

Unidade 4

Camada de Rede

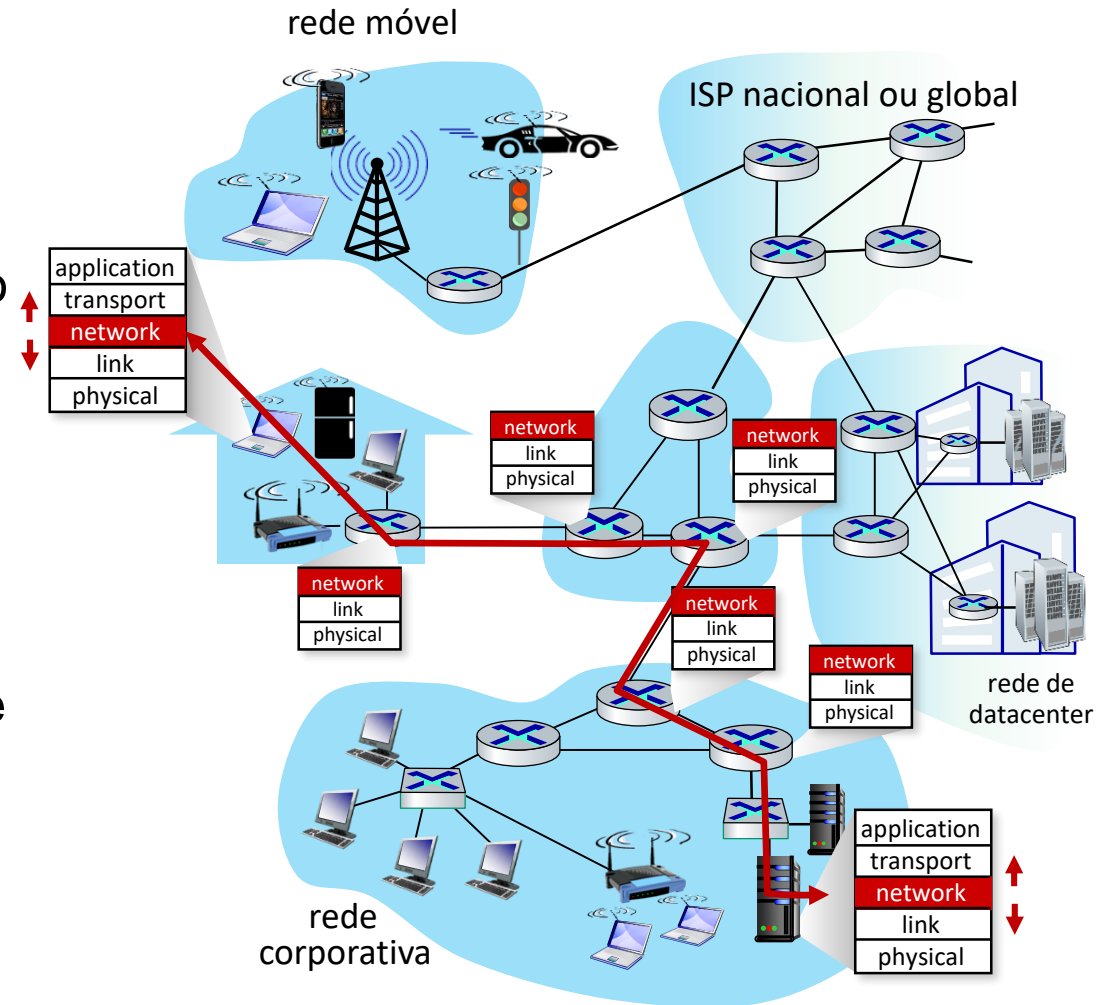
Prof. Ewerton Salvador

Baseado nos slides elaborados por J. F. Kurose e K. W. Ross



Serviços e Protocolos da Camada de Rede

- segmento de transporte do host origem a destino
 - **emissor:** encapsula segmentos em datagramas, passa para a camada de enlace
 - **receptor:** entrega segmentos ao protocolo da camada de transporte
- protocolos da camada de rede em *todos os dispositivos de Internet*: hosts, roteadores
- **roteadores:**
 - examina campos de cabeçalho em todos os datagramas IP que passam
 - move datagramas de portas de entrada para portas de saída para transferir datagramas ao longo do caminho



Duas funções-chave da Camada de Rede

funções de camada de rede:

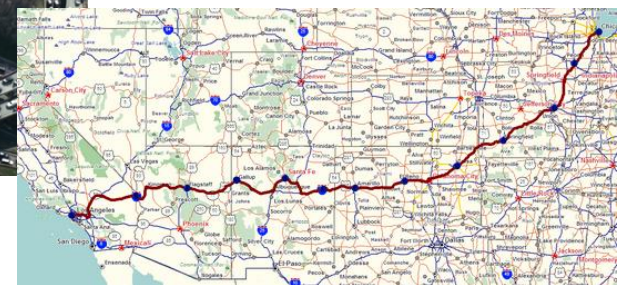
- *repasse*: mover pacotes do enlace de entrada de um roteador para o enlace de saída apropriado do roteador
- *roteamento*: determinar a rota tomada pelos pacotes da origem para o destino
 - *algoritmos de roteamento*

analogia: fazer uma viagem

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



encaminhamento



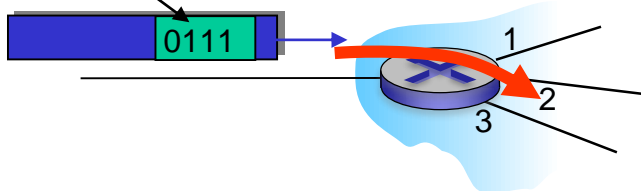
roteamento

Plano de Dados, Plano de Controle

Plano de dados:

- função *local* no roteador
- determina como o datagrama que chega na porta de entrada do roteador é repassado para a porta de saída

valores no cabeçalho do pacote de chegada

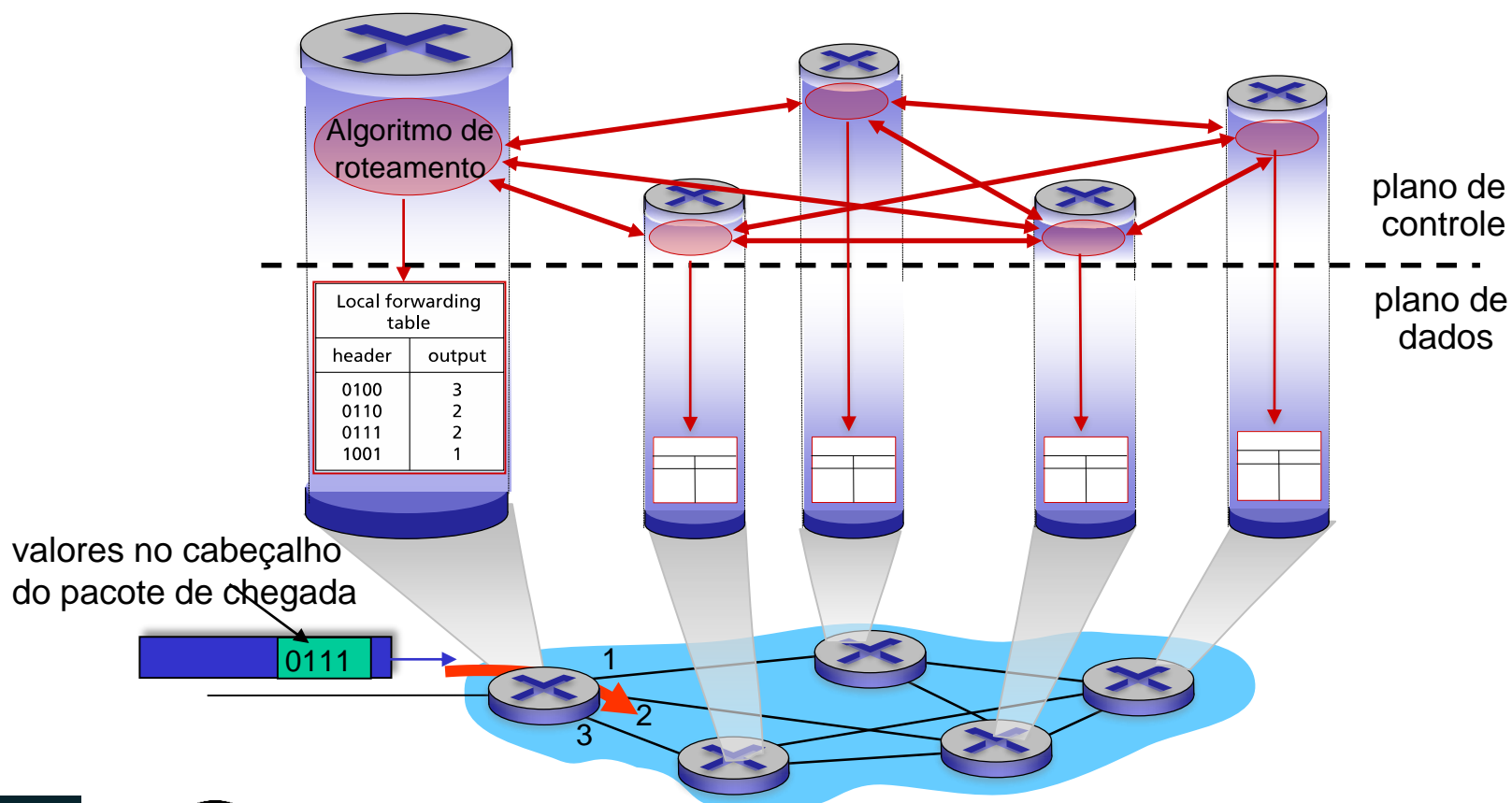


Plano de controle

- lógica em *toda a rede*
- determina como o datagrama é roteado entre roteadores ao longo do caminho final do host de origem para o host de destino
- duas abordagens de plano de controle:
 - *algoritmos de roteamento tradicionais*: implementado nos roteadores
 - *software-defined networking (SDN)*: implementado em servidores (remotos)

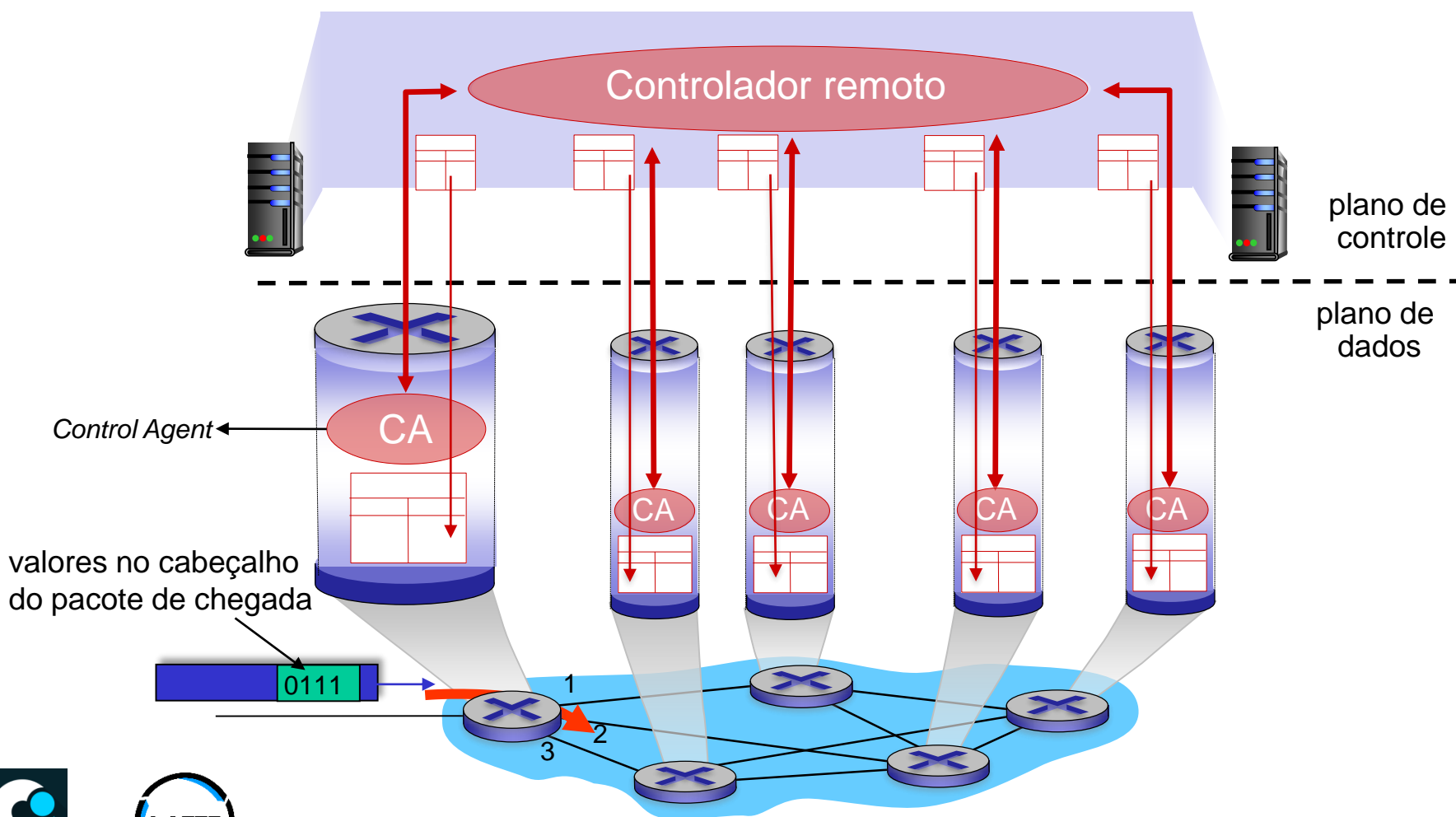
Plano de controle no roteador

componentes individuais do algoritmo de roteamento em *todo e qualquer roteador* interagem no plano de controle



Plano de controle SDN

controlador remoto calcula e instala tabelas de encaminhamento em roteadores



Modelo de serviço de rede

Pergunta: Qual **modelo de serviço** para "canal" transportando datagramas do remetente para o receptor?

Exemplos de serviços para datagramas individuais:

- entrega garantida
- entrega garantida com menos de 40 mseg de atraso

Exemplos de serviços para um fluxo de datagramas:

- entrega ordenada de datagramas
- largura de banda mínima garantida para fluxo
- restrições a alterações no espaçamento entre pacotes

Modelo de serviço de camada de rede

| Arquitetura de rede | Modelo de Serviço | Garantias de Quality of Service (QoS)? | | | |
|---------------------|-------------------|--|------|-------|--------|
| | | Bandwidth | Loss | Order | Timing |
| Internet | melhor esforço | nenhuma | não | não | não |

Modelo de serviço de "melhor esforço" da Internet

Sem garantias sobre:

- i. entrega bem-sucedida do datagrama ao destino
- ii. Entrega no tempo e em ordem
- iii. largura de banda disponível para o fluxo final

Modelo de serviço de camada de rede

| Arquitetura de rede | Modelo de Serviço | Garantias de Quality of Service (QoS)? | | | |
|---------------------|-------------------------------|--|----------|----------|--------|
| | | Bandwidth | Loss | Order | Timing |
| Internet | melhor esforço | nenhuma | não | não | não |
| ATM | Taxa de bit constant. | Taxa constant. | sim | sim | sim |
| ATM | Taxa de bit dispon. | Garantia minim. | não | sim | não |
| Internet | Intserv Guaranteed (RFC 1633) | sim | sim | sim | sim |
| Internet | Diffserv (RFC 2475) | possível | possível | possível | não |

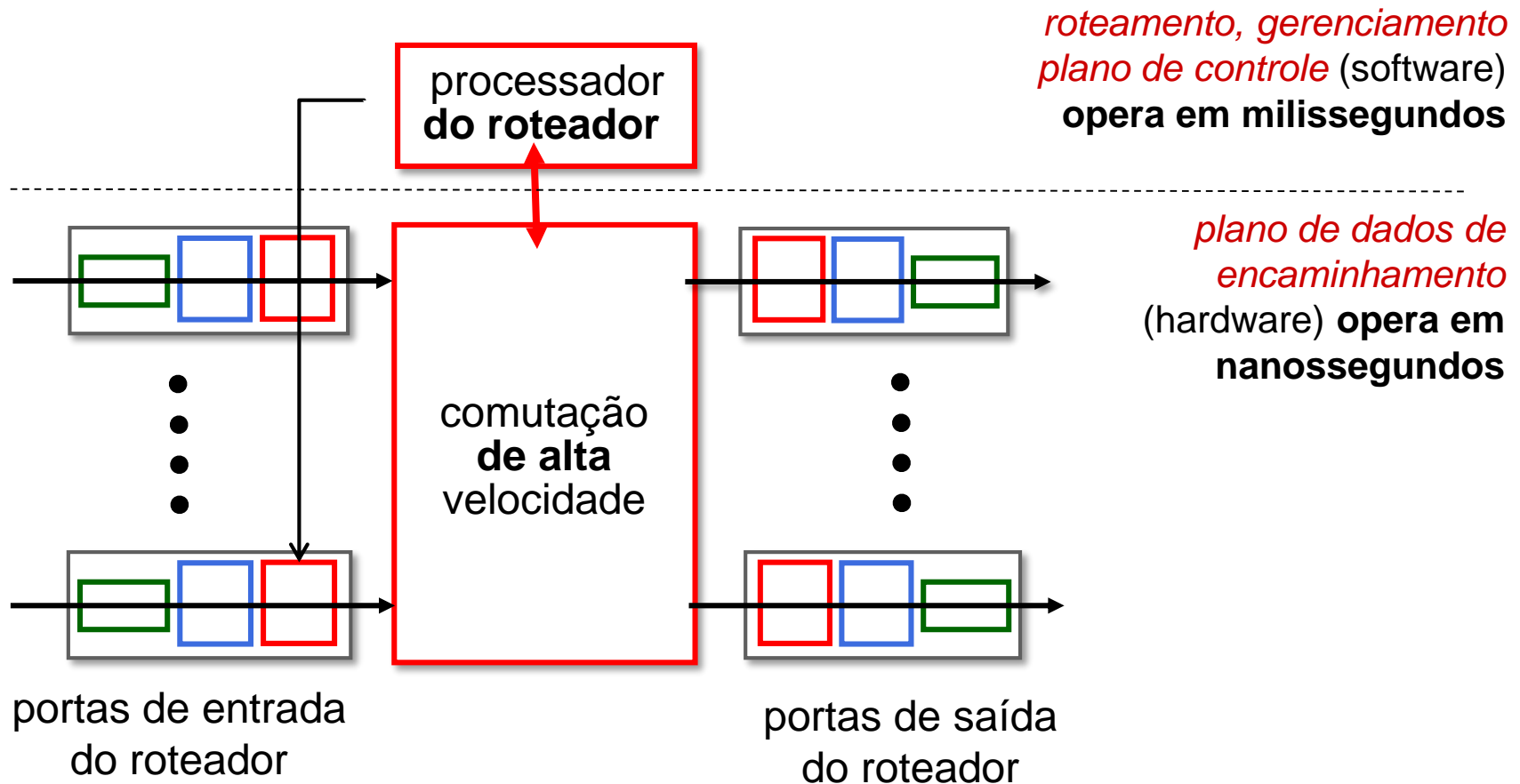
Reflexões sobre o “melhor esforço”

