

Camada de Rede

Objetivos:

Entender os princípios por trás dos serviços de rede:

- modelos de serviço da camada de rede
- repasse vs roteamento
- funcionamento de um roteador
- endereçamento IPv4
- algoritmos de roteamento
- arquitetura da Internet
- instanciação e implementação na Internet
 - Protocolos OSPF, BGP
- IPv6
- NAT
- ICMP

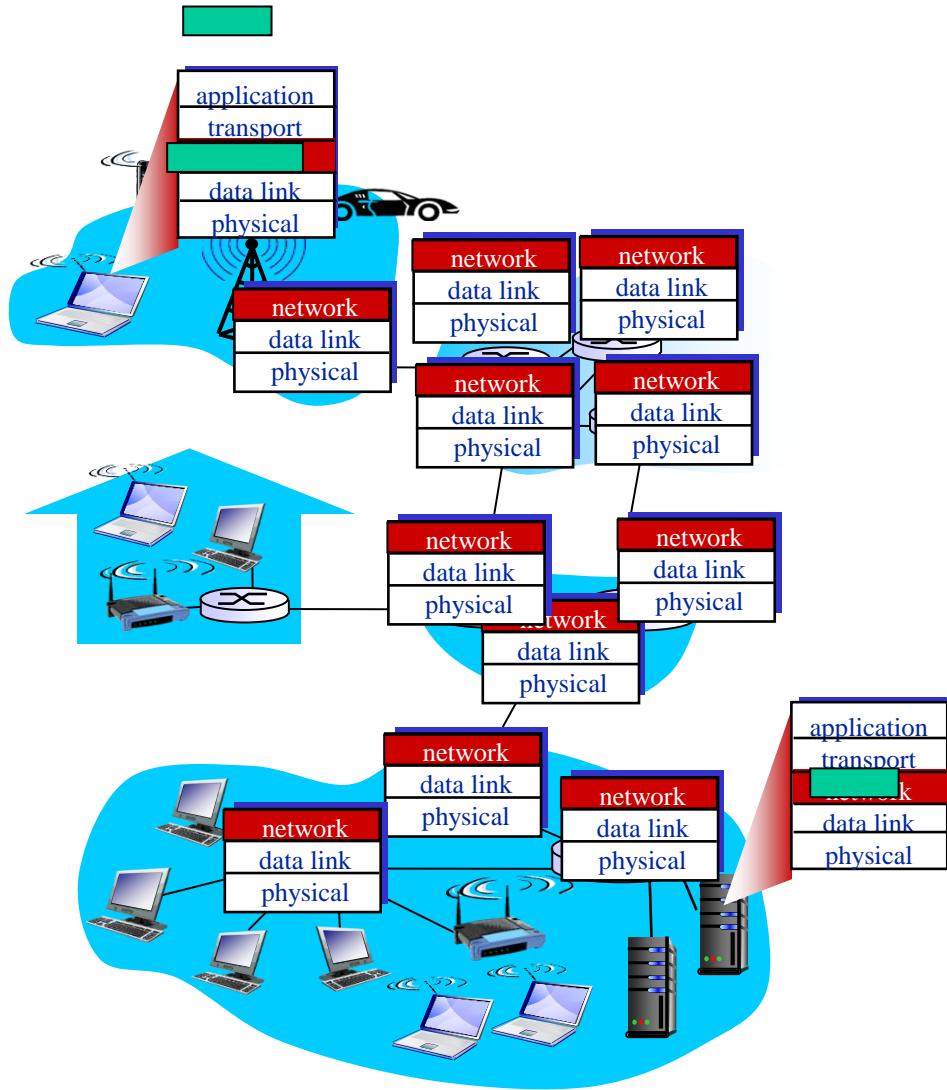
Serviços oferecidos pela Camada de Rede

Requisitos fundamentais:

- Independentes da tecnologia da camada subjacente
- Isolar a camada de transporte das questões relativas ao tamanho, tipo e topologia das redes locais
- Usar um esquema de endereçamento uniforme

Camada de rede: serviços e protocolos

- Transporta segmentos do hospedeiro transmissor para o receptor
- No lado transmissor encapsula os segmentos em datagramas e os passa para a camada de enlace
- No lado receptor, entrega os segmentos à camada de transporte
- Protocolos da camada de rede estão em **todos** os hospedeiros e roteadores
- Roteadores:
 - Examinam campos de cabeçalho em todos os datagramas IP que passam por ele
 - Movem os datagramas das portas de entrada para as portas de saída do caminho fim-a-fim



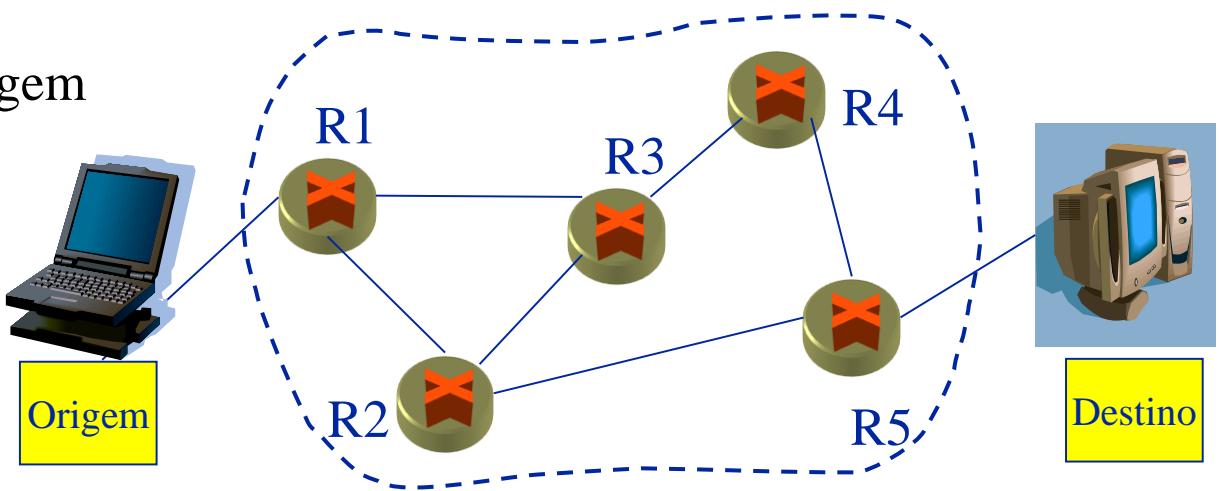
Funções-chave da camada de rede

- **Roteamento:** determinar rotas para a transmissão de pacotes da origem ao destino (fim-a-fim)
→ Algoritmos de roteamento
- **Repasso (ou comutação):** mover pacotes dentro do roteador, da porta de entrada para a porta de saída apropriada

Analogia:

✓ **Roteamento:**

processo de planejar a viagem
da origem ao destino



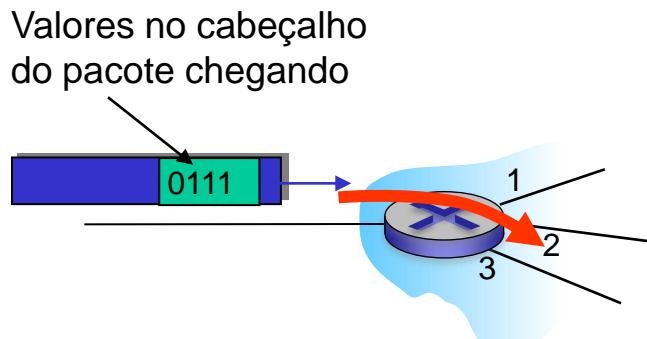
✓ **Comutação:**

processo de passar por
uma única conexão

Funções-chave da Camada de Rede

O Repasse é uma função local, por roteador

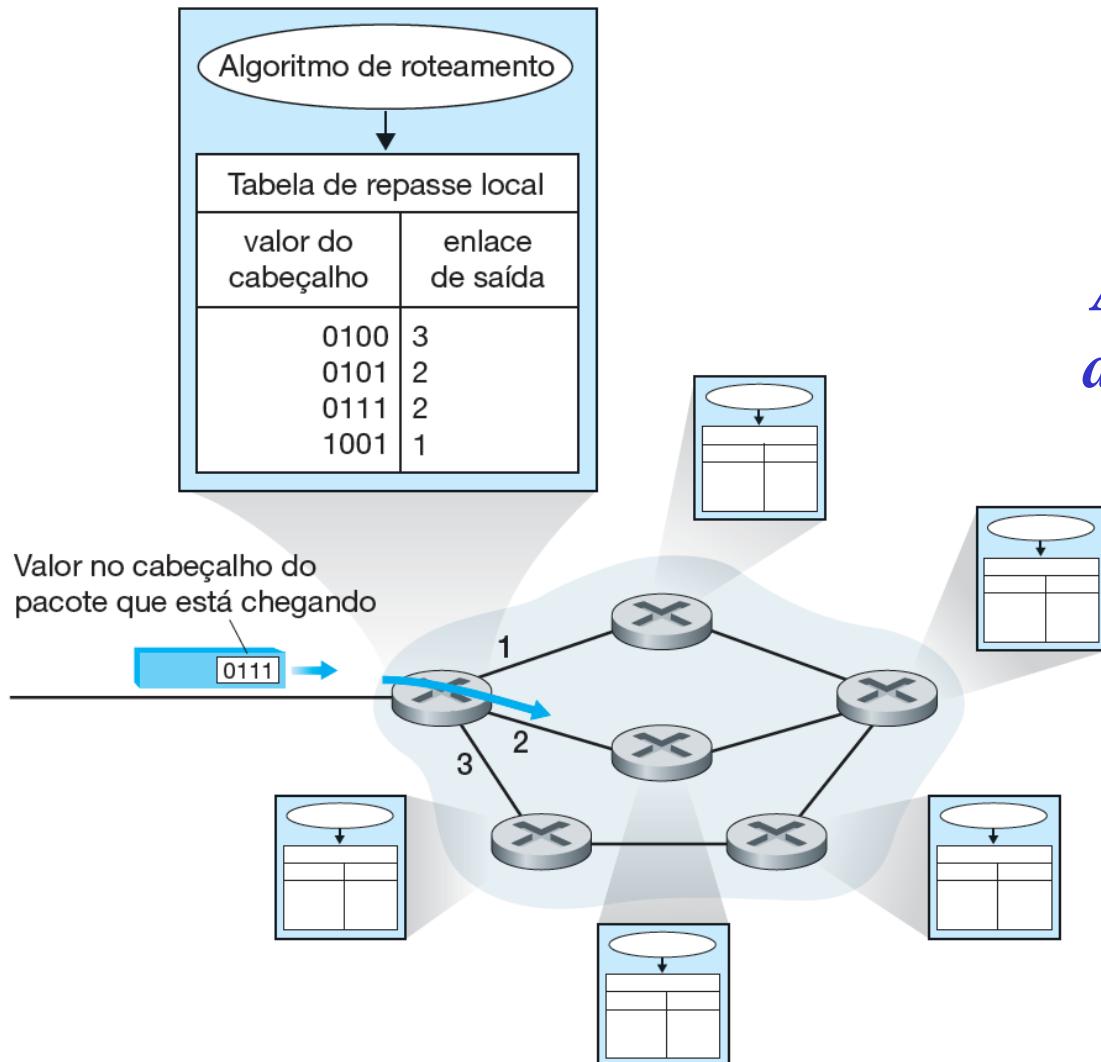
- Determina como um datagrama chegando por uma porta de entrada é repassado para uma porta de saída analisando o cabeçalho



O Roteamento considera uma lógica global da rede

- Determina como datagramas são encaminhados do host fonte até o host de destino (caminho fim-a-fim) através dos roteadores intermediários
- Existem duas abordagens:
 - *Algoritmos de roteamento tradicionais*: rotas calculadas e implementadas nos roteadores
 - *Software-Defined Networking (SDN)*: rotas calculadas nos servidores (remotos) e “descarregadas” nas tabelas de repasse dos roteadores

Algoritmos de roteamento determinam os valores nas tabelas de repasse



Modelo de serviço da camada de rede

O modelo de serviço define as características do transporte de datagramas do transmissor ao receptor

Exemplo de serviços para datagramas individuais:

- Garantia de distribuição
- Garantia de distribuição com menos de 40 mseg de atraso

Exemplo de serviços para um fluxo de datagramas:

- Distribuição em ordem dos datagramas
- Garantia de banda mínima para o fluxo
- Jitter máximo garantido

Modelo de serviço da camada de rede

Arquitetura re rede	Modelo de serviço	Garantias de Qualidade de Serviço ?			
		Largura de banda	Perda	Ordem	Atraso máx
Internet	melhor esforço	nenhuma	não	não	não

Modelo de serviço de “melhor esforço” da Internet

Sem garantias de:

- i. entrega do datagrama ao destinatário
- ii. entrega em ordem e com um atraso máximo
- iii. garantia de banda para o fluxo fim-a-fim

Modelo de serviço da camada de rede

Arquitetura da rede	Modelo de serviço	Garantias de Quality of Service (QoS)?			
		Largura de banda	Perda	Ordem	Atraso máx
Internet	Melhor esforço	nenhuma	não	não	não
ATM	Constant Bit Rate	taxa constante	sim	sim	sim
ATM	Available Bit Rate	garantia min	não	sim	não
Internet	Intserv Guaranteed (RFC 1633)	sim	sim	sim	sim
Internet	Diffserv (RFC 2475)	possível	possível	possível	não

Reflexões sobre o serviço de melhor esforço

- *Simplicidade do mecanismo* permitiu que a Internet fosse amplamente adotada e implantada
- *Provisionamento suficiente de largura de banda* permite que o desempenho de aplicativos em tempo real (por exemplo, voz interativa e vídeo) seja “bom o suficiente” “na maior parte do tempo”
- *Serviços distribuídos e replicados da camada de aplicação (datacenters, content distribution networks)* permite que os serviços sejam fornecidos a partir de vários locais, próximos às redes dos clientes
- Controle de congestionamento de serviços “elásticos” também ajuda

