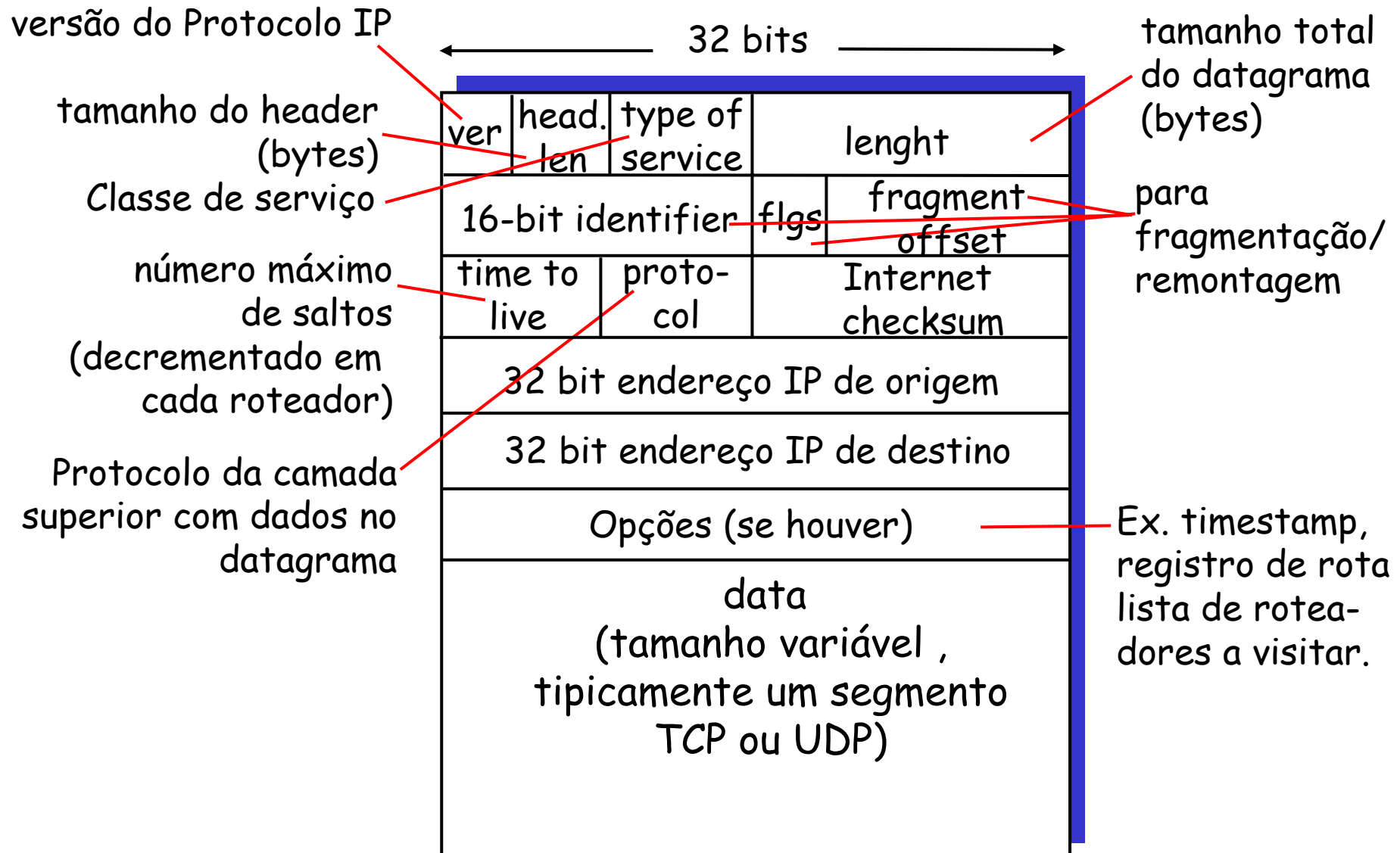
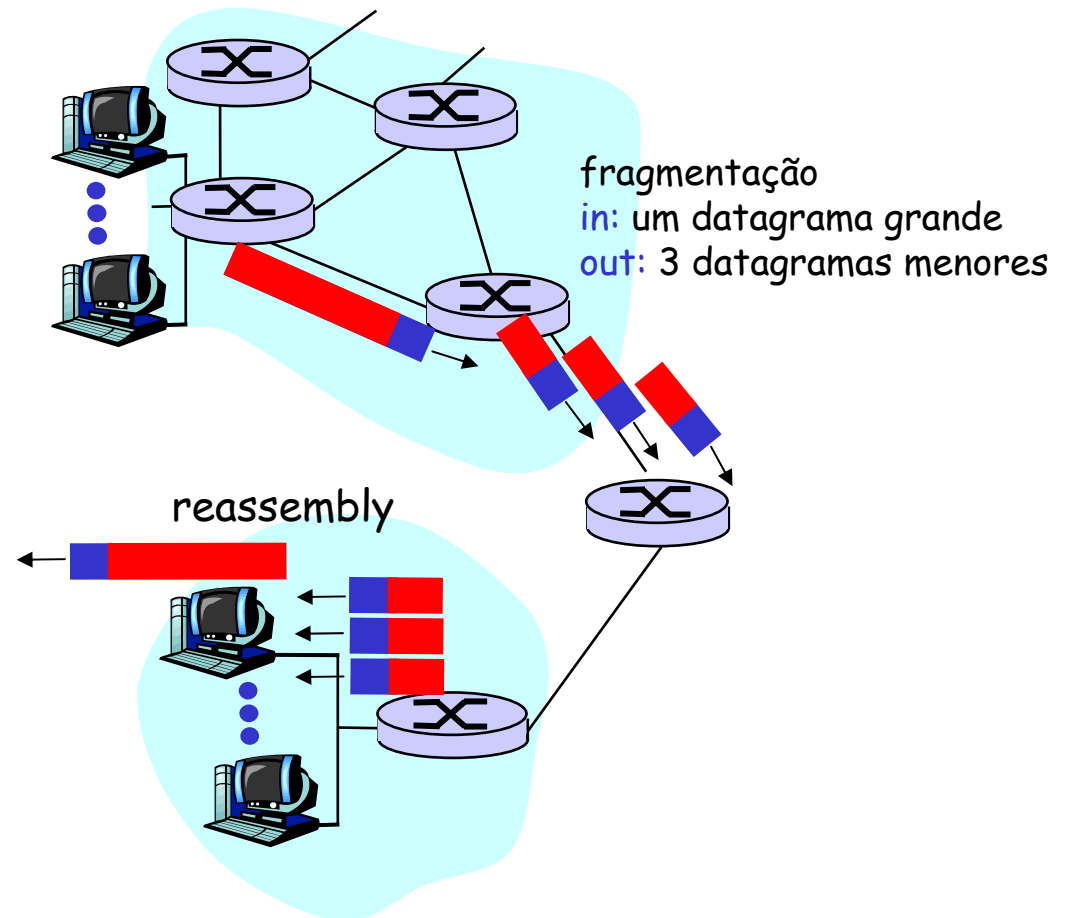


Formato do Datagrama IP



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: 1518 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



IP Fragmentação e Remontagem

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

Um grande datagrama se torna
vários datagramas menores



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

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

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

ICMP: Internet Control Message Protocol

- ❑ usado por computadores e roteadores para troca de informação de controle da camada de rede
 - error reporting: host, rede, porta ou protocolo
 - echo request/reply (usado pela aplicação ping)
- ❑ transporte de mensagens:
 - mensagens ICMP transportadas em datagramas Ip
- ❑ **ICMP message:** tipo, código, mais primeiros 8 bytes do datagrama IP que causou o erro