- **Simplicidade do mecanismo** permitiu que a Internet fosse amplamente implantada
- Provisionamento suficiente de **largura de banda** permite que o desempenho de aplicações de tempo real (ex.:, voz interativa, vídeo) seja “bom o suficiente” para “a maior parte do tempo”
- **Serviços distribuídos replicados na camada de aplicação** (datacenters, redes de distribuição de conteúdo) conectando-se perto das redes dos clientes. Permite que serviços sejam fornecidos a partir de vários locais
- Controle de congestionamento de serviços “elásticos” ajuda

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

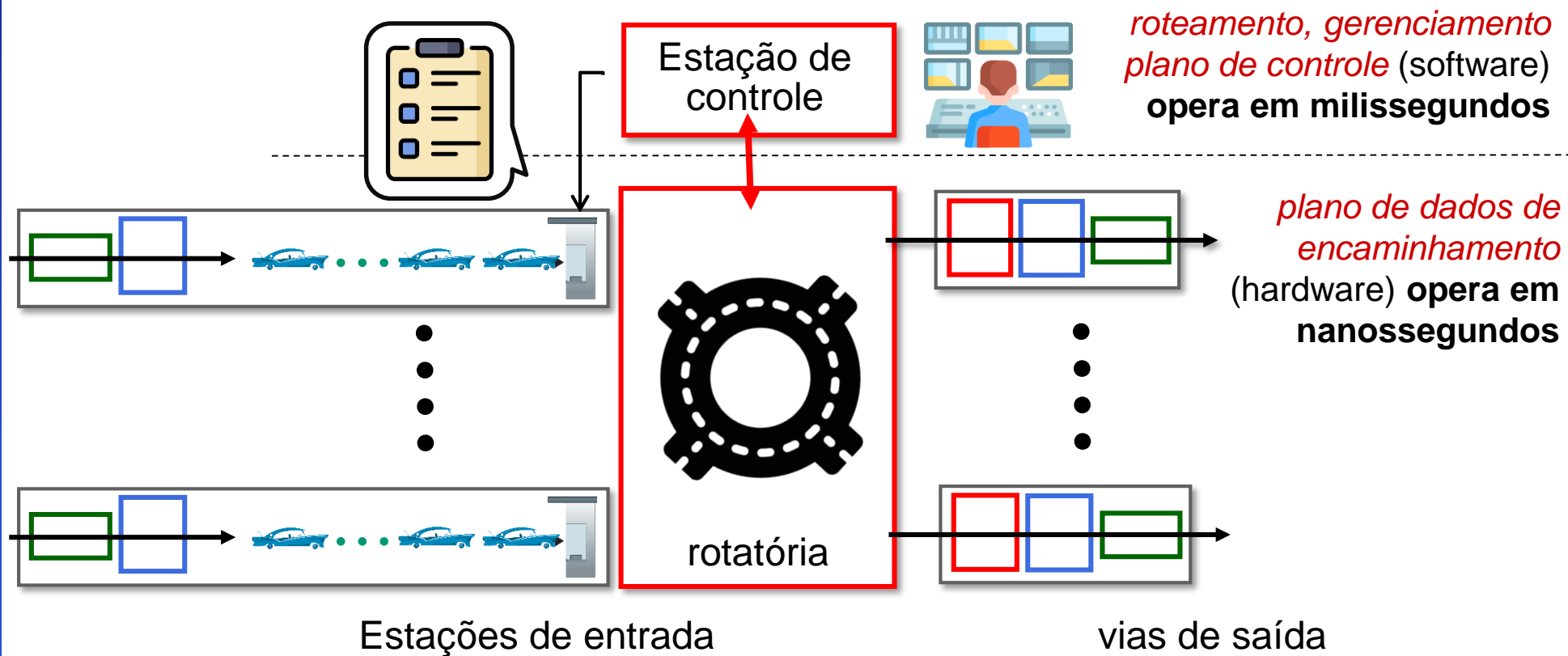
Visão geral da arquitetura do roteador

visão de alto nível da arquitetura genérica do roteador:

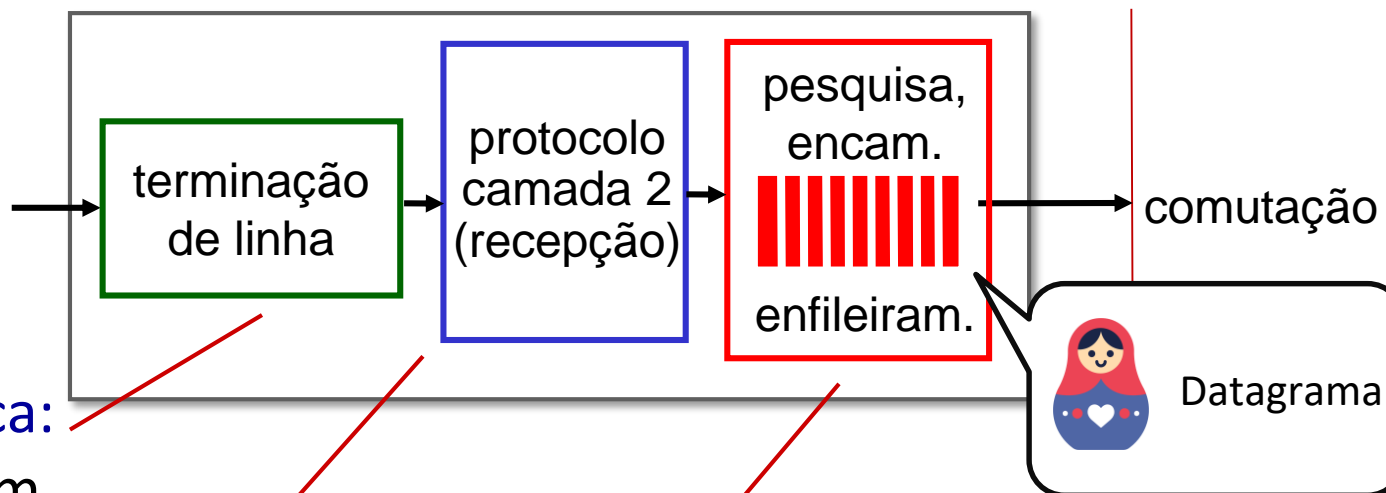


Visão geral da arquitetura do roteador

Visão de uma analogia da arquitetura genérica do roteador:



Funções da porta de entrada



camada física:
recepção em
nível de bits

camada de enlace:
e.g., Ethernet

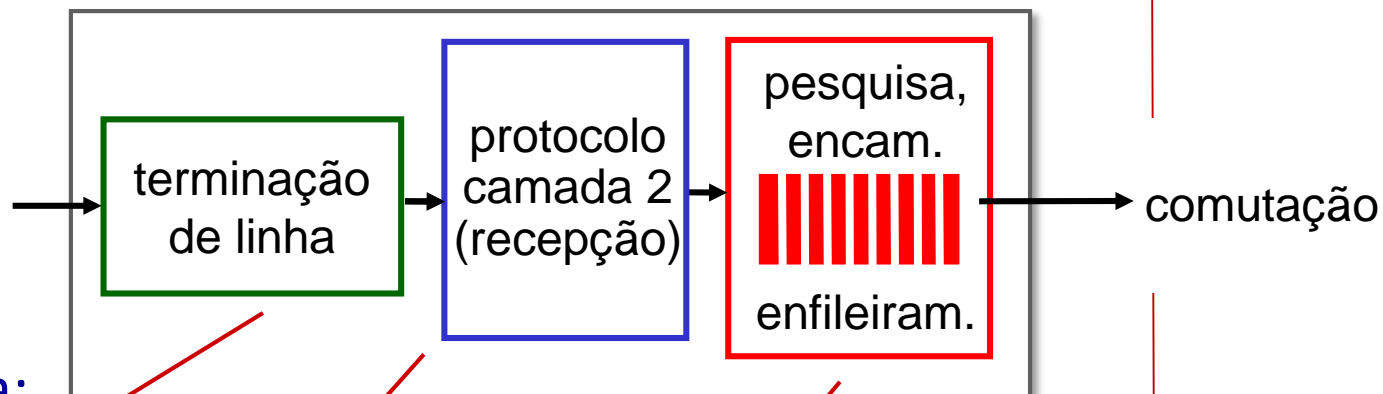
comutação descentralizada:

- usa valores de campo de cabeçalho, procura por porta de saída usando tabela de repasse na porta de entrada
- objetivo: processamento completo da porta de entrada na 'velocidade da linha'
- **enfileiramento de portas de entrada:** se datagramas chegam mais rápido do que a taxa de encaminhamento para comutação



Quadro

Funções da porta de entrada



camada física:
recepção em
nível de bits

camada de enlace:
e.g., Ethernet

comutação descentralizada :

- usa valores de campo de cabeçalho, procura por porta de saída usando tabela de repasse na porta de entrada
- **encaminhamento baseado em destino:** encaminhamento com base apenas no endereço IP de destino (tradicional)
- **encaminhamento generalizado:** encaminhar com base em qualquer conjunto de valores de campo de cabeçalho

Encaminhamento baseado em destino

forwarding table

| Destination Address Range | Link Interface |
|---|----------------|
| 11001000 00010111 00010000 00000000 through 11001000 00010111 00010000 00000100 through 11001000 00010111 00010000 00000111 | n 3 |
| 11001000 00010111 00011000 11111111 through 11001000 00010111 00011001 00000000 through 11001000 00010111 00011111 11111111 | 2 |
| otherwise | 3 |

P: Mas o que acontece se os intervalos não se dividirem tão bem?

Correspondência de prefixo mais longo

longest prefix match

ao procurar a entrada da tabela de repasse para endereço de destino, use o prefixo de endereço **mais longo** que corresponda ao endereço de destino.

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

exemplos:

11001000 00010111 00010110 10100001 qual interface?

11001000 00010111 00011000 10101010 qual interface?

Correspondência de prefixo mais longo

longest prefix match

ao procurar a entrada da tabela de repasse para endereço de destino, use o prefixo de endereço **mais longo** que corresponda ao endereço de destino.

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

exemplos:

| | |
|-------------------------------------|-----------------|
| 11001000 00010111 00010110 10100001 | qual interface? |
| 11001000 00010111 00011000 10101010 | qual interface? |

Correspondência de prefixo mais longo

longest prefix match

ao procurar a entrada da tabela de repasse para endereço de destino, use o prefixo de endereço **mais longo** que corresponda ao endereço de destino.

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

match!

exemplos:

11001000 00010111 00010110 10100001 qual interface?

11001000 00010111 00011000 10101010 qual interface?

Correspondência de prefixo mais longo

longest prefix match

ao procurar a entrada da tabela de repasse para endereço de destino, use o prefixo de endereço **mais longo** que corresponda ao endereço de destino.

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

match!

exemplos:

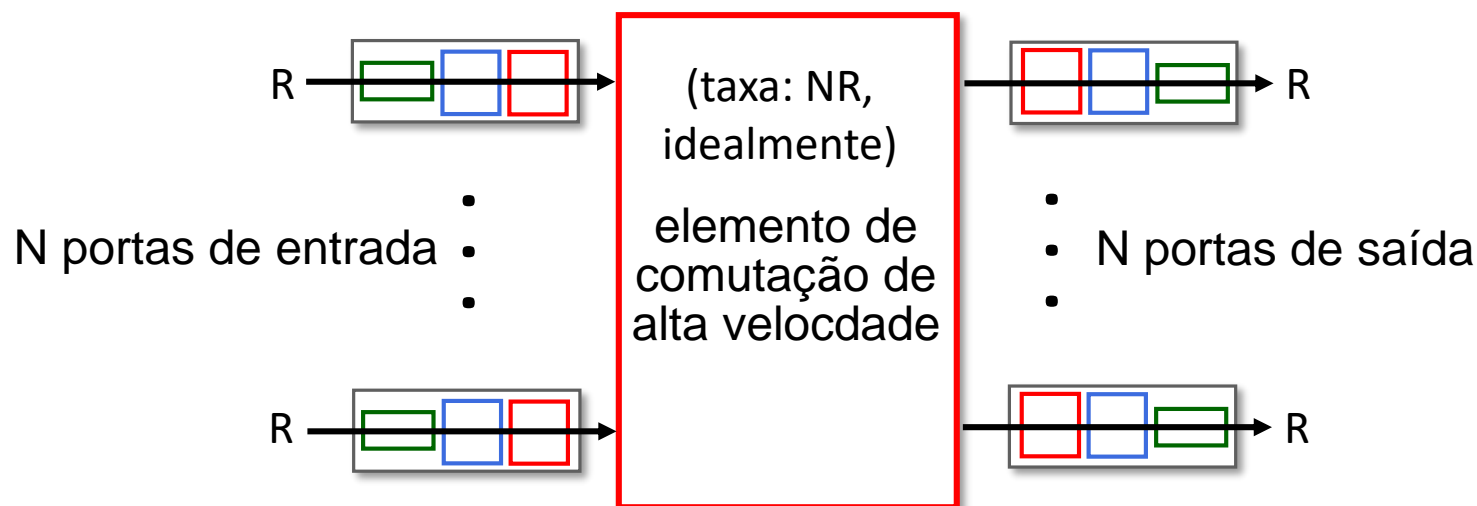
| | | | | |
|----------|----------|----------|----------|-----------------|
| 11001000 | 00010111 | 00010110 | 10100001 | qual interface? |
| 11001000 | 00010111 | 00011000 | 10101010 | qual interface? |

Correspondência de prefixo mais longo

- Veremos por que a correspondência de **prefixo mais longa** é usada em breve, quando estudarmos o endereçamento
- Correspondência de prefixo mais longa: geralmente executada usando memórias endereçáveis de conteúdo ternário (TCAMs)
 - necessárias outras técnicas além da busca linear simples por uma tabela grande
 - **conteúdo endereçável**: endereço presente para TCAM: recuperar endereço em um ciclo de relógio, independentemente do tamanho da tabela
 - Cisco Catalyst: ~1M de entradas na tabela de roteamento no TCAM

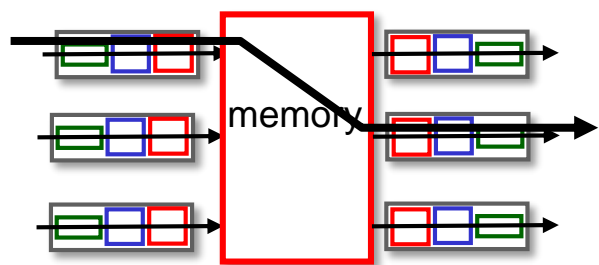
Elemento de comutação

- transfere pacote do enlace de entrada para o enlace de saída apropriado
- **taxa de comutação:** taxa na qual os pacotes podem ser transferidos de entradas para saídas
 - frequentemente medido como múltiplo da taxa de linha de entrada/saída
 - N entradas: taxa de comutação N vezes taxa de linha desejável

