pacote para o processador de roteamento central
  - Processador central consulta a tabela e encaminha pacote para a porta de saída apropriada
  - Abordagem realizada quando uma estação de trabalho ou servidor atua como roteador
    - Processador de roteamento é a própria CPU da máquina
    - Porta de entrada é a placa de interface de rede

# Funções da porta de entrada

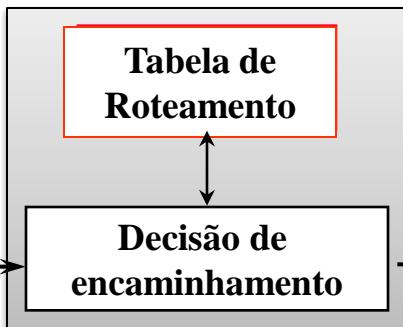


- Uma vez determinada a porta de saída apropriada
  - O pacote pode ser encaminhado via o elemento de comutação
  - Mas pacote pode ter a entrada temporariamente bloqueada
    - Se o elemento de comutação estiver ocupada enviando outro pacote
  - Pacotes bloqueados são colocados em uma fila na porta de entrada
- Pacotes de controle (RIP, OSPF ou BGP)
  - São encaminhados para o processador de roteamento

# Componentes arquiteturais básicos

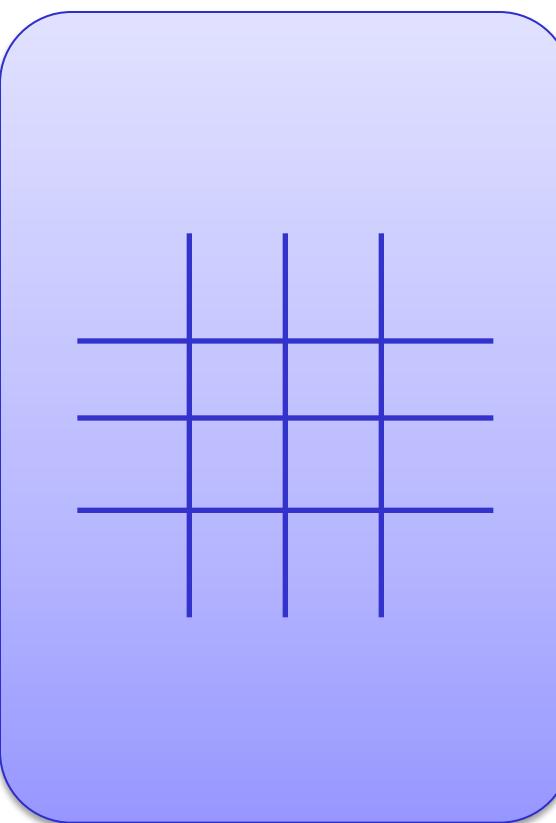
## Datapath: Processamento por pacote

1.



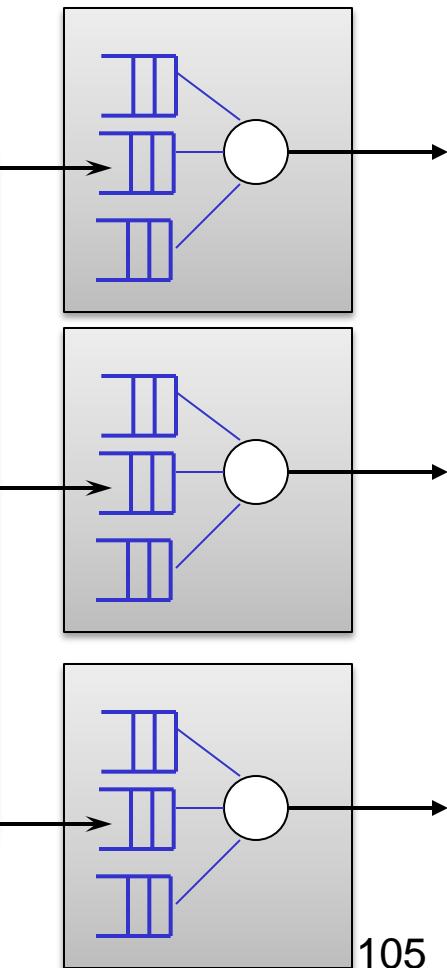
2.

Elemento de comutação



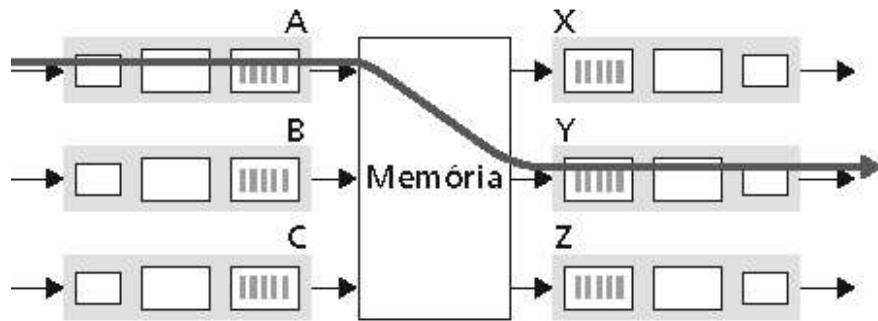
3.

Escalonamento na saída

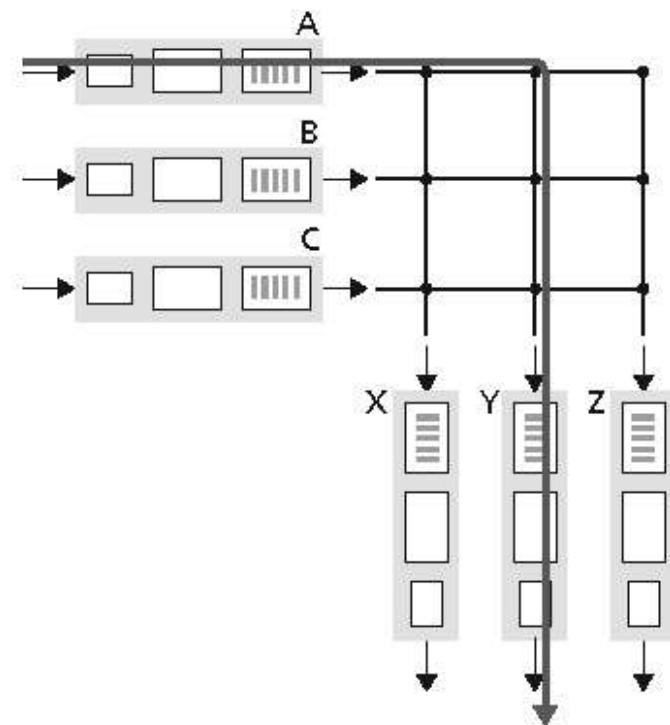


# Três tipos de Elementos de Comutação

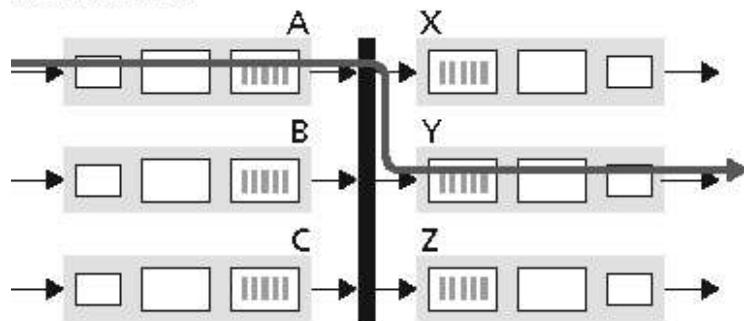
Memória



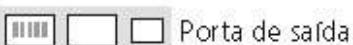
Crossbar



Barramento



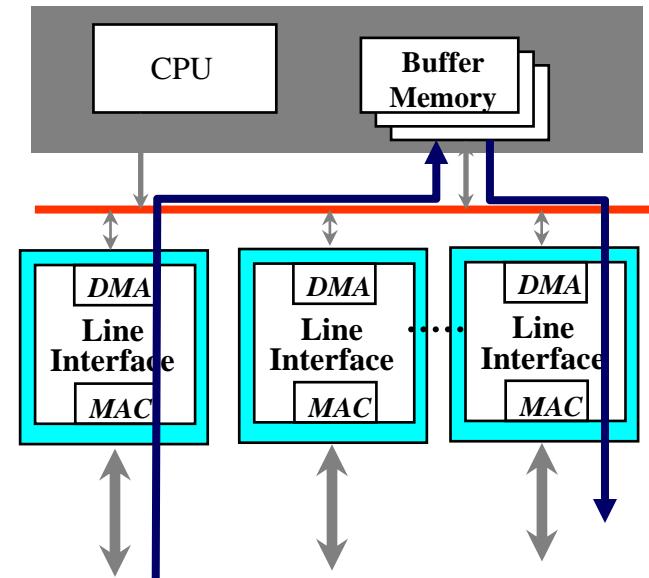
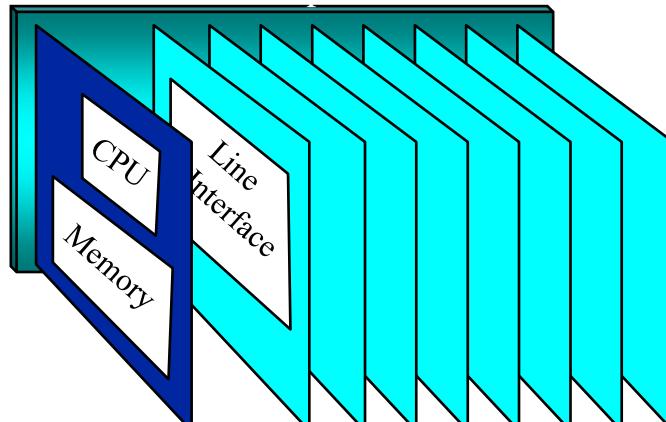
Legenda:



# Switching Via Memória

Primeira geração de roteadores:

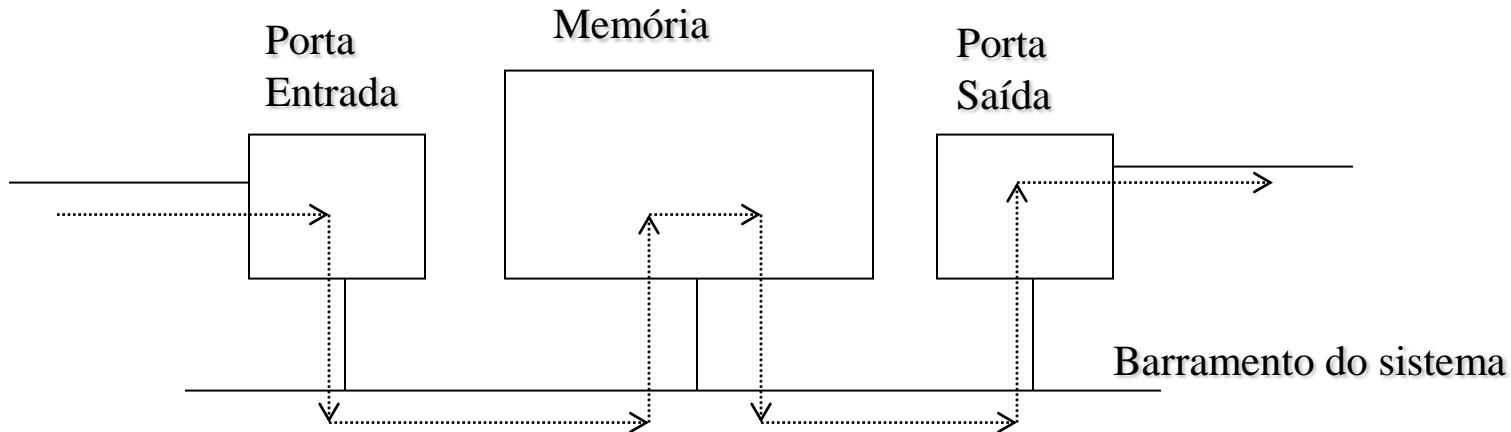
- Pacote é copiado pela CPU do sistema
- Velocidade limitada pela largura de banda da memória (2 passagens de barramento por datagrama)



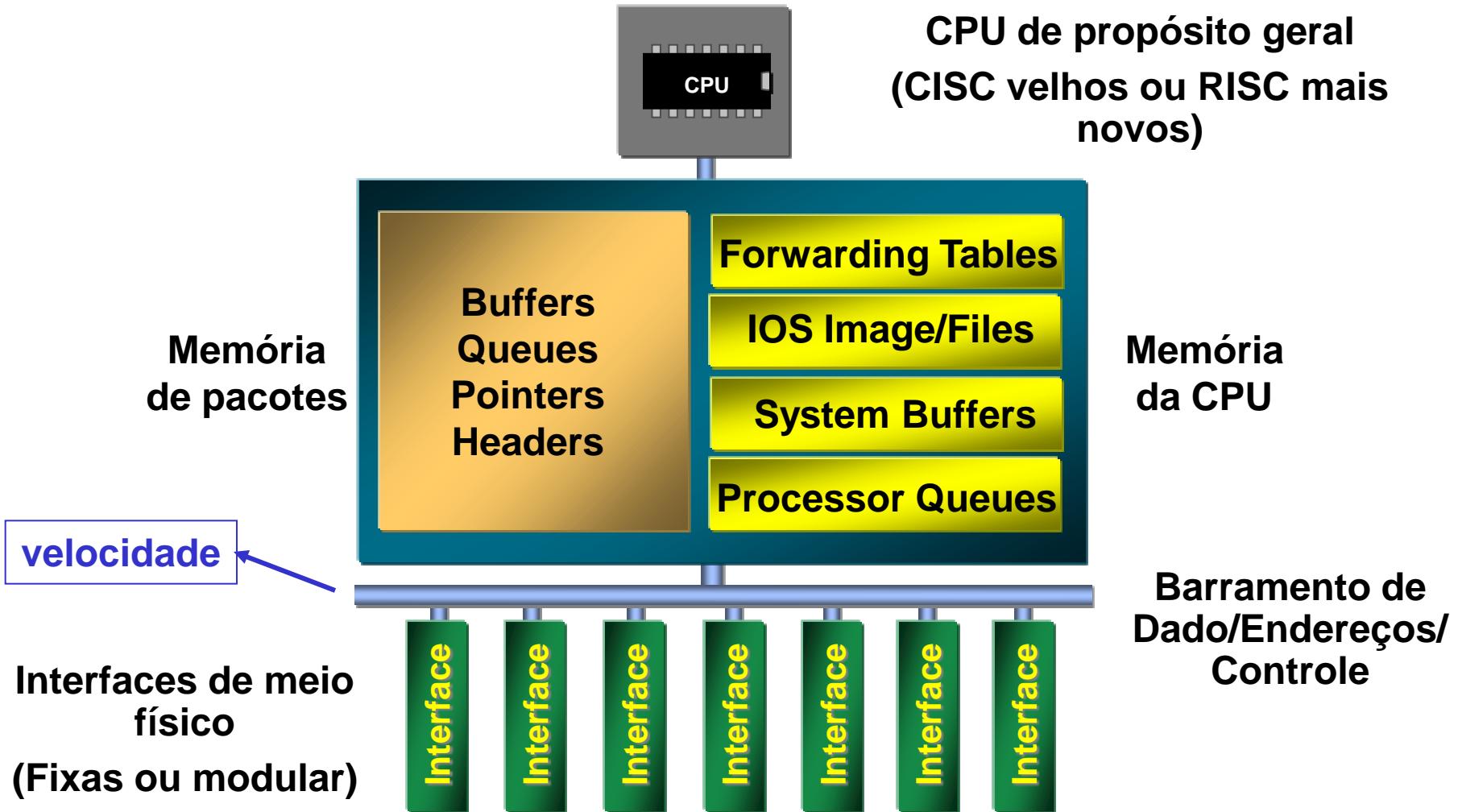
# Switching Via Memória

Primeira geração de roteadores:

- Pacote é copiado pela CPU do sistema
- Velocidade limitada pela largura de banda da memória (2 passagens de barramento por datagrama)



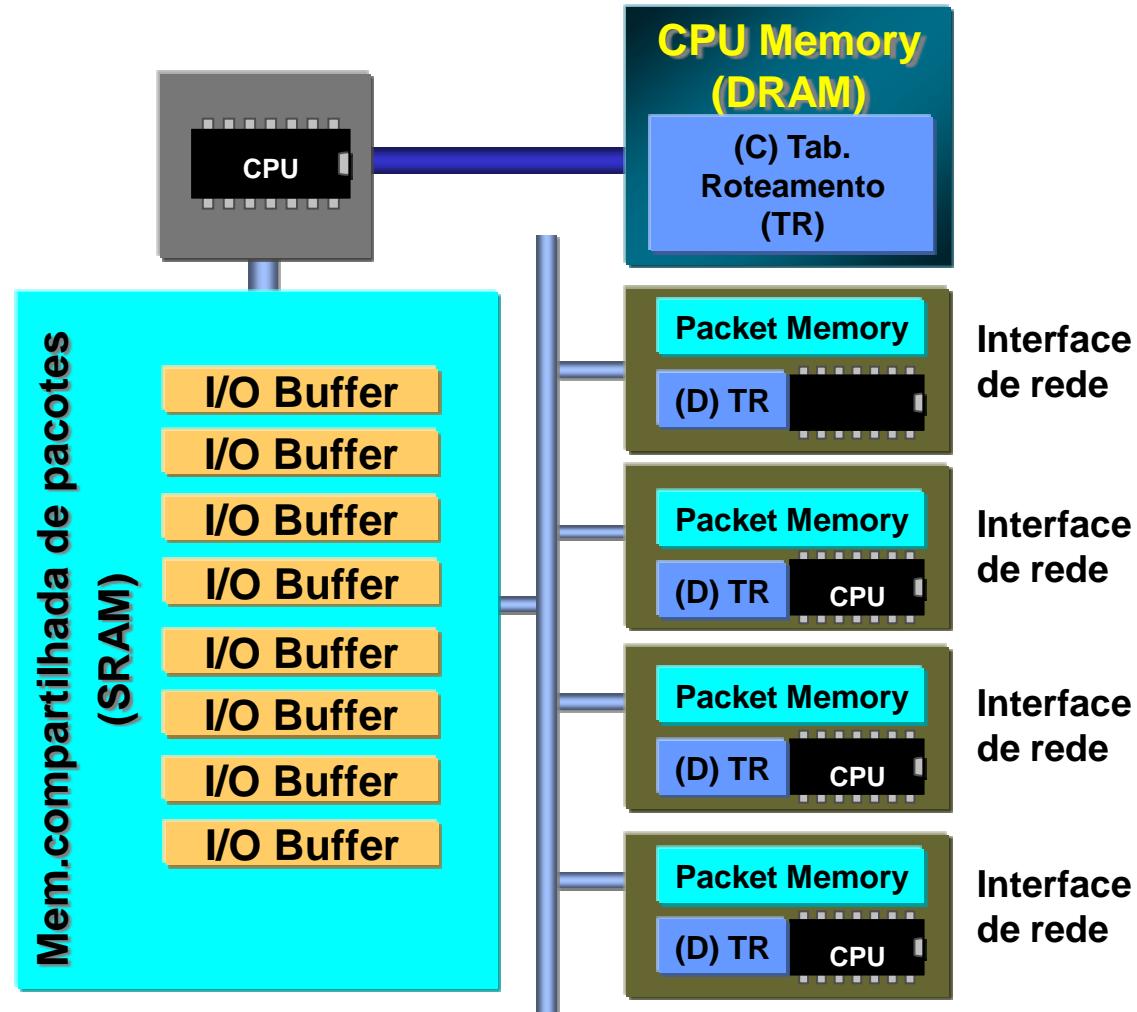
# Memória compartilhada (Processador único)



# Memória compartilhada (Memória/Processadores Distribuídos)

Cada porta tem uma memória de pacotes, memória de tabela de roteamento e uma CPU discreta

Uma cópia da tabela de roteamento é propagada do Processador de roteamento central para placas de linha para comutação local de pacotes



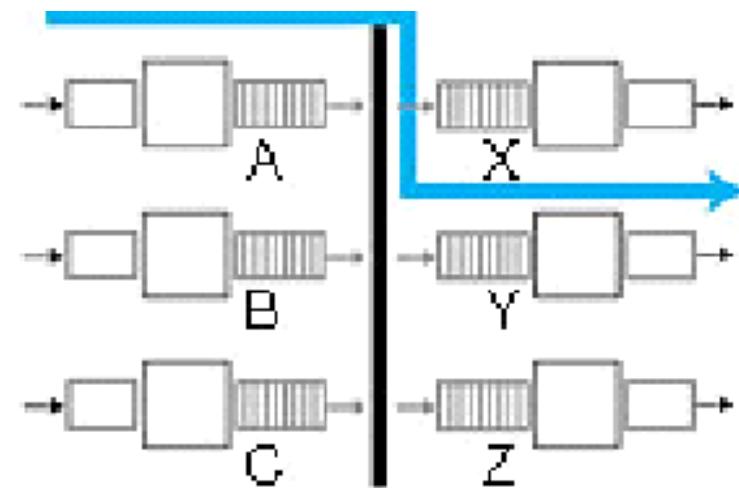
# Comutação Via Barramento

## □ Comutação

- Datagrama é encaminhado da memória da porta de entrada para a memória da porta de saída via barramento compartilhado

## □ Contenção no barramento

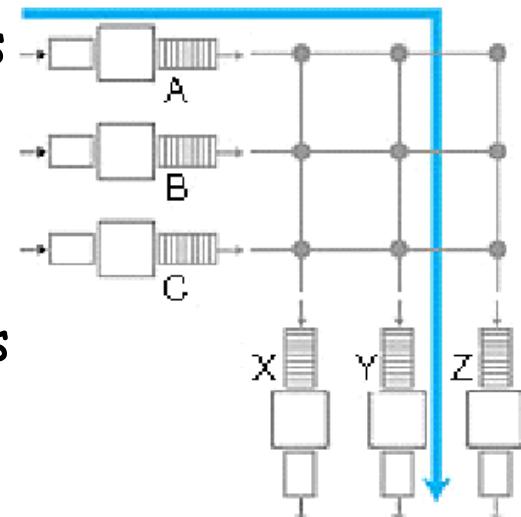
- Velocidade de comutação limitada pela velocidade do barramento
- Barramento de 1 Gbps é suficiente para roteadores de acesso e de empresas (Não para regional ou backbones)
  - Cisco 1900 é de 1 Gbps
  - 3Com CoreBuilder 5000 systems é de 2 Gbps



# Comutação Cross-bar

## (rede ou matriz de interconexão)

- Supera as limitações de largura de banda do barramento
- Redes de Banyan
  - interconexão inicialmente desenvolvidas para conectar processadores em multiprocessamento
- Portas interconectadas
  - $2N$  barramentos que conectam  $N$  portas de entrada a  $N$  portas de saída
  - Portas podem se comunicar ao mesmo tempo (com paralelismo)

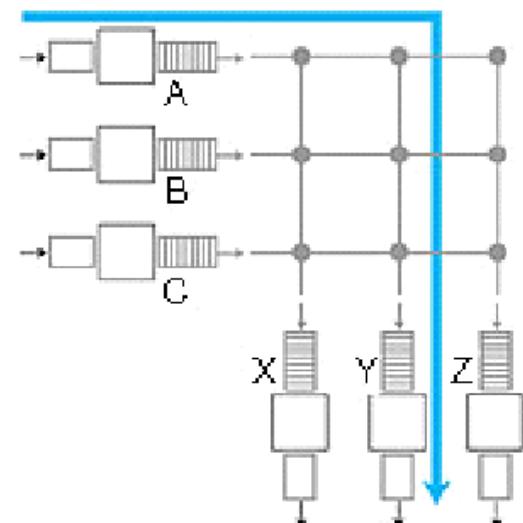


# Comutação Cross-bar

## (rede ou matriz de interconexão)

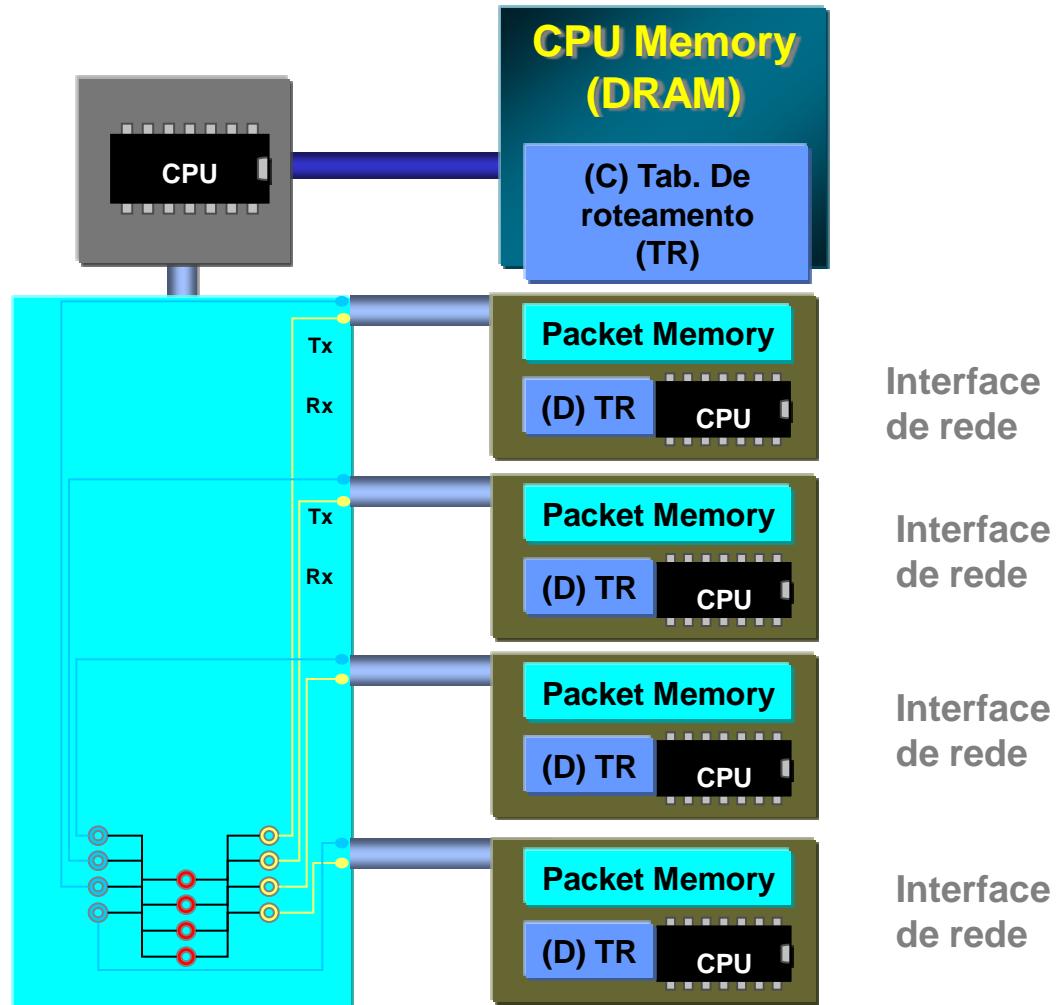
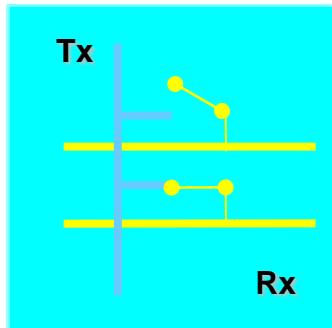
### □ Comutação

- Pacote que chega na porta de entrada atravessa o barramento horizontal ligado a porta de entrada até a intersecção com o barramento vertical levando a porta de saída
  - Se o barramento vertical é livre o pacote é transmitido para a porta de saída
  - Se ocupado o pacote é bloqueado e deve ser enfileirado na porta de entrada



# Cross Bar

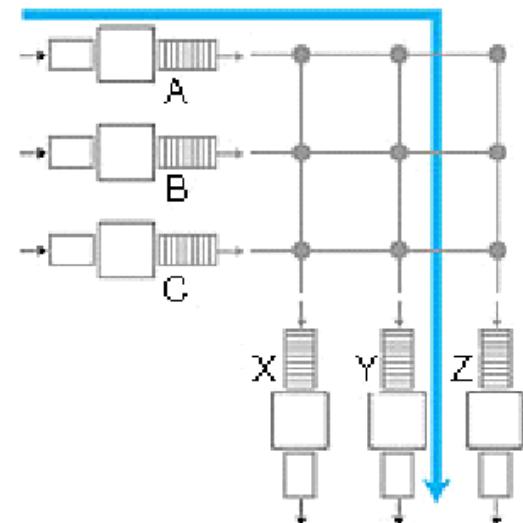
Linhas seriais de Entrada/Saída, “para” e “da” Fábrica



ASIC X-Bar Fabric

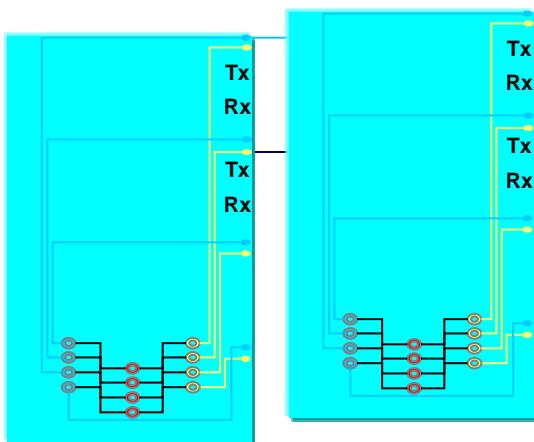
# Comutação Cross-bar (rede ou matriz de interconexão)

- Projeto avançado:
  - fragmentar datagramas em células de tamanho fixo,
    - comutar as células através do switch.
- Exemplo
  - Família de roteadores Cisco 12000
    - Usam uma rede interconectada fornecendo até 60 Gbps

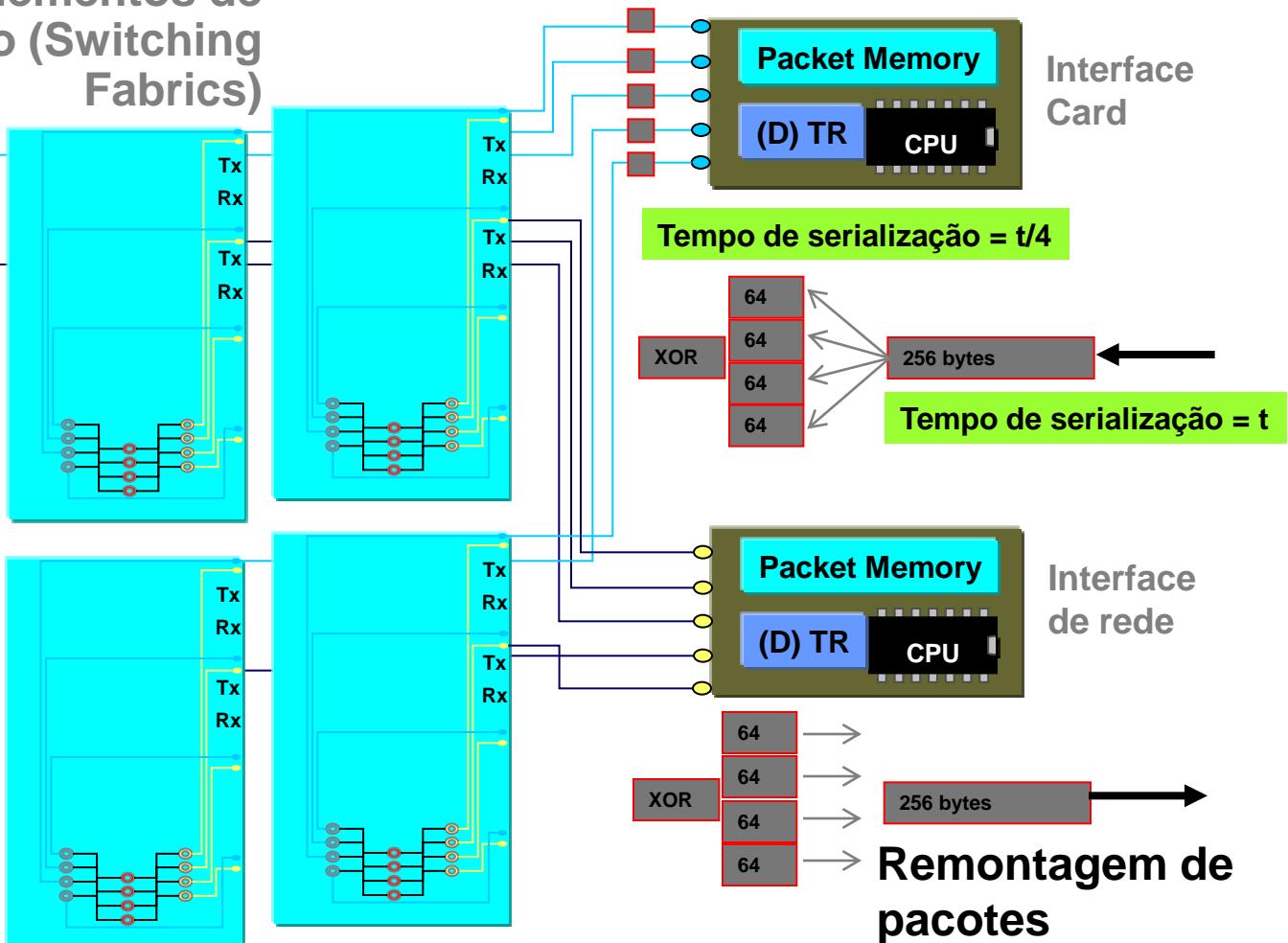


# Cross Bar (Multiple Fabrics)

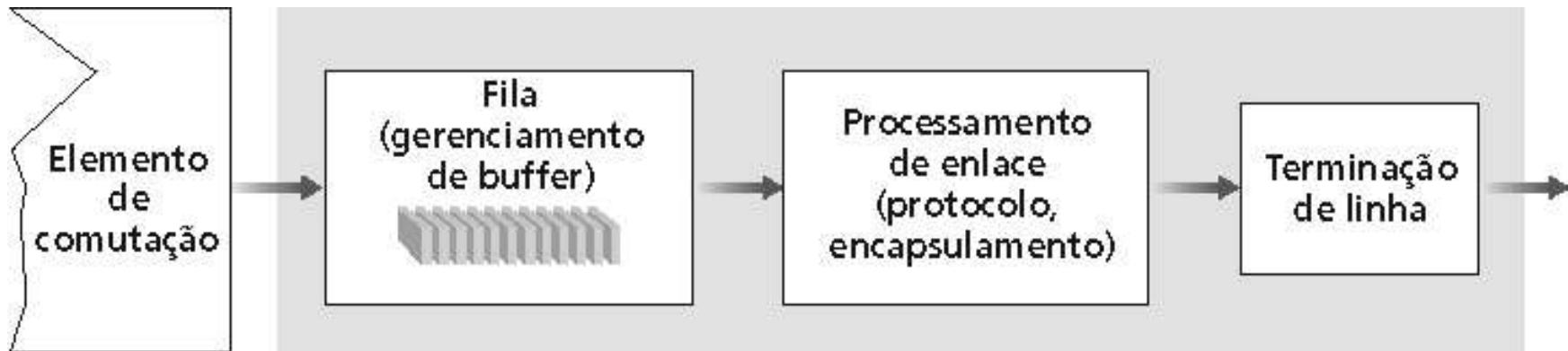
Fragmentação de pacotes permite vários elementos de comutação (Switching Fabrics)



Pacotes cortados em células (64Bytes)



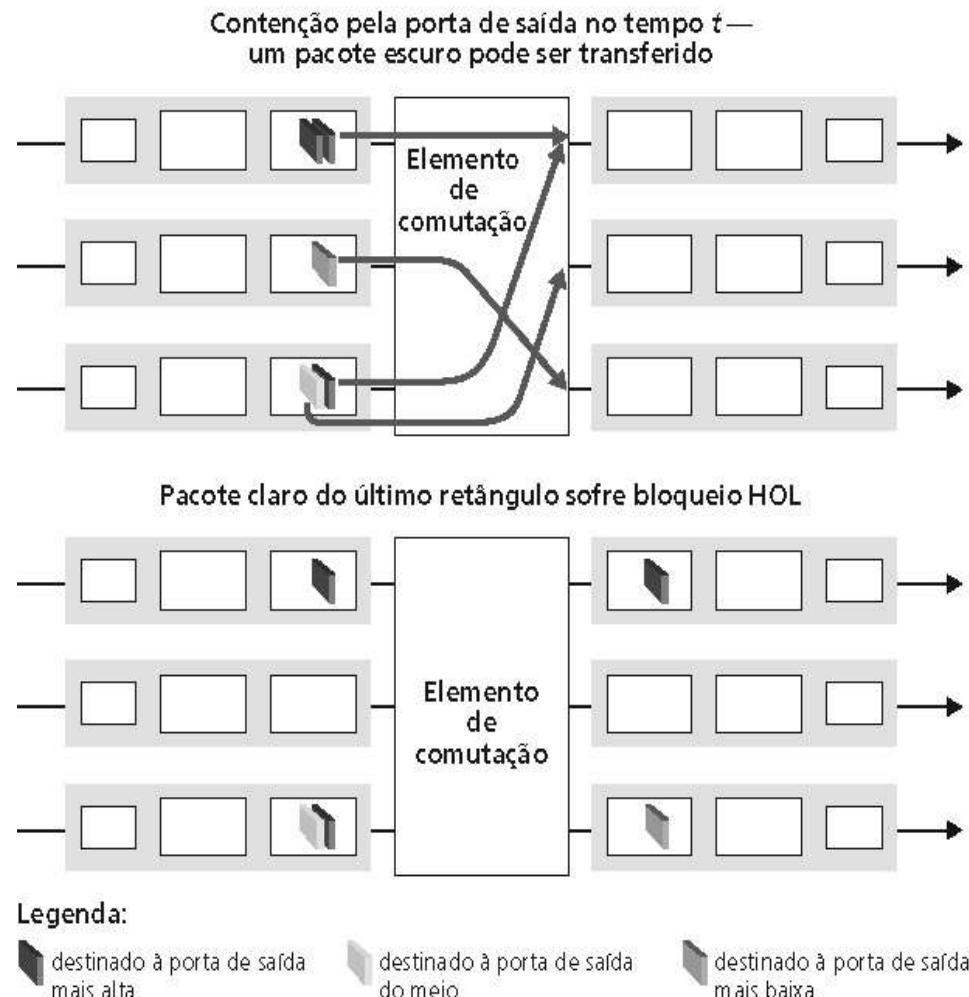
# Portas de Saída



- Bufferização é necessária quando datagramas chegam do elemento de comutação mais rápido que a taxa de transmissão
  - Enfileiramento (atraso) e perda devido ao overflow de buffers de saída
- Disciplinas de escalonamento escolhem entre datagramas na fila para transmissão

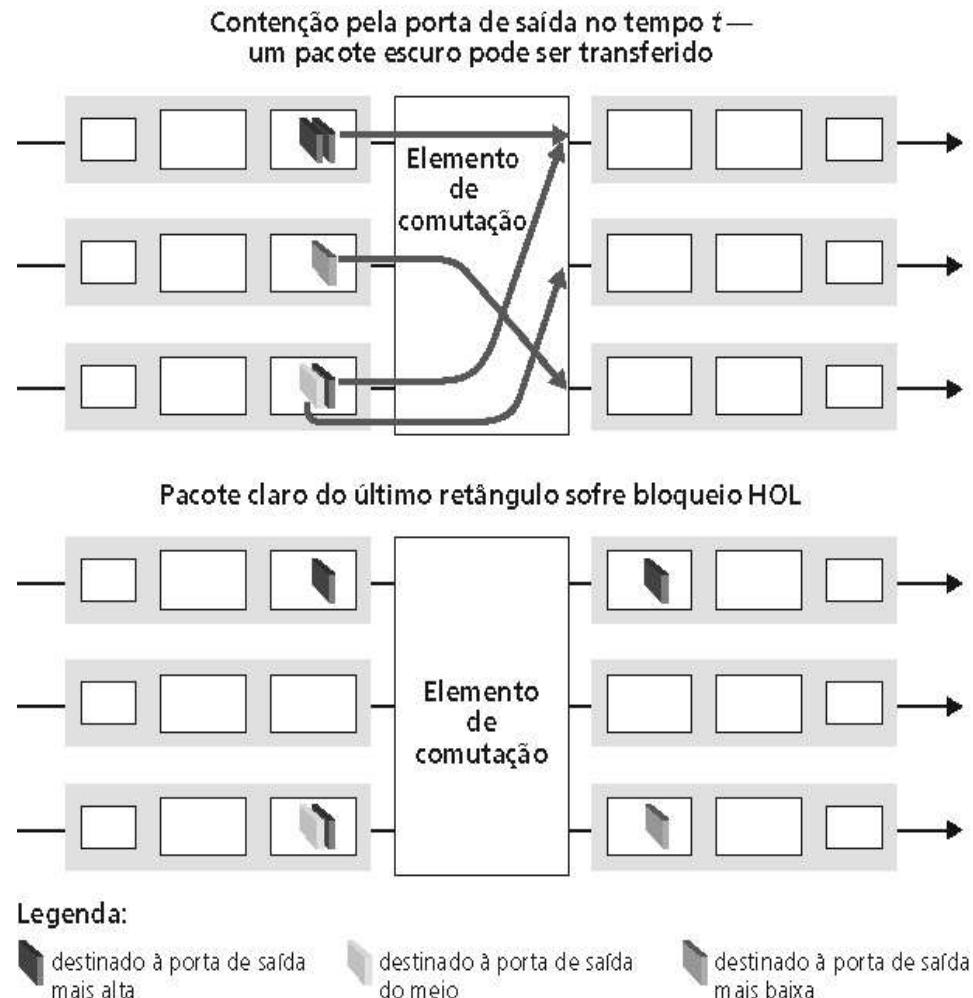
# Enfileiramento na porta de saída

- **Buferização:** quando a taxa de chegada pelo roteador excede a velocidade da linha de saída
- **Enfileiramento (atraso) e perda** devido ao overflow da fila da porta de saída!