<u>Tipo</u>	<u>Código</u>	<u>descrição</u>
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

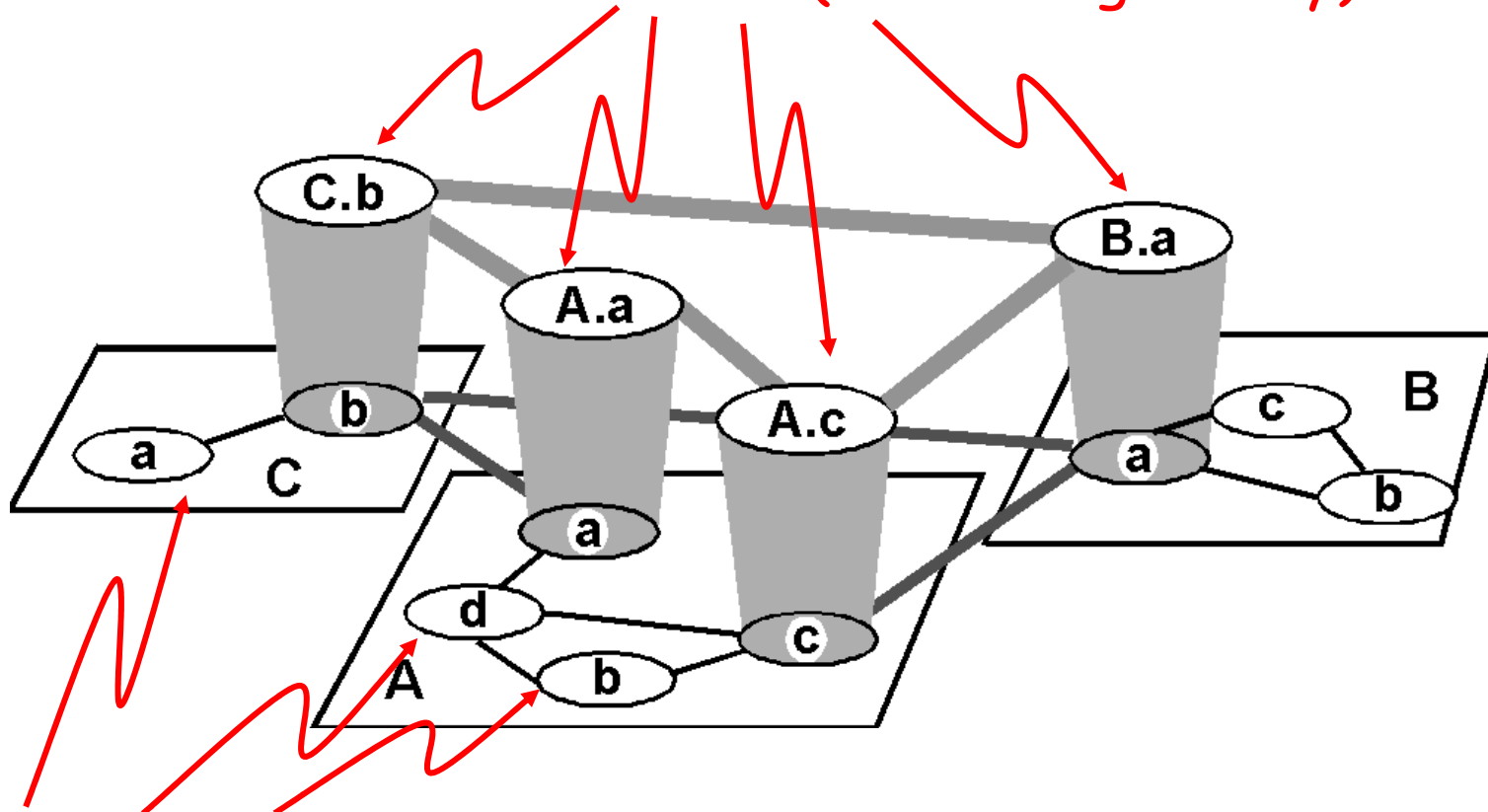
Roteamento na Internet

- ❑ A Internet consiste de **Sistemas Autônomos (AS)** interconectados entre si:
 - **Stub AS**: pequena corporação
 - **Multihomed AS**: grande corporação (sem tráfego de trânsito)
 - **Transit AS**: provedor

- ❑ Dois níveis de roteamento:
 - **Intra-AS**: o administrador é responsável pela definição do método de roteamento
 - **Inter-AS**: padrão único

Hierarquia de AS

Roteador de borda Intra-AS (exterior gateway)



Roteador interno Inter-AS (gateway)

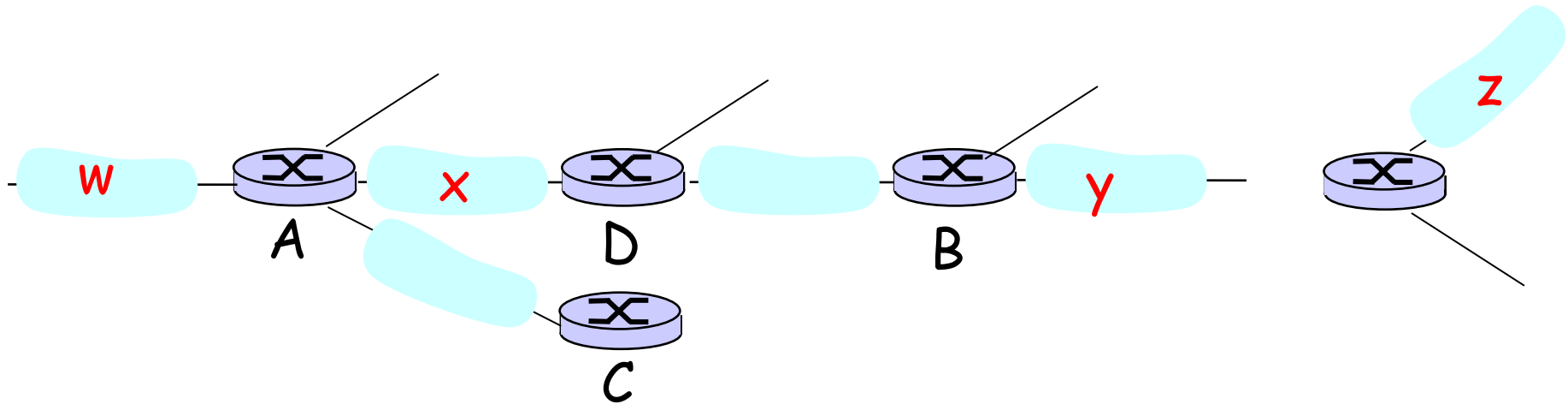
Roteamento Intra-AS

- ❑ Também conhecido como **Interior Gateway Protocols (IGP)**
- ❑ IGP's mais comuns:
 - RIP: Routing Information Protocol
 - OSPF: Open Shortest Path First
 - IGRP: Interior Gateway Routing Protocol (proprietário da Cisco)

RIP (Routing Information Protocol)

- ❑ Algoritmo do tipo vetor distância
- ❑ Incluso na distribuição do BSD-UNIX em 1982
- ❑ Métrica de distância: # of hops (max = 15 hops)
 - *motivo: simplicidade*
- ❑ Vetores de distância: trocados cada 30 sec via Response Message (também chamado **advertisement**, ou **anúncio**)
- ❑ Cada anúncio: indica rotas para até 25 redes de destino

RIP (Routing Information Protocol)



Rede de Destino	Next Router	Num. de saltos para dest.
W	A	2
Y	B	2
Z	B	7
X	--	1
....

