Instituto de computação

Programa de Pós-Graduação
Instituto de Computação
Universidade Federal Fluminense

Aulas 14, 15 e 16

Camada de Rede

Conceitos, modelos de serviço e IP

Igor Monteiro Moraes

Redes de Computadores

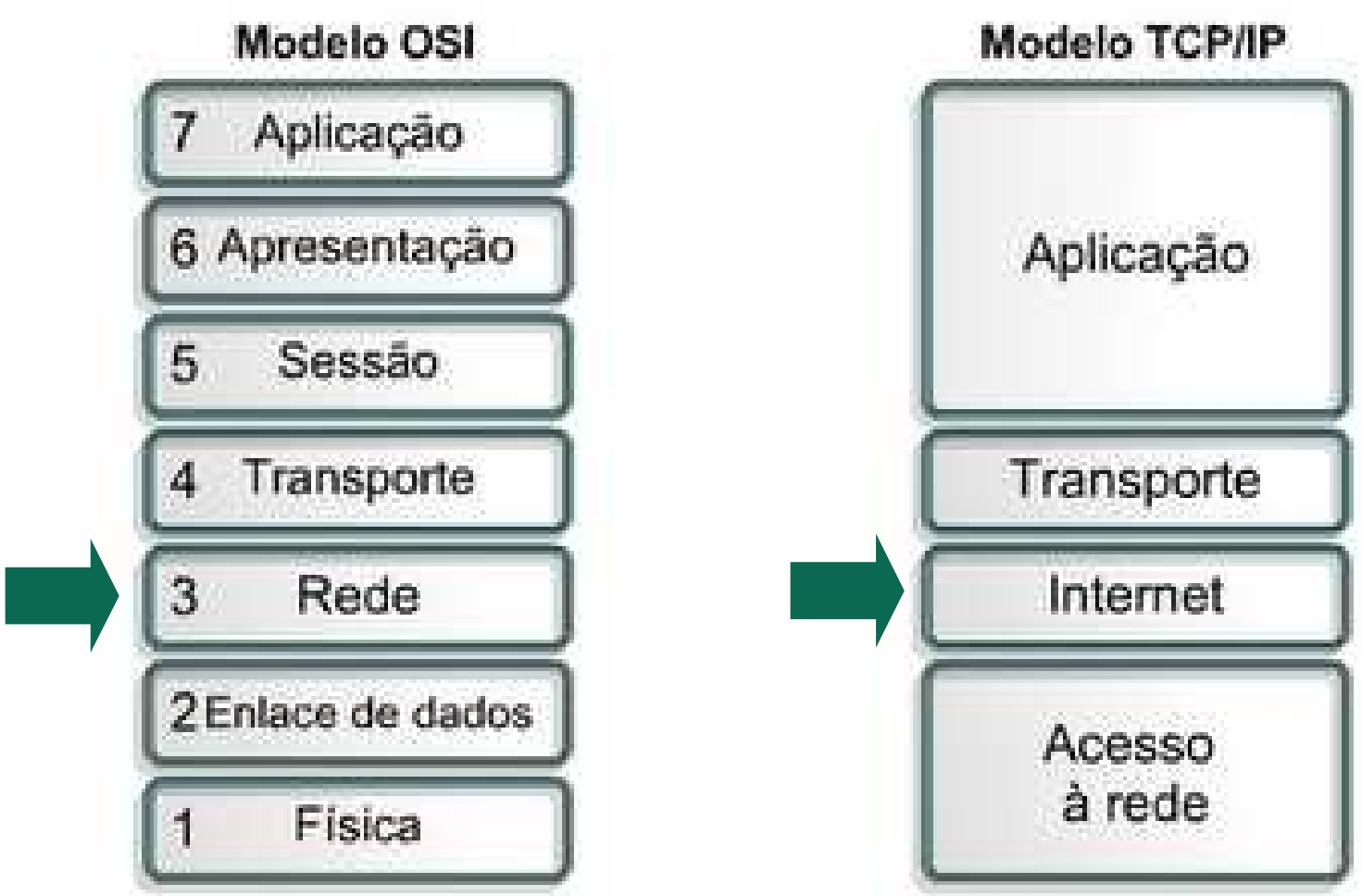
ATENÇÃO!

Instituto de Computação

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

Camada de Rede

Instituto de computação



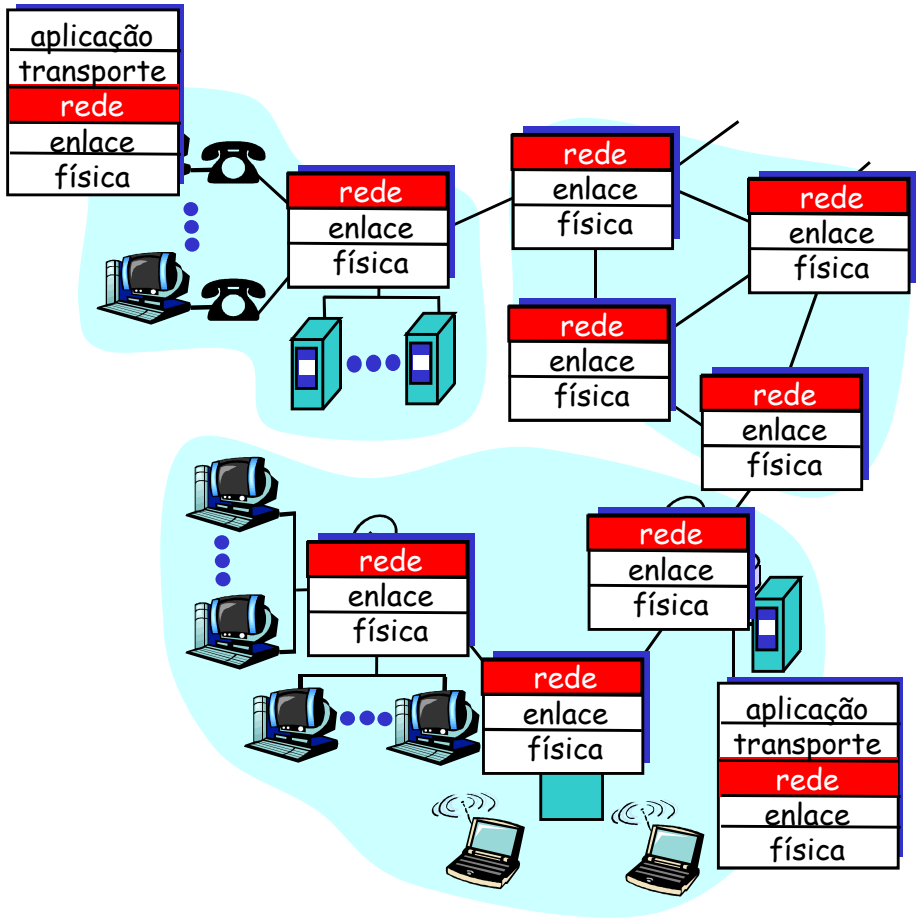
Camada de Rede

Instituto de computação

- Responsável por
  - **Determinar o melhor caminho** para o envio dos pacotes
    - É função dos protocolos de roteamento
  - **Encaminhar** os pacotes até o destino
    - É função do protocolo IP
  - **Interconectar** redes de diferentes tecnologias
    - É função do protocolo IP

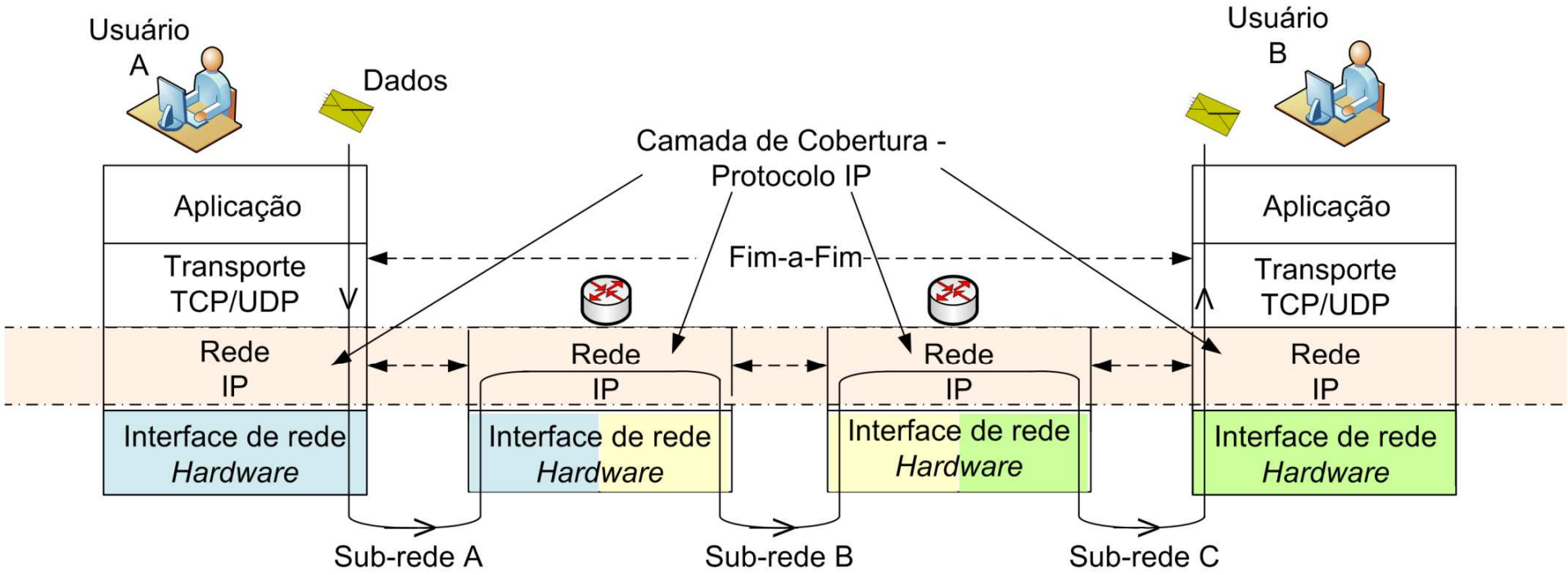
Camada de Rede

- Protocolos da camada de rede
  - Executados nos **sistemas finais** e nos **roteadores**



Camada de Rede

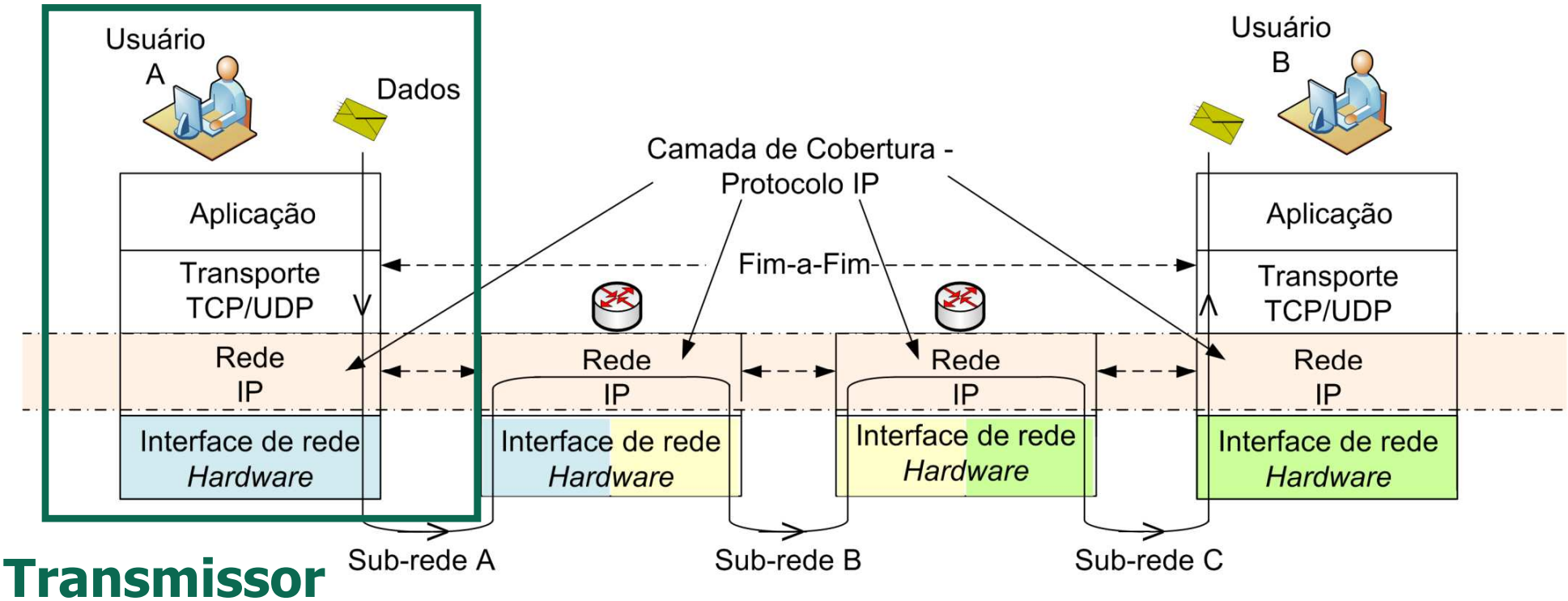
- Protocolos da camada de rede
  - Executados nos **sistemas finais** e nos **roteadores**



Transporta segmentos da estação remetente à receptora

Camada de Rede

- Protocolos da camada de rede
  - Executados nos **sistemas finais** e nos **roteadores**

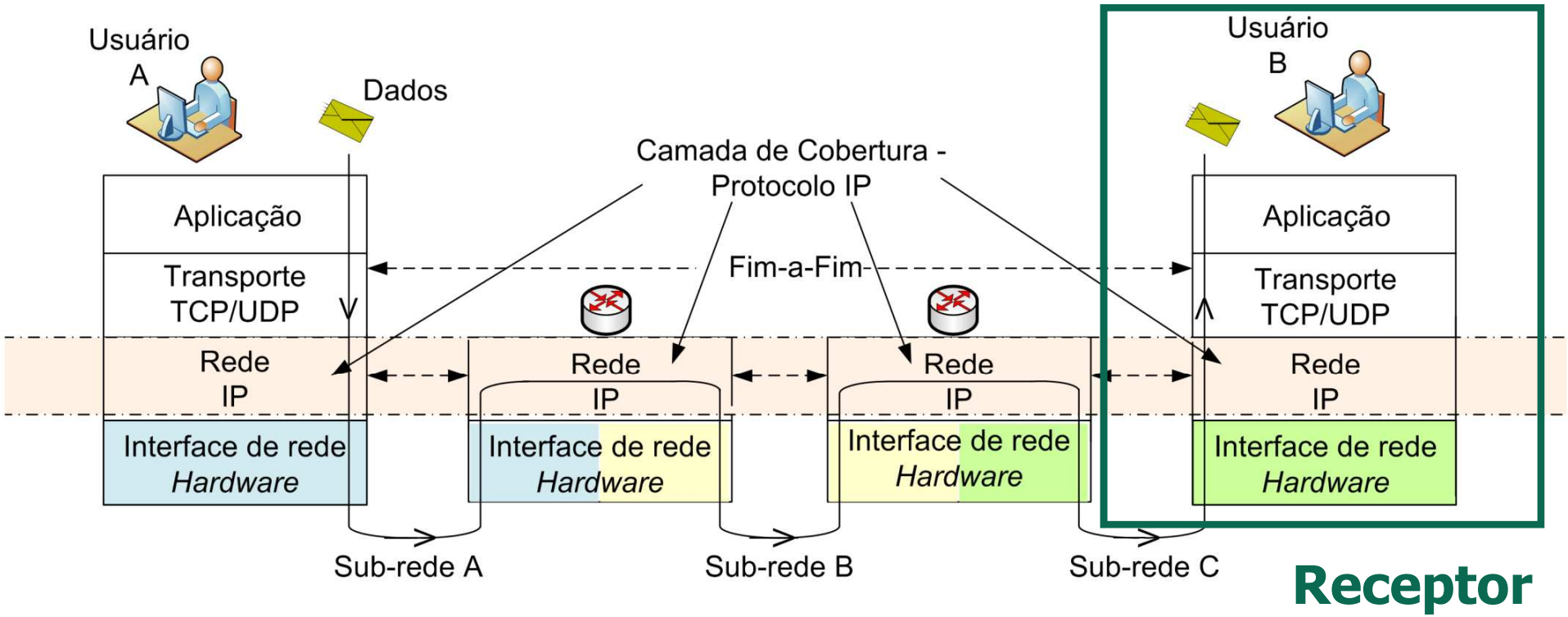


Transmissor

No lado transmissor, encapsula segmentos dentro de datagramas

Camada de Rede

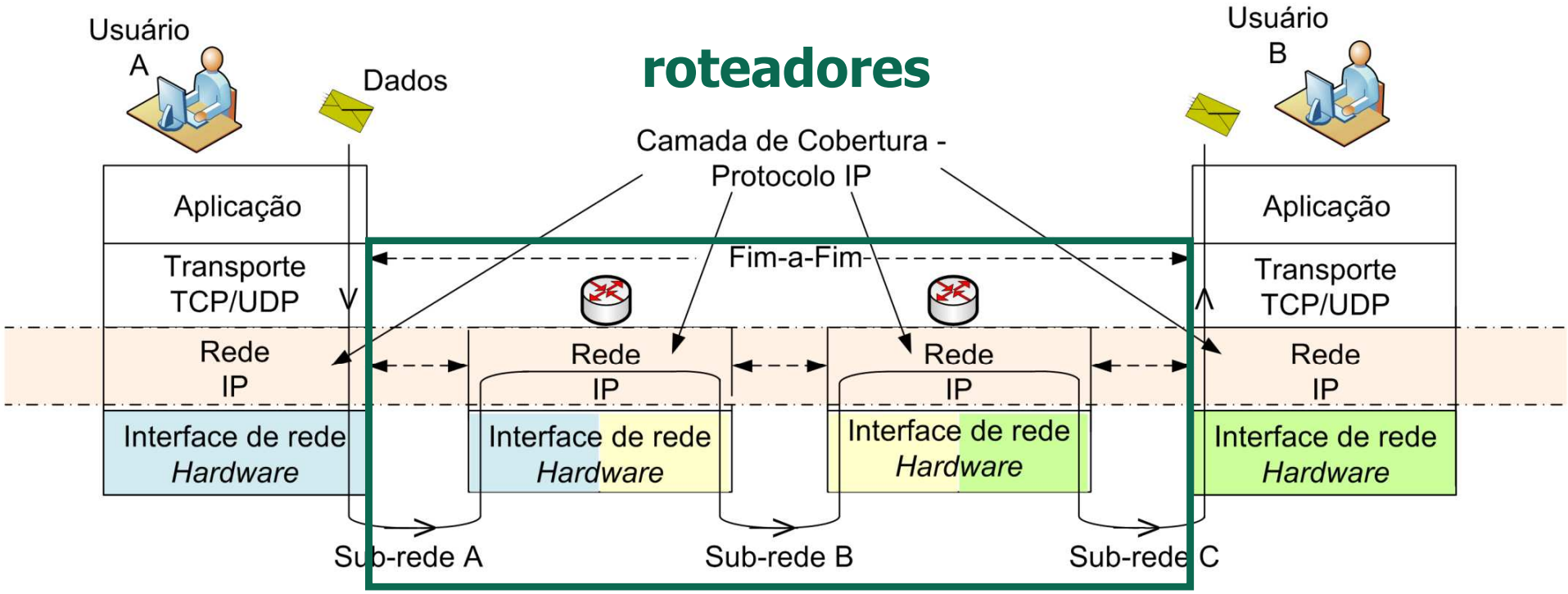
- Protocolos da camada de rede
  - Executados nos **sistemas finais** e nos **roteadores**



No lado receptor, entrega os segmentos para a camada de transporte

Camada de Rede

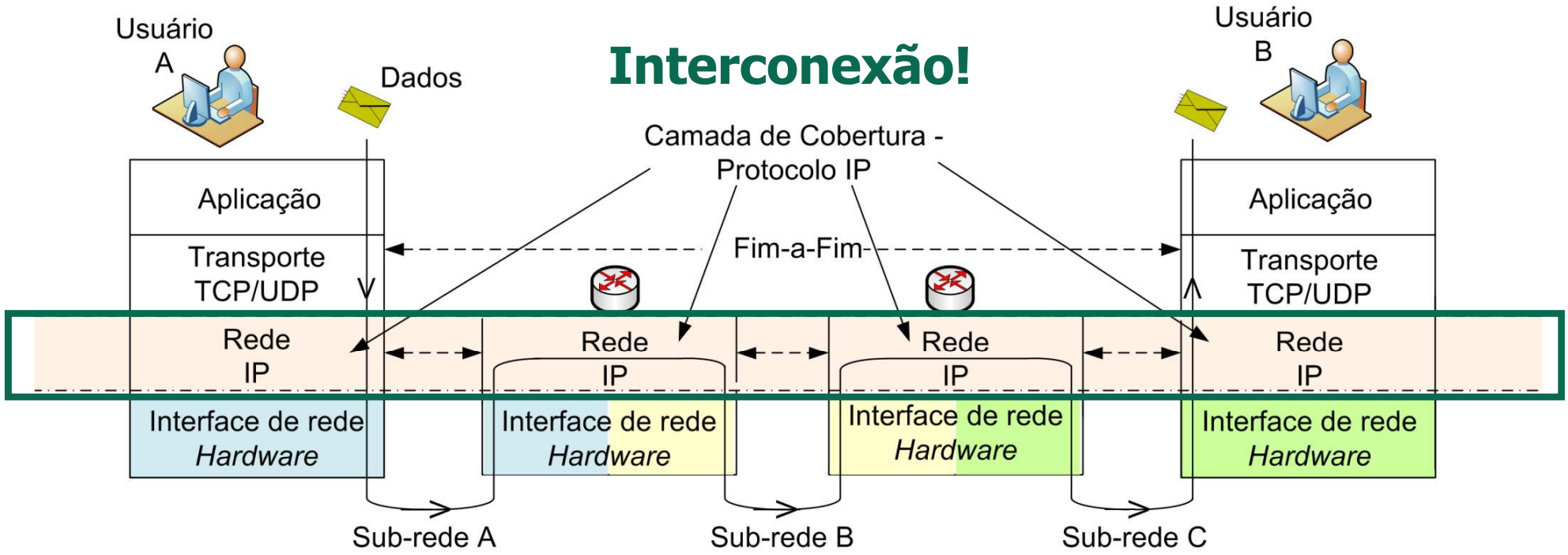
- Protocolos da camada de rede
  - Executados nos **sistemas finais** e nos **roteadores**



Roteadores examinam campos de cabeçalho de **todos** os datagramas IP que passam por eles

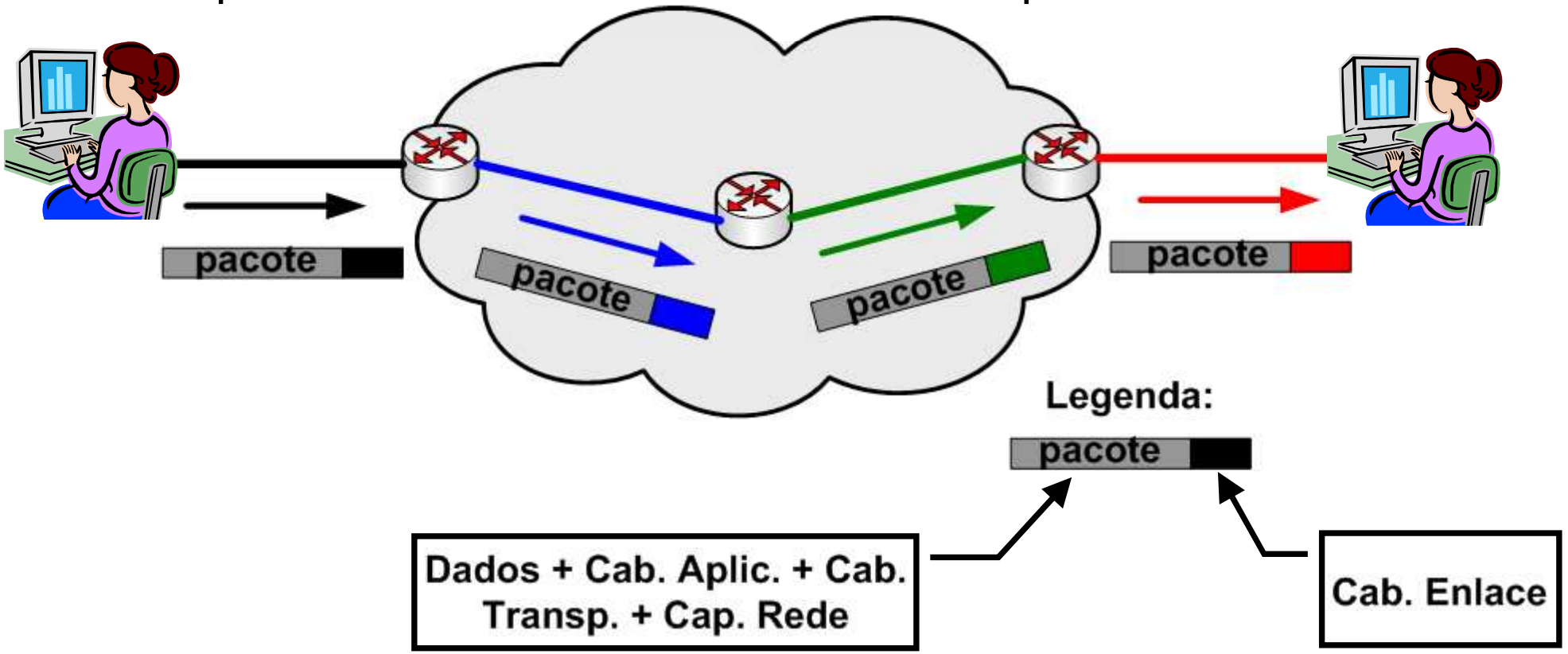
Camada de Rede

- Protocolos da camada de rede
  - Executados nos **sistemas finais** e nos **roteadores**



Transparência

- Transparência sintática
  - Pacotes são transferido da origem ao destino sem que a rede modifique os dados
    - Apenas erros de transmissão modificam pacotes

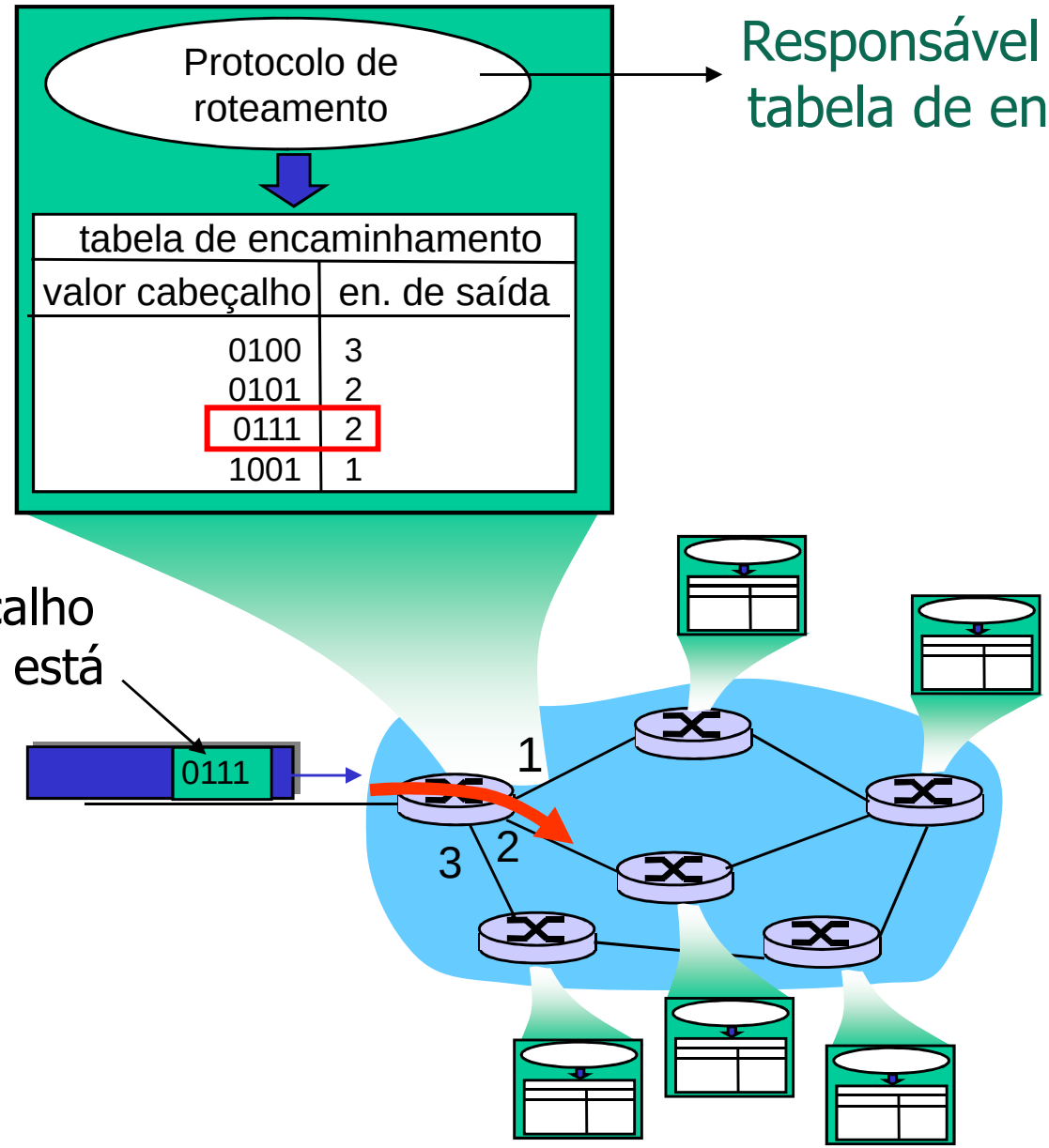


Encaminhamento x Roteamento

Instituto de computação

- Encaminhamento
  - "Mover" pacotes de uma entrada do roteador para a saída apropriada
    - É função do protocolo IP
- Roteamento
  - Determinar a rota a ser seguida pelos pacotes da fonte até o destino
    - É função dos protocolos de roteamento

Encaminhamento x Roteamento



Modelos de Serviço

Instituto de computação

- Definem as características do transporte de pacotes fim-a-fim entre transmissor e receptor

- Para pacotes individuais
  - Entrega garantida
    - Um pacote irá chegar ao destino
  - Entrega garantida com atraso limitado
    - Ex.: Um pacote irá chegar com atraso menor que 100 ms

Modelos de Serviço

Instituto de Computação

- Para fluxos de pacotes
  - Entrega ordenada de pacotes
  - Largura de banda mínima garantida
  - *Jitter* máximo garantido
  - Serviços de segurança
    - Usando uma chave secreta de sessão o transmissor poderia cifrar o conteúdo de todos os pacotes enviados para o destinatário
- Na Internet
  - Apenas um protocolo, o IP
    - Apenas um serviço oferecido → **melhor esforço**

Melhor Esforço

Instituto de computação

- Roteadores se esforçam ao máximo para entregar os pacotes
  - Da melhor maneira possível e sem distinção
- Nós simples e de baixo custo – sem estados na rede
    - Encaminhamento de pacote independente um dos outros
    - Sem reserva de recursos, recuperação de erros, garantia de acesso
    - Atraso dependente do tamanho da fila
    - Sem garantia de entrega do pacote ao destino
      - Pacote é descartado se fila cheia