# Enfileiramento na porta de entrada

- Elemento de comutação mais lento que as portas de entrada combinadas -> pode ocorrer filas na entrada
- Bloqueio Head-of-the-Line (HOL): datagrama na frente da fila impede os outros na fila de se moverem para adiante
- Atraso e perda na fila devido ao overflow no buffer de entrada!

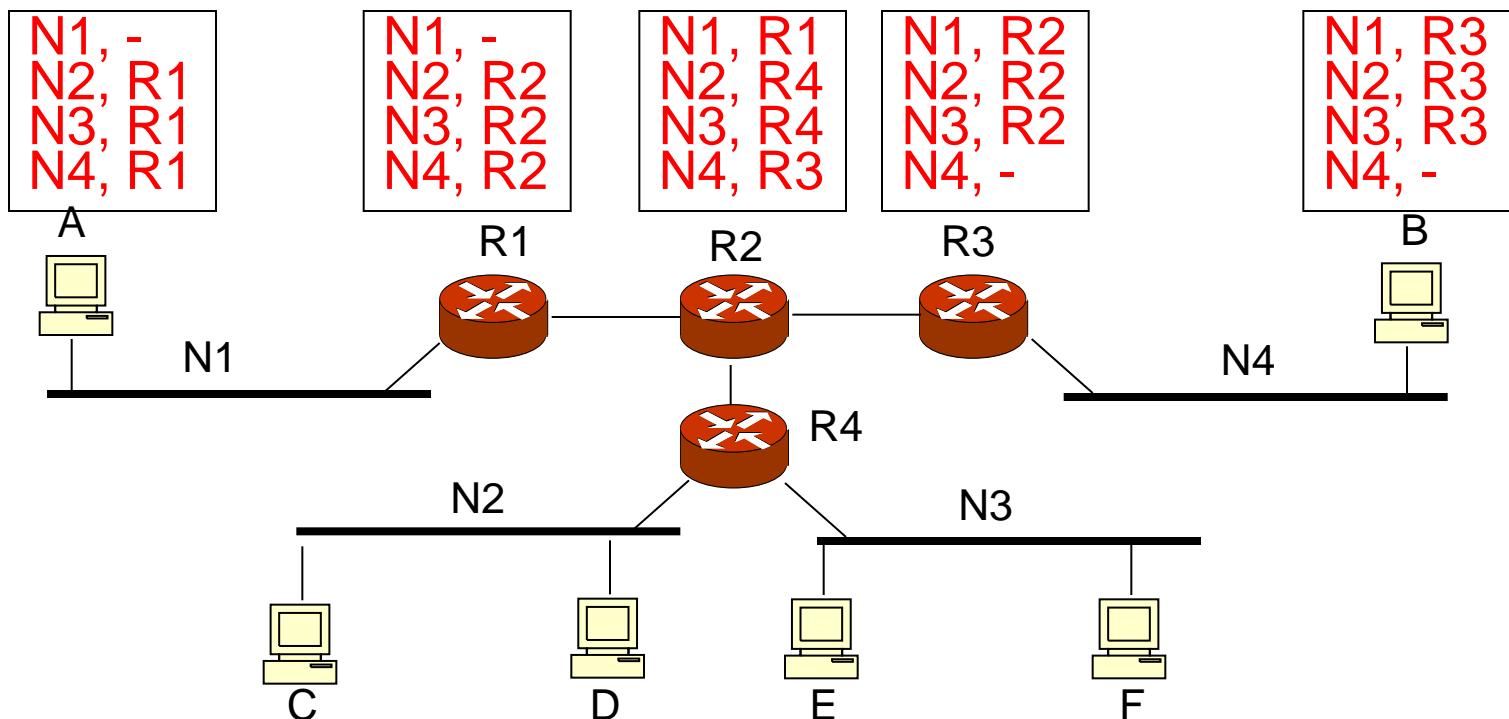


# Roteamento

- Roteamento inter-redes é a principal função do protocolo IP
  - Protocolo assume que um host sabe enviar datagramas para qualquer host da mesma rede local
  - Roteador entra em ação no momento que um datagrama tem destino fora da rede local
    - Host origem envia o datagrama para o gateway (roteador)

# Roteamento e Tabelas de Roteamento

- Como manter informações sobre rotas de A para qualquer outro host?
  - $A \rightarrow R1 \rightarrow R2 \rightarrow R3 \rightarrow B$
- Cada nó armazena endereço de rede e próximo hop



# Tipos de Roteamento

## Estático

- Configurado pelo administrador no roteador
- Não se adapta a mudanças na topologia da rede
- Informação não é trocada com outros roteadores, por default (+ segurança)

## Dinâmico

- Administrador configura parâmetros iniciais
- Protocolos de roteamento são usados para trocar informações entre roteadores
- Adaptação automática em caso de mudança de topologia

# Protocolos de Roteamento

## □ Objetivos

- Gerenciar a tabela de roteamento dinamicamente, com rota para “todas” as redes
- Se há mais de uma rota para uma rede, a com o menor custo é colocada na tabela de roteamento
- Adicionar novas rotas, trocar por melhores naquele momento
- Prevenir loops de roteamento

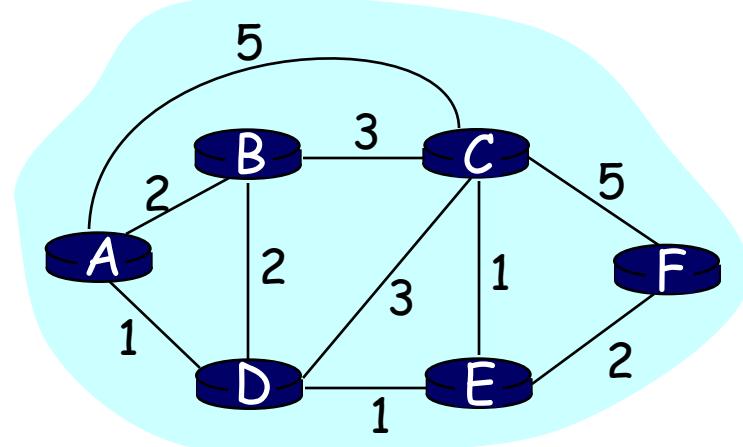
# Roteamento

## Protocolo de Roteamento

**OBJ:** determinar “bons” caminhos (sequência de roteadores) através da rede da fonte ao destino.

Algoritmos de roteamento são descritos por grafos:

- Nós do grafo são roteadores
- Arestas do grafo são enlaces
  - Custo do enlace: atraso, preço ou nível de congestionamento



- “bons” caminhos:
  - tipicamente corresponde aos caminhos de menor custo
  - caminhos redundantes

# Roteamento IP

- Estratégia de roteamento:
  - se host destino está na mesma rede => envia datagrama diretamente (endereço MAC)
  - senão => envia a um gateway local (endereço MAC)
- Pesquisa por uma rota:
  - extrai o endereço da rede respeitando as classes padrões
    - se é a rede local => aplica máscara de sub-rede
  - pesquisa tabela de roteamento pelo endereço da rede
    - se encontrou rota => envia ao gateway
    - senão => envia ao gateway default

# Levando um Datagrama da Fonte ao Destino

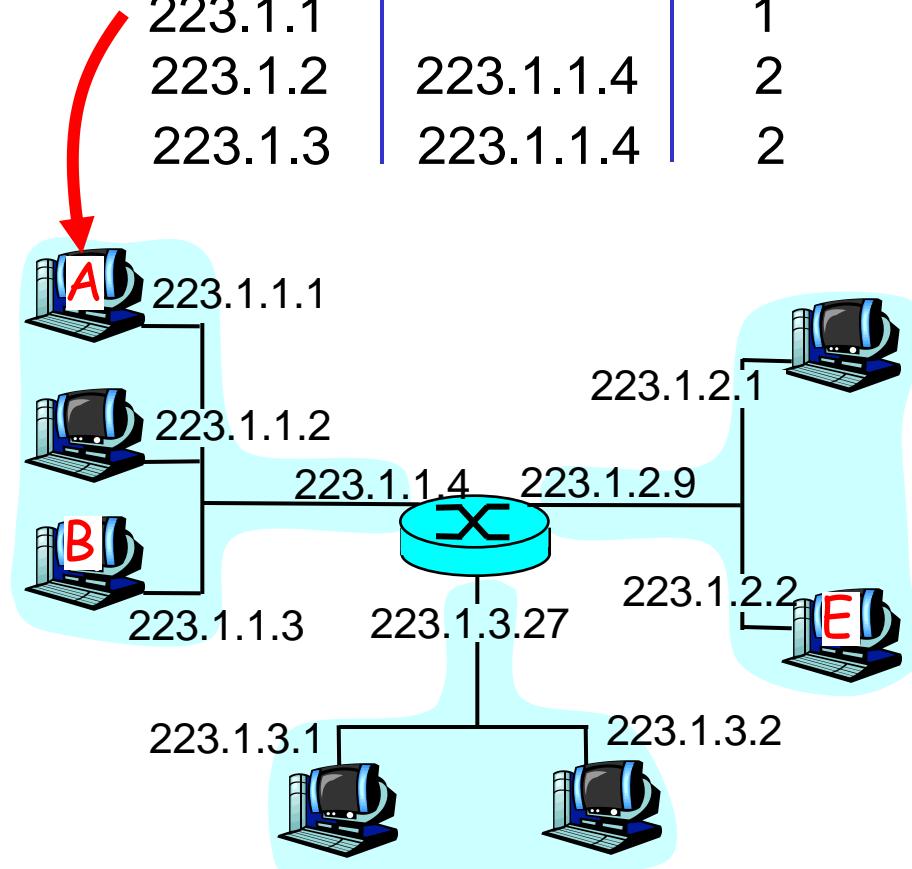
datagrama IP:



- os endereços do datagrama não mudam ao viajar da fonte ao destino

tabela de roteamento  
em A

Rede destino	próx. roteador	N. saltos
223.1.1		1
223.1.2	223.1.1.4	2
223.1.3	223.1.1.4	2



# Roteamento IP

- Comando "route print" no host 192.168.0.194

```
C:\Users\willrich>route print -4
=====
Lista de interfaces
14...00 1c bf 1c 53 7b ....Intel(R) PRO/Wireless 3945ABG Network Connection
11...00 1b 24 95 70 bf ....Broadcom NetLink(TM) Gigabit Ethernet
1.....00 00 00 00 00 00 e0 Software Loopback Interface 1
15...00 00 00 00 00 00 e0 Microsoft 6to4 Adapter #2
12...00 00 00 00 00 00 e0 Microsoft 6to4 Adapter
13...00 00 00 00 00 00 e0 Teredo Tunneling Pseudo-Interface
17...00 00 00 00 00 00 e0 Microsoft 6to4 Adapter #3
16...00 00 00 00 00 00 e0 Microsoft ISATAP Adapter #2
26...00 00 00 00 00 00 e0 Microsoft ISATAP Adapter
=====

Tabela de rotas IPv4
=====
Rotas ativas:
Endereço de rede      Máscara    Ender. gateway      Interface   Custo
          0.0.0.0        0.0.0.0      192.168.0.1      192.168.0.194  25
          127.0.0.0       255.0.0.0     No vínculo        127.0.0.1   306
          127.0.0.1       255.255.255.255  No vínculo        127.0.0.1   306
          127.255.255.255 255.255.255.255  No vínculo        127.0.0.1   306
          192.168.0.0       255.255.255.0     No vínculo      192.168.0.194  281
          192.168.0.194     255.255.255.255  No vínculo      192.168.0.194  281
          192.168.0.255     255.255.255.255  No vínculo      192.168.0.194  281
          224.0.0.0         240.0.0.0     No vínculo        127.0.0.1   306
          224.0.0.0         240.0.0.0     No vínculo      192.168.0.194  281
          255.255.255.255  255.255.255.255  No vínculo        127.0.0.1   306
          255.255.255.255  255.255.255.255  No vínculo      192.168.0.194  281
=====
Rotas persistentes:
Nenhuma

C:\Users\willrich>_
```

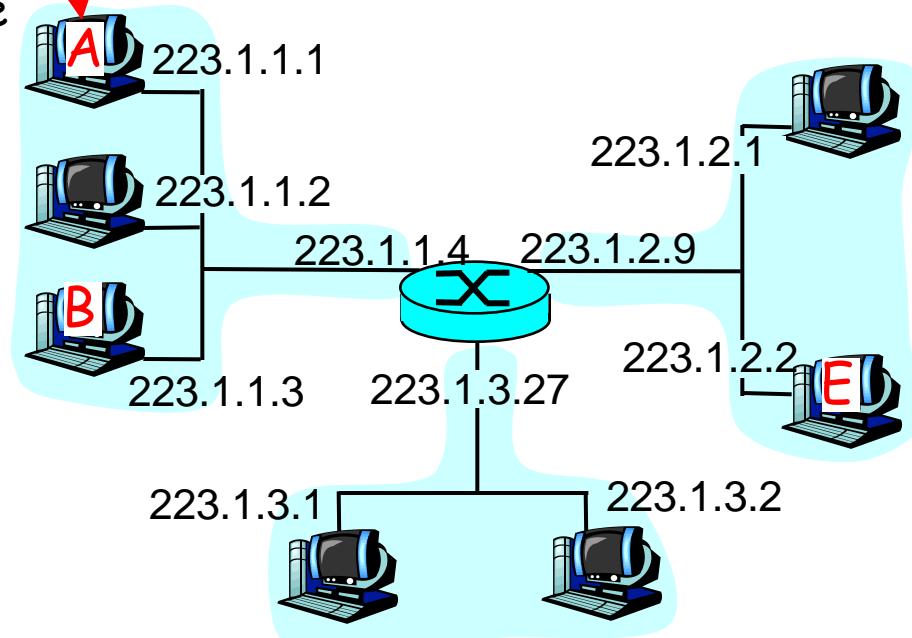
# Levando um Datagrama da Fonte ao Destino

outros campos	223.1.1.1	223.1.1.3	dados
---------------	-----------	-----------	-------

A envia datagrama IP para B:

- examine endereço de rede de B
- descobre que B está na mesma rede de A
- camada de enlace envia datagrama diretamente para B num quadro da camada de enlace
- Se necessário descobre endereço físico de B (usando ARP)

Rede destino	Próx. roteador	Núm. saltos
223.1.1		1
223.1.2	223.1.1.4	2
223.1.3	223.1.1.4	2



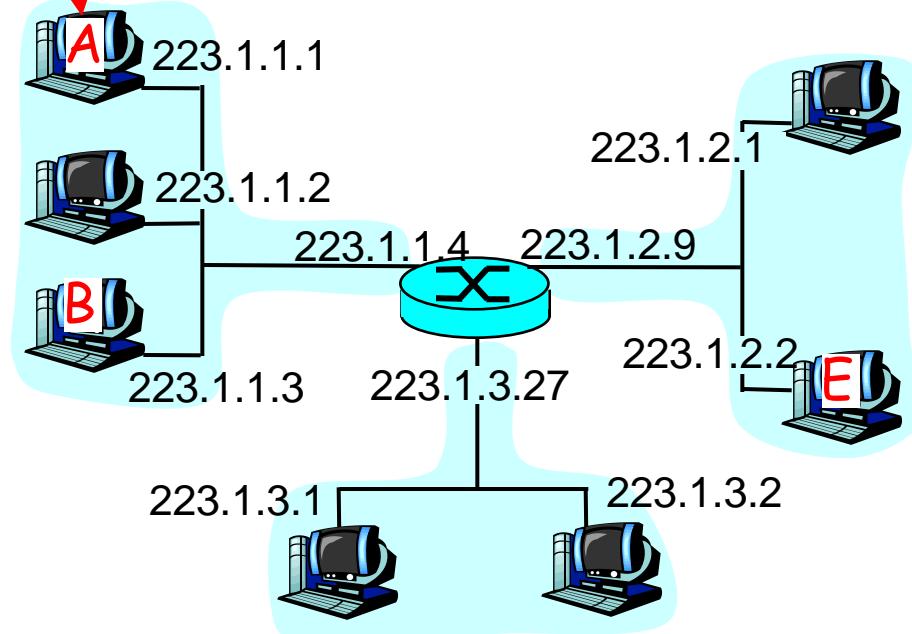
# Levando um Datagrama da Fonte ao Destino

outros campos	223.1.1.1	223.1.2.2	dados
---------------	-----------	-----------	-------

A envia datagrama IP para E:

- examina endereço de rede de E
- E está numa rede diferente
  - A, E não estão diretamente conectados
- tabela de roteamento: próximo roteador para E é 223.1.1.4
- encontra endereço físico de 223.1.1.4 e envia o datagrama num quadro de enlace
- datagrama chega em 223.1.1.4
- continua.....

Rede destino	Próx. roteador	N. saltos
223.1.1		1
223.1.2	223.1.1.4	2
223.1.3	223.1.1.4	2



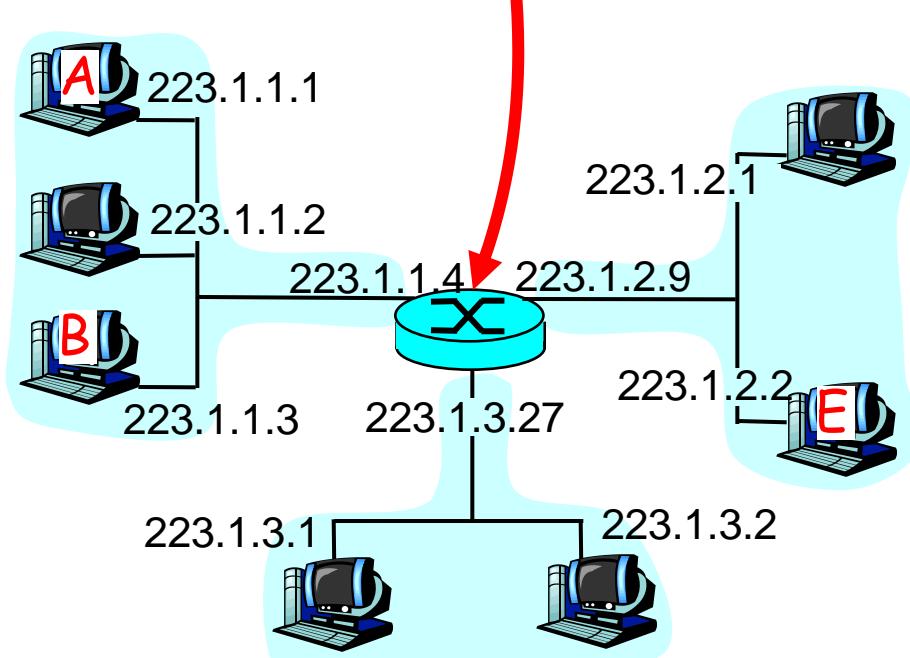
# Levando um Datagrama da Fonte ao Destino

outro campos	223.1.1.1	223.1.2.2	dados
--------------	-----------	-----------	-------

Chegando em 223.1.1.4,  
destinado para 223.1.2.2

- examina endereço de rede de E
- E está na mesma rede da interface 223.1.2.9 do roteador
  - roteador e E estão diretamente ligados
- descobre endereço físico de 223.1.2.2 e envia o datagrama num quadro da camada de enlace
- datagrama chega em 223.1.2.2!!! (ufa!)

Rede destino	Próx. roteador	Nºm. saltos	Endereço Interface
223.1.1	-	1	223.1.1.4
223.1.2	-	1	223.1.2.9
223.1.3	-	1	223.1.3.27



# Roteamento

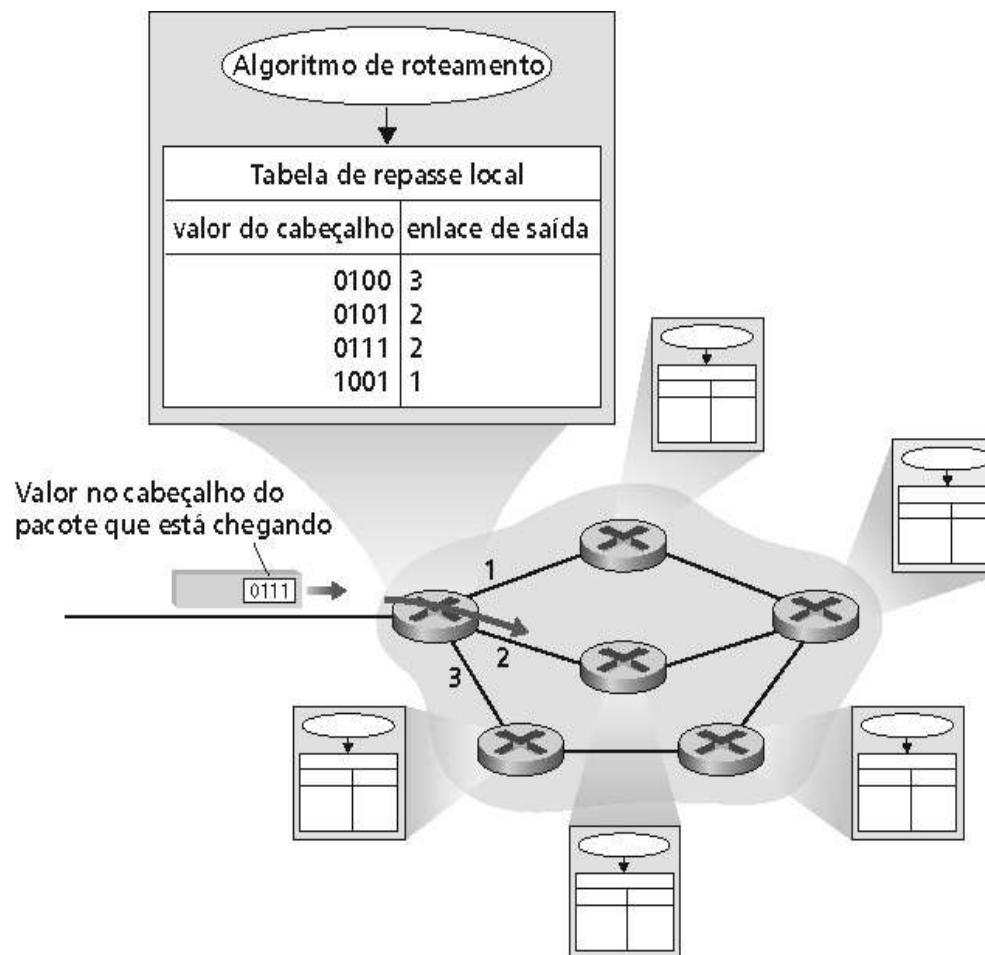
## □ Algoritmos de roteamento

- Link state
- Distance vector
- Roteamento hierárquico

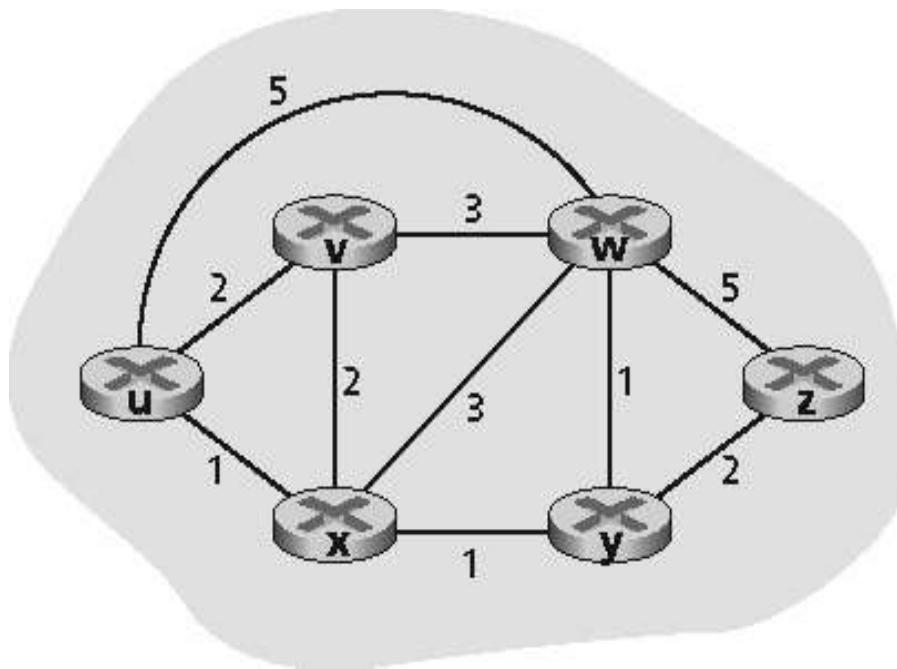
## □ Roteamento na Internet

- RIP
- OSPF
- BGP

# Interação entre roteamento e comutação



# Abstração da Rede: Grafo

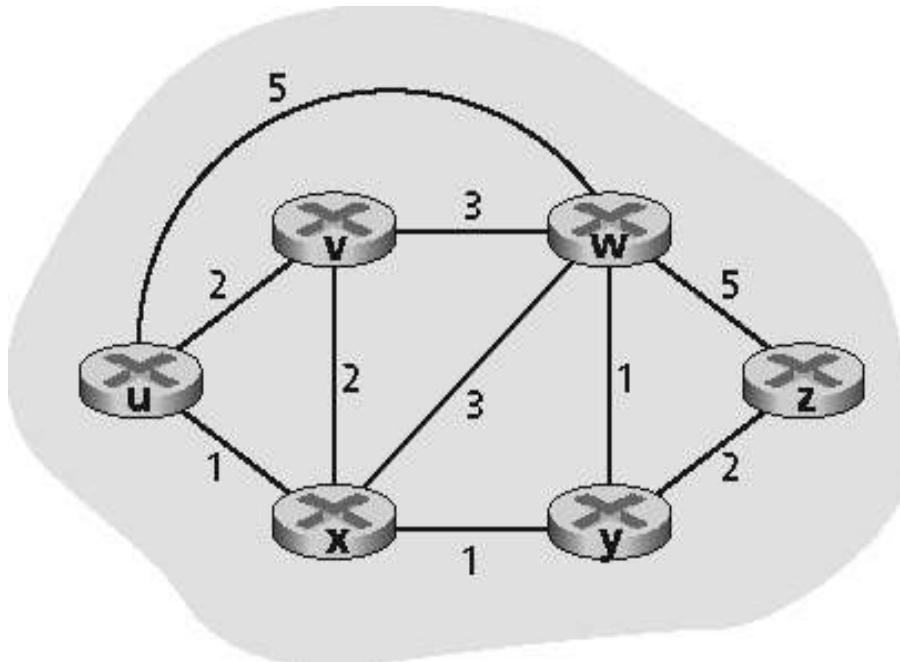


□ Grafo:  $G = (N, E)$

- $N$  = conjunto de roteadores = {  $u, v, w, x, y, z$  }
- $E$  = conjunto de enlaces = {  $(u,v), (u,x), (v,x), (v,w), (x,w), (x,y), (w,y), (w,z), (y,z)$  }

# Abstração da Rede: Grafo

## Custo



- $c(x, x') = \text{custo do enlace } (x, x')$ 
  - ex.,  $c(w, z) = 5$
- Custo poderia ser sempre 1, ou inversamente relacionado à largura de banda ou ao congestionamento

Custo do caminho  $(x_1, x_2, x_3, \dots, x_p) = c(x_1, x_2) + c(x_2, x_3) + \dots + c(x_{p-1}, x_p)$

**Questão: Qual é o caminho de menor custo entre  $u$  e  $z$  ?**

**Algoritmo de roteamento:** algoritmo que encontra o caminho de menor custo

# Classificação dos algoritmos de roteamento

- Informação global ou descentralizada
  - Global:
    - Todos os roteadores têm informações completas da topologia e dos custos dos enlaces
    - Algoritmos "link state"
  - Descentralizada:
    - Roteadores só conhecem informações sobre seus vizinhos e os enlaces para eles
    - Processo de computação interativo
      - troca de informações com os vizinhos
    - Algoritmos "distance vector"

# Roteamento

## □ Algoritmos de roteamento

- Link state
- Distance vector
- Roteamento hierárquico

## □ Roteamento na Internet

- RIP
- OSPF
- BGP

# Algoritmo de roteamento link-state

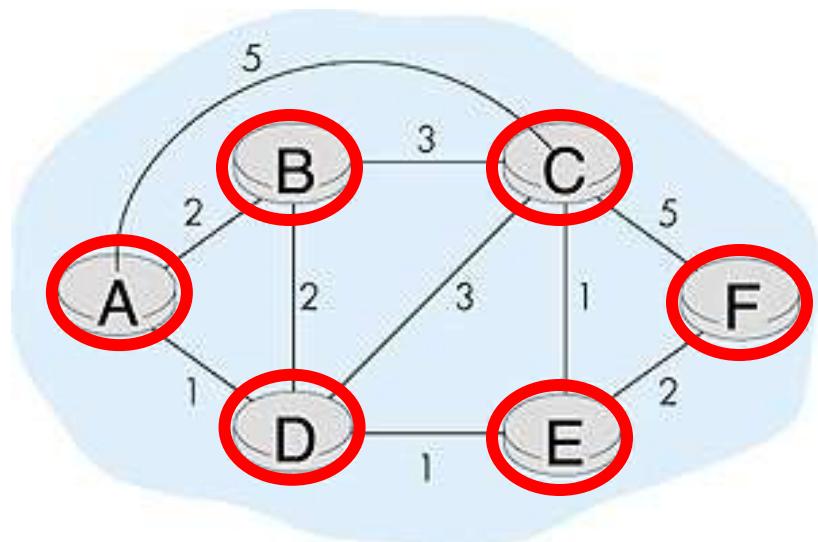
- Algoritmo de Dijkstra
  - Topologia de rede e custo dos enlaces são conhecidos por todos os nós
    - Implementado via "link state broadcast"
    - Todos os nós têm a mesma informação
  - Computa caminhos de menor custo de um nó (fonte) para todos os outros nós
    - Fornece uma tabela de roteamento para aquele nó
  - Convergência: após k iterações, conhece o caminho de menor custo para k destinos
- Notação:
  - $C(i,j)$ : custo do enlace do nó i ao nó j. Custo é infinito se não houver ligação entre i e j
  - $D(v)$ : valor atual do custo do caminho da fonte ao destino v
  - $P(v)$ : nó predecessor ao longo do caminho da fonte ao nó v, isto é, antes do v
  - $N'$ : conjunto de nós cujo caminho de menor custo é definitivamente conhecido

# Algoritmo de Dijkstra

- 1 Inicialização (rodando em um nó u):
- 2  $N' = \{u\}$
- 3 para todos os nós v
- 4   se v é adjacente a u
- 5     então  $D(v) = c(u,v)$
- 6     senão  $D(v) = \infty$
- 7
- 8 **Loop**
- 9   ache w não em  $N'$  tal que  $D(w)$  é um mínimo
- 10   acrescente w a  $N'$
- 11   atualize  $D(v)$  para todo v adjacente a w e não em  $N'$ :
- 12      $D(v) = \min( D(v), D(w) + c(w,v) )$
- 13   /\* novo custo para v é ou o custo anterior para v ou o menor
- 14    custo do caminho conhecido para w mais o custo de w a v \*/
- 15 **até que todos os nós estejam em  $N'$**

# Exemplo: Algoritmo de Dijkstra

Passo	N'	D(B),p(B)	D(C),p(C)	D(D),p(D)	D(E),p(E)	D(F),p(F)
0	A	2,A	5,A	1,A	$\infty$	$\infty$
1	AD	2,A	4,D		2,D	$\infty$
2	ADE	2,A	3,E			4,E
3	ADEB		3,E			4,E
4	ADEBC					4,E
5	ADEBCF					



# Algoritmo de Dijkstra: exemplo

Passo	N	D(B),p(B)	D(C),p(C)	D(D),p(D)	D(E),p(E)	D(F),p(F)
0	A	2,A	5,A	1,A	$\infty$	$\infty$
1	AD	2,A	4,D		2,D	$\infty$
2	ADE	2,A	3,E			4,E
3	ADEB		3,E			4,E
4	ADEBC					4,E
5	ADEBCF					

Árvore de caminhos mínimos resultante originada em A:

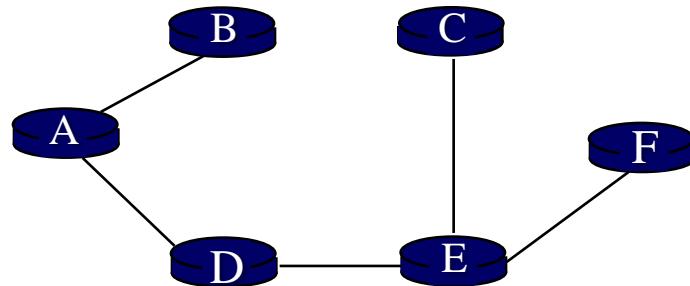
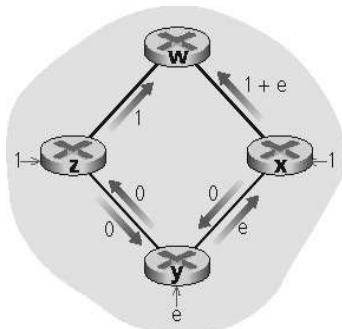


Tabela de encaminhamento resultante em A:

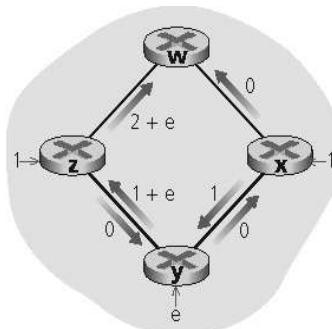
destino	enlace
B	(A,B)
C	(A,D)
D	(A,D)
E	(A,D)
F	(A,D)

# Discussão do algoritmo de Dijkstra

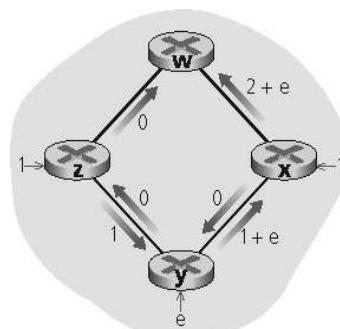
- Complexidade do algoritmo:  $n$  nós
  - Cada iteração: precisa verificar todos os nós  $w$ , que não estão em  $N$ 
    - 1a iteração  $n$  nós, 2a iteração  $n-1$  nós, 3a iteração  $n-2$  nós,...
    - Número de nós buscados em todas as interações:  $n(n+1)/2$
    - Complexidade:  $O(n^2)$
  - Implementações mais eficientes:  $O(n \log n)$
- Oscilações possíveis:
  - Ex.: custo do enlace = quantidade de tráfego transportado



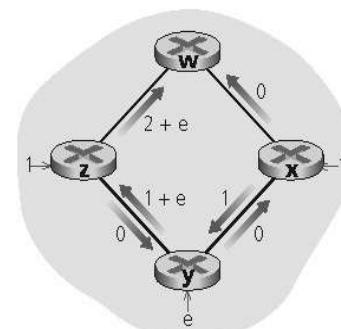
a. Roteamento inicial



b.  $x, y$  detectam melhor caminho até  $w$  em sentido horário



c.  $x, y, z$  detectam melhor caminho até  $w$  em sentido anti-horário

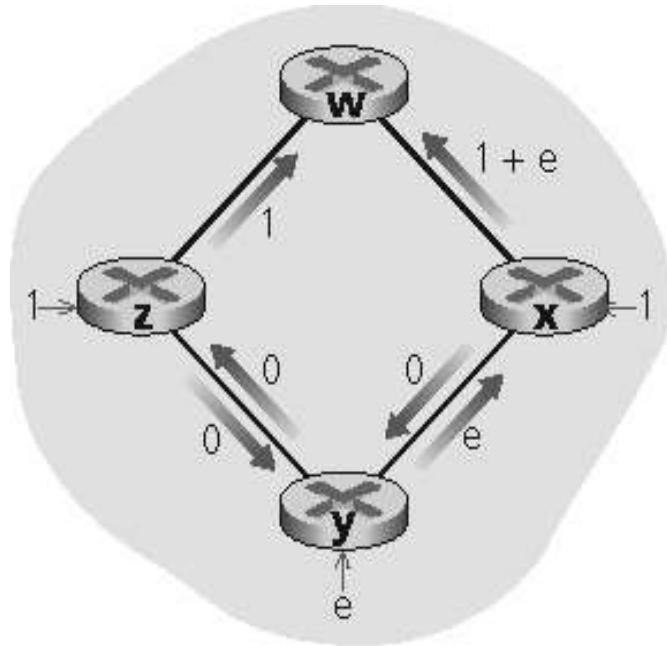


d.  $x, y, z$  detectam melhor caminho até  $w$  em sentido horário

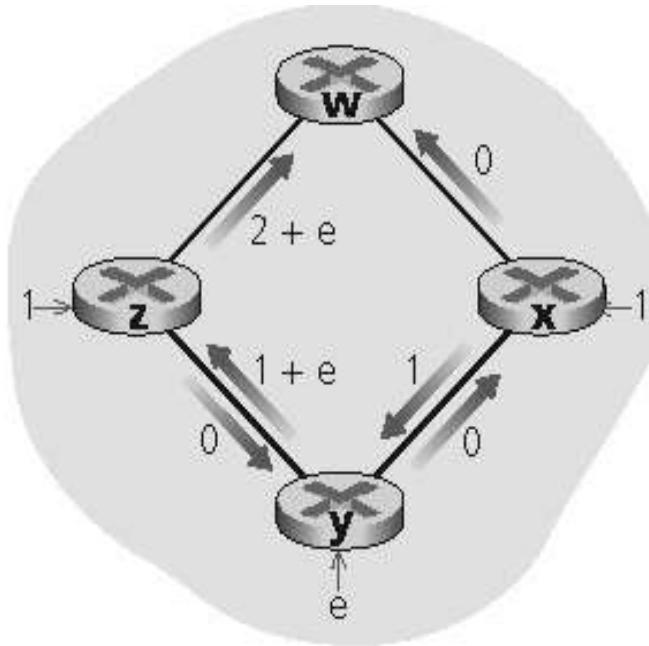
# Discussão do algoritmo de Dijkstra

- Oscilações possíveis:

- Ex.: custo do enlace = quantidade de tráfego transportado



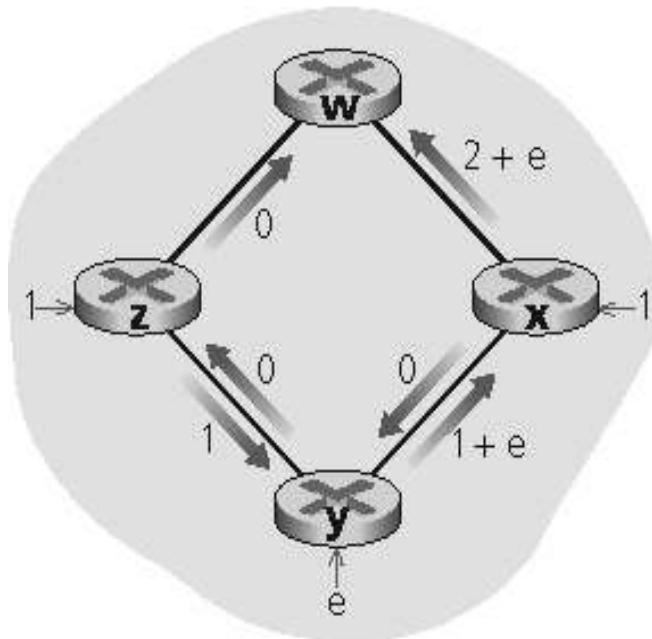
a. Roteamento inicial



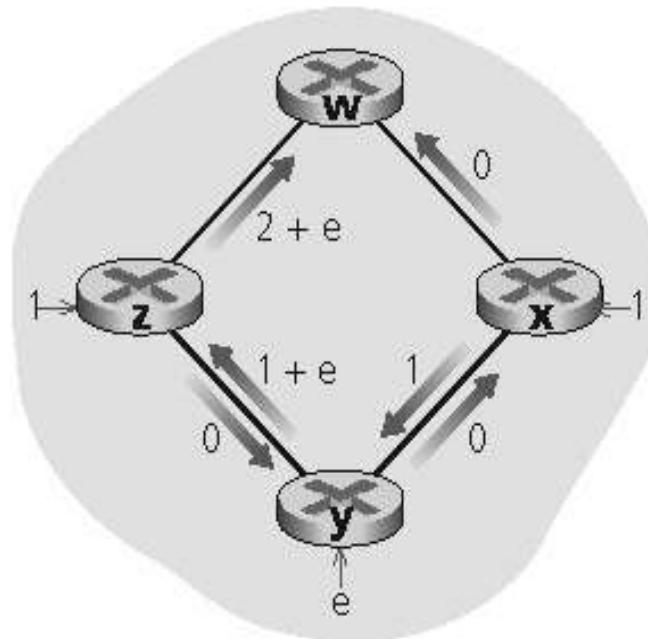
b.  $x, y$  detectam melhor caminho até  $w$  em sentido horário

# Discussão do algoritmo de Dijkstra

- Oscilações possíveis:
  - Ex.: custo do enlace = quantidade de tráfego transportado



c.  $x, y, z$  detectam melhor caminho até  $w$  em sentido anti-horário



d.  $x, y, z$  detectam melhor caminho até  $w$  em sentido horário

# Roteamento

## □ Algoritmos de roteamento

- Link state
- Distance vector
- Roteamento hierárquico

## □ Roteamento na Internet

- RIP
- OSPF
- BGP

# Algoritmo vetor de distância (1)

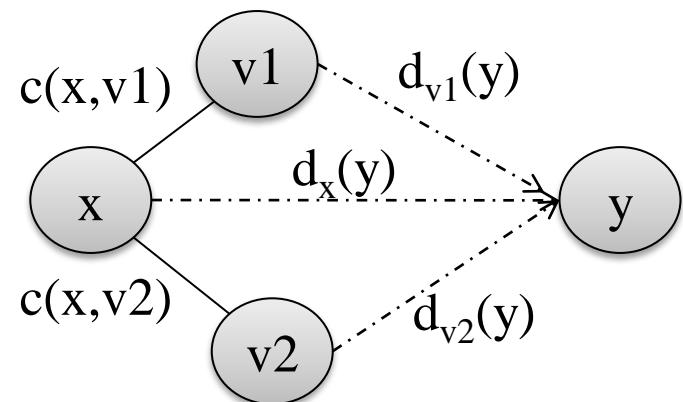
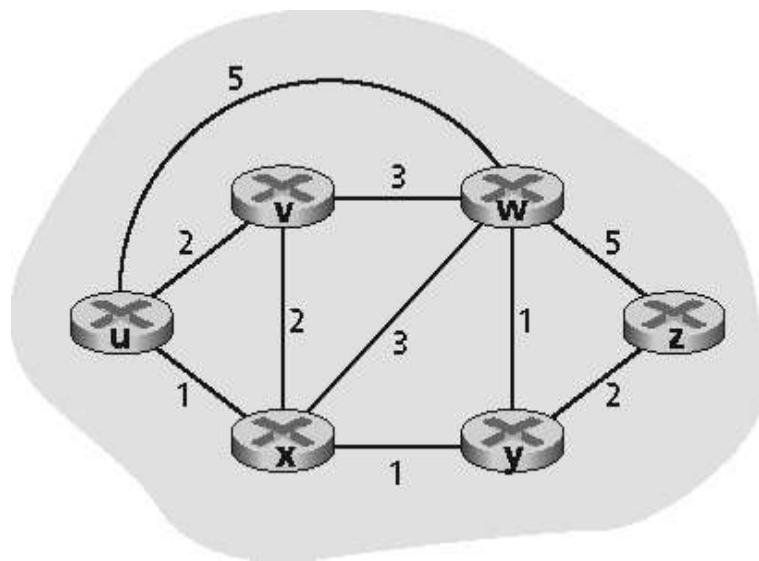
## □ Equação de Bellman-Ford (programação dinâmica)