Tabela de roteamento em D

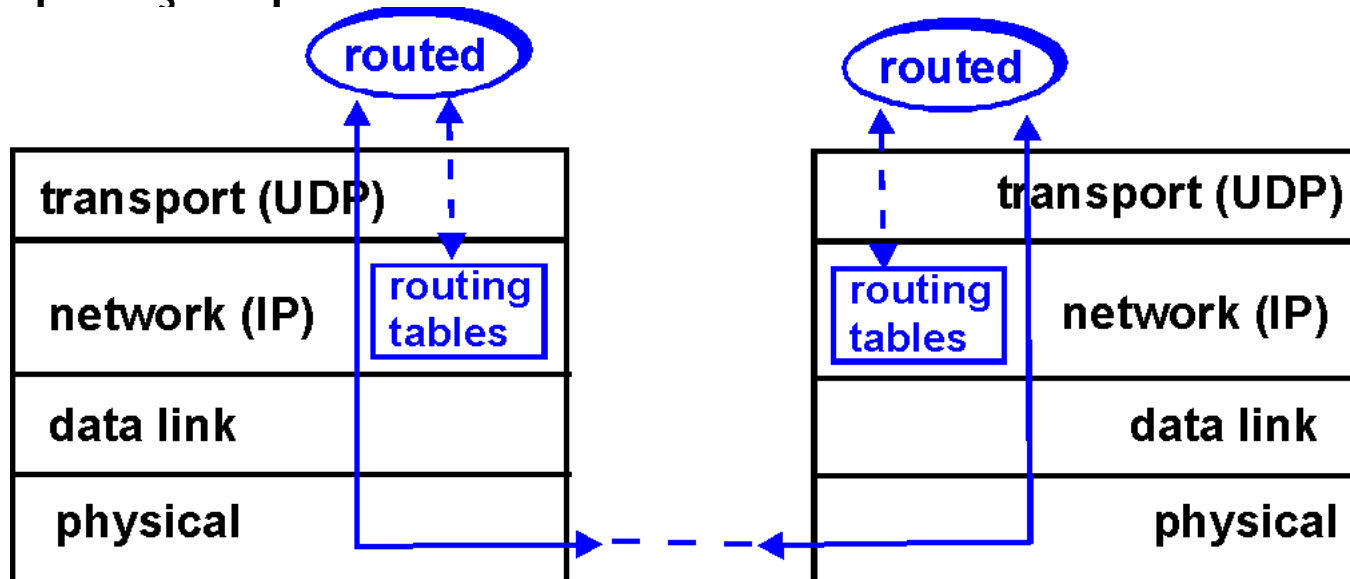
RIP: Falha de Enlaces e Recuperação

Se não há depois de 180 sec --> 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
- poison reverse é usado para prevenir loops, isto é, evitar que a rota para um destino passe pelo próprio roteador que está enviando a informação de distância (distância infinita= 16 hops)

RIP Processamento da tabela de rotas

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



Exemplo de tabela RIP

Roteador: *giroflée.eurocom.fr*

Destination	Gateway	Flags	Ref	Use	Interface
-----	-----	-----	-----	-----	-----
127.0.0.1	127.0.0.1	UH	0	26492	lo0
192.168.2.	192.168.2.5	U	2	13	fa0
193.55.114.	193.55.114.6	U	3	58503	le0
192.168.3.	192.168.3.5	U	2	25	qaa0
224.0.0.0	193.55.114.6	U	3	0	le0
default	193.55.114.129	UG	0	143454	

- ❑ 3 redes classe C diretamente conectadas (LANs)
- ❑ Roteador somente conhece rotas para as LANS locais (nesse caso particular)
- ❑ rota Default usada para mandar mensagens para fora
- ❑ endereço de rota multicast: 224.0.0.0
- ❑ Loopback interface (para depuração): 127.0.0.1

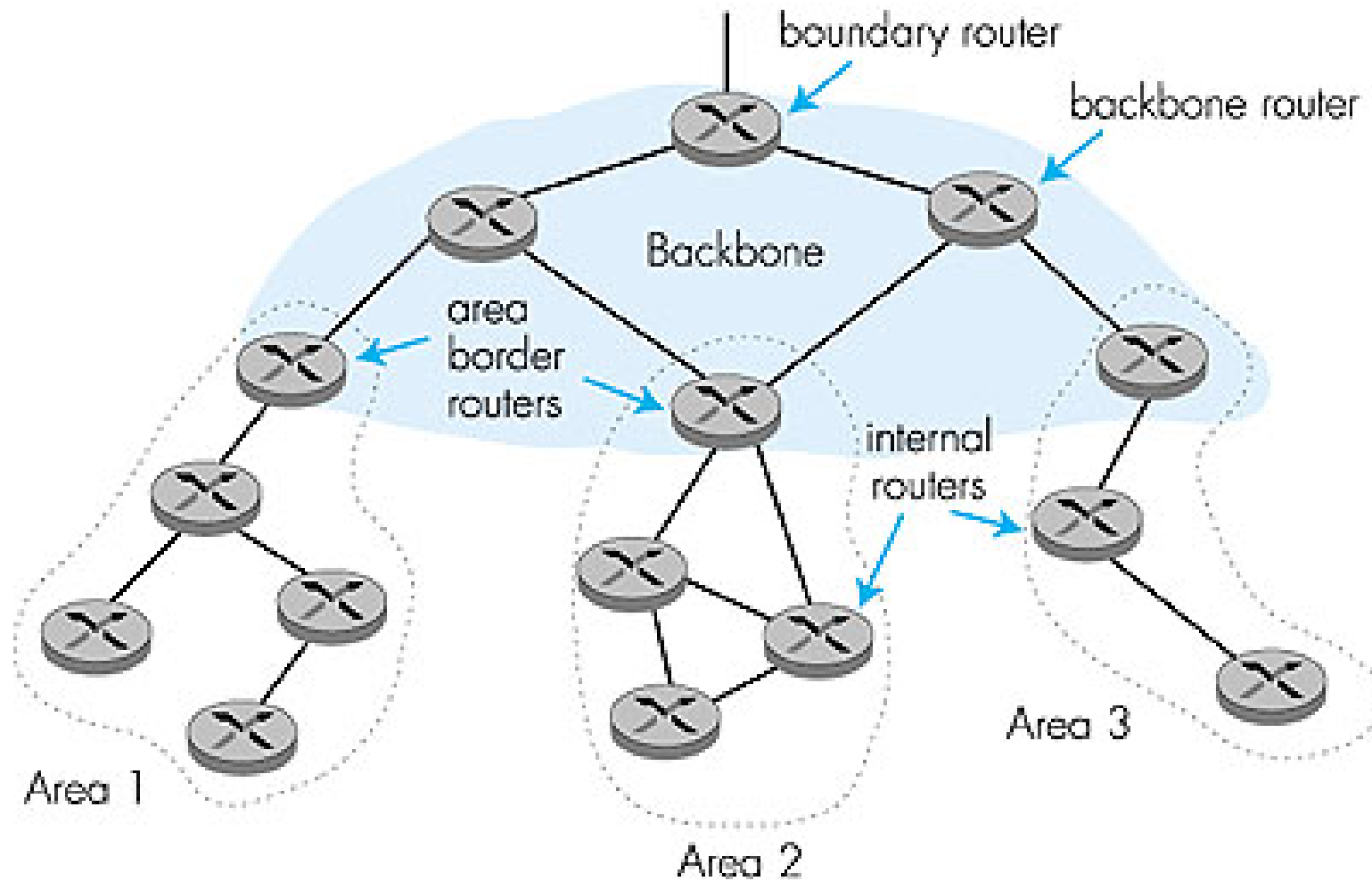
OSPF (Open Shortest Path First)