É difícil argumentar com o sucesso do modelo de serviço de melhor esforço

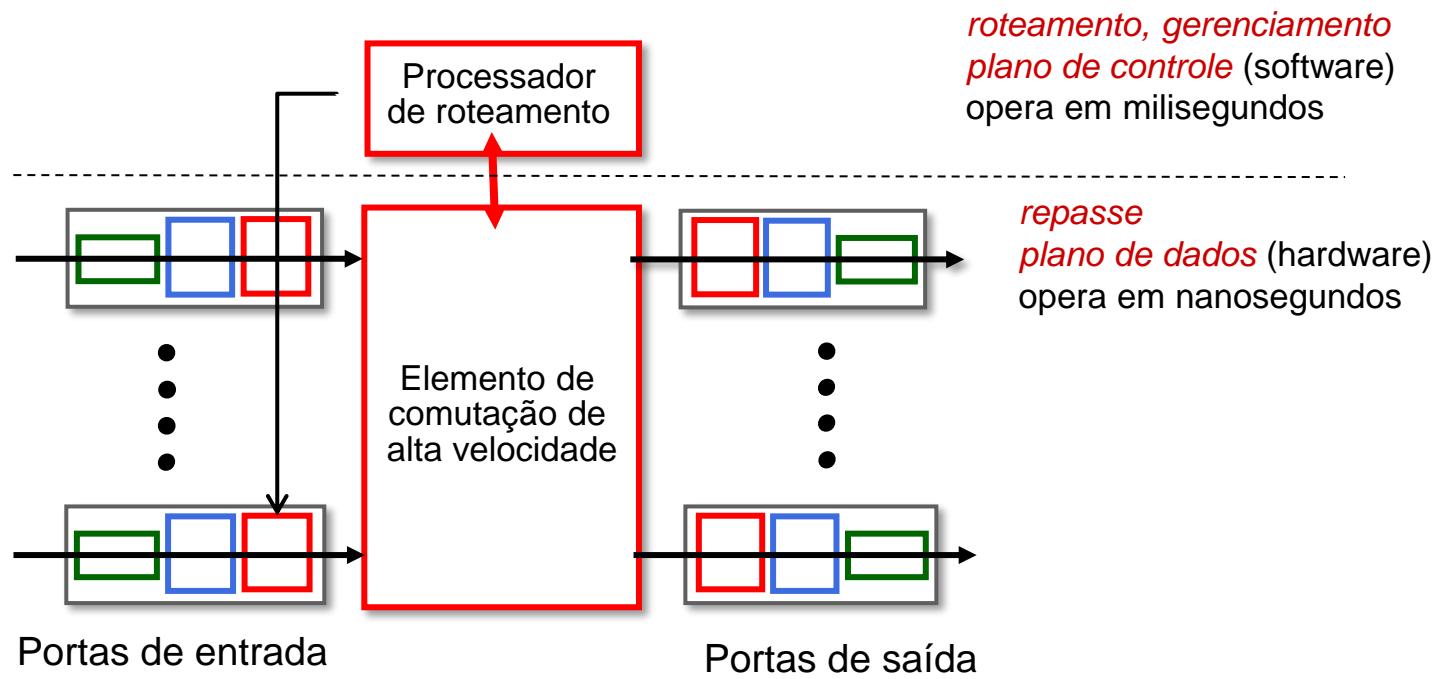
Visão da arquitetura de roteadores

Duas funções chave dos roteadores:

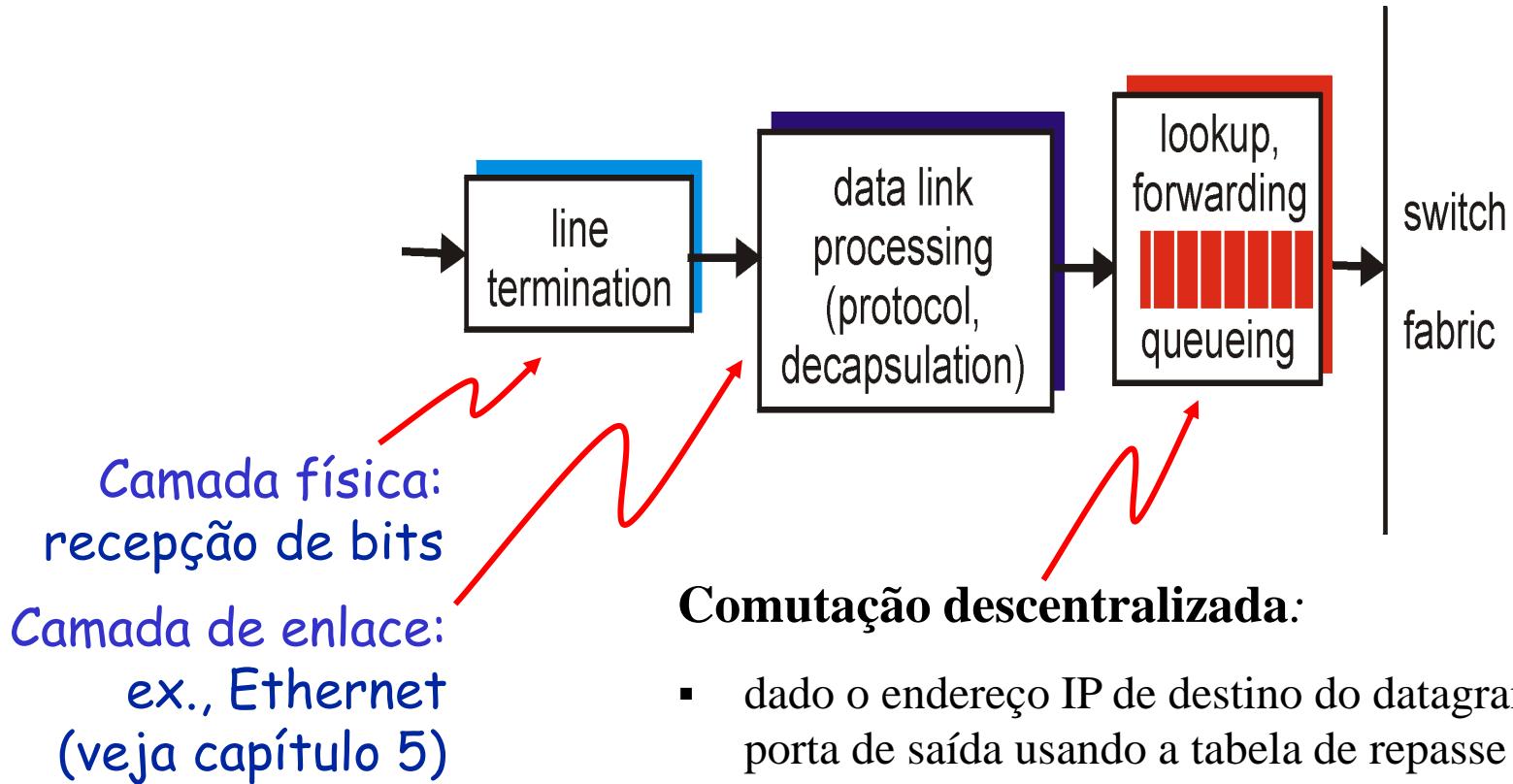
- *rodar algoritmos e protocolos de roteamento (RIP, OSPF, BGP)*
- *repassar datagramas do enlace de entrada para o enlace de saída*

Visão da arquitetura de roteadores

Arquitetura de um roteador:

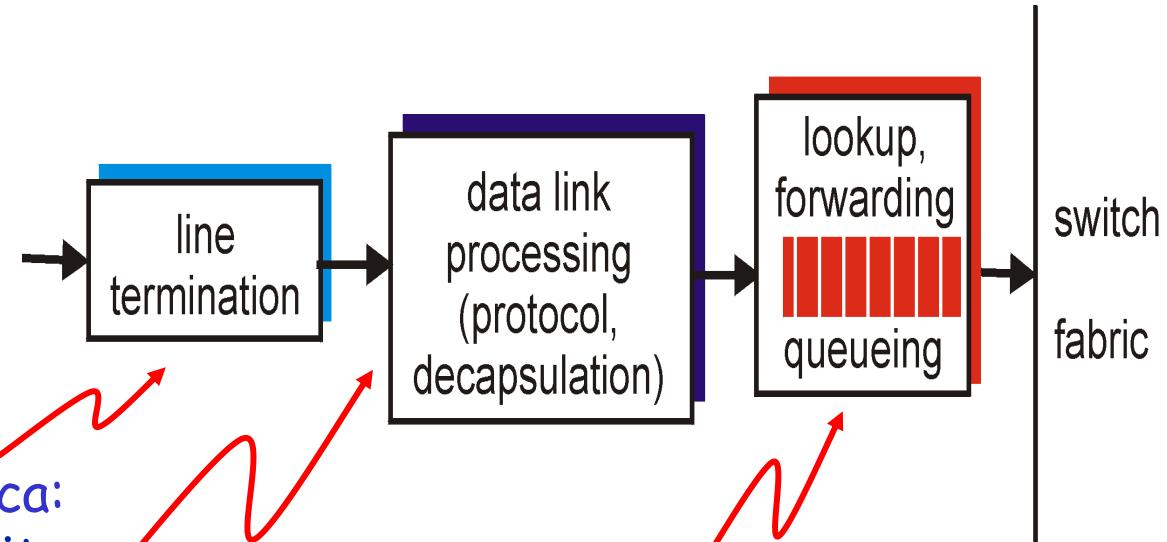


Funções na porta de entrada



- dado o endereço IP de destino do datagrama, busca porta de saída usando a tabela de repasse na memória da porta de entrada
- objetivo: completar o processamento da porta de entrada na ‘velocidade da linha’ (velocidade de recebimento de pacotes)
- filas: ocorrem se o datagramas chegam mais rapidamente que a taxa de repasse da estrutura de comutação

Funções na porta de entrada



Camada física:
recepção de bits

Camada de enlace:
ex., Ethernet
veja capítulo 5

Comutação descentralizada:

- cópia da tabela de repasse calculada pelo processador de roteamento
- em roteadores de backbone a porta de entrada tem que operar à velocidade da linha
- milhões de consultas por segundo → inviável realizar uma busca linear → árvores