### ○ Define

- $d_x(y) = \text{custo do caminho de menor custo de } x \text{ para } y$

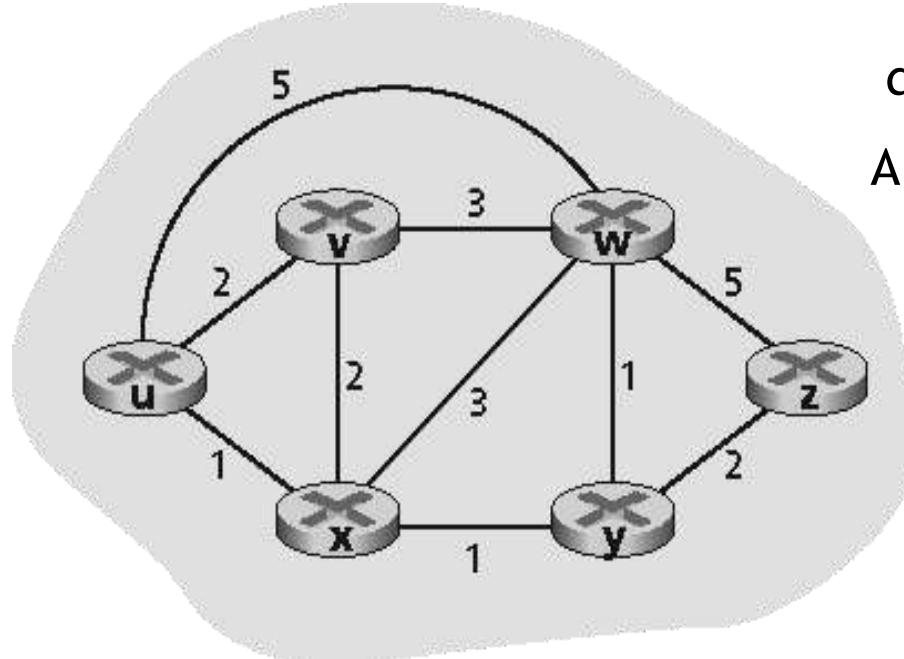
### ○ Então

- $d_x(y) = \min_v \{c(x,v) + d_v(y)\}$
- Em que  $\min_v$  é calculado para todos os vizinhos de  $x$
- No exemplo:  $d_x(y) = \min_v \{c(x,v1) + d_{v1}(y), c(x,v2) + d_{v2}(y)\}$



# Exemplo: Bellman-Ford (2)

**Exemplo: custo entre o nó u e o nó z**



$$d_v(z) = 5, d_x(z) = 3, d_w(z) = 3$$

A equação B-F diz que:

$$\begin{aligned} d_u(z) &= \min \{ c(u,v) + d_v(z), \\ &\quad c(u,x) + d_x(z), \\ &\quad c(u,w) + d_w(z) \} \\ &= \min \{ 2 + 5, \\ &\quad 1 + 3, \\ &\quad 5 + 3 \} = 4 \end{aligned}$$

O nó que atinge o mínimo é o próximo salto no caminho mais curto  
→ tabela de roteamento

## Algoritmo vetor de distância (3)

- Define a forma de comunicação de vizinho para vizinho
  - Atualizando a tabela de roteamento para que o pacote seja enviado para o vizinho no caminho de menor custo
- Cada nó  $x$  mantém os seguintes dados de roteamento
  - Para cada vizinho  $v$ , ele mantém o custo  $c(x,v)$
  - O vetor de distâncias do nó  $x$ 
    - $D_x = [D_x(y) : y \in N]$
  - Os vetores de distância de seus vizinhos
    - Para cada vizinho  $v$ ,  $x$  mantém  $D_v = [D_v(y) : y \in N]$

## Algoritmo vetor de distância (4)

### □ Idéia básica:

- Cada nó envia periodicamente sua própria estimativa de vetor de distâncias aos vizinhos
- Quando o nó  $x$  recebe nova estimativa de DV do vizinho, ele atualiza seu próprio DV usando a equação B-F:
  - $D_x(y) = \min_v \{c(x,v) + D_v(y)\}$  para cada nó  $y \in N$
- Ao menos em condições naturais, a estimativa  $D_x(y)$  converge para o menor custo atual  $d_x(y)$

# Algoritmo vetor de distância (5)

- **Iterativo**, assíncrono:  
cada iteração local é  
causada por:

- Mudança no custo do enlace local
  - Mensagem de atualização DV do vizinho

- **Distribuído**:
  - Cada nó notifica os vizinhos apenas quando seu DV mudar
  - Os vizinhos então notificam seus vizinhos, se necessário

Cada nó:

espera por (mudança no custo do enlace local na mensagem do vizinho)

recalcula estimativas

se o DV para qualquer destino mudou, notifica os vizinhos

# Algoritmo vetor de distância

Tabela do nó x

		Custo até					
		x	y	z	x	y	z
D <sub>x</sub>	x	0	2	7	2	0	1
	y	∞	∞	∞	7	1	0
	z	∞	∞	∞	3	1	0

$$D_x(y) = \min\{c(x,y) + D_y(y), c(x,z) + D_z(y)\} \\ = \min\{2+0, 7+1\} = 2$$

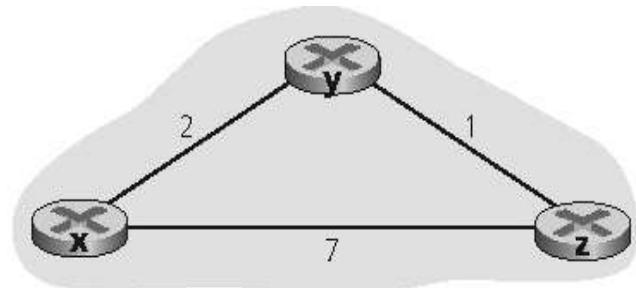
$$D_x(z) = \min\{c(x,y) + D_y(z), c(x,z) + D_z(z)\} \\ = \min\{2+1, 7+0\} = 3$$

Tabela do nó y

		Custo até					
		x	y	z	x	y	z
D <sub>y</sub>	x	∞	∞	∞	0	2	7
	y	2	0	1	2	0	1
	z	∞	∞	∞	7	1	0

Tabela do nó z

		Custo até					
		x	y	z	x	y	z
D <sub>z</sub>	x	∞	∞	∞	0	2	7
	y	∞	∞	∞	2	0	1
	z	7	1	0	3	1	0

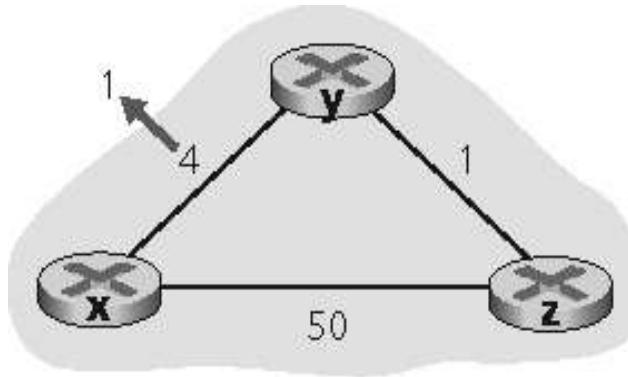


Tempo

# Vetor de distância: mudanças no custo do enlace

Mudanças no custo do enlace:

- Nó detecta mudança no custo do enlace local
- Atualiza informações de roteamento, recalcula o vetor de distância
- Se o DV muda, notifica vizinhos



a.

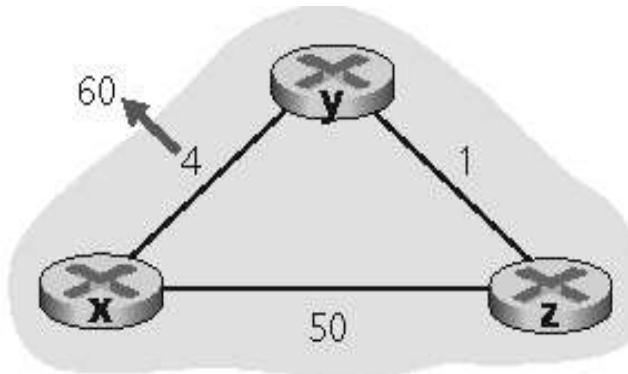
No tempo  $t_0$ , y detecta a mudança no custo do enlace, atualiza seu DV e informa seus vizinhos.

No tempo  $t_1$ , z recebe a atualização de y e atualiza sua tabela. Ele calcula o menor custo novo para x e envia seu DV para os vizinhos. No tempo  $t_2$ , y recebe a atualização de z e atualiza sua tabela de distância. O menor custo de y's não muda e então y não envia nenhuma mensagem para z.

“boas notícias viajam depressa”

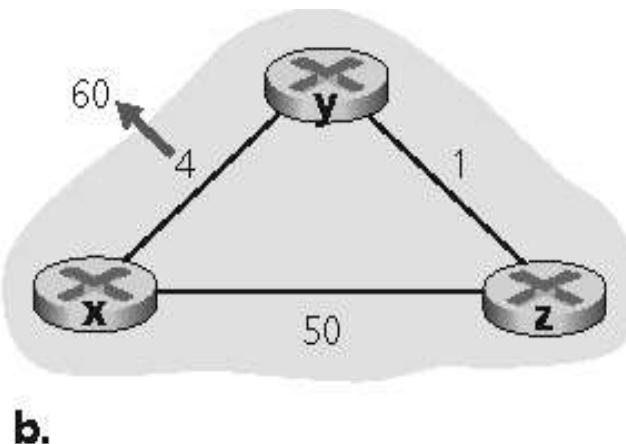
# Vetor de distância: mudanças no custo do enlace

- Mudanças no custo do enlace:
  - Boas notícias viajam rápido
  - Más notícias viajam devagar — problema da “contagem ao infinito”!
- No inicio (antes da mudança do custo)
  - $D_y(x)=4, D_y(z)=1, D_z(y)=1, D_z(x)=5$
- Y detecta a mudança de custo
  - $D_y(x) = \min\{c(y,x) + D_x(x), c(y,z) + D_z(x)\} = \min\{60, 6\}=6$
  - Um equivoco! Gera um loop de roteamento
- Atualizando DV, y informa a z
  - Y informa que  $D_y(x)=6$  e como z sabe que  $c(z,y)=1$ , então  $D_z(x)=7$
  - Z informa novo DV a y
- 44 iterações antes de o algoritmo estabilizar: veja o texto



# Vetor de distância: mudanças no custo do enlace

- Solução: Reversão envenenada:
  - Se Z roteia por Y para alcançar X :
    - Z diz a Y que sua distância (de Z) para X é infinita (então Y não roteará até X via Z)
- Isso resolverá completamente o problema da contagem ao infinito?
  - Não é solução geral para loops envolvendo três ou mais nós



# Comparação dos algoritmos LS e VD

## Complexidade

- **LS:** com  $n$  nós,  $E$  links,  $O(NE)$  mensagens enviadas
- **DV:** trocas somente entre vizinhos
  - Tempo de convergência varia

## Tempo de convergência

- **LS:** algoritmo  $O(N^2)$  exige mensagens  $O(NE)$ 
  - Pode ter oscilações
- **DV:** tempo de convergência varia
  - Pode haver loops de roteamento durante a convergência
  - Problema da contagem ao infinito

**Robustez:** o que acontece se um roteador funciona mal?

**Ls:**

- Nós podem informar custos de **link** incorretos
- Cada nó calcula sua própria tabela de roteamento: aumenta a robustez

**Dv:**

- Nós DV pode informar custo de **caminho** incorreto
- Tabela de cada nó é usada por outros
  - Propagação de erros pela rede

# Roteamento

## □ Algoritmos de roteamento

- Link state
- Distance vector
- Roteamento hierárquico

## □ Roteamento na Internet

- RIP
- OSPF
- BGP

# Roteamento hierárquico

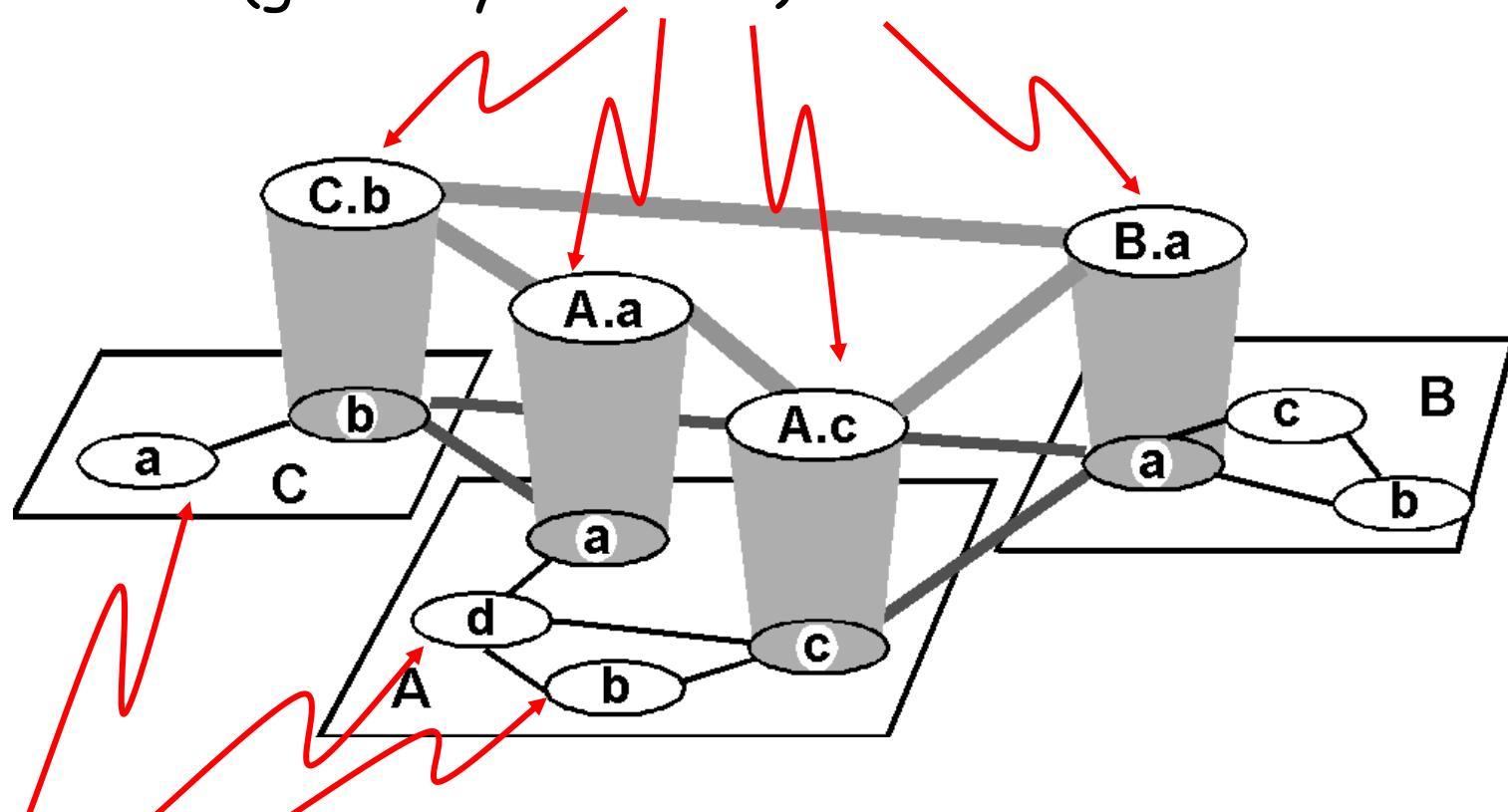
- Nosso estudo até aqui foi uma idealização
  - Roteadores são todos idênticos (executam o mesmo algoritmo de roteamento)
  - ... na prática, isso não é verdade
- Na prática existem mais problemas
  - Escala: com 200 milhões de destinos:
    - Não é possível armazenar todos os destinos numa única tabela de rotas!
    - As mudanças na tabela de rotas irão congestionar os enlaces!
  - Autonomia administrativa
    - Internet = rede de redes
    - Cada administração de rede pode querer controlar o roteamento na sua própria rede

# Roteamento hierárquico

- Agrega roteadores em regiões: “**Sistemas Autônomos**” (AS)
  - Roteadores no mesmo AS rodam o mesmo protocolo de roteamento
    - Protocolo de roteamento “**intra-AS**”
  - Roteadores em diferentes AS podem rodar diferentes protocolos de roteamento
- Roteador de Borda (*Gateway*)
  - Link direto para um roteador em outro AS
  - Protocolo de roteamento “**inter-AS**”

# Internet como Hierarquia

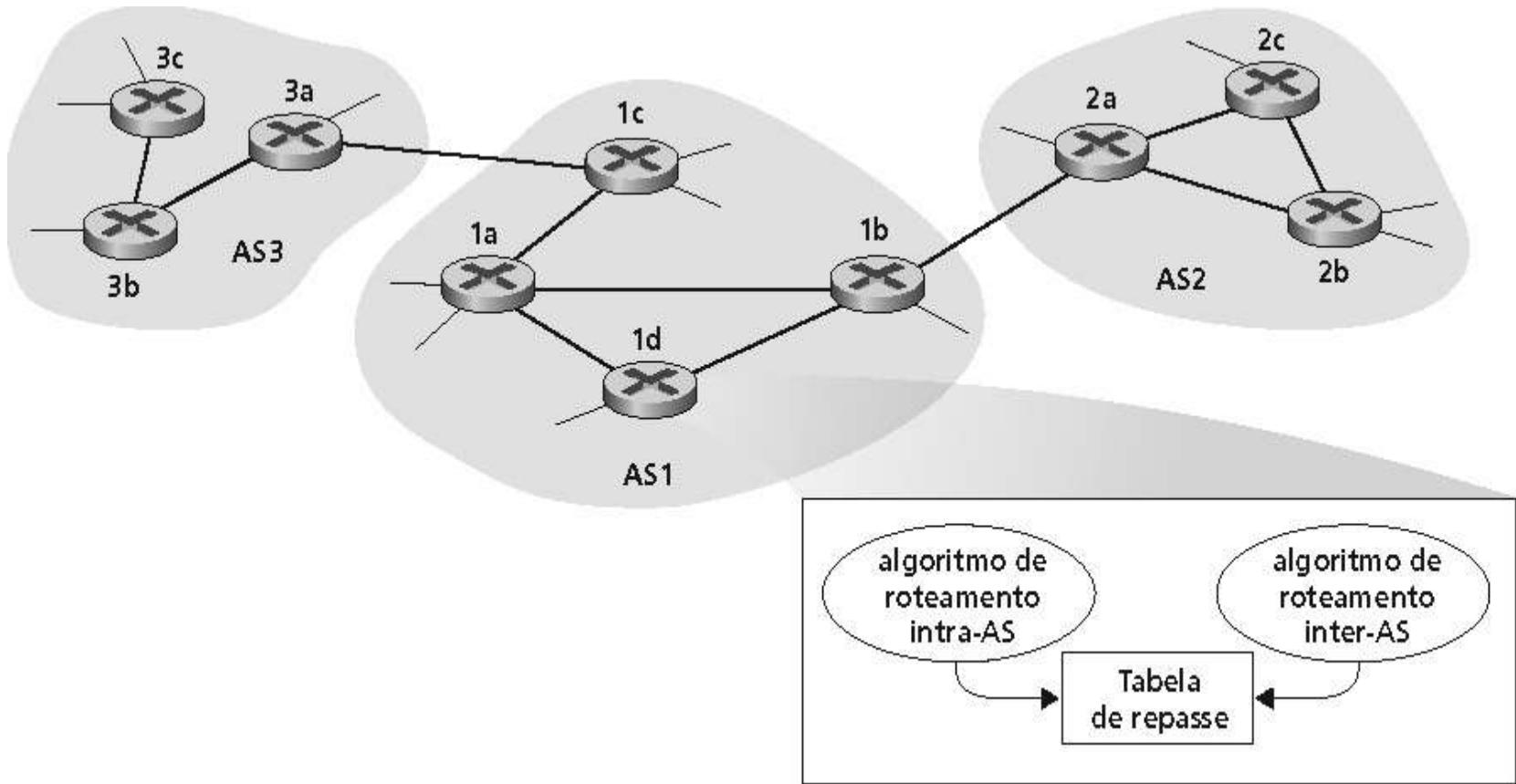
# Roteador (gateway exterior) Inter-AS



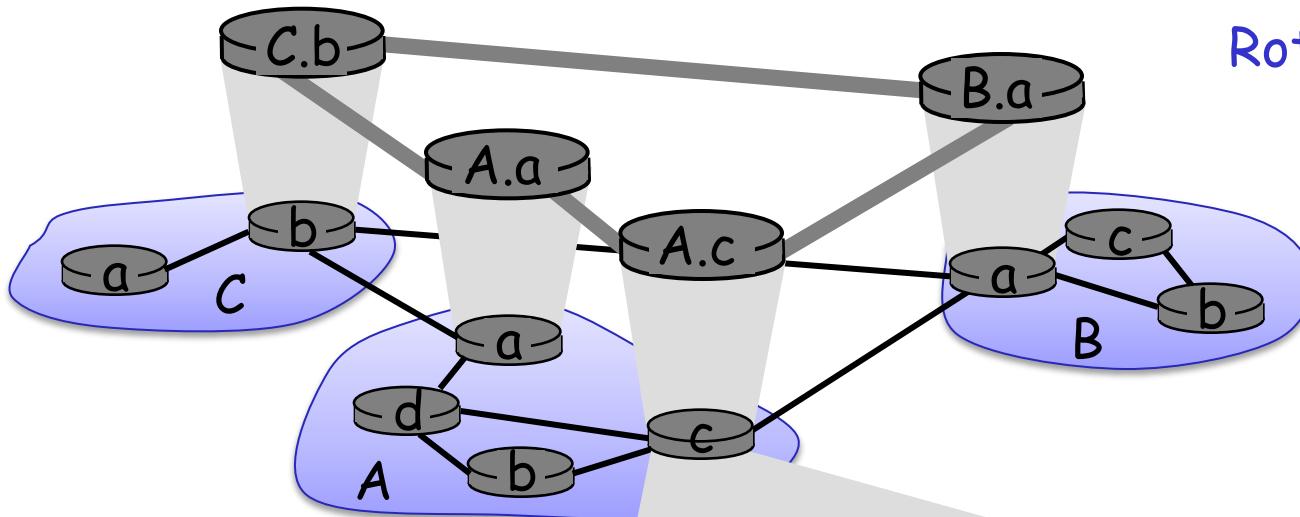
# Roteador (gateway interior) Intra-AS

# ASs interconectadas

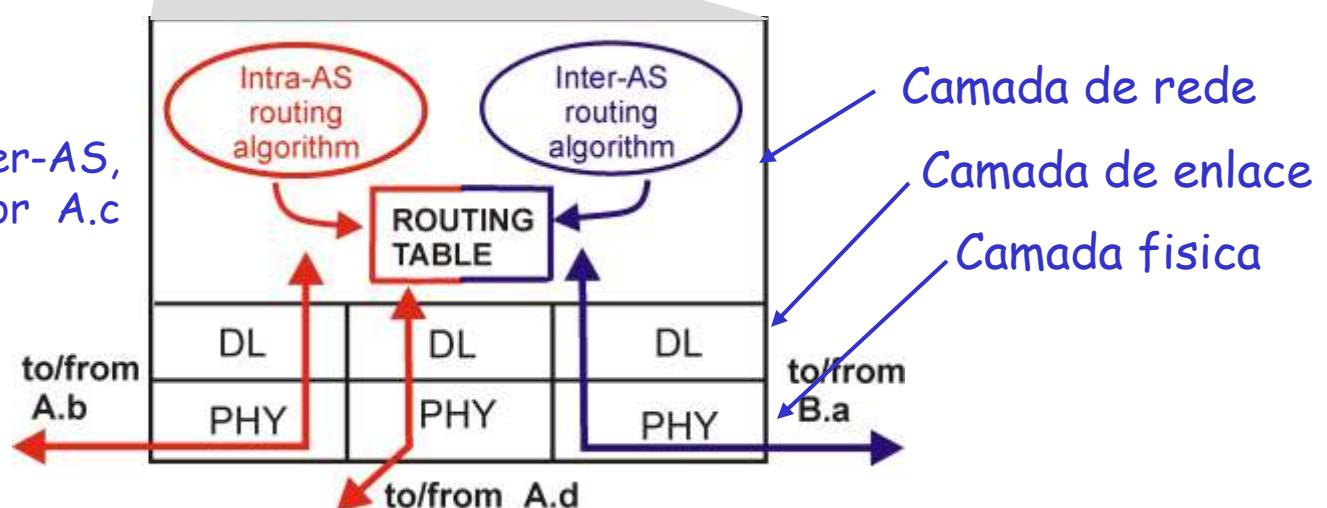
- Tabela de roteamento é configurada por ambos os algoritmos, intra e inter-AS
  - Intra-AS estabelece entradas para destinos internos
  - Inter-AS e intra-AS estabelecem entradas para destinos externos



# Roteamento Intra-AS e Inter-AS



Roteamento inter-AS,  
intra-AS no roteador A.c

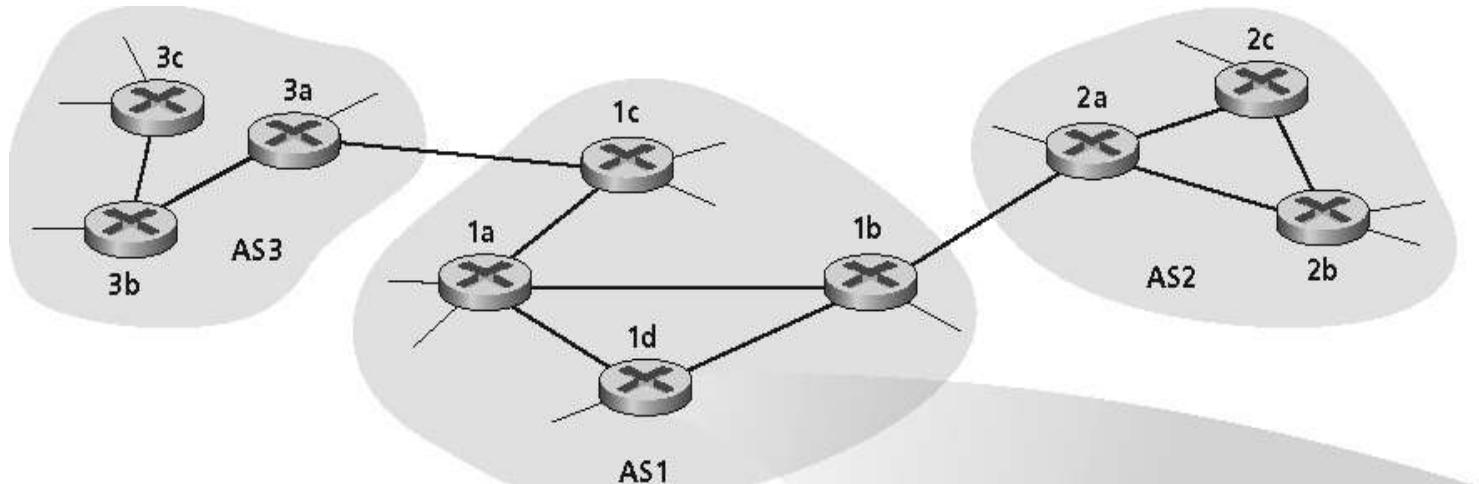


Roteadores de Borda

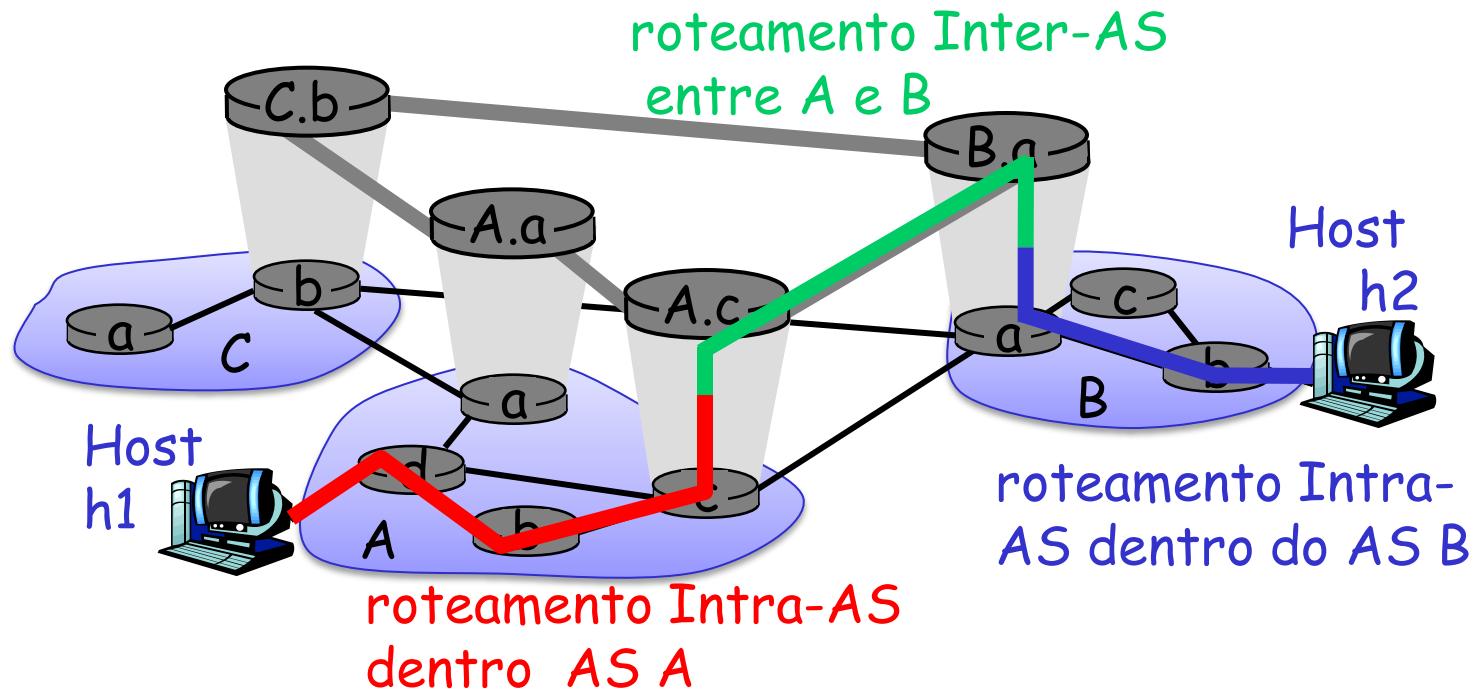
- realizam roteamento inter-AS entre si
- realizam roteamento intra-AS com outros roteadores do mesmo AS

# Tarefas Inter-AS

- Suponha que um roteador no AS1 receba um datagrama cujo destino seja fora do AS1
  - O roteador deveria encaminhar o pacote para os roteadores gateway, mas qual deles?
  - AS1 precisa:
    - 1. Aprender quais destinos são alcancáveis através de AS2 e através de AS3.
    - 2. Propagar suas informações de alcance para todos os roteadores em AS1.
  - Tarefa para o roteamento inter-AS!

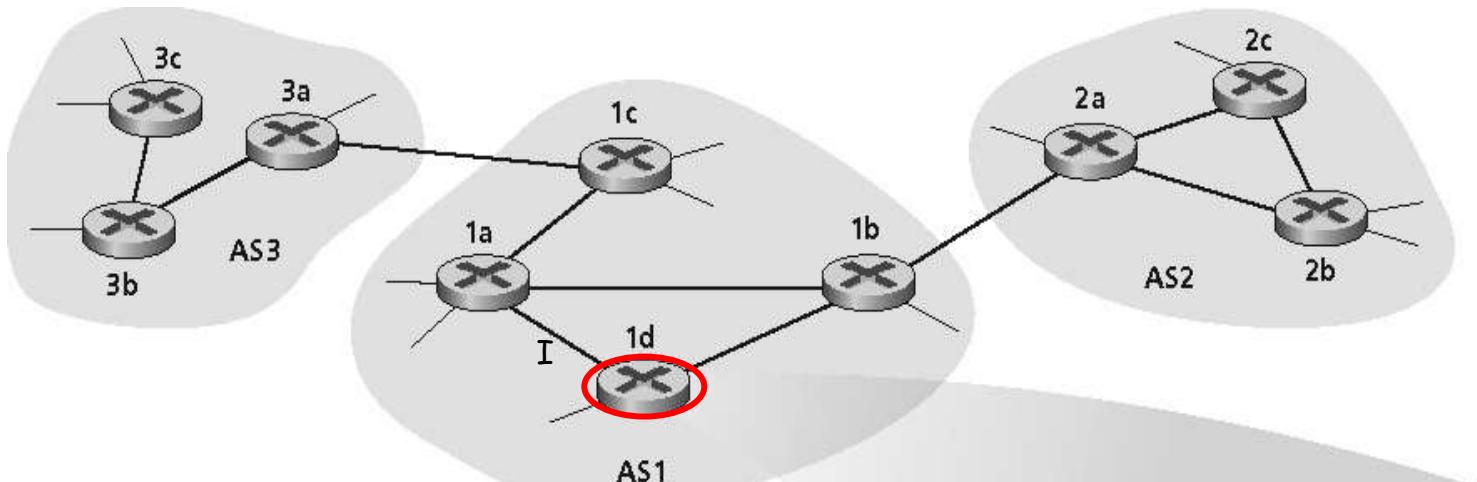


# Roteamento Intra-AS e Inter-AS



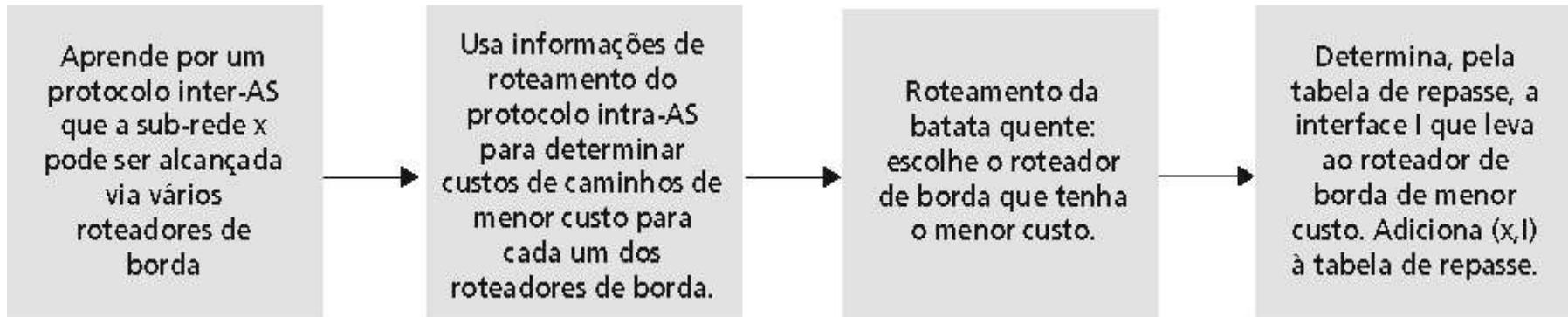
# Exemplo: Ajustando a tabela de roteamento no roteador 1d

- Suponha que AS1 aprende pelo protocolo inter-AS que a sub-rede x é alcancável através de AS3 (gateway 1c) mas não através de AS2
  - O protocolo inter-AS propaga informações de alcance para todos os roteadores internos
  - Baseado nas informações de roteamento intra-AS, o roteador 1d determina que sua interface I está no caminho de menor custo para 1c
  - Coloca na tabela de roteamento a entrada ( $x, I$ )



# Exemplo: Escolhendo entre múltiplas ASs

- Agora suponha que AS1 aprende pelo protocolo inter-AS que a sub-rede x é alcançável através de AS3 e através de AS2.
  - Para configurar a tabela de roteamento, o roteador 1d deve determinar por qual gateway ele deve encaminhar os pacotes para o destino x.
  - Roteamento de "batata quente": envia o pacote para o mais próximo de dois roteadores
    - Para o roteador com o caminho de menor custo na AS1 (via informações do protocolo intra-AS)



# Roteamento

## □ Algoritmos de roteamento

- Link state
- Distance vector
- Roteamento hierárquico

## □ Roteamento na Internet

- RIP
- OSPF
- BGP

# Roteamento na Internet:

## Intra-AS

- Também conhecidos como IGP (Interior Gateway Protocols)
  - atuam dentro de um único domínio de roteamento, ou um único Autonomous System (AS)
- IGPs mais comuns:
  - RIP: Routing Information Protocol
  - OSPF: Open Shortest Path First
  - IGRP: Interior Gateway Routing Protocol (Cisco)

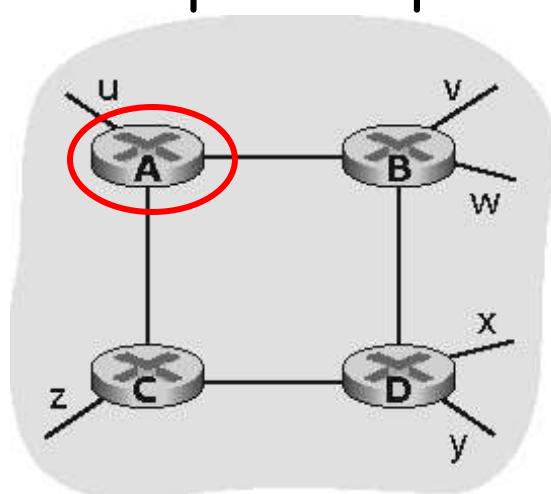
# Roteamento na Internet:

## Inter-AS

- Também conhecidos como EGP (External Gateway Protocols)
  - atuam entre Autonomous System (AS)
- Protocolo único
  - EGP (Obsoleto)
  - BGP-4

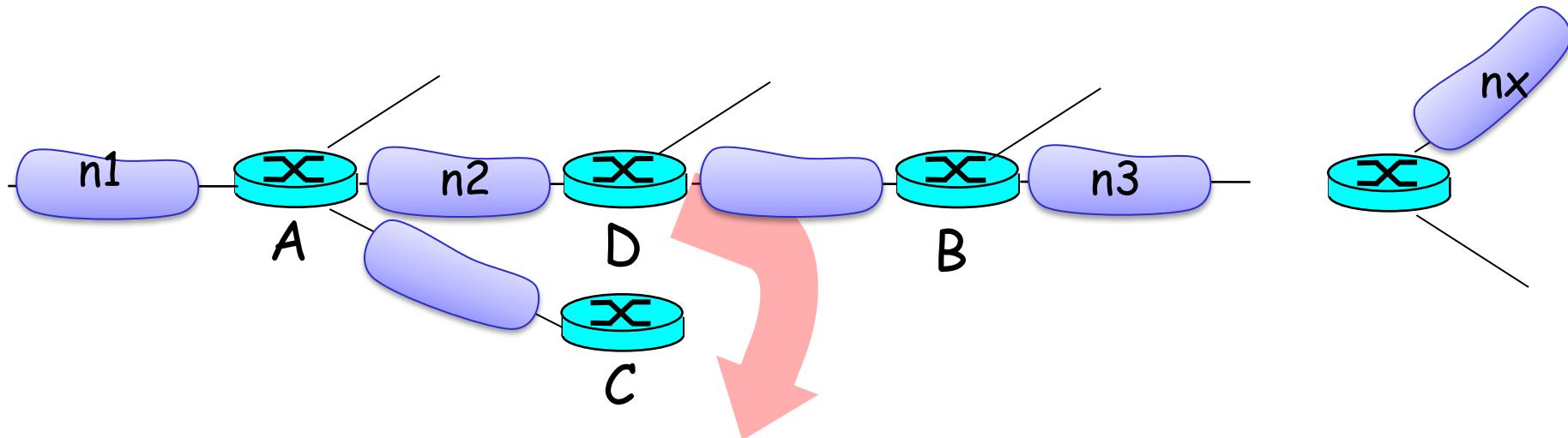
# RIP ( Routing Information Protocol)

- RIP-1 (RFC 1058), RIP-2 (RFC 1388)
- Métrica do custo é a distância
  - calculada através da soma do número de roteadores existentes no caminho (# de hops, max=15 hops)
    - 1: Diretamente conectado
    - 16: infinito
  - RIP não pode suportar redes com diâmetro > 15.