- ❑ “open”: publicamente disponível
- ❑ Usa algoritmo do tipo Link State
 - disseminação de pacotes LS
 - Mapa topológico em cada nó
 - usa algoritmo de Dijkstra's para cálculo de rotas
- ❑ anúncios do OSPF transportam um registro para cada roteador vizinho
- ❑ Anúncios são distribuídos para todo o AS (via flooding)

OSPF características avançadas

- ❑ **Segurança:** todas as mensagens do OSPF são autenticadas (para prevenir intrusão de hackers); usa conexões TCP para as suas mensagens
- ❑ Múltiplos caminhos de mesmo custo são permitidos (o RIP só permite um caminho para cada destino)
- ❑ Para cada enlace podem ser calculadas múltiplas métricas uma para cada tipo de serviço (**TOS**) (ex, custo de enlace por satélite definido baixo para tráfego de "melhor esforço" e alto para serviços de tempo real)
- ❑ Integra tráfego uni- e **multicast** :
 - Multicast OSPF (MOSPF) usa a mesma base de dados topológica do OSPF
- ❑ **Hierarchical** OSPF: dois níveis de roteamento para domínios grandes.

OSPF Hierárquico



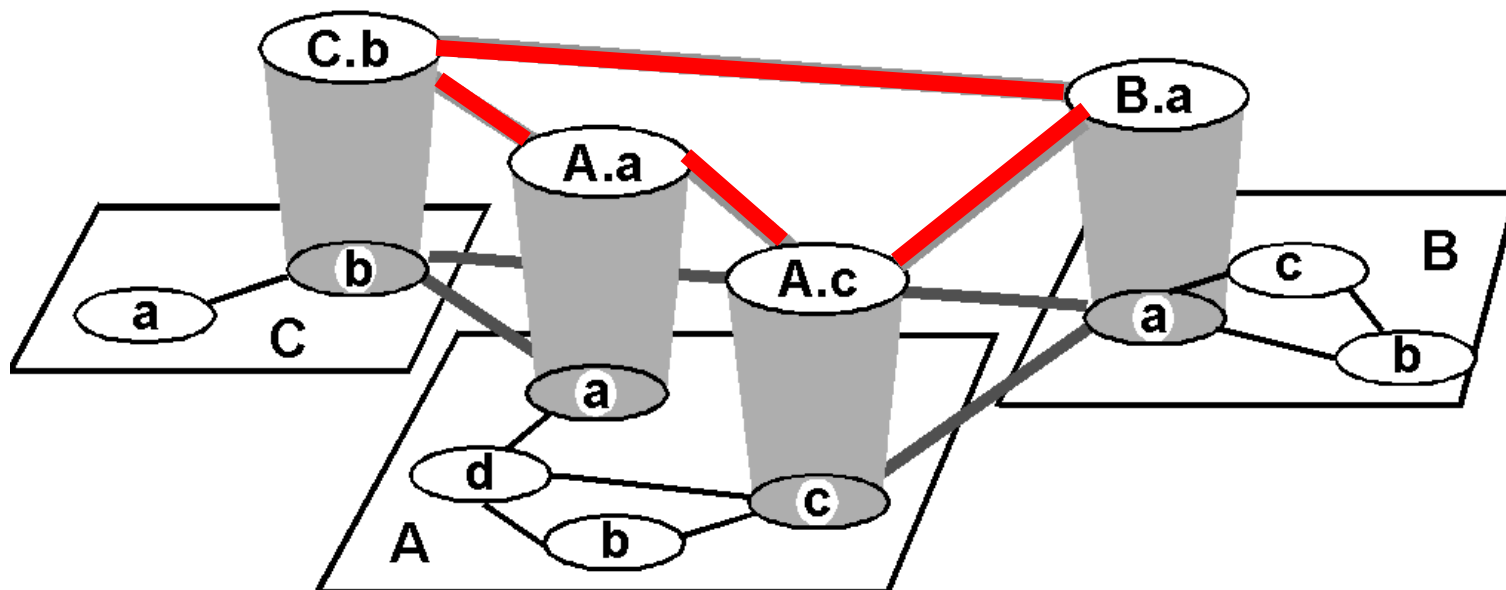
OSPF Hierárquico

- ❑ **Hierarquia de dois níveis:** área local e backbone.
 - anúncios de Link-state são enviados apenas nas áreas
 - cada nó tem a topologia detalhada da área; mas somente direções conhecidas (caminhos mais curtos) para redes em outra áreas.
- ❑ **Area border routers:** “resumem” distâncias para redes na própria área e enviam para outros roteadores de borda de área
- ❑ **Backbone routers:** executam o roteamento OSPF de forma limitada ao backbone.
- ❑ **Boundary routers:** realizam as funções de interconexão com outros sistemas autônomos.

IGRP (Interior Gateway Routing Protocol)

- ❑ Protocolo proprietário da CISCO; sucessor do RIP (meados dos anos 80)
- ❑ Vetor distância, como RIP
- ❑ várias métricas de custo (atraso, banda, confiabilidade, carga, etc.)
- ❑ usa o TCP para trocar informações de novas rotas
- ❑ Loop-free routing via Distributed Updating Algorithm (DUAL) baseado em técnicas de *computação difusa*

Inter-AS routing



Internet inter-AS routing: BGP

- ❑ **BGP (Border Gateway Protocol):** é o padrão de fato para uso na Internet
- ❑ **Algoritmo Path Vector :**
 - similar ao protocolo Distance Vector
 - cada Border Gateway envia em broadcast aos seus vizinhos (peers) o caminho inteiro (isto é a seqüência de ASs) até o destino
 - Exemplo: Gateway X deve enviar seu caminho até o destino Z:

$$\text{Path (X,Z)} = X, Y_1, Y_2, Y_3, \dots, Z$$

Internet inter-AS routing: BGP

Suponha: roteador X envia seu caminho ao roteador parceiro W

- ❑ W pode escolher ou não o caminho oferecido por X
 - critérios de escolha: custo, regras (não rotear através de AS rivais), prevenção de loops.