Modelos de Serviço

Arquitetura de Rede	Modelo de serviço	Garantias ?				Indicação de congestion.?
		Banda	Perdas	Ordem	Tempo	
Internet	melhor esforço	nenhuma	não	não	não	não (inferido via perdas)
ATM	CBR	taxa constante	sim	sim	sim	sem congestion.
ATM	VBR	taxa garantida	sim	sim	sim	sem congestion.
ATM	ABR	mínima garantida	não	sim	não	sim
ATM	UBR	nenhuma	não	sim	não	não

Intserv e Diffserv → extensões do "melhor esforço"

Serviços da Camada de Rede

Instituto de computação

- Orientado à conexão
  - Redes de **circuitos virtuais**
- Não-orientado à conexão
  - Redes de **datagramas**

- Análogos aos serviços da camada de transporte, porém
  - É um serviço estação-a-estação
  - Sem escolha
    - A rede provê ou um tipo ou o outro
  - É implementado no núcleo da rede

Circuitos Virtuais

Instituto de computação

- Emular uma rede de comutação de circuitos
  - Caminho da origem ao destino "se comporta" como um circuito telefônico
    - Em termos de desempenho
    - Em ações da rede ao longo do caminho

Circuitos Virtuais

Instituto de computação

- Funcionamento
  - Estabelecimento de uma **chamada** antes do envio dos dados
  - Cada pacote carrega a identificação do circuito virtual
    - Ao invés de endereços de origem e destino
  - Cada roteador no caminho origem-destino mantém **estado** para cada conexão que o atravessa
    - Cada conexão está associada a um CV
  - Recursos de enlace, roteador (banda, *buffers*) podem ser **alocados** ao circuito virtual

Circuitos Virtuais

Instituto de computação

- Um CV consiste de
  1. Caminho da origem para o destino
  2. Números (identificadores) de CV, um número para cada enlace ao longo do caminho
     - Menor complexidade de gerenciamento
  3. Entradas nas tabelas de repasse dos roteadores ao longo do caminho
- Pacote que pertence a um CV carregam o número do CV
- Número do CV deve ser trocado a cada enlace
- Novo número do CV vem da tabela de encaminhamento

Circuitos Virtuais: encaminhamento

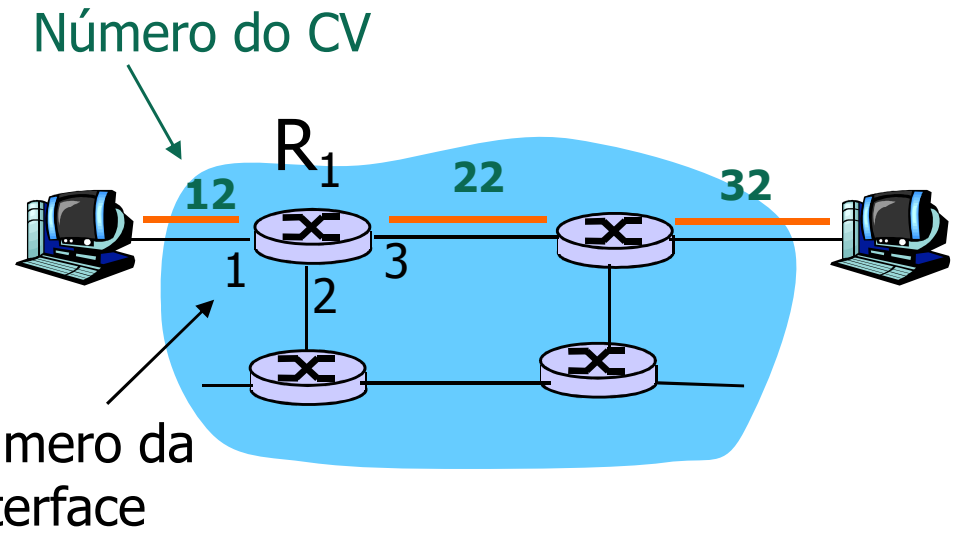


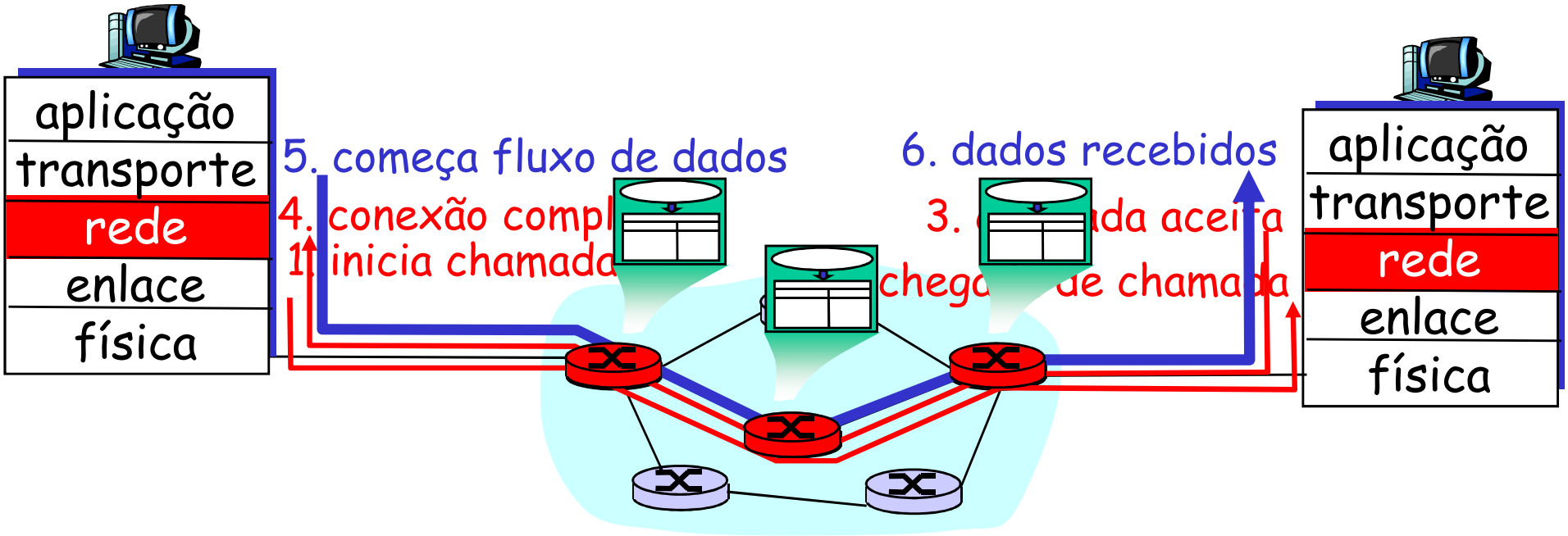
Tabela de encaminhamento no roteador R_1

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

Roteadores mantêm informação sobre o estado da conexão!

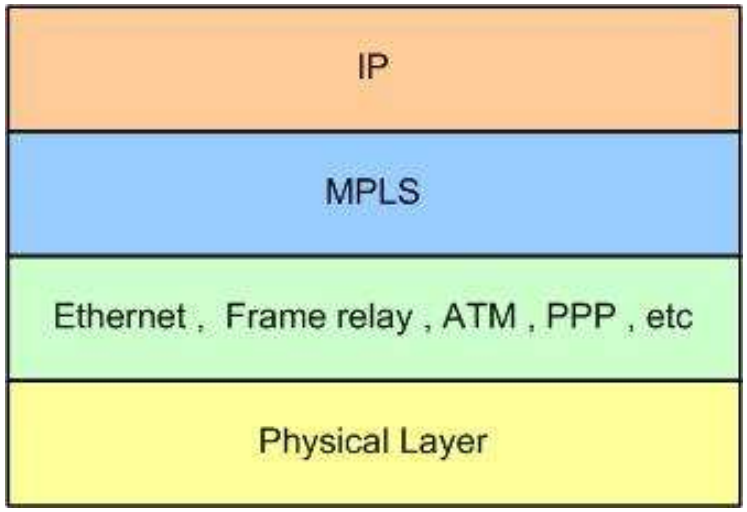
Circuitos Virtuais: Protocolos de Sinalização

- Responsáveis por estabelecer, manter e destruir um CV
  - Usados em ATM, *frame-relay,* X.25
  - Não usados na Internet convencional



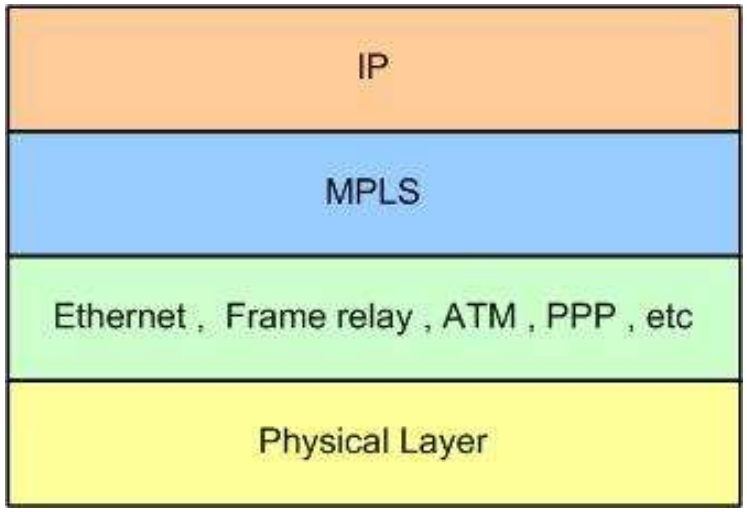
MPLS

- *Multiprotocol Label Switching*
  - Comutação por rótulos
  - Construção de circuitos virtuais
  - Diminuir o custo computacional do roteamento



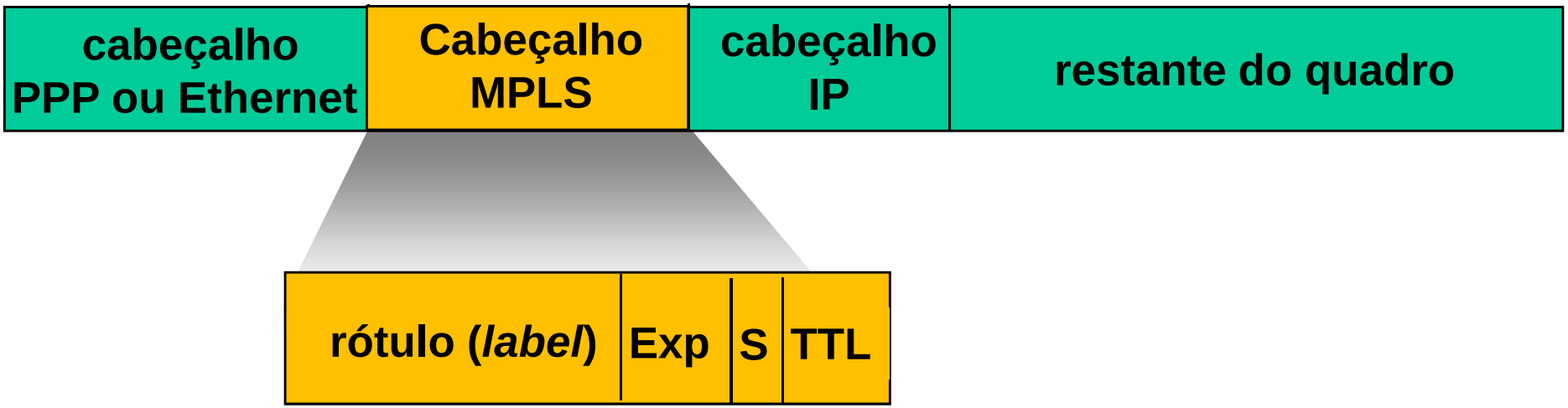
MPLS

- *Multiprotocol Label Switching*
  - Garantir encaminhamento rápido dos pacotes e QoS
    - Indexação rápida em uma tabela de comutação
    - Circuitos virtuais possibilitam a reserva de recursos
  - Análise do cabeçalho IP
    - Muita informação para escolher somente o próximo salto



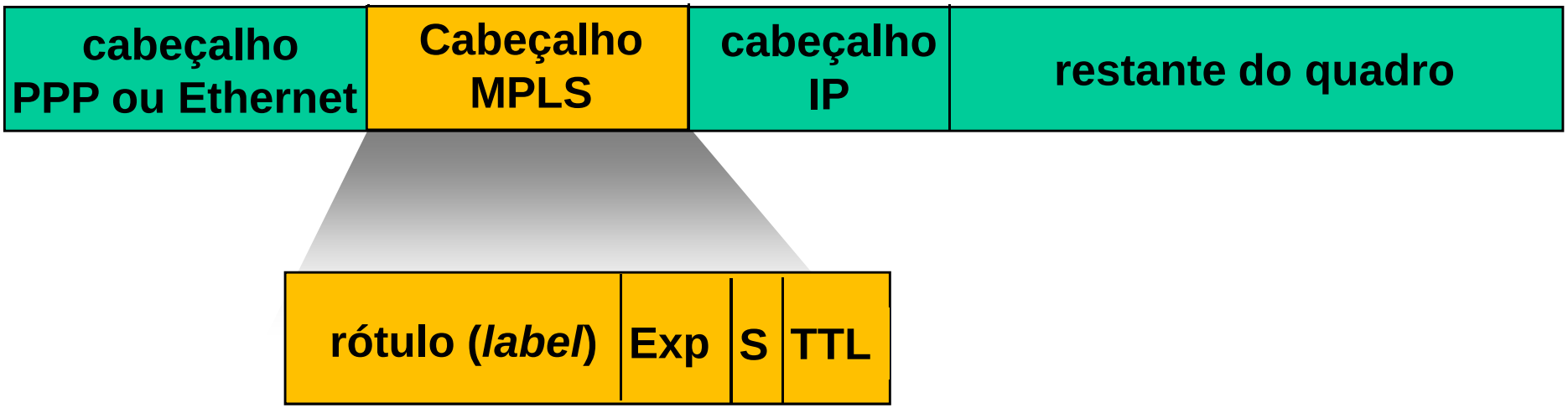
MPLS

- Comutação de rótulos
- Objetivo inicial
  - Acelerar o encaminhamento IP
    - Uso de **rótulo de comprimento fixo** ao invés de endereço IP



MPLS

- Idéias similares às da abordagem de circuitos virtuais
  - Mas os datagramas ainda mantêm o endereço IP
- Usado para fazer **engenharia de tráfego**



Roteador com Suporte ao MPLS

Instituto de computação

- Chamado de LSR (*Label-switched Router*)
- Encaminha os pacotes para a interface de saída baseada apenas no valor do rótulo
  - Não verifica o endereço IP
- A tabela de encaminhamento do MPLS é distinta da tabela de encaminhamento do IP

Roteador com Suporte ao MPLS

Instituto de computação

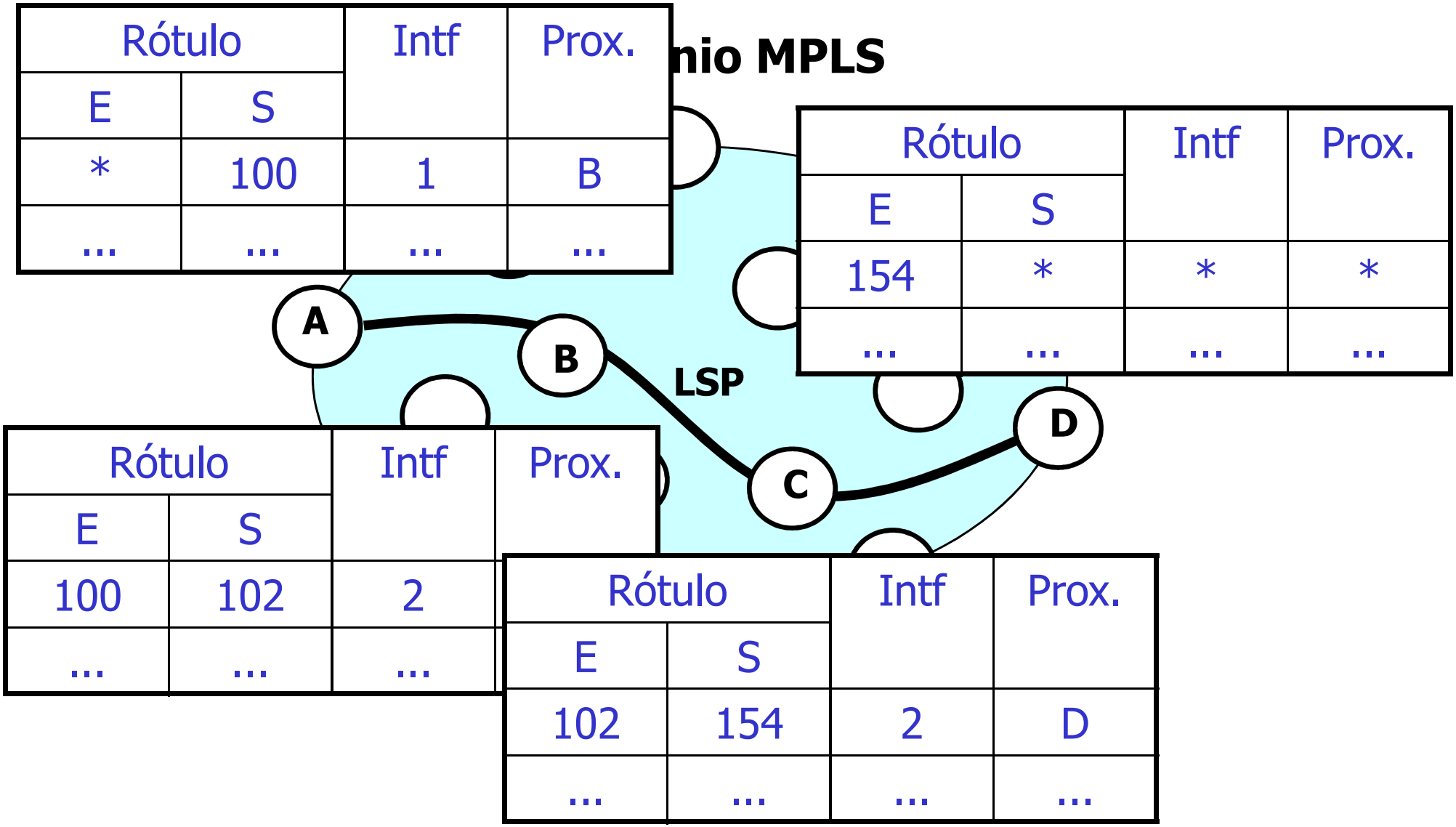
- É necessário protocolo de sinalização para criar as rotas
  - Chamadas de LSPs (*Label Switched Paths*)
  - Sinalização usando o RSVP-TE
    - Repasse possível através de caminhos que o IP sozinho não permitiria
      - Ex.: roteamento específico da origem
    - Engenharia de tráfego
- Deve coexistir com roteadores apenas IP

Domínio MPLS

Instituto de computação

- Roteador comutado por rótulo – LSRs
  - Encaminhamento de acordo com o rótulo e interface
  - Rótulo trocado a cada salto
  - Mapeamento constante
- Caminho dos pacotes – LSPs
- Pacotes com mesmo rótulo pertencem a mesma classe de encaminhamento (FEC)

Encaminhamento no MPLS



Rótulo		Intf	Prox.
E	S		
*	100	1	B
...	...	...	...

Rótulo		Intf	Prox.
E	S		
154	*	*	*
...	...	...	...

Rótulo		Intf	Prox.
E	S		
100	102	2	
...	...	...	

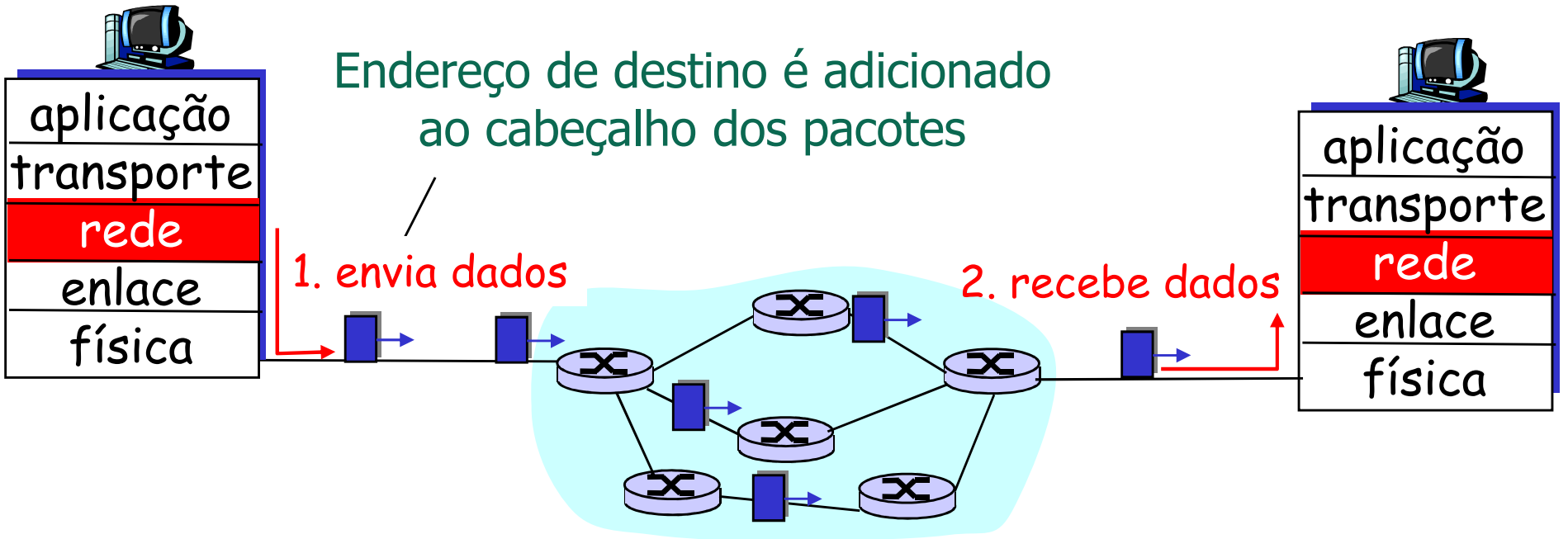
Rótulo		Intf	Prox.
E	S		
102	154	2	D
...	...	...	...

Datagramas

Instituto de computação

- Serviço não confiável
- Sem estabelecimento prévio de conexão
- Roteadores não guardam estado sobre conexões
- Pacotes são encaminhados
  - Com base no endereço de destino
  - De acordo com o modelo de melhor esforço
- Dois pacotes entre o mesmo par origem-destino podem seguir caminhos diferentes

Datagramas


Endereço de destino é adicionado
ao cabeçalho dos pacotes
aplicação
transporte
rede
enlace
física
1. envia dados
2. recebe dados
aplicação
transporte
rede
enlace
física

Datagramas: Encaminhamento

Instituto de computação

- Endereço IP: 32 bits
  - 4 bilhões de endereços → **4 bilhões de entradas!**
  - **Agregação de endereços**

Como resumir a tabela?

Maior Prefixo

Instituto de computação

Faixa de endereços de destino	Interface de saída
11001000.00010111.00010000.00000000 a 11001000.00010111.00010111.11111111	0
11001000.00010111.00011000.00000000 a 11001000.00010111.00011000.11111111	1
11001000.00010111.00011001.00000000 a 11001000.00010111.00011111.11111111	2
Caso contrário	3

Maior Prefixo

Instituto de computação

Faixa de endereços de destino	Interface de saída
11001000.00010111.00010000.00000000 a 11001000.00010111.00010111.11111111	0
11001000.00010111.00011000.00000000 a 11001000.00010111.00011000.11111111	1
11001000.00010111.00011001.00000000 a 11001000.00010111.00011111.11111111	2
Caso contrário	3

Maior Prefixo

Instituto de computação

Faixa de endereços de destino	Interface de saída
11001000.00010111.00010000.00000000 a 11001000.00010111.00010111.11111111	0
11001000.00010111.00011000.00000000 a 11001000.00010111.00011000.11111111	1
11001000.00010111.00011001.00000000 a 11001000.00010111.00011111.11111111	2
Caso contrário	3

Maior Prefixo

Instituto de computação

Faixa de endereços de destino	Interface de saída
11001000.00010111.00010	0
11001000.00010111.00011000	1
11001000.00010111.00011	2
Caso contrário	3

Maior Prefixo

Instituto de computação

Faixa de endereços de destino	Interface de saída
11001000.00010111.00010	0
11001000.00010111.00011000	1
11001000.00010111.00011	2
Caso contrário	3

Exemplos

ED: 11001000 00010111 00010110 10100001 Qual interface?

ED: 11001000 00010111 00011000 10101010 Qual interface?

Maior Prefixo

Instituto de computação

Faixa de endereços de destino	Interface de saída
11001000.00010111.00010	0
11001000.00010111.00011000	1
11001000.00010111.00011	2
Caso contrário	3

Exemplos

ED: 11001000 00010111 00010110 10100001 Interface 0

ED: 11001000 00010111 00011000 10101010 Interface 1

Circuitos Virtuais x Datagramas

Características	Circuito Virtual	Datagrama
Estabelecimento de conexão	É necessário	Não é necessário
Endereçamento	Identificador do CV	Endereços da fonte e do destino
Estados	Por conexão	Sem estado
Roteamento	Rota escolhida na conexão e seguida posteriormente	Cada pacote é "independente"
Falha de roteadores	Todos os circuitos fechados	Perda de pacotes durante a falha
Qualidade de serviço	Mais fácil	Difícil
Controle de congestionamento	Mais fácil	Difícil

Instituto de computação

Arquitetura de Roteadores

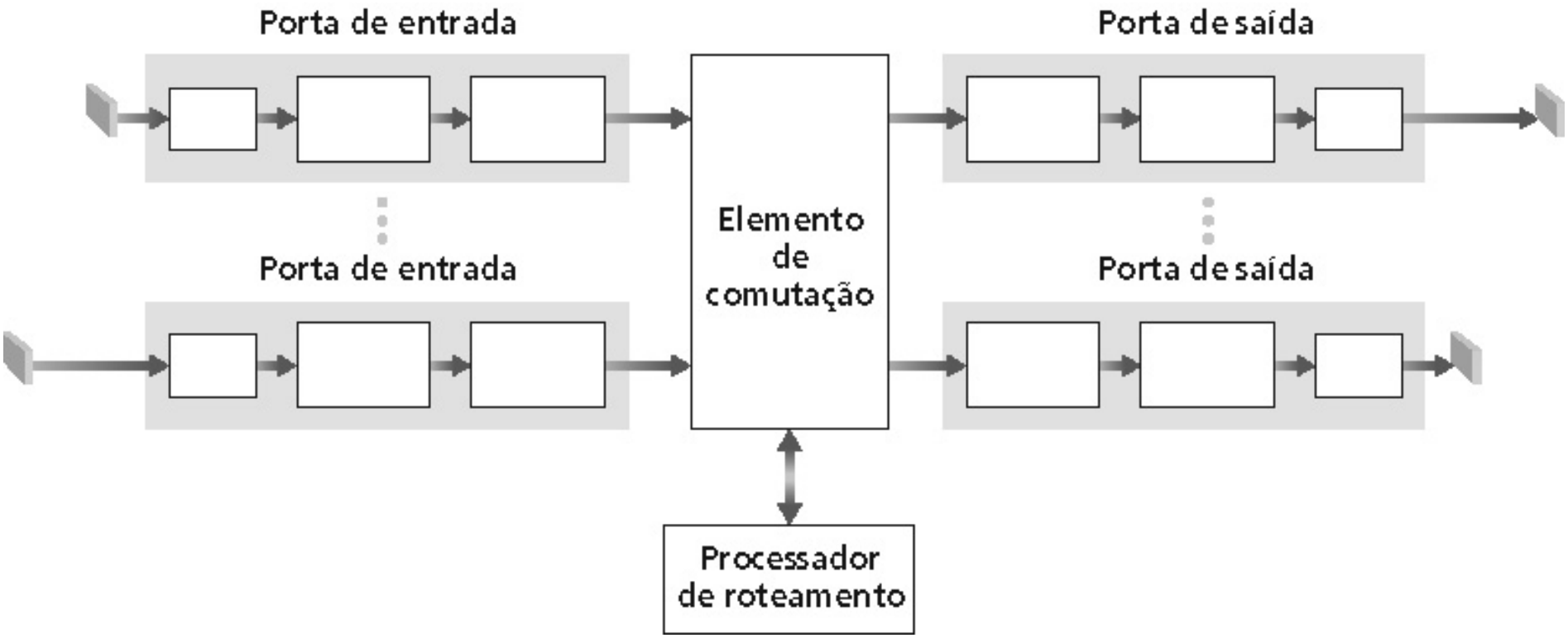
Roteador

Instituto de computação

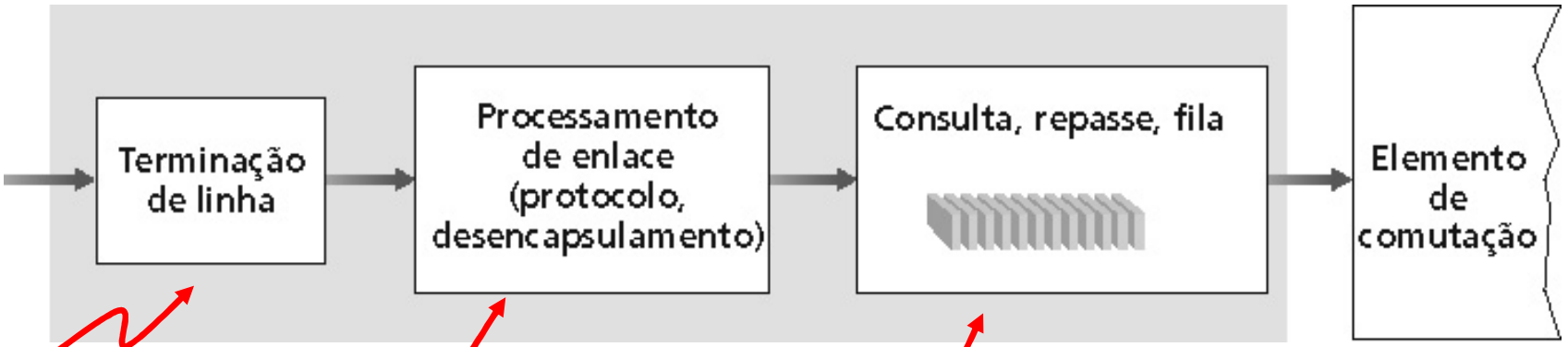
- Elemento responsável
  - Determinar o caminho entre um par origem-destino
    - Ação distribuída
  - Encaminhar pacotes
  - Interconectar redes distintas

- Cada pacote ao chegar a um roteador
  - Tem seu endereço de destino analisado
    - Se o endereço for igual ao de uma das interfaces do roteador
      - Pacote é enviado para camada de transporte
    - Caso contrário
      - Pacote é encaminhado a outro roteador pela interface mais indicada

Roteador


Porta de entrada
Porta de entrada
Elemento de comutação
Porta de saída
Porta de saída
Processador de roteamento

Funções das Portas de Entrada



Camada física:
recepção de bits

Camada de enlace:
p.ex., Ethernet
veja capítulo 5

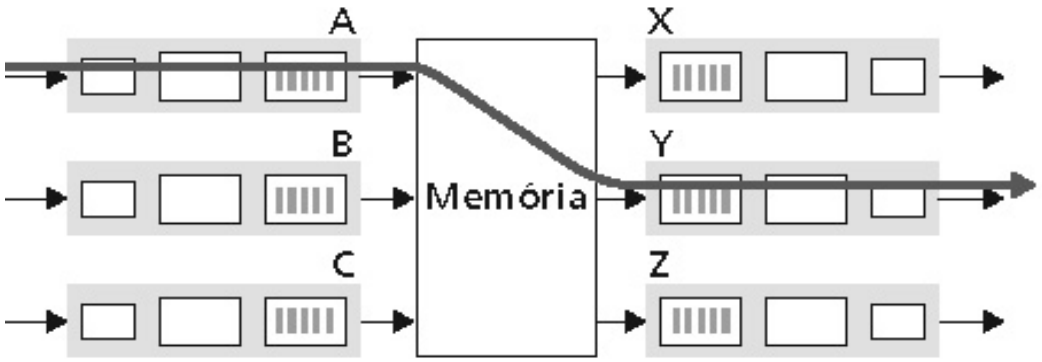
Comutação descentralizada*:*

- dado o dest. do datagrama, procura porta de saída usando tab. de rotas na memória da porta de entrada
- meta: completar processamento da porta de entrada na '**velocidade da linha**'
- filas: se datagramas chegam mais rápido que taxa de re-envio para matriz de comutação