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

# RIP (Routing Information Protocol)



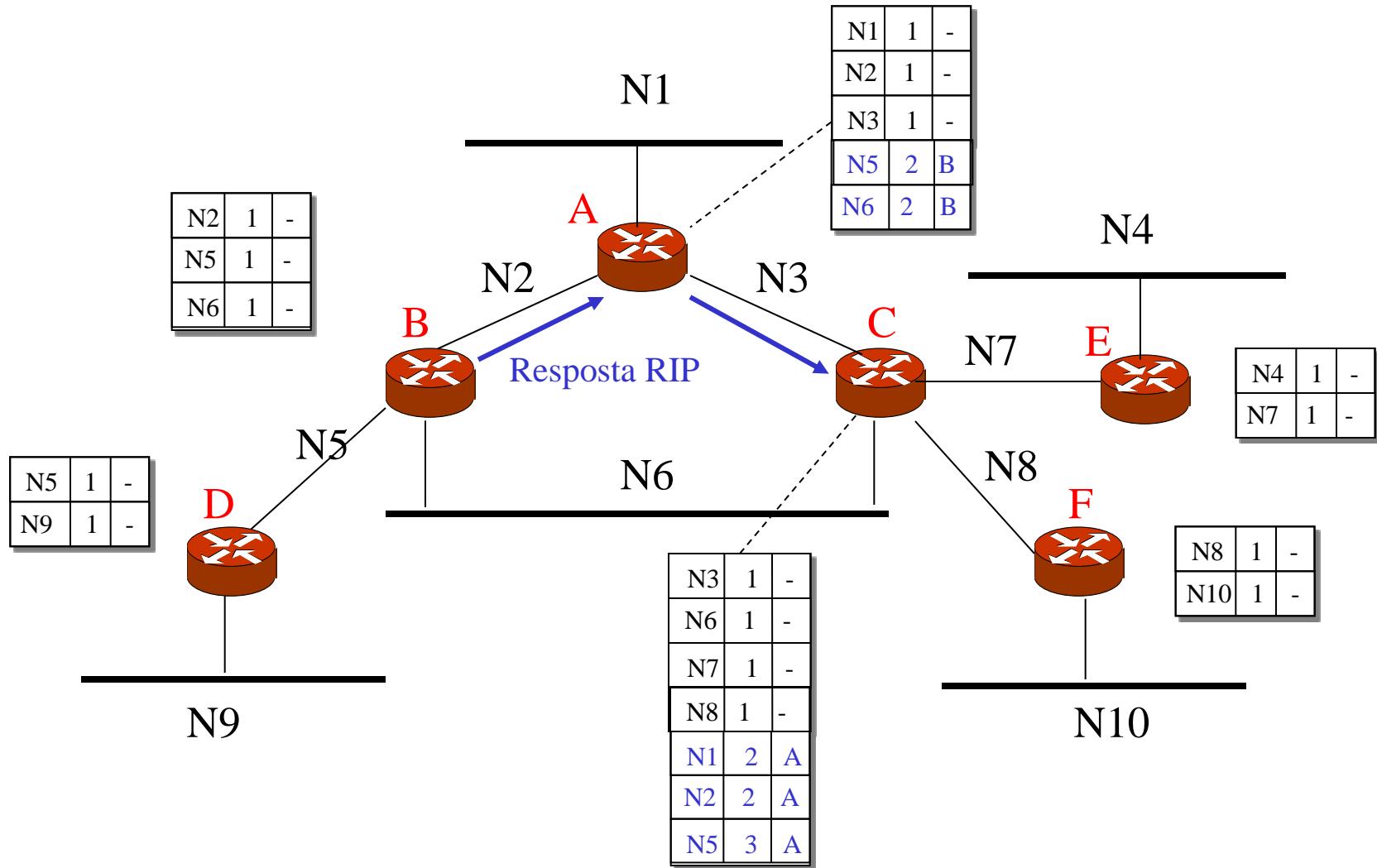
Rede de Destino	Próx. Roteador	Num hops até destino
n1	A	2
n3	B	2
nx	B	7
n2	--	1
....	....	....

Tabela de Roteamento em D

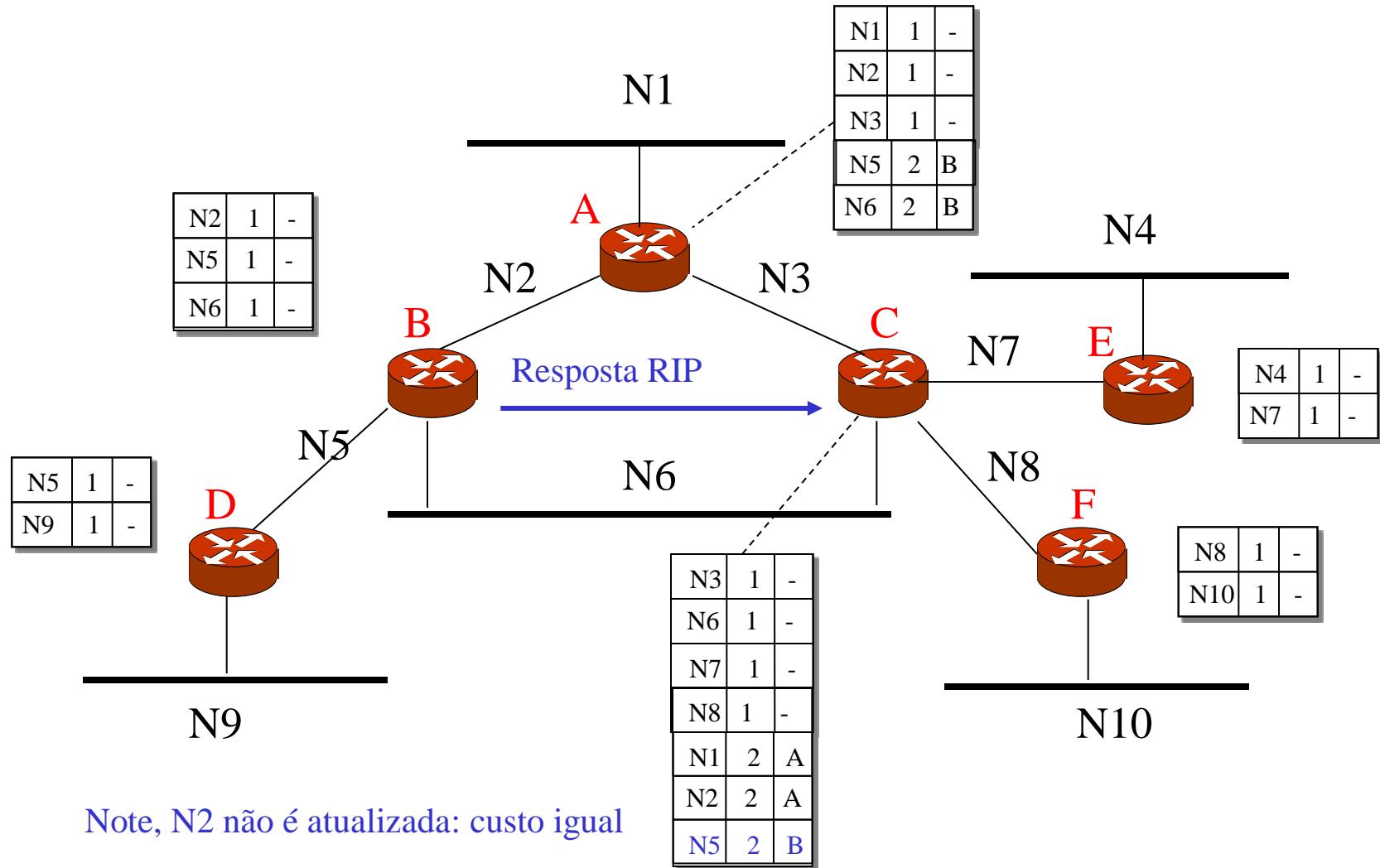
# RIP

- RIP usa vetor de distâncias
  - Todo nó envia suas rotas para seus vizinhos
    - A cada 30 segundos via Response Message (também chamado **advertisement**, ou anúncio)
    - Contendo a rota e o custo (número de hops)
  - Informações de rota gradualmente se espalham através da rede
  - Todo nó seleciona a rota com a menor métrica
- Mensagens rip são enviadas via datagramas UDP
  - IP Multicast (RIP-2) or Broadcast (RIP-1)

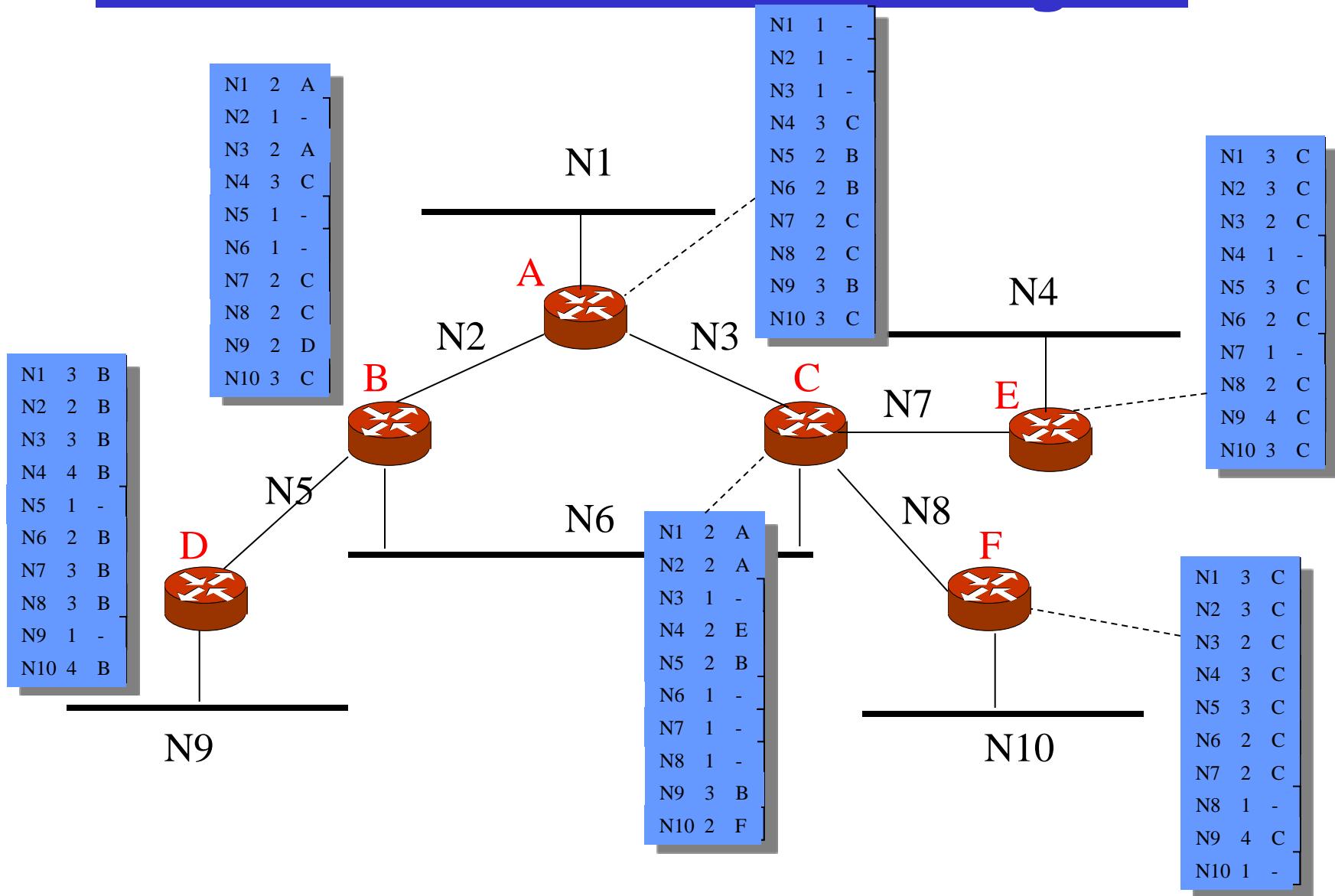
# RIP: Vetor de Distância



# Vetor Distância



# Vetor Distância - Convergido

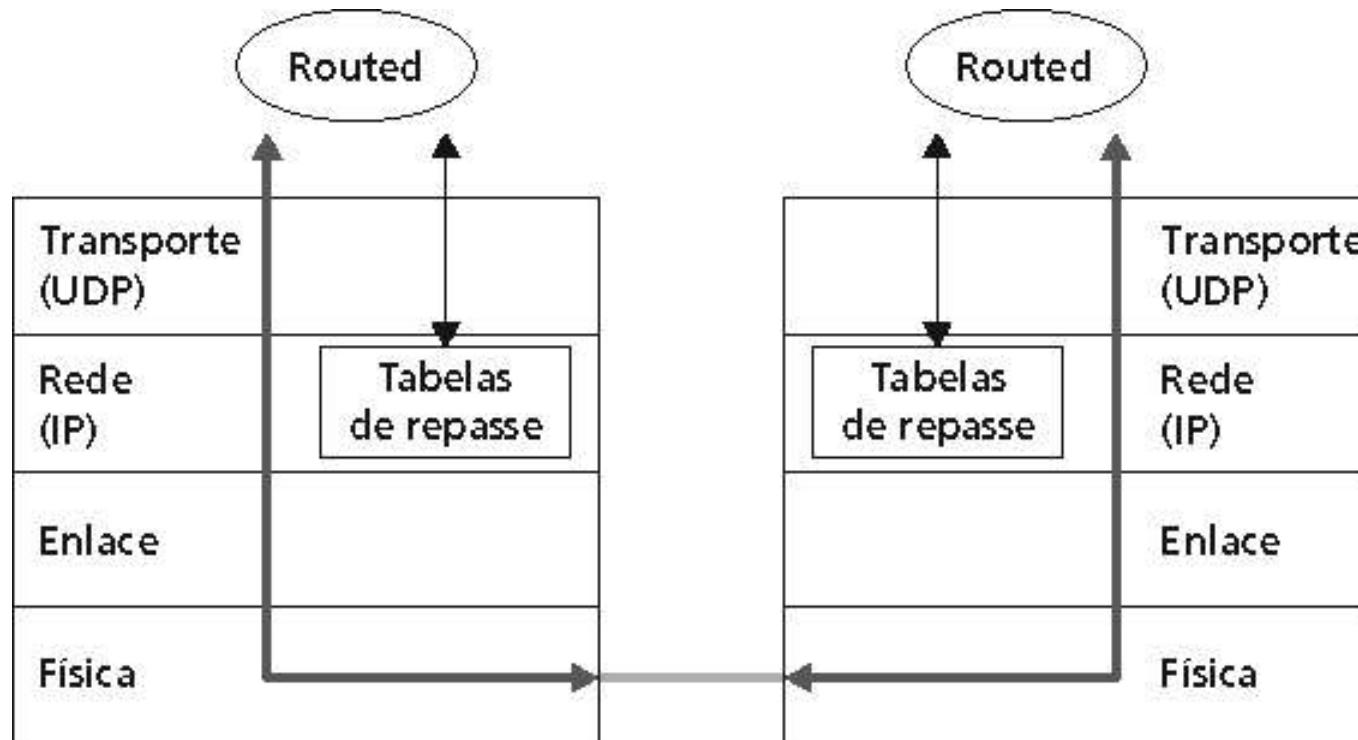


# RIP: falha de enlaces e recuperação

- 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)

# Processamento da tabela RIP

- Tabelas de roteamento RIP são gerenciadas por um processo no nível aplicação chamado route-d (daemon)
- Avisos são enviados em pacotes UDP a cada 30 segundos



# Problemas do RIP

- Uso de contagem de hosts pode não ser a métrica adequada para definir a rota de menor custo
  - Nem sempre traz bons resultados
    - # saltos = 3 passando por Ethernets
    - # saltos = 2 passando por linhas seriais lentas!
  - Muitas implementações do RIP permitem que o gerenciador sete artificialmente os custos
- Após a falha de uma rota ou link
  - RIP pode levar minutos para estabilizar
  - Cada vizinho fala apenas todo 30s
    - O tempo para a informação propagar por vários hops é na ordem de minutos

# Problemas do RIP

- A métrica máxima útil é 15
  - Diâmetro da rede dever ser menor ou igual a 15
- RIP aceita atualizações de qualquer um
  - Um dispositivo mal configurado pode gerar problemas para toda a rede
- RIP usa muita largura de banda
  - Ele envia tabelas de roteamento completas para atualizações

# Porque RIP ainda é usado?

## Resposta

- Pois RIP geralmente é disponível
- É simples de configurar
- É gratuito

# OSPF: Open Shortest Path First

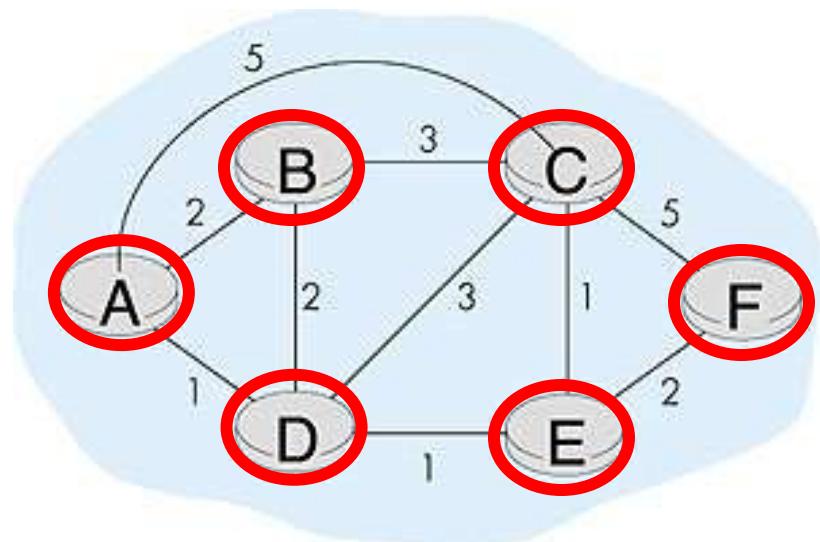
- Protocolo recomendado pela IETF para ser usado com IGP
  - Desenvolvido para substituir o RIP (1988-1991)
  - OPEN por ser um padrão aberto
- Protocolo SPF (Link State)
  - Cada roteador tem informações completas da topologia
    - Um mapa indicando os roteadores e as redes às quais estão conectados
  - Roteador executa duas tarefas
    - Testa o status de todos os roteadores próximos (se conectam a uma rede comum)
    - Difunde periodicamente informações de status de enlace para os demais roteadores
- OSPF usa IP diretamente (Campo Protocol = 89)
  - Não usa UDP ou TCP.
  - Usa multicast

# OSLP: Protocolo Link-State (SPF)

- No SPF, todo roteador faz o seguinte:
  - Testa ativamente o status de todos os vizinhos
    - Constrói um pacote Link State (LSP) com esta informação e propaga para todos os outros roteadores
  - Usando LSPs vindo de todos os outros roteadores, computa uma árvore de envio de caminho de menor custo usando o algoritmo de Dijkstra.
- Vantagens (sobre o vetor de distância):
  - Mais funcionalidade devido ao cálculo na origem do dado e não dependência dos roteadores intermediários
    - Conhecimento da topologia completa
    - Facilidade de recuperação de falhas
  - Convergência rápida

# Exemplo: Algoritmo de Dijkstra

Passo	N'	D(B),p(B)	D(C),p(C)	D(D),p(D)	D(E),p(E)	D(F),p(F)
0	A	2,A	5,A	1,A	$\infty$	$\infty$
1	AD	2,A	4,D		2,D	$\infty$
2	ADE	2,A	3,E			4,E
3	ADEB		3,E			4,E
4	ADEBC					4,E
5	ADEBCF					



# Algoritmo de Dijkstra: exemplo

Passo	N	D(B),p(B)	D(C),p(C)	D(D),p(D)	D(E),p(E)	D(F),p(F)
0	A	2,A	5,A	1,A	$\infty$	$\infty$
1	AD	2,A	4,D		2,D	$\infty$
2	ADE	2,A	3,E			4,E
3	ADEB		3,E			4,E
4	ADEBC					4,E
5	ADEBCF					

Árvore de caminhos mínimos resultante originada em A:

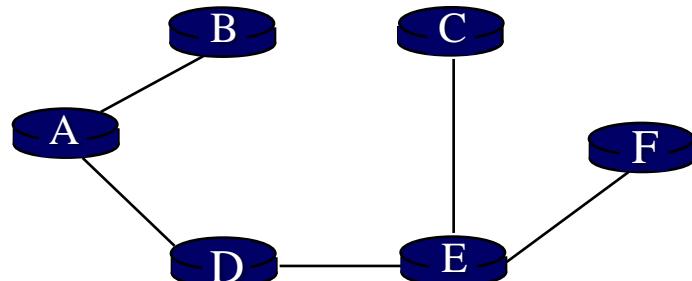


Tabela de encaminhamento resultante em A:

destino	enlace
B	(A,B)
C	(A,D)
D	(A,D)
E	(A,D)
F	(A,D)

# OSLP: Protocolo Link-State (SPF)

- Especifica que toda troca de informações entre roteadores seja autenticada
  - Apenas roteadores confiáveis difundem as informações sobre o roteamento
- Desvantagem
  - Usa mais memória

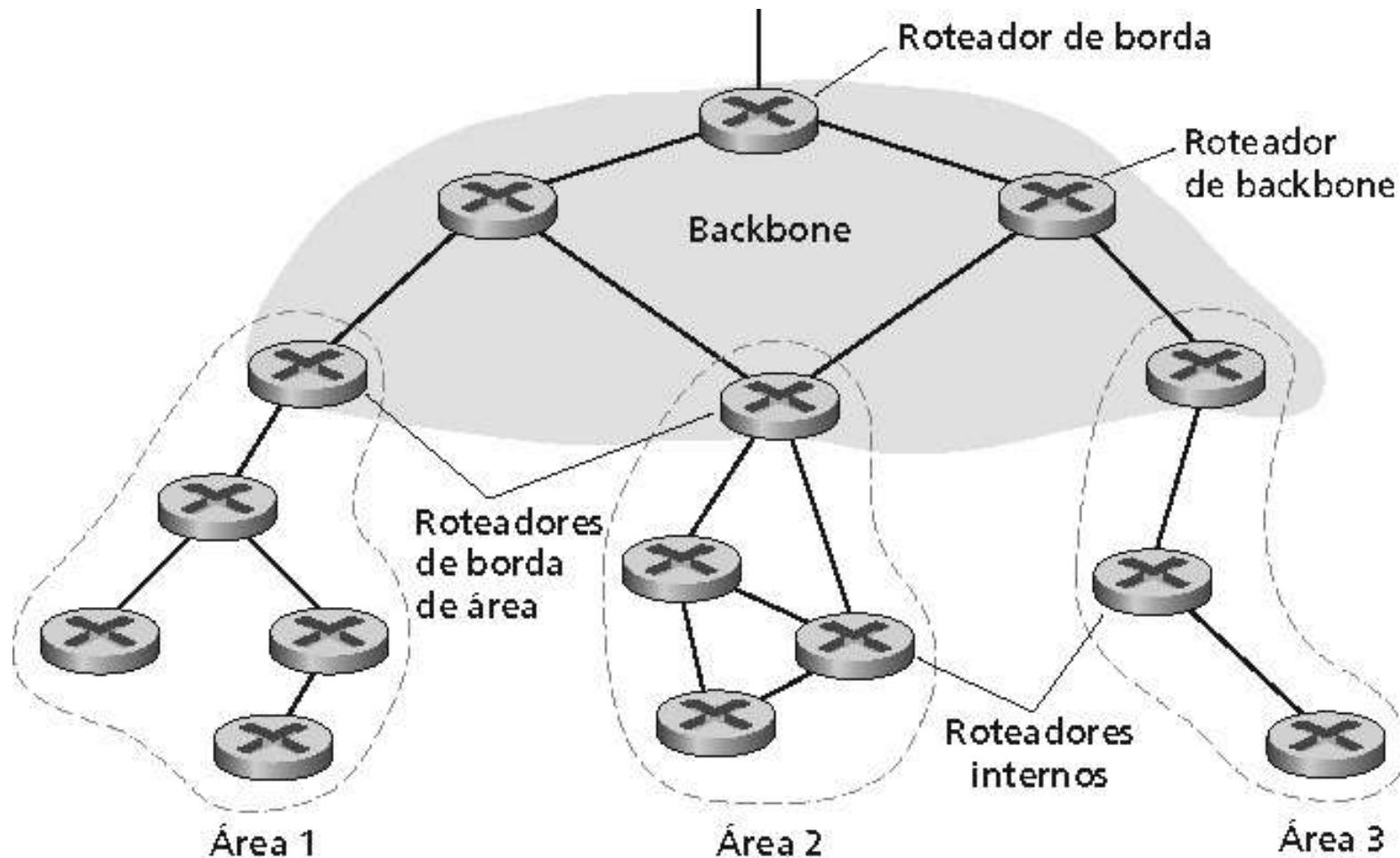
# OSPF: Open Shortest Path First

- Roteamento por tipo de serviço
  - Gerenciadores podem instalar várias rotas para um destino específico
    - uma para cada tipo de serviço
      - Retardo de transmissão baixo ou vazão alta, etc...
  - Define a rota a partir dos campos TOS (DS) e o endereço de destino

# OSPF: Open Shortest Path First

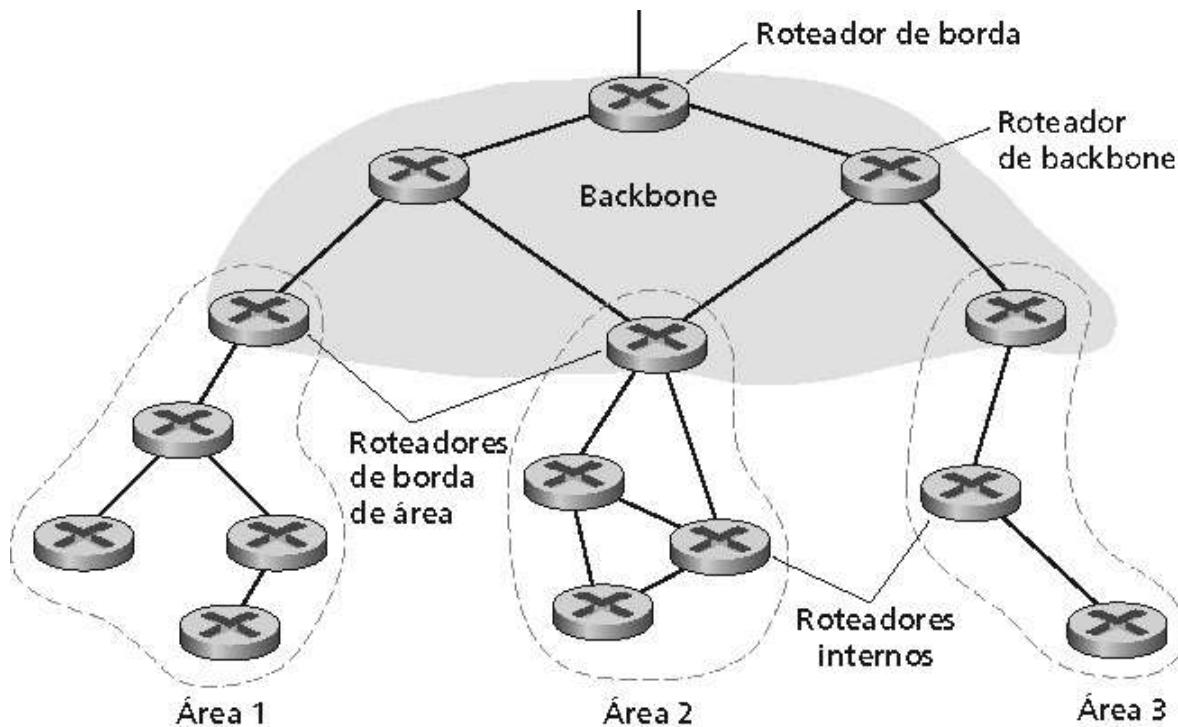
- Fornece balanceamento de carga
  - Se o gerenciador especificar várias rotas para determinado destino ao mesmo custo
    - OSPF distribui o tráfego por todas as rotas igualmente
    - RIP calcula uma rota por destino
- Permite o crescimento e faz com que as redes sejam fáceis de gerenciar
  - Permite que um domínio particione suas redes e roteadores em subconjuntos chamados de áreas

# Hierarquia OSPF



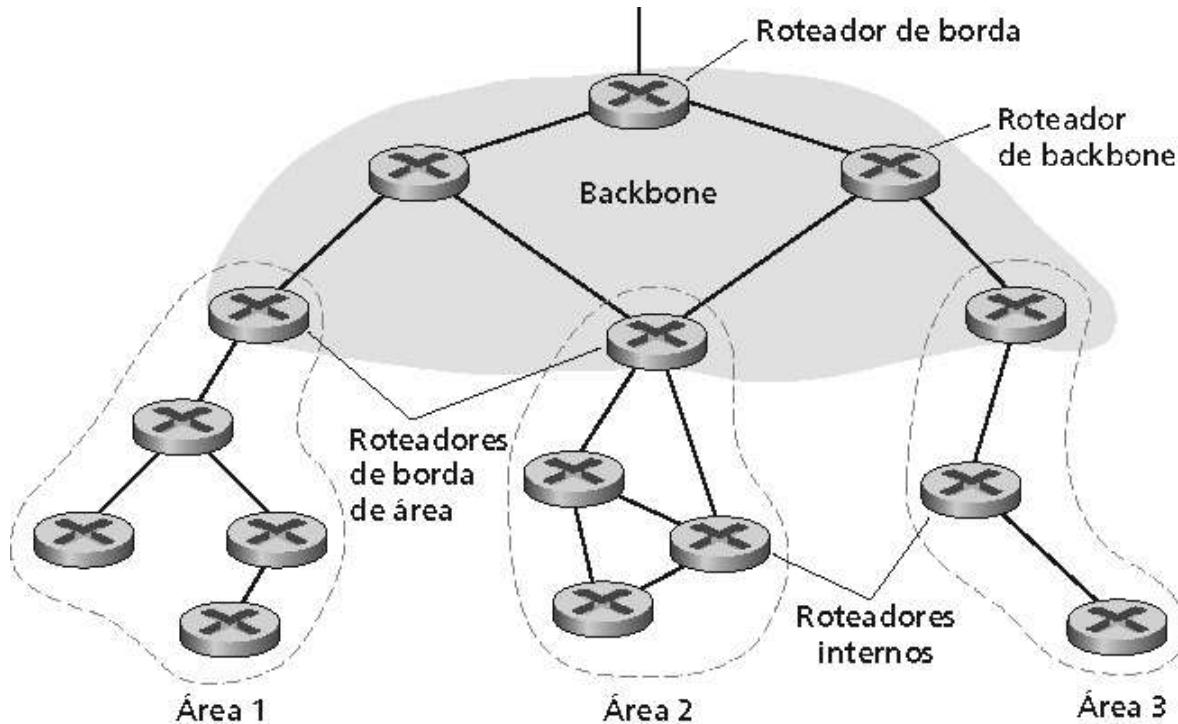
# OSPF Hierarquia

- Hierarquia de dois nível: local area, backbone.
- Avisos de estado dos links apenas na área
- Cada nó conhece a topologia da área; apenas conhece a direção (caminho mais curto) para as outras áreas



# OSPF Hierarquia

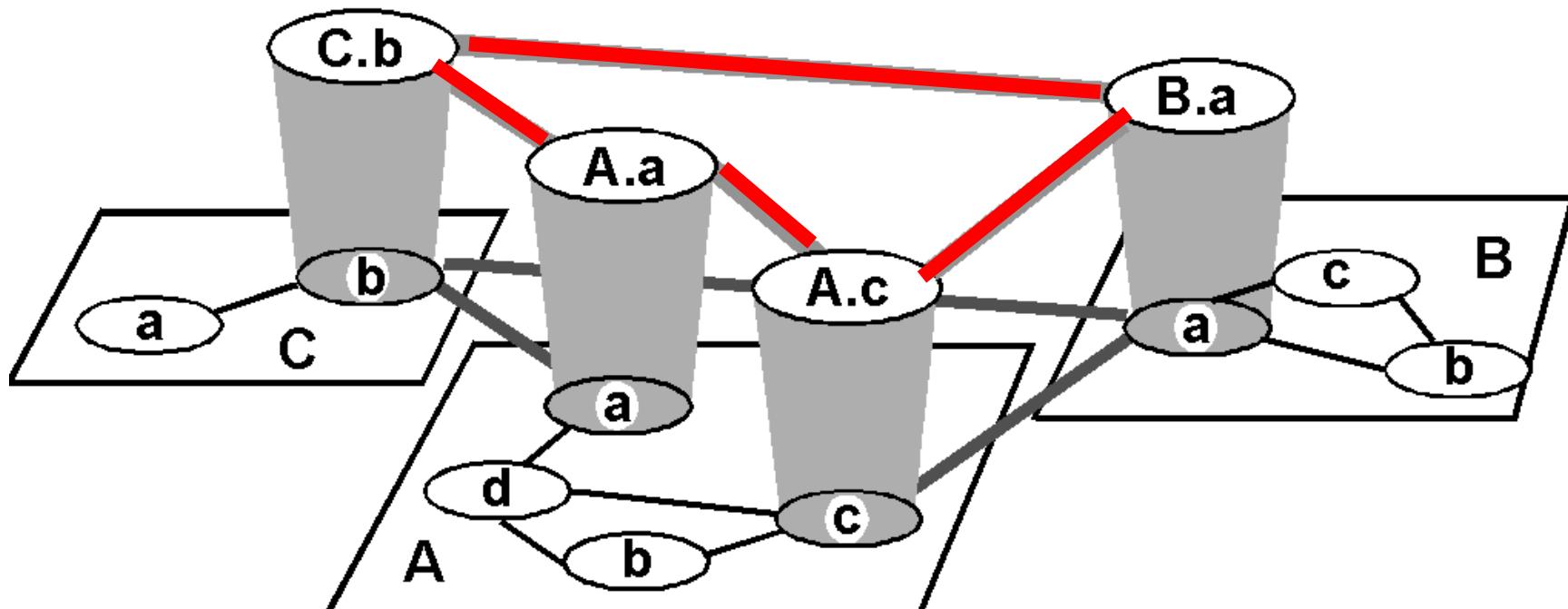
- Roteadores na borda da área: "resumem" distâncias para redes na sua própria área, avisa aos outros roteadores de borda
- Roteadores Backbone: executam o roteamento OSPF limitado no backbone
- Roteadores fronteira: conecta a outras ASs.



# IGRP (Interior Gateway Routing Protocol)

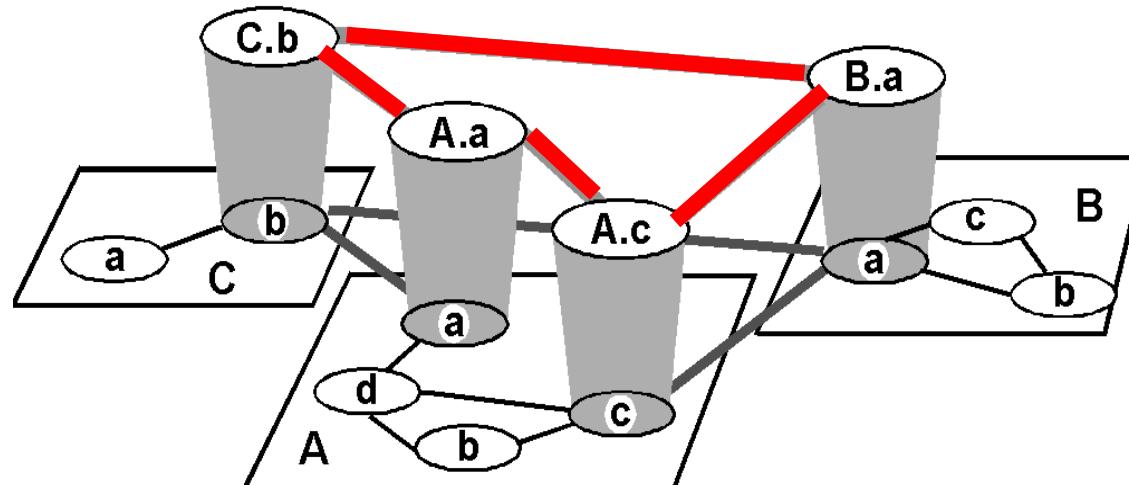
- Protocolo proprietário da CISCO
  - sucessor do RIP (meados dos anos 80)
- Vetor de distância, como RIP
  - várias métricas de custo (atraso, banda, confiabilidade, carga, etc.)
- Usa o TCP para trocar informações de novas rotas

# Roteamento Inter-AS



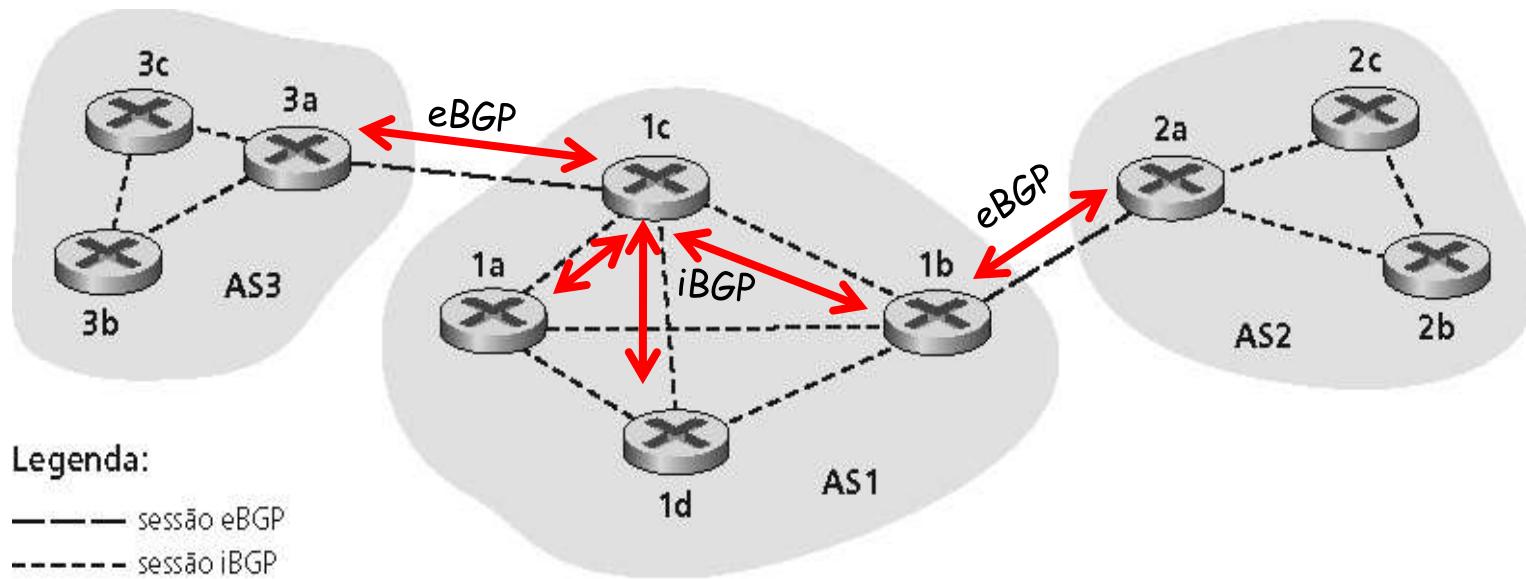
# Roteamento Internet inter-AS: BGP

- BGP (Border Gateway Protocol): é o padrão de fato para uso na Internet
  - BGP provê a cada AS 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 subrede comunique sua existência para o resto da Internet: “Estou aqui”



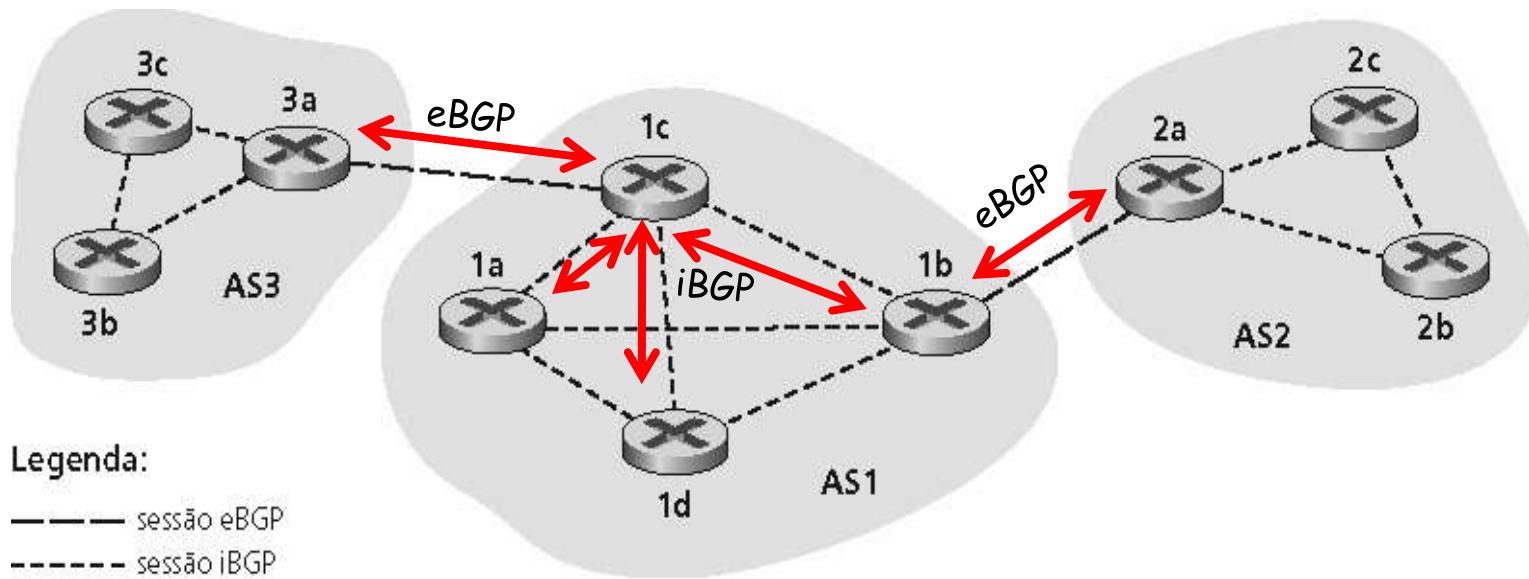
# BGP: conceitos básicos

- Pares de roteadores (BGP peers) trocam informações de roteamento por conexões TCP semi-permanentes: sessões BGP
- Quando AS2 comunica um prefixo ao AS1, AS2 está prometendo que encaminhará todos os datagramas destinados a esse prefixo
  - AS2 pode agragar prefixes em seu comunicado



# 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.

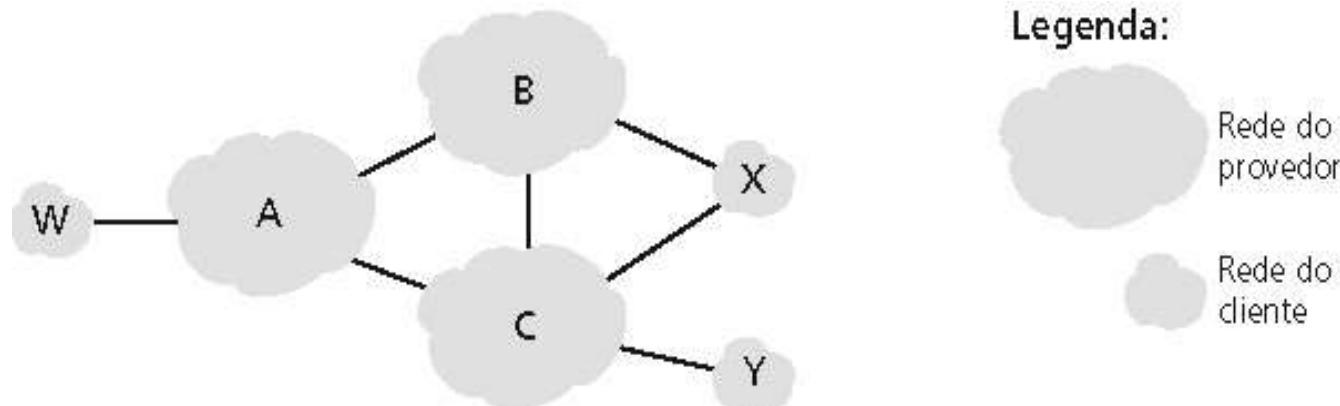


# Atributos de caminho e rotas BGP

- Quando um roteador de borda comunica um prefixo, ele informa também alguns atributos do BGP.
- Dois atributos importantes:
  - AS-PATH: contém os ASs (AS number) 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.