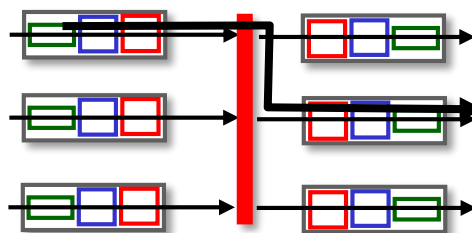


Elemento de comutação

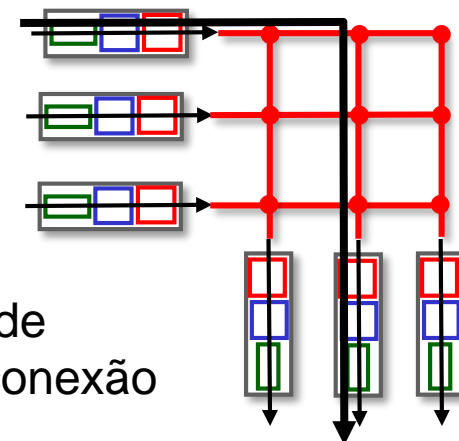
- transfere pacote do enlace de entrada para o enlace de saída apropriado
- taxa de comutação:** taxa na qual os pacotes podem ser transferidos de entradas para saídas
 - frequentemente medido como múltiplo da taxa de linha de entrada/saída
 - N entradas: taxa de comutação N vezes taxa de linha desejável
- três tipos principais de elementos de comutação:



memória



barramento

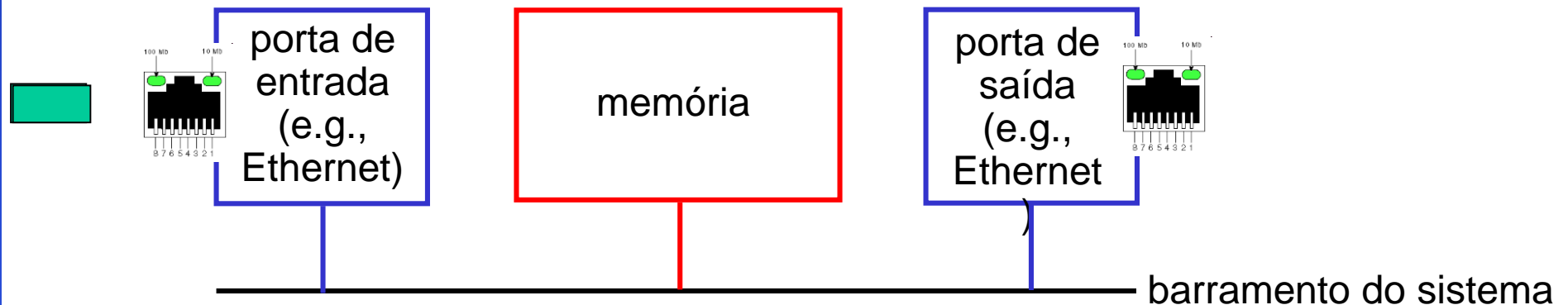


rede de
interconexão

Comutação via memória

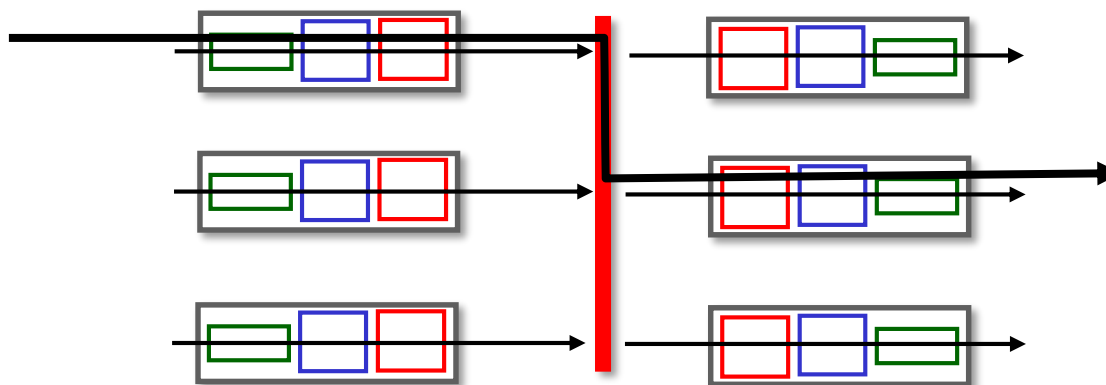
Roteadores de primeira geração:

- computadores tradicionais com comutação sob controle direto da CPU
- pacote copiado para a memória do sistema
- velocidade limitada pela largura de banda da memória (datagrama cruza o barramento duas vezes)



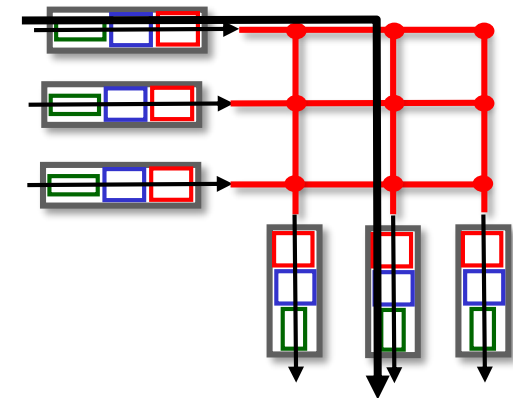
Comutação via barramento

- datagrama da memória da porta de entrada para a memória da porta de saída por meio de um barramento compartilhado
- **contenção de barramento:** velocidade de comutação limitada pela largura de banda do barramento
- barramento de 32Gbps, Cisco 5600: velocidade suficiente para roteadores de acesso

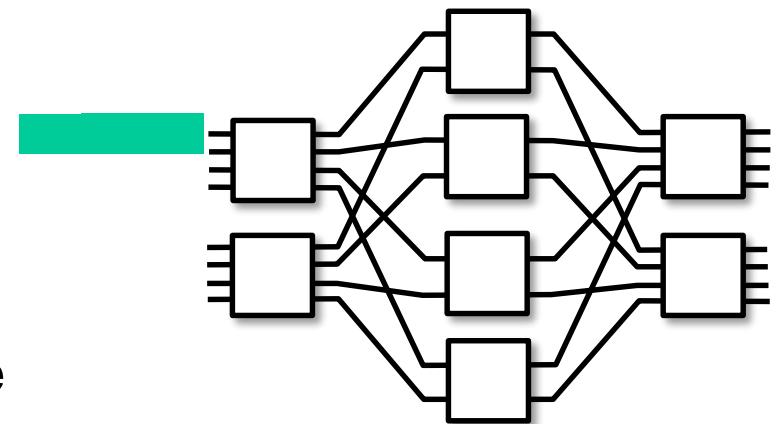


Comutação via rede de interconexão

- Crossbar, redes Clos e outras redes de interconexão inicialmente desenvolvidas para conectar processadores em sistemas multiprocessador
- **switch multiestágio:** $n \times n$ comuta entre múltiplos estágios de switches menores
- **explora o paralelismo:**
 - fragmenta datagrama em células de comprimento fixo na entrada
 - Comuta células através do elemento de comutação, remonta datagrama na saída



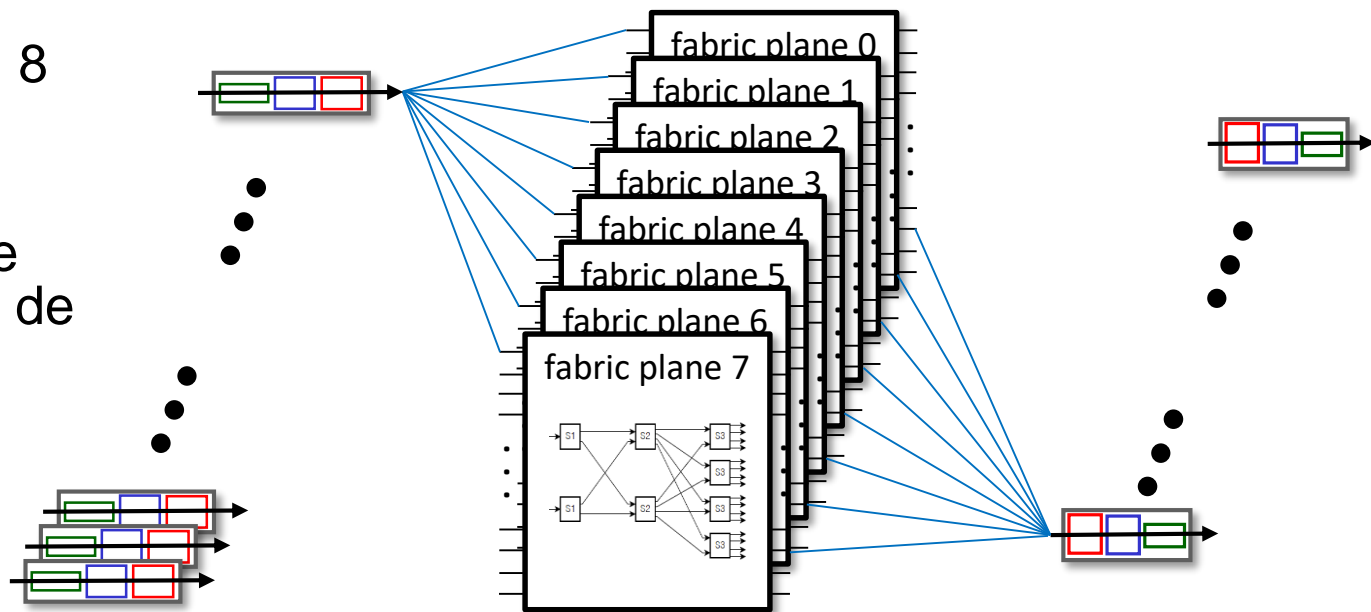
crossbar 3x3



switch multiestágio 8x8
construído a partir de
switches menores

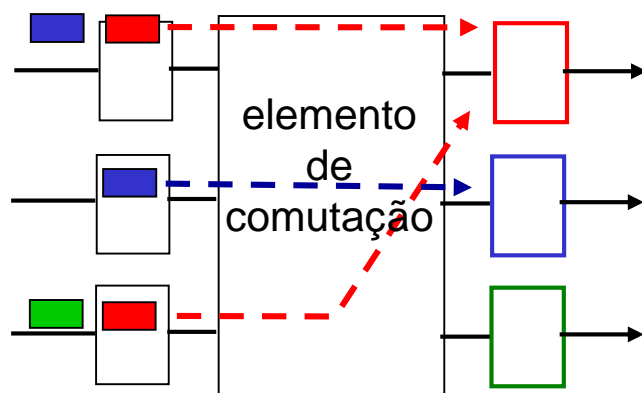
Comutação via rede de interconexão

- escala usando vários “planos” de comutação em paralelo:
 - expansão via paralelismo
- roteador Cisco CRS:
 - Unidade básica: 8 planos de comutação
 - cada plano: rede de interconexão de 3 estágios
 - até centenas de Tbps de capacidade de comutação

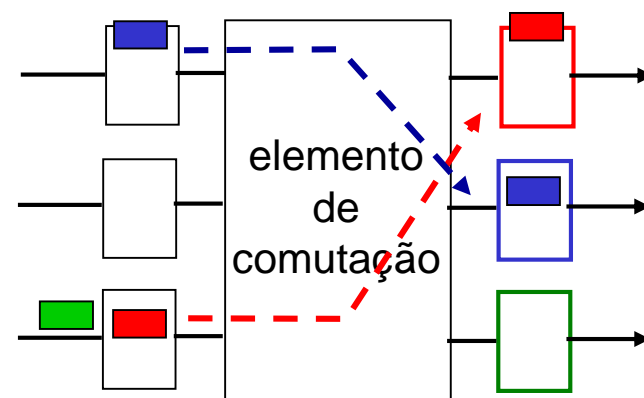


Enfileiramento nas portas de entrada

- Se o elemento de comutação for mais lento do que as portas de entrada combinadas, poderá ocorrer filas de entrada
 - atraso e perda de fila devido ao estouro de buffer de entrada!
- **Bloqueio Head-of-the-Line (HoL):** datagrama na frente da fila impede que outros na fila avancem

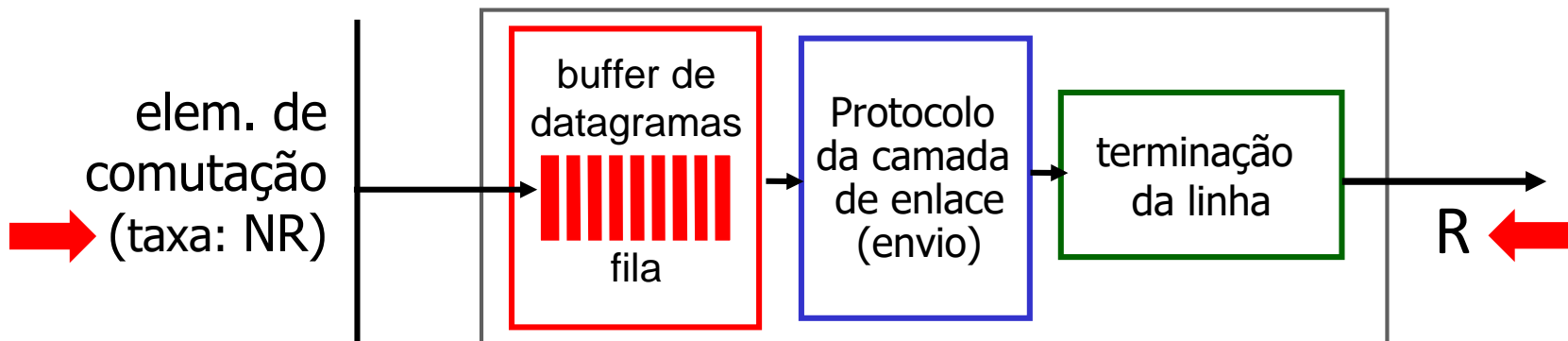


contenção da porta de saída: apenas um datagrama vermelho pode ser transferido. O pacote vermelho inferior está **bloqueado**



um pacote depois: o pacote verde experimenta bloqueio HOL

Enfileiramento de portas de saída

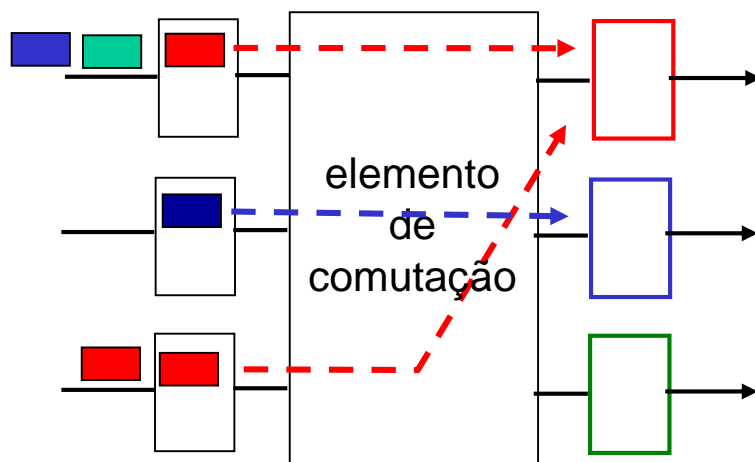


- **Buffering** necessário quando os datagramas chegam da comutação mais rápido do que a taxa de transmissão do link.
- **Política de descarte:** quais datagramas descartar se não houver buffers livres?
- **Disciplina de escalonamento** escolhe entre datagramas enfileirados para transmissão

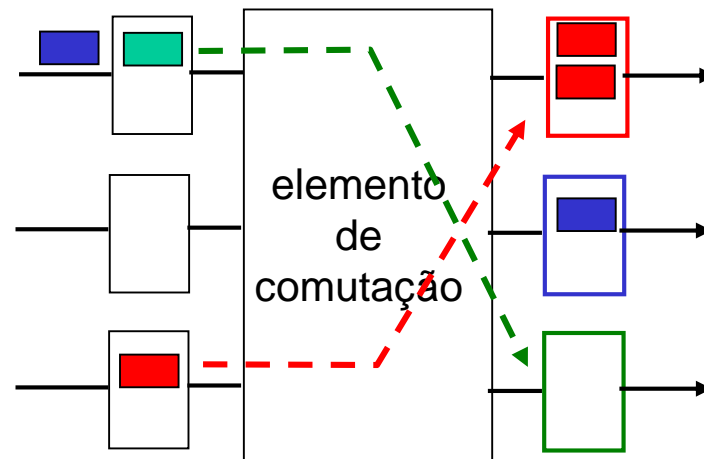
→ Datagramas podem ser perdidos devido a congest., falta de buffers

→ Escalonamento prioritário – quem obtém melhor desempenho, neutralidade de rede

Enfileiramento de portas de saída



no instante t, pacotes de
mais de uma entrada
para uma saída



um pacote depois

- bufferização quando a taxa de chegada via comutador excede a velocidade da linha de saída
- enfileiramento (atraso) e perda devido ao overflow do buffer da porta de saída!

Quanto buffer é “suficiente”?