- ❑ Se W seleciona o caminho oferecido por X, então:

$$\text{Path}(W,Z) = w, \text{Path}(X,Z)$$

- ❑ Nota: X pode controlar o tráfego de entrada controlando as rotas que ele informa aos seus parceiros:
 - ex., se X não quer rotear tráfego para Z, X não informa nenhuma rota para Z

Internet inter-AS routing: BGP

- ❑ As mensagens do BGP são trocadas encapsuladas no TCP.
- ❑ mensagens BGP:
 - **OPEN**: inicia a conexão TCP com um roteador parceiro e autentica o transmissor
 - **UPDATE**: anuncia novo caminho (ou retira um velho)
 - **KEEPALIVE** mantém a conexão viva em caso de ausência de atualizações; também reconhece mensagens OPEN
 - **NOTIFICATION**: reporta erros nas mensagens anteriores; também usado para encerrar uma conexão

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

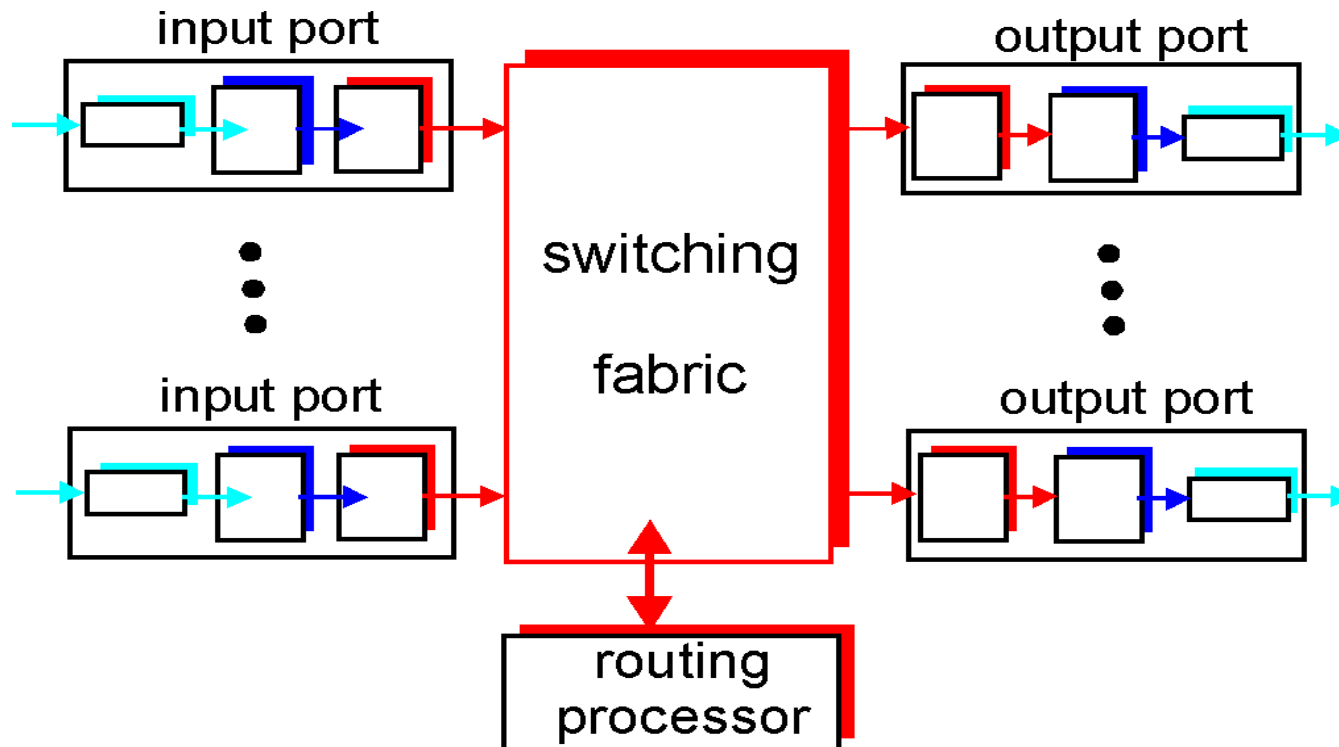
Performance:

- ❑ Intra-AS: preocupação maior é desempenho
- ❑ Inter-AS: regras de mercado podem ser mais importantes que desempenho

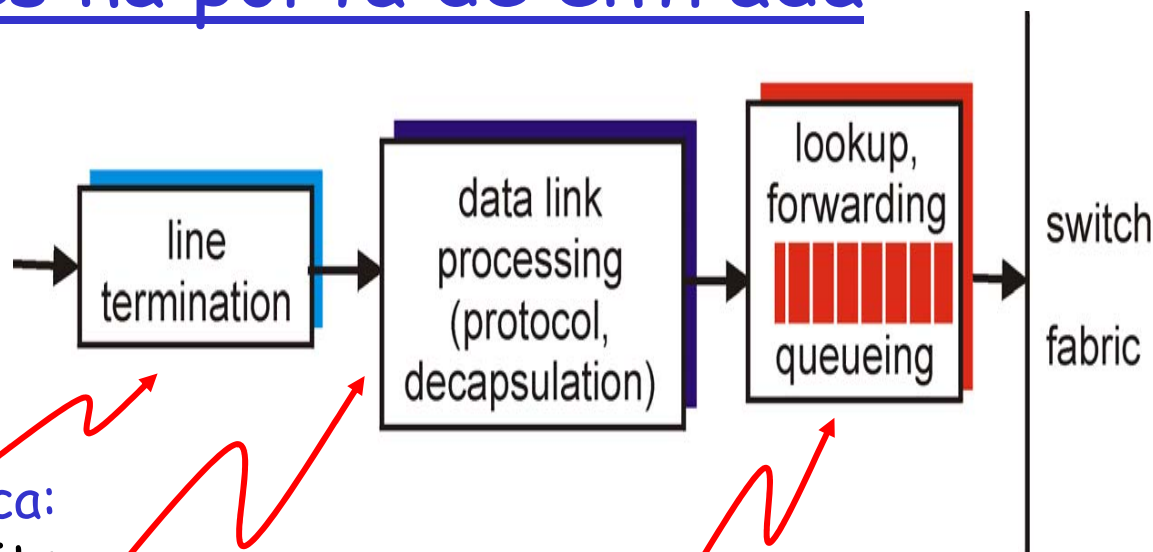
Visão da Arquitetura de Roteadores

Duas funções chave dos roteadores:

- ❑ rodar algoritmos e protocolos de roteamento (RIP, OSPF, BGP)
- ❑ *comutar* datagramas do enlace de entrada para o enlace de saída



Funções na porta de entrada



Camada física:
recepção de bits

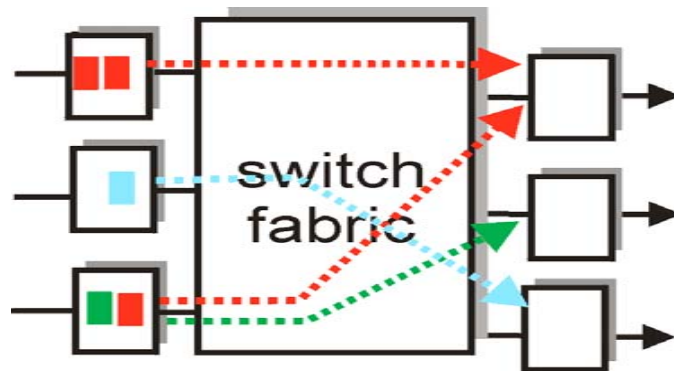
Camada de enlace:
ex., Ethernet
veja capítulo 5

Comutação descentralizada:

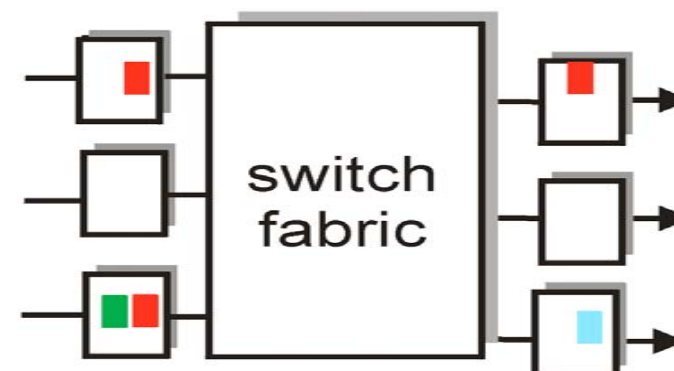
- ❑ dado o destino do datagrama, busca porta de saída, usando a tabela de roteamento na memória da porta de entrada
- ❑ objetivo: completar o processamento da porta de entrada na 'velocidade da linha'
- ❑ filas: se o datagramas chegam mais depressa que a taxa de envio para a estrutura de comutação