# BGP: seleção de rota

- Um roteador pode aprender mais de uma 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



# Roteamento Internet inter-AS:

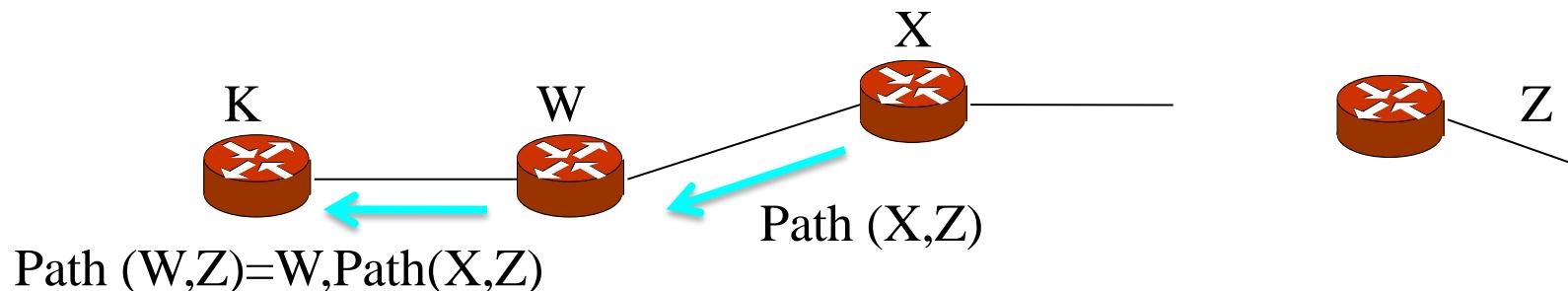
## BGP

### □ Protocolo Path Vector

- Similar ao protocolo Distance Vector
- Cada Border Gateway difunde para seus vizinhos (pares) o caminho completo (seqüência de ASs) para o destino  $x$
- Exemplo: Gateway  $X$  pode enviar seu caminho para  $Z$ :
  - $\text{Path}(X,Z) = X, Y_1, Y_2, Y_3, \dots, Z$

# Internet inter-AS routing: BGP

- Suponha que o roteador X enviou seu caminho para Z ao roteador par W ( $\text{Path}(X,Z)$ )
  - W pode ou não aceitar o caminho oferecido por X
    - custo, política (não rotear via AS de competidor), razões de prevenção de loops.
  - Se W seleciona o caminho avisado por X, então:
  - $\text{Path}(W,Z) = w, \text{Path}(X,Z)$
  - Nota: X pode controlar o tráfego que chega controlando seus avisos de rota para os pares
    - e.g., não quer receber o tráfego para Z  $\rightarrow$  não avisar nenhuma rota para Z



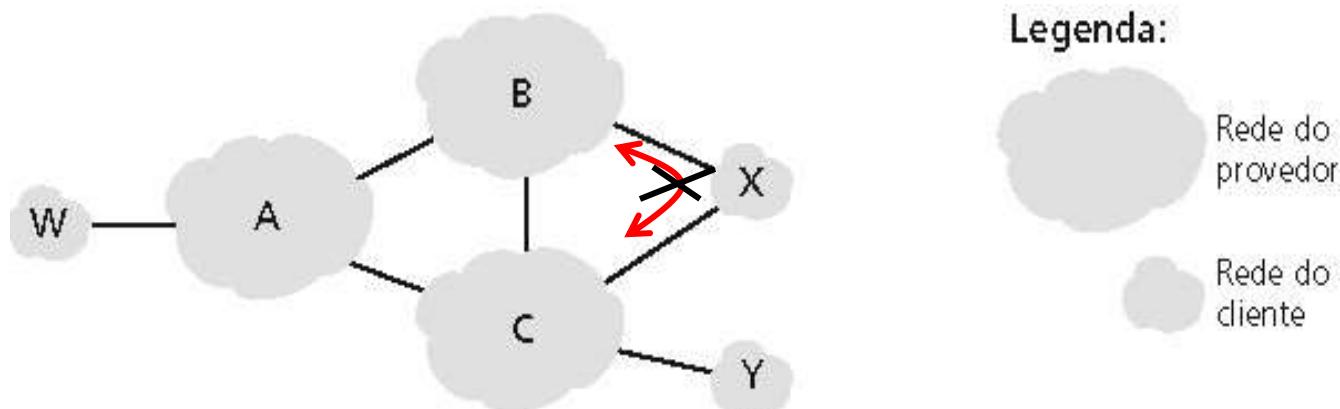
# BGP: política de roteamento

## □ Legenda

- A, B, C são redes de provedores
- X, W, Y são clientes (das redes dos provedores)

## □ X é dual-homed: anexados a duas redes

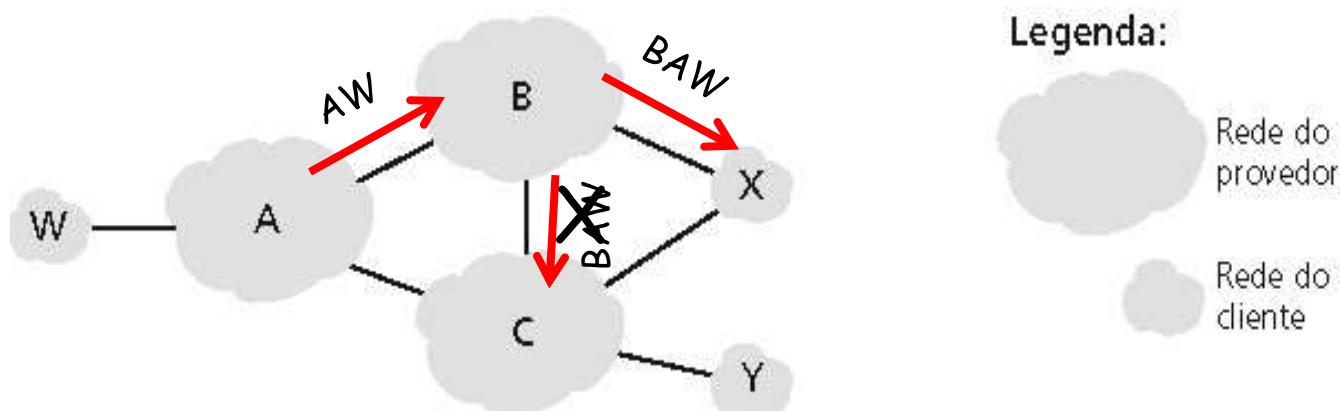
- X não quer rotear de B para C
- ... então X não comunicará ao B uma rota para C



# BGP: política de roteamento (2)

## □ Anúncios

- 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 rotear somente de/para seus clientes!
  - B quer forçar C a rotear para W via A
  - Não anuncia esta rota



# Internet inter-AS routing: BGP

- Mensagens BGP são trocadas usando o TCP
- Mensagens:
  - OPEN: abre conexão TCP para o par e autentica o emissor
  - UPDATE: avisa novo caminho (ou relembra antigos)
  - KEEPALIVE mantém conexões vivas na ausência de UPDATES; também confirma pedidos OPEN
  - NOTIFICATION: reporta erros na mensagem anterior; também usado para fechar a conexão

# Porque os protocolos Intra- e Inter-AS são diferentes ?

Problemas do mundo real

- roteadores não são todos idênticos
- as redes não são "flat" na prática

**Escala:** com 50 milhões de destinos:

- Não é possível armazenar todos os destinos numa única tabela de rotas!
- As mudanças na tabela de rotas irão congestionar os enlaces!

**Autonomia Administrativa**

- Internet = rede de redes
- Cada administração de rede pode querer controlar o roteamento na sua própria rede

# Porque 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: as decisões políticas são mais simples

## 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: regras de mercado podem ser mais importantes que desempenho

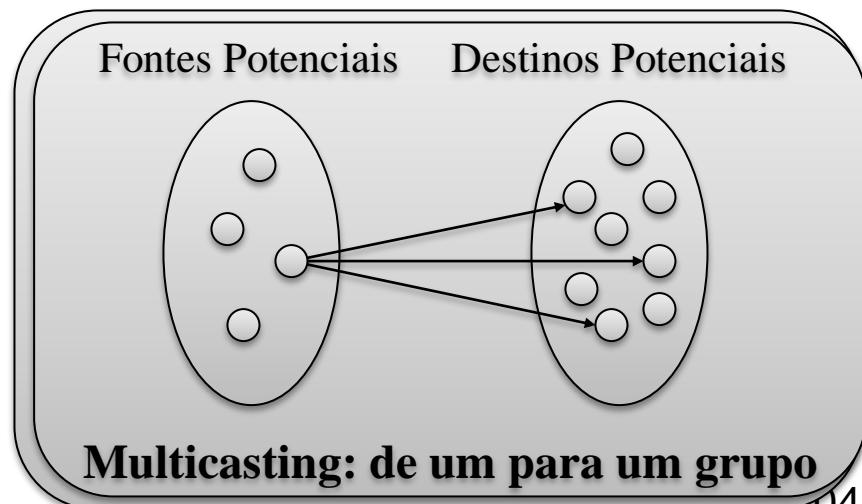
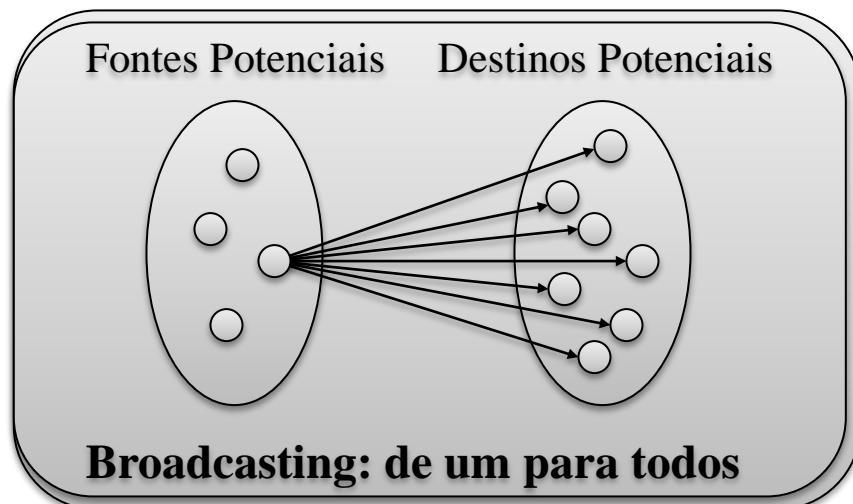
# Multicast

- Protocolos **unicast**: envolvem apenas um transmissor e um receptor.
- **Multicast**: envio de um pacote de um transmissor para múltiplos receptores com apenas uma operação de transmissão.
  - Exemplos:
    - transmissão de uma aula para diversos participantes distribuídos;
    - alimentação de dados: cotações da bolsa de valores;
    - ambientes virtuais interativos distribuídos, etc.

# Multicast

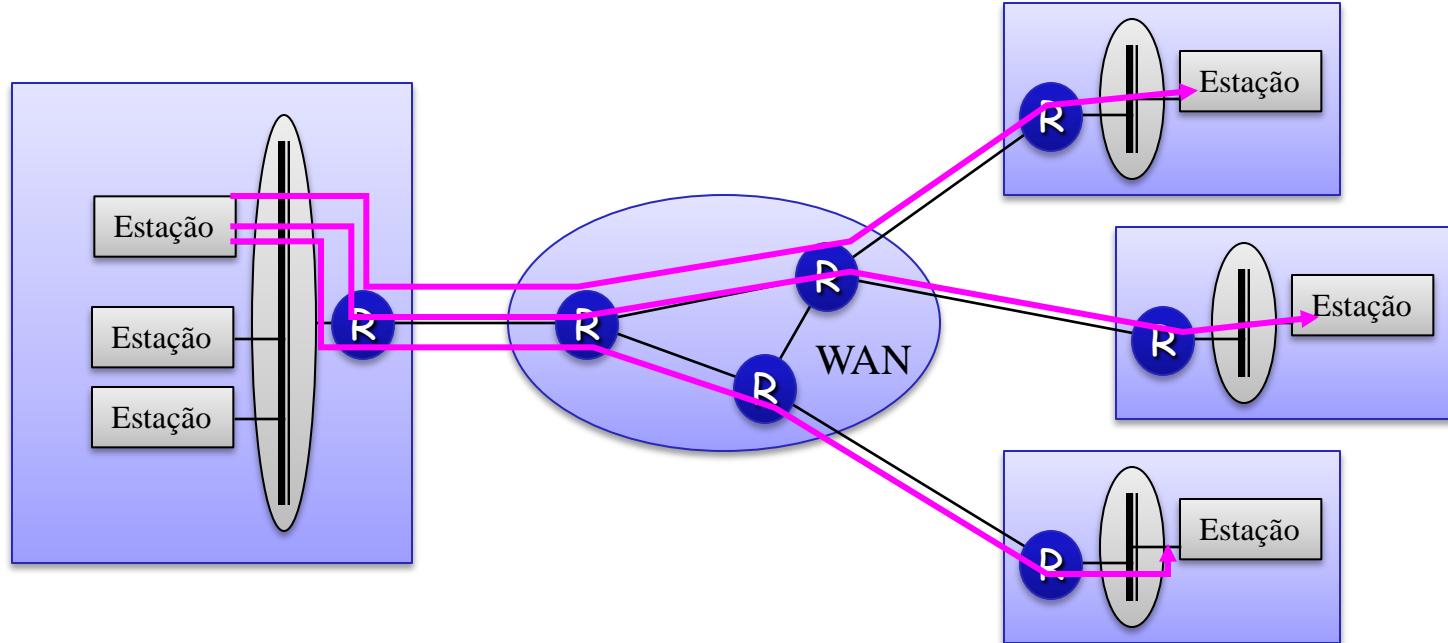
## □ Broadcast e Multicast

- Broadcast: comunicação ponto-a-multiponto onde todos os receptores potenciais podem receber a mensagem
- Multicast: comunicação ponto-a-multiponto onde um grupo de receptores podem receber a mensagem
  - Multicast para grupos fechados: a lista de receptores é predefinida e sobre o controle de uma autoridade central
  - Multicast para grupos abertos: qualquer receptor potencial pode espontaneamente se juntar ou sair do grupo



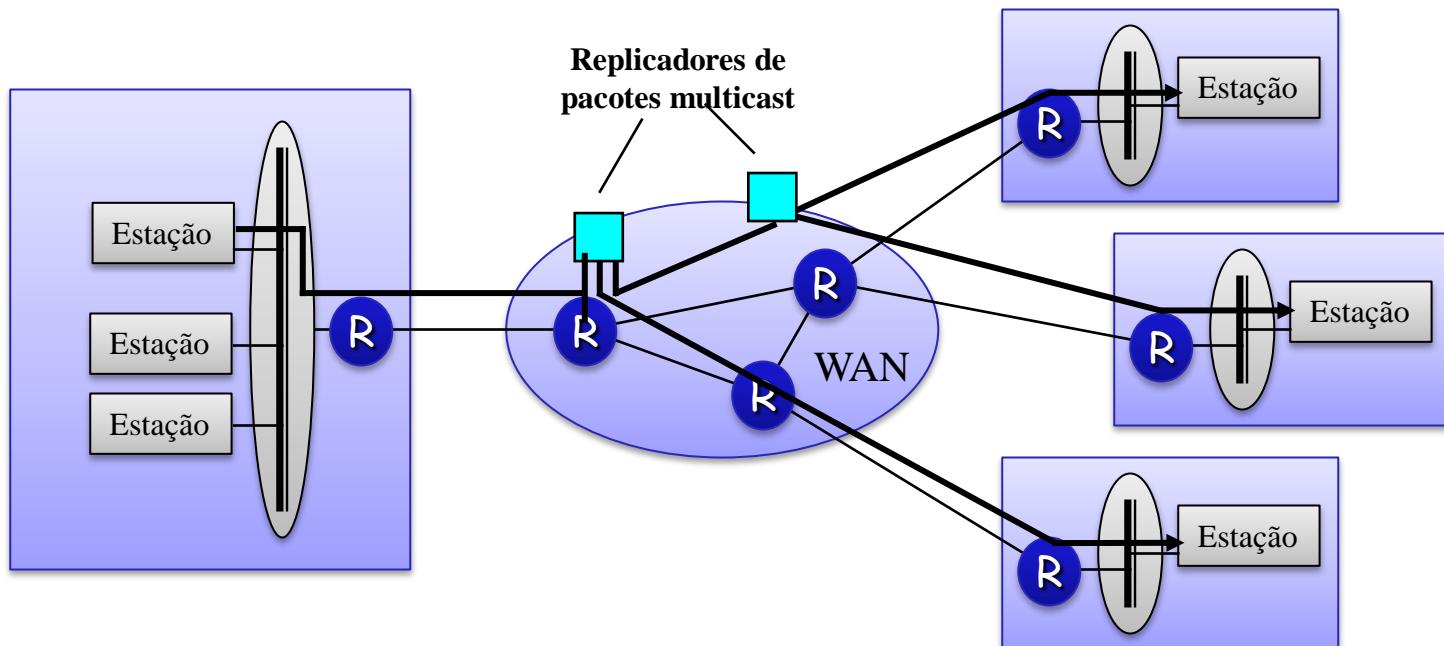
# Multicast

- Múltiplas conexões ponto-a-ponto
  - Conexões ponto-a-ponto são estabelecidas entre os participantes
  - Vantagem
    - confidencialidade é mais alta
  - Desvantagens
    - fonte deve gerar múltiplos fluxos de dados idênticos (um para cada destino)
      - não econômico em termos de processamento na fonte e de comunicação
    - apenas envolve membros especificamente aceitos



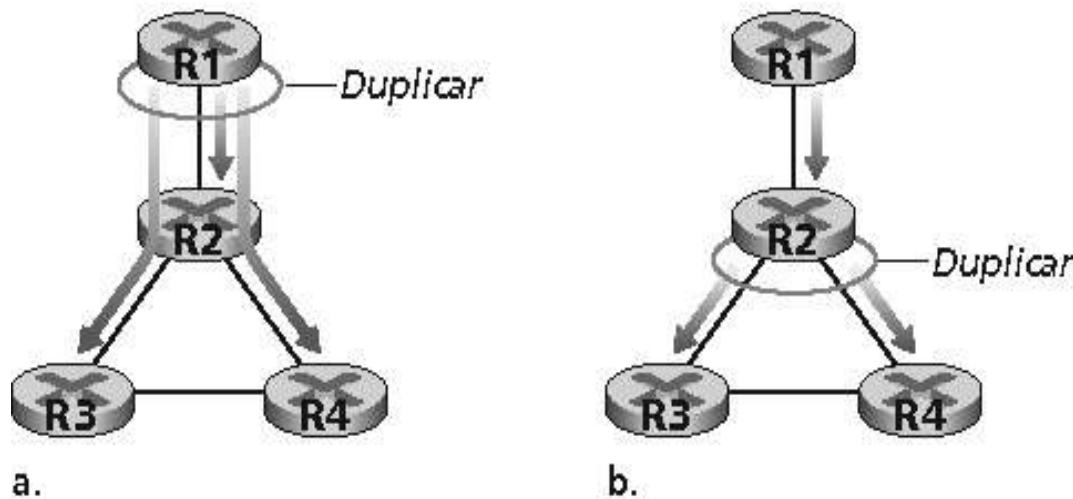
# Multicasting

- Tecnologia multicast LAN/WAN
  - Suportar multicast em LAN é teoricamente simples
    - explorando a capacidade multicast inerente de LAN de meios-compartilhados (Ethernet, Token Ring, FDDI)
  - Sobre WANs, esta funcionalidade é disponível sobre vários tipos de rede sem conexão (IP)
  - Nem todas implementações asseguram confidencialidade
    - uso de técnicas de criptografia



# Multicasting

Criação/transmissão de duplicatas



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

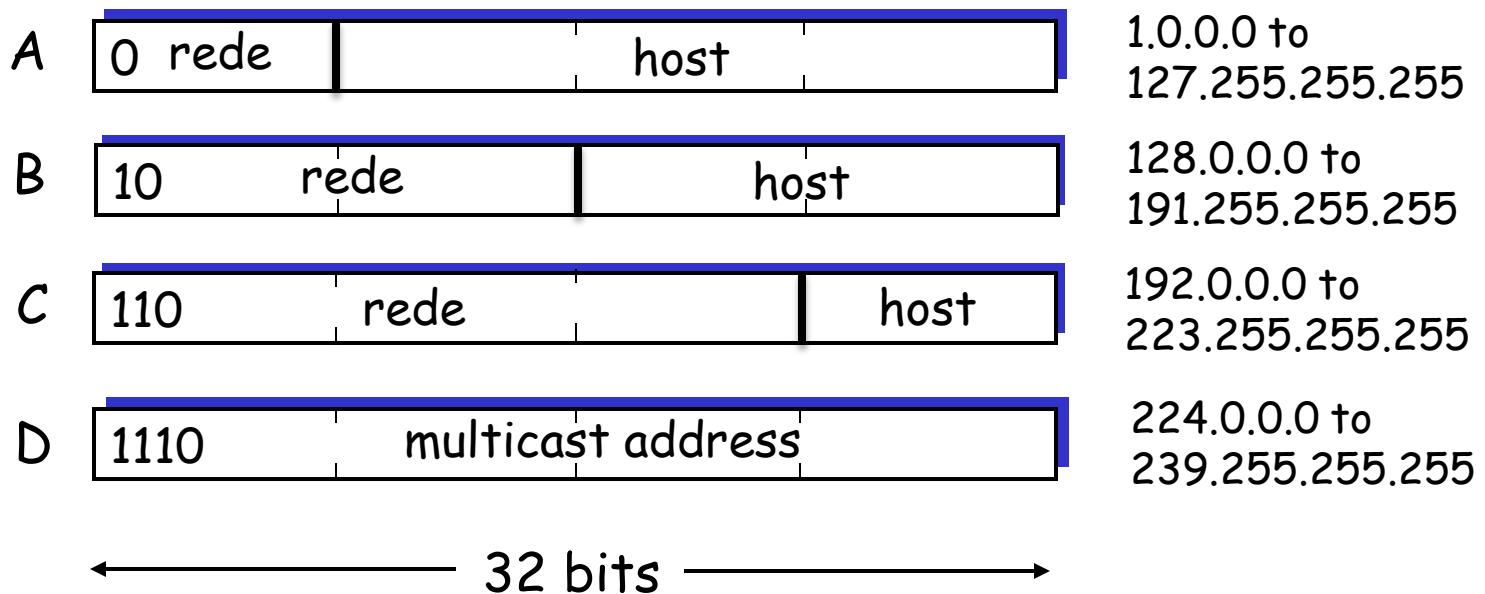
# IP Multicast

- IP multicast é uma extensão do IPv4
  - Permite envio de datagramas IP para um conjunto de máquinas
    - formam um grupo multicast identificado por um endereço IP Classe "D"
- Uso do IP Multicast
  - Videoconferência
  - Rádio e TV Internet
  - Distribuição de informações de roteamento
  - Jogos Distribuídos
- UDP é preferível para aplicações multicast (impossível com TCP)
- Orientado a receptor:
  - Emissores enviam para endereços de grupo
  - Hosts se juntam a grupo multicast
  - Emissores e receptores são distintos: emissores não necessitam ser membros

# Endereços IP

- Endereçamento “class-full”:

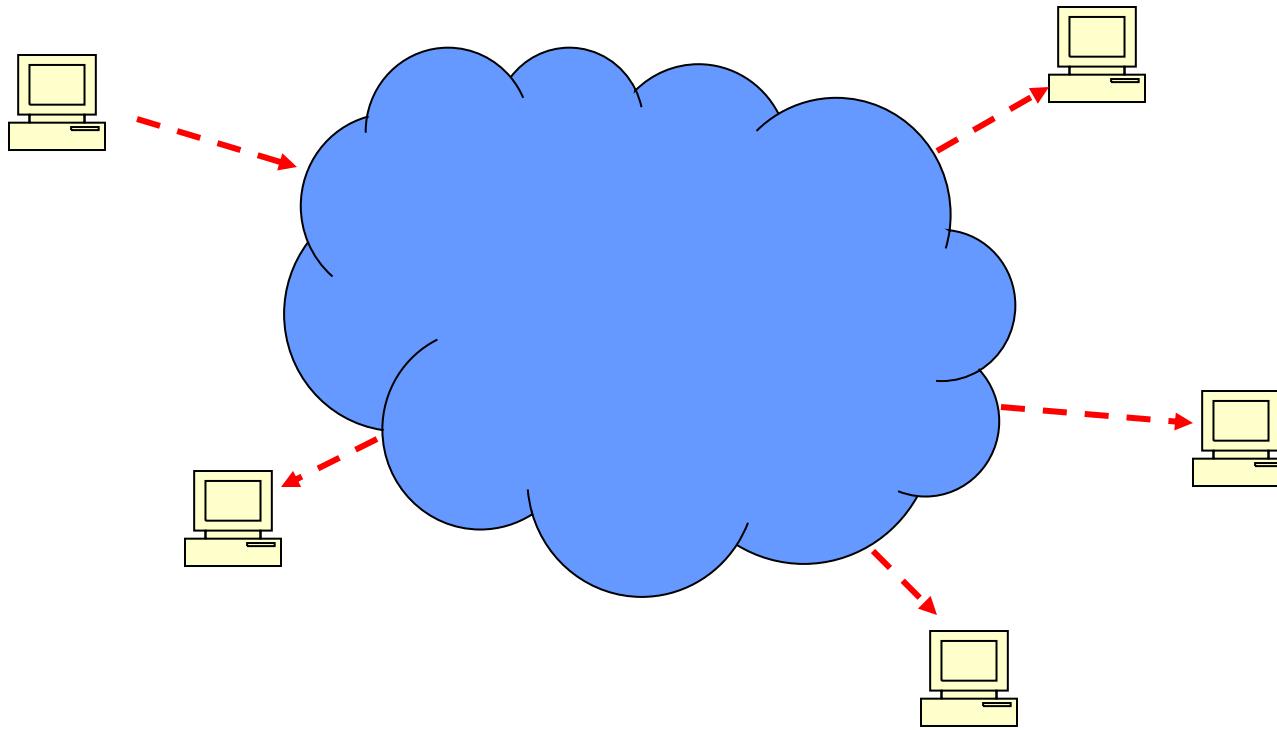
class



# Multicast: aspectos da camada de rede

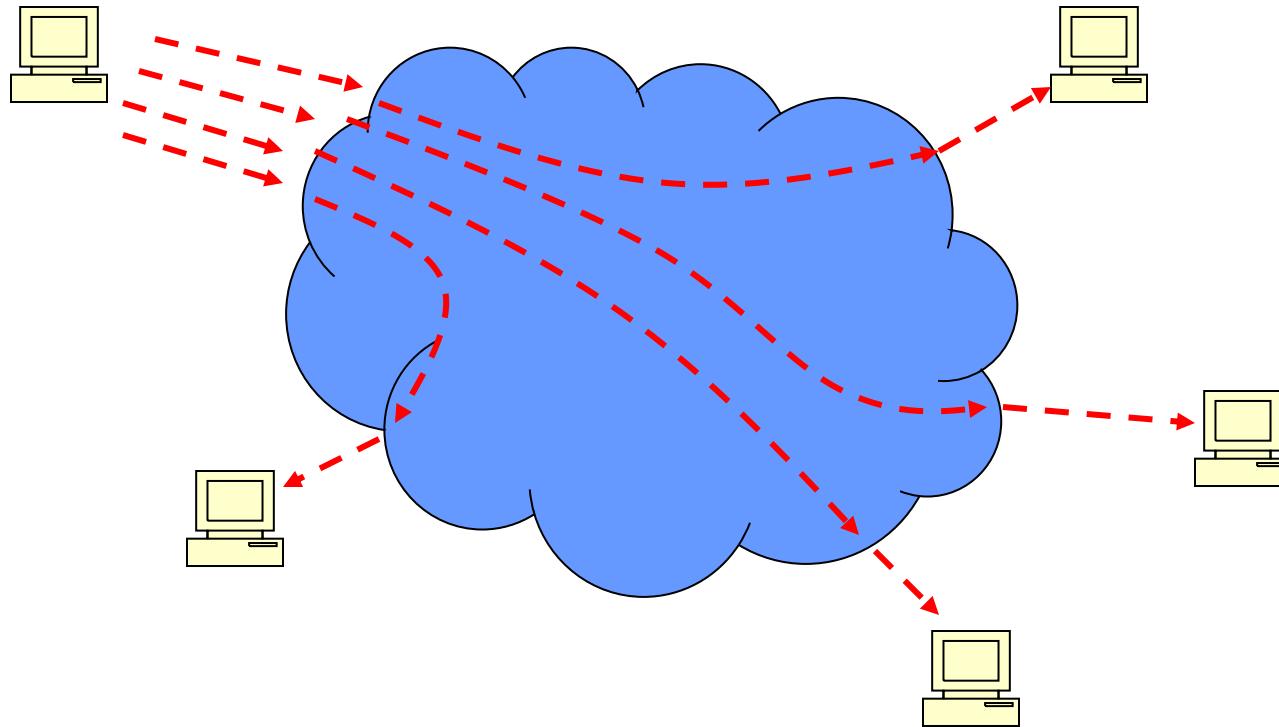
- Algoritmos de roteamento
- Multicast na Internet **não** é um serviço sem conexões:
  - devem ser estabelecidas conexões multicast
  - devem ser mantidas informações de estado das conexões multicast em cada roteador participante da mesma.
  - Necessita de uma combinação de protocolos de sinalização e de roteamento.

# Multicast



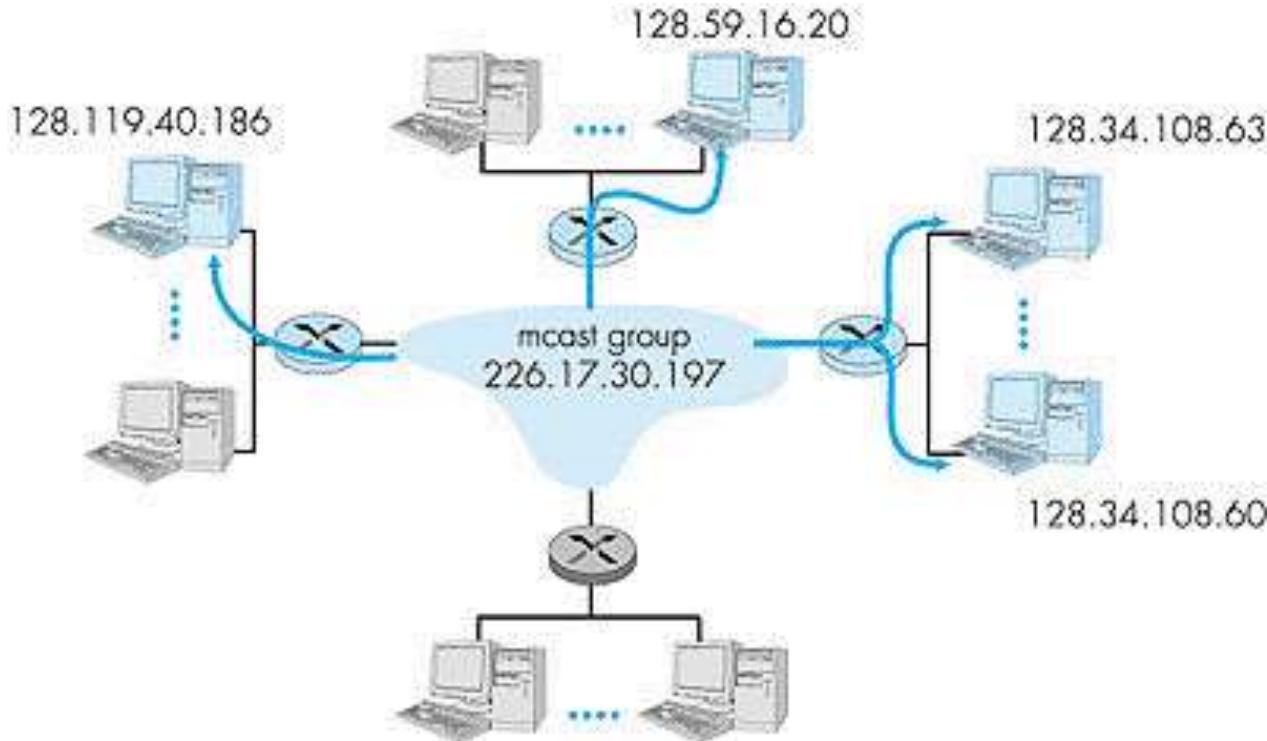
- Um grupo: um emissor e vários receptores
- A rede replica pacotes
- Grupos são dinâmicos: membros entram e saem do grupo

# Usando unicast



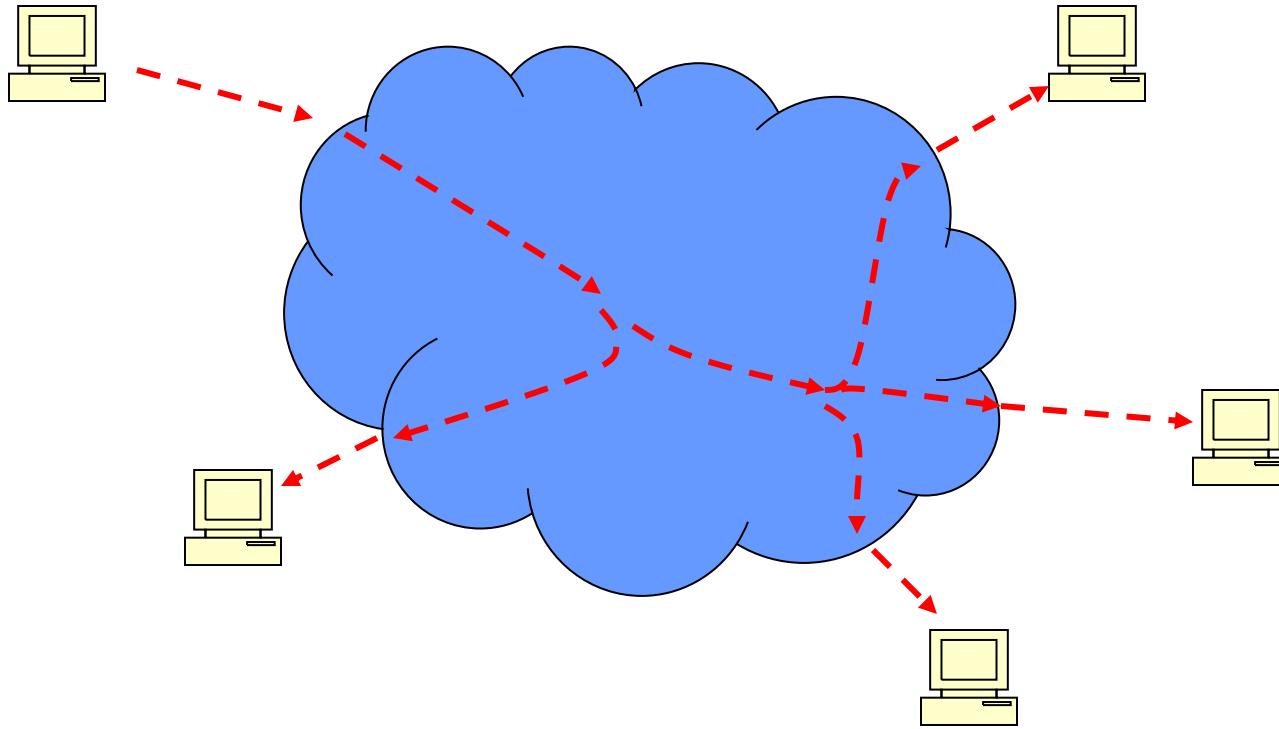
- Usando unicast, o emissor deve replicar as mensagens e enviar o mesmo dado várias vezes pela rede
- Má uso de recursos de rede

# Grupo Multicast



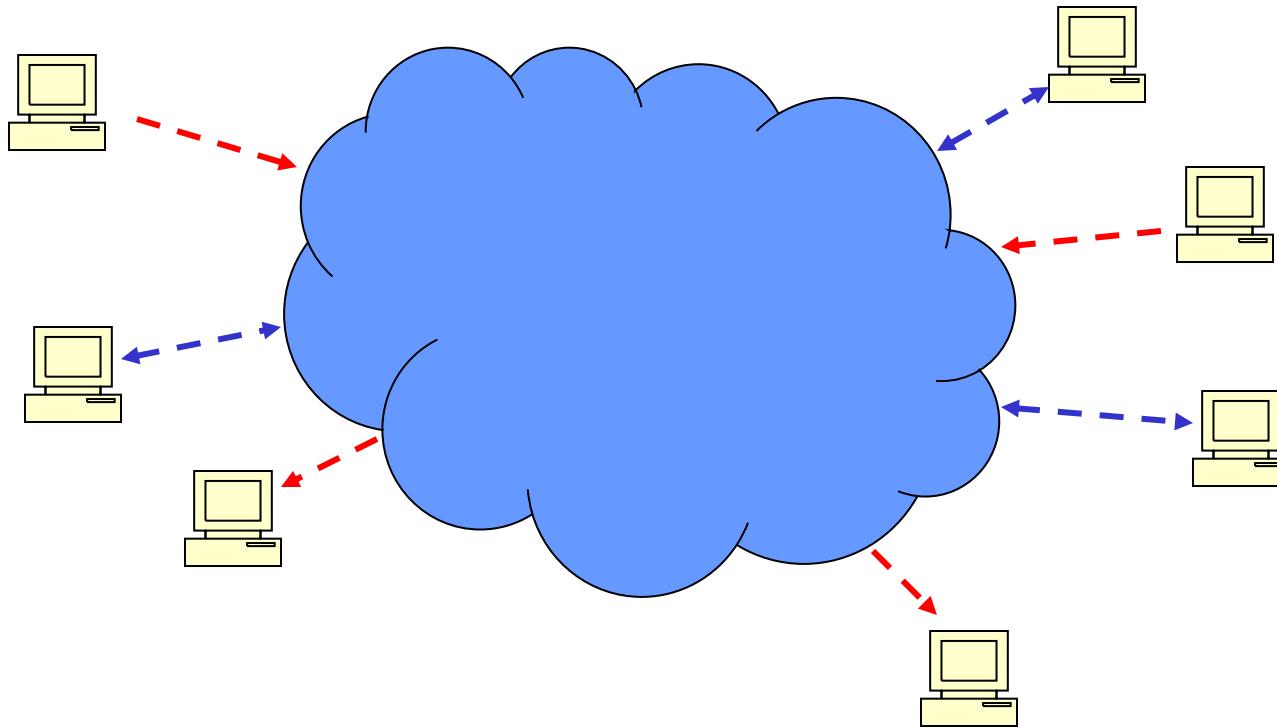
Um datagrama endereçado para o grupo é entregue  
a cada um dos membros do grupo Multicast.

# Multicast suportado pela rede



- Com multicast suportado pela rede
  - Datagramas são replicados na rede
- Uso ótimo de recursos de rede

# Multicast - Múltiplos Emissores



- Vários emissores para o mesmo grupo
  - Emissores simultâneos

# Grupos Multicast: questões

- Como um grupo é iniciado e como ele é encerrado?
- Como é escolhido o endereço do grupo?
- Como são adicionados novos hosts ao grupo?
- Qualquer um pode fazer parte (ativa) do grupo ou a participação é restrita?
- Como os roteadores interoperam para entregar um datagrama multicast a todos os membros do grupo?