Tabela de repasse para um datagrama

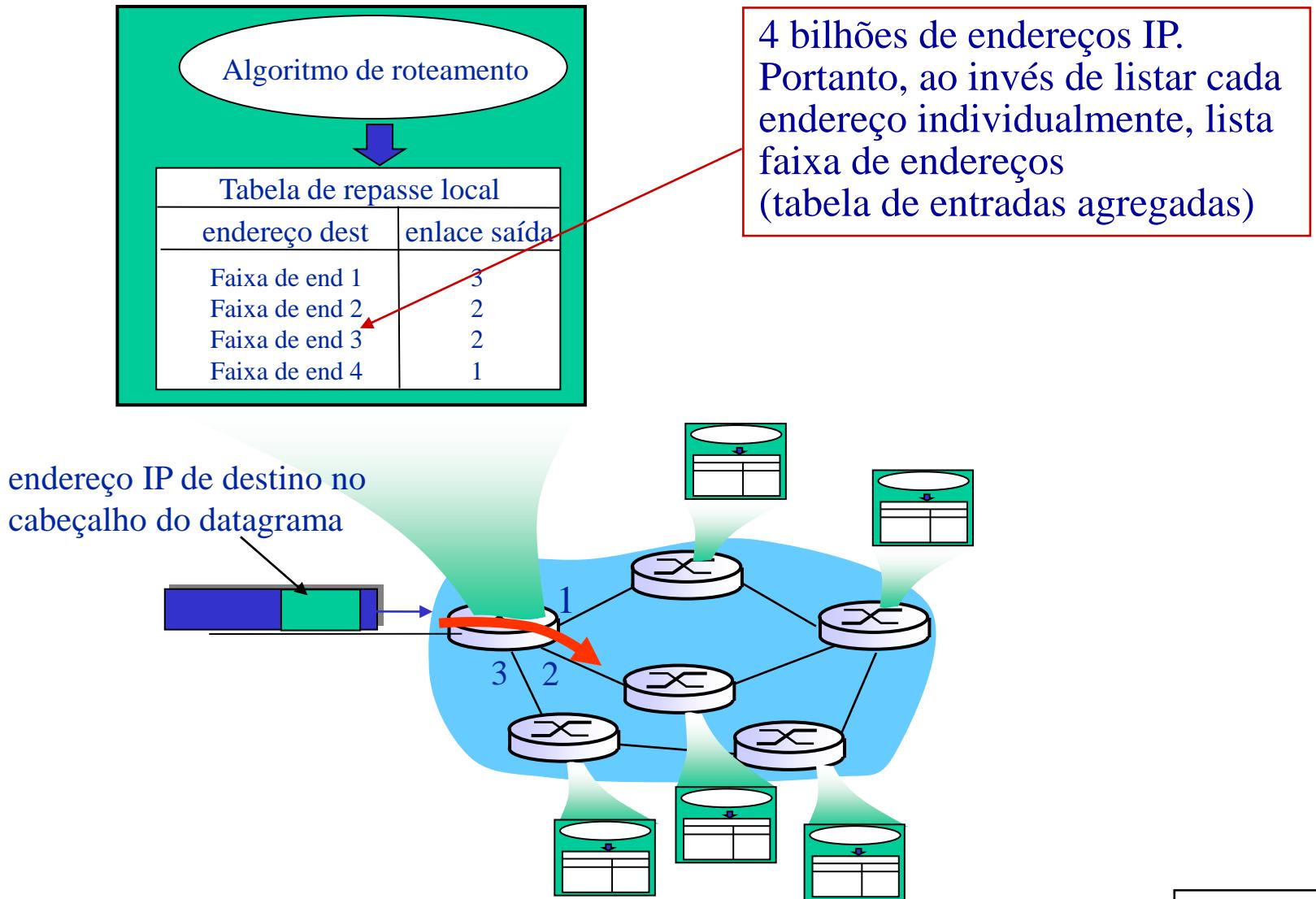


Tabela de repasse

4 bilhões de entradas possíveis

<u>Faixa de endereços de destino</u>	<u>Interface</u>
11001000 00010111 00010000 00000000 até 11001000 00010111 00010111 11111111	0
11001000 00010111 00011000 00000000 até 11001000 00010111 00011000 11111111	1
11001000 00010111 00011001 00000000 até 11001000 00010111 00011111 11111111	2
caso contrário	3

Coincidência de prefixo mais longo

Prefixo mais longo

Quando procurando por uma entrada na tabela de repasse para um determinado endereço de destino, use o prefixo de endereço mais longo que “casa” com o endereço de destino.

Destination Address Range	Link interface
11001000 00010111 00010***	0
11001000 00010111 000110000 *	1
11001000 1 00011**	2
otherwise	3

examples:

11001000 00010111 00010110 10100001 which interface?
11001000 00010111 00011000 10101010 which interface?

Coincidência de prefixo mais longo

Prefixo mais longo

Quando procurando por uma entrada na tabela de repasse para um determinado endereço de destino, use o prefixo de endereço mais longo que “casa” com o endereço de destino.

Destination Address Range	Link interface
11001000 00010111 00010** *****	0
11001000 00010111 00011000 *****	1
11001000 00010111 00011*** *****	2
otherwise *	3

↑
match!

11001000 00010111 00010110 10100001 which interface?
11001000 00010111 00011000 10101010 which interface?

examples:

Coincidência de prefixo mais longo

Prefixo mais longo

Quando procurando por uma entrada na tabela de repasse para um determinado endereço de destino, use o prefixo de endereço mais longo que “casa” com o endereço de destino.

Destination Address Range	Link interface
11001000 00010111 00010** *****	0
11001000 00010111 00011000 *****	1
11001000 00010111 00011** *****	2
otherwise *	3

examples:

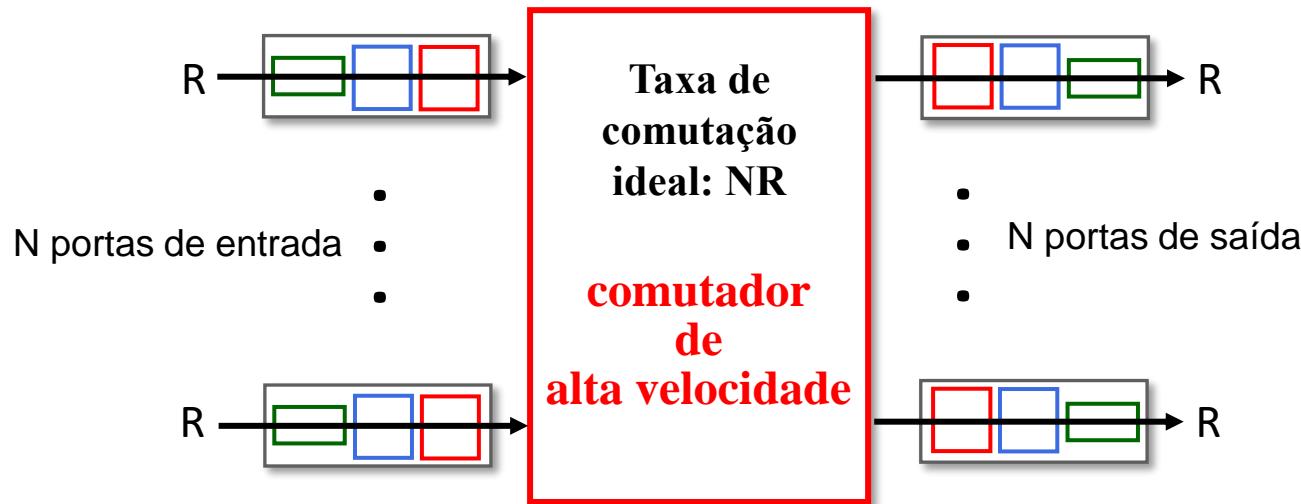
11001000 00010111 00010110	10100001	which interface?
11001000 00010111 00011000	10101010	which interface?