Enfileiramento na Porta de Entrada

- ❑ Se a estrutura de comutação for mais lenta que a capacidade combinada das portas de entrada -> pode ocorrer filas nas portas de entrada
- ❑ **Bloqueio Head-of-the-Line (HOL):** datagramas enfileirados no início da fila bloqueiam aqueles que estão atrás na fila
- ❑ *atrasos de filas e perdas são provocados pela saturação do buffer de entrada!*

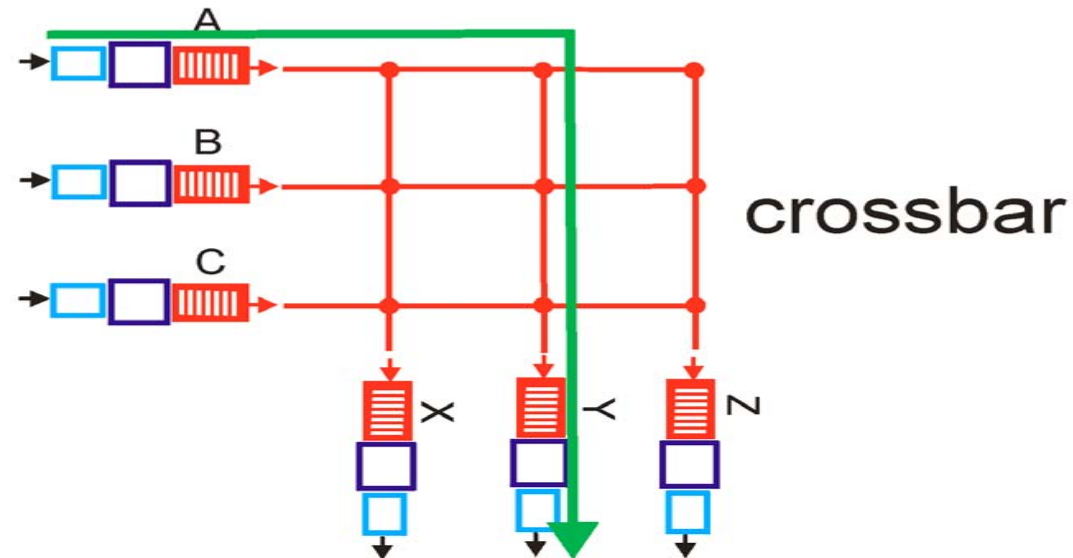
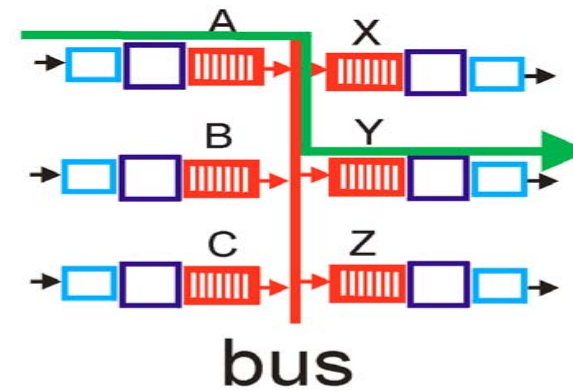
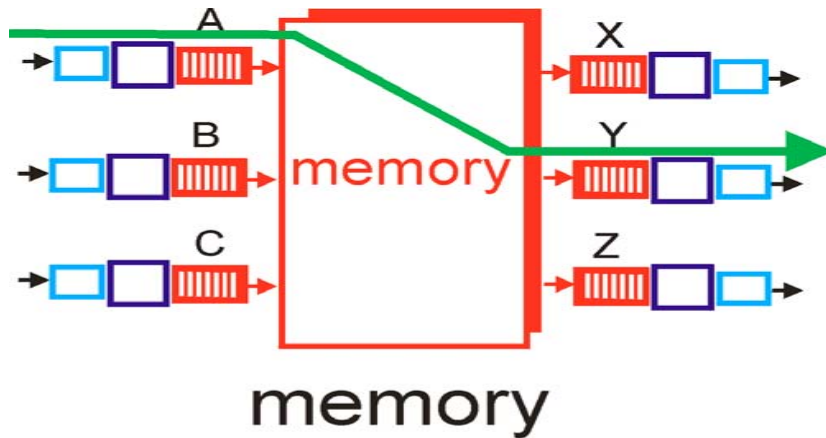


output port contention
at time t - only one red
packet can be transferred



green packet
experiences HOL blocking

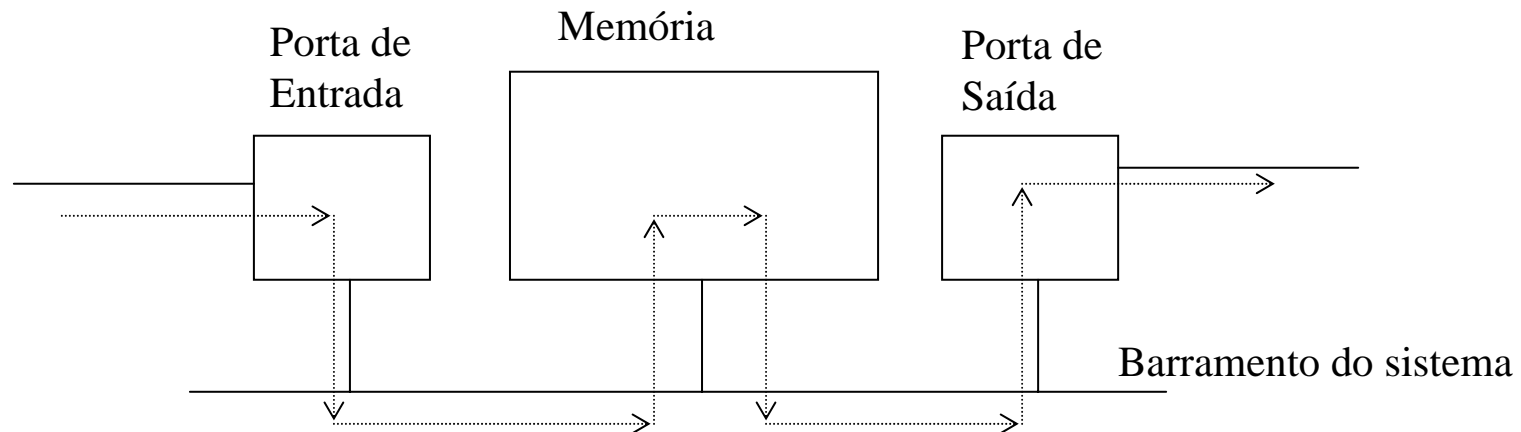
Três tipos de estruturas de comutação



Comutação via Memória

Empregada nos roteadores de primeira geração:

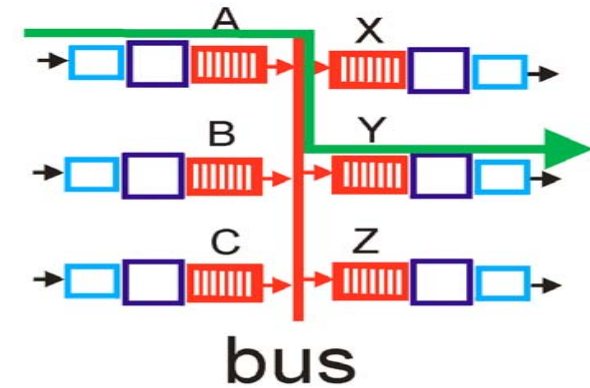
- ❑ pacotes são copiados pela única CPU do sistema
- ❑ velocidade é limitada pela banda passante da memória (2 cruzamentos do bus por datagrama)