# Grupo Multicast

## □ Criação de um Grupo

- Antes de um multicast ser iniciado no nível IP
  - emissor tem que reservar um endereço de grupo disponível
- Receptores têm que encontrar endereço IP Multicast
  - protocolos SAP (Session Announcement Protocol) e SDP (Session Description Protocol) foram definidos para estes propósitos

# Endereços IP Multicast

- São endereços classe D (prefixo binário: 1110)
  - 224.0.0.0 - 239.255.255.255
- Alguns endereços são reservados pela IANA para propósitos especiais

Endereços	Descrição
224.0.0.0 – 224.0.0.255	Bloco de controle da LAN (não para roteamento)
224.0.0.1	Todos os sistemas na sub-rede
224.0.0.2	Todos os roteadores na sub-rede
224.0.0.4	Roteadores DVMRP
224.0.0.9	Roteadores RIP
224.0.1.0 – 238.255.255.255	Escopo Global
239.0.0.0 – 239.255.255.255	Escopo Limitado administrativamente

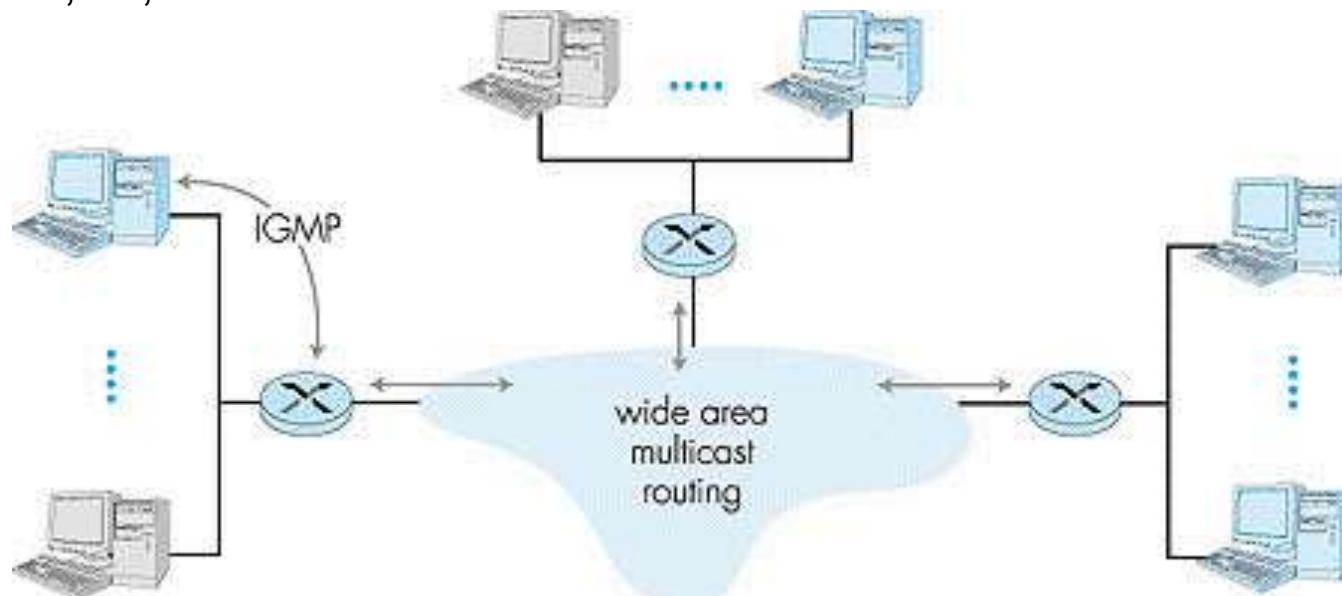
# IP Multicast

- Entrada em um grupo multicast
  - Host desejando se juntar ou deixar um grupo não propaga ele mesmo sua informação de membro
    - ele faz isto indiretamente, via roteador multicast que propaga o informação do membro para outros roteadores multicast
      - procedimento iniciado pelo receptor
  - Um host pode fazer parte de vários grupos
    - Exemplo: som é distribuído para um dado grupo - ou para vários grupos com diferentes graus de qualidade
      - e o vídeo é em um grupo diferente

# Modelo de serviço Multicast

Protocolo Host para roteador:

- IGMP v1, v2, v3



Multicast a nível de enlace/hardware

- Ethernet

Protocolos de roteamento multicast

- PIM, CBT, DVMRP, MOSPF, MBGP

# IGMP (Internet Group Management Protocol)

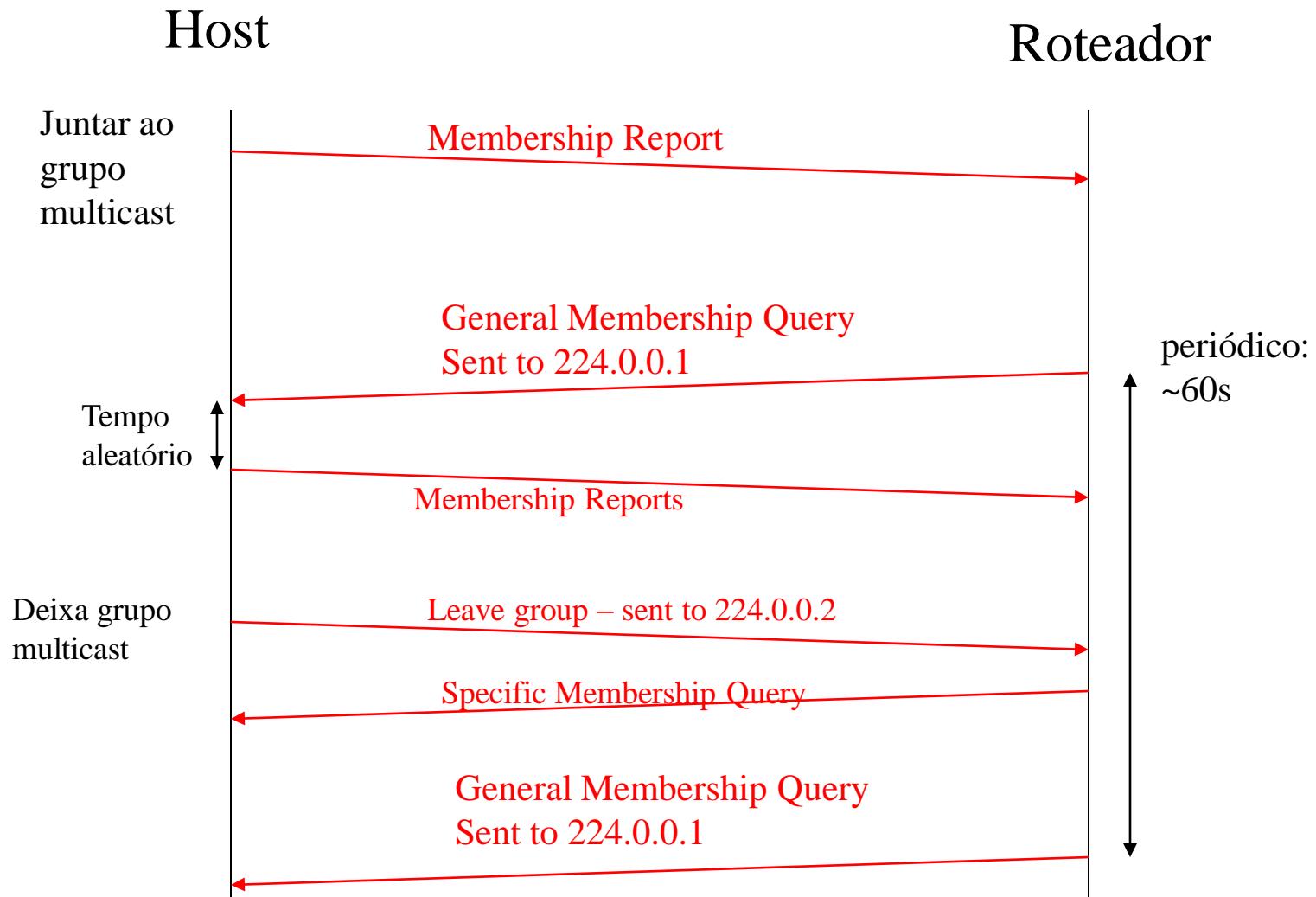
- Usado para controle do grupo entre os hosts e os roteadores multicast
  - Não é para roteamento multicast
- Parte da camada Rede
  - Encapsulado no datagrama IP (como o ICMP)
- Sempre endereçado para um endereço multicast
  - Com freqüência para todos os sistemas (224.0.0.1), todos os roteadores (224.0.0.2)
  - Ou para um grupo multicast específico

# Mensagens IGMP

## □ Tipos de mensagens

- Consulta de Membros em intervalos regulares
  - Enviado por roteadores para consultar (query) membros na rede
- Relatório de membros
  - Enviados pelos hosts para relatar grupos que participa
  - Enviado quando usuário entra em um grupo
- Deixar grupo
  - Enviados pelos hosts para deixar um grupo

# Mensagens IGMP

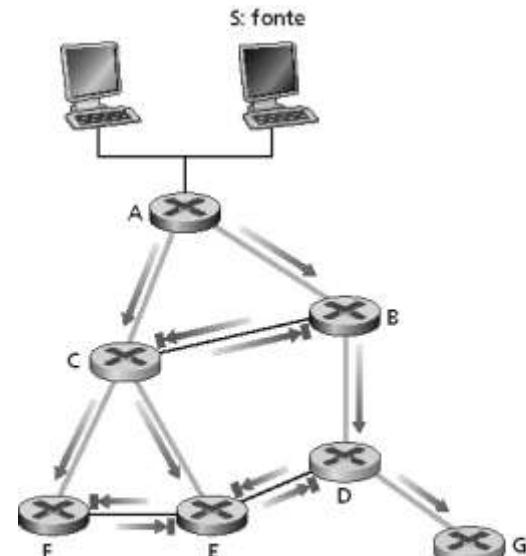
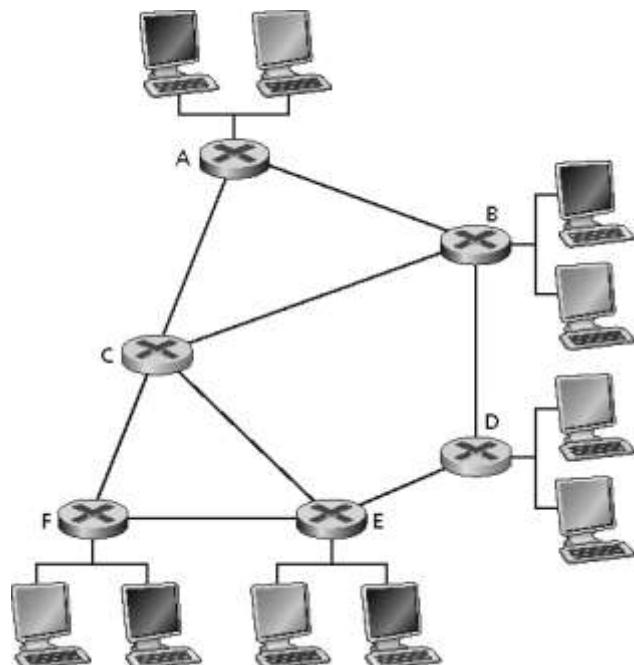


# Roteadores Multicast

- Um roteador habilitado com multicast é referenciado como um roteador multicast
  - Detecta todo o tráfego multicast e roteia se necessário
  - Serve hosts diretamente conectados
- Roteador multicast detecta todos os endereços multicast
  - Captura no modo promíscuo todo o tráfego multicast na LAN

# Roteamento Multicast

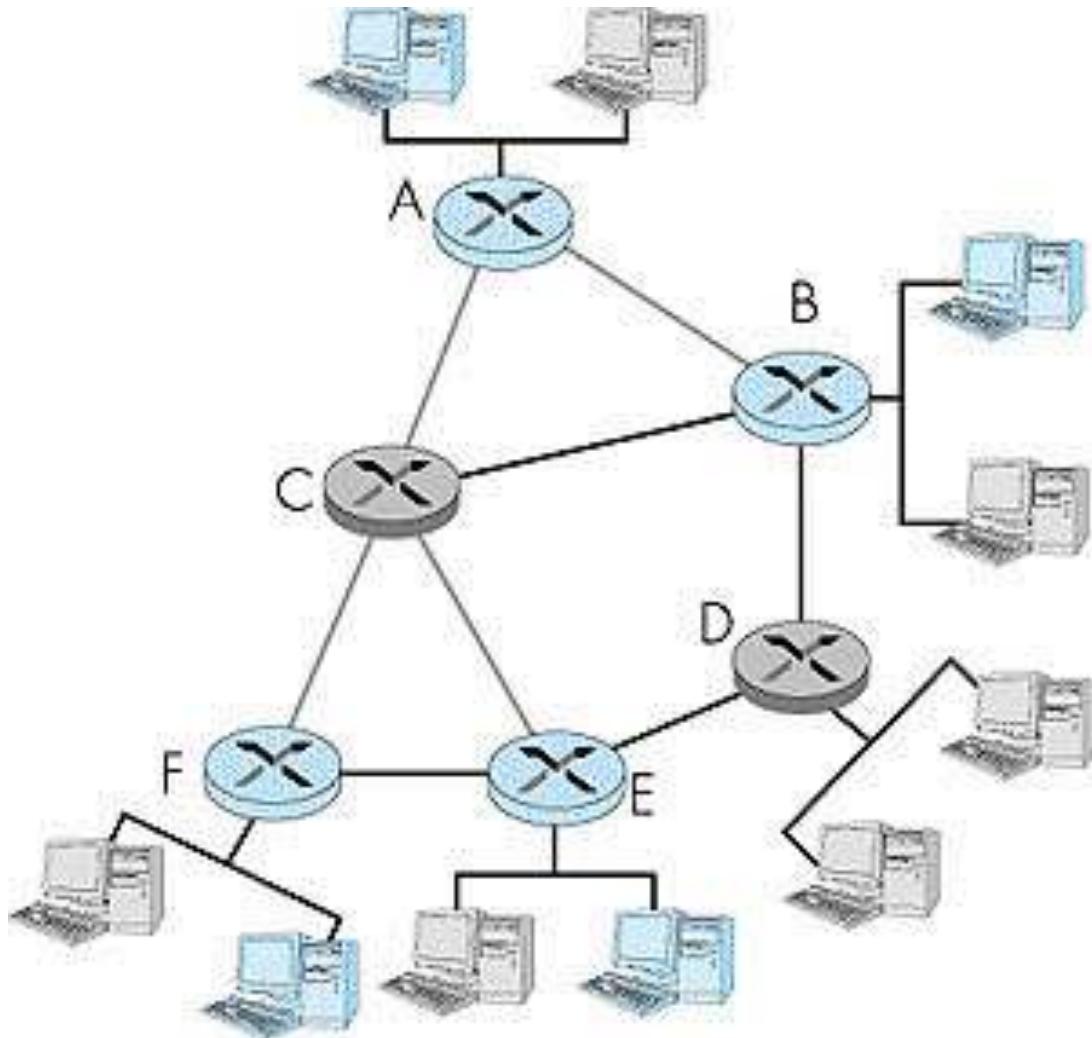
- Objetivo: encontrar uma árvore (ou árvores) conectando roteadores que possuem membros de grupo multicast local
  - Árvore: não são todos os caminhos entre os roteadores usados
  - Baseada na fonte: uma árvore diferente de cada transmissor para os receptores
  - Árvore compartilhada: a mesma árvore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



Legenda:

→ pacote (pkt) que será repassado  
→ pacote (pkt) que não será repassado além do roteador destinatário

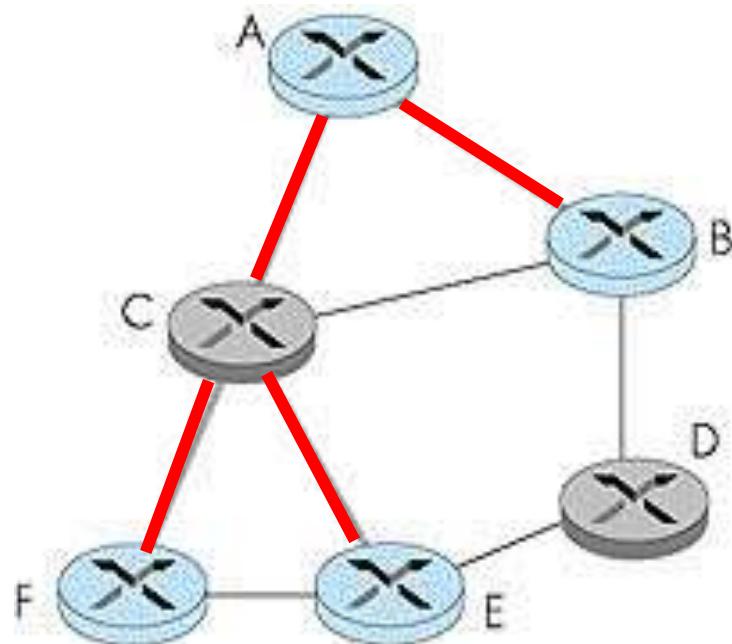
# Roteamento Multicast



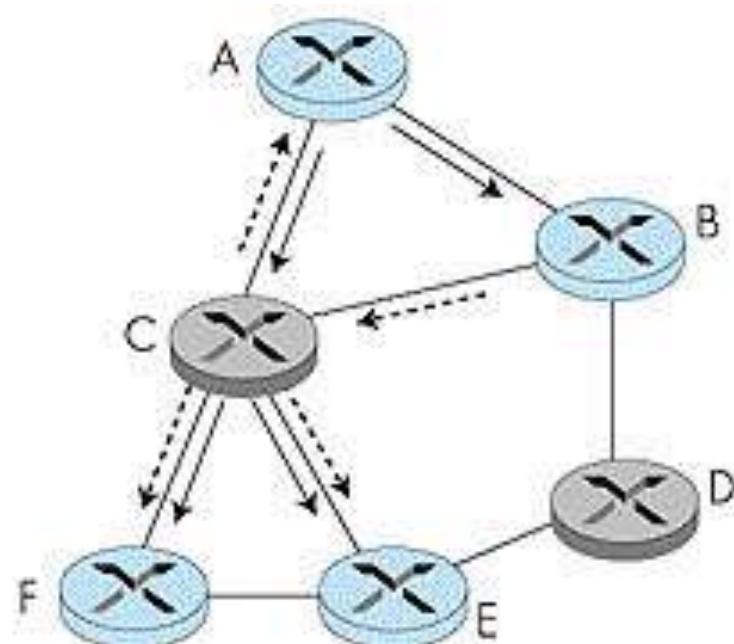
- Um único grupo multicast.
- Estão coloridos os hosts que pertencem ao grupo e os roteadores aos quais eles estão conectados.
- Apenas estes roteadores (A, B, E e F) necessitam receber este tráfego multicast.

# Árvores de Roteamento

## Multicast

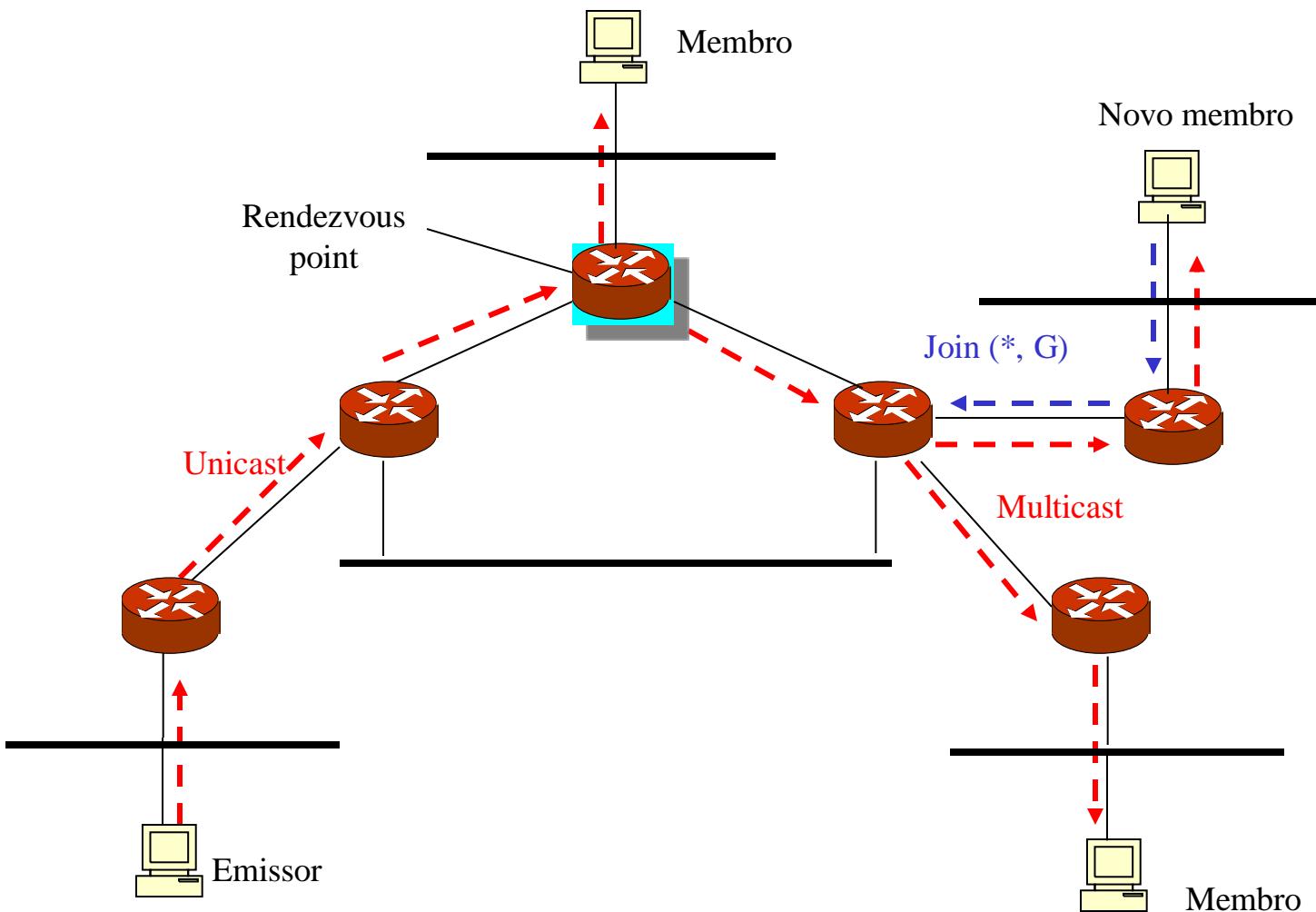


Árvore única  
compartilhada pelo grupo.

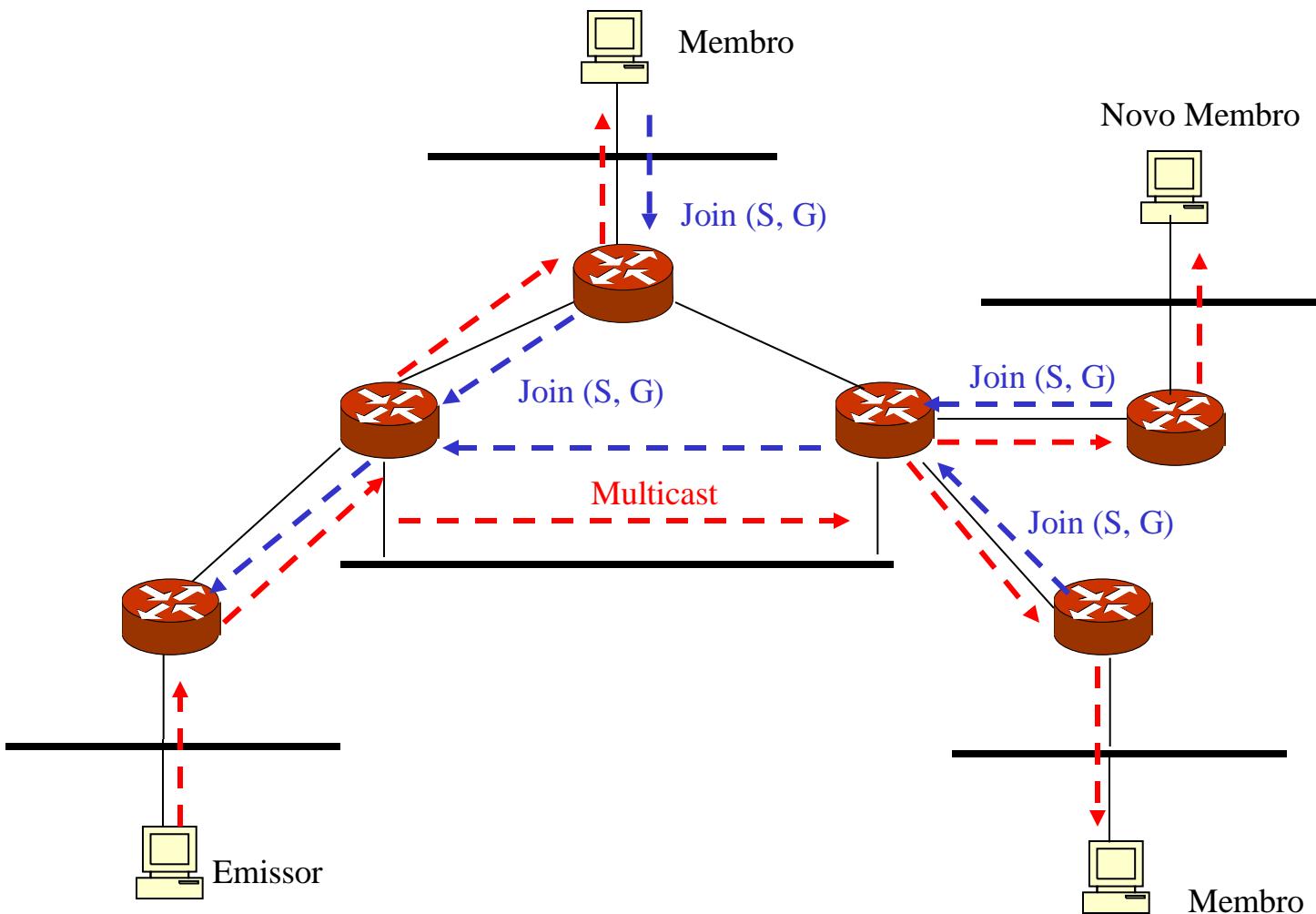


Árvores baseadas  
nas origens.

# Árvore Multicast Compartilhada

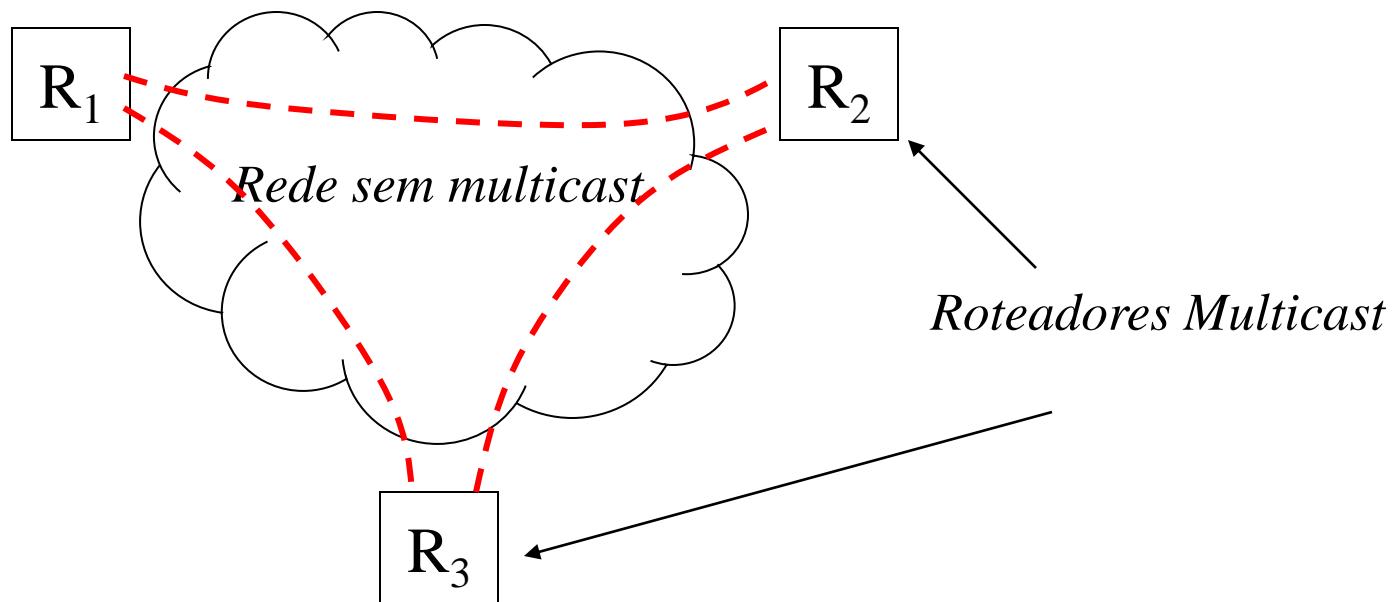


# Árvore Multicast Específico da Fonte



# Tunelamento

- Não é necessário que todos os roteadores sejam roteadores multicast
  - Como enviar o pacote multicast para todos os membros do grupo?
  - Resp.: via tunelamento nas sub-redes sem multicast
- Pacotes IP multicast são colocados em pacotes IP unicast para o roteador multicast.



# IPv6



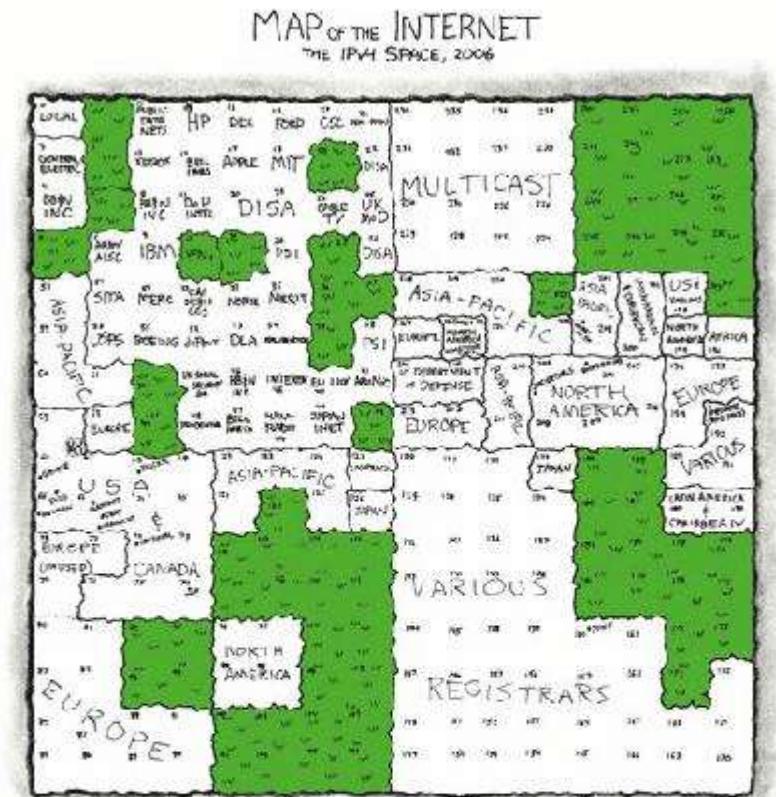
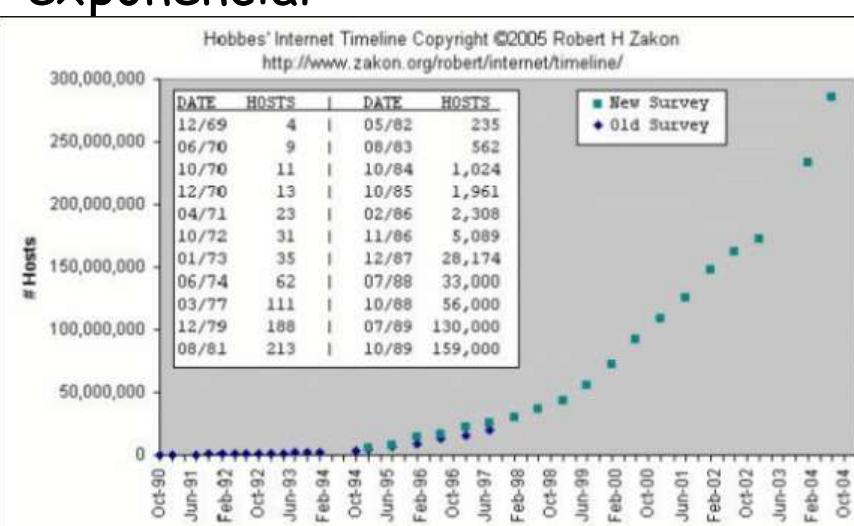
- Uma “nova versão” do protocolo Internet IP
  - IETF decidiu em 1992 desenvolver uma nova versão do IP pois o espaço de endereçamento disponível do IPv4 provavelmente terminaria no início do século 21
  - Baseado nos principais paradigmas IPv4
    - sem conexão, sem controle de erro e de fluxo na camada de rede
  - Projetado para ser um passo evolucionário do IPv4
    - Aumento do espaço de endereçamento, autenticação e criptografia
    - Extensões para fluxos de dados multimídia
    - Mais suporte à mobilidade
- Compatível com IPv4
  - Uma importante meta de projeto do IPv6 é a compatibilidade com IPv4
    - novos hosts e roteadores executando IPv6 são capazes de coexistir com hosts IPv4
      - habilitando assim uma migração gradativa da Internet

# Motivos que levam à substituição do IPv4 pra IPv6

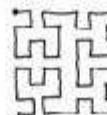
- O espaço de endereçamento do IPv4 é insuficiente (32 bits)
  - "Esgotamento no IANA, que é a entidade que controla mundialmente esse recurso ocorreu em fev/2011 e nos Registros Regionais, como o LACNIC, que controla os números IP para a América Latina e Caribe, em algum momento entre 2012 e 2014"  
[\(http://www.ipv6.br/IPV6/ArtigoEsgotamentoIPv4\)](http://www.ipv6.br/IPV6/ArtigoEsgotamentoIPv4)

# Motivos que levam à substituição do IPv4 para IPv6

- O espaço de endereçamento do IPv4 é insuficiente (32 bits)
    - $2^{32} = 4.294.967.296$   
(4 bilhões de endereços)
    - Por ser hierárquico, o limite prático é de 250 milhões de hosts
  - O crescimento da Internet é exponencial



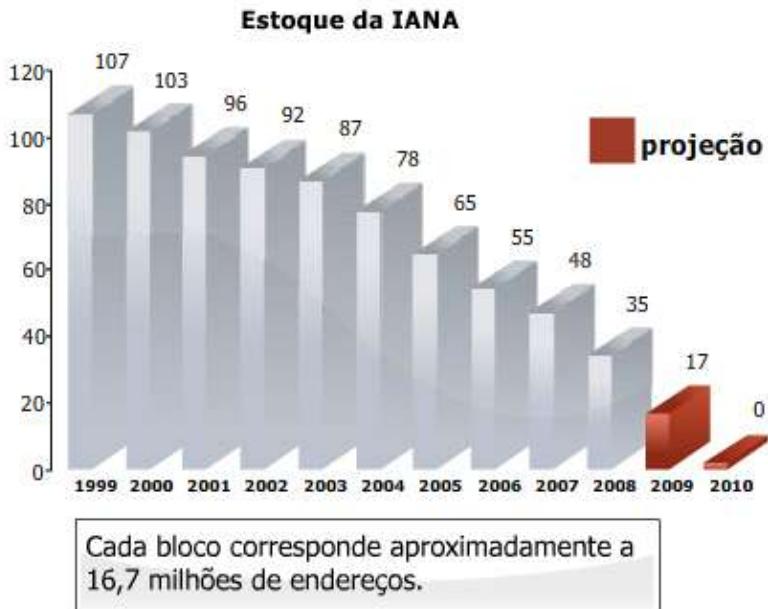
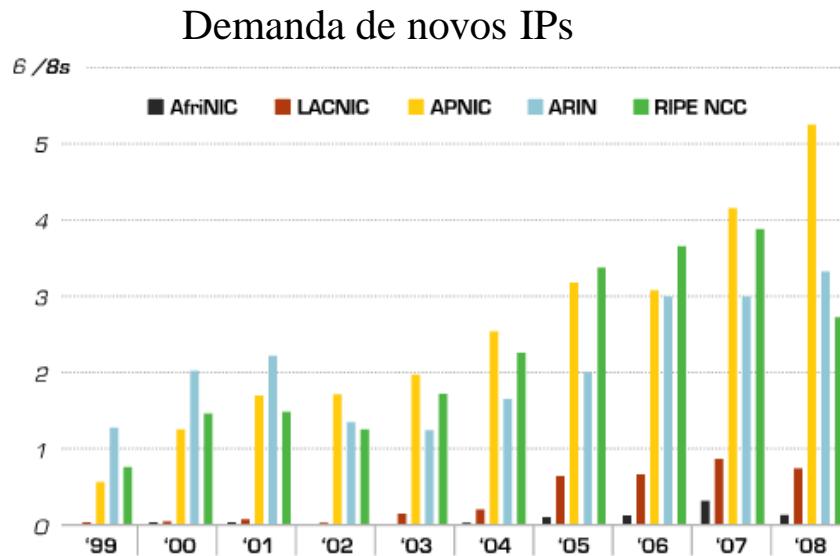
THIS CHART SHOWS THE IP ADDRESS SPACE ON A PLANE USING A FRACTAL MAPPING WHICH PRESERVES GROUPING--ANY CONSECUTIVE STRING OF IPs WILL TRANSLATE TO A SINGLE COMPACT, CONTIGUOUS REGION ON THE MAP. EACH OF THE 256 NUMBERED BLOCKS REPRESENTS ONE /8 SUBNET (CONTAINING ALL IPs THAT START WITH THAT NUMBER). THE UPPER LEFT SECTION SHOWS THE BLOCKS SOLD DIRECTLY TO CORPORATIONS AND GOVERNMENTS IN THE 1980'S BEFORE THE RIPE TOOK OVER ALLOCATION.



+ UNALLOCATED  
BLOCK

# Motivos que levam à substituição do IPv4 para IPv6

- O espaço de endereçamento do IPv4 é insuficiente (32 bits)
  - Demanda anual por novos números IPs nos Registros Regionais



# Adoção do IPv6 é em breve....

- <http://tecnologia.uol.com.br/ultimas-noticias/reuters/2011/02/03/ultimos-blocos-de-enderecos-da-internet-ipv4-sao-entregues-pelo-iana.jhtm>

03/02/2011 - 18h00 / Atualizada 03/02/2011 - 18h04

## **Últimos blocos de endereços da Internet IPv4 são entregues pelo IANA**



REUTERS



0



0

Recomendar 2 recomendações. Cadastre-se para ver o que seus amigos recomendam.

SÃO PAULO (Reuters) - A IANA, entidade que controla a distribuição de protocolos de Internet (IPs), entregou nesta quinta-feira os últimos blocos disponíveis de endereços IPv4. Com o fim do IPv4, que permite 4 bilhões de endereços, um novo protocolo será adotado, o IPv6, que pode formar até 340 decílhões de combinações.

Cada dispositivo conectado à Internet possui seu próprio endereço de IP, número para que as informações encontrem seu caminho na rede. O crescimento de equipamentos com acesso à Web no mercado exige maior disponibilidade de combinações.

Em 31 de janeiro, dois blocos do IPv4 foram entregues para a região Ásia/Pacífico, com distribuição igualitária dos blocos restantes para todas as regiões, incluindo a da América Latina e Caribe, da qual o Brasil faz parte.

Estima-se que o Brasil ainda distribua endereços ".br" do seu estoque no padrão IPv4 ao longo de mais um ano, por meio do Núcleo de Informação e Coordenação do Ponto BR (NIC.br). O IPv4 deve conviver com o IPv6 de 15 a 20 anos, até que um padrão substitua o outro, afirmou o NIC.br.

# Motivos que levam à substituição do IPv4 para IPv6

## Medidas Paliativas: CIDR

- A política de alocação inicial não foi favorável a uma utilização racional dos mesmos. Dividiu-se esse espaço em 3 classes:
  - **Classe A:** com 128 segmentos atribuídos individualmente às entidades que deles necessitassem, com 16 milhões de endereços cada. Ela utilizava o espaço compreendido entre os endereços 00000000.\*.\*.\* (0.\*.\*.\*.) e 0111111.\*.\*.\* (127.\*.\*.\*.).
  - **Classe B:** com 16 mil segmentos de 64 mil endereços cada. Essa classe era classificada como /16. Ela utilizava o espaço compreendido entre os endereços 10000000.0000000.\*.\* (128.0.\*.\*.) e 10111111.11111111.\*.\* (191.255.\*.\*.).
  - **Classe C:** 2 milhões de segmentos de 256 endereços cada. Essa classe era classificada como /24. Ela utilizava o espaço compreendido entre os endereços 11000000.0000000.00000000.\* (192.0.0.\*.) e 11011111.11111111.11111111.\* (213.255.255.\*).
- Além disso, 32 blocos /8 restantes foram reservados para Multicast e para a IANA.

# Motivos que levam à substituição do IPv4 para IPv6

## Medidas Paliativas: CIDR

### CIDR (de *Classless Inter-Domain Routing*) RFC 1519, 1993.

- Permitindo flexibilidade quando dividindo margens de endereços IP em redes separadas
- Usa máscaras de comprimento variável - VLSM (de Variable Length Subnet Masks)
  - para alocar endereços IP em subredes de acordo com as necessidades individuais e não nas regras de uso generalizado em toda a rede
  - a divisão de rede/host pode ocorrer em qualquer fronteira de bits no endereço
- Promoveu assim um uso mais eficiente para os endereços IP cada vez mais escassos.