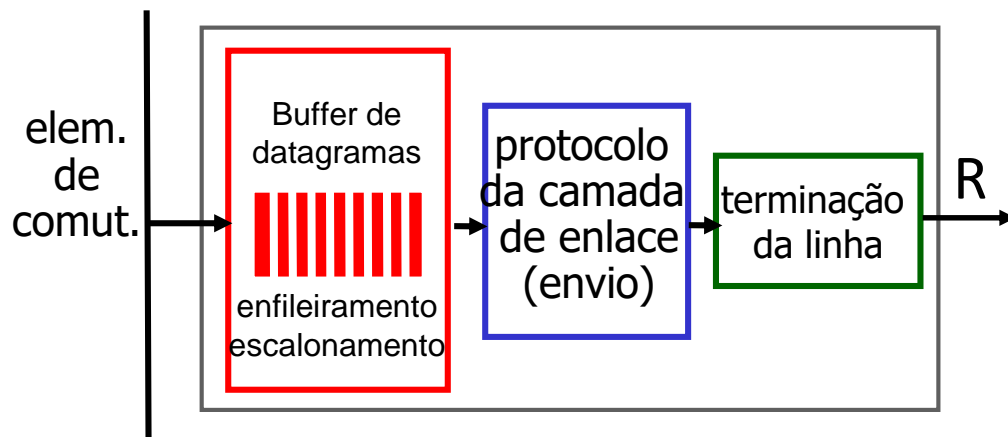
- Regra prática da RFC 3439: buferização media é igual ao RTT “típico” (por exemplo 250 msec) vezes capacidade C do enlace
 - e.g., C = enlace de 10 Gbps: 2.5 Gbit buffer

- recomendação mais recente: para N fluxos, buffer igual a

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

- buffer em excesso pode aumentar atrasos (particularmente em roteadores domésticos)
 - RTTs longos: baixo desempenho para aplicações em tempo real, responsividade lenta do TCP frente a um congestionamento
 - Lembre-se do controle de congestionamento baseado em atraso: “manter o canal cheio, mas não mais do que cheio”

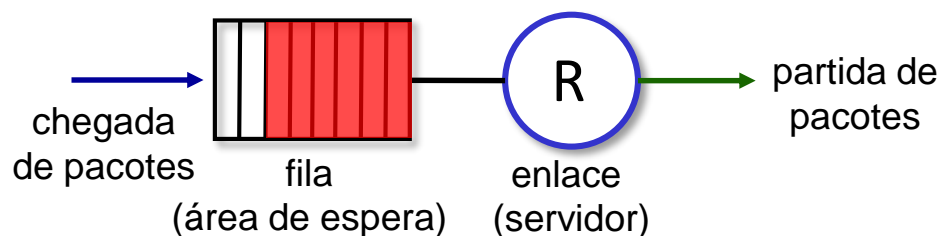
Gerenciamento de buffer



Gerenciamento de buffer:

- **descarte**: qual pacote descartar quando os buffers estiverem cheios
 - **Descarte de cauda**: descarta pacote chegando
 - **Prioridade**: descartar com base na prioridade
- **marcação**: quais pacotes marcar para sinalizar congestionamento (ECN, RED)

Abstração: fila



Política de escalonamento: FCFS

Escalonamento de pacotes:

Decide qual pacote enviar em seguida no link

- *First Come, First Served*
- Prioridade
- *Round Robin*
- *Weighted Fair Queueing* (WFQ)

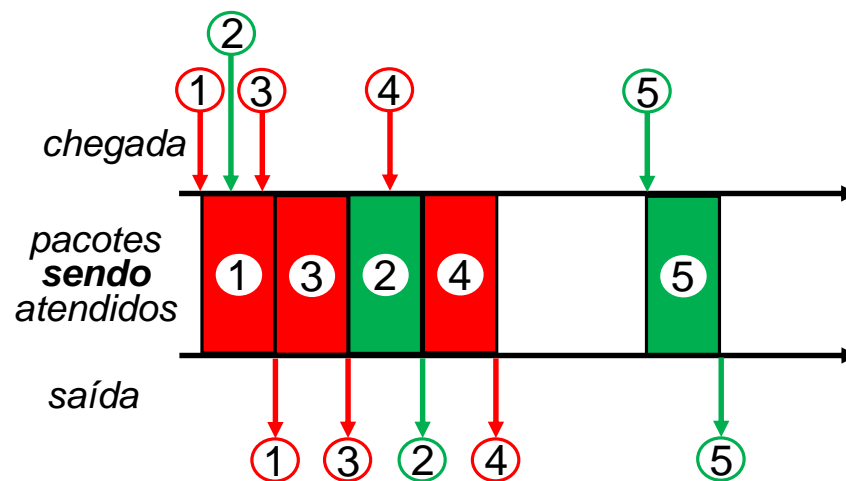
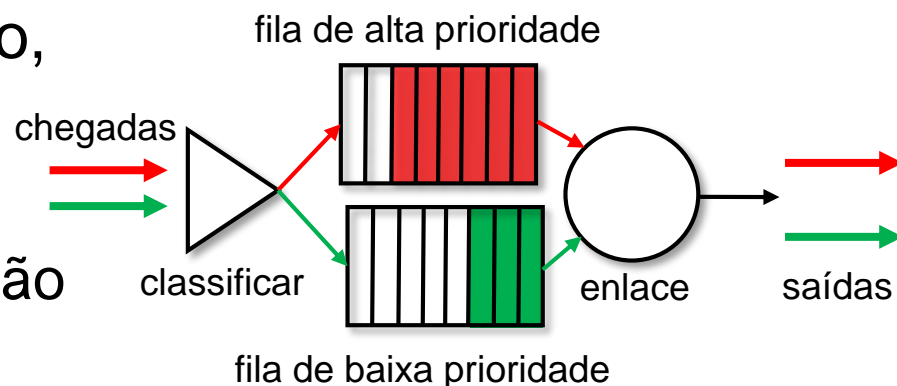
FCFS: pacotes transmitidos por ordem de chegada à porta de saída

- também conhecido como: *First-in-first-out* (FIFO)

Política de escalonamento: Prioridade

Escalonamento por prioridade:

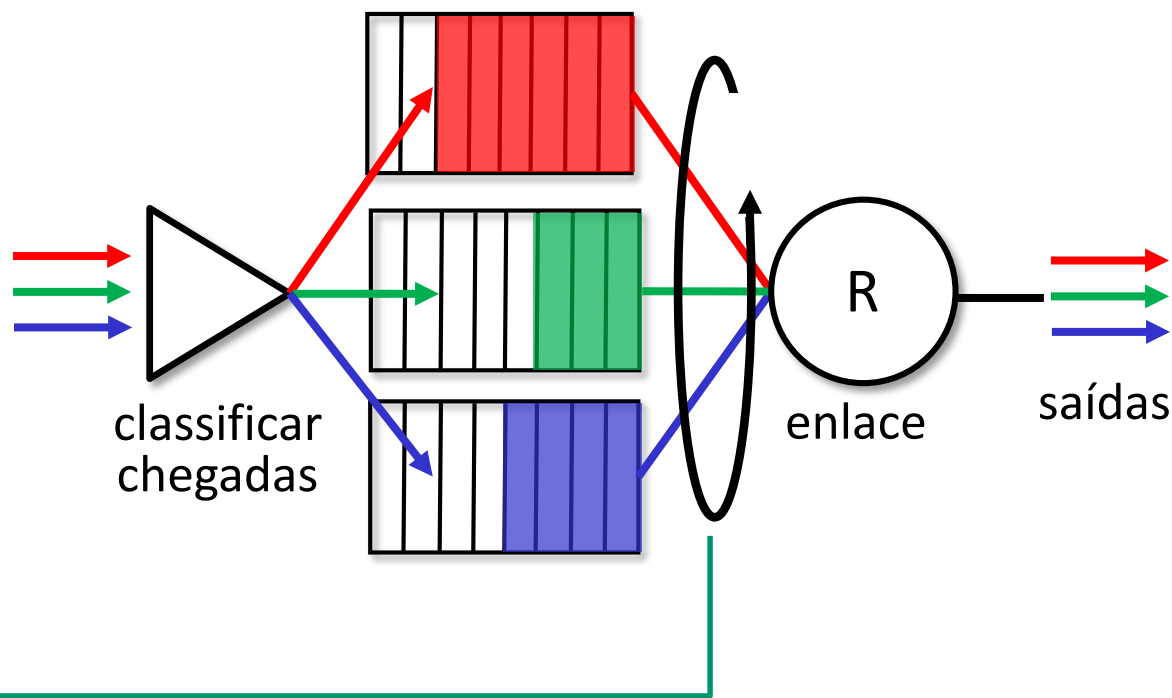
- tráfego de chegada classificado, enfileirado por classe
 - qualquer campo de cabeçalho pode ser usado para classificação
- enviar pacote da fila de prioridade mais alta que tenha pacotes armazenados em buffer
 - FCFS dentro da classe de prioridade



Política de escalonamento: Round Robin

Escalonamento Round Robin (RR):

- tráfego de chegada classificado, enfileirado por classe
 - qualquer campo de cabeçalho pode ser usado para classificação
- verifica ciclicamente as filas de classe, enviando um pacote completo de cada classe (se disponível) na sua vez



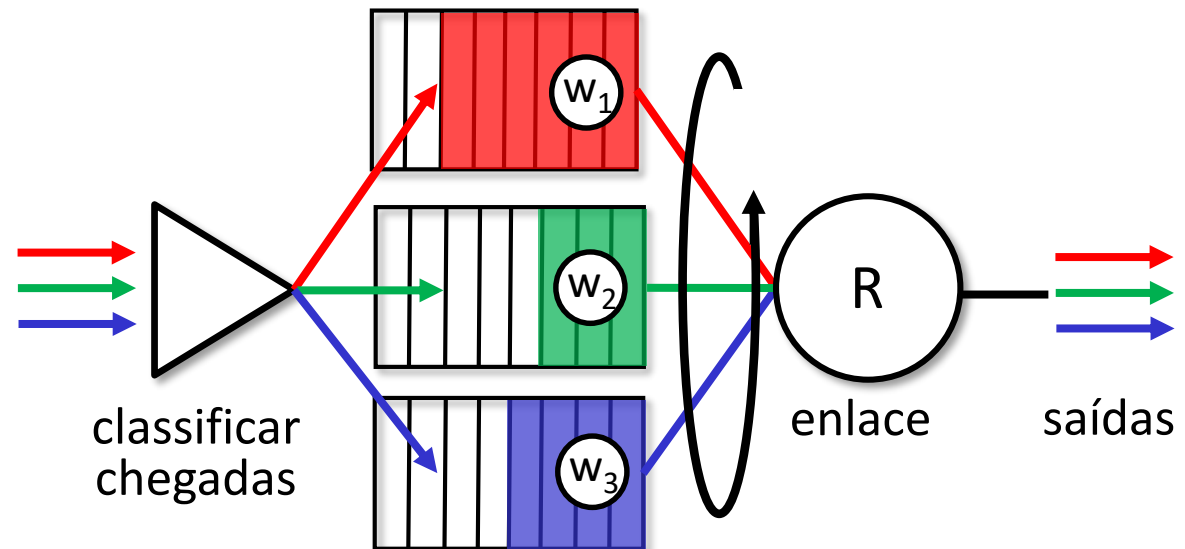
Política de escalonamento: Weighted Fair Queueing

Enfileiramento Justo Ponderado (WFQ):

- Round Robin generalizado
- cada classe i tem peso w_i , e recebe quantidade ponderada de serviço em cada ciclo

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

- garantia de largura de banda mínima (por classe de tráfego)



Neutralidade da Rede

O que é Neutralidade da Rede?

- **Aspecto técnico:** como um ISP deve dividir/alocar seus recursos
 - escalonamento de pacote e gerenciamento de buffer são os **mecanismos**
- **Aspectos sociais e econômicos**
 - Proteger liberdade de expressão
 - Incentivar inovação/competição
- Aplicação de leis e políticas

Países diferentes possuem interpretações diferentes sobre
Neutralidade da Rede

Neutralidade da Rede

- Nos Estados Unidos

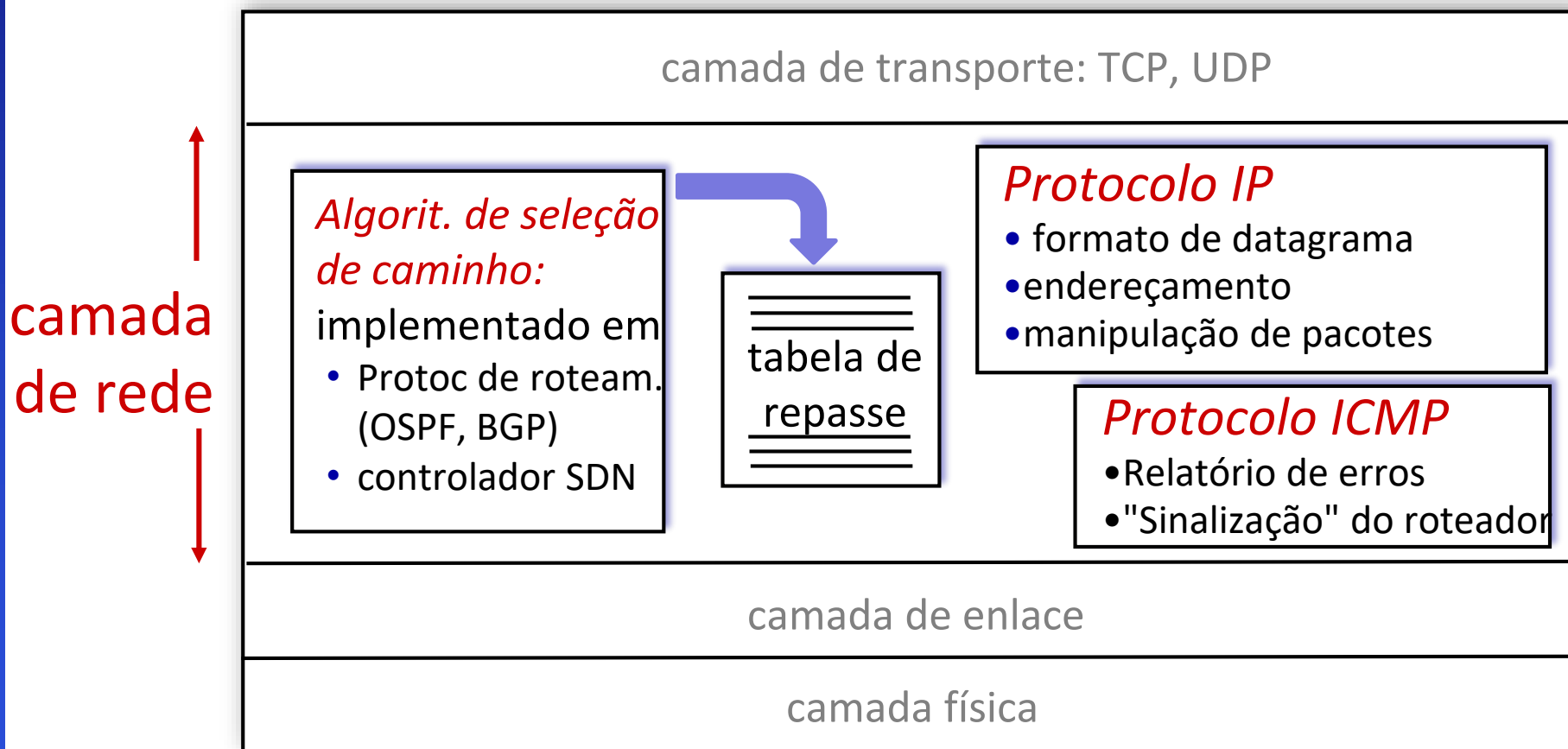
- Neutralidade da Rede imposta em 2015
- Revogada em 2017
- FCC iniciou em 2023 um processo para restauração da Neutralidade da Rede

- No Brasil

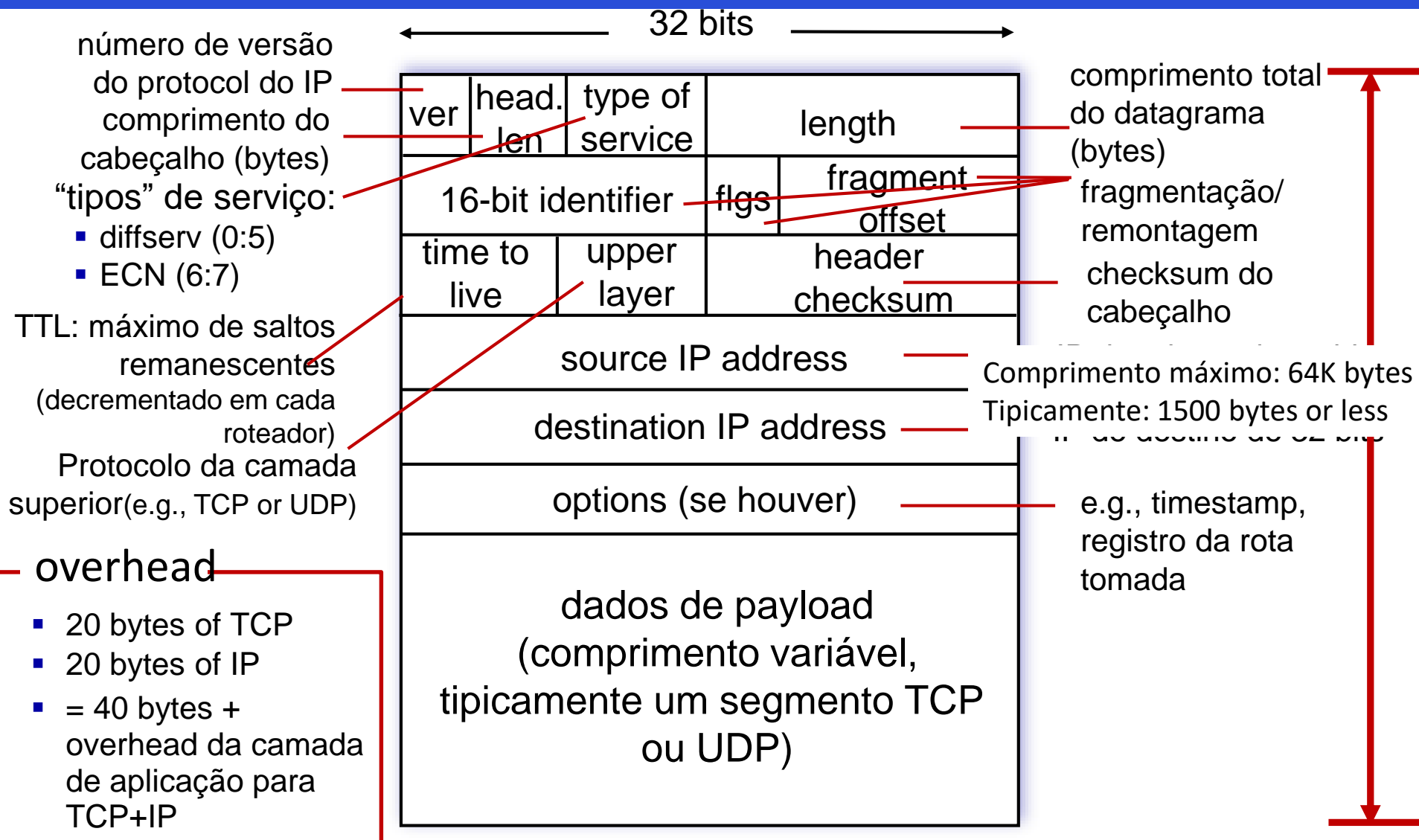
- Marco Civil da Internet (2014) estabelece a Neutralidade da Rede como um de seus princípios
- Permite discriminar tráfego apenas em casos de ser um “requisito indispensável à prestação do serviço” ou em caso de “priorização de serviço de emergência”

Camada de rede: Internet

- Funções da camada de rede do host ou roteador:

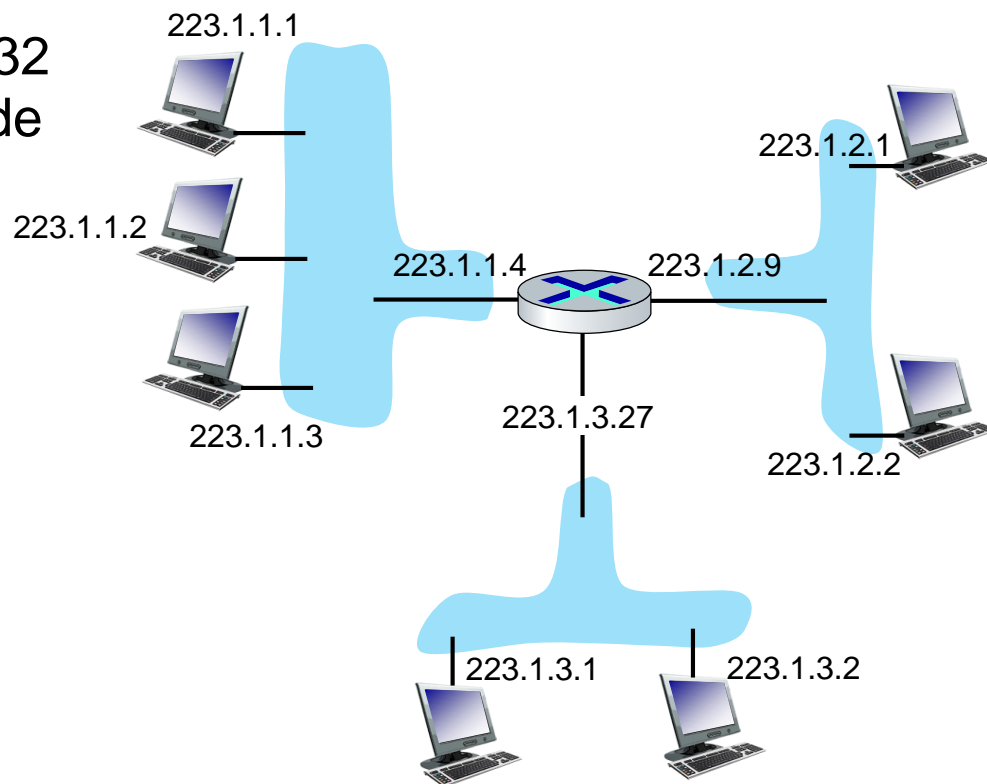


Formato de datagrama IP



Endereçamento IP: introdução

- **endereço IP:** identificador de 32 bits associado a cada interface de host ou roteador
- **interface:** conexão entre host/roteador e enlace físico
 - roteadores normalmente têm várias interfaces
 - host normalmente tem uma ou duas interfaces (e.g., ethernet, wireless 802.11)

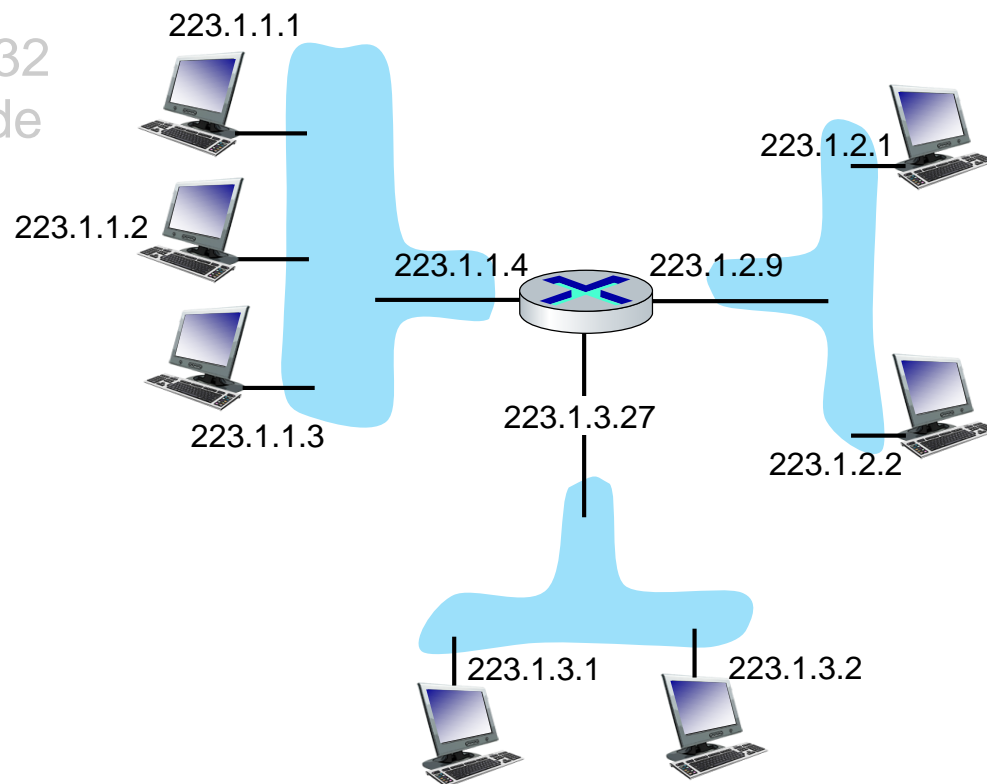


notação decimal separada por pontos:

$$223.1.1.1 = \underbrace{11011111}_{223} \underbrace{00000001}_{1} \underbrace{00000001}_{1} \underbrace{00000001}_{1}$$

Endereçamento IP: introdução

- **endereço IP:** identificador de 32 bits associado a cada interface de host ou roteador
- **interface:** conexão entre host/roteador e enlace físico
 - roteadores normalmente têm várias interfaces
 - host normalmente tem uma ou duas interfaces (e.g., ethernet, wireless 802.11)



notação de endereço IP decimal :

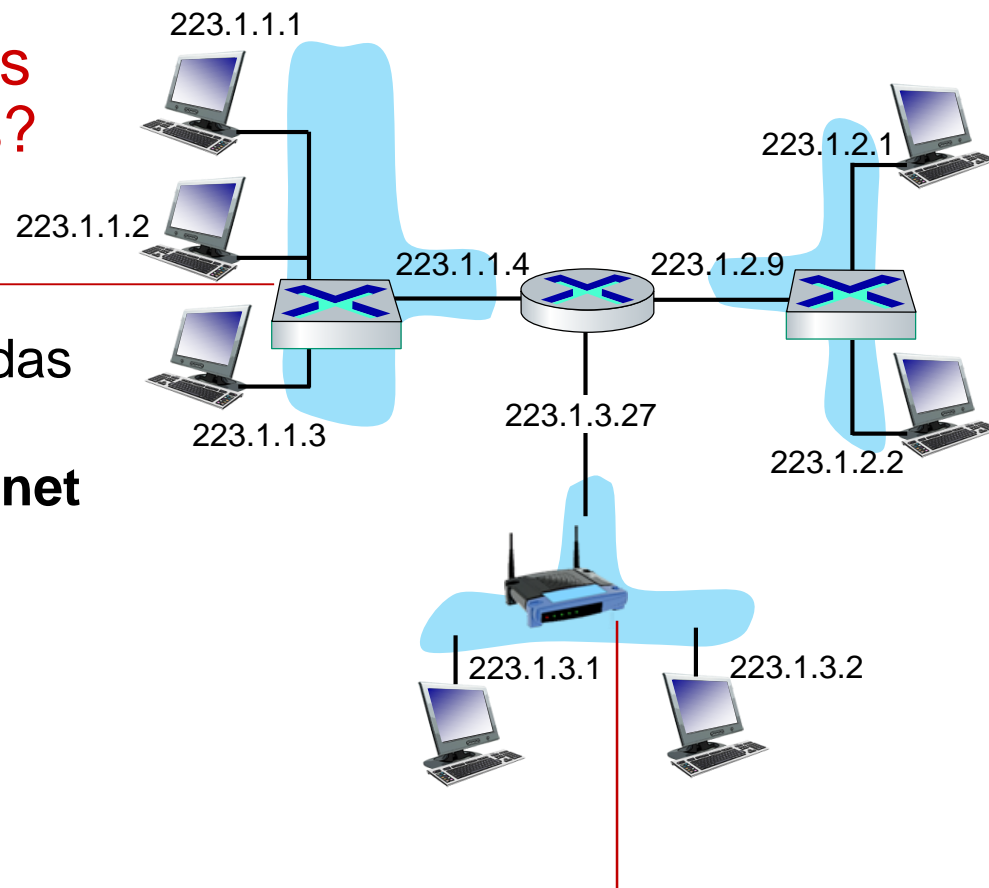
223.1.1.1 = 11011111 00000001 00000001 00000001

223 1 1 1

Endereçamento IP: introdução

Pergunta: como as interfaces estão realmente conectadas?

R: interfaces Ethernet cabeadas conectadas por **switches Ethernet**



R: interfaces WiFi sem fio conectadas pela **estação base WiFi**

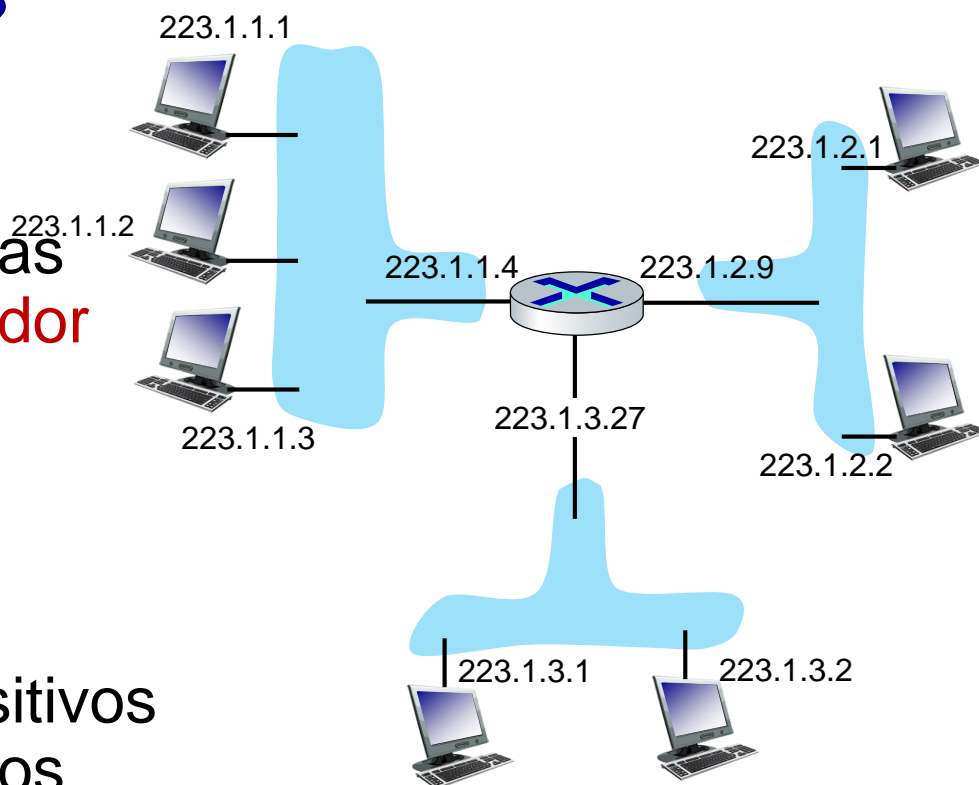
Sub-redes

■ *O que é uma sub-rede?*

- interfaces de dispositivos que podem alcançar fisicamente umas às outras **sem passar por um roteador interveniente**

■ Os endereços IP têm estrutura:

- **parte da sub-rede:** dispositivos na mesma sub-rede têm os mesmos bits mais significativos
- **parte do host:** bits menos significativos **restantes**

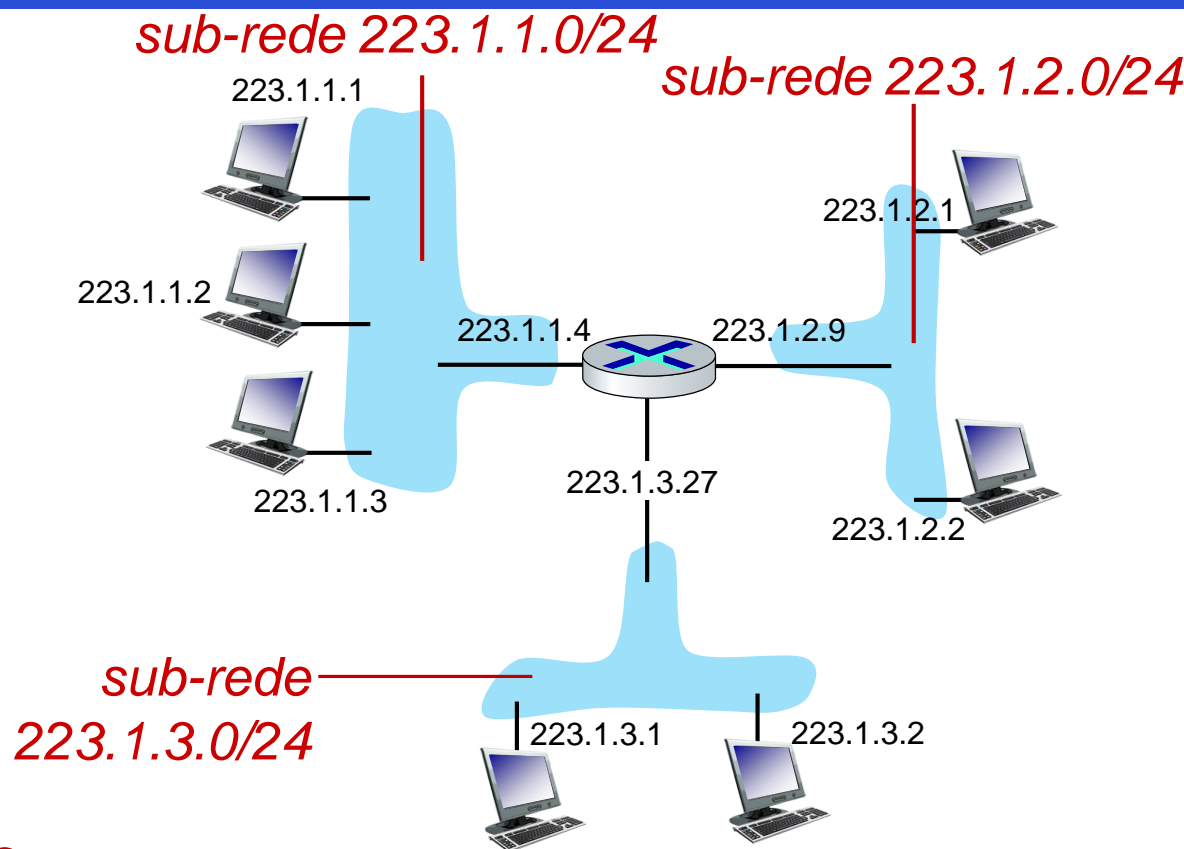


rede composta por 3 sub-redes

Sub-redes

Receita para definir sub-redes:

- Desanexar cada interface de seu host ou roteador, criando "ilhas" de redes isoladas
- Cada rede isolada é chamada de **sub-rede**

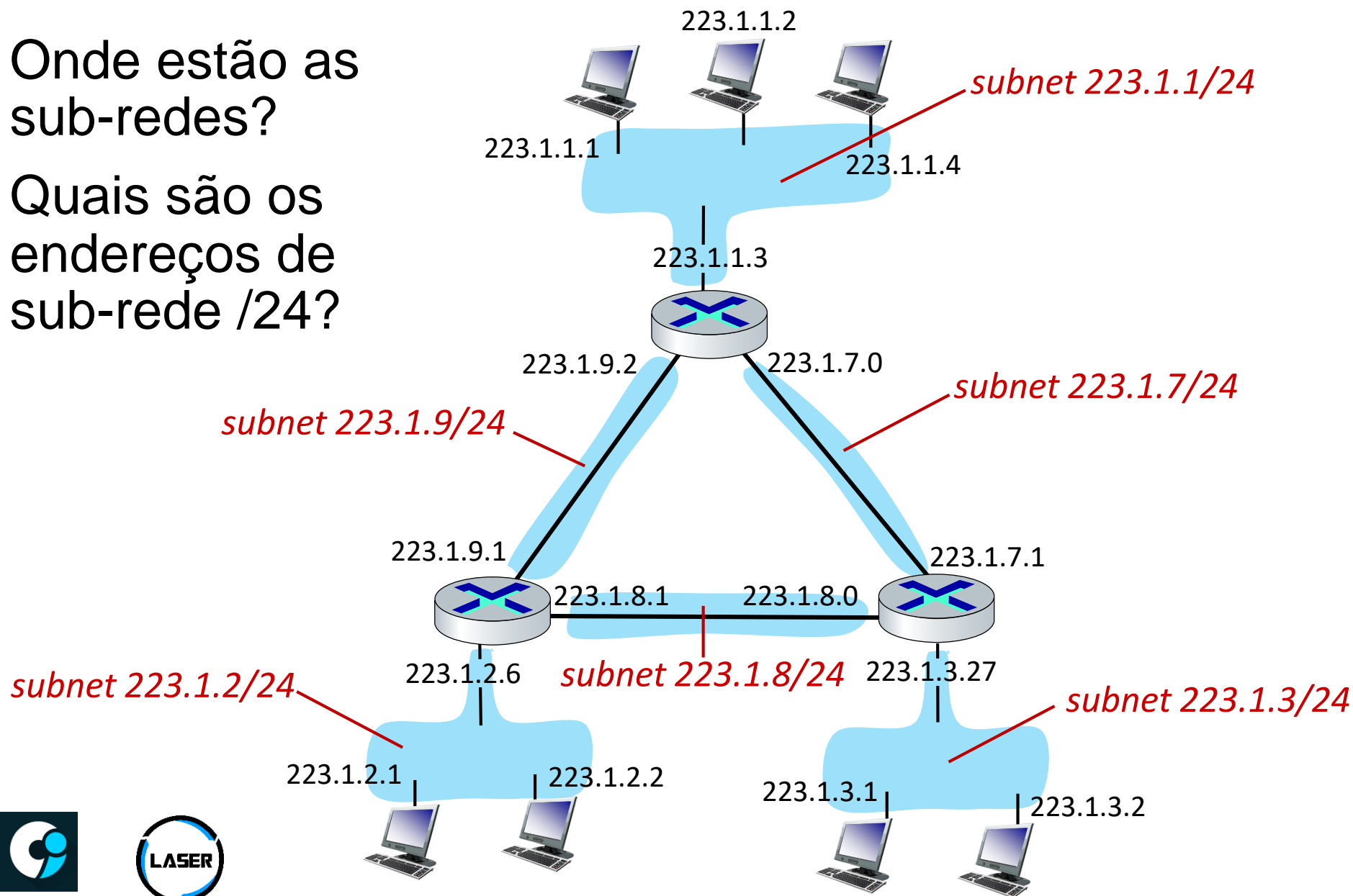


máscara de sub-rede: /24

(24 bits mais significativos: parte da sub-rede do endereço IP)

Sub-redes

- Onde estão as sub-redes?
- Quais são os endereços de sub-rede /24?