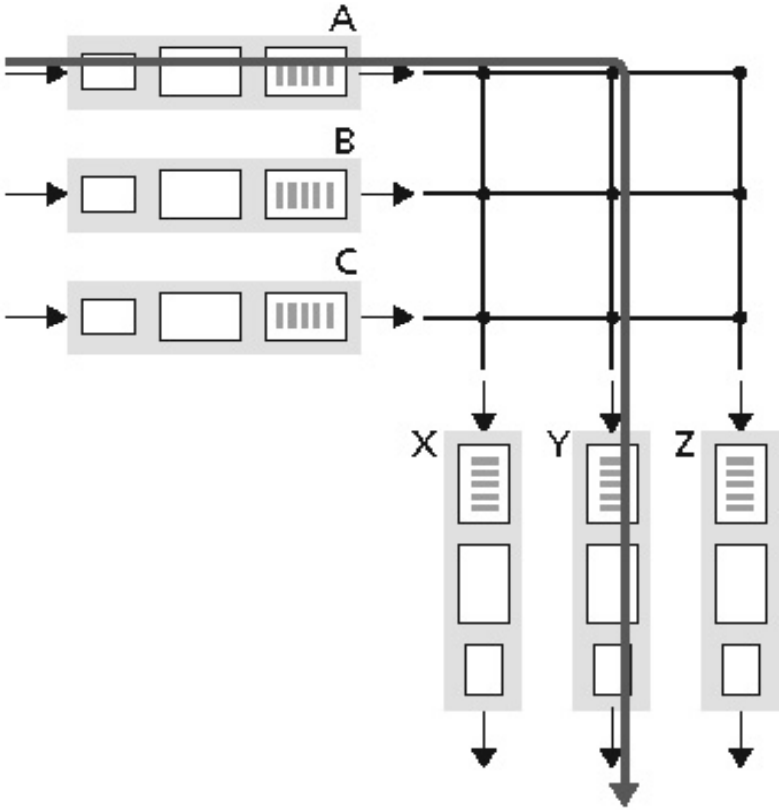
Três Técnicas de Comutação

Instituto de computação

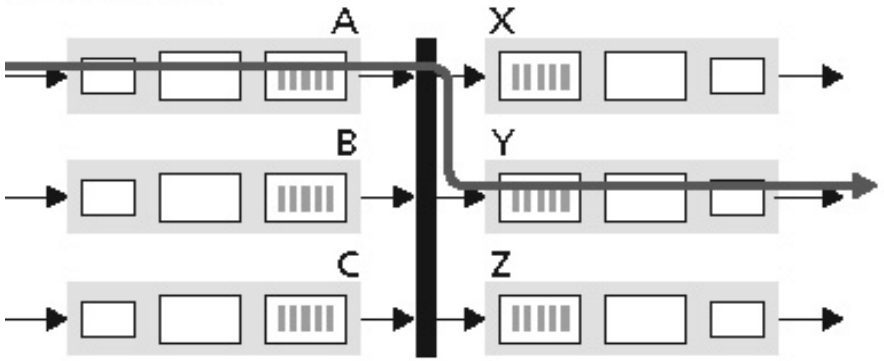
Memória



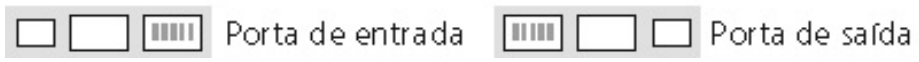
Crossbar



Barramento

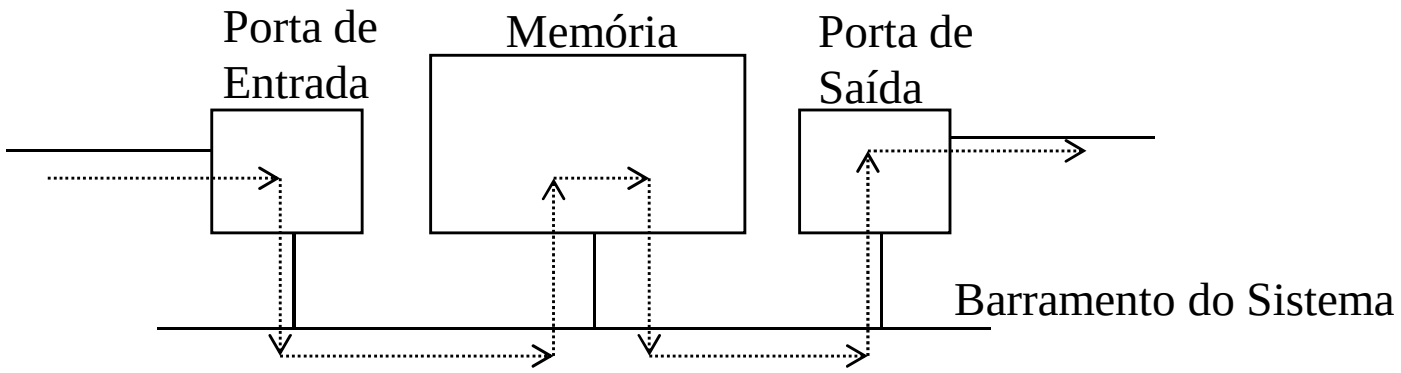


Legenda:



Comutação por Memória

- Roteadores da primeira geração
- Pacote copiado pelo processador (único) do sistema
  - Velocidade limitada pela largura de banda da memória
    - **Duas travessias** do barramento por datagrama



Comutação por Barramento

Instituto de computação

- Datagrama da memória da porta de entrada à memória da porta de saída via um barramento compartilhado
- Disputa (contenção) pelo barramento
  - Taxa de comutação limitada pela largura de banda do barramento

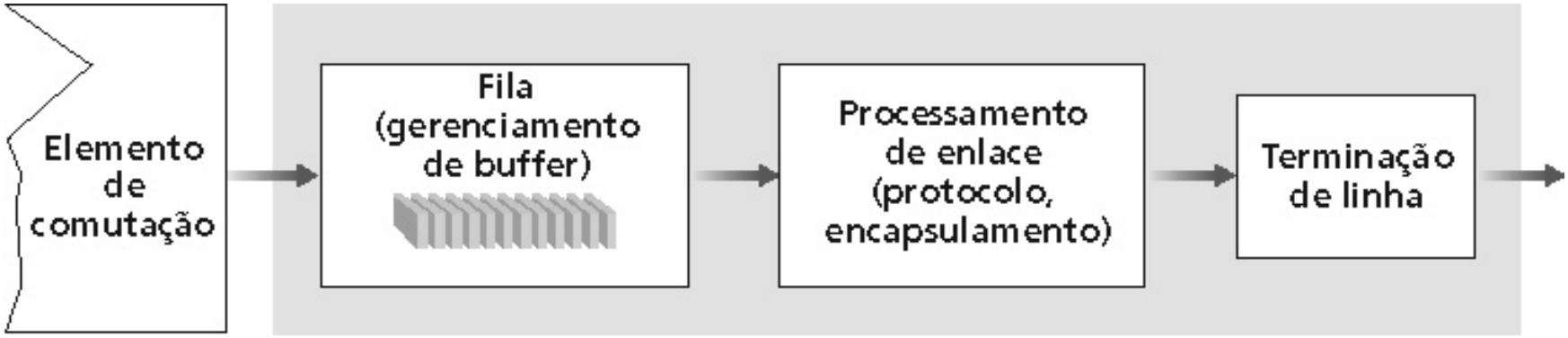
Comutação por *Crossbar*

Instituto de computação

- Reduzir a disputa
  - "Por porta de saída"
- $2n$ barramentos
  - Interconectar n portas de entrada a n portas de saída

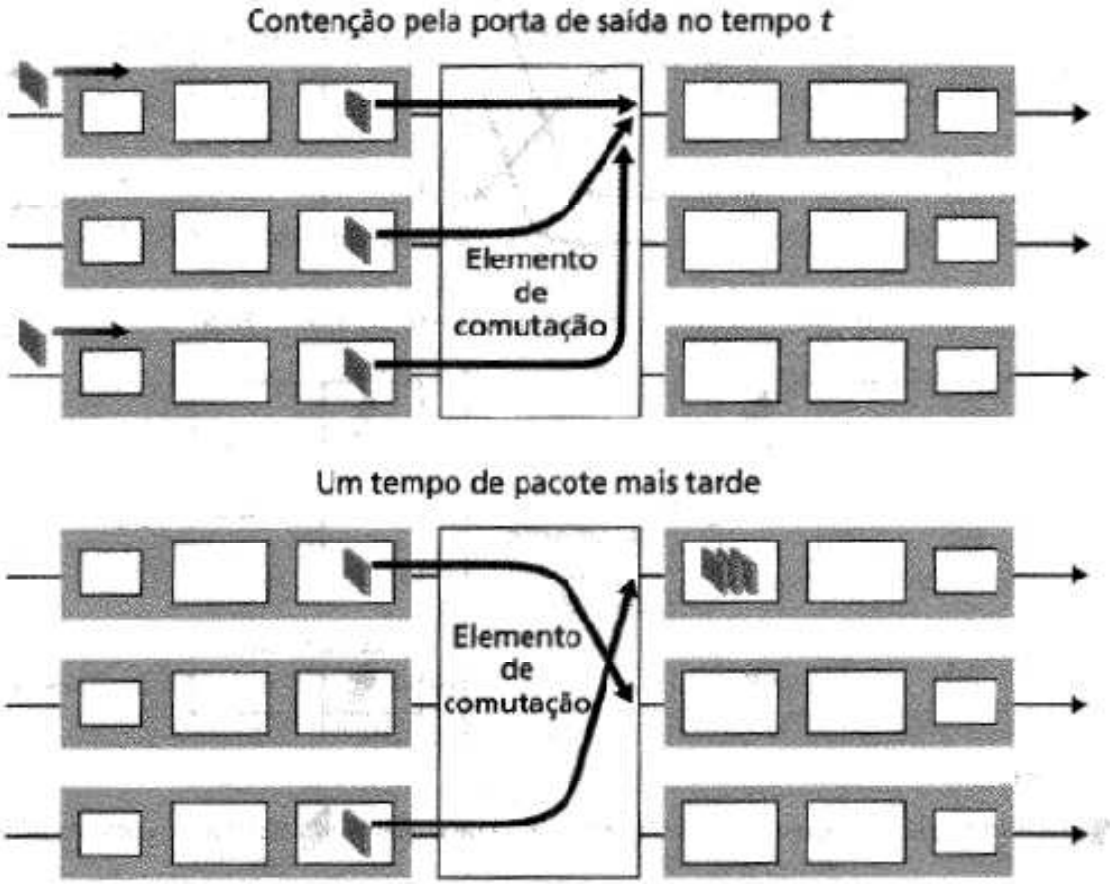
Funções das Portas de Saída

- Filas
  - Necessárias quando datagramas chegam do elemento de comutação mais rapidamente do que a taxa de transmissão
- Escalonador de pacotes escolhe um dos datagramas enfileirados para transmissão



Como Ocorrem as Filas?

- Portas de saída
  - Usam *buffers* quando taxa de chegada através do comutador excede taxa de transmissão de saída
  - *enfileiramento (retardo), e perdas devidas ao transbordo do buffer da porta de saída!*

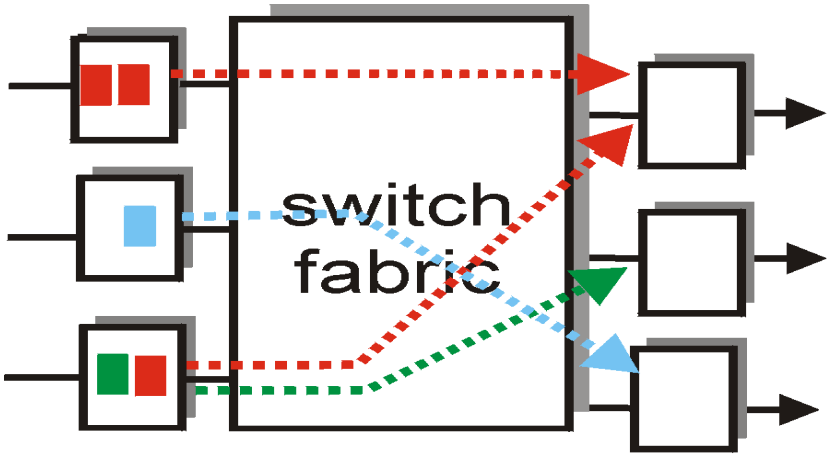


Como Ocorrem as Filas?

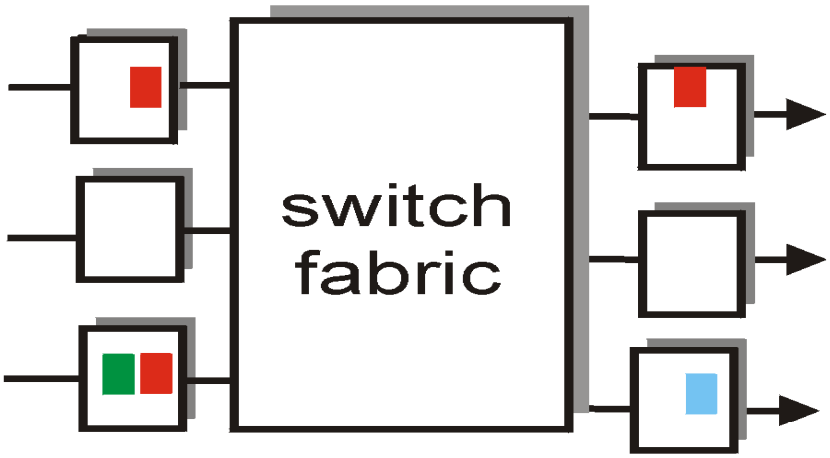
Instituto de computação

- Portas de entrada
  - Se o elemento de comutação for mais lento do que a soma das portas de entrada juntas
    - Pode haver filas nas portas de entrada
  - Bloqueio de cabeça de fila
    - Datagrama na cabeça da fila impede outros na mesma fila de avançarem
  - Retardo de enfileiramento e perdas devido ao transbordo do buffer de entrada!

Como Ocorrem as Filas?



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



green packet
experiences HOL blocking

Instituto de computação

Internet Protocol (IP)

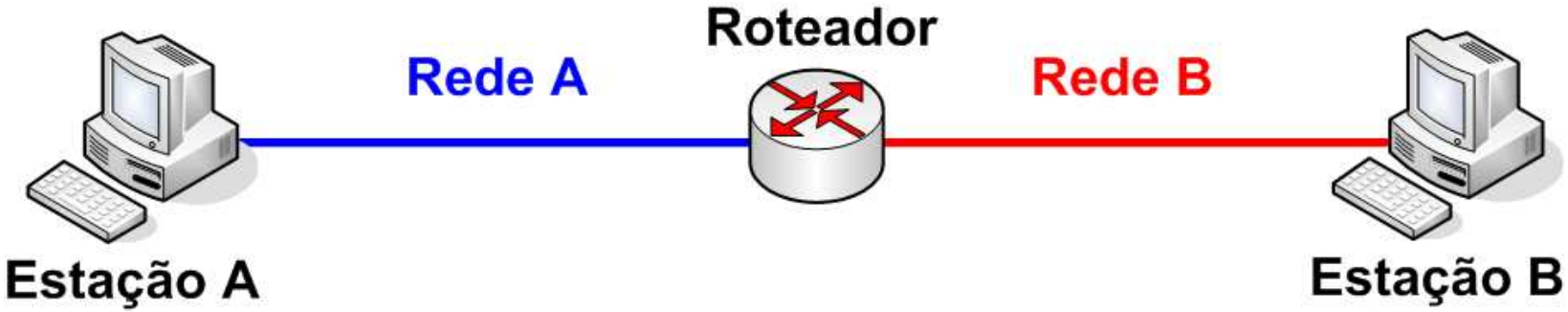
Internet Protocol

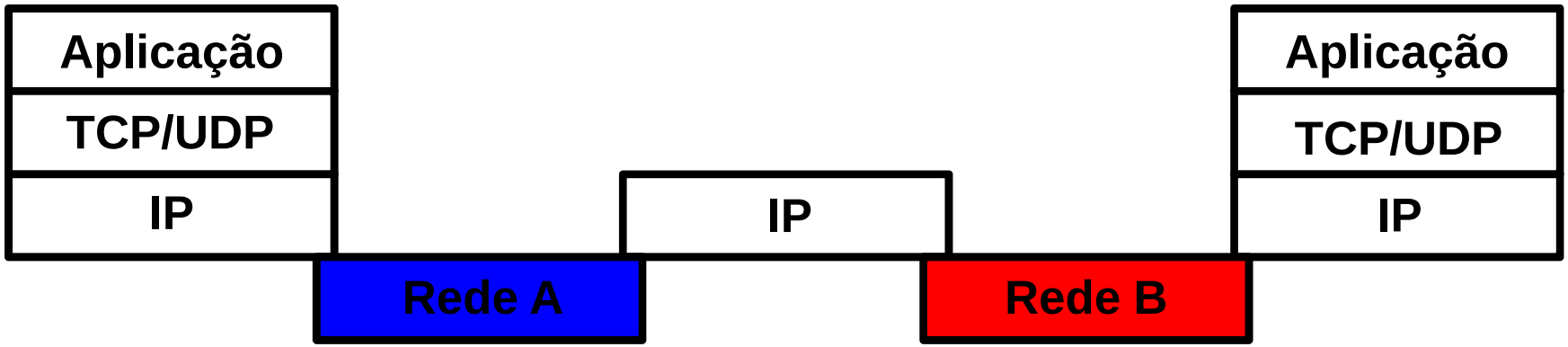
Instituto de computação

- Definido pela RFC 791

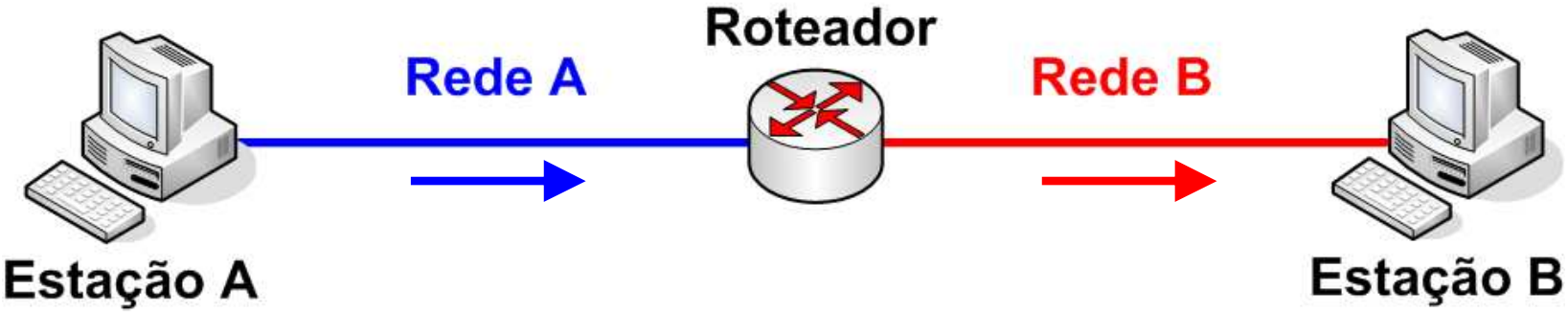
- É o responsável pelo
  - **Encaminhamento de pacotes**
    - **Não pelo roteamento!**
  - Endereçamento e identificação de estações e roteadores
    - Semântica sobrecarregada

Operação do IP





Transmissão de um Pacote IP



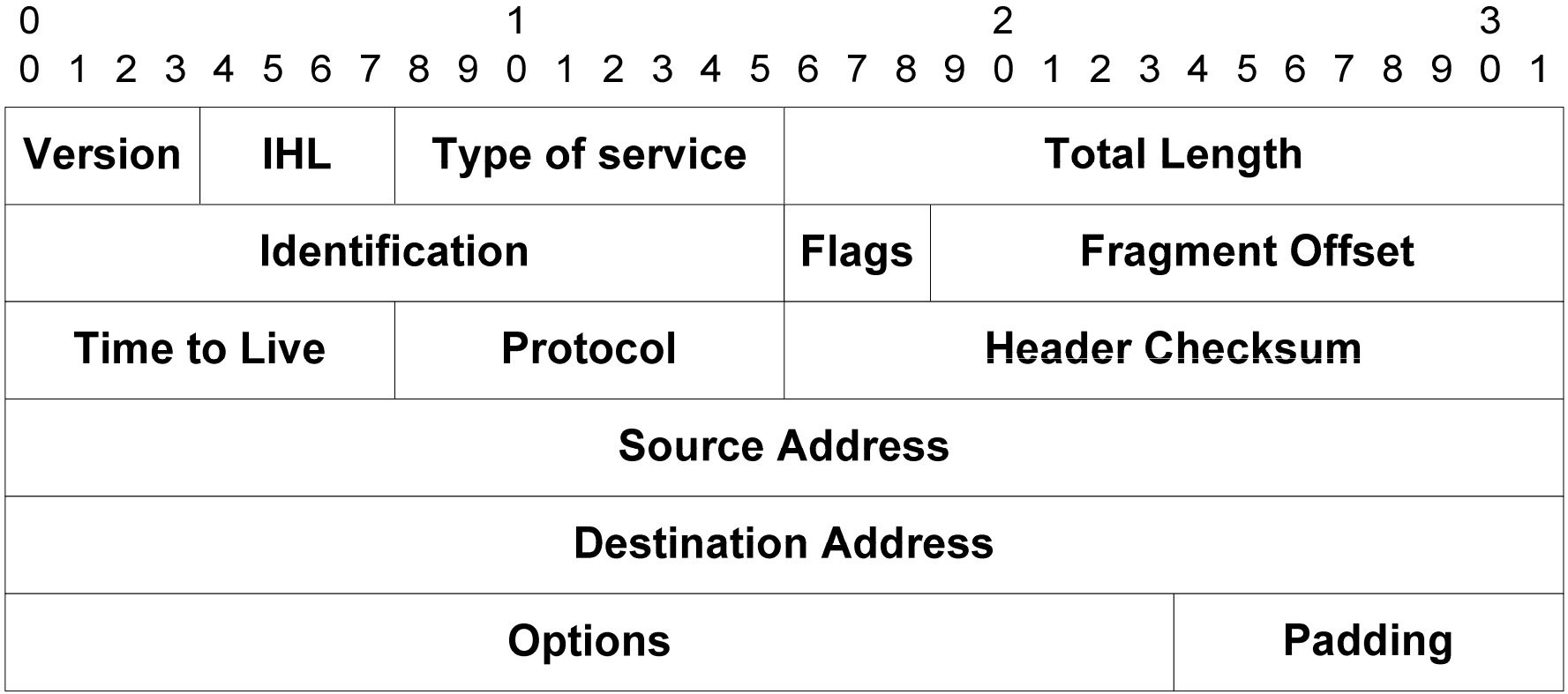
A1 → C1, IP	A → B, TCP	cabeçalho TCP + dados
Cabeçalho Ethernet	Cabeçalho IP	

C2 → B2, IP	A → B, TCP	cabeçalho TCP + dados
Cabeçalho Ethernet	Cabeçalho IP	

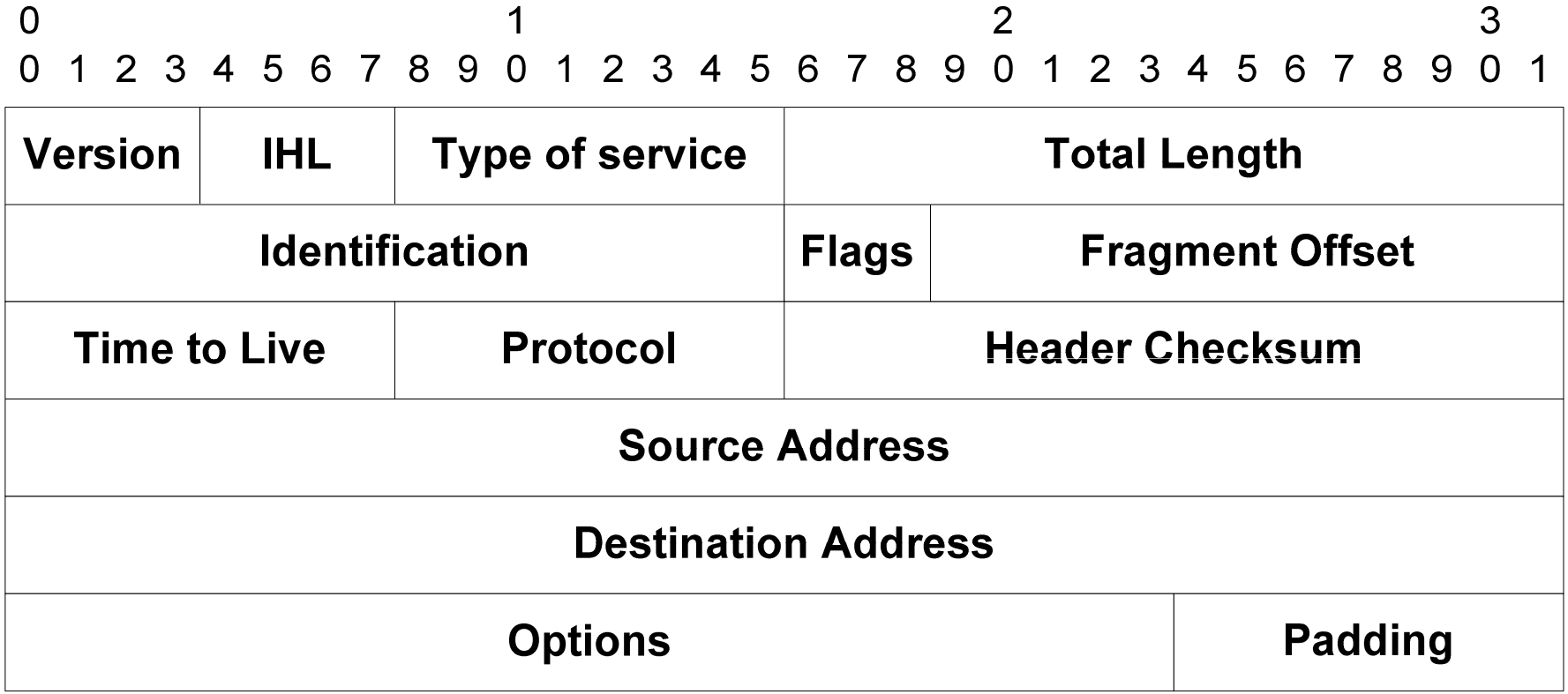
Instituto de Computação

IPv4

O Cabeçalho IP

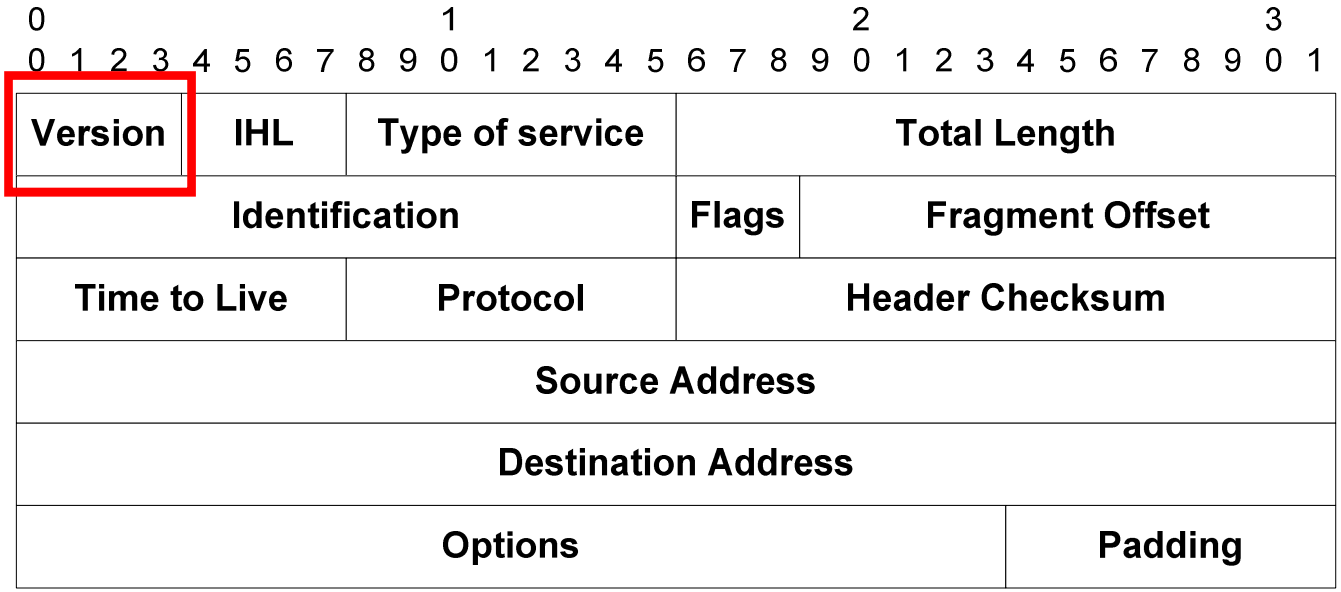


O Cabeçalho IP



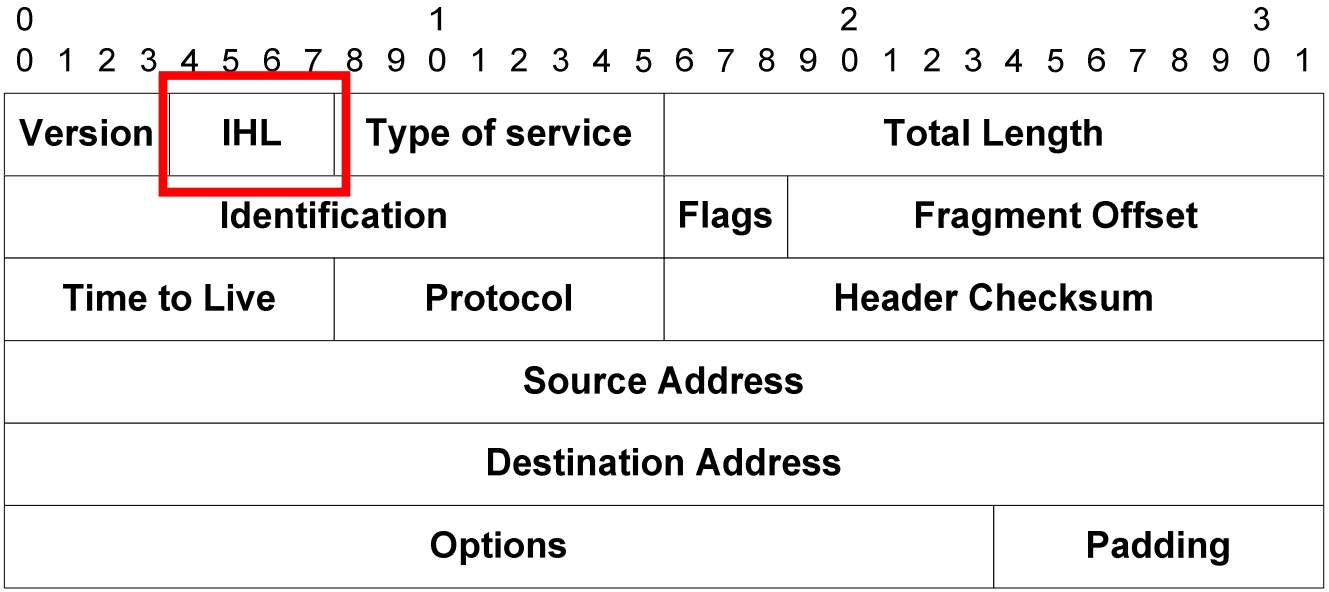
Todos os campos possuem tamanho fixo, exceto o campo de opções