# Motivos que levam à substituição do IPv4 para IPv6

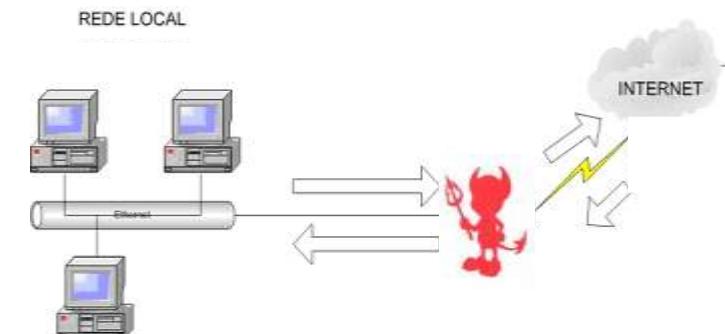
## Medidas Paliativas:

- RFC 1918 (endereços privados)
  - Permite o uso de endereços não válidos na Internet nas redes corporativas
- NAT (tradução de endereços)
  - Permite que com um endereço válido na Internet apenas, toda uma rede de computadores usando endereços privados seja conectada (mas com restrições)
- DHCP (alocação dinâmica de endereços IP)
  - Permite que provedores reutilizem endereços Internet para conexões não permanentes

# Motivos que levam à substituição do IPv4 para IPv6

## □ Medidas Paliativas:

- ... mas também colaborando para a demora em sua adoção!
- Alguns questionam porque não utilizar o NAT indefinidamente, mas ele foi concebido como uma solução provisória!
  - O NAT acaba com o modelo de funcionamento fim a fim, trazendo complicações ou impedindo o funcionamento de uma série de aplicações.
  - O NAT tem alguns problemas técnicos:
    - Não é fácil manter o estado do NAT no caso de falha em um dos hosts.
    - O NAT não funciona bem com o IPsec.
    - O NAT não escala bem



# A razão principal para o IPv6 é, então...

- A necessidade de mais endereços Internet!
  - Para suportar seu crescimento:
    - Possibilitando a interligação de mais redes, de forma que a expansão da economia, com novas empresas, novos negócios seja suportada.
    - A fim de que todos possam ser incluídos digitalmente, em especial nos países em desenvolvimento
    - Com o uso de novas aplicações, como sua utilização em dispositivos móveis com tecnologia 3G, por exemplo, ou em eletrônicos-domésticos e outros aparelhos com eletrônica embarcada
    - Com a eliminação de tecnologias como o NAT, que dificultam o funcionamento de várias aplicações

# A razão principal para o IPv6 é, então...

- Um endereço IPv4 é formado por 32 bits.
  - $2^{32} = 4.294.967.296$  endereços  
aproximadamente 4 bilhões de endereços
- Um endereço IPv6 é formado por 128 bits.
  - $2^{128} =$   
340.282.366.920.938.463.463.374.607.431.768.211.456  
endereços
  - ~ 79 trilhões de trilhões de vezes mais que no IPv4.
  - $\sim 5.6 \times 10^{28}$  endereços IP por ser humano.
  - Distribuídos na terra:  
 $665,570,793,348,866,943,898,599/m^2$
  - Estimativa pessimista com hierarquias: ~1,564  
endereço/ $m^2$

# Outras razões

Além dessa quantidade quase ilimitada de IPs, o novo formato dos endereços permitirá:

- Definir uma arquitetura hierárquica na Internet, possibilitando um encaminhamento mais eficiente dos pacotes de dados;
- Facilitar a distribuição de IPs fixos e válidos para conexões DSL, *Cable Modems* e telefones móveis;
- Fornecer endereços válidos na Internet para todos os dispositivos conectados a ela;
- Utilizar a arquitetura fim-a-fim;
- Eliminar os problemas associados ao NAT.

# Outras razões

Versão [Version]	Classe de Tráfego [Traffic Class]	Identificador de Fluxo [Flow Label]		
Tamanho dos dados [Payload Length]	Próximo Cabeçalho [Next Header]	Limite de Encaminhamento [Hop Limit]	Endereço de Origem [Source Address]	
Endereço de Destino [Destination Address]				

Outra mudança importante em relação à versão anterior do protocolo IP é o formato do cabeçalho IPv6.

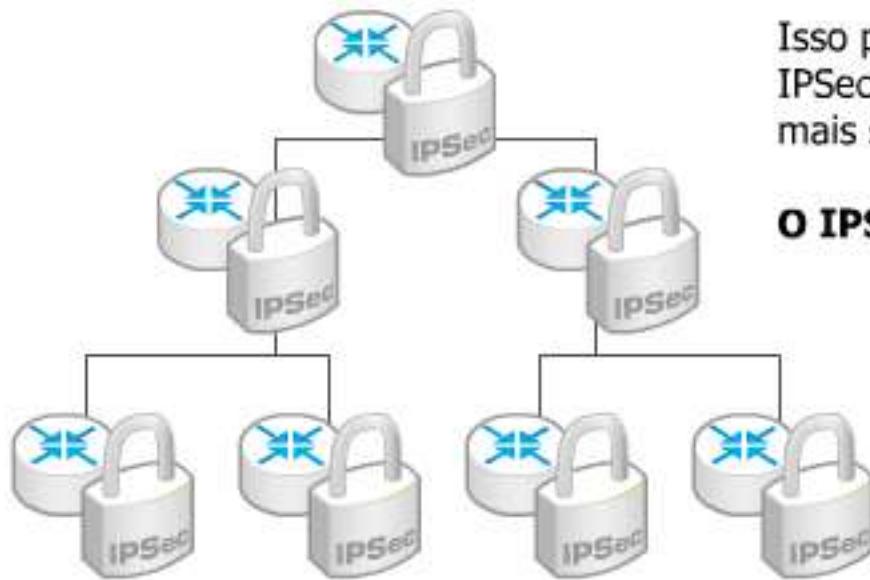
A nova versão foi simplificada, tornando-se mais eficiente, reduzindo o processamento dos roteadores.

# IPv6

# Outras razões

**Questões relacionadas a segurança também foram revistas.**

O suporte ao protocolo IPSec passa a ser obrigatório, fazendo parte do próprio protocolo IPv6.

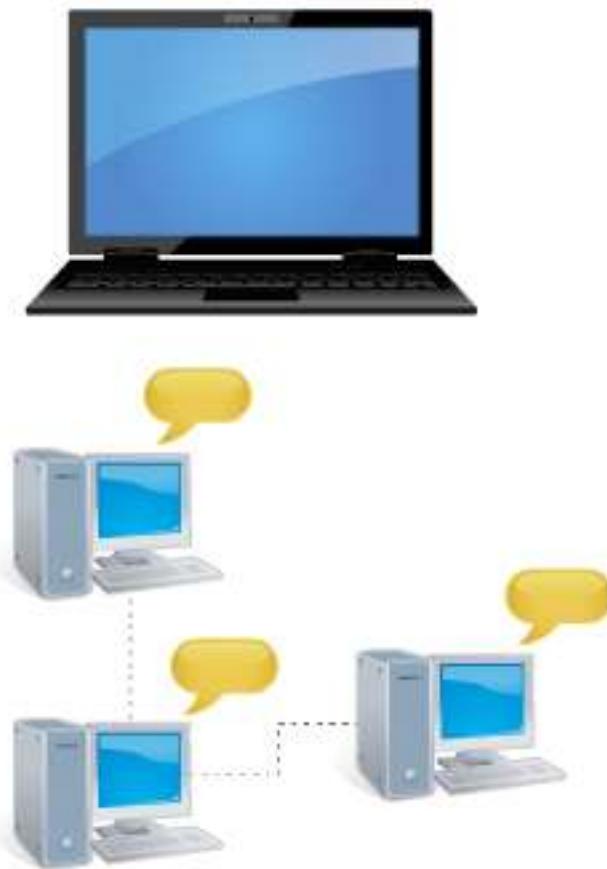


Isso permite aos administradores de rede ativar o IPSec em todos os dispositivos da rede tornando-a mais segura.

**O IPSec é capaz de garantir:**

- autenticidade;
- privacidade;
- integridade dos dados na comunicação.

# Outras razões



O protocolo ICMP (*Internet Control Message Protocol*) também foi modificado tornando-se mais eficaz. Isso permitiu a inclusão de novas funcionalidades ao IPv6 e o aprimoramento de outras, como:

- Mecanismos de autoconfiguração de endereços;
- Descoberta de Vizinhança (*Neighbour Discovery*);
- Gerenciamento de grupos *multicast*.

# Outras razões

- Protocolo de Descoberta de Vizinhanças
  - Utilizado por hosts e roteadores para os seguintes fins:
    - Divulgar o endereço MAC dos nós da rede;
    - Encontrar roteadores vizinhos;
    - Autoconfiguração de endereços.
      - Determinar prefixos e outras informações de configuração de rede;
    - Detectar endereços duplicados;

# Outras razões

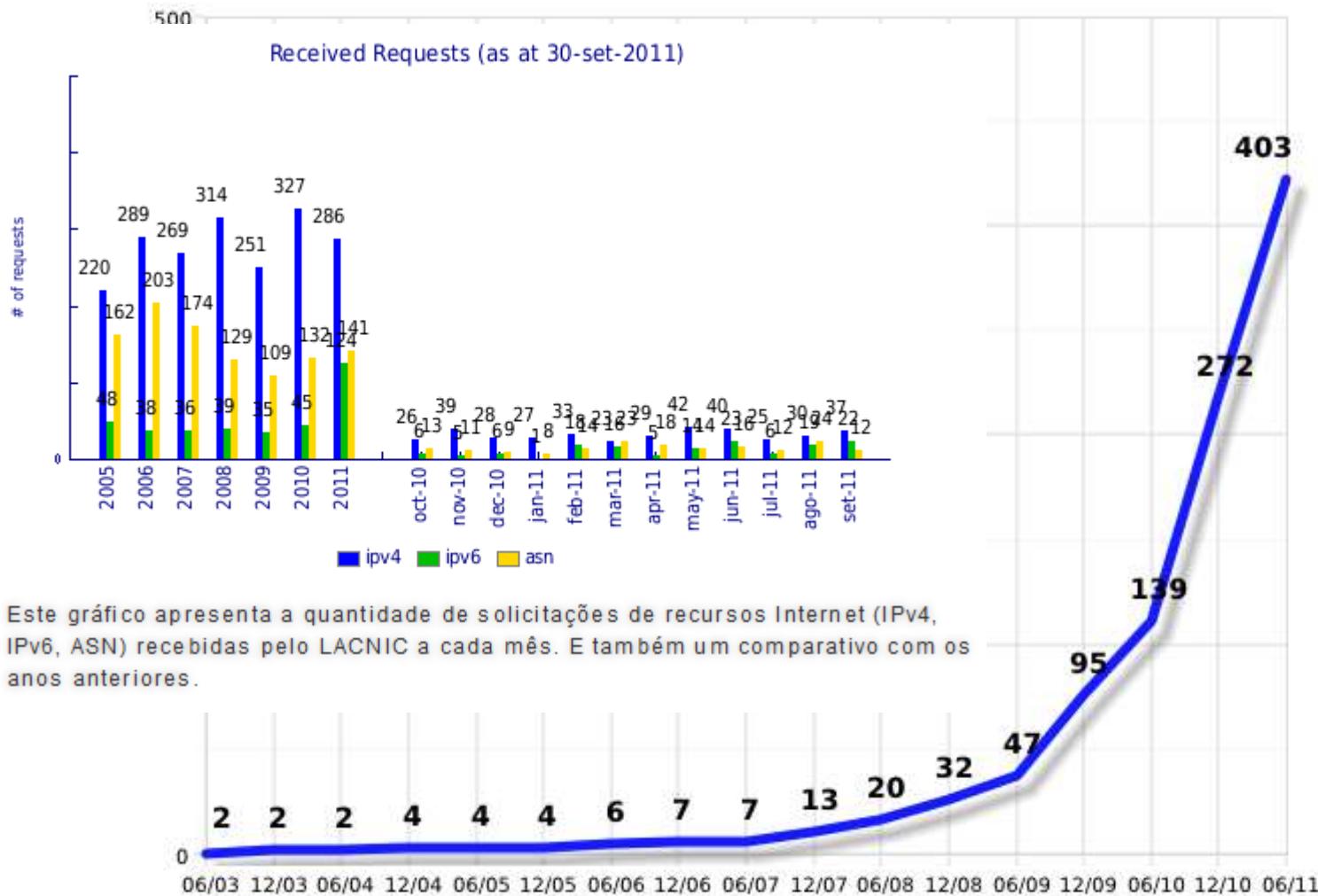
Outras vantagens apresentadas pelo IPv6 que podem ser destacadas são:



- O suporte a conexões móveis foi aprimorado e agora passa a fazer parte integrada do protocolo IPv6. Esta funcionalidade permite a um usuário se deslocar de uma rede para outra sem necessidade de alterar seu endereço;
- Com o IPv4, cada roteador pode fragmentar os pacotes de dados durante seu trajeto, sendo que esse processo pode ser realizado diversas vezes dependendo do desenho da rede. No IPv6, a fragmentação é realizada apenas na origem, com o intuito de agilizar o roteamento dos pacotes.

# A razão principal para o IPv6 é, então...

Mais de 400 Sistemas Autônomos brasileiros já têm alocações IPv6!



Este gráfico apresenta a quantidade de solicitações de recursos Internet (IPv4, IPv6, ASN) recebidas pelo LACNIC a cada mês. E também um comparativo com os anos anteriores.

# Cabeçalho IPv6

## □ Cabeçalho IPv4

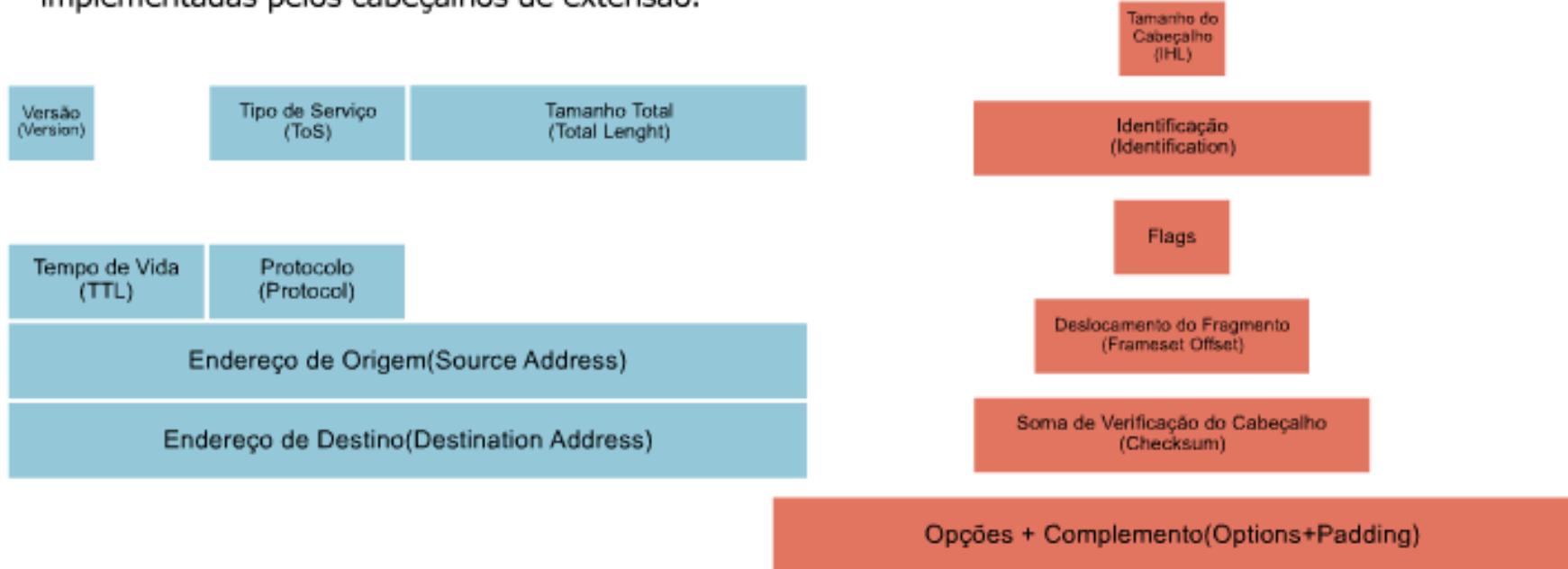
O cabeçalho IPV4 é composto por 12 campos fixos, podendo conter ou não opções, fazendo com que seu tamanho possa variar entre 20 e 60 bytes:

Versão (Version)	Tamanho do Cabeçalho (IHL)	Tipo de Serviço (ToS)	Tamanho Total (Total Length)	
Identificação (Identification)		Flags	Deslocamento do Fragmento (Fragment Offset)	
Tempo de Vida (TTL)	Protocolo (Protocol)	Soma de Verificação do Cabeçalho (Checksum)		
Endereço de Origem(Source Address)				
Endereço de Destino(Destination Address)				
Opções + Complemento(Options+Padding)				

# Cabeçalho IPv6

## □ Cabeçalho IPv4 => Cabeçalho IPv6

Seis campos do cabeçalho IPv4 foram removidos, pois suas funções não são mais necessárias ou são implementadas pelos cabeçalhos de extensão.



# Cabeçalho IPv6

- Formato dos cabeçalhos fixos
  - Usa cabeçalho de extensão em vez de options
- Remove o header checksum
  - Confiabilidade da camada de enlace e camadas mais altas para verificar a integridade dos dados
- Remove segmentação hop-a-hop
  - Sem fragmentação devido a descoberta do MTU do caminho

# Cabeçalho IPv6

- Mais simples
  - 8 campos (40 bytes) => IPv4 tem normalmente 20 bytes
- Mais flexível
  - Prevê sua extensão, através do uso cabeçalhos adicionais

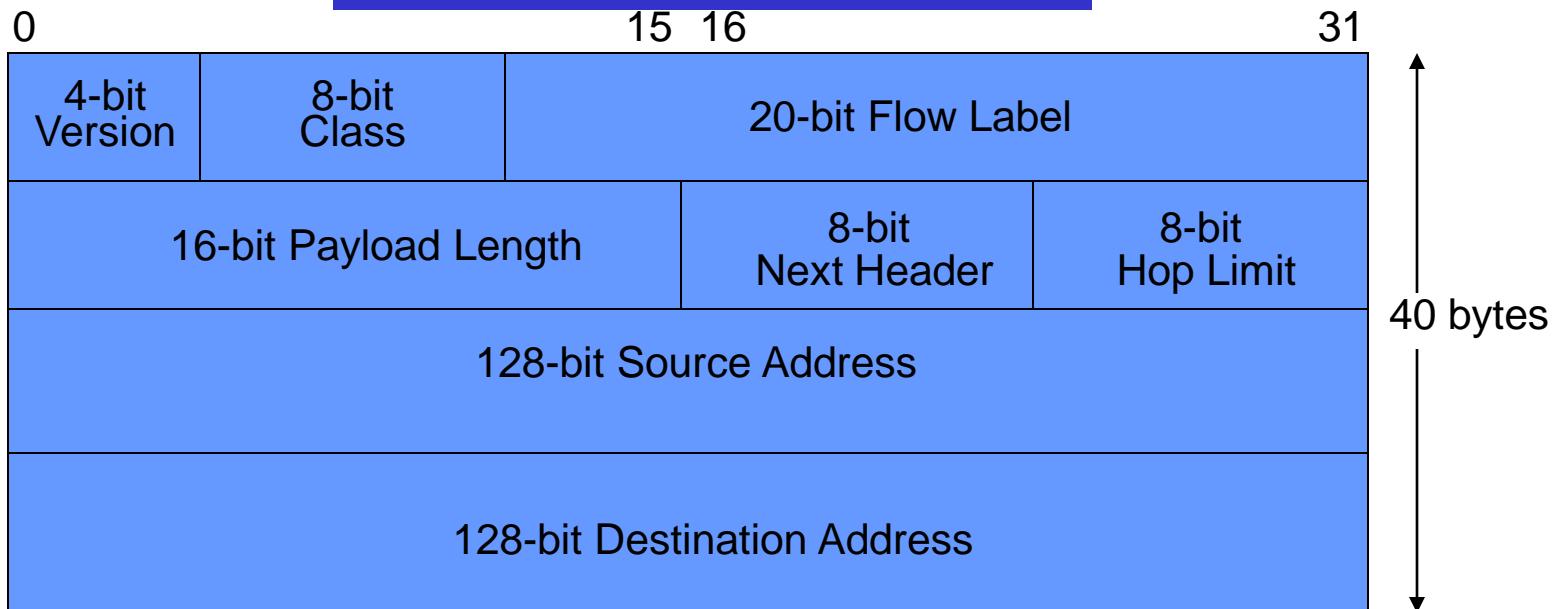
IPv4 Header			
Version	IHL	Type of Service	Total Length
Identification		Flags	Fragment Offset
Time to Live	Protocol	Header Checksum	
Source Address			
Destination Address			
Options			
Padding			

IPv6 Header	
Version	Traffic Class
Flow Label	
Payload Length	
Next Header	Hop Limit
Source Address	
Destination Address	

## Legend

- Field's Name Kept from IPv4 to IPv6
- Fields Not Kept in IPv6
- Name and Position Changed in IPv6
- New Field in IPv6

# Cabeçalho IPv6



- Version** Único campo idêntico ao IPv4. Código é 6 em IPv6
- Class** Facilita manipulação do tráfego tempo real
- Flow Label** Distingue pacotes requerendo o mesmo tratamento
- Payload Length** Substitui campo *length* do IPv4. Dá o tamanho do dado seguindo o cabeçalho IPv6
- Next Header** Substitui campo *protocol* do IPv4. Cabeçalhos de extensão pode ser usado.
- Hop Limit** Substitui campo *TTL* do IPv4. Limite de hop reflete melhor o uso.
- Src Address** 128 bits no IPv6 vs 32 bits no IPv4.
- Dst Address** 128 bits no IPv6 vs 32 bits no IPv4.

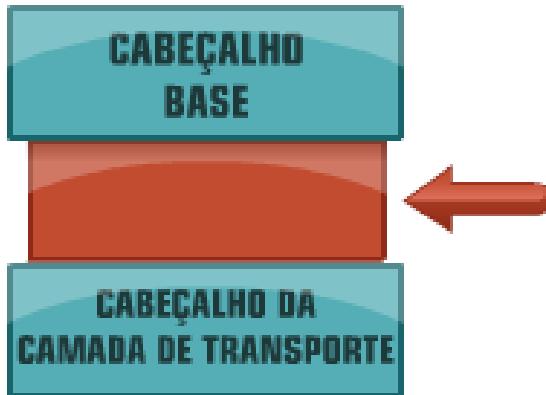
# Formato do cabeçalho IPv6

## □ Campo *flow label*

- Permite a identificação de todos os pacotes de um mesmo fluxo de dados
  - fluxo é uma seqüência de pacotes enviados por um host para um endereço unicast ou multicast
  - todos os roteadores no caminho podem identificar os pacotes de um fluxo e tratar eles de um modo específico ao fluxo
  - Por exemplo, eles podem escalar pacotes de um fluxo de áudio com uma mais alta prioridade que aqueles pertencente a um fluxo de transferência de arquivo

# Cabeçalhos de extensão

Outra importante mudança na estrutura do cabeçalho IPv6 é que, diferentemente do IPv4, que inclui em seu cabeçalho todas as opções adicionais, agora essas informações são tratadas por meio de **cabeçalhos de extensão**.



Os cabeçalhos de extensão localizam-se entre o cabeçalho base e o cabeçalho da camada de transporte.

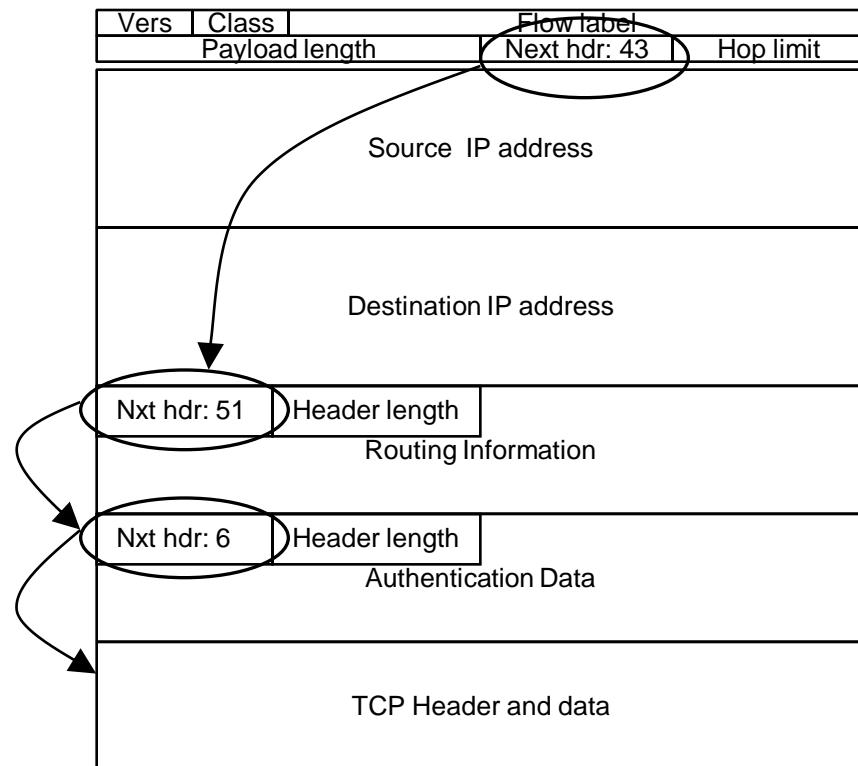
Atenção



Não há uma quantidade fixa de cabeçalhos de extensão que podem ser anexados ao cabeçalho base, porém, caso existam múltiplos cabeçalhos de extensão no mesmo pacote, eles serão adicionados em série formando uma "cadeia de cabeçalhos".

# Cabeçalhos de extensão

- Dá mais funcionalidade ao IP
- Vários cabeçalhos de extensão do IPv6 são opções no IPv4 (campo option do cabeçalho IPv4)
- Cabeçalhos de extensão são colocados entre o cabeçalho base IPv6 e o cabeçalho do nível de transporte (TCP/UDP)



# Cabeçalhos de extensão

- Cada cabeçalho tem um tamanho múltiplo de 8 bytes
- Os seguintes cabeçalhos já foram definidos

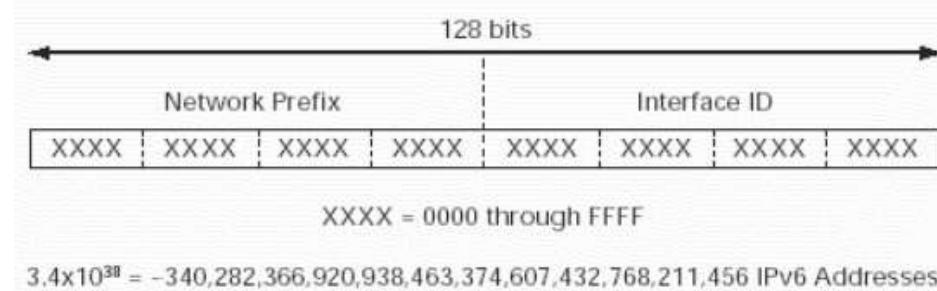
Valor	Nome do cabeçalho	Definição
0	Hop-By-Hop	Transporta informações opcionais que são processadas em cada nó ao longo do caminho do pacote, incluindo a origem e o destino.
60	Destination Options	Transporta informações opcionais que são processadas apenas pelo destino final do pacote.
43	Routing	Utilizado no suporte a mobilidade do IPv6, ele armazena o endereço original de um nó móvel (Type 2).
44	Fragmentation	Utilizado pela origem para enviar pacotes maiores do que a Maximum Transmit Unit (MTU) de um caminho. Ao contrário do IPv4, a fragmentação no IPv6 não ocorre nos roteadores encontrados ao longo do caminho do pacote, apenas na origem, sendo re-agrupados no destino final.
51	Authentication	Utilizado pelo serviço IPSec (IP Security) para prover autenticação e garantia de integridade aos pacotes IPv6. Esse cabeçalho é idêntico ao utilizado no IPv4.
50	Encapsulating Security Payload	Também utilizado pelo IPSec, provê integridade e confidencialidade para os pacotes.

# Endereços IPv6

- Notação decimal em coluna com oito inteiros hexadecimais de 16 bits
  - 68E8:1480:0022:0000:ABC1:0000:0000:01FE
- Zeros podem ser resumidos
  - 68E8:1480:22:0:ABC1:0:0:1FE
- Compressão de zeros: zeros podem ser substituído por "::"
  - 68E8:1480:22:0:ABC1:0:0:1FE substituído por  
68E8:1480:22:0:ABC1::1FE
  - Só pode ser realizada uma única vez, caso contrário poderia haver ambigüidade:
    - 2001:0000:0000:0058:0000:0000:0000:0230 poderia ser anotado:
      - 2001::58:0:0:320 ou 2001:0:0:58::320
      - Mas nunca 2001::58::320

# Endereços IPv6

- Não há classes de endereço
  - A rede é indicada por um prefixo
  - Prefixo é um número decimal que indica quantos bits de mais alta ordem representam o prefixo do endereço



- Exemplos de nós com 64 bits superiores indicando a rede:
  - FE80:2B00:23:201::1/64
  - FE80:2B00:23:201::2/64
- Ambos pertencem a rede:
  - FE80:2B00:23:201::/64

# Endereços IPv6

- Para ambientes mistos com IPv4 e IPv6
  - X:X:X:X:X:X. d.d.d.d
  - "X" são os grupos de 4 números hexadecimais
  - "d" são valores decimais de 8 bits que variam de 0 a 255, como na notação do IPv4.
  - Exemplos
    - 0:0:0:0:0:152.84.253.35
      - abreviando => ::152.84.253.35
    - 0:0:0:0:FFFF:152.84.253.35
      - abreviando => ::FFFF:152.84.253.35

# Endereços IPv6

## □ Endereço de Rede

- Similar ao IPv6: parte de host é 0
- Exemplo: endereço de sub-rede

2001:0000:0004:CFE, possui 60 bits de prefixo  
pode ser representado das seguintes formas:

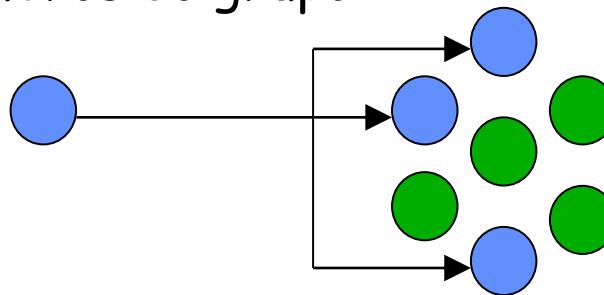
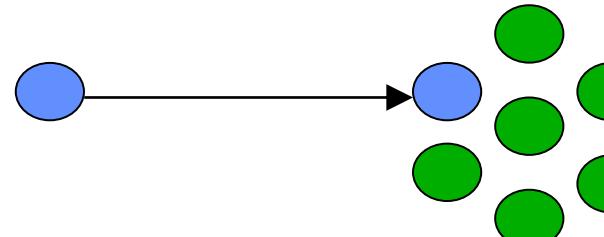
- 2001:0000:0004:CFE0:0000:0000:0000:0000/60
- Retirando os zeros a esquerda:  
2001:0:4:CFE0:0:0:0:0/60
- Abreviando: 2001:0:4:CFE0::/60

# Endereços IPv6

## □ Alguns prefixes alocados

Alocação	Prefixo (Binário)
Reservado	0000 0000
Reservado para Alocação NSAP (Network Service Access Point address)	0000 001
Reservado para Alocação IPX (Internetwork Packet Exchange)	0000 010
Aggregatable Global Unicast Address	001
Site-local Unicast Address	1111 1110 10
Link -local Unicast Address	1111 1110 11
Multicast Address	1111 1111

# Endereços IPv6

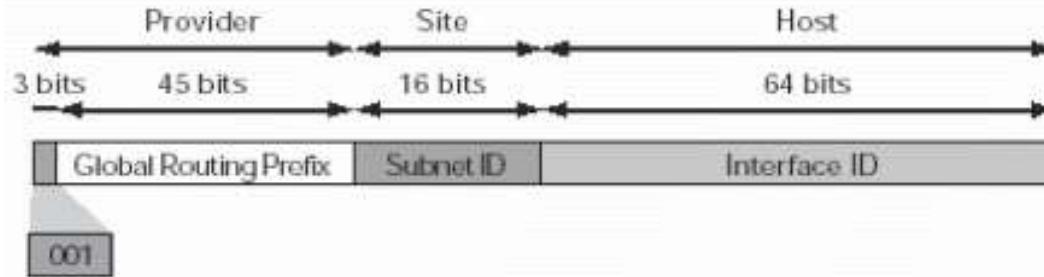
- IPv6 tem três categorias de endereçamento:
  - **unicast** - identifica uma interface
  - **multicast** - identifica um grupo; pacote é transmitido para todos os membros do grupo
  - **anycast** - identifica um grupo; pacote normalmente é transmitido ao membro mais próximo do grupo, respeitando os critérios de roteamento

# Endereços IPv6 Unicast

- Identifica apenas uma única interface
- Foram definidos pela RFC 2374 vários tipos de endereços Unicast :
  - Aggregatable Global Unicast Address
  - Loopback Address (::1)
  - Unspecified Address (::)
  - NSAP Address: suporte para endereçamento OSI NSAP (prefixo 0000001)
  - IPX Address (prefixo 0000010)
  - Site-local Unicast Address
  - Link-local Unicast Address
  - IPv4-compatible IPv6 Address (::172.16.1.2)

# Endereços IPv6 Unicast

- Aggregatable Global Unicast Address (visão simplificada)
  - Equivalente ao endereço global unicast usado em IPv4
    - Estrutura de endereços globais permite uma agregação de prefixos de roteamento que limitam o número de entradas nas tabelas de rotas



- FP - Format Prefix, indica que se trata de um endereço do tipo Global Unicast.
- Global Routing Prefix, destinado a identificação dos ISP's - Internet Service Provider
- Subnet ID, o campo Site ID da estrutura de hierarquização do endereço IPv6
- Interface ID, identifica a interface do host destino.

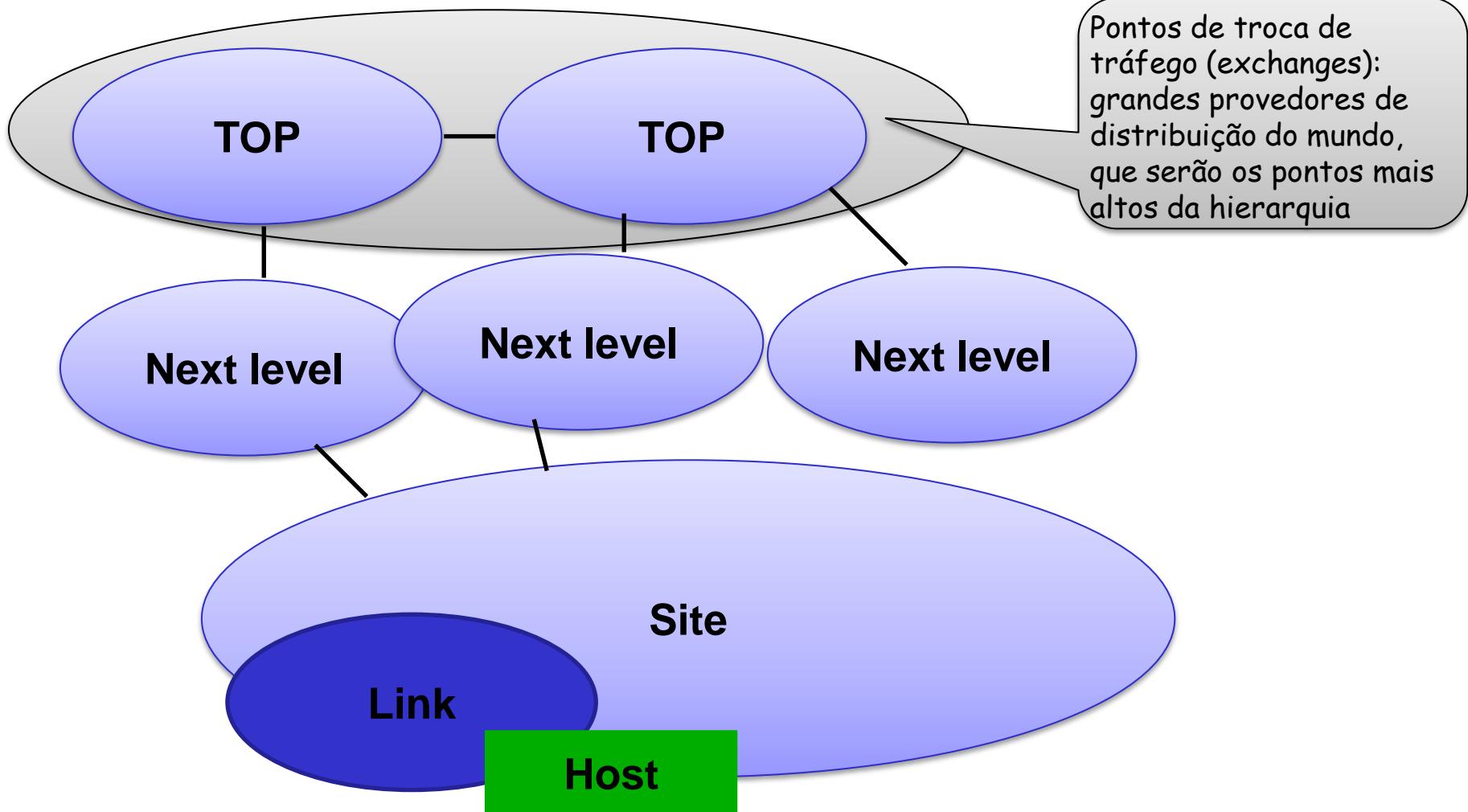
# Endereços IPv6 Unicast

- Endereços unicast globais agregáveis  
(Aggregatable Global Unicast Addresses)
  - Endereços unicast agregáveis são concebidos para serem agregados (ou resumidos para) criar uma infra-estrutura de encaminhamento eficiente
    - endereços possuem uma hierarquia definida por seus provedores de acesso à rede
  - Estrutura hierárquica existem 4 níveis:
    - TLA ID - Identificador Top-Level Aggregation;
    - NLA ID - Identificador Next-Level Aggregation;
    - SLA ID - Identificador Site-Level Aggregation;
    - Interface ID - Identificador de Interface



# Endereços IP Unicast

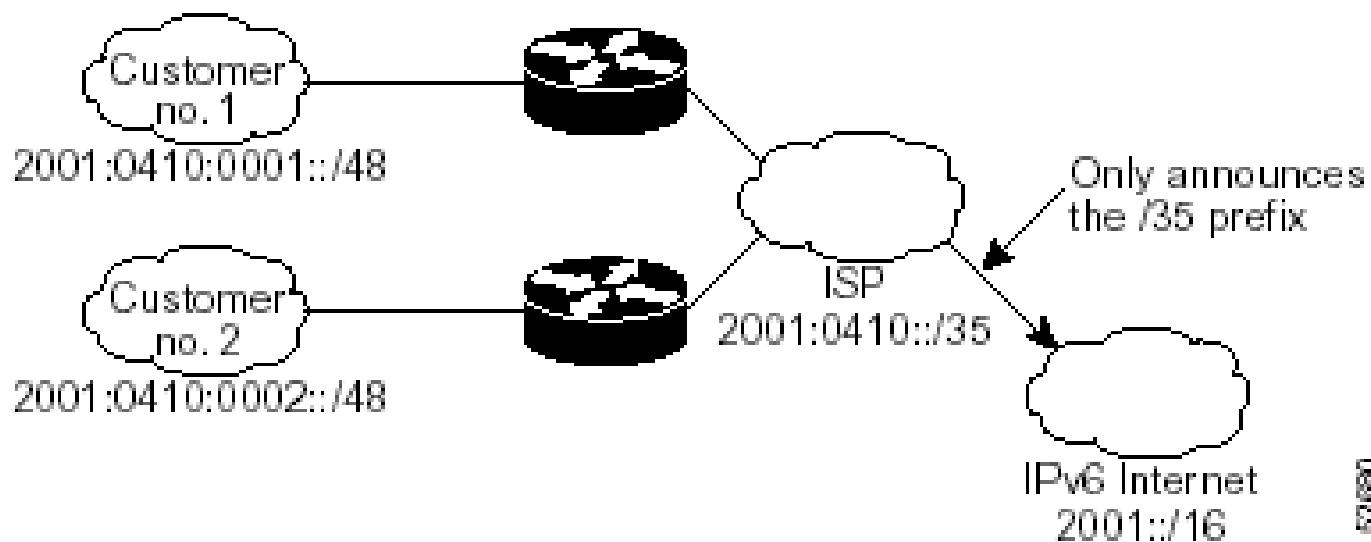
- Estrutura hierárquica existem 4 níveis:



# Endereço IPv6 Unicast

## □ Exemplo de agregação

- um ISP pode dividir o seu prefixo pelos seus clientes agregando-os quando os anuncia na Internet



# Endereços IPv6 Unicast

## □ Top-Level Aggregation ID

- Os identificadores TLA são o topo da hierarquia de roteamento.
- Suporta 8.192 ou ( $2^{13}$ ) identificadores TLA.
  - Esse campo pode ser aumentado através de um espaço previamente reservado
- Administrados pela IANA e atribuídos a registros de Internet locais que, por sua vez, atribuem IDs de TLA individuais a fornecedores de serviços Internet (ISP, Internet Service Provider) globais de grandes dimensões
- Os roteadores devem ter uma entrada na tabela de roteamento para cada TLA ID ativo



# Endereços IPv6 Unicast

## □ Next-Level Aggregation ID

- Organização que recebe um TLA ID tem um espaço de endereçamento de 24 bits para o campo NLA
  - 16.777.216 ou  $2^{24}$  endereços
- Permite a um ISP criar vários níveis de hierarquia de endereçamento para organizar o endereçamento e encaminhamento e para identificar locais.