Endereçamento IP: CIDR

CIDR: Classless **I**nter**D**omain **R**outing (pronuncia “cider”) –
Roteamento Interdomínio sem classe

- parte da sub-rede do endereço de comprimento arbitrário
- formato de endereço: **a.b.c.d/x**, onde x é o número de bits na parte de sub-rede do endereço

← sub-rede → ← host →

11001000 00010111 00010000 00000000

200.23.16.0/23

Endereçamento IP: CIDR

- Estrutura de endereço IP

223.1.1.1/23 = $\underbrace{11011111}_{223} \underbrace{00000001}_1 \underbrace{00000001}_1 \underbrace{00000001}_1$

Endereçamento IP: CIDR

- Estrutura de endereço IP

223.1.1.1/23 = $\underbrace{11011111}_{223} \underbrace{00000001}_{1} \underbrace{0000000}_{?} \underbrace{* \text{ *****}}_{?}$

- Endereço de rede

223.1.0.0/23 = $\underbrace{11011111}_{223} \underbrace{00000001}_{1} \underbrace{00000000}_{0} \underbrace{00000000}_{0}$

- Endereço de broadcast

223.1.1.255/23 = $\underbrace{11011111}_{223} \underbrace{00000001}_{1} \underbrace{00000001}_{1} \underbrace{11111111}_{255}$

Endereços IP: como obter um?

- Na verdade, são **duas** perguntas:
 1. Como um host obtém o endereço IP dentro de sua rede (parte do host do endereço)?
 2. Como uma rede obtém o endereço IP para si mesma (parte da rede do endereço)
- Como o host obtém o endereço IP?
 - Codificado pelo sysadmin no arquivo de config (e.g., /etc/rc.config no UNIX)
 - **DHCP: Dynamic Host Configuration Protocol**: obtém o endereço dinamicamente do servidor
 - “plug-and-play”

DHCP: Dynamic Host Configuration Protocol

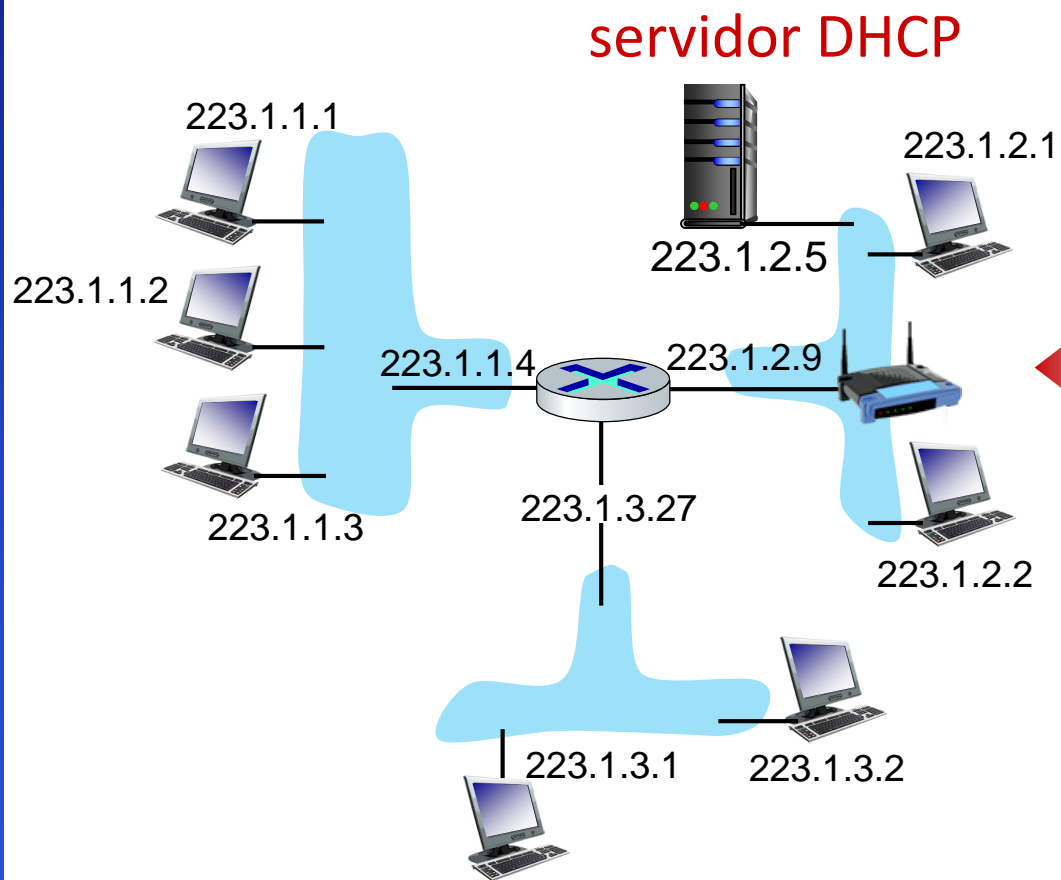
Objetivo: host obtém dinamicamente o endereço IP do servidor de rede quando ele “ingressa” na rede

- pode renovar sua “locação” no endereço em uso
- permite a reutilização de endereços (mantém o endereço apenas enquanto estiver conectado/ligado)
- suporte para vários usuários móveis que ingressam/saem da rede

- Visão geral do DHCP:

- hosts fazem broadcast de mensagem de **descoberta DHCP** (opcional)
- servidor DHCP responde com a mensagem de **oferta DHCP** (opcional)
- host solicita endereço IP: mensagem de **requisição DHCP**
- servidor DHCP envia o endereço: mensagem **ack DHCP**

Cenário cliente-servidor DHCP



Normalmente, o servidor DHCP está co-localizado no roteador, servindo todas as sub-redes às quais o roteador está conectado



cliente DHCP necessita de endereço nesta rede

Cenário cliente-servidor DHCP

servidor DHCP: 223.1.2.5

Descoberta DHCP

Broadcast: existe um servidor DHCP por aí?

cliente

Oferta DHCP

Broadcast: Eu sou um servidor DHCP! Segue um endereço IP que você pode usar

Requisição DHCP

Broadcast: OK. Gostaria de usar este endereço IP!

ACK DHCP

Broadcast: OK. Você tem esse endereço IP!

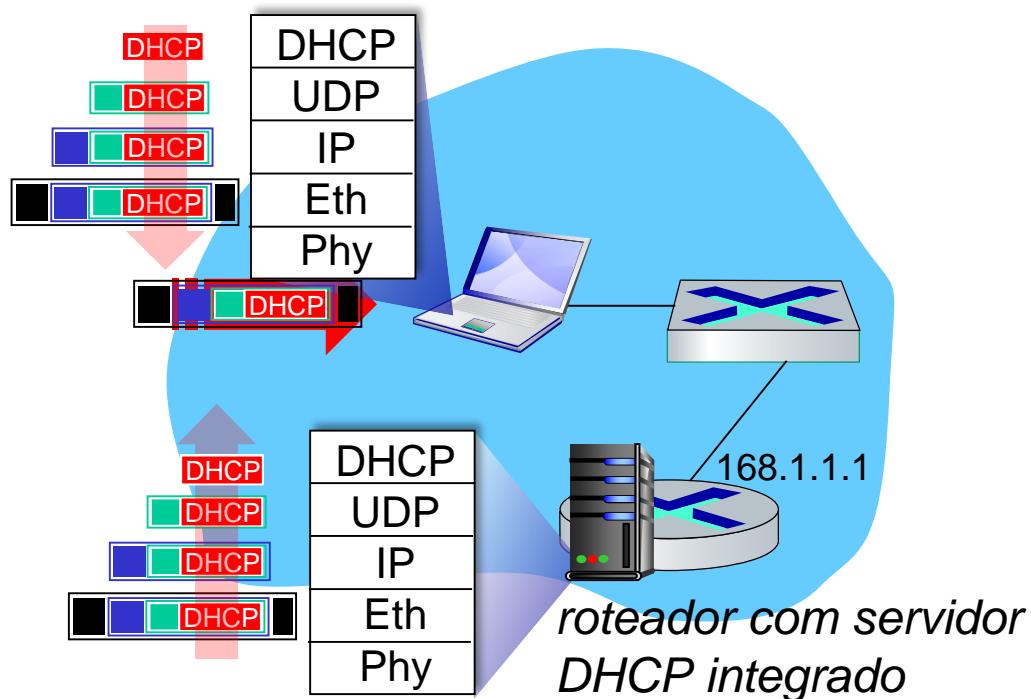
As duas etapas acima podem ser ignoradas "se um cliente se lembrar e desejar reutilizar um endereço de rede alocado anteriormente" [RFC 2131]

DHCP: mais do que endereços IP

O DHCP pode retornar mais do que apenas o endereço IP alocado na sub-rede:

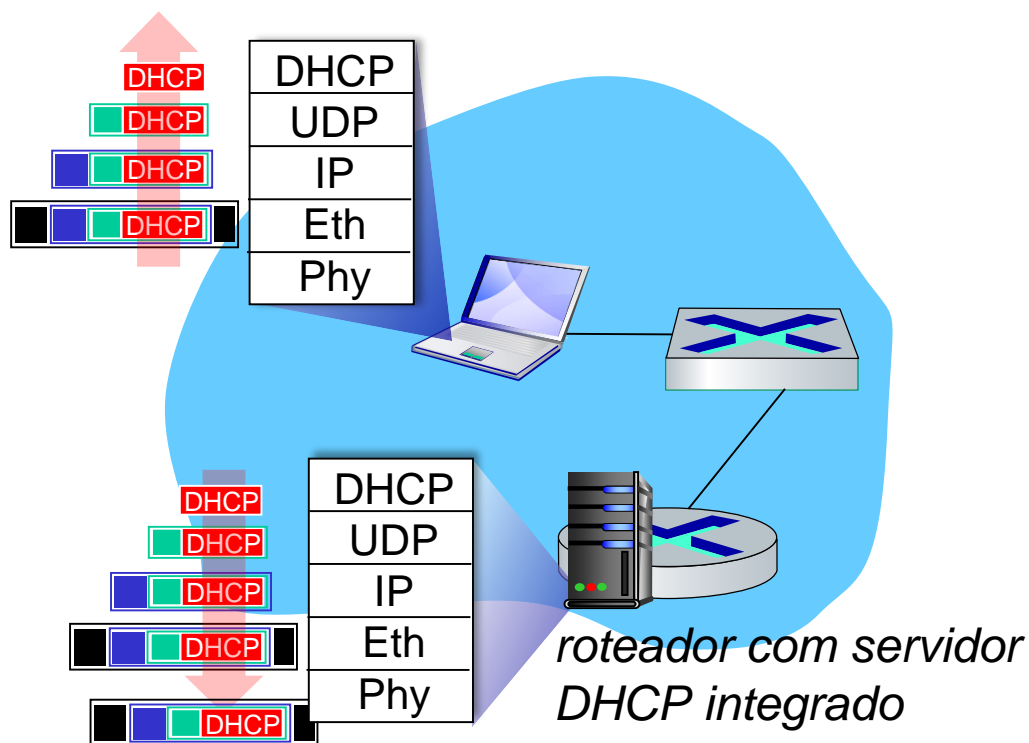
- endereço do roteador de primeiro salto para o cliente
- nome e endereço IP do servidor DNS
- máscara de rede (indicando partes, redes x host, do endereço)

DHCP: exemplo



- Laptop usará DHCP para obter endereço IP, endereço do roteador de primeiro salto, endereço do servidor DNS.
- Mensagem DHCP REQUEST encapsulada em UDP, encapsulada em IP, encapsulada em Ethernet
- Transmissão de quadro Ethernet (dest: FFFFFFFFFFFFFFFF) na LAN, recebida no roteador que executa o servidor DHCP
- Ethernet de-muxe para IP de-muxe, UDP de-muxe para DHCP

DHCP: exemplo



- O servidor DHCP cria o ACK DHCP contendo o endereço IP do cliente, o endereço IP do roteador de primeiro salto para o cliente, o nome e o endereço IP do servidor DNS
- resposta encapsulada do servidor DHCP encaminhada para o cliente, de-muxing até DHCP no cliente
- o cliente agora sabe seu endereço IP, nome e endereço IP do servidor DNS, endereço IP de seu roteador de primeiro salto

Endereços IP: como obter um?

Pergunta: como a rede obtém a parte da sub-rede do endereço IP?

Resposta: obtém a parte alocada do espaço de endereços do ISP

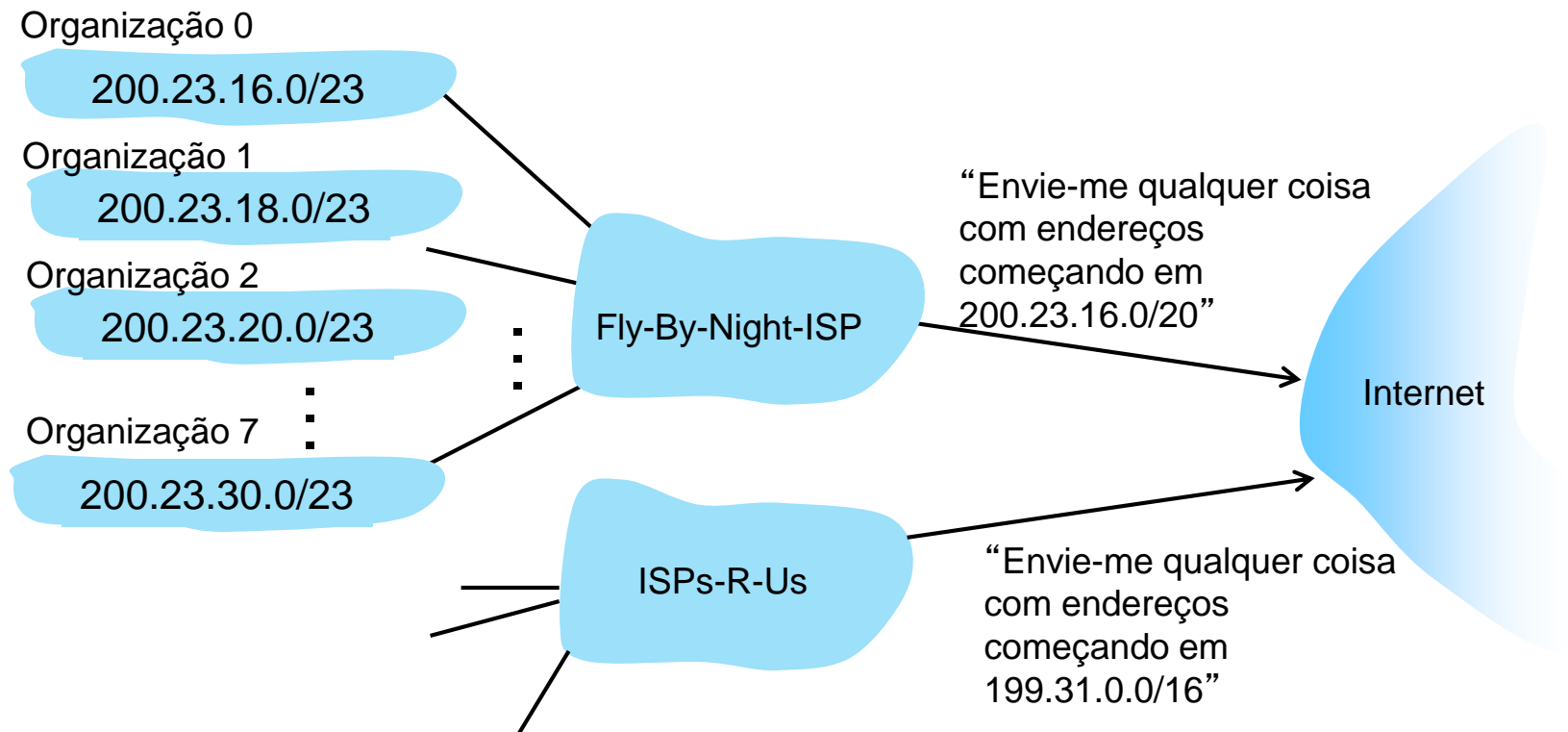
Bloco do ISP 11001000 00010111 00010000 00000000 200.23.16.0/20

O ISP pode então alocar seu espaço de endereço em 8 blocos:

| | | | |
|---------------|-----------------------------------|----------|----------------|
| Organização 0 | <u>11001000 00010111 00010000</u> | 00000000 | 200.23.16.0/23 |
| Organização 1 | <u>11001000 00010111 00010010</u> | 00000000 | 200.23.18.0/23 |
| Organização 2 | <u>11001000 00010111 00010100</u> | 00000000 | 200.23.20.0/23 |
| ... | | | |
| Organização 7 | <u>11001000 00010111 00011110</u> | 00000000 | 200.23.30.0/23 |