Campos do Cabeçalho IP



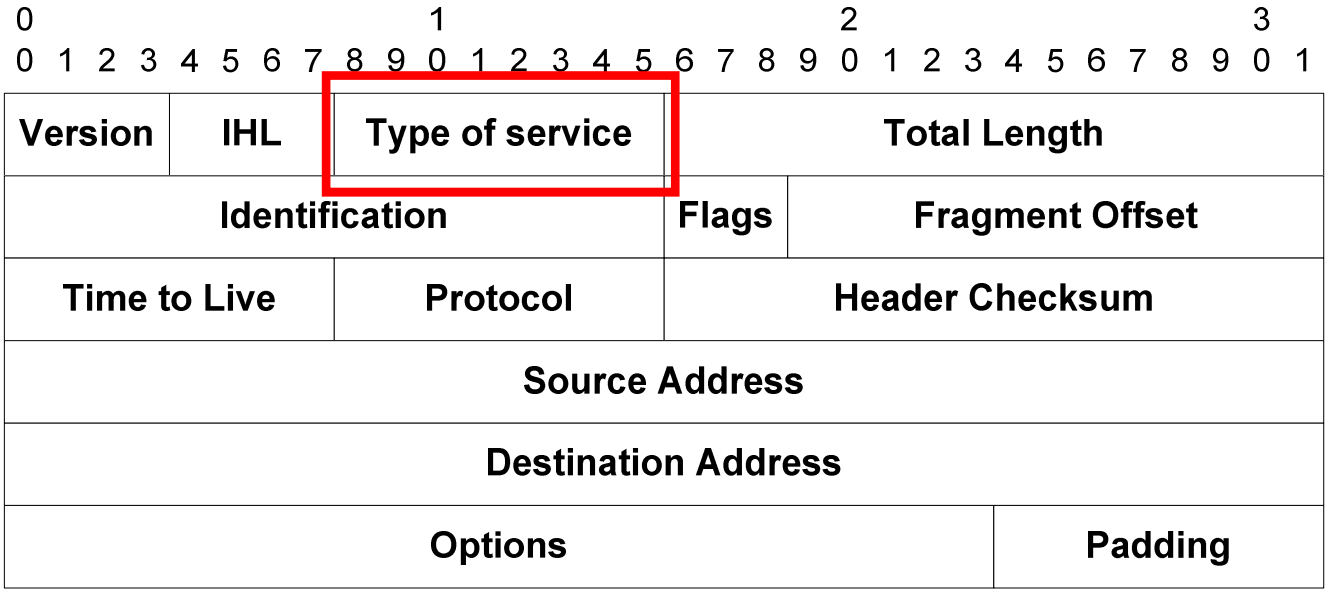
- Versão (4bits)
  - Versão atual = 4
  - Versão 5 = Protocolo ST-2
  - Versão 6 = “A próxima geração”
  - Versões 7 e 8

Campos do Cabeçalho IP



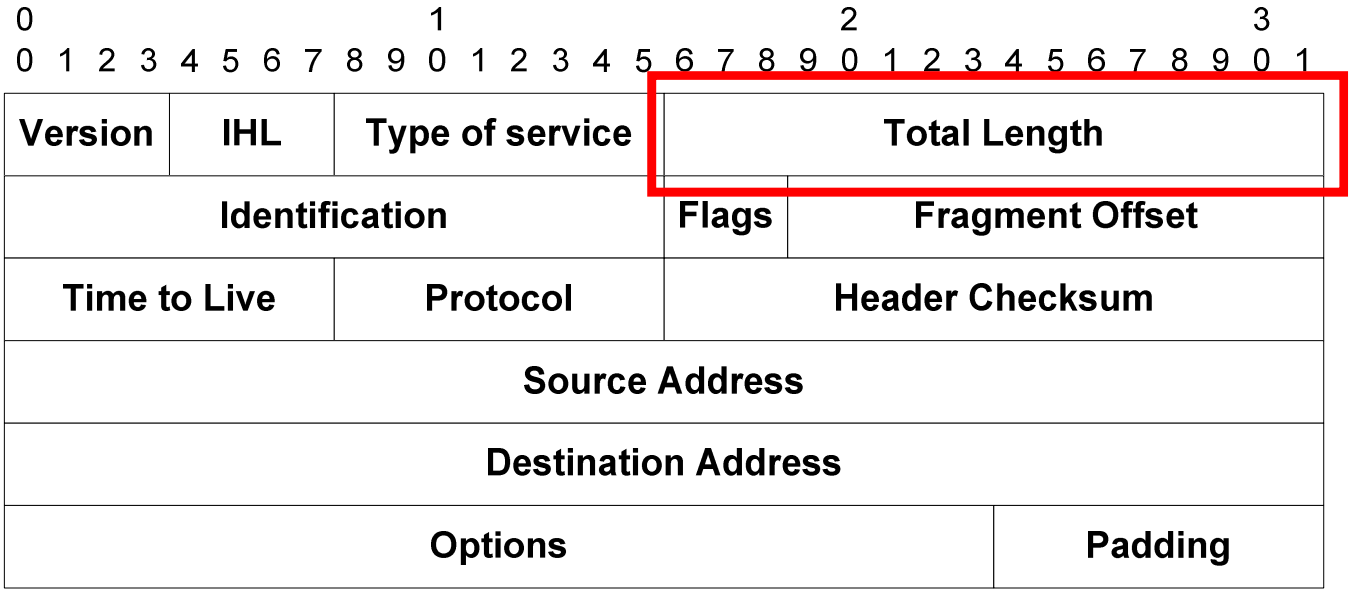
- IHL (*Internet header's length*) (4 bits)
  - Comprimento do cabeçalho, em palavras de 32 bits
  - Varia de 5 (quando não há opções) a 15
    - Ou seja, podem haver 40 bytes de opções, no máximo

Campos do Cabeçalho IP



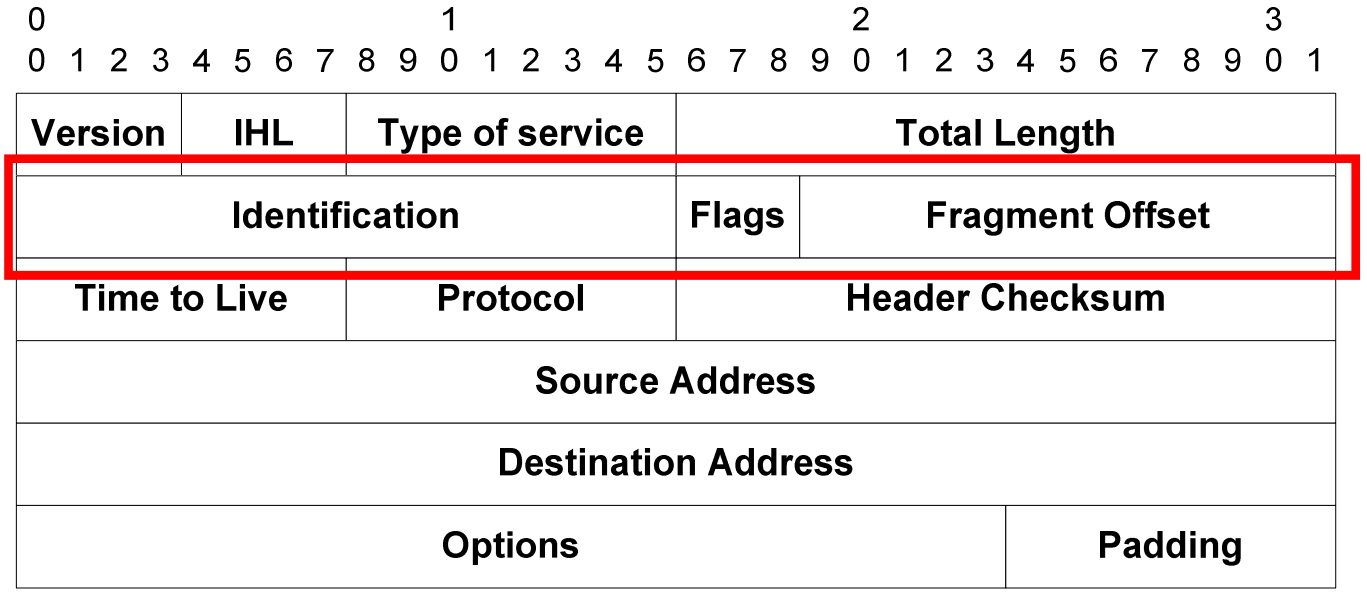
- Tipo de serviço (*Type of Service*) (8 bits)
  - Define a precedência e o tipo de roteamento desejado para o pacote
    - Utilizado para qualidade de serviço (QoS)

Campos do Cabeçalho IP



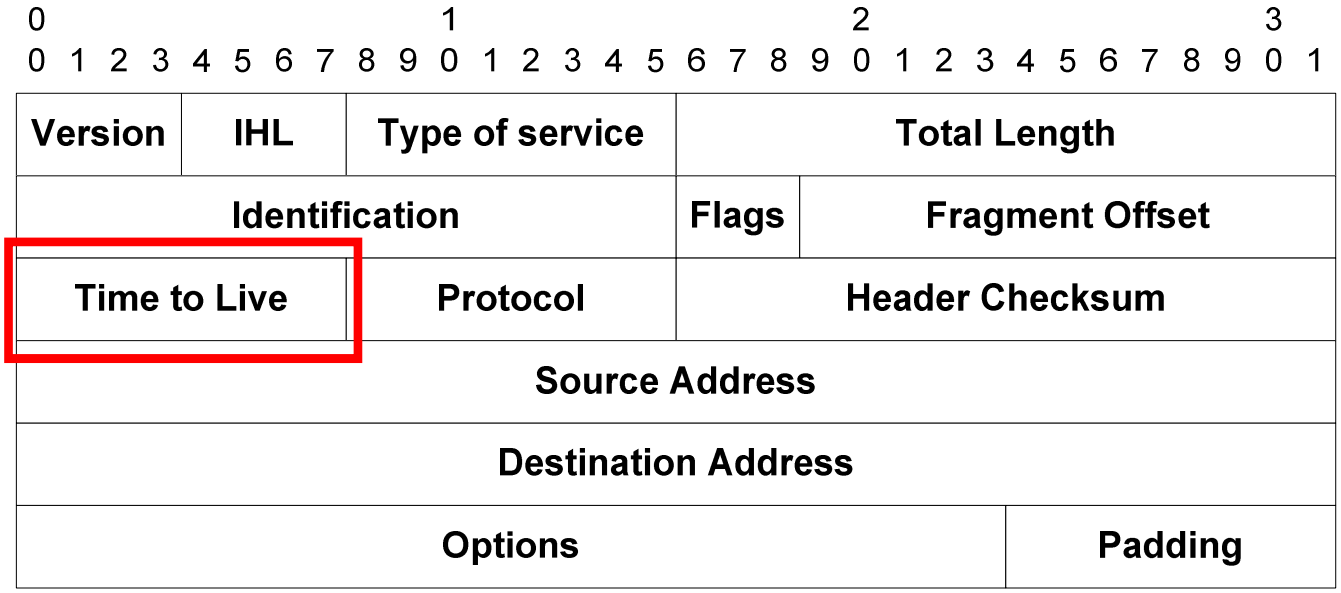
- Comprimento total (*Total Length*) (16 bits)
  - Comprimento total do pacote, incluindo o cabeçalho
  - Limita o tamanho do pacote a 65.535 bytes

Campos do Cabeçalho IP



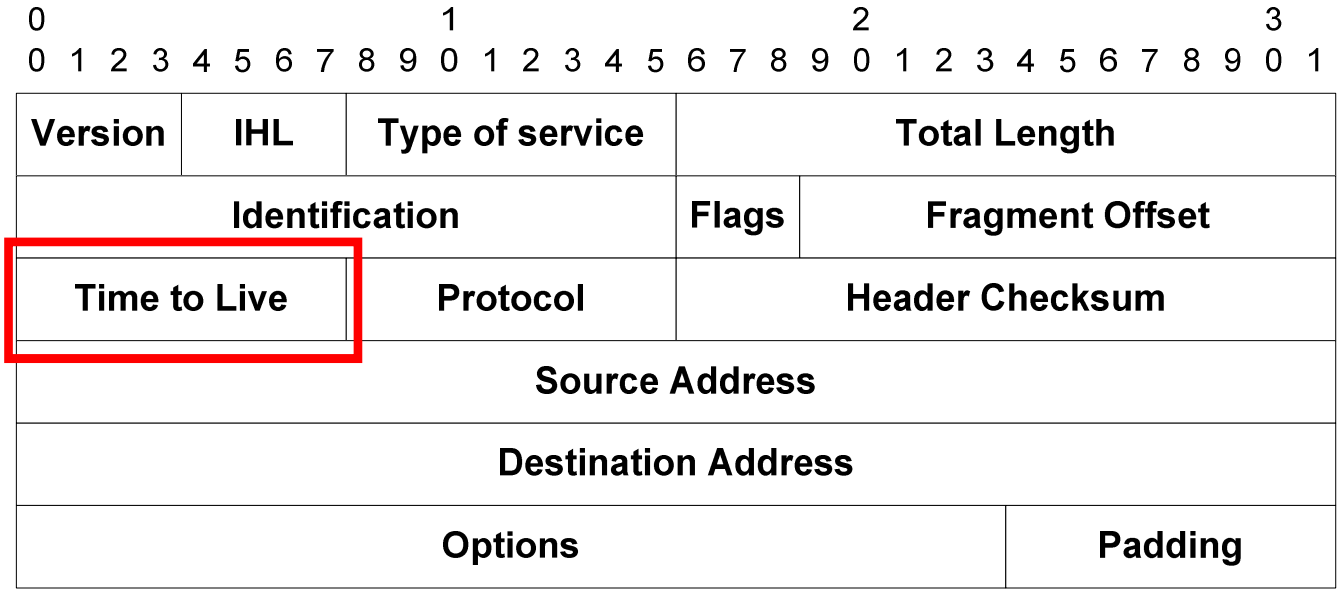
- Identification, Flags e Fragment Offset
  - Utilizados no processo de fragmentação e remontagem

Campos do Cabeçalho IP



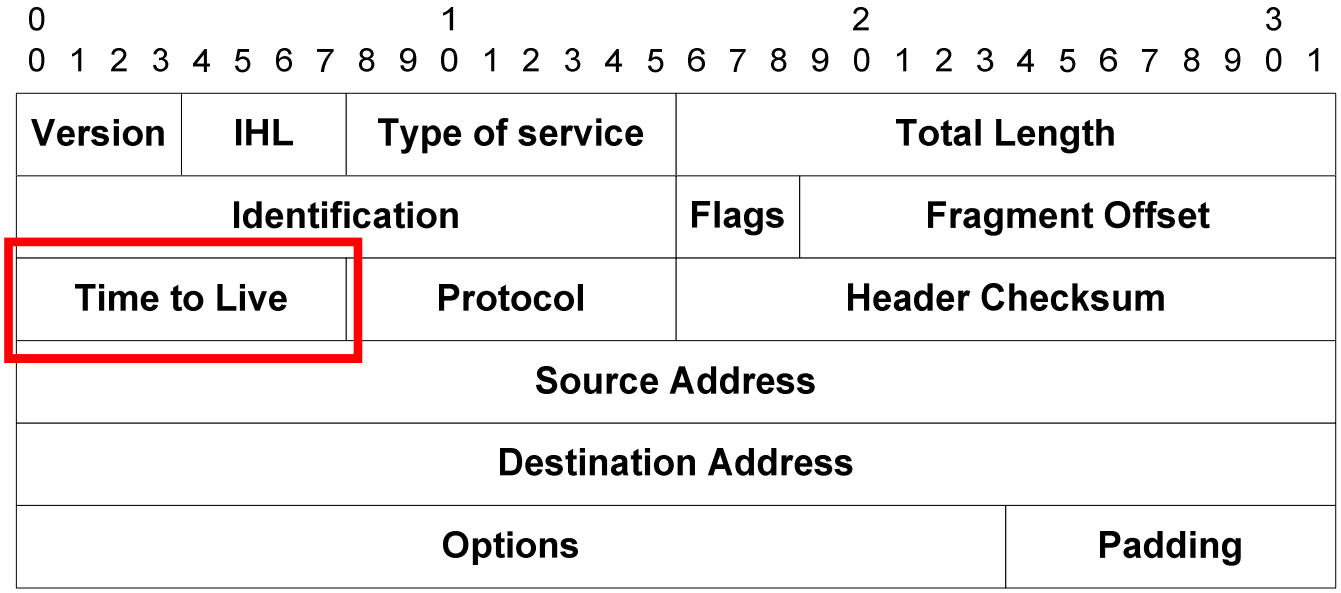
- Tempo de Vida (*Time to Live* -TTL) (8 bits)
  - Tempo de vida máximo do pacote na rede, em segundos
    - Um dos objetivos era saber que depois do TTL máximo, nenhum outro pacote daquela comunicação estaria em trânsito
      - Evita-se misturar pacotes de fluxos de dados diferentes

Campos do Cabeçalho IP



- Tempo de Vida (*Time to Live* -TTL) (8 bits)
  - RFC-791: Um roteador deve sempre decrementar o TTL antes de retransmitir um pacote
    - O TTL deve ser decrementado de 1, se o tempo gasto nas filas e na transmissão ao próximo nó for menor que 1 segundo
    - Ou do número de segundos estimado

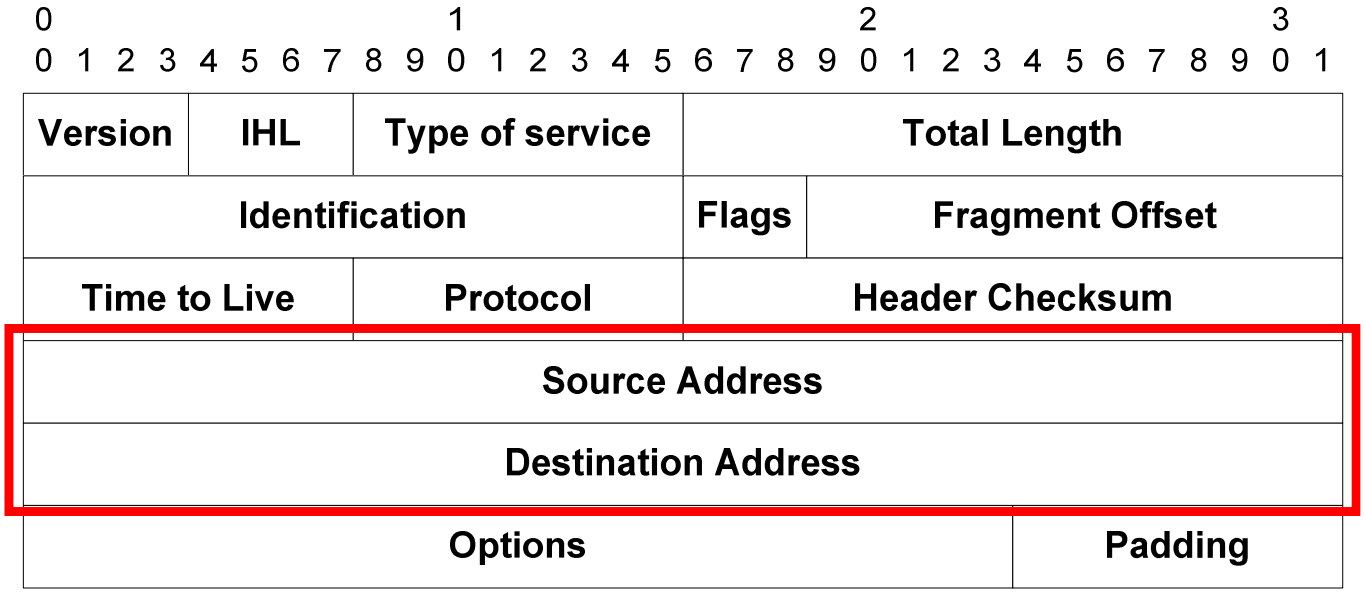
Campos do Cabeçalho IP



- Tempo de Vida (Time to Live -TTL) (8 bits)
  - Na prática, estimar este tempo é difícil e o tempo de transmissão nos enlaces dificilmente ultrapassa 1 segundo
    - Maioria dos roteadores decrementa o TTL de 1
  - Se o TTL atinge o valor 1, o pacote deve ser descartado
    - Sinal de que o pacote já trafegou mais tempo que o devido...

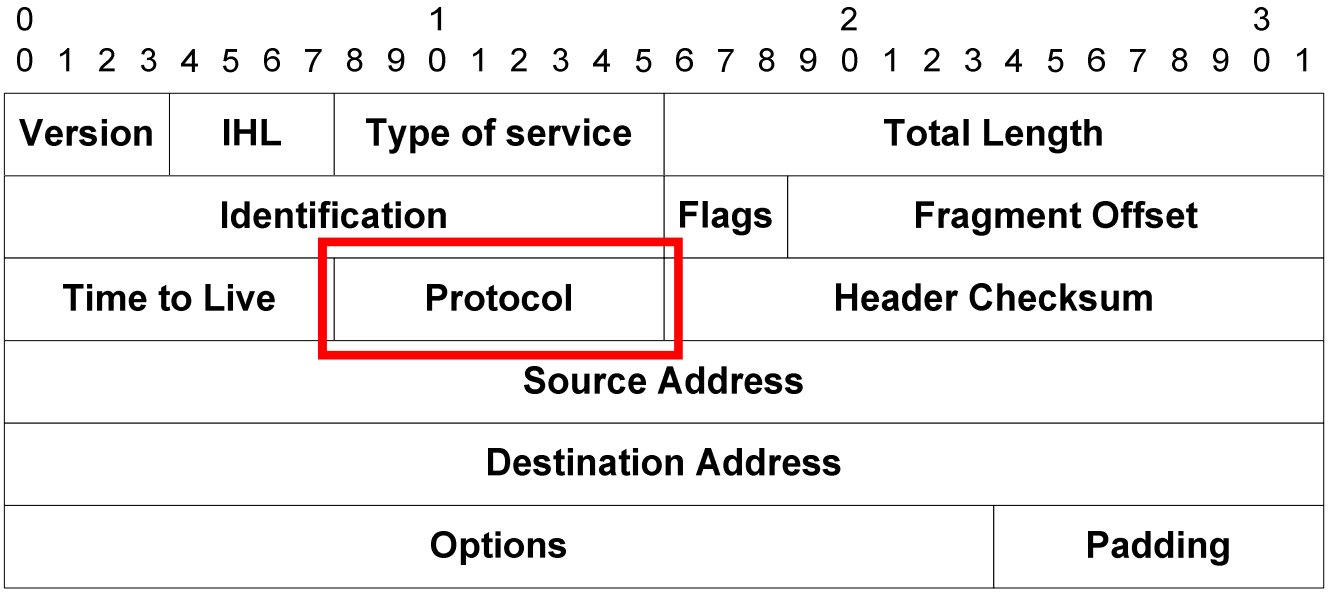
Valor padrão: TTL = 64

Campos do Cabeçalho IP



- Source Address e Destination Address (32 bits cada)
  - Identificam a fonte e destino do pacote

Campos do Cabeçalho IP



- Protocol (8 bits)
  - Determina o programa para o qual o pacote é passado, no destino

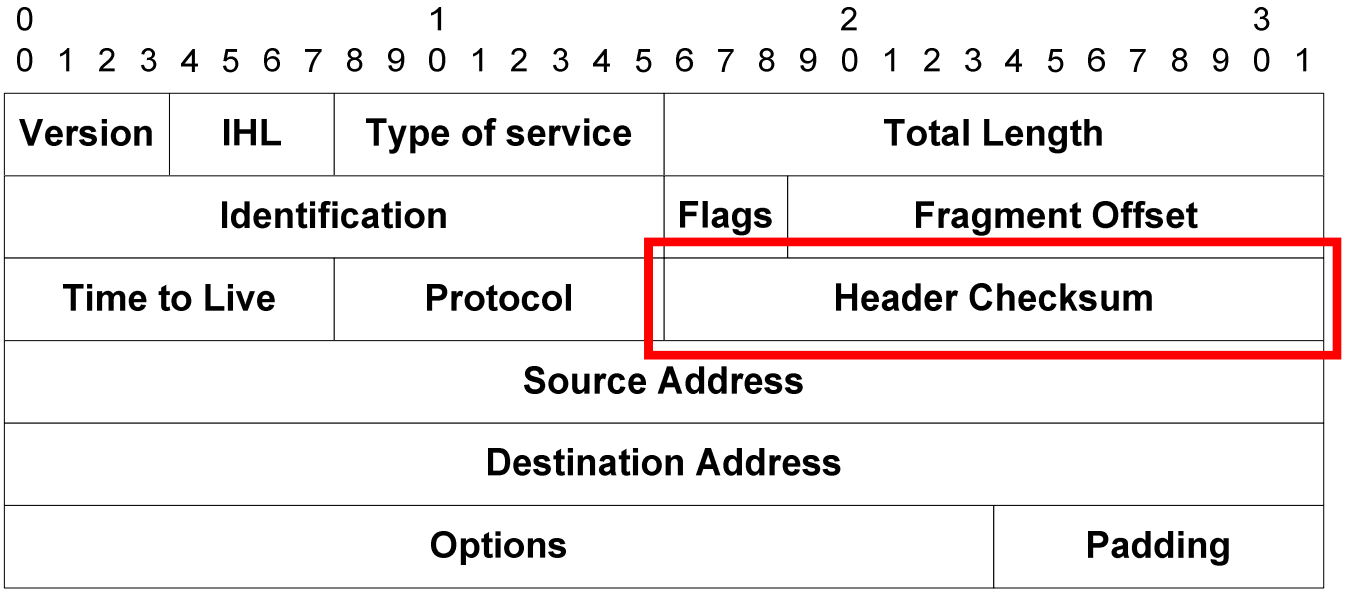
Campos do Cabeçalho IP

- Diferentes protocolos

Decimal	Sigla	Protocolo
0		Reservado
1	ICMP	Internet Control Message
2	IGMP	Internet Group Management
4	IP	IP em IP (encapsulação)
6	TCP	Transmission Control

Decimal	Sigla	Protocolo
17	UDP	User Datagram
29	ISO-TP4	ISO Transport Prot Class 4
80	ISO-IP	ISO Internet Protocol (CLNP)
89	OSPF	Open Shortest Path First
255		Reservado

Campos do Cabeçalho IP



- Header Checksum (16 bits)
  - Proteção do cabeçalho contra erros

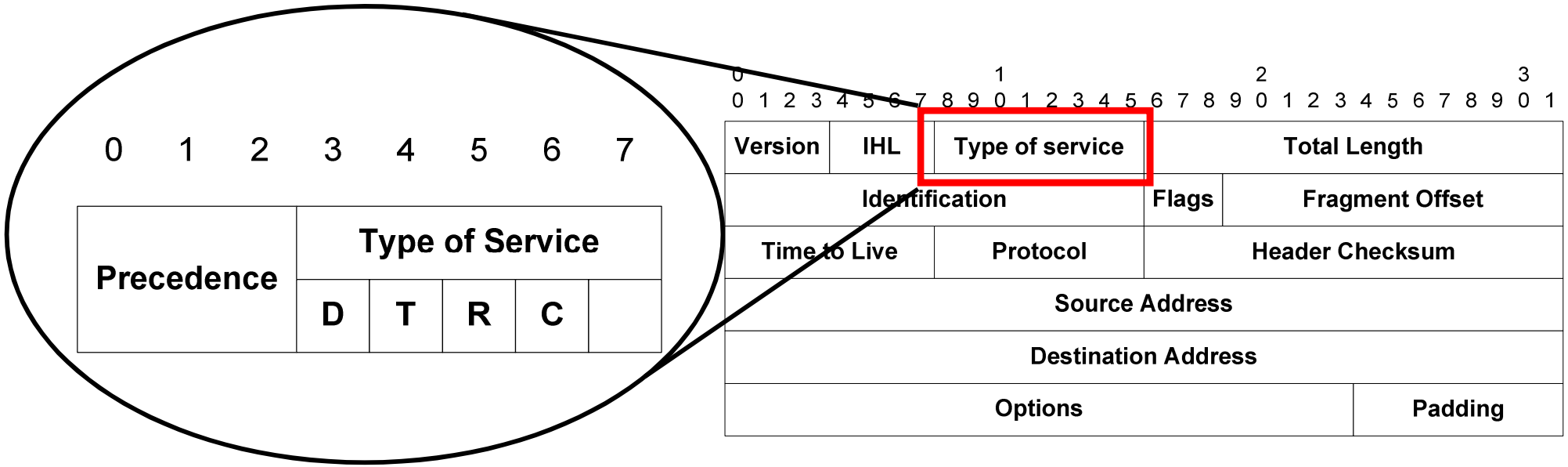
Campos do Cabeçalho IP

Instituto de computação

- *Header Checksum*
  - Calculado como:
    - Complemento a 1 da soma em complemento a 1 de todas as palavras de 16 bits do cabeçalho
      - Considera os bits do *checksum* em 0
      - Considera o campo de opção
  - Compromisso
    - Não protege contra inserção de palavras em zero (16 bits iguais a zero) ou inversão de palavras...
    - Mas é de simples implementação
  - Calculado a cada salto
    - Caso a verificação falhe, a mensagem é descartada
    - Se não falhar, o *checksum* é recalculado
      - Campo TTL é decrementado a cada salto

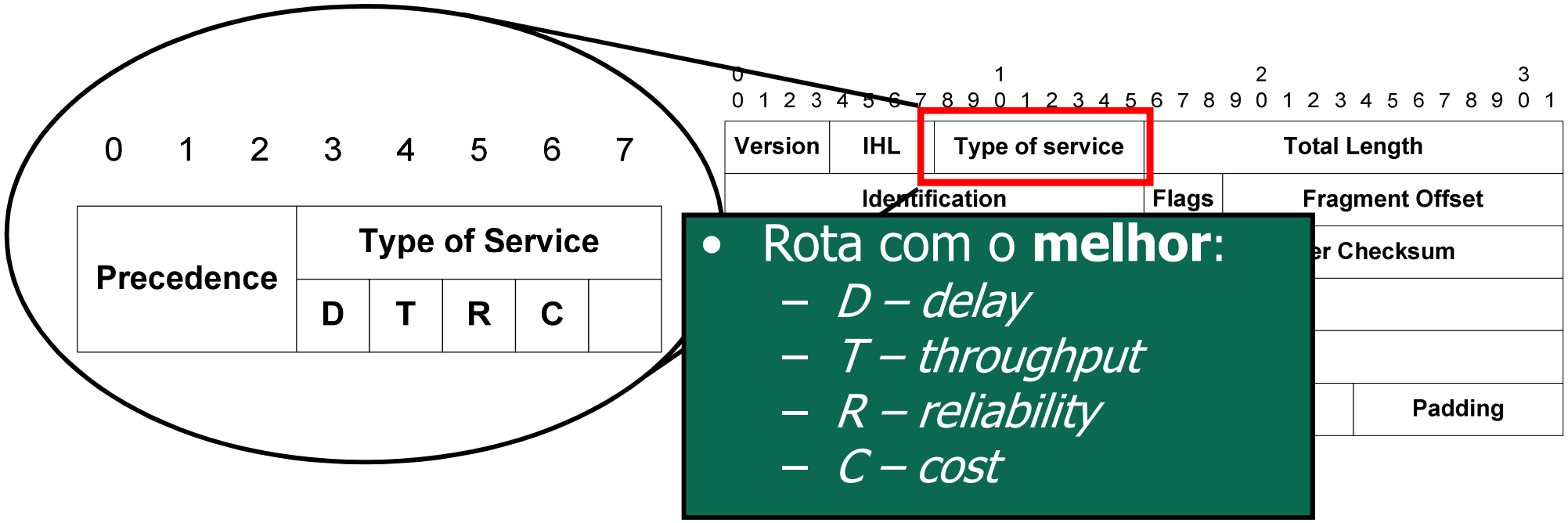
Precedência e Tipo de Serviço

- *Precedence (3 bits)*
  - Indica a prioridade de transmissão do pacote em fila
  - Valores maiores, maior prioridade
  - RFC791 diz que a precedência é válida apenas dentro de uma rede
    - Evita usuários mal-intencionados