Elemento de comutação

Função: transferir pacotes do enlace de entrada p/ o enlace de saída apropriado

Taxa de comutação: taxa na qual os pacotes podem ser transferidos das entradas p/ as saídas

- Normalmente medida como múltipla da taxa da linha de entrada/saída
- N entradas: desejável uma taxa de comutação N vezes a taxa da linha



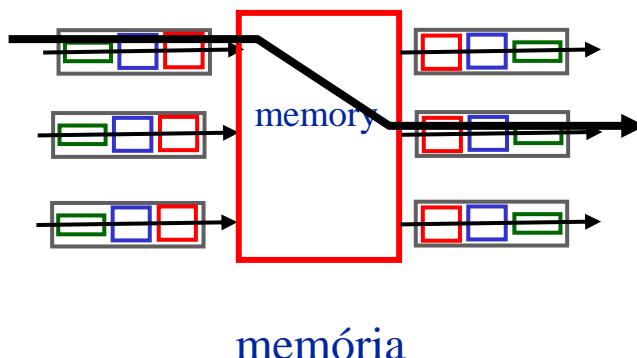
Elemento de comutação

Função: transferir pacotes do enlace de entrada p/ o enlace de saída apropriado

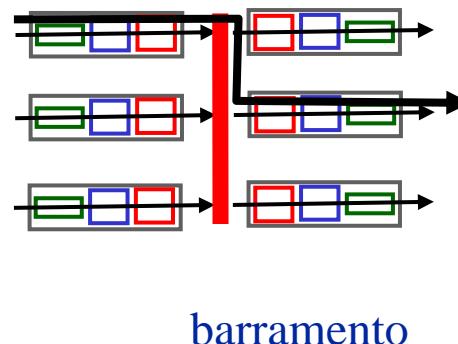
Taxa de comutação: taxa na qual os pacotes podem ser transferidos das entradas p/ as saídas

- Normalmente medida como múltipla da taxa da linha de entrada/saída
- N entradas: desejável uma taxa de comutação N vezes a taxa da linha

Três tipos de elementos de comutação:

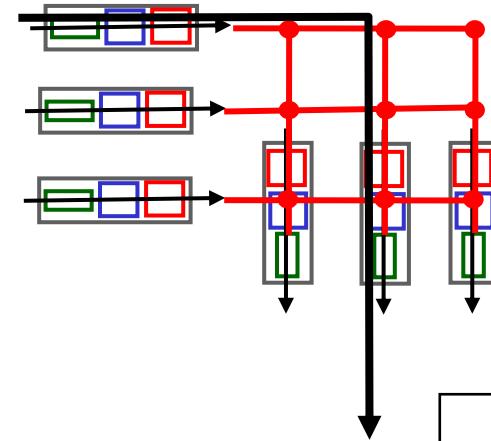


memória



barramento

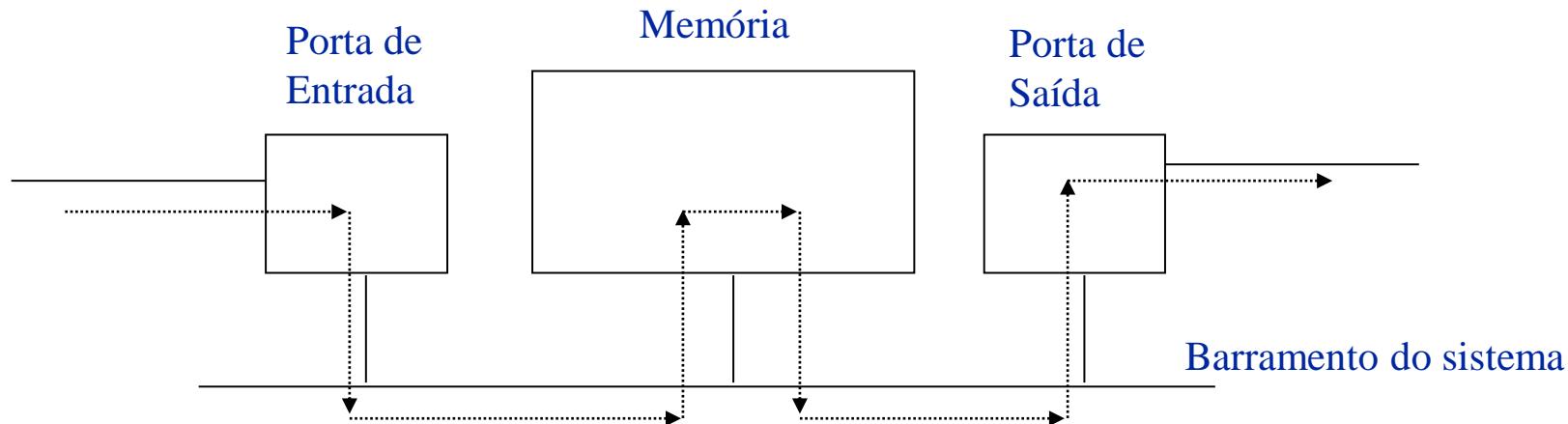
rede de interconexão



Comutação via memória

Empregada nos roteadores de primeira geração:

- computadores tradicionais com comutação sob controle direto da CPU
- pacotes são copiados pela CPU para a memória do sistema
- velocidade é limitada pela largura de banda da memória (2 cruzamentos do barramento por datagrama). Vazão de repasse máxima $B/2$.

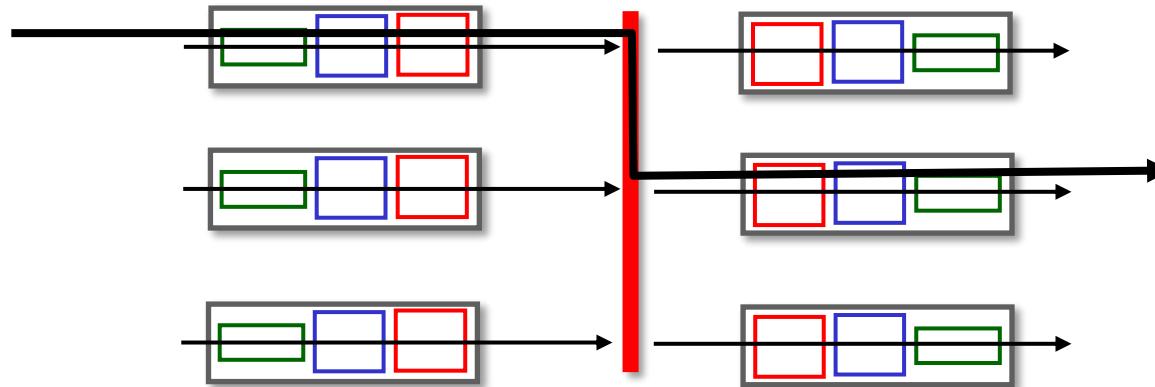


Roteadores modernos:

Processador localizado na porta de entrada realiza busca e cópia para a memória. Ex: Cisco Catalyst 8500

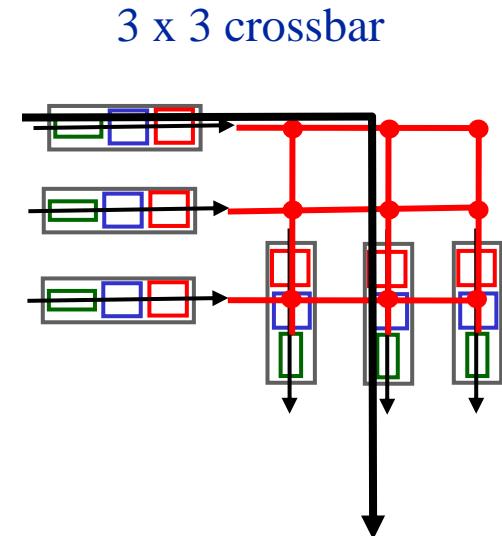
Comutação via barramento

- datagrama é transferido da memória da porta de entrada para a memória da porta de saída via um barramento compartilhado
- somente um pacote por vez é transferido
- ***concorrência no barramento:*** velocidade de comutação é limitada pela capacidade do barramento
- ***Cisco 5600:*** 32 Gbps - velocidade suficiente para roteadores de acesso e de empresas