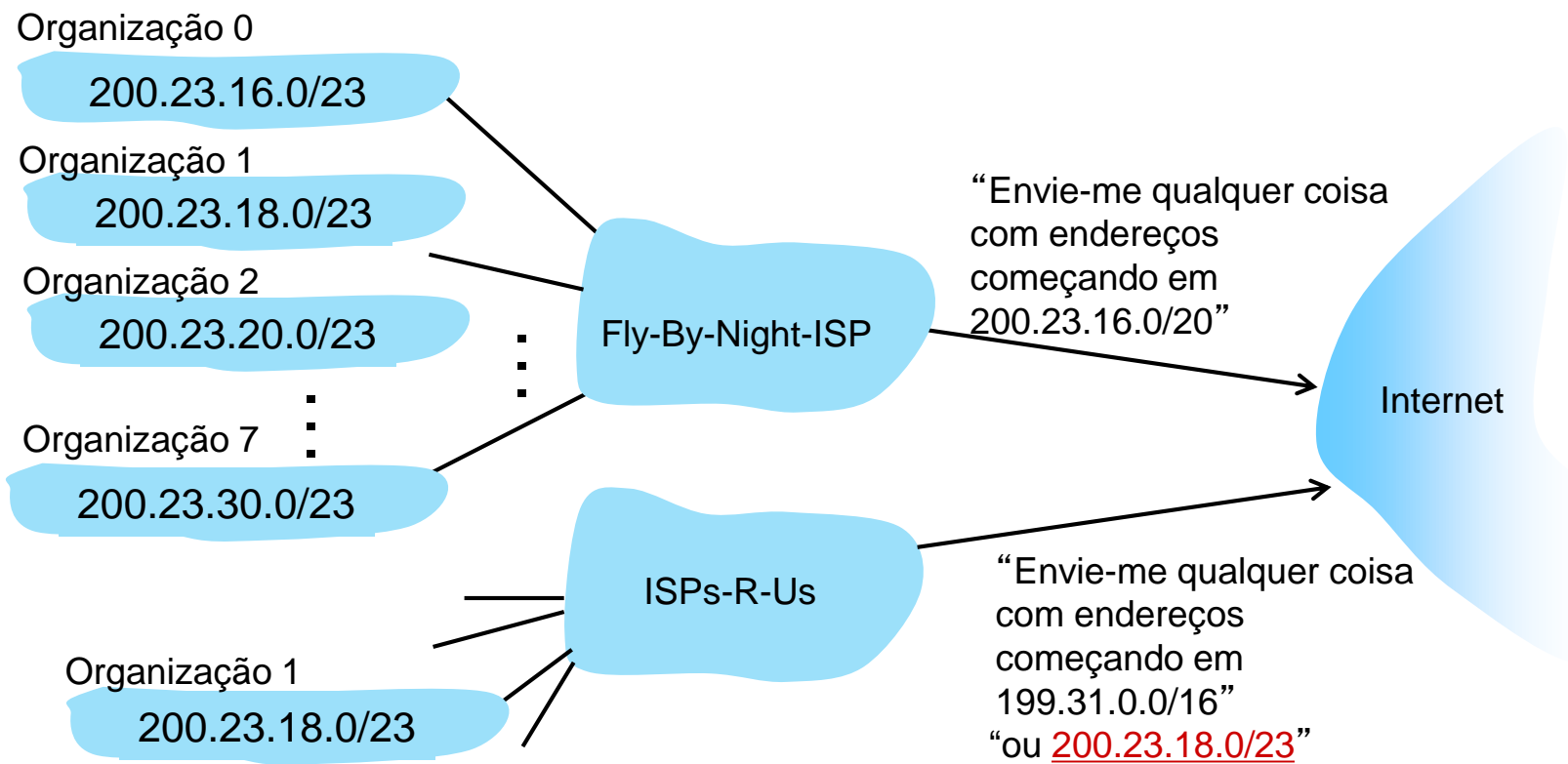
Endereços hierárquico: agregação de rotas

O endereçamento hierárquico permite anúncio eficiente das informações de roteamento:



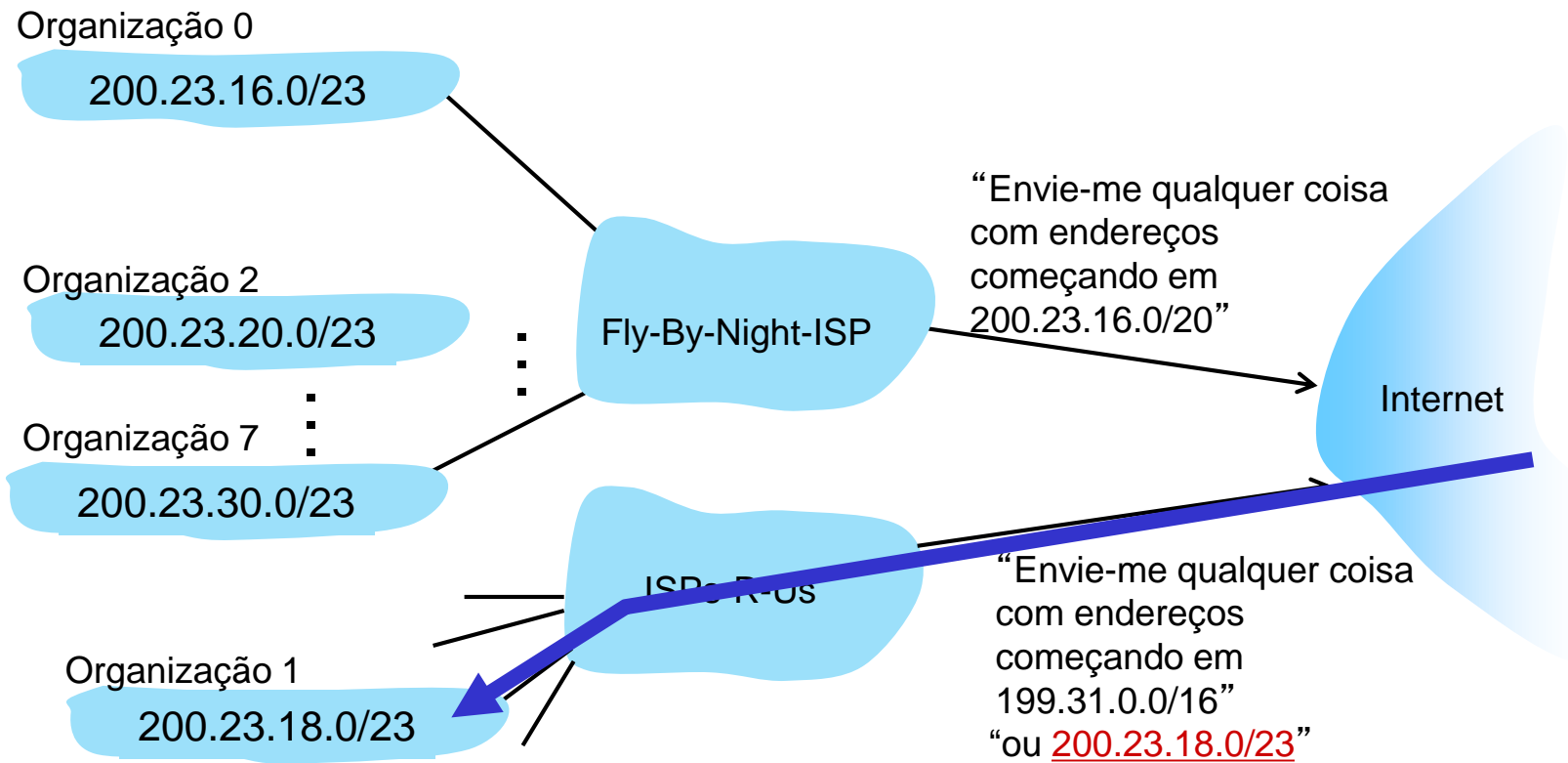
Endereços hierárquico: rotas mais específicas

- A organização 1 muda de Fly-By-Night-ISP para ISPs-R-Us
- ISPs-R-Us agora anuncia rota mais específica para a Organização 1



Endereços hierárquico: rotas mais específicas

- A organização 1 muda de Fly-By-Night-ISP para ISPs-R-Us
- ISPs-R-Us agora anuncia rota mais específica para a Organização 1



Endereçamento IP: considerações finais

Pergunta: como um ISP obtém um bloco de endereços?

Resposta: ICANN: Internet Corporation for Assigned Names and Numbers

<http://www.icann.org>

- aloca endereços IP por 5 ***regional registries (RRs)*** (os quais podem fazer alocações para registros locais)
- gerencia zona raiz do DNS, incluindo delegação do gerenciamento de TLDs individuais (.com, .edu, ...)

Endereçamento IP: considerações finais

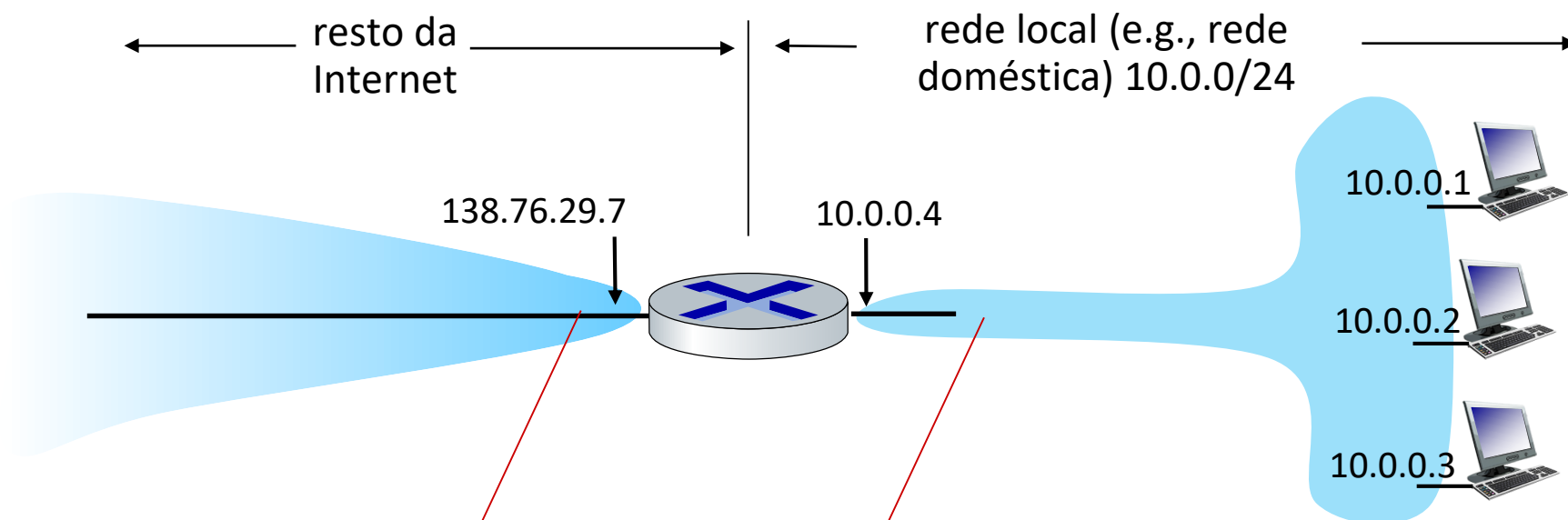
Pergunta: existem endereços IP de 32 bits suficientes?

- ICANN alocou o último bloco de endereços IPv4 para RRs em 2011
- NAT (estudado na próxima aula) ajuda com o problema de exaustão do espaço de endereços IPv4
- IPv6 possui espaço de endereços de 128 bits

“Quem diabos sabia o quanto de espaço de endereços precisávamos?” Vint Cerf (refletindo sobre a decisão de tornar o tamanho do endereço IPv4 32 bits)

NAT: conversão de endereços de rede

NAT (Network Address Translation): todos os dispositivos na rede local compartilham apenas **um** endereço IPv4 no que diz respeito ao mundo externo



*todos os datagramas que **saem** da rede local têm o **mesmo** endereço IP NAT origem: 138.76.29.7, mas nº de porta origem diferentes*

datagramas com origem ou destino nesta rede têm endereço 10.0.0/24 para origem, destino (como de costume)

NAT: conversão de endereços de rede

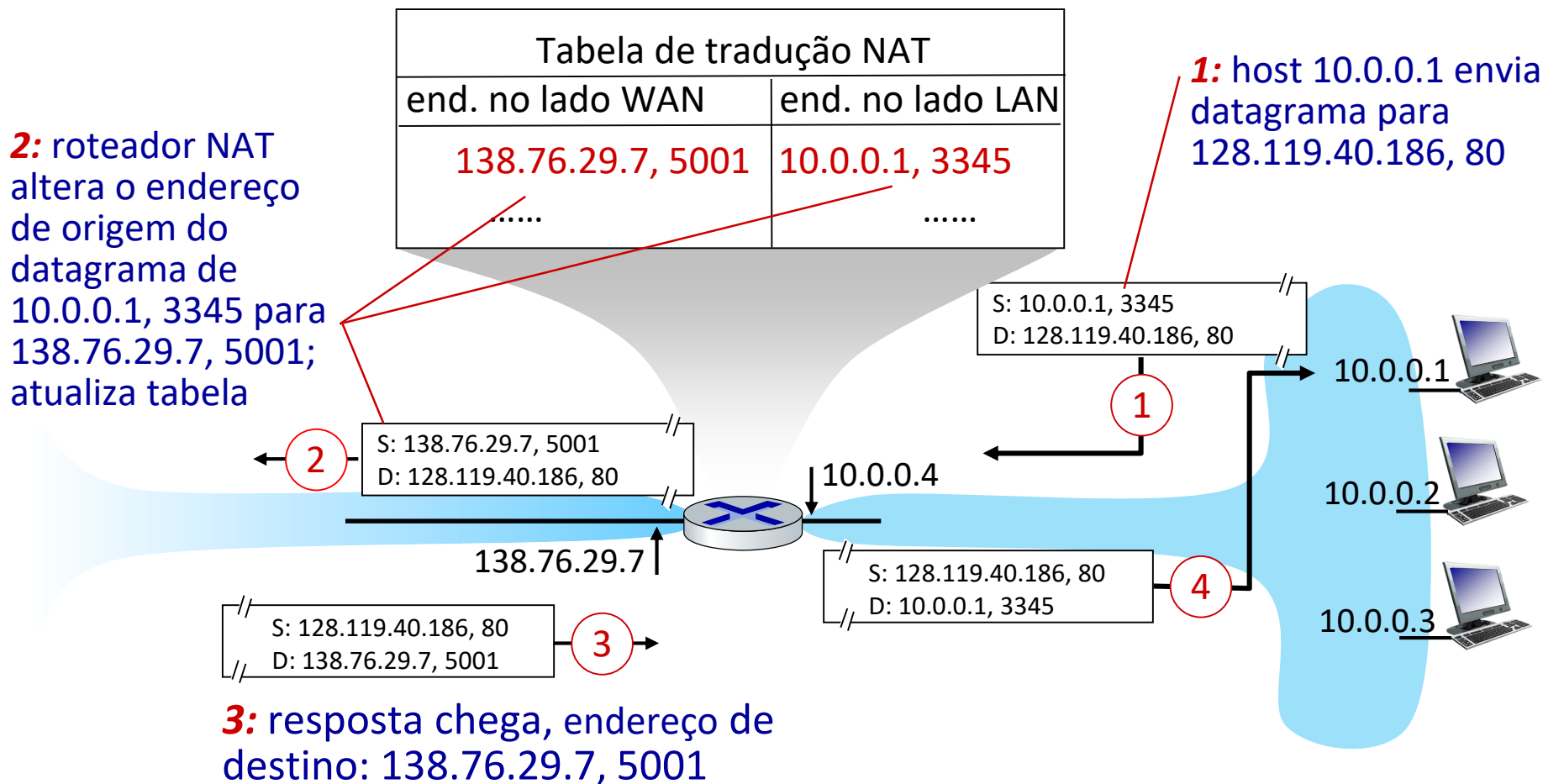
- Todos os dispositivos na rede local têm endereços de 32 bits em um espaço de endereço IP “privado” (prefixos 10/8, 172.16/12, 192.168/16) que só podem ser usados na rede local
- Vantagens:
 - Necessário apenas **um** endereço IP do ISP para todos os dispositivos
 - Pode alterar endereço de host na rede sem notificar o mundo externo
 - Pode alterar o ISP sem alterar endereços de dispositivos de rede local
 - Segurança: dispositivos dentro da rede local não são endereçáveis/visíveis diretamente, pelo mundo exterior

NAT: conversão de endereços de rede

Implementação: o roteador NAT deve (de forma transparente):

- **Datagramas de saída: substituir** (endereço IP de origem, número de porta) de cada datagrama de saída para (endereço IP NAT, novo número de porta)
 - Clientes/servidores remotos responderão usando (endereço IP NAT, novo número de porta) como endereço de destino
- **Lembrar (na tabela de tradução NAT)** cada par (endereço IP de origem, número de porta) para (endereço IP NAT, novo número de porta)
- **Datagramas de entrada: substituir** (endereço IP NAT, novo número de porta) nos campos de destino de cada datagrama de entrada com correspondente (endereço IP de origem, número de porta) armazenado na tabela NAT

NAT: conversão de endereços de rede



NAT: conversão de endereços de rede

- NAT é controverso:
 - roteadores “devem” processar apenas até a camada 3
 - “escassez” de endereços deve ser resolvida pelo IPv6
 - viola o argumento de ponta a ponta (manipulação do número de porta pelo dispositivo da camada de rede)
 - NAT Transversal (NAT-T): e se um cliente externo quiser se conectar ao servidor atrás de um NAT? (e.g. P2P, videoconferência, jogos eletrônicos)
- Mas o NAT veio para ficar:
 - amplamente utilizado em redes domésticas e institucionais, redes celulares 4G/5G