Precedência e Tipo de Serviço

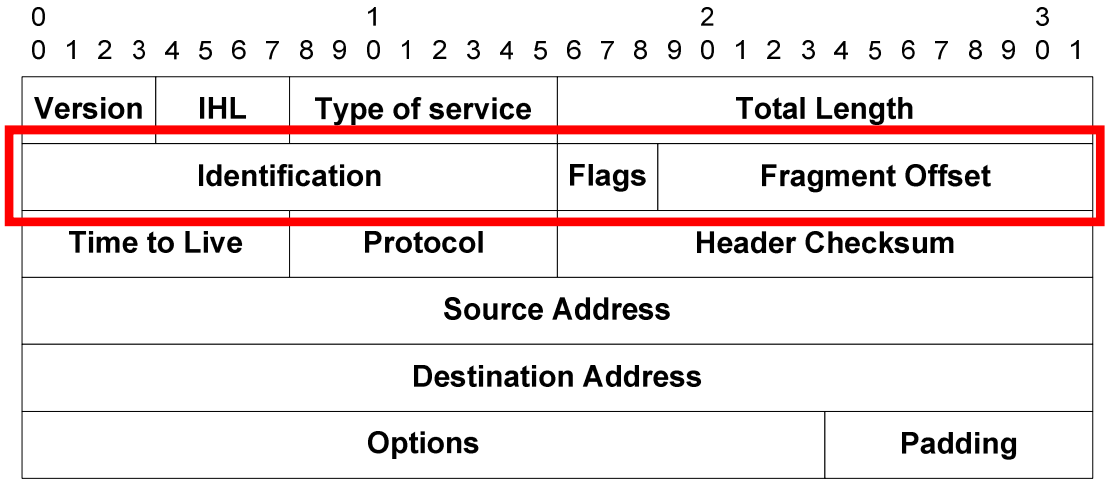
- *Type of Service (5 bits)*
  - Útil quando existem múltiplas rotas
  - Indicação para o roteamento
    - Nunca são utilizados mais de um campo
      - Combinação ilegal



- Rota com o **melhor**:
  - *D – delay*
  - *T – throughput*
  - *R – reliability*
  - *C – cost*

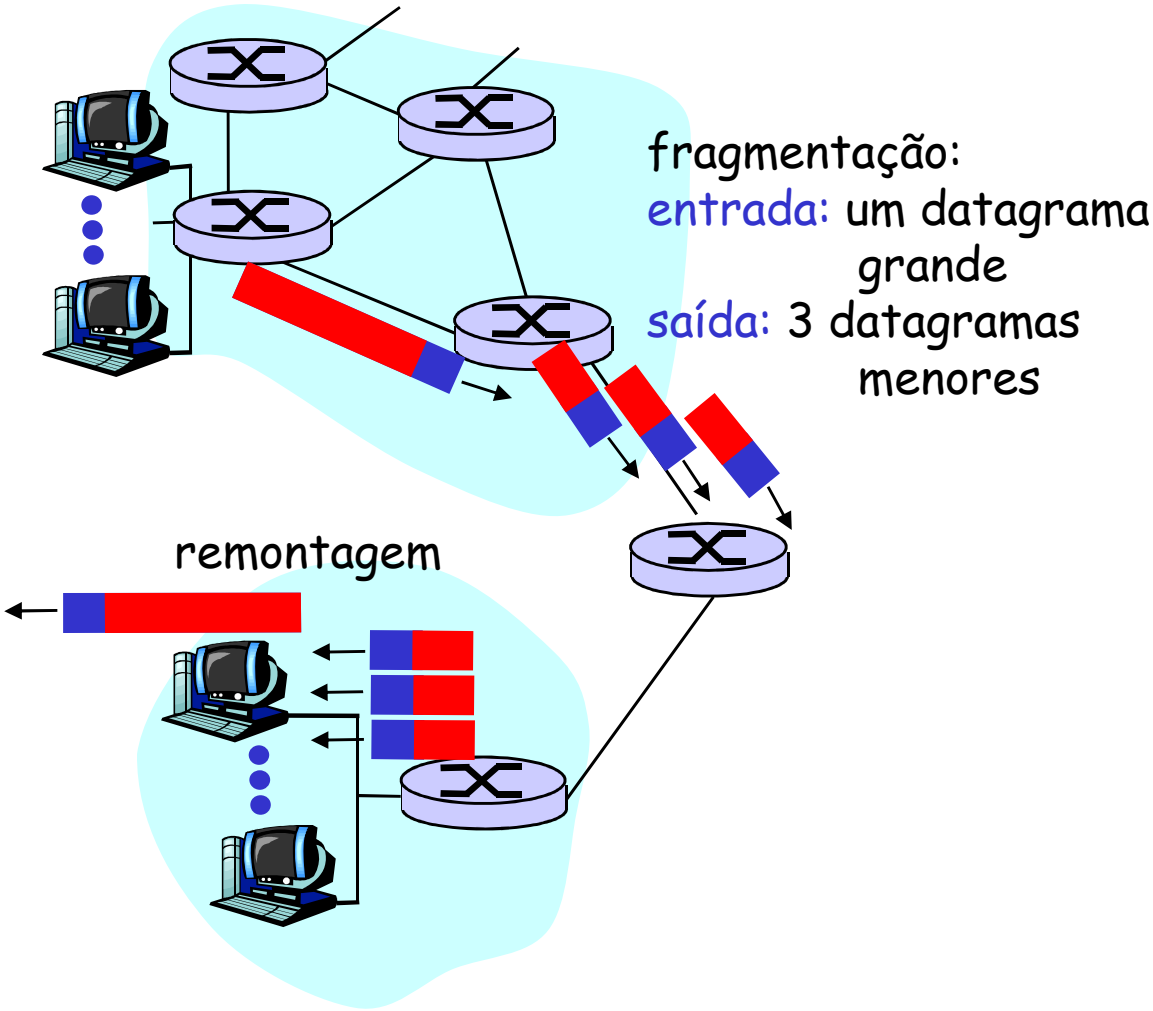
Fragmentação e Remontagem

- A fragmentação é necessária
  - Roteador conecta duas tecnologias de rede diferentes
    - Cada uma possui um tamanho máximo de pacote
      - Ex.: Rede com alta perda → pacotes devem ser pequenos
        Rede com baixa perda → pacotes podem ser grandes



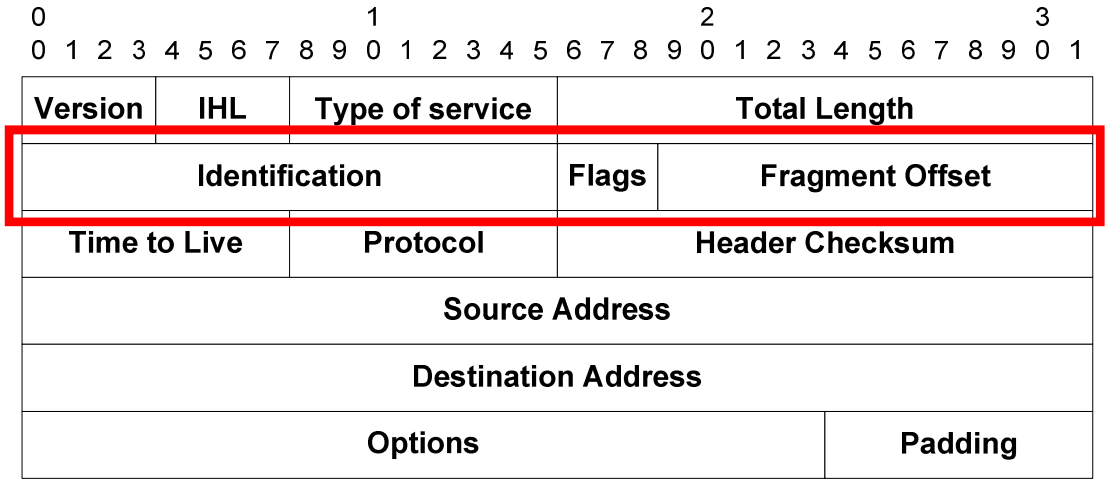
Fragmentação e Remontagem

- Cada enlace de rede tem MTU (*max.transmission unit*) - maior tamanho possível de quadro neste enlace
  - Tipos diferentes de enlace têm MTUs diferentes
- Datagrama IP muito grande dividido ("fragmentado") dentro da rede
  - Um datagrama vira vários datagramas
  - "Remontado" apenas no destino final
  - Bits do cabeçalho IP usados para identificar, ordenar fragmentos relacionados



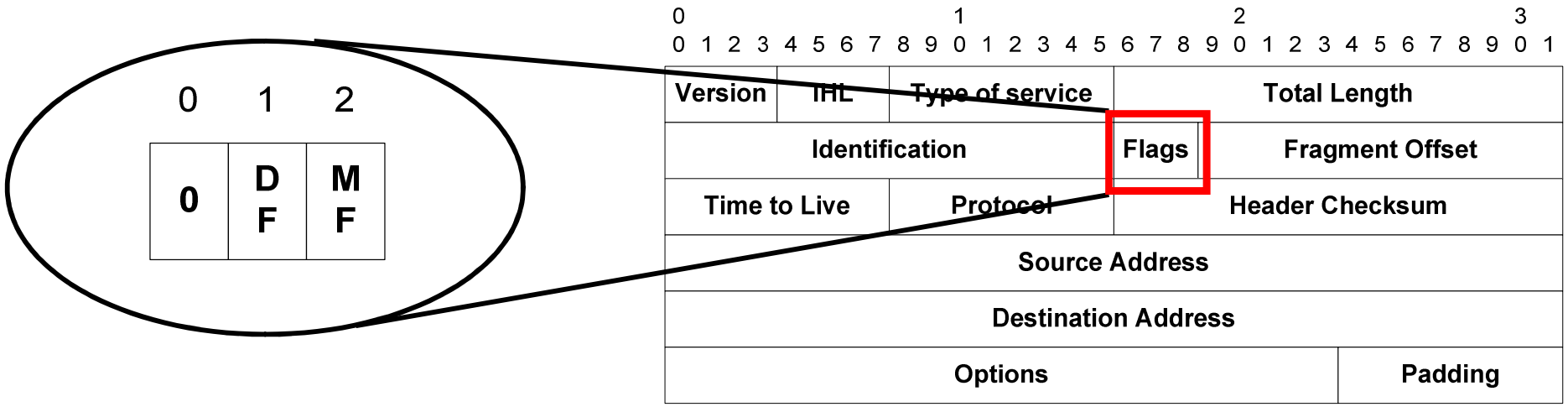
Fragmentação e Remontagem

- *Identification* (16 bits)
  - Junto ao campo endereço de origem, identifica a qual pacote pertence o fragmento
- *Fragment Offset* (13 bits)
  - Identifica a posição do fragmento no pacote
  - Palavras de 8 bits



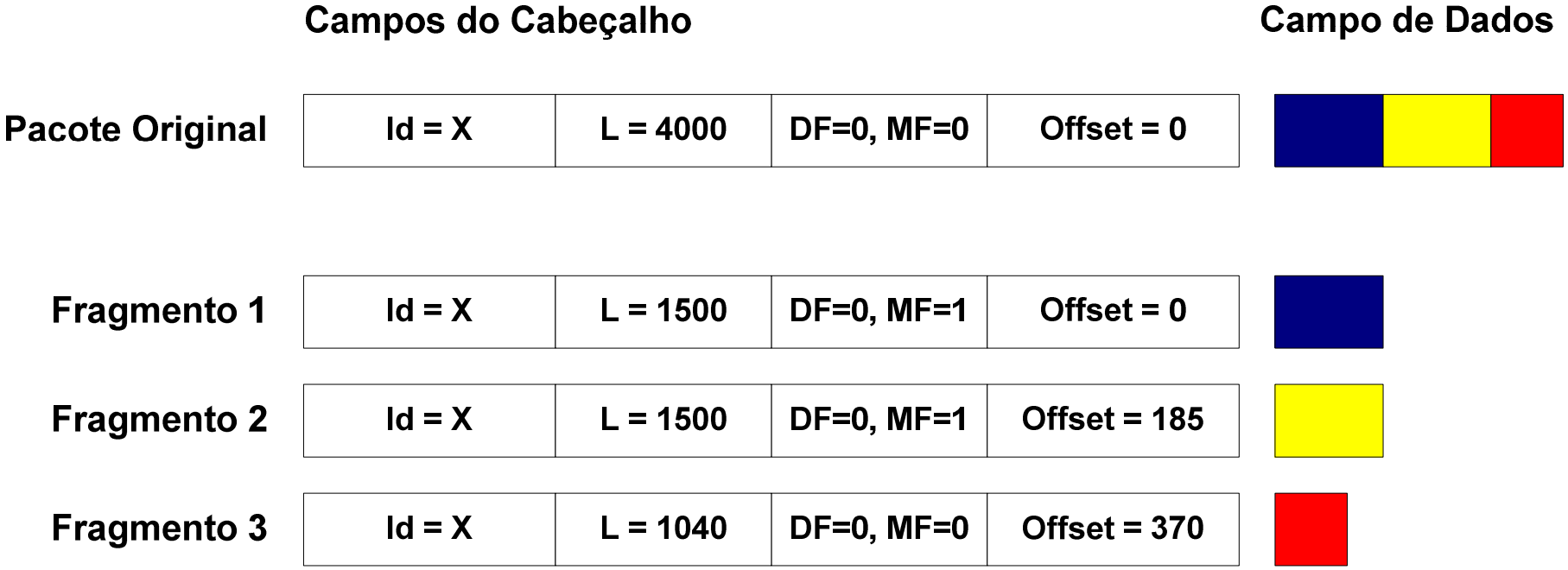
Fragmentação e Remontagem

- *Flags* (3 bits)
  - Informa se o pacote pode ser fragmentado (DF) e se ainda existem mais fragmentos a serem recebidos (MF)
    - Bit 0 – reservado
    - Bit 1 – *don´t fragment* (DF)
    - Bit 2 – *more fragments* (MF)



Fragmentação e Remontagem

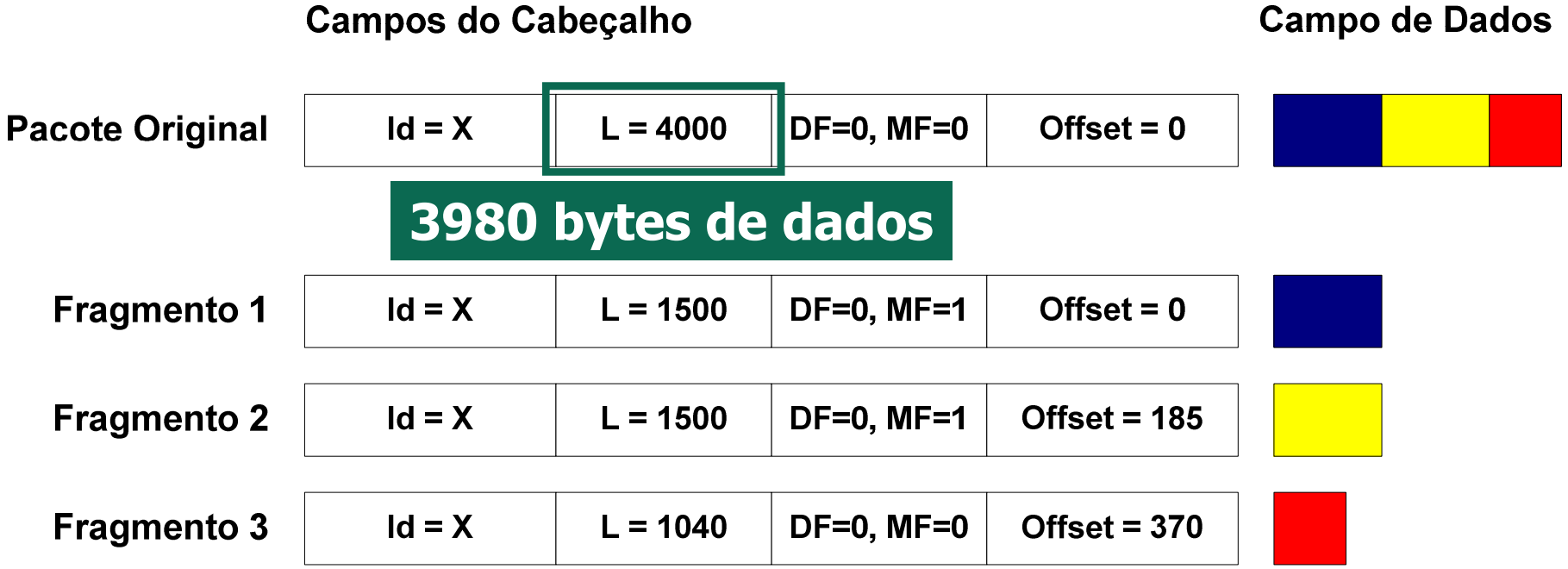
- Cada fragmento possui um cabeçalho completo
  - Igual ao do pacote original, exceto pelos campos de comprimento, *offset* e o bit MF



Datagrama de 4000 bytes e MTU = 1500 bytes

Fragmentação e Remontagem

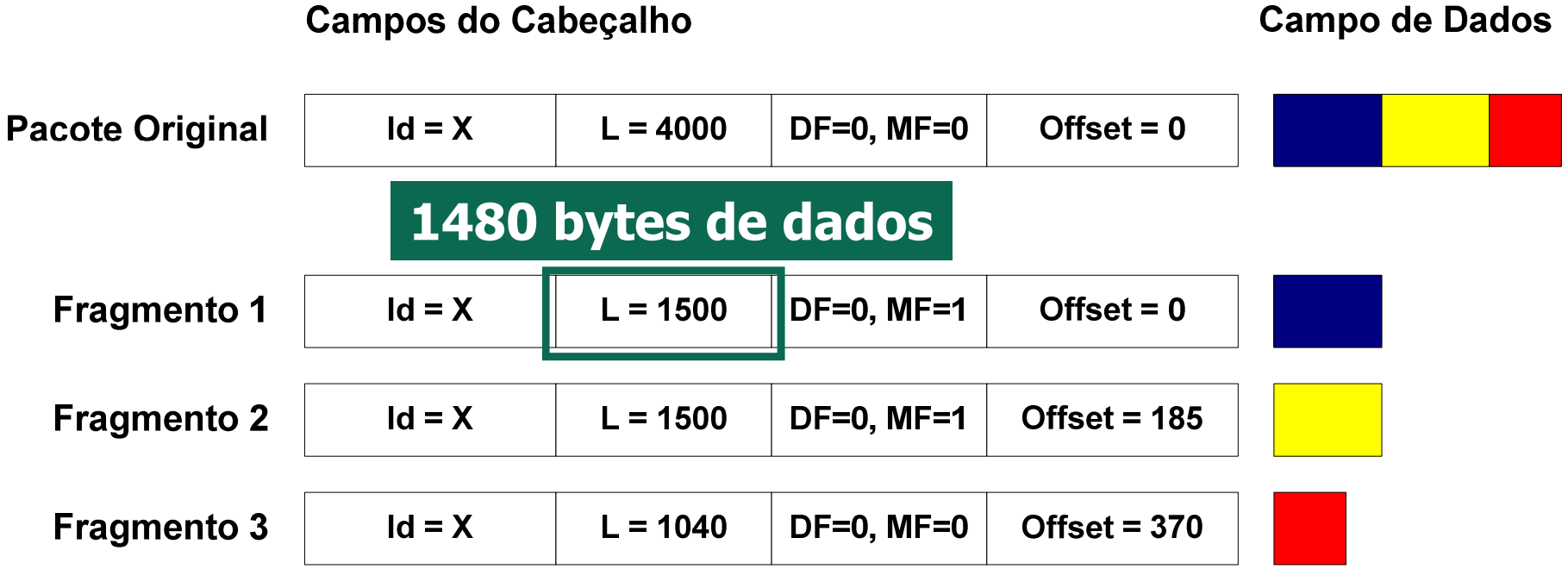
- Cada fragmento possui um cabeçalho completo
  - Igual ao do pacote original, exceto pelos campos de comprimento, *offset* e o bit MF



Datagrama de 4000 bytes e MTU = 1500 bytes

Fragmentação e Remontagem

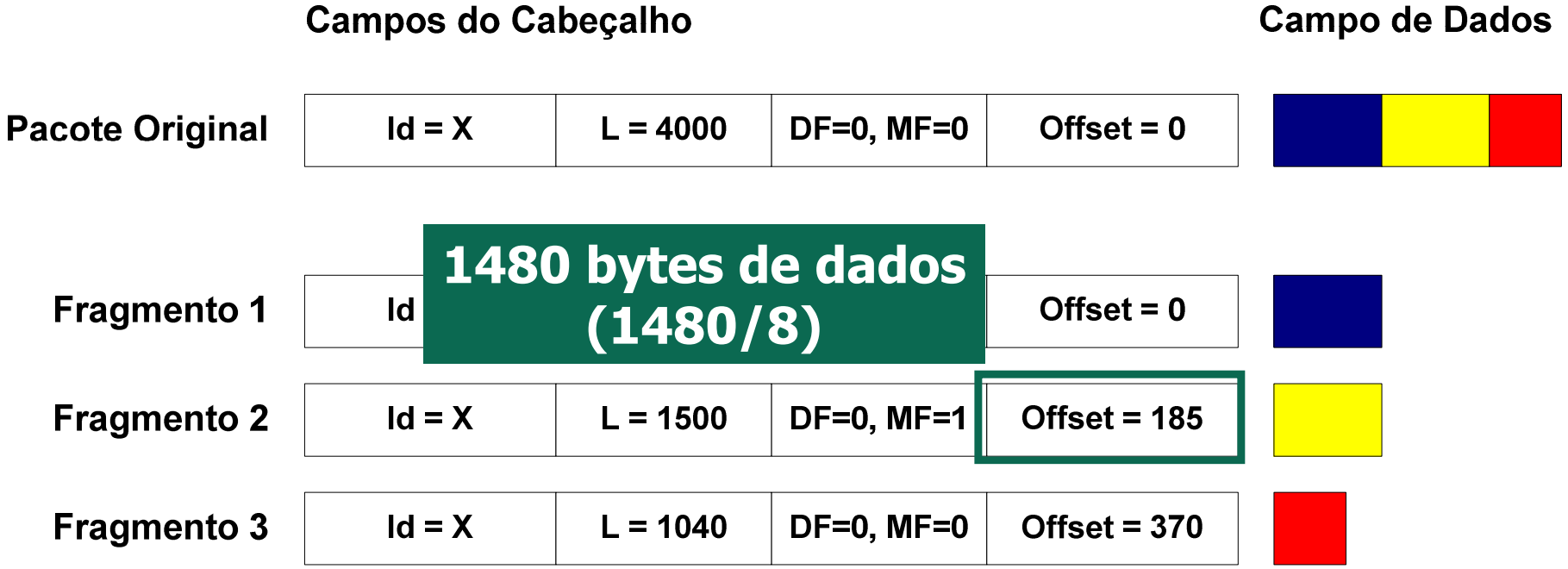
- Cada fragmento possui um cabeçalho completo
  - Igual ao do pacote original, exceto pelos campos de comprimento, *offset* e o bit MF



Datagrama de 4000 bytes e MTU = 1500 bytes

Fragmentação e Remontagem

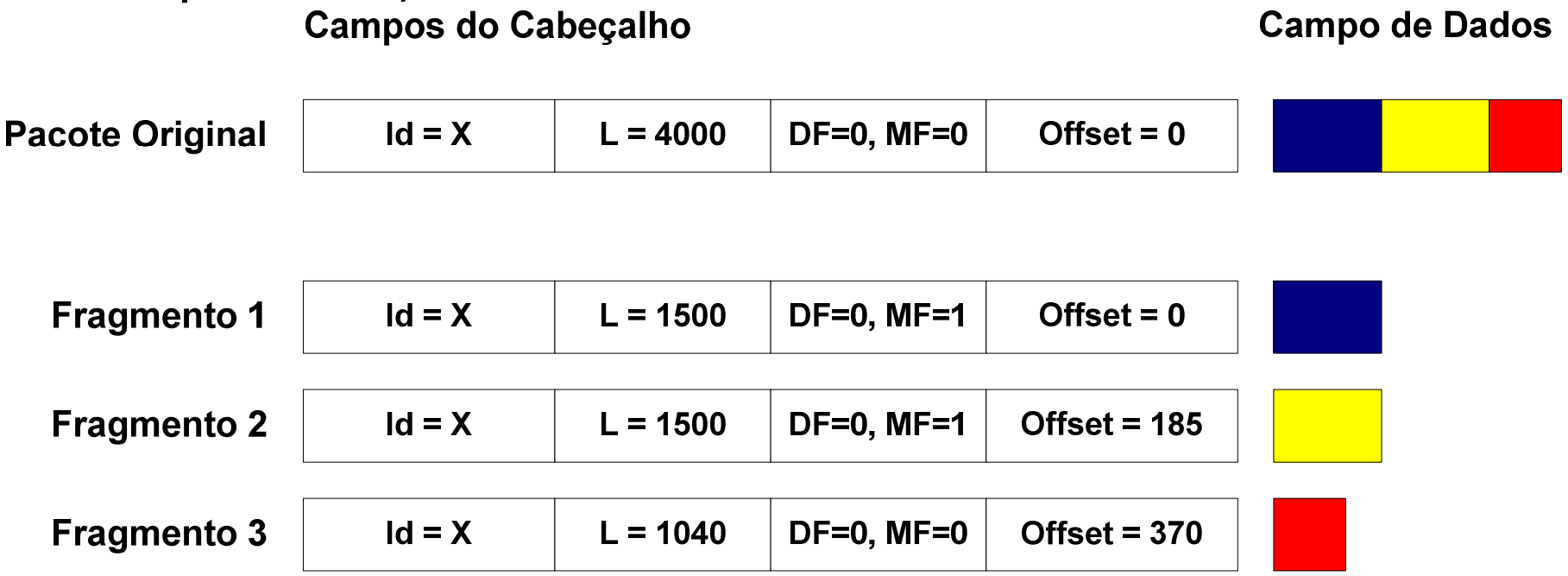
- Cada fragmento possui um cabeçalho completo
  - Igual ao do pacote original, exceto pelos campos de comprimento, *offset* e o bit MF



Datagrama de 4000 bytes e MTU = 1500 bytes

Fragmentação e Remontagem

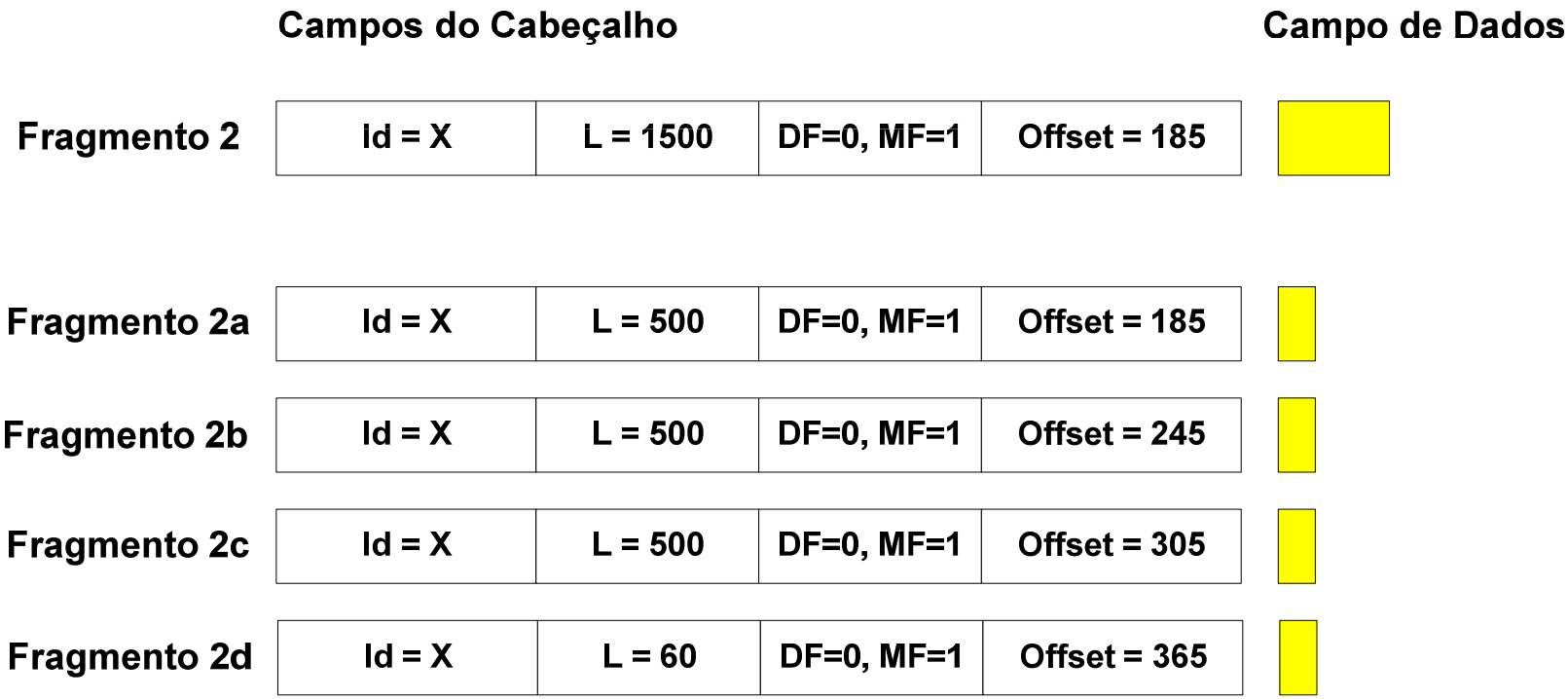
- Cada fragmento possui um cabeçalho completo
  - Igual ao do pacote original, exceto pelos campos de comprimento, *offset* e o bit MF



O bit MF é sempre 1, exceto no último fragmento

Fragmentação e Remontagem

- Em caso de nova fragmentação
  - MF e *offset* são calculados com relação ao pacote original



Fragmentação e Remontagem

Instituto de computação

- O campo identificação (16 bits) associado ao endereço de origem identifica a qual pacote pertence o fragmento

- Pacotes são remontados no destino
  - O receptor deve "expirar" pacotes **parcialmente** remontados, após um certo período de espera
    - Ex.: decrementando o campo TTL a cada segundo
  - O emissor só pode reutilizar um identificador após o período igual ao TTL utilizado

Fragmentação e Remontagem

Instituto de computação

- A reutilização dos identificadores limita a taxa de transmissão possível
  - 16 bits = 65.536 pacotes por TTL
  - TTL recomendado pelo TCP = 2 min
  - Limite de 544 pacotes por segundo
    - 17 Mb/s com pacotes de 4 kB

Fragmentação e Remontagem

Instituto de computação

- A fragmentação é ineficiente combinada com o TCP
  - Perda de um fragmento implica retransmissão do pacote inteiro
- A memória dos roteadores pode ser desperdiçada
  - Os fragmentos de um determinado pacote ficam armazenados antes de serem retransmitidos

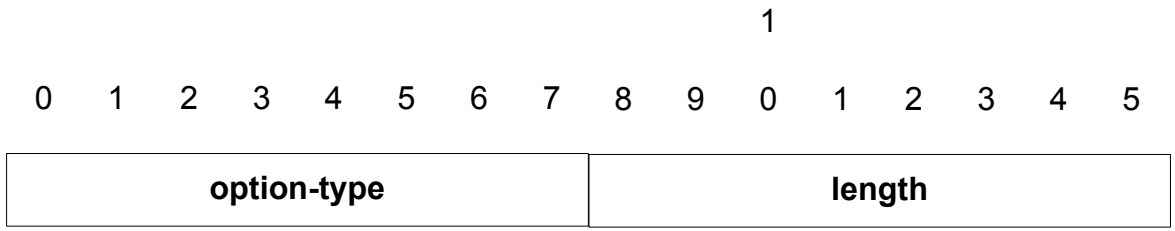
Como Evitar a Fragmentação?