# Endereços IPv6 Unicast

- Site-Level Aggregation ID
  - Utilizado por uma organização individual para identificar sub-redes no respectivo local
    - Organização pode utilizar estes 16 bits no respectivo local para criar 65.536 sub-redes ou vários níveis de hierarquia de endereçamento e uma infra-estrutura de encaminhamento eficiente
    - Equivalente a uma rede Classe A no IPv4



# Endereços IPv6 Unicast

## □ Interface ID

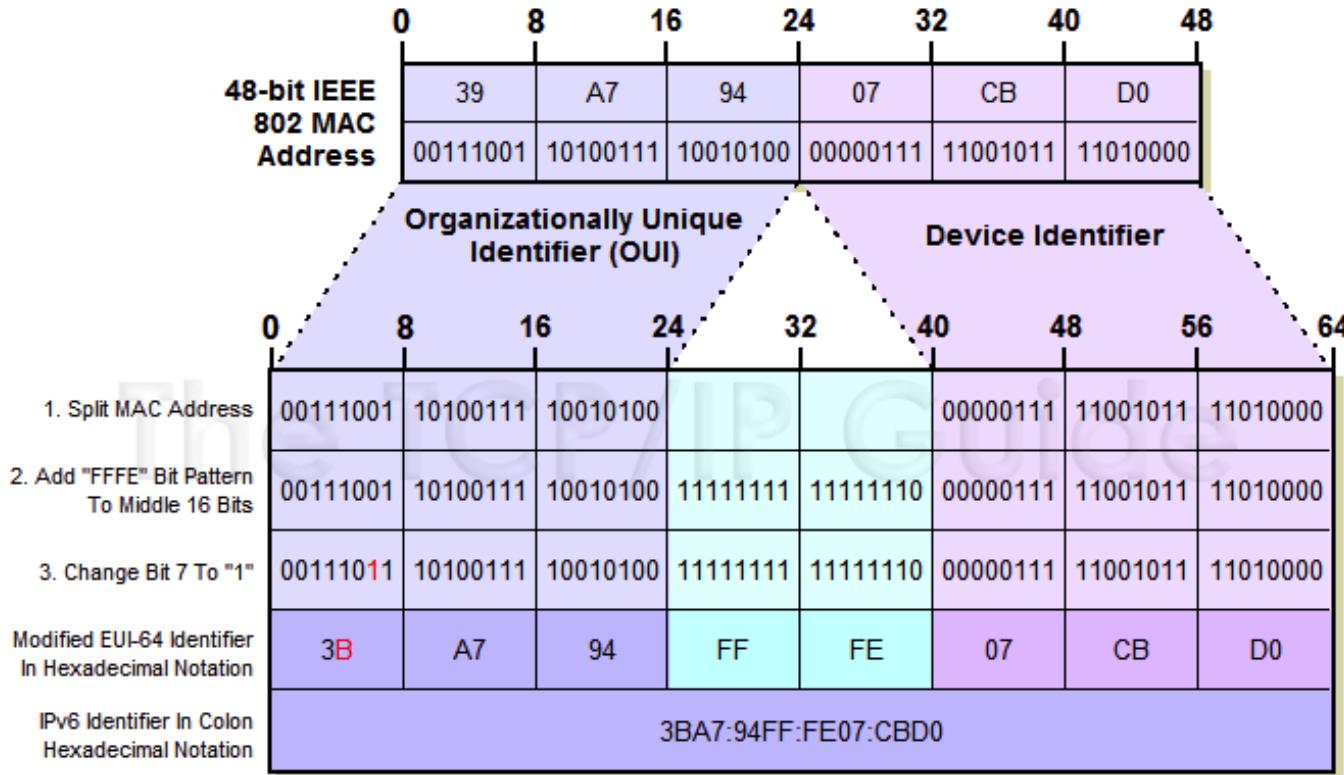
- Indica a interface de um nó numa sub-rede
- Para IPv6 sobre Ethernet, existe uma combinação entre o endereço MAC e alguns algarismos característicos. Exemplo:
  - Para um endereço MAC: 00:A0:C9:C8:E0:C2
  - E um prefixo de rede : 2001::/16
  - Seria obtida a seguinte Interface ID:  
**02A0:C9FF:FEC8:E0C2**
  - E o endereço completo seria:  
2001::02A0:C9FF:FEC8:E0C2/128



# Endereços IPv6 Unicast

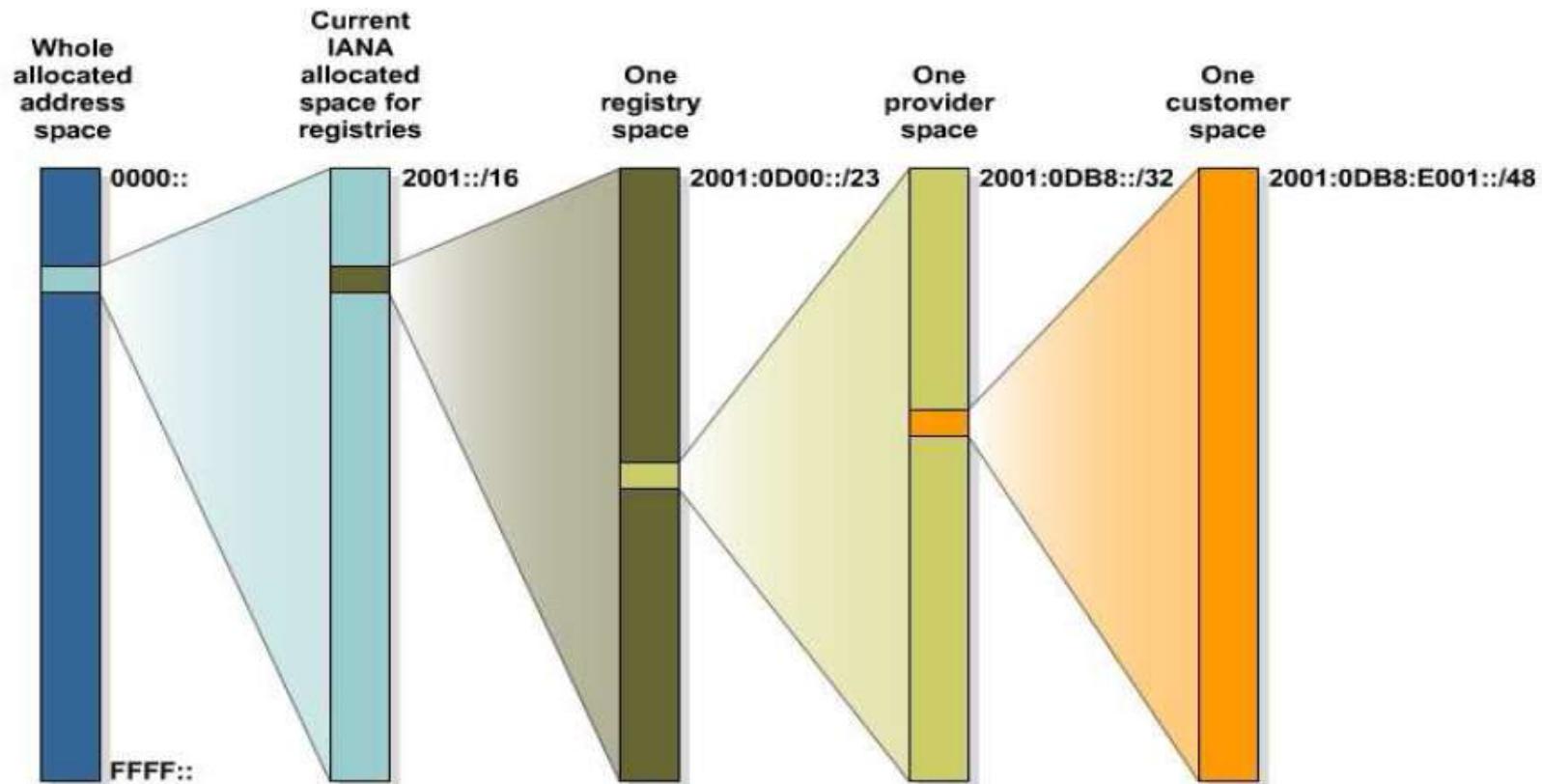
## □ Interface ID

### ○ Convertendo MAC em endereço IPv6



# Endereços IPv6 Unicast

- Essa foi a política de alocação adotada até 2006.



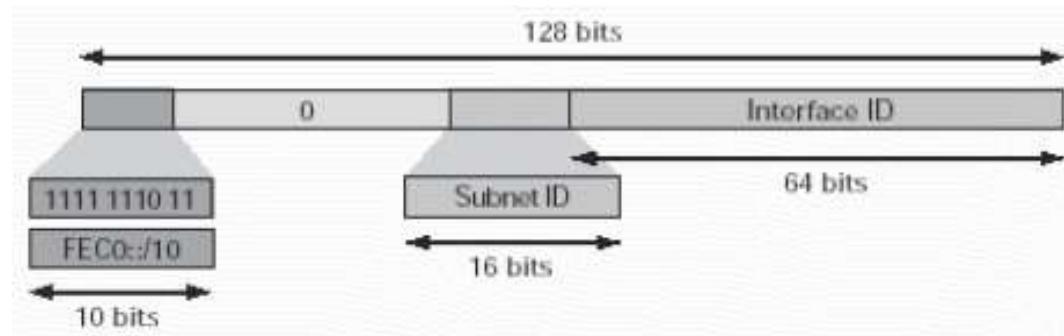
# Endereços IPv6 Unicast

- Política de alocação de endereços IPv6
  - Cada RIR recebe da IANA um bloco /12
  - Os provedores recebem dos RIRs blocos /32
  - Provedores devem entregar aos seus clientes blocos variando entre /48 e /56, dependendo de suas necessidades:
    - Um bloco /48 pode ser dividido em até 65.536 redes diferentes, cada uma com 18.446.744.073.709.551.616 endereços;
    - Um bloco /56 pode ser dividido em até 256 redes diferentes, cada uma com 18.446.744.073.709.551.616 endereços
    - Um /64 pode ser designado a um usuário se houver certeza que apenas uma rede atende às suas necessidades (usuários domésticos)

# Endereços IPv6 Unicast

## □ Endereços locais de site

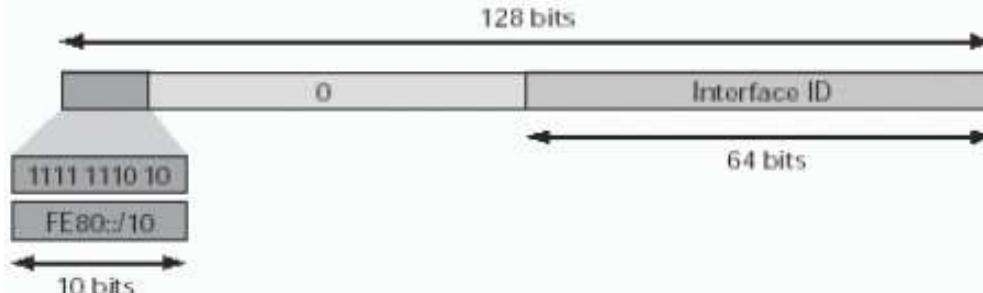
- São equivalentes ao espaço de endereço privado IPv4 (10.0.0.0/8, 172.16.0.0/12 e 192.168.0.0/16).
  - Não são configurados automaticamente
- Endereços podem ser usados para uma comunicação restrita dentro de um domínio específico
  - Desta forma ele não pode ser anunciado externamente por roteadores



# Endereços IPv6 Unicast

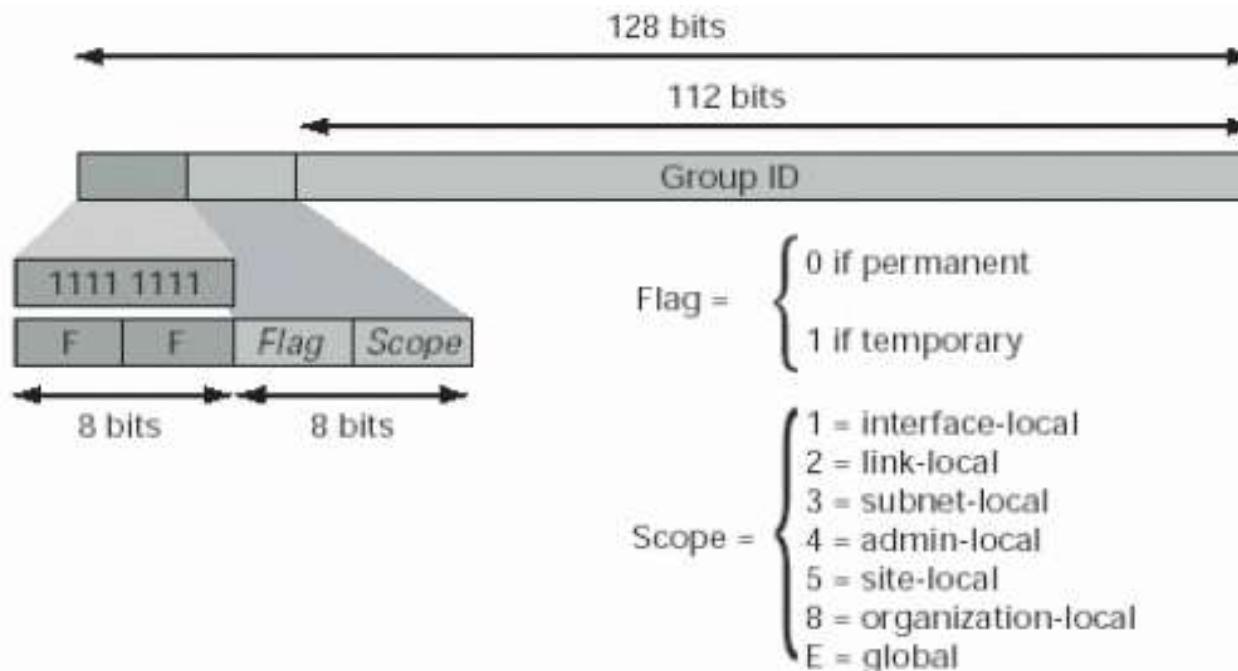
## □ Endereços de enlace local

- automaticamente configurado em qualquer *host* IPv6
- Utilizados pelos nós ao comunicar com nós vizinhos no mesmo enlace
  - Numa rede IPv6 de um único enlace sem roteador, os endereços locais de ligação são utilizados para comunicar entre nós do enlace
    - Equivalentes aos endereços IPv4 de endereçamento IP privado automático
  - Um endereço local de enlace é necessário para os processos de identificação de vizinhança
- Um roteador do IPv6 nunca reencaminha o tráfego local de ligação para além da liaison.



# Endereços IPv6 Multicast

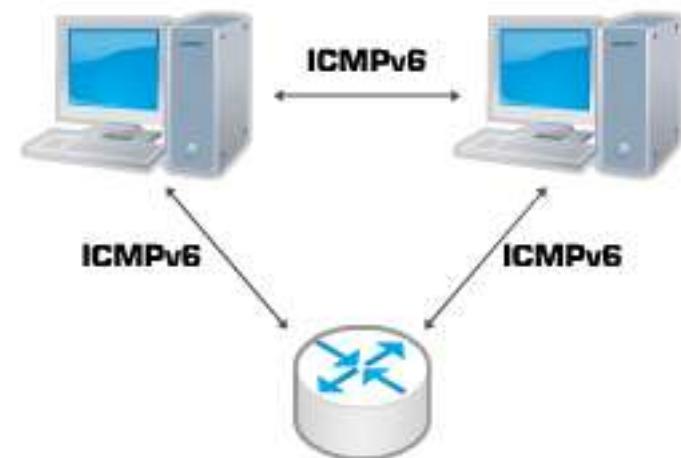
- Flags: 000T
  - T=0: indica endereço multicast reservado pela IANA
  - T=1: indica endereço multicast não permanente
- Valor Scope limita o escopo do grupo multicast



# ICMPv6

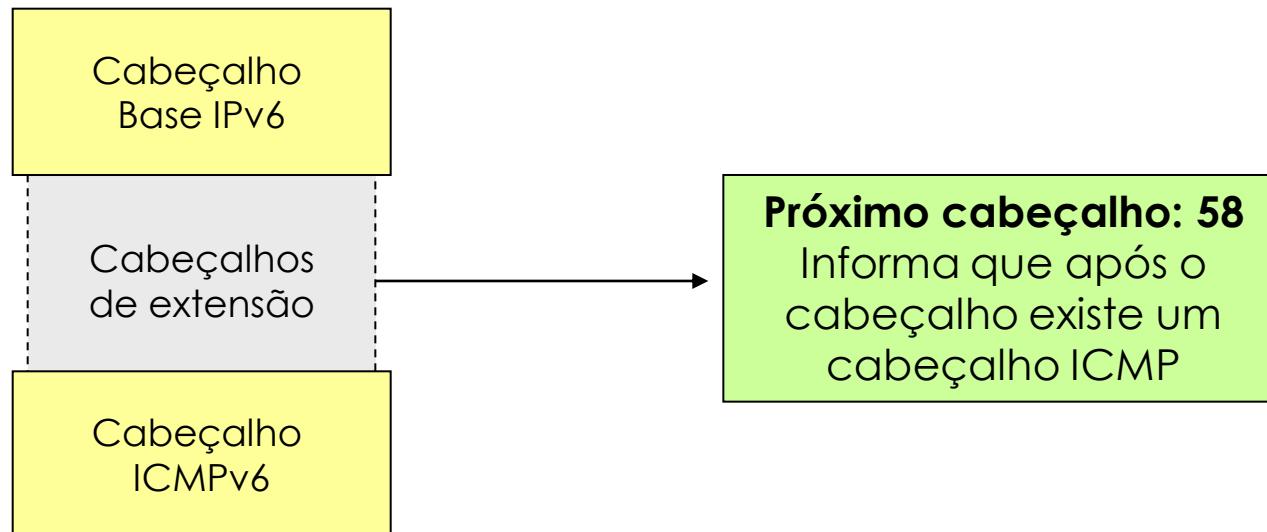
## □ ICMPv6

- Possui as mesmas funções básicas do ICMPv4
  - Informar características da rede
  - Diagnósticos
  - Informar erros no processamento e envio dos pacotes
- Dois tipos (classes) de mensagens
  - Mensagens de informação
  - Mensagens de erro



# ICMPv6

- O cabeçalho ICMPv6 é precedido pelos cabeçalhos de extensão (se houver) e pelo cabeçalho base do IPv6

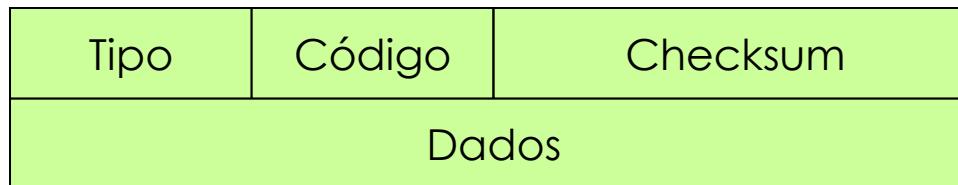


# ICMPv6

## □ Cabeçalho ICMP

**Tipo:** informa o tipo da mensagem.  
Possui 8 bits.

**Código:** fornece informações adicionais para alguns tipos de mensagens.  
Possui 8 bits.



**Checksum:** utilizado para encontrar erros e dados corrompidos no cabeçalho ICMP.  
Possui tamanho de 16 bits.

**Dados:** possuem informações referentes ao erro ocorrido, dependendo do tipo da mensagem. De acordo com a mensagem, o tamanho deste campo pode variar.

# IPv6

- Apresenta uma quantidade maior de mensagens que a versão ICMP v4
  - Além das funções básicas do ICMP são incluídas:
    - Descoberta de vizinhança
      - Incorporadas funções dos protocolos ARP/RARP
    - Gerenciamento de grupo Multicat
      - IGMP (Internet Group Management Protocol)
    - Mobilidade IPv6
    - Descoberta do Path MTU

# ICMPv6

## □ Algumas mensagens de erro :

Mensagens de Erro:		
Tipo	Nome	Descrição
1	<i>Destination Unreachable</i>	Indica falhas na entrega do pacote como endereço ou porta desconhecida ou problemas na comunicação.
2	<i>Packet Too Big</i>	Indica que o tamanho do pacote é maior que a Unidade Máxima de Transito (MTU) de um enlace.
3	<i>Time Exceeded</i>	Indica que o Limite de Roteamento ou o tempo de remontagem do pacote foi excedido.
4	<i>Parameter Problem</i>	Indica erro em algum campo do cabeçalho IPv6 ou que o tipo indicado no campo Próximo Cabeçalho não foi reconhecido.
100-101		Uso experimental.
102-126		Não utilizado.
127		Reservado para expansão das mensagens de erro ICMPv6.

# ICMPv6

## □ Algumas mensagens de informação:

Mensagens de Informação:		
Tipo	Nome	Descrição
128	<i>Echo Request</i>	Utilizadas pelo comando ping.
129	<i>Echo Reply</i>	
130	<i>Multicast Listener Query</i>	
131	<i>Multicast Listener Report</i>	Utilizadas no gerenciamento de grupos multicast.
132	<i>Multicast Listener Done</i>	
133	<i>Router Solicitation</i>	
134	<i>Router Advertisement</i>	
135	<i>Neighbor Solicitation</i>	Utilizadas com o protocolo Descoberta de Vizinhança.
136	<i>Neighbor Advertisement</i>	
137	<i>Redirect Message</i>	
138	<i>Router Renumbering</i>	Utilizada no mecanismo de Re-endereçamento (Renumbering) de roteadores.
139	<i>ICMP Node Information Query</i>	Utilizadas para descobrir informações sobre nomes e endereços, são atualmente limitadas a ferramentas de diagnóstico, depuração e gestão de redes.
140	<i>ICMP Node Information Response</i>	
141	<i>Inverse Neighbor Discovery Solicitation Message</i>	Utilizadas em uma extensão do protocolo de Descoberta de Vizinhança.
142	<i>Inverse Neighbor Discovery Advertisement Message</i>	

# ICMPv6

## □ Algumas mensagens de informação:

144	<i>Home Agent Address Discovery Request Message</i>	Utilizadas no mecanismo de Mobilidade IPv6.
145	<i>Home Agent Address Discovery Reply Message</i>	
146	<i>Mobile Prefix Solicitation</i>	
147	<i>Mobile Prefix Advertisement</i>	
148	<i>Certification Path Solicitation Message</i>	Utilizadas pelo protocolo SEND.
149	<i>Certification Path Advertisement Message</i>	
150		Utilizada experimentalmente com protocolos de mobilidade como o Seamoby.
151	<i>Multicast Router Advertisement</i>	Utilizadas pelo mecanismo Multicast Router Discovery.
152	<i>Multicast Router Solicitation</i>	
153	<i>Multicast Router Termination</i>	
154	<i>FMIPv6 Messages</i>	Utilizada pelo protocolo de mobilidade Fast Handovers
200-201		Uso Experimental
255		Reservado para expansão das mensagens de erro ICMPv6

# IPv6

## □ Descoberta de vizinhança

### ○ Utiliza 5 mensagens ICMPv6

- **Router Solicitation** (tipo 133): utilizada pelos hosts para pedir uma mensagem do tipo *Router Advertisement*
- **Router Advertisement** (tipo 134): são mensagens enviadas periodicamente (ou em resposta a uma mensagem de *Router Solicitation*) pelos roteadores para avisar que encontram-se presentes no enlace
- **Redirect** (tipo 137): são enviadas pelos roteadores para informar ao host qual é o roteador mais indicado para que seu pacote chegue ao destino



RA packet definitions:

ICMP Type = 134

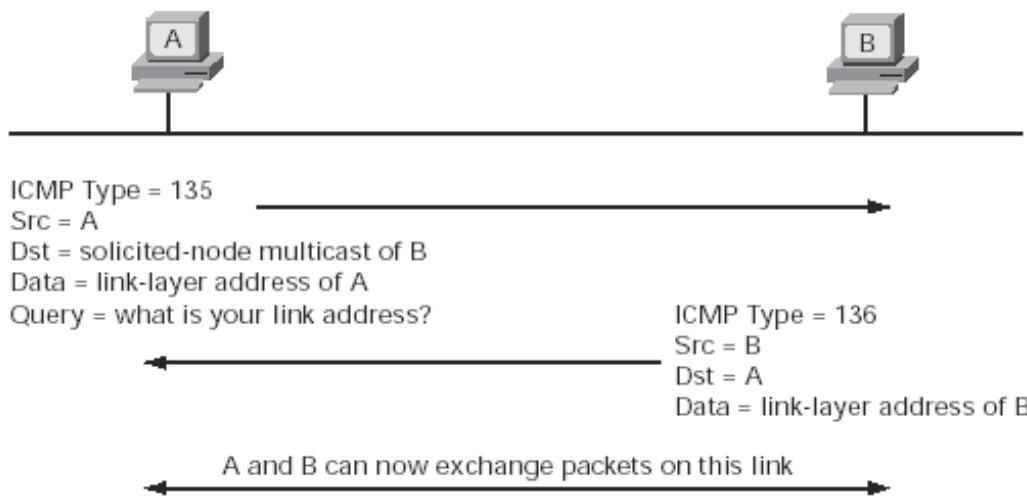
Src = router link-local address

Dst = all-nodes multicast address

Data = options, prefix, lifetime, autoconfig flag

# IPv6

- Descoberta de vizinhança
  - Descoberta de endereços da camada de enlace
    - Determina o endereço MAC dos vizinhos do mesmo enlace.
      - **Neighbor Solicitation** (tipo 135): mensagem *multicast* enviada pelos nós para determinar endereço MAC e acessibilidade de um vizinho. Também pode detectar endereços duplicados.
      - **Neighbor Advertisement** (tipo 136): é uma mensagem enviada como resposta a uma *Neighbor Solicitation*. Quando há mudança em algum endereço MAC, esta mensagem também é enviada
    - Substitui o protocolo ARP do IPv4, utilizando um endereço *multicast* como destino



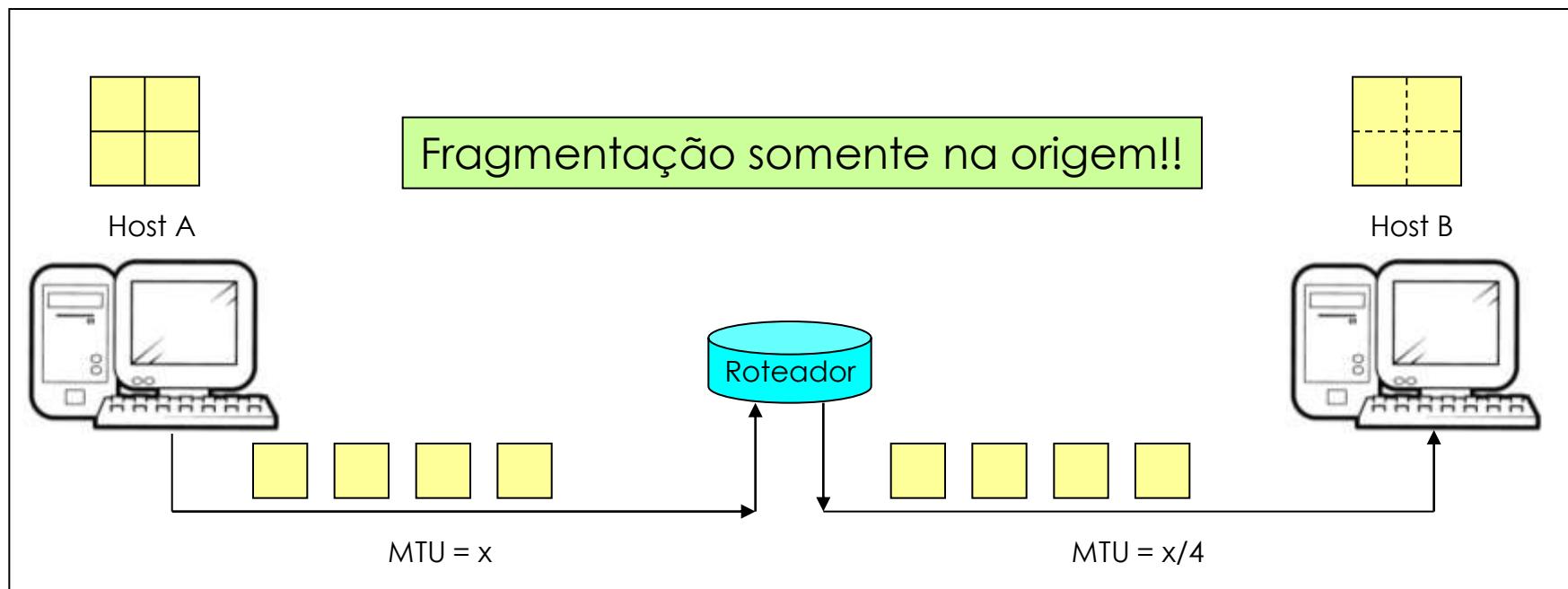
# IPv6

## □ Path MTU Discovery

- Assume que o MTU é o mesmo MTU do enlace inicial
- Se, no caminho, o tamanho de qualquer pacote for maior que o MTU informado pelo roteador do próximo enlace, o roteador descarta o pacote, e retorna um *ICMPv6 packet too big*
- Este mecanismo continua até que o tamanho do pacote seja igual ou menor ao menor MTU do caminho, realizando quantas reduções forem necessárias

# IPv6

- Assim, diminui-se o overhead dos roteadores, pois a fragmentação é realizada na origem



# Mobilidade IPv6

## □ Necessidades, Metas e Aplicações

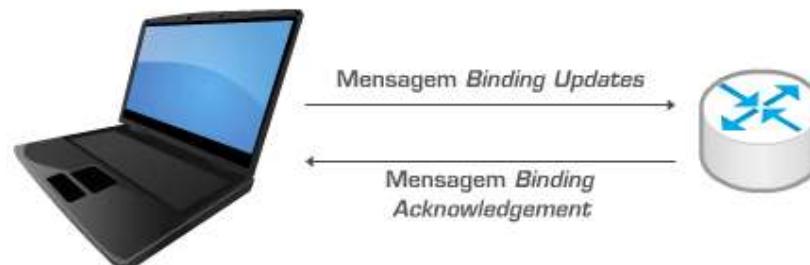
- Suporte a mobilidade no IPv6 móvel provê mecanismos para que um nó possa mudar de uma sub-rede para outra de forma transparente ao usuário.
  - Estas sub-redes não precisam necessariamente ser homogêneas: pode ocorrer do *host* mudar de sub-rede e de forma de acesso.
    - Nó ligado num momento a um segmento Ethernet, e em um próximo instante estar conectado via ondas de rádio
- Protocolo deve manter a comunicação com outros nós após a mudança de sub-rede do nó móvel
  - O *host* deve sempre ser acessível através de seu endereço de origem (*home address*), os pacotes enviados para este endereço devem ser repassados para a posição atual do *Host Móvel* (HM).
  - A mudança de sub-rede do host deve ser transparente para a camadas de transporte e superiores.

# Mobilidade IPv6

- Todo dispositivo móvel possui três endereços
  - **home address**, associado permanentemente a ele, da mesma forma que ocorre com qualquer outro nodo.
  - **local-Link address**, que não é roteável, mas é garantidamente único naquela sub-rede.
    - nodos da mesma sub-rede podem entrar em contato com o HM sem intermédio de nenhum roteador.
  - **care-of address**, associado ao dispositivo enquanto ele pertencer a uma outra sub-rede em particular
    - Este endereço pode ser roteado normalmente pelos mecanismos de roteamento da Internet.

# Mobilidade IPv6

- Quando nó móvel se desloca para fora de sua rede de origem
  - Obtêm um care-of address
  - Para que os pacotes cheguem neste novo endereço é necessária uma associação entre o home address e o care-of address
    - Associação chamada de binding
    - Feita por um Agente de Origem (roteador na rede de origem)
  - Criação do binding:



# Segurança IPv6

## □ Segurança IPv4

- Foi projetado para redes acadêmicas, que após utilizada como estrutura de comunicação comercial produziu problemas de segurança
- IPSec
  - Protocolo que implementa criptografia e autenticação no IPv4, garantindo:
    - Que a mensagem recebida não tenha sido adulterada;
    - A identidade do remetente;
    - A confidencialidade da mensagem, criptografando seu conteúdo
  - Mecanismo de autenticação não pode ser utilizado em conexões que estejam atrás de NAT

# Segurança IPv6

## □ Segurança IPv6

- Implantação do IPSec é mais simples e tem as mesmas funções do IPv4:
  - Não há necessidade de se usar NAT permitindo que o IPSec funcione sem restrições;
  - Mecanismos de autenticação e encapsulamento do IPSec fazem parte do protocolo
    - Seu suporte é obrigatório em todos os nós (que não ocorre no IPv4)

# Transição de IPv4 para IPv6

- Impossível atualizar todos os roteadores/hosts simultaneamente
  - Impossível marcar uma data para a troca de IPv4 por IPv6
  - Como as redes operam mixando IPv4 e IPv6?
    - Via migração suave do IPv4 ao IPv6



# Transição de IPv4 para IPv6

Estas técnicas de transição podem ser classificadas nas seguintes categorias:



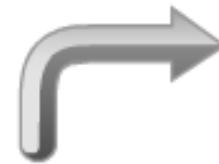
**Pilha Dupla**

Provê o suporte a ambos os protocolos no mesmo dispositivo.



**Tunelamento**

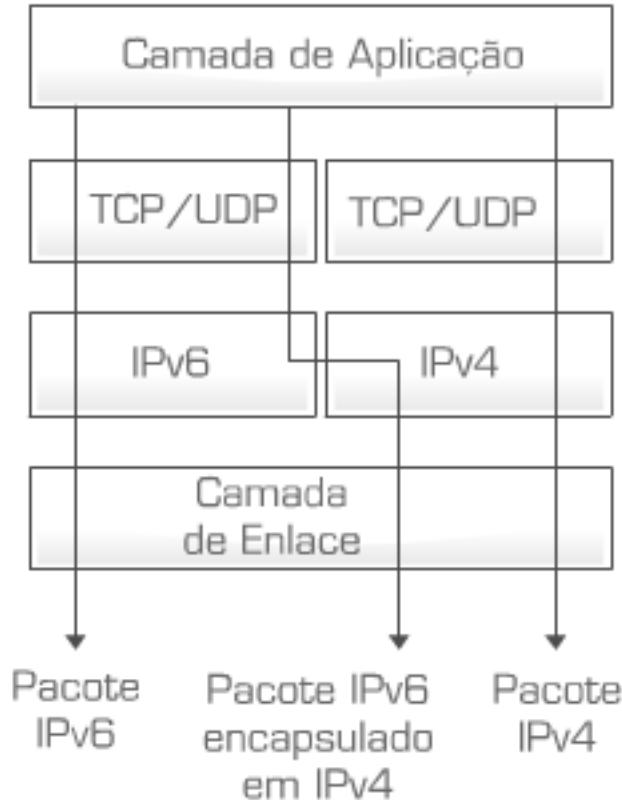
Permite o tráfego de pacotes IPv6 sobre estruturas de rede IPv4, ou o inverso.



**Tradução**

Permite a comunicação entre nós com suporte apenas a IPv6 com nós que suportem apenas IPv4 e vice-versa.

# Transição de IPv4 para IPv6



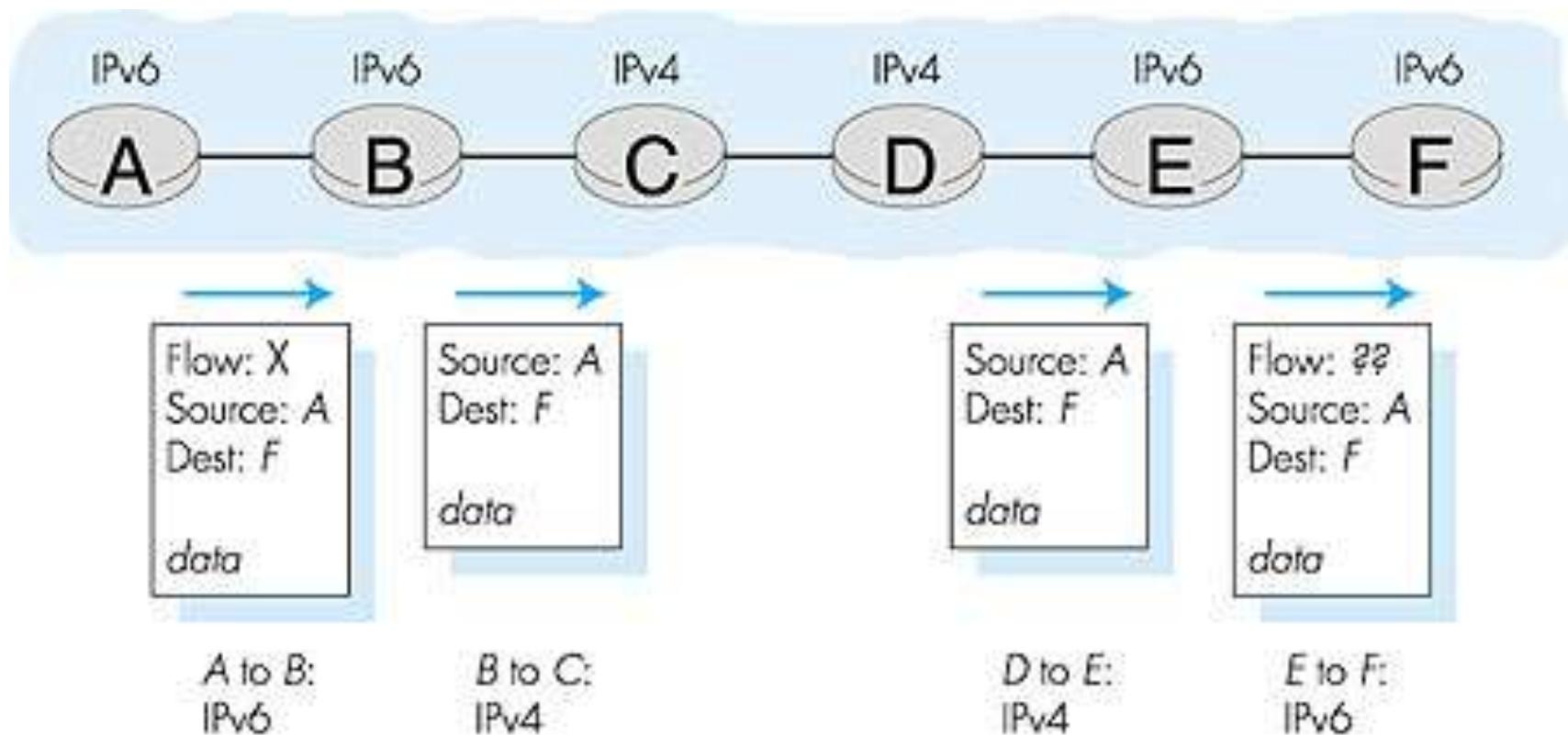
## Pilha Dupla

Com o método de transição da Pilha Dupla, *hosts* e roteadores tornam-se capazes de enviar e receber pacotes tanto para o IPv4, quanto para o IPv6.

Esse método permitirá que um nó Pilha Dupla (IPv6/IPv4), ao se comunicar com um nó IPv6, se comporte como um nó IPv6 e na comunicação com um nó IPv4, como nó IPv4.

Para isso, cada nó IPv4/IPv6 é configurado com ambos os endereços, utilizando mecanismos IPv4, como por exemplo DHCP, para adquirir endereços IPv4, e mecanismos do IPv6 para endereços IPv6.

# Abordagem Pilha Dual



# Transição de IPv4 para IPv6

## Tunelamento

A técnica de criação de túneis, ou tunelamento, permite transmitir pacotes IPv6 através da infraestrutura IPv4 já existente, sem a necessidade de realizar qualquer mudança nos mecanismos de roteamento, encapsulando o conteúdo do pacote IPv6 em um pacote IPv4.



De uma forma geral, o funcionamento de um túnel é bem simples. O nó de entrada do túnel, cria um cabeçalho IPv4 com o pacote IPv6 encapsulado e o transmite através da rede IPv4. Este processo de encapsulamento, conhecido como 6in4, é identificado como protocolo do tipo 41.

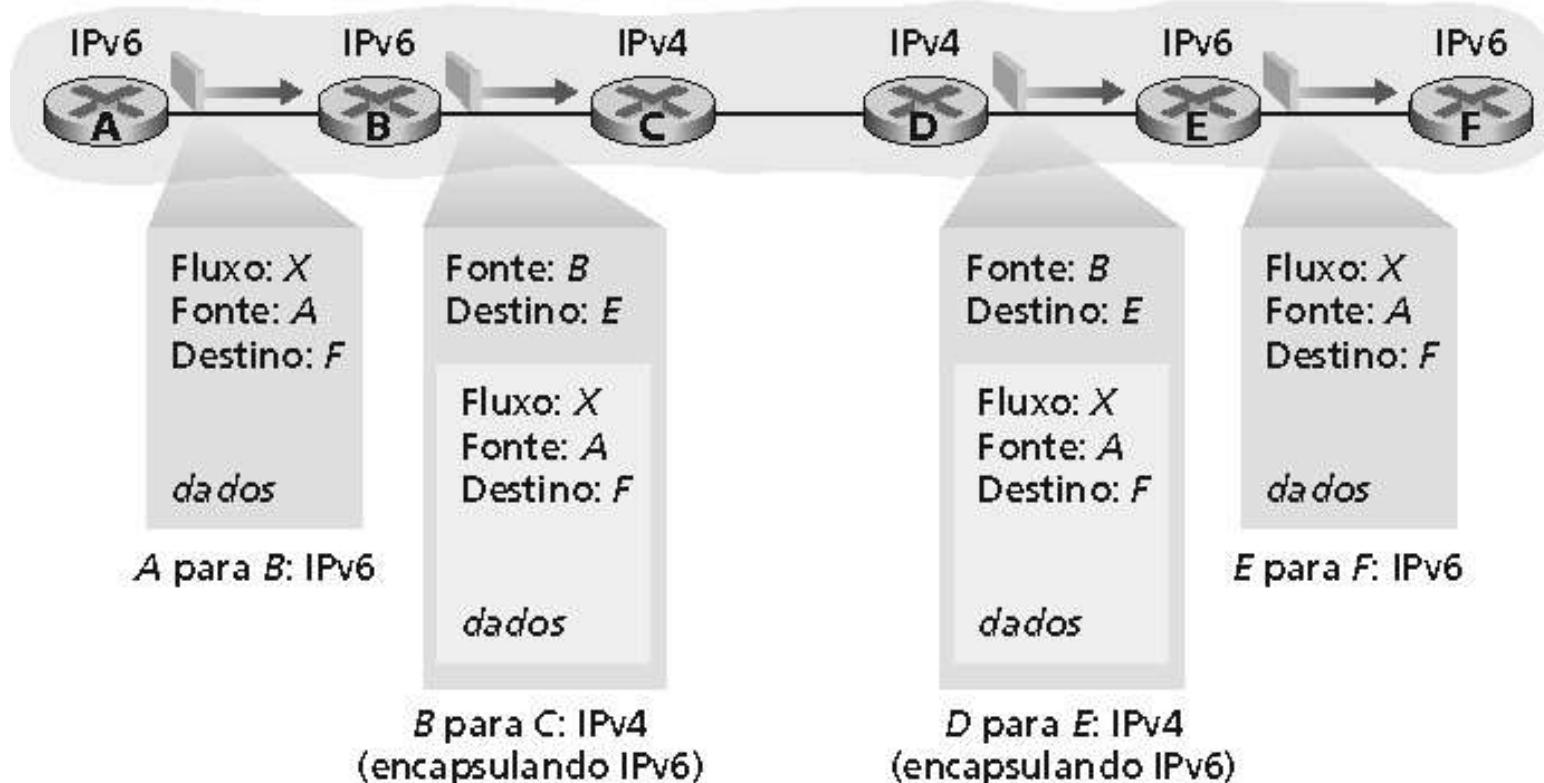
A utilização do protocolo 41 é comum em algumas técnicas de tunelamento, como 6to4, *Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)* e *Tunnel Broker*.

# Tunelamento

Visão lógica



Visão física



# Túneis

- Permitem que
  - Hoje: Ilhas IPv6 conectem-se através de redes IPv4
  - No futuro: Ilhas IPv4 conectem-se através de redes IPv6