Roteadores modernos:

- ❑ processador da porta de entrada realiza busca e cópia para a memória
- ❑ Cisco Catalyst 8500

Comutação Via Bus

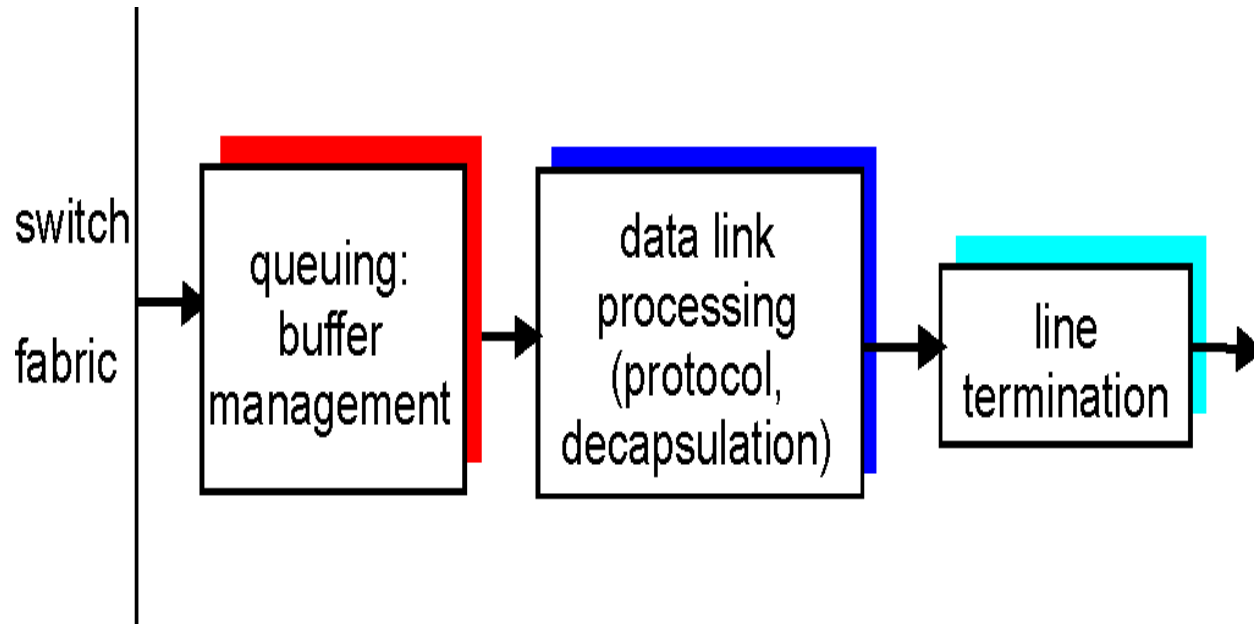


- ❑ datagrama é transferido da memória da porta de entrada para a memória da porta de saída via um barramento compartilhado
- ❑ **contenção no bus:** velocidade de comutação limitada pela capacidade do barramento
- ❑ 1 Gbps bus, Cisco 1900: velocidade suficiente para roteadores de acesso e de empresas (não para roteadores regionais e de backbone)

Comutação via Rede de Interconexão

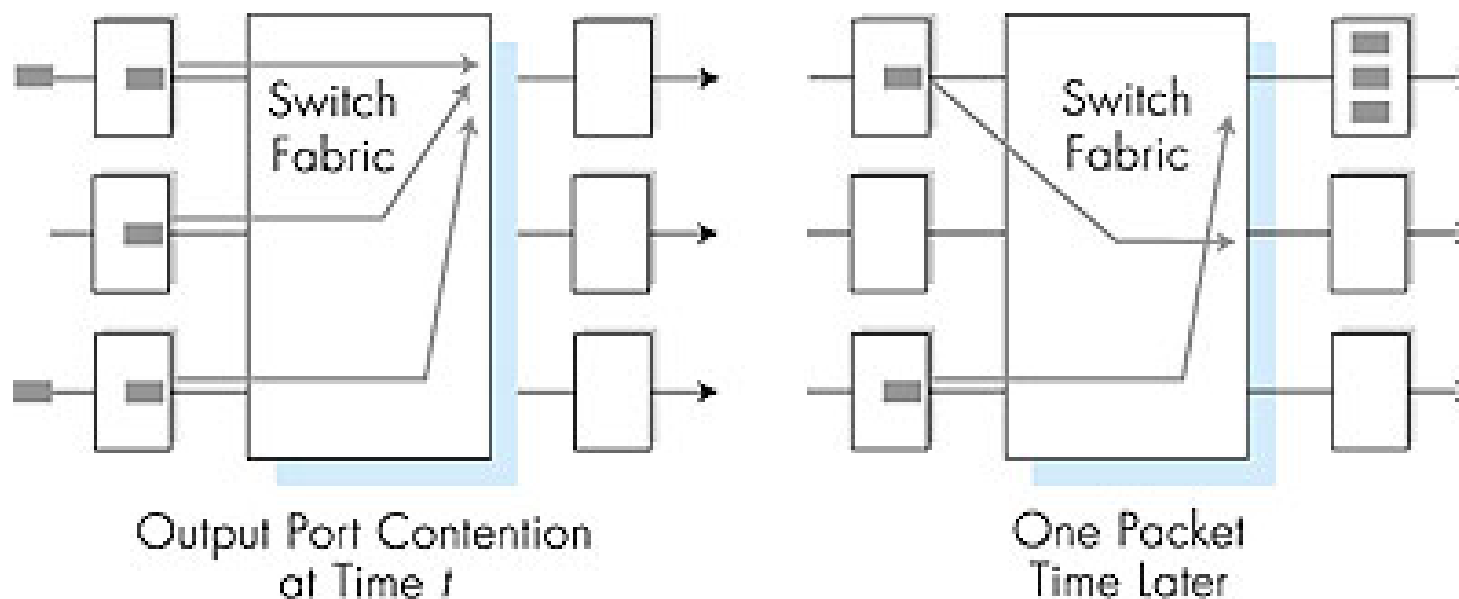
- ❑ supera limitações da banda do barramento
- ❑ redes de Banyan, outras redes de interconexão originalmente desenvolvidas para conectar processadores num sistema multi-processador
- ❑ projeto avançado: fragmentar datagramas em células de comprimento fixo e comutar as células por uma rede de comutação.
- ❑ Cisco 12000: comuta vários gigabis por segundo através de uma rede de interconexão

Portas de Saída



- ❑ *Armazenamento:* exigido quando os datagramas chegam da estrutura de comutação mais depressa que a taxa de transmissão do enlace de saída
- ❑ *Disciplina de fila:* escolhe entre os datagramas enfileirados um deles para transmissão

Filas na porta de saída



- armazenamento quando a taxa de chegada pelo comutador excede a velocidade da linha de saída
- *filas(atrasos) e perdas são provocados por um overflow do buffer da porta de saída!*