Instituto de computação

- O TCP implementa um mecanismo de descoberta da MTU (*Maximum Transmission Unit*) do caminho
  - Tentativas com diferentes tamanhos de pacote e com o campo DF em 1
    - O TCP utiliza como MTU o maior tamanho entregue

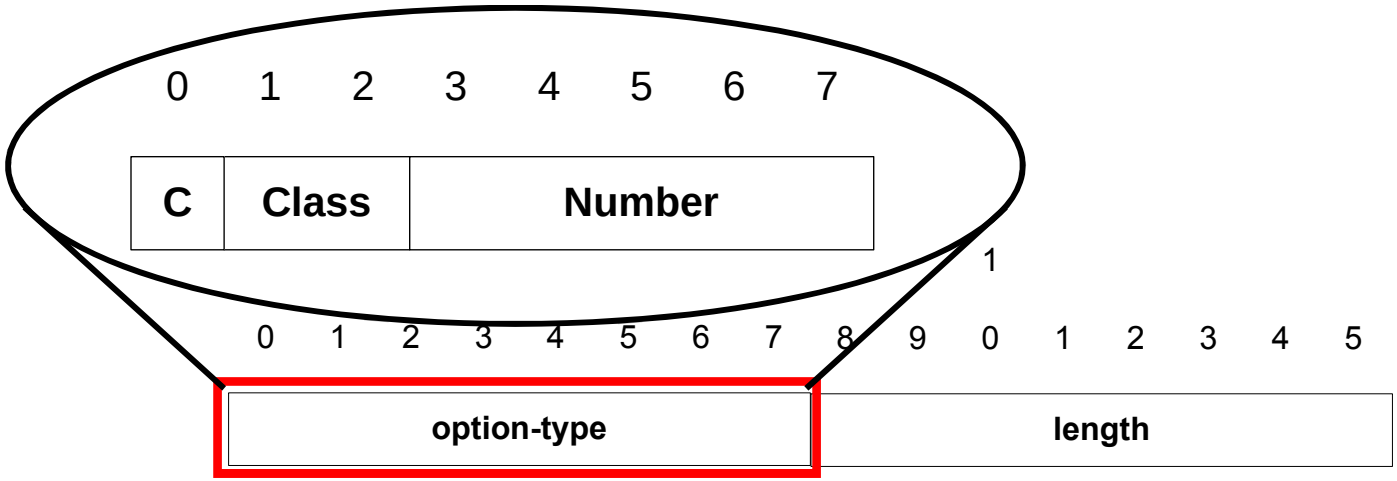
Opções do IP

- Definido para criação de funcionalidades especiais, através do roteamento específico de alguns pacotes
- *Options*
  - Pode transportar vários parâmetros
  - Cada opção começa por um byte de "tipo de opção"
  - O segundo byte normalmente indica o comprimento da opção



Opções do IP

- Flag C (*Copied*)
  - Indica que a opção deve ser copiada em todos os fragmentos ou apenas no primeiro
- Class
  - 0: opções de controle e 2: opções de debug e medidas
- Number
  - Identifica uma opção dentro de cada classe



Opções do IP

Classe	Número	Compr.	Significado
0	0	-	**End of Option list.** Indica o fim da lista de opções, possui apenas 1 byte. Não há byte de comprimento.
0	1	-	**No Operation.** Possui apenas 1 byte. Não há byte de comprimento.
0	2	11	**Security.** Utilizada para carregar parâmetros de segurança definidos pelo dep. de defesa americano.
0	3	var.	**Loose Source Routing.** Utilizada para rotear o pacote IP de acordo com a informação fornecida pela fonte.
0	7	var.	**Record Route.** Utilizada para registrar a rota atravessada pelo pacote IP.
0	8	4	**Stream ID.** Utilizada para carregar o identificador do stream.
0	9	var.	**Strict Source Routing.** Utilizada para rotear o pacote IP de acordo com a informação fornecida pela fonte.
2	4	var.	**Internet Timestamp.**

Opções do IP

Classe	Número	Compr.	Significado
0	0	-	**End of Option list.** Indica o fim da lista de opções, possui apenas 1 byte. Não há byte de comprimento.
0	1	-	**No Operation.** Possui apenas 1 byte. Não há byte de comprimento.
0	2	11	**Security.** Utilizada para carregar parâmetros de segurança definidos pelo dep. de defesa americano.
0	3	var.	**Loose Source Routing.** Utilizada para rotear o pacote IP de acordo com a informação fornecida pela fonte.
0	7	var.	**Record Route.** Utilizada para registrar a rota atravessada pelo pacote IP.
0	8	4	**Stream ID.** Utilizada para carregar o identificador do stream.
0	9	var.	**Strict Source Routing.** Utilizada para rotear o pacote IP de acordo com a informação fornecida pela fonte.
2	4	var.	**Internet Timestamp.**

Opções do IP

Instituto de computação

- *No operation*
  - Utilizada para enchimento entre opções, de forma que o início da opção esteja alinhado em 32 bits
- *End of option*
  - Indica o ponto onde a opção termina, mesmo se o campo IHL indicar mais espaço alocado para opções
- A maioria das opções não é usada
  - Stream ID foi usada apenas no experimento Satnet
  - Security codifica necessidades militares dos anos 70
  - *Timestamp* e *route record* visavam serviços que o programa **`traceroute`** implementa

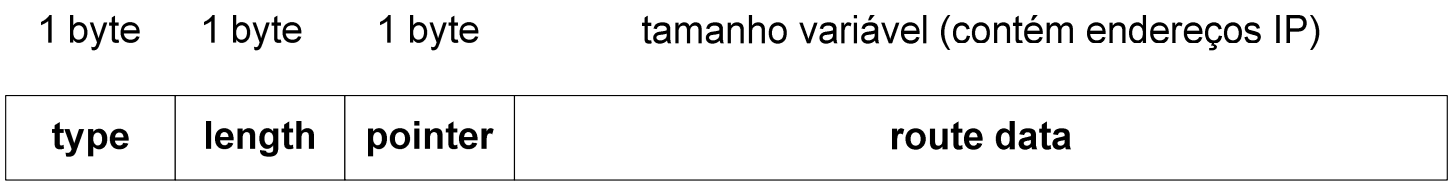
Roteamento pela Fonte

Instituto de computação

- Caminho do pacote é definido no nó de origem

- Duas possibilidades
  - *Strict Routing*
    - Define a caminho completo
  - *Loose Routing*
    - Define alguns nós do caminho

Roteamento pela Fonte



Campo de opções

- *Route data*
  - Contém a lista de endereços pelos quais o pacote deve passar
- *Pointer*
  - Aponta para o próximo endereço da lista a ser utilizado

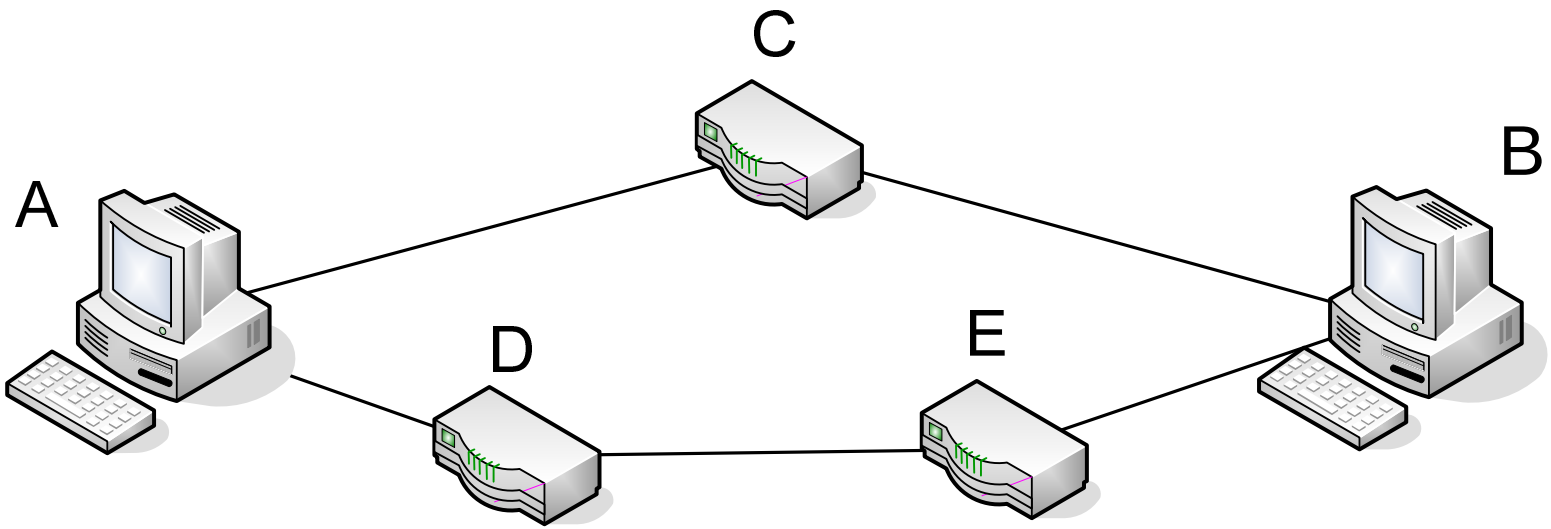
Roteamento pela Fonte

Instituto de computação

- Funcionamento
  - O campo *Destination Address* do cabeçalho possui o endereço IP do próximo nó pelo qual o pacote deve passar
  - Quando este destino é atingido, a opção é examinada
  - O campo *pointer* indica um número de octetos a partir do início da opção, de onde deve ser lido o próximo endereço
  - Se *pointer* maior que o comprimento da opção
    - O destino final foi atingido

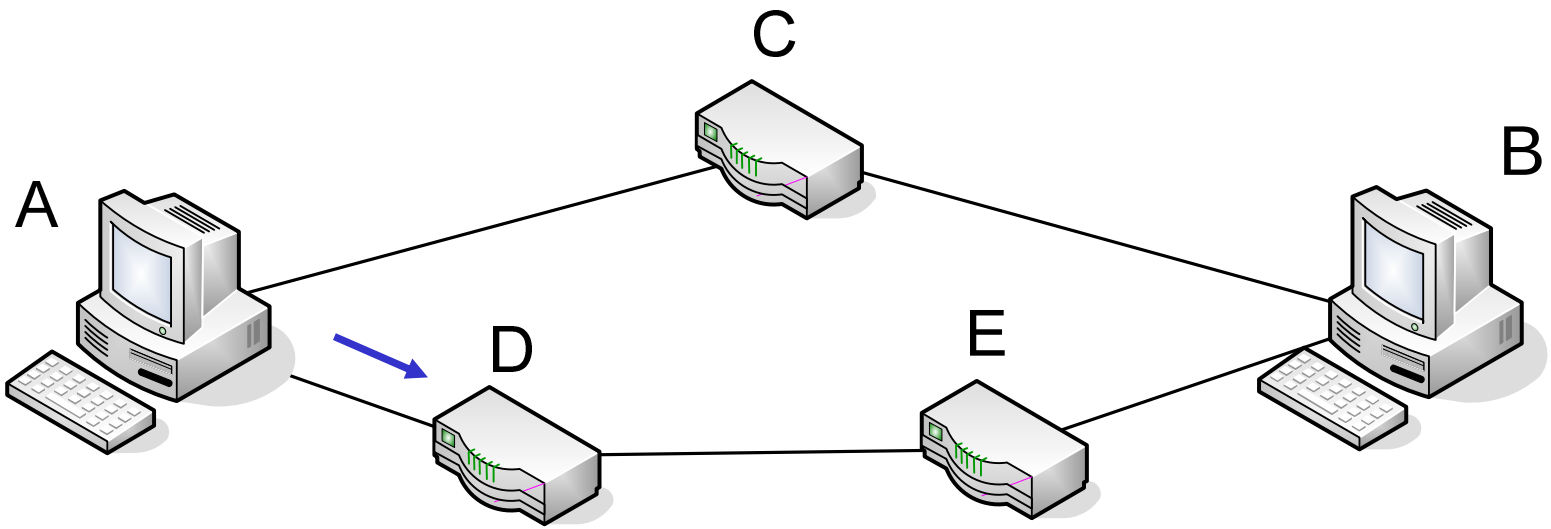
Como Evitar o *Source Routing*?

- Como enviar um pacote de A para B, passando pelos roteadores D e E
  - Encapsulamento IP sobre IP → tunelamento



Como Evitar o *Source Routing*?

- Como enviar um pacote de A para B, passando pelos roteadores D e E
  - Encapsulamento IP sobre IP → tunelamento

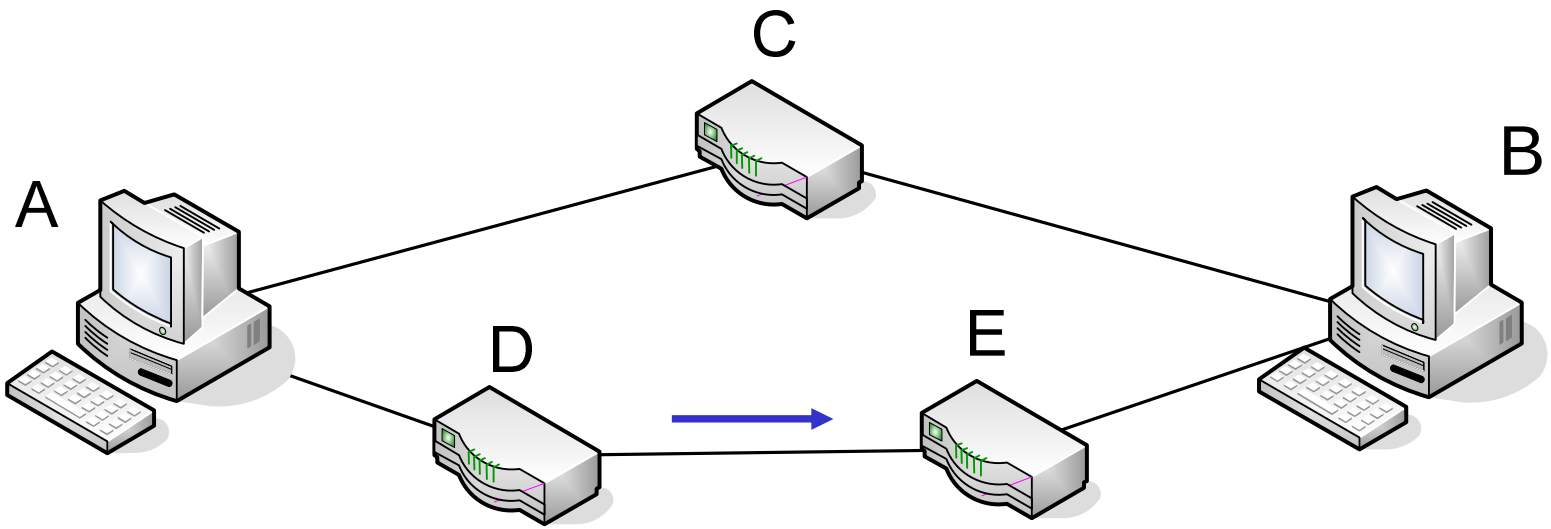


A → D, IP, IP	A → E, IP,IP	A → B, TCP	cabeçalho TCP + dados

Cabeçalho IP(1) Cabeçalho IP(2) Cabeçalho IP(3)

Como Evitar o *Source Routing*?

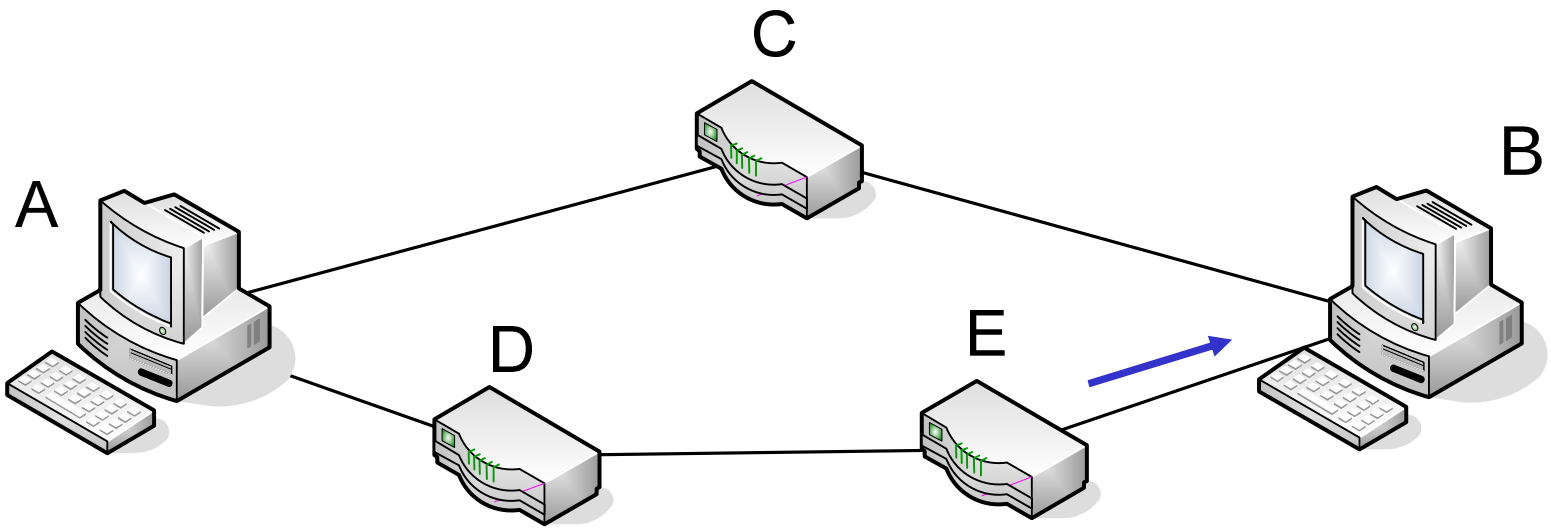
- Como enviar um pacote de A para B, passando pelos roteadores D e E
  - Encapsulamento IP sobre IP → tunelamento



Cabeçalho IP(1) **Cabeçalho IP(2)**

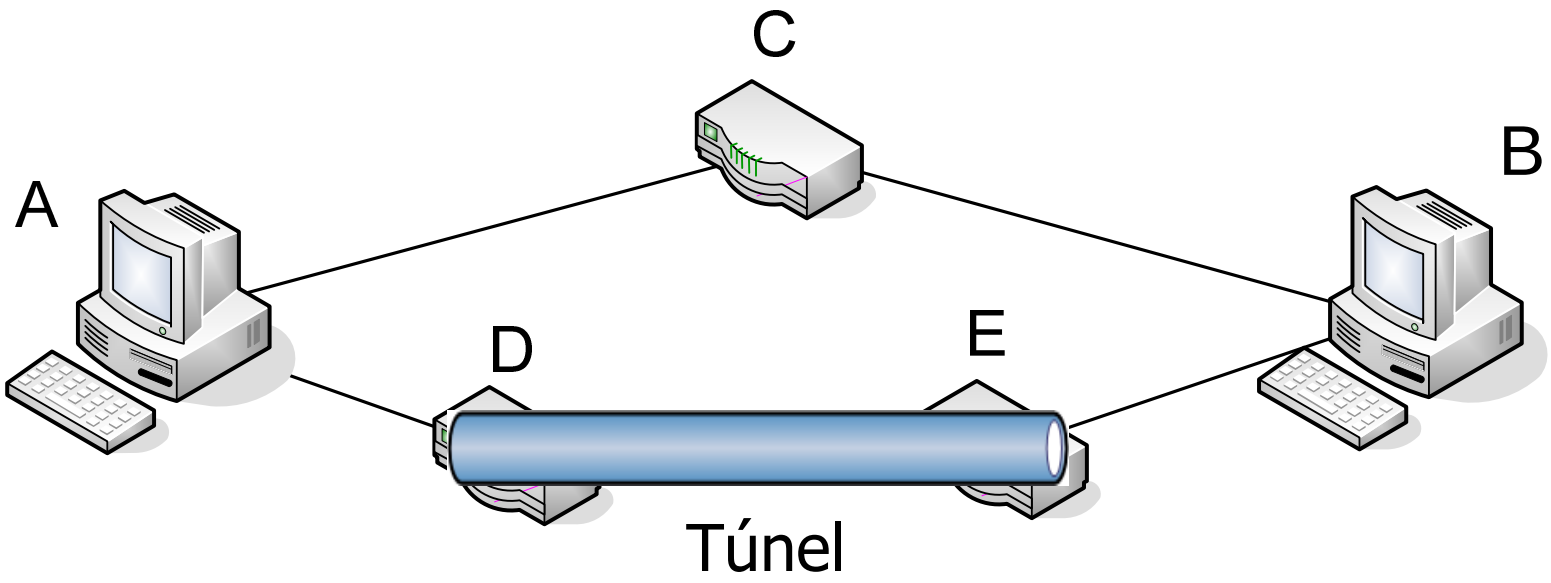
Como Evitar o *Source Routing*?

- Como enviar um pacote de A para B, passando pelos roteadores D e E
  - Encapsulamento IP sobre IP → tunelamento



Como Evitar o *Source Routing*?

- Como enviar um pacote de A para B, passando pelos roteadores D e E
  - Encapsulamento IP sobre IP → tunelamento



Processamento do Cabeçalho IP

Instituto de computação

- Operações para encaminhar um pacote
  1. Verificação da versão, do *checksum*, tamanho do pacote, e leitura das opções (se houver)
  2. Consultar a tabela de roteamento para o destino e tipo de serviço do pacote
  3. Obter a interface e endereço no meio físico

Processamento do Cabeçalho IP

Instituto de computação

- Operações para encaminhar um pacote
  1. Verificação da versão, do *checksum*, tamanho do pacote, e leitura das opções (se houver)
  2. Consultar a tabela de roteamento para o destino e tipo de serviço do pacote
  3. Obter a interface e endereço no meio físico

Número grande de operações!

Como encaminhar pacotes a taxas da ordem de Gb/s?

Processamento do Cabeçalho IP

Instituto de computação

- Roteadores otimizam as operações mais comuns (***fast-path***)
  - Ex.: *caches* com rotas mais utilizadas, processamento em paralelo de múltiplos campos
- Pacotes **sem opções**
  - Possuem cabeçalho de tamanho fixo
    - Passam pelo *fast-path*
- Pacotes **com opções**
  - Seguem o caminho "normal"
    - Além disso, em alguns roteadores, pacotes com opções possuem menos prioridade para aumentar o desempenho global

Endereçamento IP

- Cada **interface** de rede é identificada por um **endereço IP** de 32 bits

$$\underbrace{223}_{11011111}.\underbrace{1}_{00000001}.\underbrace{1}_{00000001}.\underbrace{1}_{00000001}$$

Endereçamento IP

Instituto de computação

- Formato do Endereço IP
  - Dividido em duas partes:
    - "identificador de rede" e "identificador de estação"

- 3 classes de "números de rede": A, B e C
- Mais tarde, classe D definida para endereços *multicast*
- A classe E possui endereços reservados para utilização experimental

Classes de Endereços IP

Instituto de computação

Classe	Bits mais significativos	Formato	
A	0	7 bits de redes	24 bits de estações
B	10	14 bits de redes	16 bits de estações
C	110	21 bits de redes	8 bits de estações
D	1110	28 bits de endereços de grupo multicast	
E	1111	reservados para testes	

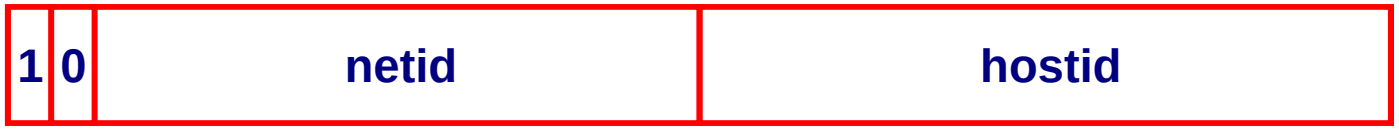
Classes A, B e C

Classe A:

7 bits **24 bits**

2^7 = 128 prefixos de classe A (0.x.x.x a 127.x.x.x)
(2^{24} - 2) = 16.777.214 estações em cada rede

Classe B:



14 bits **16 bits**

2^{14} = 16.384 prefixos de classe B (128.x.x.x a 191.x.x.x)
(2^{16} - 2) = 65.534 estações em cada rede

Classe C:

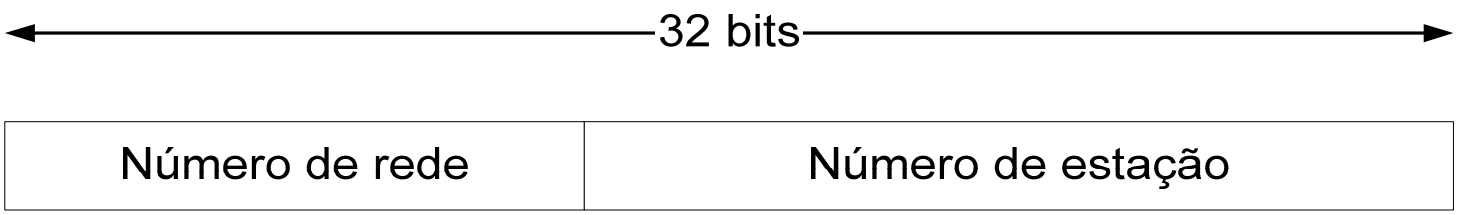
21 bits **8 bits**

2^{21} = 2.097.152 prefixos de classe C (192.x.x.x a 223.x.x.x)
(2^8 - 2) = 254 estações em cada rede

Estrutura de Endereçamento

Instituto de computação

- Quando o IP foi padronizado, em 1981

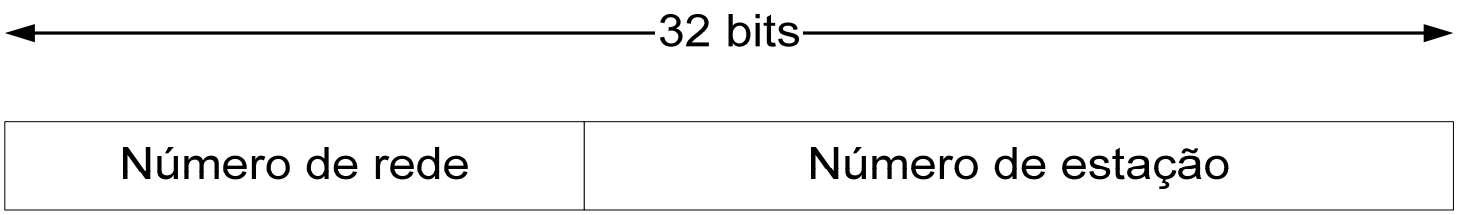


- Números de rede (*netid*)
  - Alocados pela autoridade de numeração Internet
- Números de estação (*hostid*)
  - Alocados pelo gerente de rede

Estrutura de Endereçamento

Instituto de computação

- Quando o IP foi padronizado, em 1981



- Números de rede (*netid*)
  - Alocados pela autoridade de numeração Internet
- Números de estação (*hostid*)
  - Alocados pelo gerente de rede

Unicidade do número de rede + unicidade do número da estação → Garantem a UNICIDADE GLOBAL do endereço IP

Problema das Classes de Endereço

Instituto de computação

- Número fixo de redes e estações por rede
  - Classe A
    - Número pequeno de redes
    - Número excessivo de estações por rede

→ **Desperdício de end. IP**

  - Classe C
    - Número pequeno de estações por rede
    - Número excessivo de redes

→ **Falta de end. IP**

- Resultado

Esgotamento da classe B!

Classless Inter-Domain Routing architecture (CIDR)

Instituto de computação

- Acaba com as classes
  - Introduz o conceito de **máscara de rede**
- Permite
  - Agregação de rotas
    - Aumenta a escalabilidade
      - Reduz o tamanho das tabelas de roteamento
  - Distribuição mais adequada dos endereços IP
    - Resolve o esgotamento dos endereços da classe B
    - Permite melhor planejamento de endereços
      - Número de máquinas vs. número de endereços IP

Estrutura de Endereçamento CIDR

Instituto de computação

- Número de rede de **comprimento variável**

a.b.c.d/x

- Os **x** bits mais significativos do endereço são o número de rede → **prefixo**
- Os 32-**x** bits são o número de estação

200.23.16.0/23

Máscaras de Sub-rede

Instituto de computação

- Uma máscara de sub-rede pode ser representada através da notação:
  - Endereço da rede+sub-rede/<número de bits em 1 da máscara>
    - Ex1.: 192.168.0.0/16 → essa notação é equivalente a dizer que a máscara é 255.255.0.0
    - Ex2.: 192.168.3.0/26 → essa notação é equivalente a dizer que a máscara é 255.255.255.192

Estrutura de Endereçamento CIDR

- Como obter o número de rede/prefixo a partir do endereço IP?