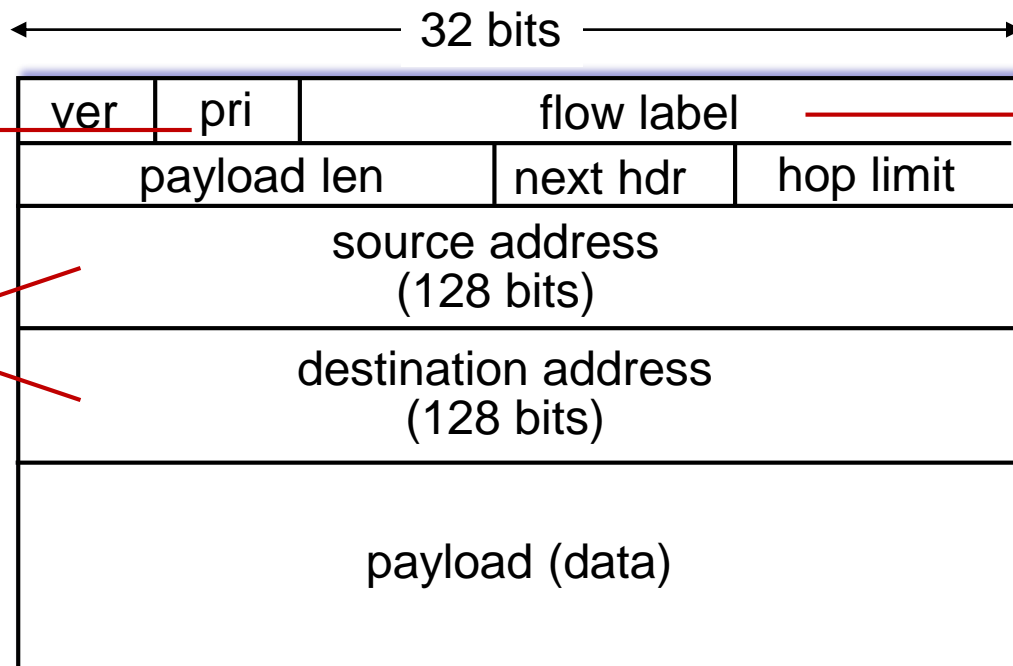
IPv6: motivação

- **Motivação inicial:** o espaço de endereços IPv4 de 32 bits seria esgotado
- Motivação adicional:
 - velocidade de processamento/encaminhamento: cabeçalhos de comprimento fixo de 40 bytes
 - permitir o tratamento de diferentes camadas de rede de “fluxos”

Formato de datagrama IPv6

prioridade: identificar
prioridade entre
datagramas no fluxo

endereços IPv6
de 128-bit



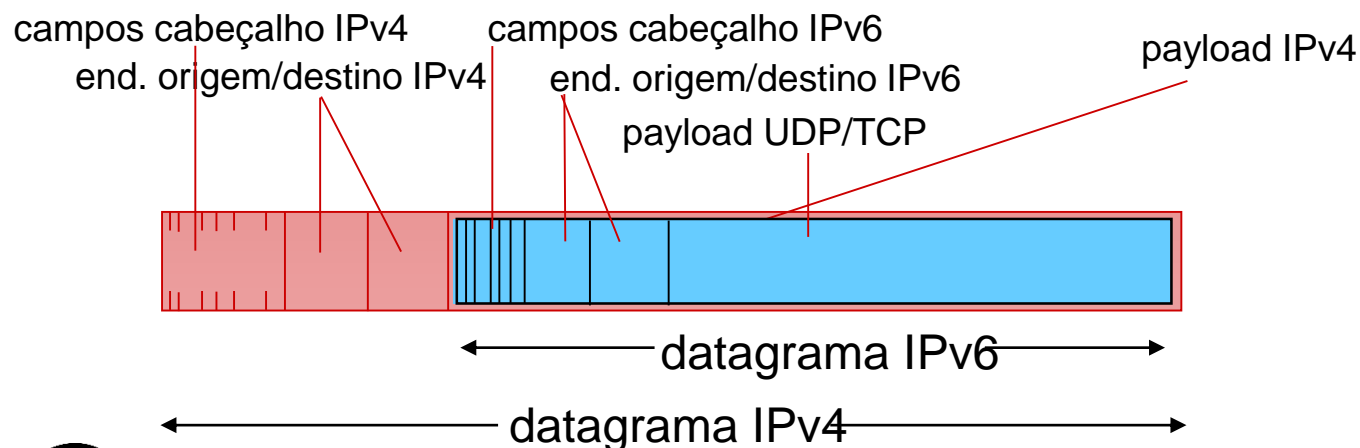
**etiqueta de
fluxo:**
identifique
datagramas
no mesmo
"fluxo".
(conceito de
"fluxo" não
bem definido).

O que está faltando (em comparação com o IPv4):

- sem soma de verificação (para acelerar o processamento em roteadores)
- sem fragmentação/remontagem
- sem opções (disponível como protocolo de camada superior)

Transição do IPv4 para o IPv6

- Nem todos os roteadores podem ser atualizados simultaneamente
 - Como a rede funcionará com roteadores IPv4 e IPv6 mistos?
- Tunelamento: datagrama IPv6 transportado como carga útil no datagrama IPv4 entre roteadores IPv4 (“pacote dentro de um pacote”)
 - Tunelamento usado extensivamente em outros contextos (4G/5G)



Tunelamento e encapsulamento

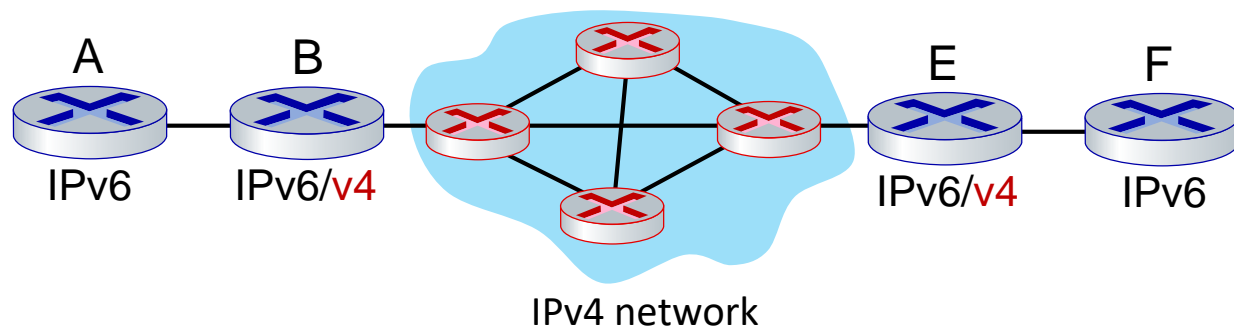
Ethernet
conectando dois
roteadores IPv6:



quadro da camada de enlace

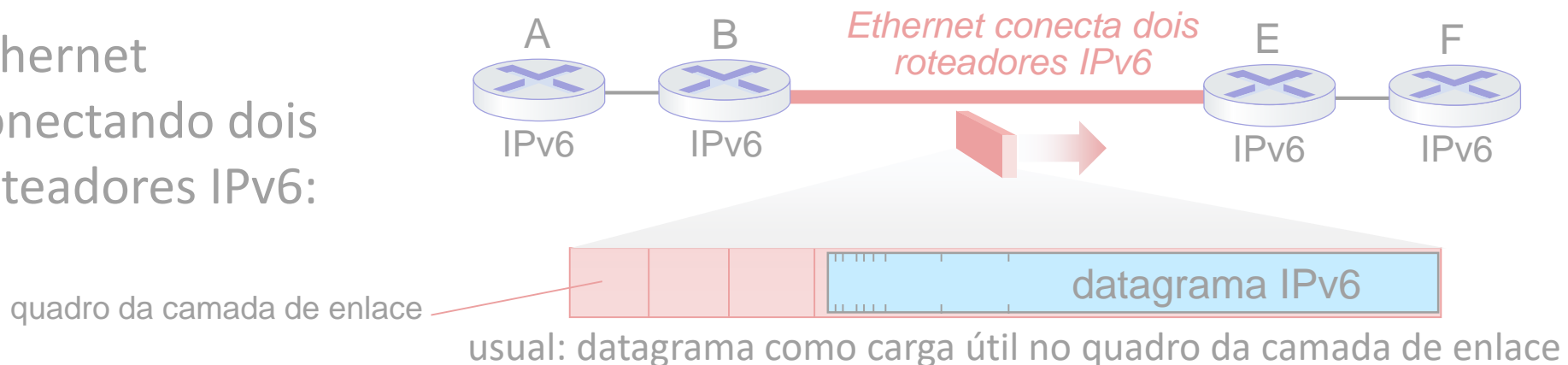
usual: datagrama como carga útil no quadro da camada de enlace

Rede IPv4
conectando dois
roteadores IPv6

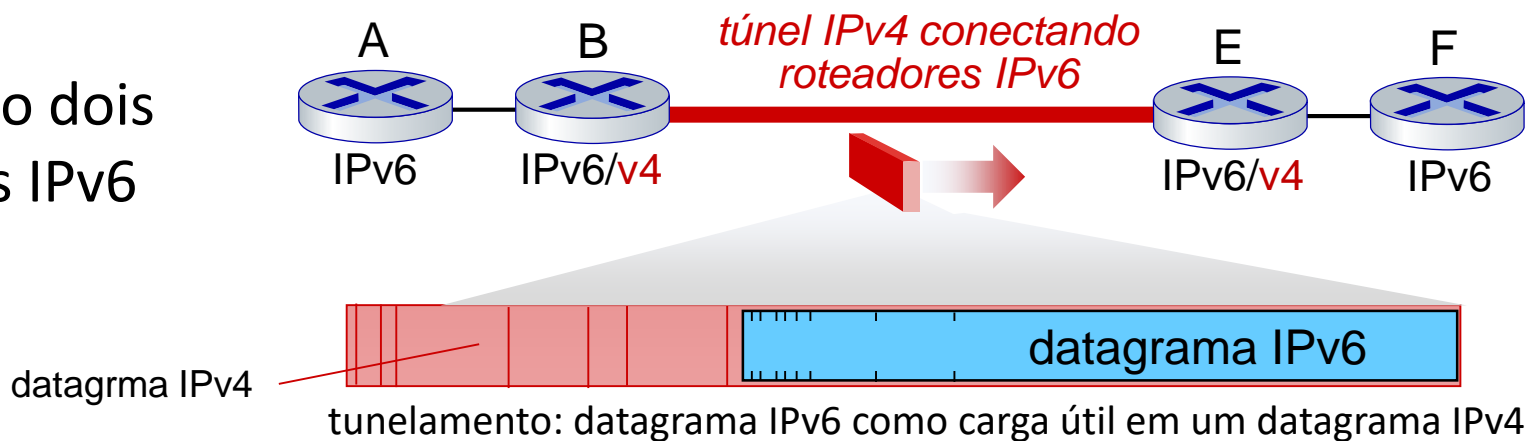


Tunelamento e encapsulamento

Ethernet
conectando dois
roteadores IPv6:



Túnel IPv4
conectando dois
roteadores IPv6

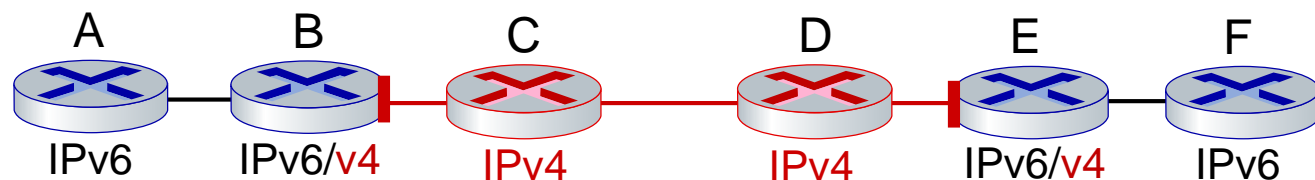


Tunelamento

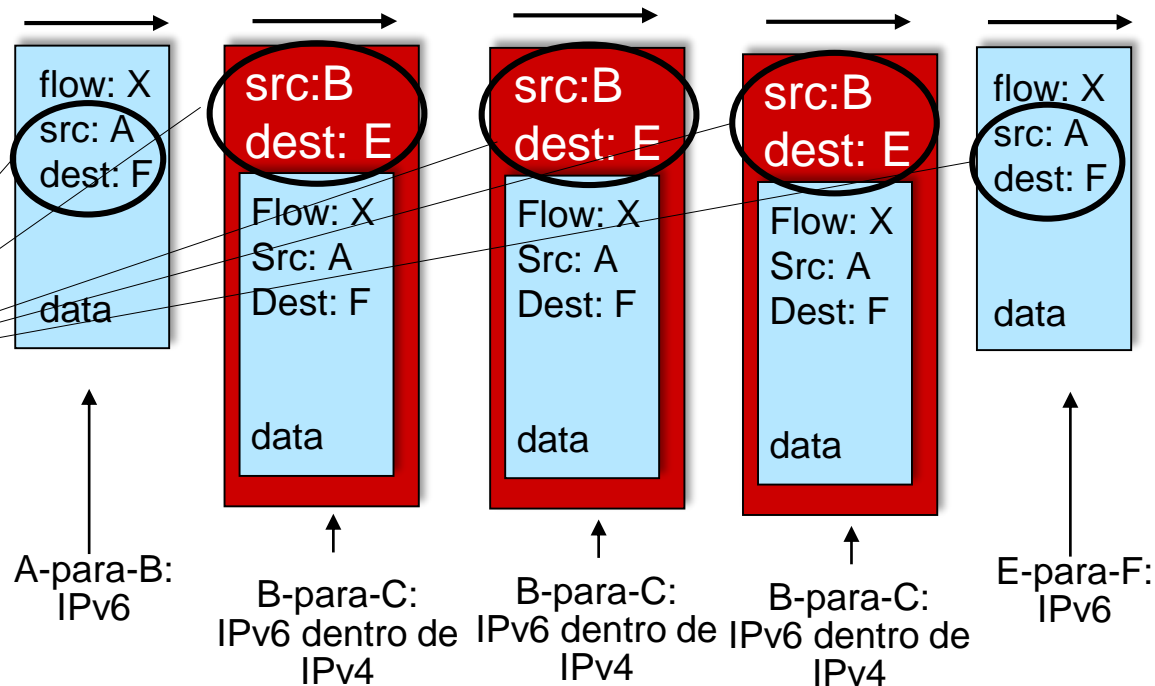
visão lógica:



visão física:



Observe os
endereços de
origem e destino!

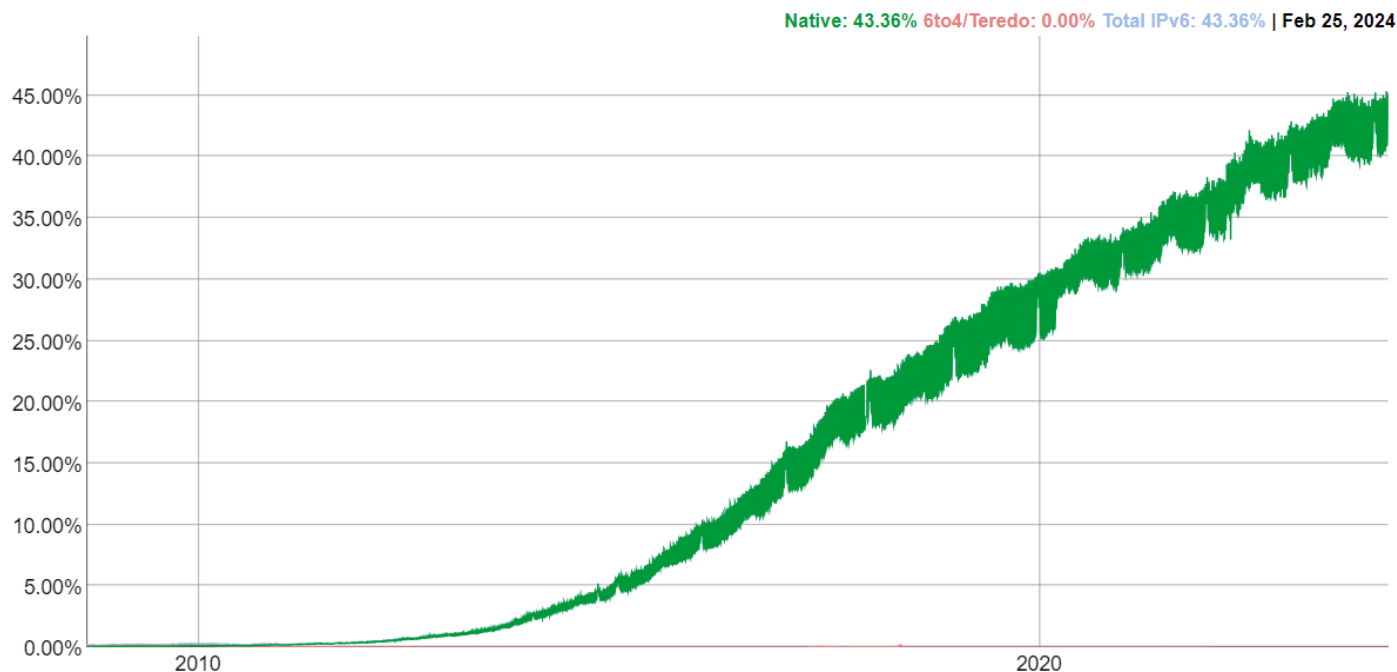


IPv6: adoção

- Google: ~40% de clientes acessam serviços via IPv6 (2023)
- NIST: 1/3 de todos os domínios do governo dos EUA são compatíveis com IPv6

IPv6 Adoption

We are continuously measuring the availability of IPv6 connectivity among Google users. The graph shows the percentage of users that access Google over IPv6.



IPv6: adoção

- Google: ~40% de clientes acessam serviços via IPv6 (2023)
- NIST: 1/3 de todos os domínios do governo dos EUA são compatíveis com IPv6
- Tempo de instalação/uso longo (longo!):
 - 28 anos e contando!
 - Pense nas mudanças em nível de aplicação nos últimos 28 anos: WWW, mídia social, streaming de mídia, gaming, telepresença, ...
 - Por quê?

Unidade 4

Camada de Rede

Baseado nos slides elaborados por J. F. Kurose e K. W. Ross

Perguntas?