IPv6

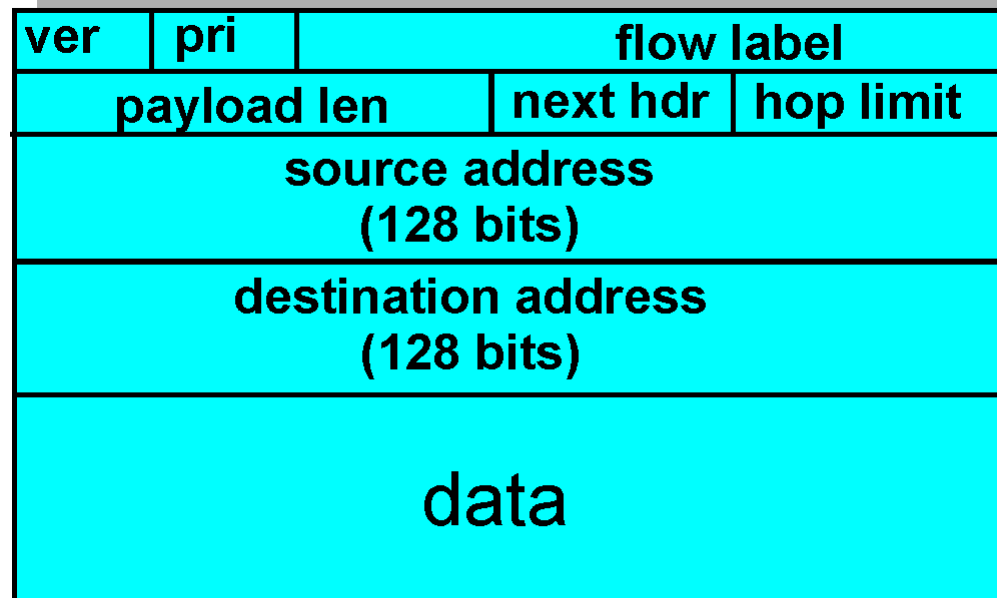
- ❑ **Motivação inicial:** o espaço de endereços de 32-bits estará completamente alocado por volta de 2008.
- ❑ **Motivação adicional:**
 - melhorar o formato do header para permitir maior velocidade de processamento e de transmissão
 - mudanças no header para incorporar mecanismos de controle de QOS
 - novo tipo de endereço: "anycast" - permite enviar uma mensagem para o melhor dentre vários servidores replicados
 - **IPv6 formato dos datagramas:**
 - cabeçalho fixo de 40 bytes
 - não é permitida fragmentação

IPv6 Header (Cont)

Priority: permitir definir prioridades diferenciadas para vários fluxos de informação

Flow Label: identifica datagramas do mesmo "fluxo."
(conceito de "fluxo" não é bem definido).

Next header: identifica o protocolo da camada superior ou um header auxiliar



← 32 bits →

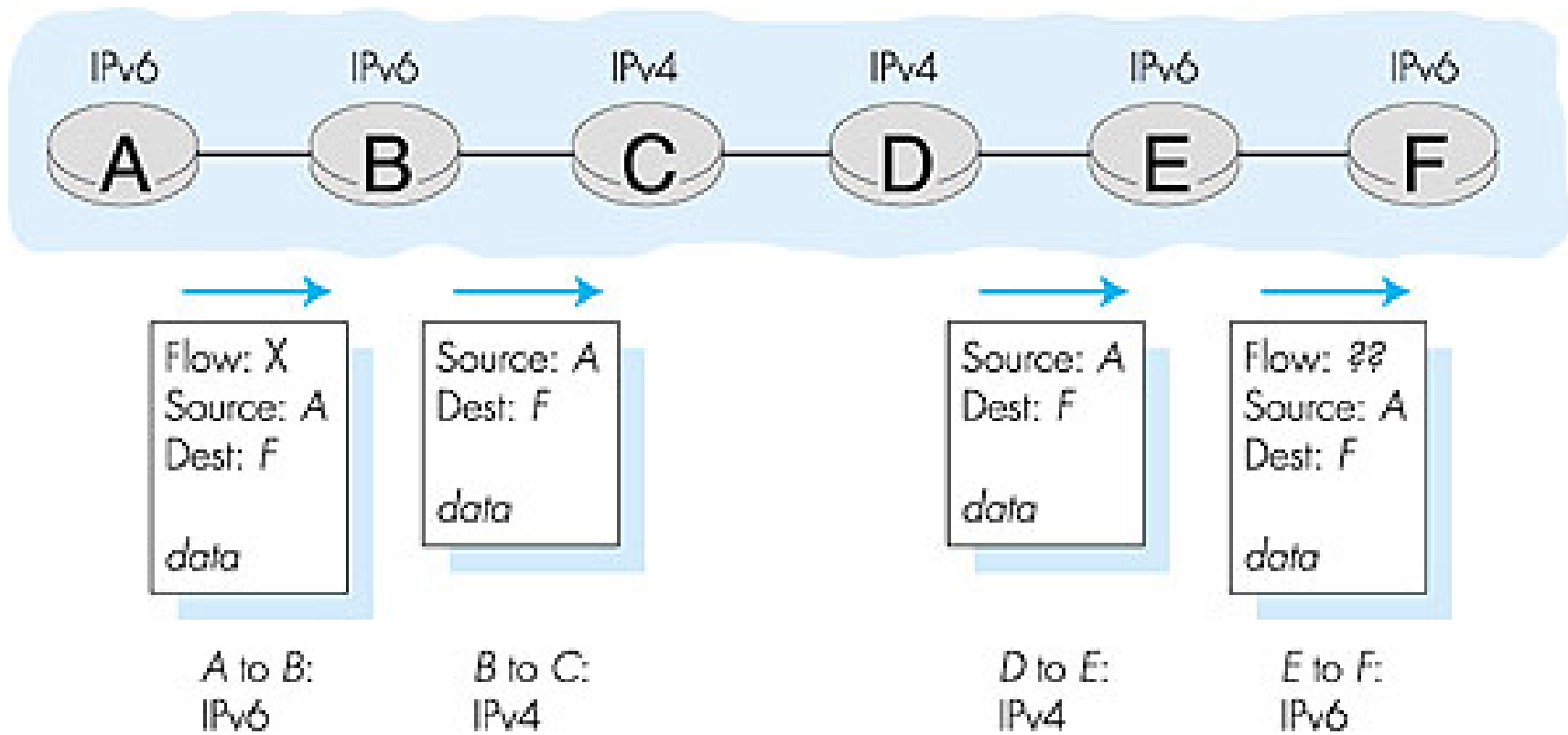
Outras mudanças do IPv4

- ❑ *Checksum*: removido inteiramente para reduzir o tempo de processamento em cada hop
- ❑ *Options*: são permitidas, mas são alocadas em cabeçalhos suplementares, indicados pelo campo "Next Header"
- ❑ *ICMPv6*: nova versão de ICMP
 - tipos de mensagens adicionais , ex. "Packet Too Big"
 - funções de gerenciamento de grupos multicast

Transição do IPv4 para IPv6

- ❑ Nem todos os roteadores poderão ser atualizados simultaneamente
 - não haverá um dia da vacinação universal
 - A rede deverá operar com os dois tipos de datagramas simultaneamente presentes
 - Duas abordagens propostas:
 - *Dual Stack*: alguns roteadores com pilhas de protocolos duais (v6, v4) podem trocar pacotes nos dois formatos e traduzir de um formato para o outro
 - *Tunneling*: IPv6 transportado dentro de pacotes IPv4 entre roteadores IPv4

Dual Stack Approach



Tunneling

Logical view



Physical view