`Prefixo = IP AND Máscara`

200.23.16.1/255.255.254.0

endereço	11001000	00010111	00010000	00000001
máscara	11111111	11111111	11111110	00000000
rede	11001000	00010111	00010000	00000000

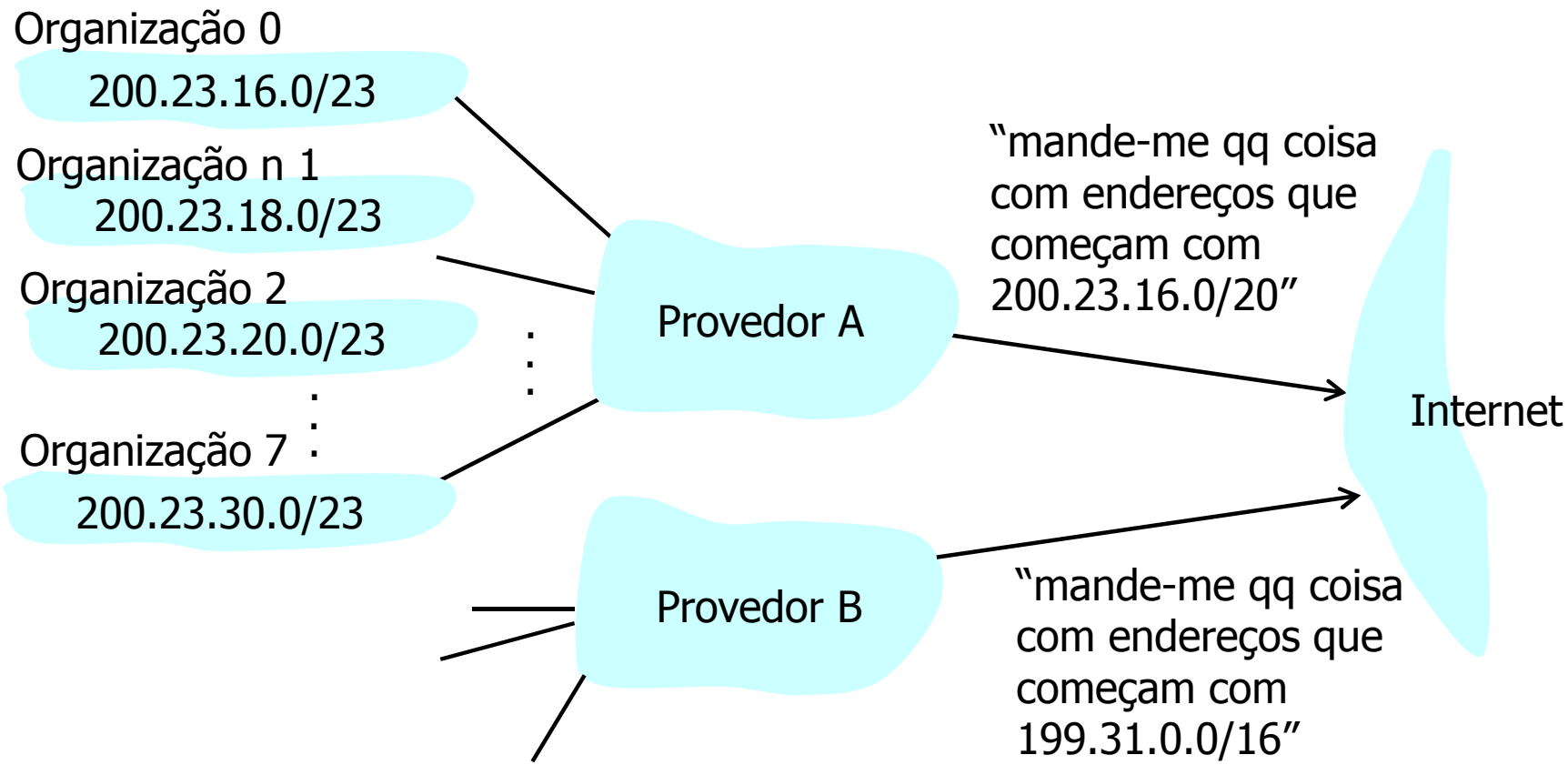
Estrutura de Endereçamento CIDR

- Como um provedor de serviços pode distribuir endereços para suas redes clientes?

Bloco do provedor	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

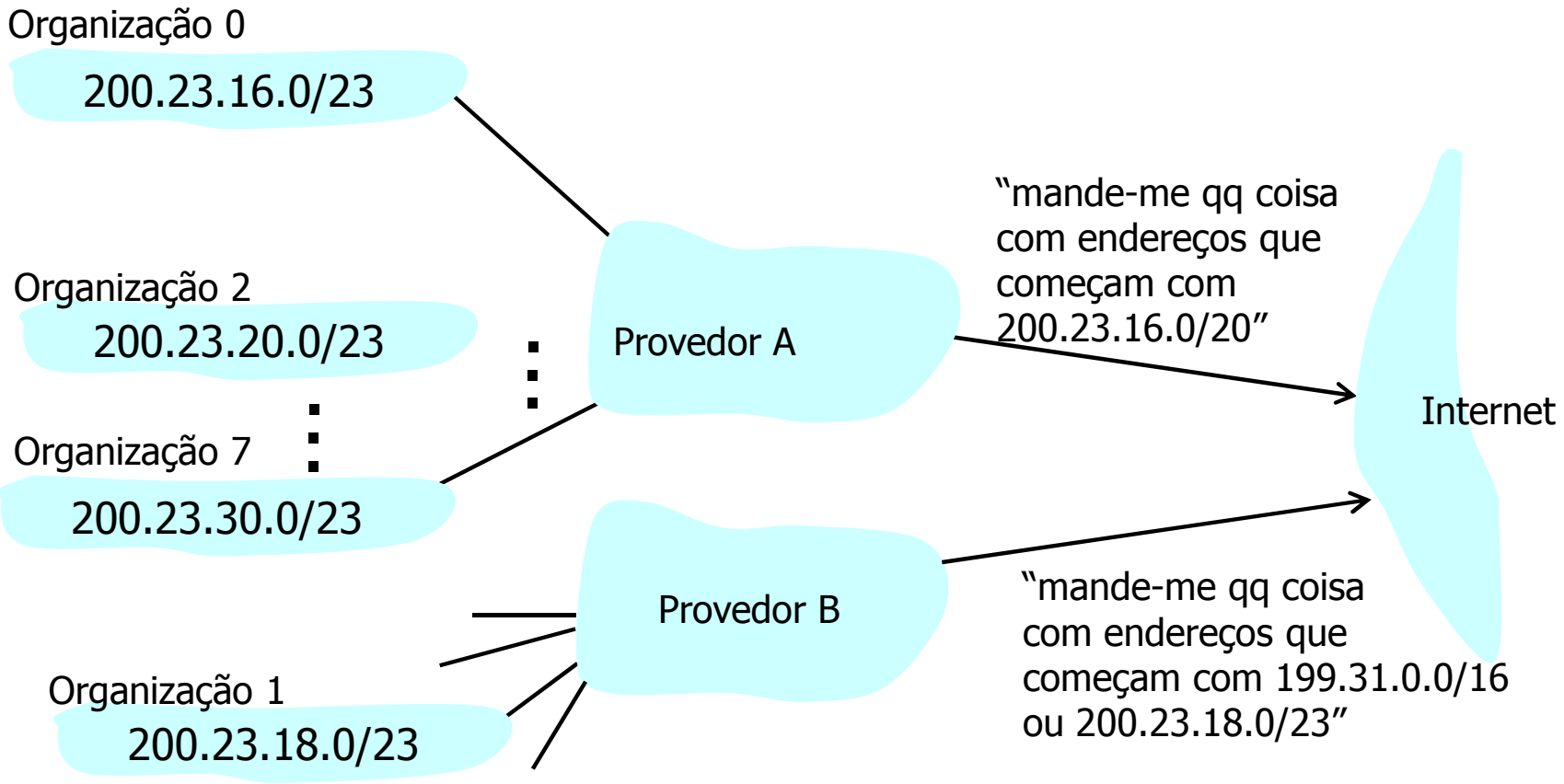
Endereçamento Hierárquico

Endereçamento hierárquico permite anunciar eficientemente informação sobre rotas ➡ **agregação**



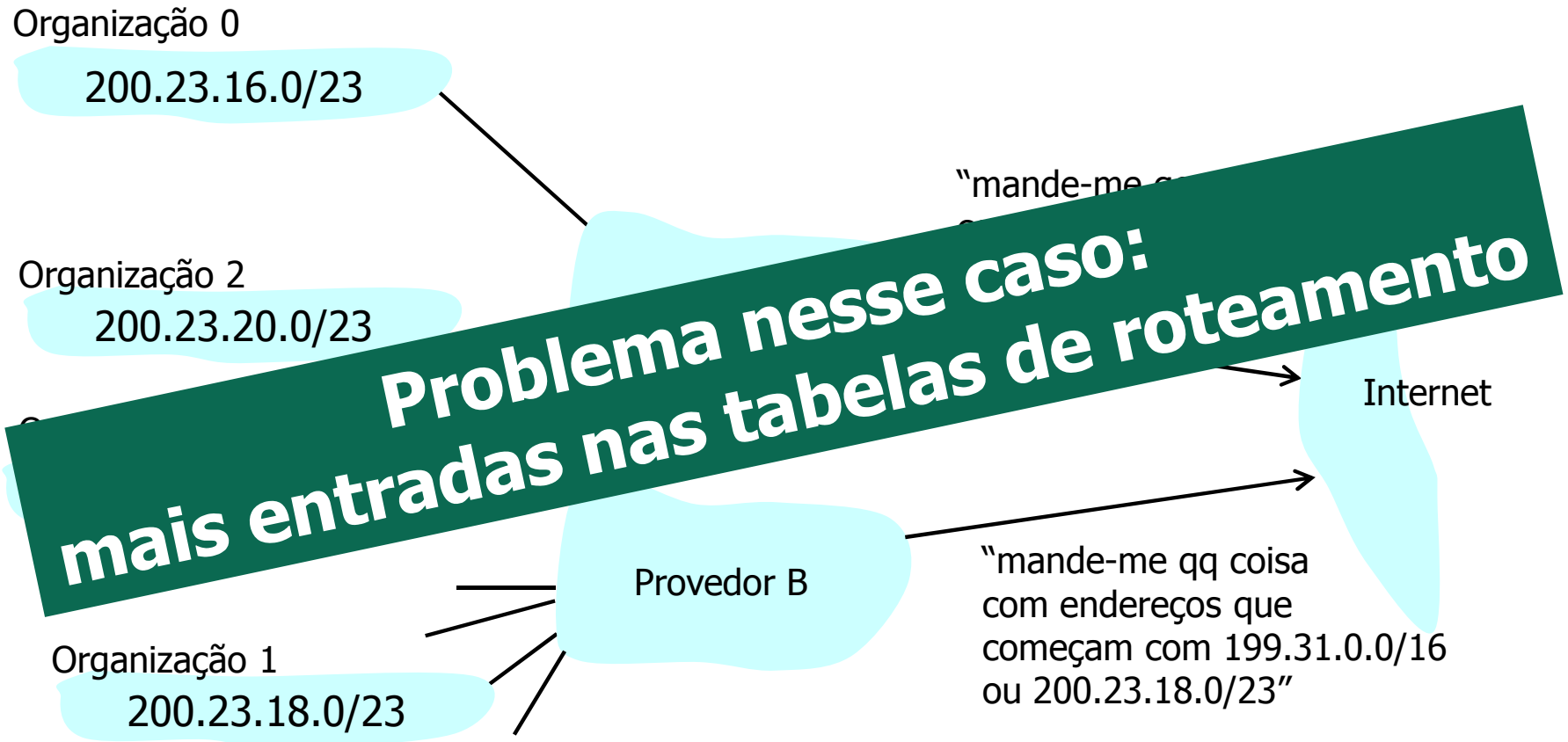
Endereçamento Hierárquico

• Provedor B tem uma rota mais específica (maior prefixo) para a Organização 1 → pacotes encaminhados corretamente



Endereçamento Hierárquico

• Provedor B tem uma rota mais específica (maior prefixo) para a Organização 1 → pacotes encaminhados corretamente



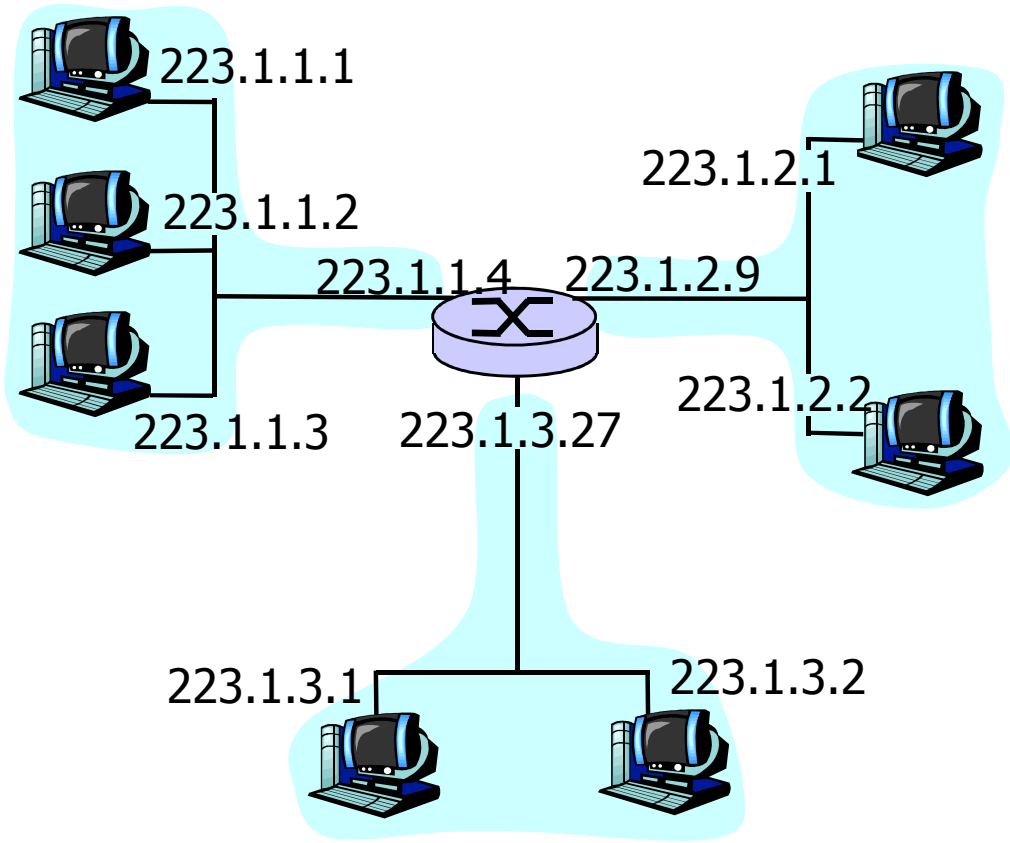
Endereços e Interfaces

Instituto de Computação

- Endereços IP identificam **interfaces de rede**
  - **NÃO** identificam estações
    - Uma única estação pode ter várias interfaces de rede
- Uma estação com várias interfaces de rede possui vários endereços IP
  - Estação *multi-homed*
    - Exs. roteadores, estações que balanceiam o tráfego entre diversas redes
- Cada endereço pertence a uma sub-rede, que geralmente corresponde a uma "rede física"

Sub-redes

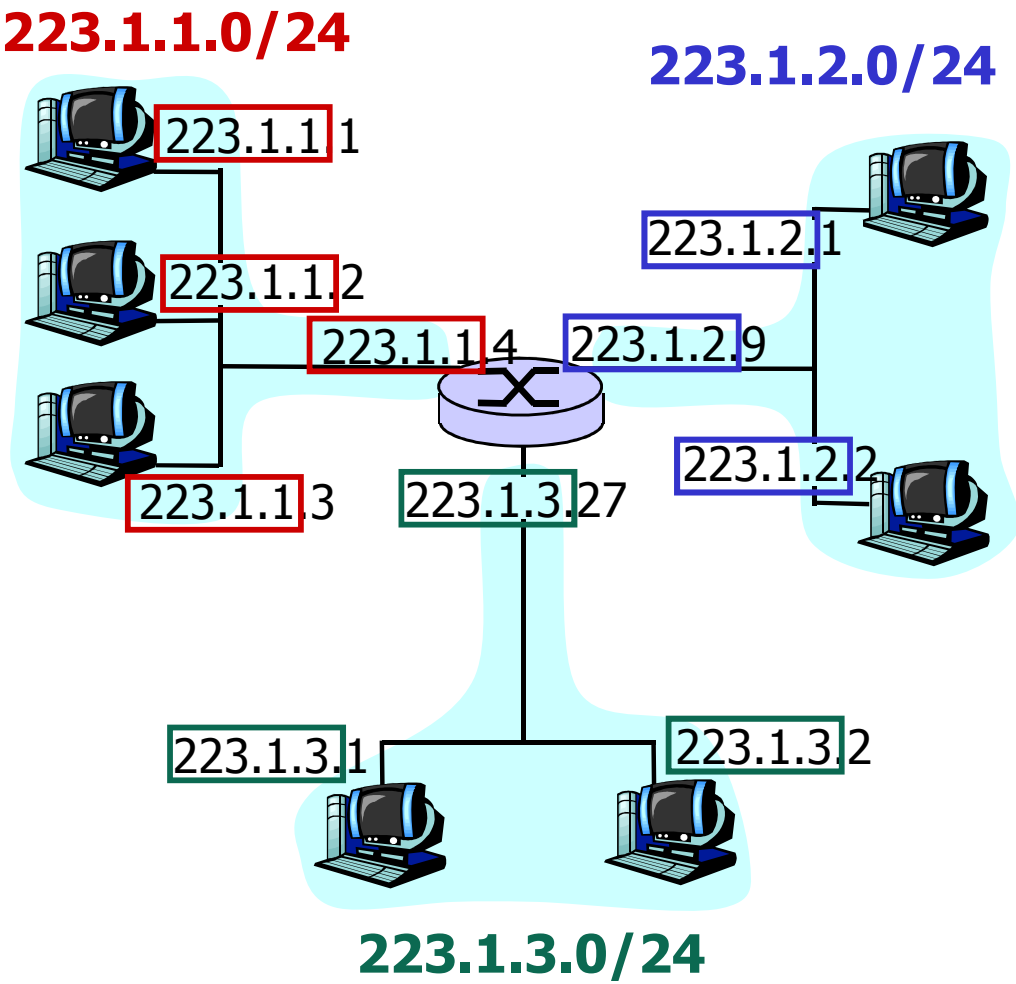
- O que é uma sub-rede IP?
  - Interfaces de dispositivos com a mesma parte de rede nos seus endereços IP
  - Podem alcançar um ao outro sem passar por um roteador



Esta rede consiste de 3 sub-redes IP

Sub-redes

- O que é uma sub-rede IP?
  - Interfaces de dispositivos com a mesma parte de rede nos seus endereços IP
  - Podem alcançar um ao outro sem passar por um roteador



Esta rede consiste de 3 sub-redes IP

Endereços e Interfaces

Instituto de computação

- Entradas na tabela de roteamento dos roteadores
  - Normalmente apontam para **sub-redes**
    - Entretanto, podem eventualmente apontar para endereços de máquinas

```
[user@exemplo ~]$ route -n
Tabela de Roteamento IP do Kernel
Destino      Roteador     MáscaraGen.     Opções Métrica Ref Uso Iface
200.20.10.64 0.0.0.0      255.255.255.224 U      0       0   0   eth0
169.254.0.0  0.0.0.0      255.255.0.0     U      0       0   0   eth0
0.0.0.0      200.20.10.65 0.0.0.0         UG     0       0   0   eth0
```

Endereços e Interfaces

Instituto de Computação

- Por que não um endereço por estação?
  - Um endereço por interface permite ***escolher o caminho*** utilizado para chegar a uma estação
    - Busca do melhor caminho e balanceamento de carga
  - Endereços por interface permitem a ***agregação de endereços*** nas tabelas de roteamento
    - Se os endereços não fossem ligados à topologia, seria necessária uma entrada na tabela de roteamento para cada estação
      - Cada interface pertence a uma sub-rede
  - Um endereço por interface permite **manter conectividade** em caso de falha de uma interface
    - Tolerância a falhas

Endereços e Interfaces

Instituto de computação

- Desvantagens
  - Todos os endereços de uma estação devem ser incluídos no servidor de nomes
    - Para se comunicar com um determinado nó, deve-se saber todos os possíveis endereços desse nó
  - O "melhor endereço" deve ser escolhido para uma conexão
    - Melhor depende de diversos fatores como caminho, requisitos da aplicação etc.
  - O endereço fonte deve ser cuidadosamente escolhido pela aplicação
    - Determina o caminho seguido pelos pacotes de resposta

Endereços Especiais

Instituto de computação

- Endereço de rede
  - Usado para identificar uma rede
  - Geralmente, o primeiro endereço IP da faixa de endereços
    - Ex.: 146.164.0.0

- O "**`0`**" pode ser utilizado como **endereço fonte**, quando o número de rede é desconhecido, portanto:
  - **`0.0.0.0`** significa "esta estação nesta rede"
  - **`0.X.Y.Z`** significa "a estação **`X.Y.Z`** nesta rede"
    - Utilizado por ex. quando uma estação está iniciando

Endereços Especiais

Instituto de computação

- Difusão limitada (*limited broadcast*)
  - Formado por todos os bits em "**1**" – **255.255.255.255**
  - Só pode ser utilizado como **endereço destino**
  - Pacote é enviado a todas as estações da sub-rede
  - Não é retransmitido por um roteador

- Difusão direcionada (*directed broadcast*)
  - Todos os bits da "parte estação" do endereço são colocados em "**1**"
    - Ex. "A.255.255.255", "C.C.C.255"
  - Com sub-redes a mesma regra é válida
    - todos os bits do complemento da máscara são colocados em "**1**"

Endereços Especiais

Instituto de computação

- Conseqüências
  - Não existe sub-rede identificada apenas por **0**'s,
  - Assim como não existe sub-rede identificada apenas por **1**'s
  - O tamanho da sub-rede é maior ou igual a 2 bits

Endereços Especiais

Instituto de computação

- Endereço de *loopback*
  - Na verdade, existe um número de rede de *loopback*:
    - Rede Classe A: "**127**"
- Qualquer endereço da forma "**127.X.Y.Z**" é:
  - Local e não é transmitido para fora da **estação**

Alocação de Endereços IP

Instituto de Computação

- Atualmente
  - ICANN (*The Internet Corporation for Assigned Names and Numbers*)
    - Organização sem fins lucrativos responsável pela
      - Alocação do espaço de endereçamento IP
      - Atribuição de parâmetros de protocolos
      - Gerenciamento do sistema de nomes de domínios
      - Gerenciamento dos servidores raiz
- Anteriormente
  - IANA (*Internet Assigned Numbers Authority*) e outras entidades através de contratos com o governo americano

Alocação de Endereços IP

Instituto de Computação

- Os endereços IP são alocados através de delegações de acordo com uma estrutura hierárquica
  1. Usuários recebem endereços IP de um provedor de serviço (ISP - *Internet Service Provider*)
  2. ISPs obtêm faixas de endereços IP de uma autoridade de registro local (LIR - *Local Internet Registry*), nacional (NIR - *National Internet Registry*), ou regional (RIR - *Regional Internet Registry*)
- O papel do ICANN é alocar faixas de endereços aos RIRs, de acordo com suas necessidades e a partir das faixas de endereços livres

Alocação de Endereços IP

Instituto de computação

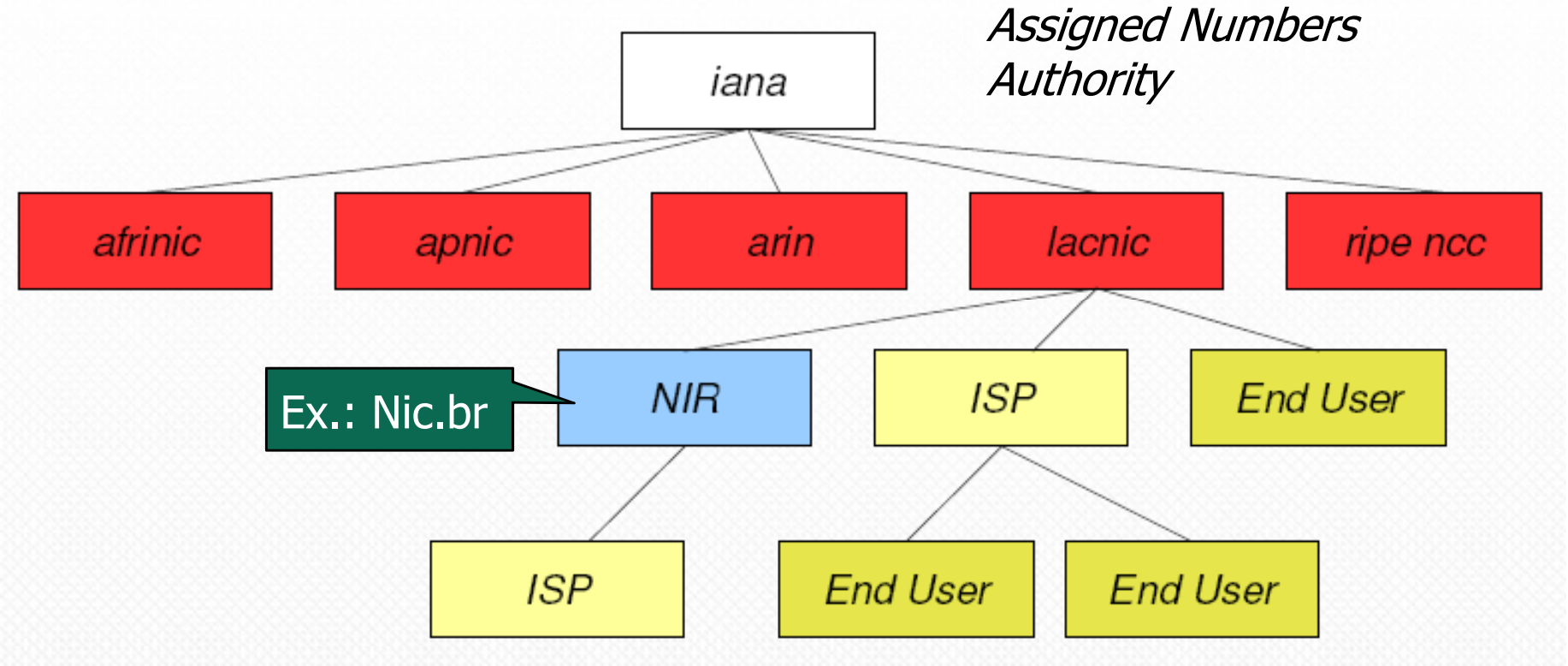
- RIR - *Regional Internet Registry*
  - APNIC (*Asia Pacific Network Information Centre*)
    - Região Ásia/Pacífico
  - ARIN (*American Registry for Internet Numbers*)
    - América do Norte e África ao Sul do Saara
  - LACNIC (*Regional Latin-American and Caribbean IP Address Registry*)
    - América Latina e algumas Ilhas Caribenhas
  - RIPE NCC (*Réseaux IP Européens*)
    - Europa, Oriente Médio, Ásia Central e África do Norte

Alocação de Endereços IP

LACNIC é a instituição responsável para a América Latina e o Caribe

Alocação de Endereços IP

IANA = *Internet Assigned Numbers Authority*



No Brasil, estas funções foram delegadas ao NIC.br pelo Comitê Gestor da Internet BR (CGI.br)

Regras disponíveis em:
http://registro.br/provedor/numeracao/regras.html

Internet Control Message Protocol (ICMP)

Instituto de computação

- Objetivo
  - Diagnóstico de condições de erro da rede
    - Simplicidade do IP dificulta diagnóstico de falhas
- Executado em cima do IP
    - *Protocol type* = 1
- Todo sistema que roda IP deve rodar o ICMP
- Não provê confiabilidade
  - Apenas informação sobre problemas na rede

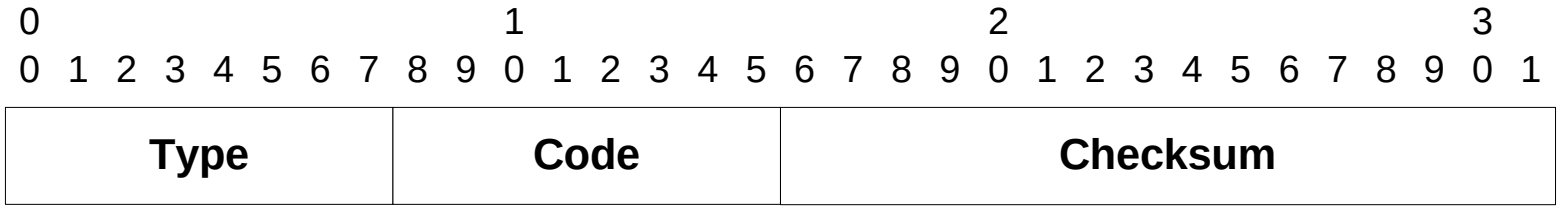
Internet Control Message Protocol (ICMP)

Instituto de computação

- Erros de transmissão de pacotes IP geram mensagens ICMP
  - Exceto erros nas próprias mensagens ICMP
  - Se as mensagens ICMP também gerassem mensagens de erro
    - Poderia haver recursividade e avalanche de mensagens de controle
      - Ex.: Problemas ligados a congestionamentos na rede

Mensagens ICMP

- Cabeçalho
  - Toda mensagem ICMP possui uma parte do cabeçalho em comum

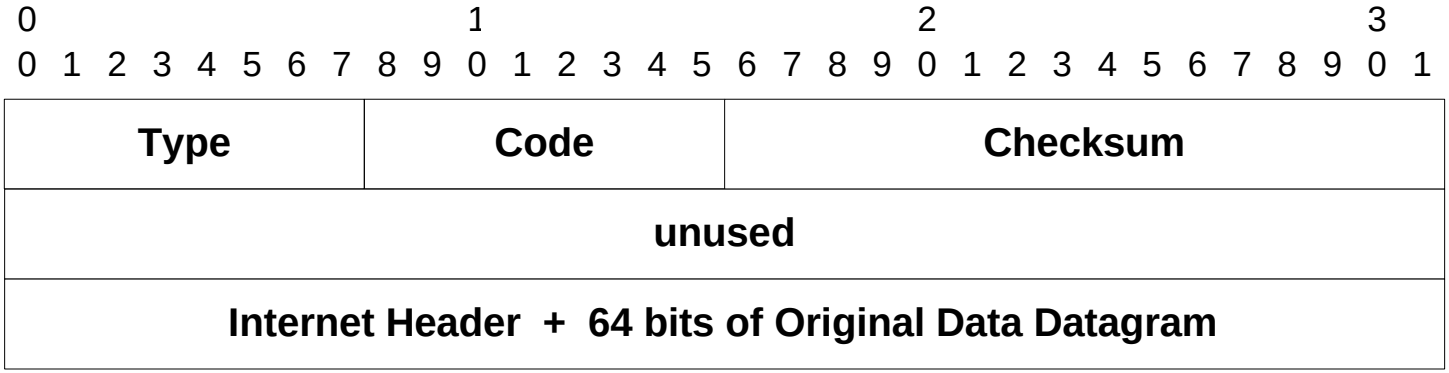


Tipo	Significado
0	Echo Reply
3	Destination Unreachable
4	Source Quench
5	Redirect
8	Echo
9	Router Advertisement
10	Router Solicitation
11	Time Exceeded
12	Parameter Problem
13	Timestamp
14	Timestamp Reply
15	Information Request
16	Information Reply

O *checksum* do cabeçalho é calculado como para o IP

Diagnóstico com o ICMP

- Problemas operacionais → Mais comuns
  - *Destination Unreachable*
  - *Time Exceeded*
  - *Source Quench*



- Formato comum

  Cabeçalho básico do ICMP **+** 32 bits de enchimento **+**

  Primeiros bytes do pacote que causou o envio do ICMP

Diagnóstico com ICMP

Instituto de computação

- ICMP envia
  - Cabeçalho IP e 8 primeiros bytes dos dados da aplicação
    - Esses dados representam informação suficiente para o nó de origem do pacote IP entender o motivo do erro

Diagnóstico com o ICMP

Instituto de computação

- *Destination Unreachable*
  - Roteador não consegue encaminhar um pacote
  - Código:
    - 0 = *net unreachable*
    - 1 = *host unreachable*
    - 2 = *protocol unreachable*
    - 4 = *fragmentaion needed but DF set*
    - 5 = *source route failed*