Comutação via rede de interconexão

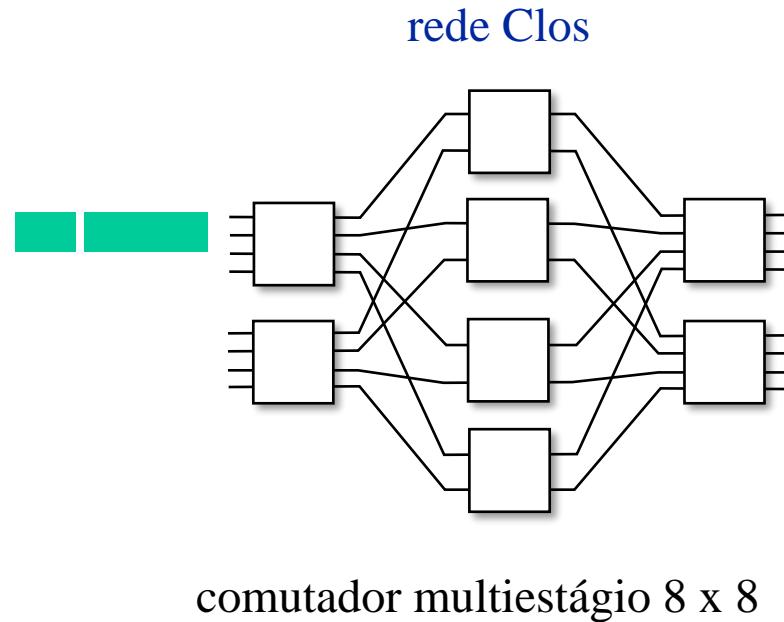
- supera limitações da banda do barramento:
rede de interconexão com **2n** barramentos,
conectando **n** portas de entrada com n portas
de saída
- inicialmente desenvolvido para conectar
processadores em sistemas multiprocessados
- **Cisco 12000:** comuta até 60 Gbps através de
uma rede de interconexão



Comutação multiestágio

Comutador multiestágio:

comutador $n \times n$ construído a partir de múltiplos estágios de comutadores menores



Tendência atual: explorar paralelismo, fragmentando os datagramas em células de comprimento fixo, marcando-as e comutando essas células por uma rede de interconexão ou multiestágio, remontando-as na porta de saída

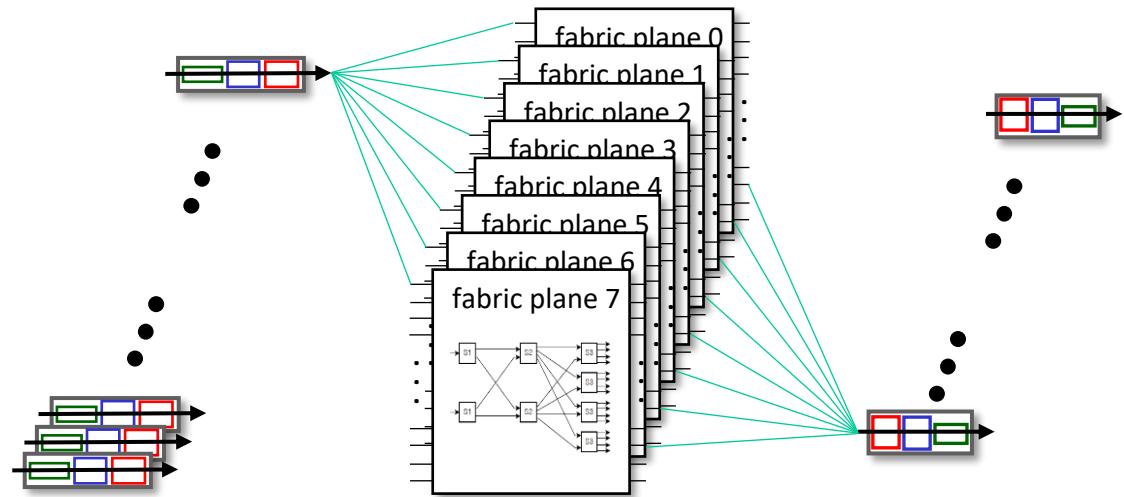
Paralelismo na comutação

Usando vários “planos” de comutação em paralelo:

- Acelerando a comutação via paralelismo

Cisco CRS:

- unidade básica: 8 planos de comutação
- cada plano: rede de interconexão de 3 estágios
- comuta até 100's Tbps



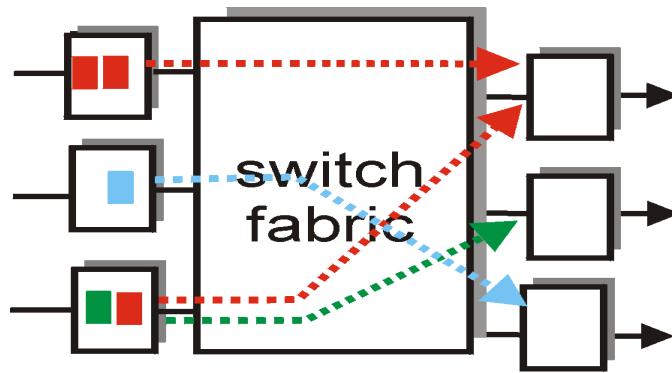
Filas: onde ocorrem?

- Tanto na porta de entrada quanto na porta de saída
- Local da perda e extensão da fila depende do tráfego, da velocidade do elemento de comutação e das taxas dos enlaces
- *Se a taxa do elemento de comutação for n vezes a taxa da linha de entrada (n entradas) não haverá filas na entrada*

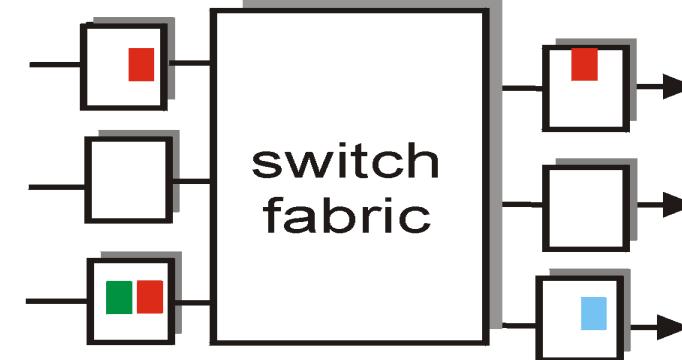
Ocorre fila na porta de entrada quando a taxa de comutação não é suficientemente grande para atender todas as portas de entrada ao mesmo tempo

Filas 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

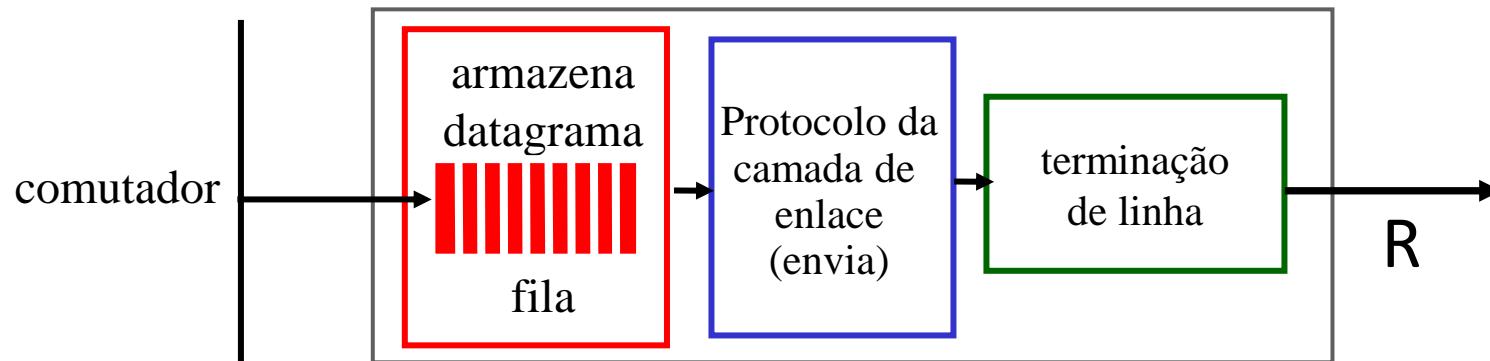


green packet
experiences HOL blocking

Processamento na porta de saída

Processamento na porta de saída:

Toma pacotes armazenados na fila e os transmite pelo enlace de saída



Filas na porta de saída

- ***Processamento na porta de saída***

- ***Armazenamento:***

Necessário quando os datagramas chegam da estrutura de comutação mais rapidamente do que a taxa de transmissão do enlace de saída (podem ocorrer perdas devido ao congestionamento ou buffer insuficiente)

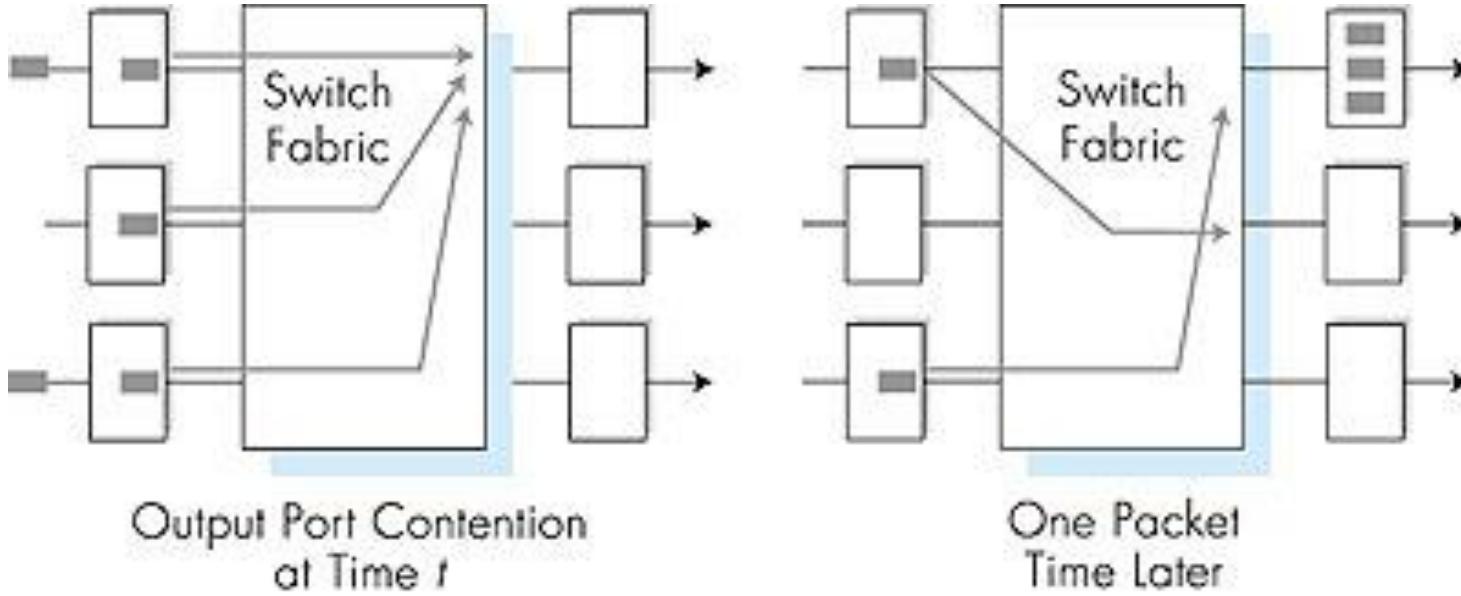
- ***Política de descarte***

Qual datagrama descartar se o buffer está cheio?

- ***Escalonamento prioritário:***

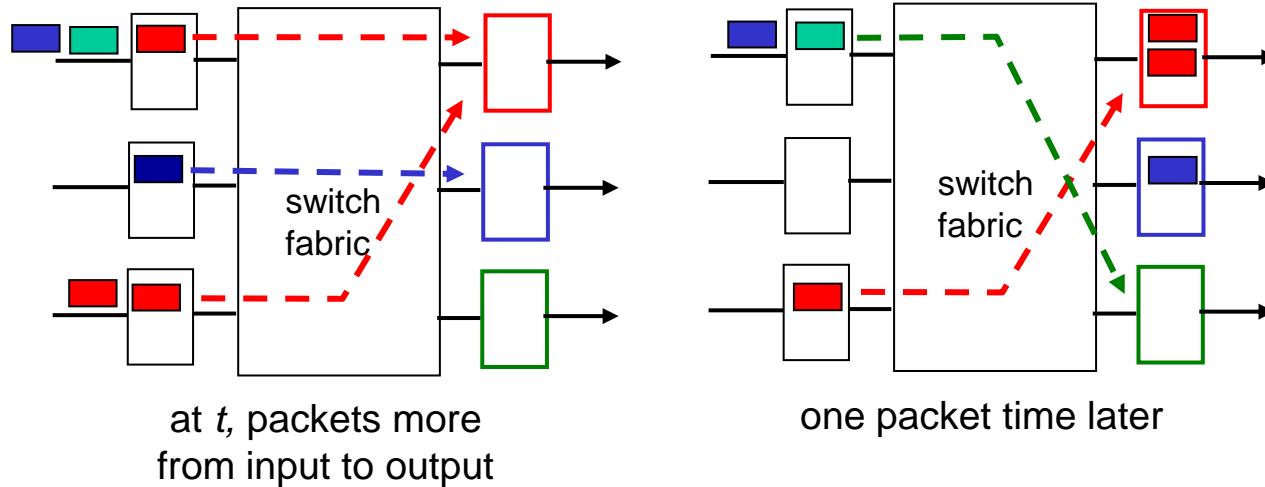
Escolhe entre os datagramas enfileirados qual deles tem prioridade para ser transmitido de forma a manter o desempenho de determinada aplicação (QoS)

Filas na porta de saída



- ❑ Pior caso: os n pcts que chegam são transferidos para a mesma porta de saída
- ❑ *Número de pcts pode crescer até o ponto de exaurir o espaço de memória na porta de saída \Rightarrow aumento*

Filas na porta de saída



- Armazenamento é necessário quando a taxa de comutação excede a velocidade do enlace de saída
- Necessidade de um escalonador de pcts: FCFS? WFQ? Implica na qualidade de serviço (QoS) oferecida
- Necessidade de política de descarte (gerenciamento ativo de fila)

A fila de transmissão causa atraso e perda quando o buffer da porta de saída se esgota!

Qual o tamanho do buffer?

Regra prática (RFC 3439): tamanho médio do buffer igual ao RTT "típico" (digamos 250 mseg) vezes a capacidade do enlace (C), para um fluxo. Por exemplo, C = 10 Gbps: buffer de 2,5 Gbit

Recomendação mais recente: com N fluxos, buffer igual a

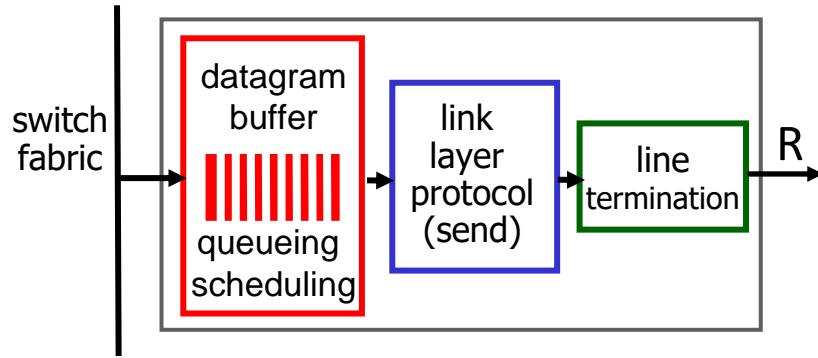
$$\frac{\text{RTT } C}{\sqrt{N}}$$

Qual o tamanho do buffer?

- excesso de buffer pode aumentar os atrasos (particularmente em roteadores de redes de acesso)
 - RTTs longos:
 - baixo desempenho para aplicativos em tempo real
 - resposta lenta do TCP
 - lembre-se do controle de congestionamento baseado no atraso:

“mantenha o enlace de gargalo cheio o suficiente (ocupado), mas não mais cheio do que o suportável”

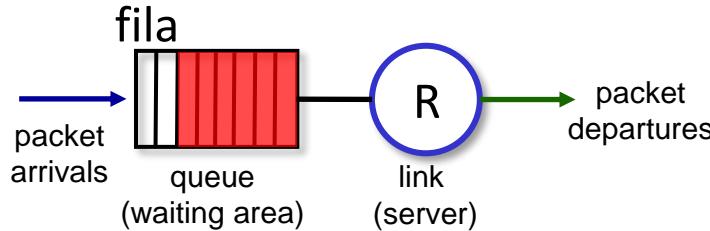
Gerenciamento do Buffer



descarte: qual pacote descartar quando os buffers estiverem cheios?

- descarta pacote chegando
- ***prioridade:*** descarta baseado na em alguma regra de prioridade

Abstração:



marcação: quais pacotes marcar para sinalizar o congestionamento (ECN, RED)

Escalonamento de pacotes: FCFS

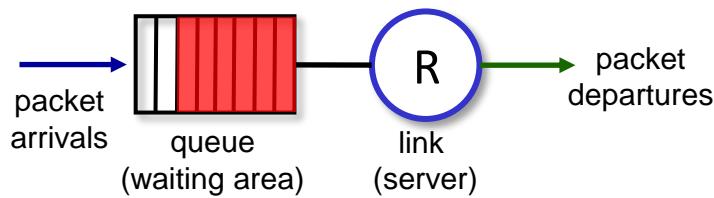
Escalonamento de pcts:

decidir qual pacote enviar em seguida no enlace:

- first come, first served (FCFS)
- prioridade
- round robin
- weighted fair queueing (WFQ)

FCFS: pacotes transmitidos em ordem de chegada à porta de saída (também conhecido como: First-in-first-out (FIFO))

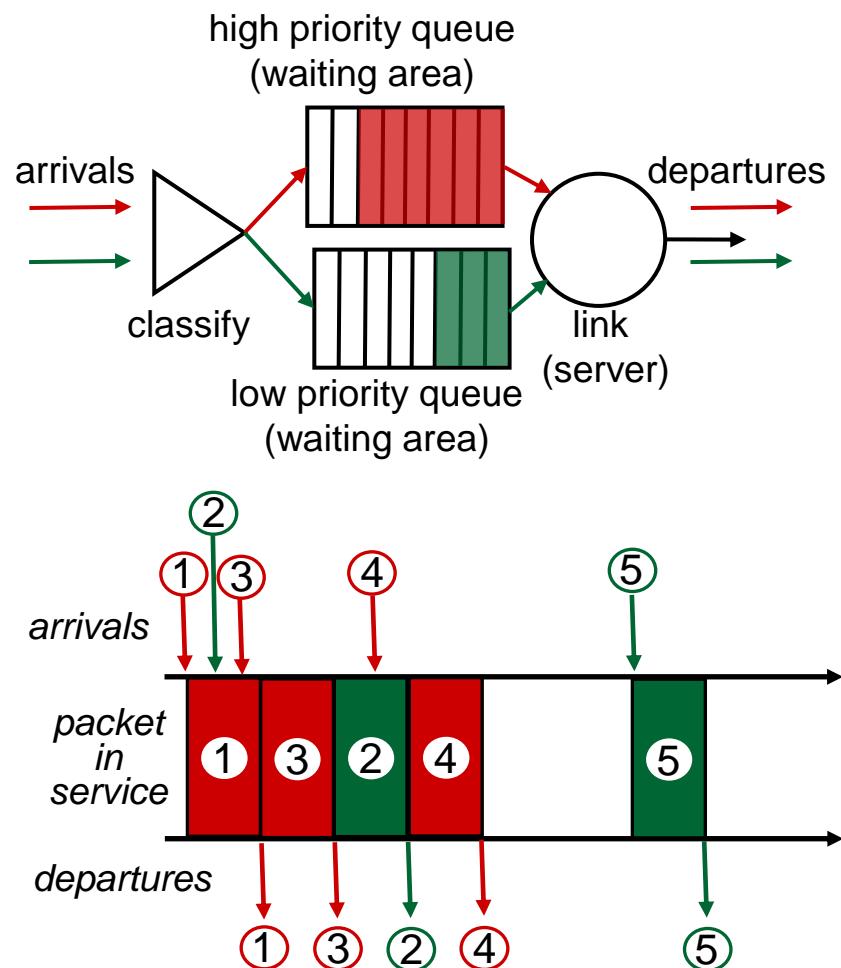
Abstração: fila



Políticas de escalonamento: prioridades

Escalonamento prioritário:

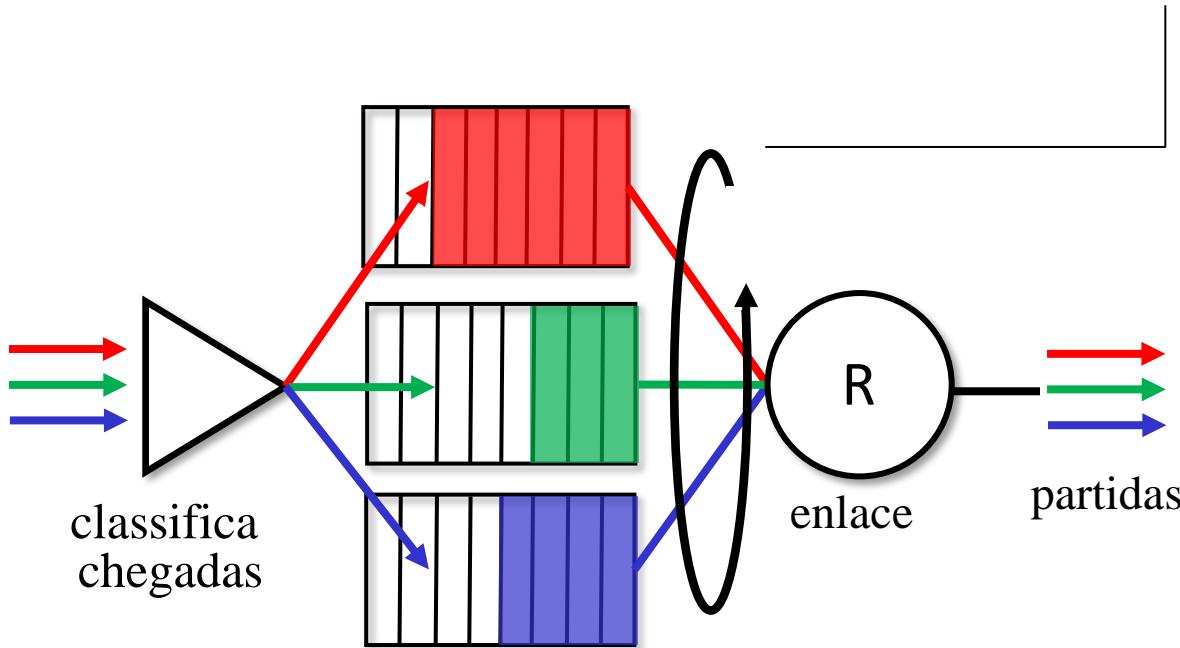
- tráfego chegando é colocado em filas de múltiplas *classes*, com diferentes prioridades
⇒ *classe* pode depender de uma info contida no cabeçalho ou, por ex., do endereço IP origem/destino, nº de porta, etc.
- envia pacote da fila com maior prioridade que contenha pacotes armazenados
 - ✓ FCFS dentro da mesma classe



Políticas de escalonamento: round robin

Escalonamento Round Robin (RR):

⇒ Filas das múltiplas classes são escaneadas ciclicamente, enviando um pacote completo de cada classe (se disponível) por vez



Políticas de escalonamento: weighted fair queueing

Weighted Fair Queueing (WFQ - enfileiramento justo ponderado):

- Generalização do método de Round Robin
- Cada classe i tem um peso w_i , porção ponderada de serviço em cada ciclo

$$\frac{w_i}{\sum_j w_j}$$

Garantia de uma banda mínima por classe de tráfego