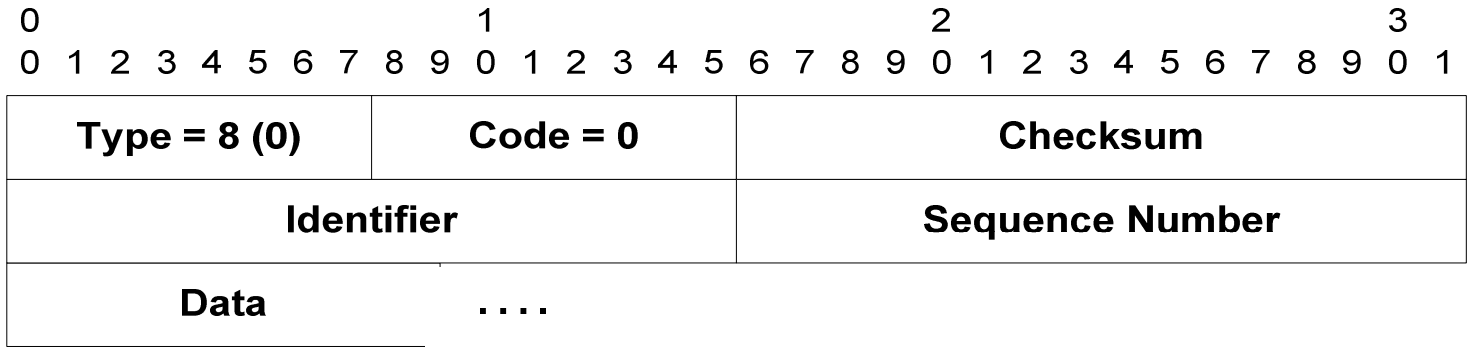
Diagnóstico com ICMP

Instituto de computação

- *Time Exceeded*
  - TTL expirado
  - Código
    - 0 = em trânsito
    - 1 = durante remontagem
- *Source Quench*
  - Enviado pelo roteador para sinalizar congestionamento
  - Não utiliza código (code = 0)

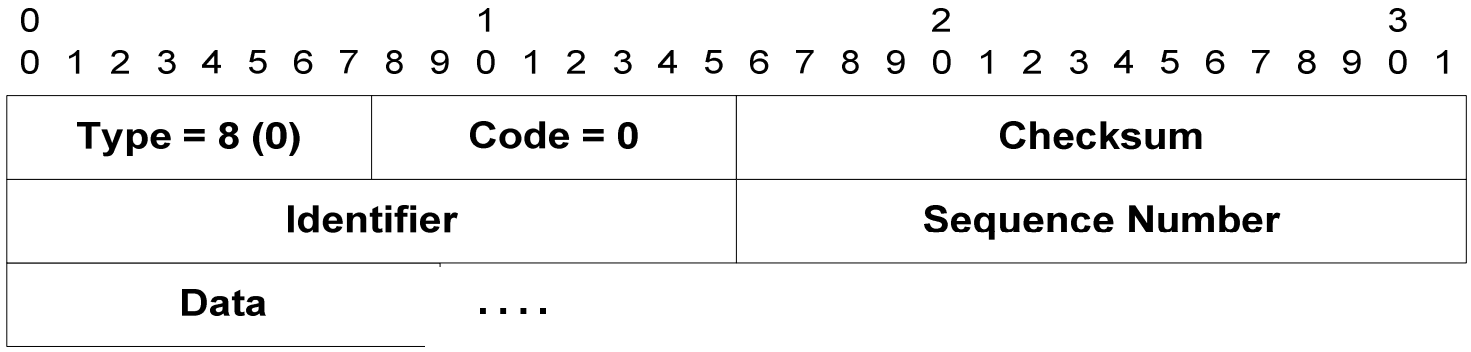
Ping

- Testa se uma estação está "viva"
  - Ou se a conectividade da rede está funcionando
- Utiliza a função **echo** do ICMP
  - Tipo:
    - 8 = Echo
    - 0 = Echo Reply



Ping

- Resposta (*Echo Reply*)
  - Endereços fonte e destino são trocados
  - Troca do valor do tipo da mensagem
  - *Checksums* IP e ICMP recalculados
  - Dados inalterados



Ping

Instituto de computação

- Campos identificação e número de seqüência possibilitam estatísticas

- Outras mensagens ICMP com funcionalidade semelhante
  - Type = 15 – Information Request
  - Type = 16 – Information Reply

Exemplo de Ping

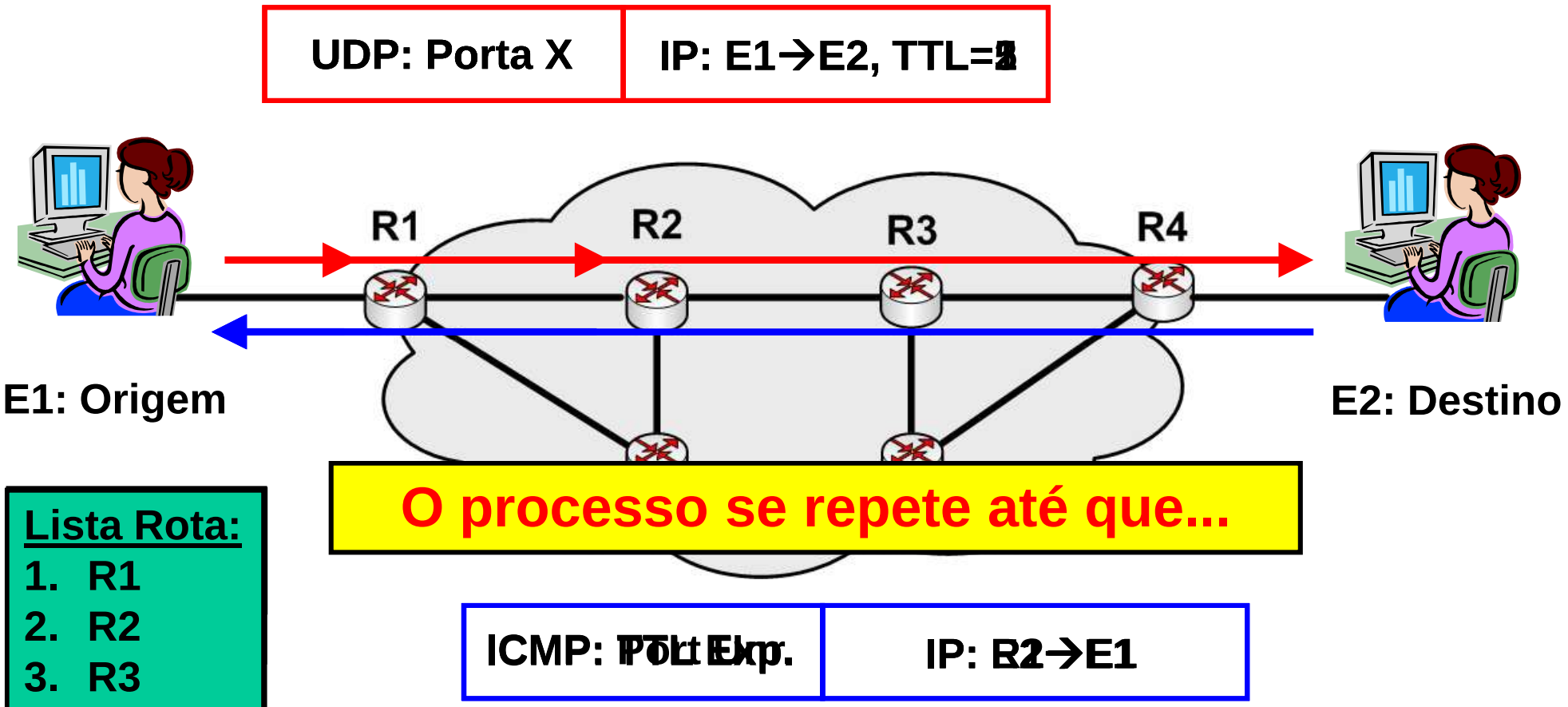
```
PING angra (146.164.69.1) from 146.164.69.2 : 56(84) bytes of data.
recreio::user [ 31 ] ping angra
64 bytes from angra (146.164.69.1): icmp_seq=1 ttl=64 time=0.471 ms
64 bytes from angra (146.164.69.1): icmp_seq=2 ttl=64 time=0.404 ms
64 bytes from angra (146.164.69.1): icmp_seq=3 ttl=64 time=0.544 ms
64 bytes from angra (146.164.69.1): icmp_seq=4 ttl=64 time=0.388 ms
64 bytes from angra (146.164.69.1): icmp_seq=5 ttl=64 time=0.398 ms
64 bytes from angra (146.164.69.1): icmp_seq=6 ttl=64 time=0.398 ms
64 bytes from angra (146.164.69.1): icmp_seq=7 ttl=64 time=0.495 ms
64 bytes from angra (146.164.69.1): icmp_seq=8 ttl=64 time=0.436 ms
64 bytes from angra (146.164.69.1): icmp_seq=9 ttl=64 time=0.413 ms
64 bytes from angra (146.164.69.1): icmp_seq=10 ttl=64 time=0.407 ms
64 bytes from angra (146.164.69.1): icmp_seq=11 ttl=64 time=0.393 ms
64 bytes from angra (146.164.69.1): icmp_seq=12 ttl=64 time=0.391 ms

--- angra ping statistics ---
12 packets transmitted, 12 received, 0% loss, time 11109ms
rtt min/avg/max/mdev = 0.388/0.428/0.544/0.049 ms
```

Traceroute

Instituto de computação

- Identifica os roteadores entre uma fonte e um destino
- Funcionamento
  - Envio sucessivo de pacotes para o destino, variando o TTL
    - UDP em uma porta não utilizada
    - TTL inicial igual a 1
  - Primeiro roteador decrementa o TTL, descarta o pacote, e envia uma mensagem ICMP TTL Exceeded
    - Roteador identificado através do Source Address da mensagem
  - A fonte continua o processo incrementando o TTL de 1 até chegar ao destino ou alcançar um enlace com problema
  - O destino é identificado, pois ele envia uma mensagem ICMP Port unreachable

Traceroute



Lista Rota:

1. R1
2. R2
3. R3
4. R4
5. E2

FIM!

Exemplo - Traceroute

```
recreio::user [ 38 ] traceroute sphinx.lip6.fr
traceroute to sphinx.lip6.fr (132.227.74.253), 30 hops max, 38 byte packets
 1  angra (146.164.69.1)  0.596 ms  0.349 ms  0.341 ms
 2  rt-ct-bloco-H.ufrj.br (146.164.5.193)  175.723 ms  203.553 ms  30.226 ms
 3  rt-nce2.ufrj.br (146.164.1.5)  51.432 ms  3.994 ms  4.137 ms
 4  rederio2-atm-cbpf.rederio.br (200.20.94.58)  3.495 ms  4.421 ms  4.664 ms
 5  200.143.254.66 (200.143.254.66)  4.184 ms  12.224 ms 200.143.254.78
    (200.143.254.78)  13.372 ms
 6  rj7507-fast6_1.bb3.rnp.br (200.143.254.93)  4.473 ms  4.135 ms  4.550 ms
 7  ds3-rnp.ampath.net (198.32.252.237)  110.658 ms  106.239 ms  107.241 ms
 8  abilene.ampath.net (198.32.252.254)  125.393 ms  135.971 ms  127.111 ms
 9  washng-atla.abilene.ucaid.edu (198.32.8.66)  143.388 ms  154.348 ms  144.619 ms
10  abilene.de2.de.geant.net (62.40.103.253)  234.914 ms  235.300 ms  239.316 ms
11  de2-1.de1.de.geant.net (62.40.96.129)  234.644 ms  238.821 ms  236.147 ms
12  de.fr1.fr.geant.net (62.40.96.50)  231.422 ms  232.743 ms  232.437 ms
13  renater-gw.fr1.fr.geant.net (62.40.103.54)  234.984 ms  234.233 ms  231.723 ms
14  jussieu-a1-1-580.cssi.renater.fr (193.51.179.154)  230.906 ms  231.090 ms
    233.714 ms
15  rap-jussieu.cssi.renater.fr (193.51.182.201)  232.602 ms  232.125 ms  238.066 ms
16  cr-jussieu.rap.prd.fr (195.221.126.77)  235.182 ms  239.903 ms  276.221 ms
17  jussieu-rap.rap.prd.fr (195.221.127.182)  234.955 ms  237.264 ms  234.210 ms
18  r-scott.reseau.jussieu.fr (134.157.254.10)  233.992 ms  238.306 ms  239.047 ms
19  olympe-gw.lip6.fr (132.227.109.1)  236.396 ms !N  235.261 ms !N  234.322 ms !N
```

Exemplo – Ping -R

```
recreio::user [ 35 ] ping -R sphinx.lip6.fr
PING sphinx.lip6.fr (132.227.74.253) from 146.164.69.2 : 56(124) bytes of data.
64 bytes from sphinx.lip6.fr (132.227.74.253): icmp_seq=1 ttl=237 time=252 ms
RR:     recreio (146.164.69.2)
        gtagw (146.164.5.210)
        rt-ct2.ufrj.br (146.164.1.3)
        ufrj-atm.rederio.br (200.20.94.9)
        200.143.254.65
        rj-fast4_1.bb3.rnp.br (200.143.254.94)
        rnp.ampath.net (198.32.252.238)
        abilene-oc3.ampath.net (198.32.252.253)
        atla-washng.abilene.ucaid.edu (198.32.8.65)
64 bytes from sphinx.lip6.fr (132.227.74.253): icmp_seq=2 ttl=237 time=289 ms
RR:     recreio (146.164.69.2)
        ...
64 bytes from sphinx.lip6.fr (132.227.74.253): icmp_seq=3 ttl=237 time=247 ms
RR:     recreio (146.164.69.2)
        ...
--- sphinx.lip6.fr ping statistics ---
3 packets transmitted, 3 received, 0% loss, time 2021ms
rtt min/avg/max/mdev = 247.821/263.167/289.150/18.477 ms
```

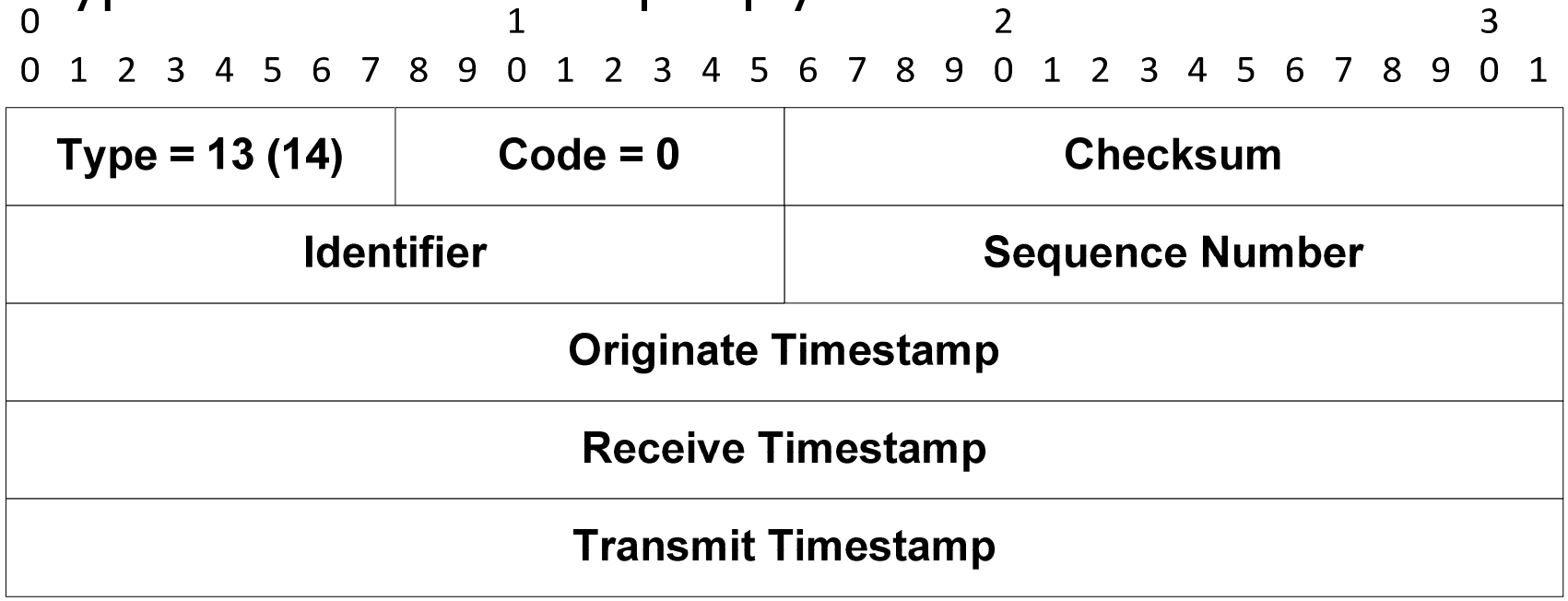
Exemplo – Ping -R

```
recreio::user [ 35 ] ping -R sphinx.lip6.fr
PING sphinx.lip6.fr (132.227.74.253) from 146.164.69.2 : 56(124) bytes of data.
64 bytes from sphinx.lip6.fr (132.227.74.253): icmp_seq=1 ttl=237 time=252 ms
RR:     recreio (146.164.69.2)
        gtagw (146.164.5.210)
        rt-ct2.ufrj.br (146.164.1.3)
        ufrj-atm.rederio.br (200.20.94.9)
        200.143.254.65
        rj-fast4_1.bb3.rnp.br (
        rnp.ampath.net (1
        abilene-
        atl                                  (198.32.8.65)
64 bytes                    (132.227.74.253): icmp_seq=2 ttl=237 time=289 ms
RR:     re                  164.69.2)
        ...
64 bytes from sphinx.lip6.fr (132.227.74.253): icmp_seq=3 ttl=237 time=247 ms
RR:     recreio (146.164.69.2)
        ...
--- sphinx.lip6.fr ping statistics ---
3 packets transmitted, 3 received, 0% loss, time 2021ms
rtt min/avg/max/mdev = 247.821/263.167/289.150/18.477 ms
```

Usa a opção Record Route do IP!
Limitado ao tamanho do campo "opções"

Gerenciamento de Tempo

- Mensagens
  - Type = 13 – Timestamp
  - Type = 14 – Timestamp reply



0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15	16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Type = 13 (14)	Code = 0	Checksum
Identifier	Sequence Number
Originate Timestamp	
Receive Timestamp	
Transmit Timestamp	

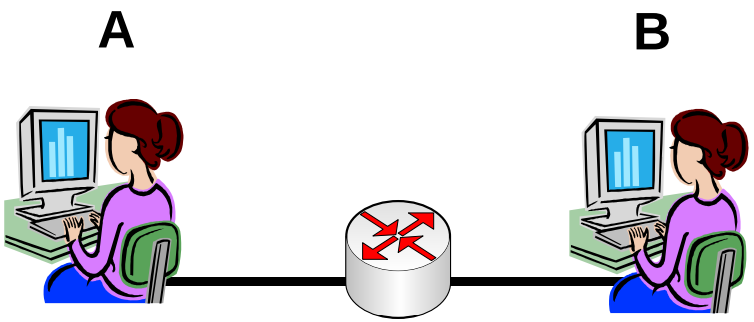
Tempos expressos em ms desde 0:00 h GMT

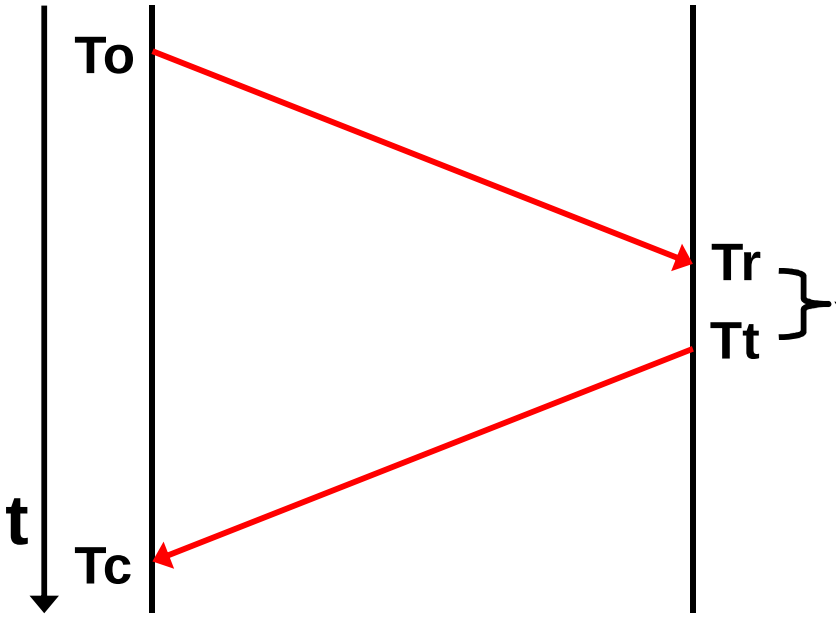
Cálculo da Defasagem entre Duas Estações

Instituto de computação

- Funcionamento
  - Estação A preenche o tempo de origem (To) pouco antes de enviar a mensagem
  - Na recepção, a estação B preenche o tempo de recepção (Tr)
    - Assim que a mensagem chega
    - Em seguida, a estação B prepara a resposta
  - Antes do envio da resposta, B preenche o tempo de transmissão (Tt)
  - Ao receber a resposta, A armazena o tempo de chegada (Tc)
    - Assim que a mensagem chega

Cálculo da Defasagem entre Duas Estações





Defasagem = Diferença medida de relógios – tempo de transmissão

RTT = *Round Trip Time*

Tempo de transmissão = RTT/2

RTT = Tc – To – (Tt – Tr)

Defasagem = Tr – (To + RTT/2)

Se Defasagem > 0, A está atrasada,
Caso contrário, A está adiantada

Tempo de processamento da mensagem

Envio de Pacotes IP

Instituto de computação

- Roteadores
  - Executam um protocolo de roteamento
- Estações
  - Não, necessariamente, executam um protocolo de roteamento

- Porque...
  - Complexidade e variedade dos protocolos de roteamento modernos
  - Poderia-se apenas "ouvir" as mensagens de roteamento
    - Algumas vezes este processo pode não ser fácil
      - Ex. mecanismos de segurança (autenticação, criptografia)

Envio de Pacotes IP

- Roteadores
  - Executam um protocolo de roteamento
- Estações
  - Não, necessari

O que é necessário para uma estação enviar um pacote?

protocolos de roteamento

uderia-se apenas "ouvir" as mensagens de roteamento

- Algumas vezes este processo pode não ser fácil
  - Ex. mecanismos de segurança (autenticação, criptografia)

Envio de Pacotes IP

- Roteadores
  - Executam um protocolo de roteamento
- Estações
  - Não, necessari

O que é necessário para uma estação enviar um pacote?

descobrir um roteador de saída

protocolos de roteamento

deria-se apenas "ouvir" as mensagens de roteamento

- Algumas vezes este processo pode não ser fácil
  - Ex. mecanismos de segurança (autenticação, criptografia)

Descoberta do Próximo Salto

Instituto de computação

- Dado um pacote IP a transmitir, a quem enviar?
  - Estação destino na rede
    - Envio direto
  - Estação destino distante
    - Envio a um roteador que encaminhará o pacote
- Para descobrir se a estação de destino está na sub-rede
  - Testa-se a mascara de rede do endereço IP do destino

Descoberta do Próximo Salto

Instituto de computação

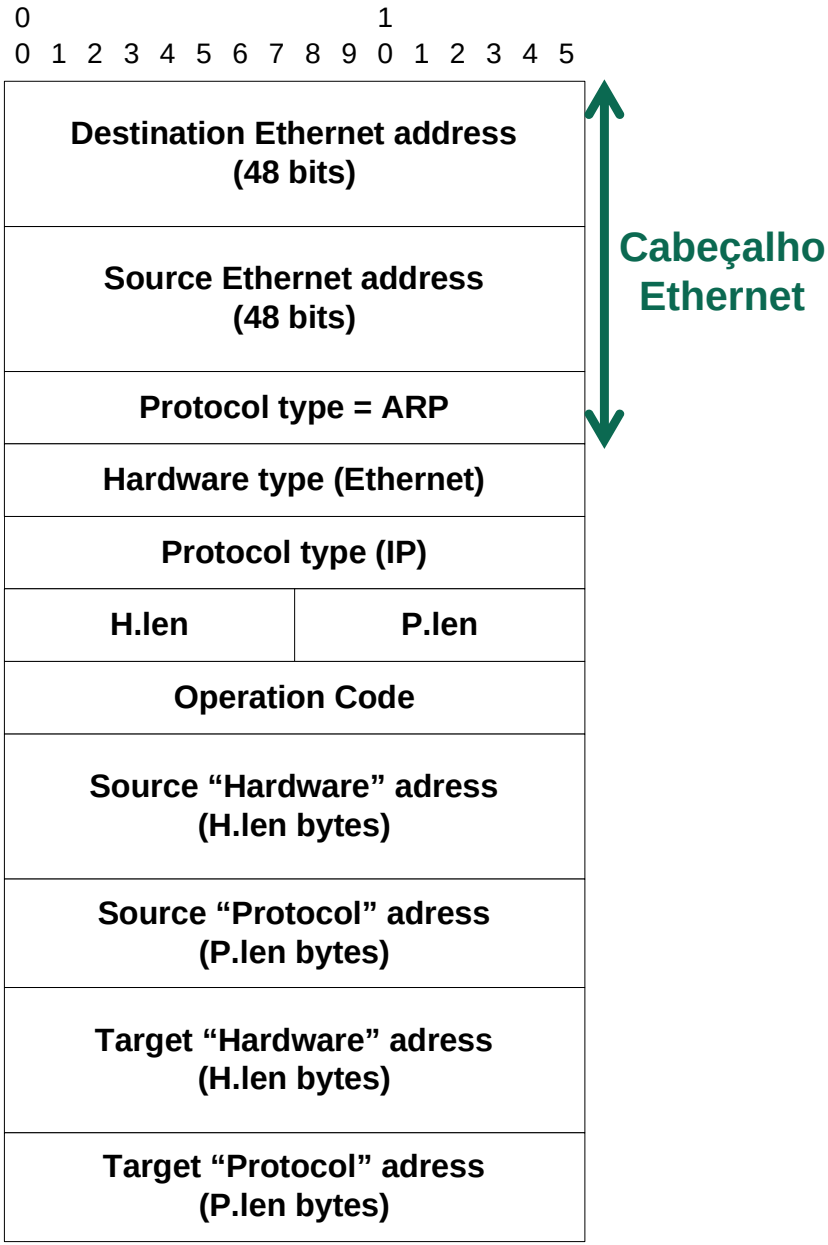
- Dado um pacote IP a transmitir, a quem enviar?
  - Estação destino na rede
    - Envio direto
  - Estação destino distante
    - Envio a um roteador que encaminhará o pacote

- Para descobrir se a estação de destino está na sub-rede
  - Testa-se a mascara de rede do endereço IP do destino

Independente se está na sub-rede, o próximo passo é descobrir o endereço físico (MAC) do próximo salto

Address Resolution Protocol (ARP)

- Envio de ARP *request*
  - Realizado em *broadcast*
    - op. code 1
- Máquina que reconhece seu IP no ARP *request*
  - Envia um ARP *response*
    - op. code 2
- Uso de *cache*
  - Utilizada para evitar o envio frequente de requisições ARPs



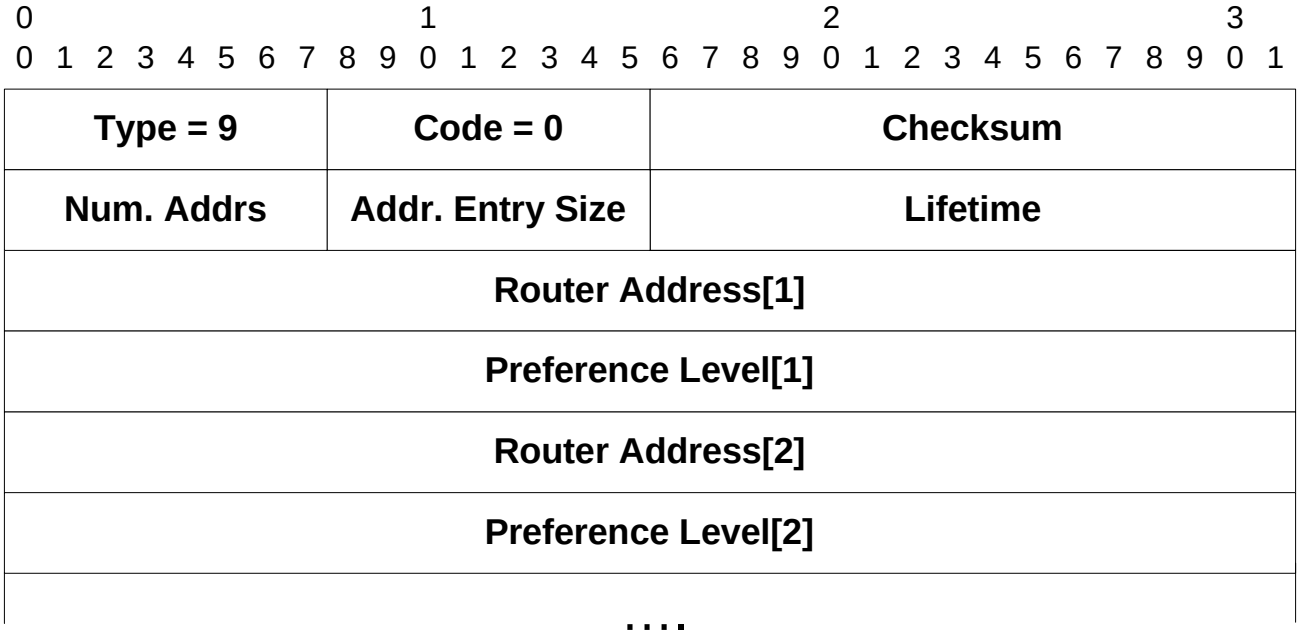
Descoberta do Roteador

Instituto de computação

- Por configuração ou
- Usando o ICMP
  - Roteadores enviam mensagens ICMP router advertisement (type = 10) **periodicamente**
  - Estações podem enviar mensagens ICMP router solicitation (type = 9) para requisitar anúncios de rotas
  - O objetivo do procedimento é descobrir **um** roteador de saída, não necessariamente **o melhor** roteador de saída
    - Mensagens ICMP *redirect* podem ser utilizadas para informar as estações de rotas melhores

Anúncios (*Router Advertisements*)

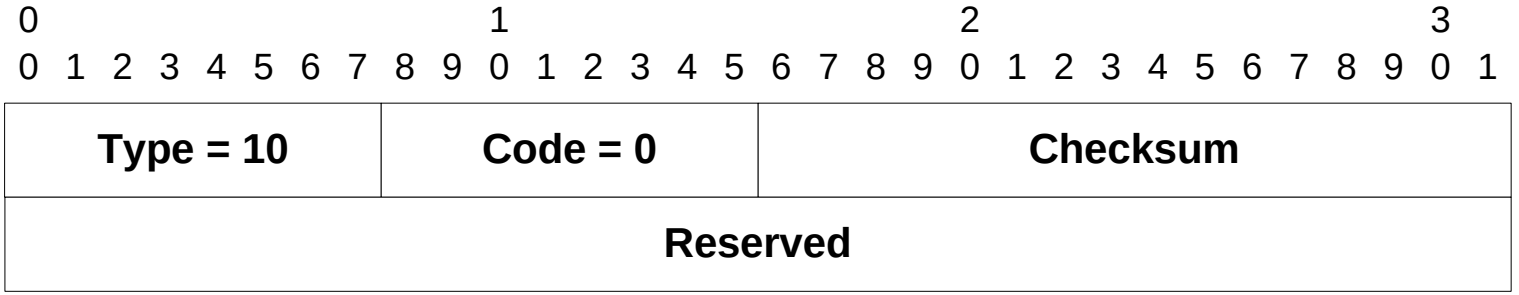
- Podem conter diversos endereços para o mesmo roteador
  - Várias interfaces conectadas à mesma rede
  - Uma interface de rede com dois endereços IP
    - Sub-redes IP na mesma rede física (ex. segmento Ethernet)
  - Preference - prioridade de escolha entre vários roteadores
    - Configurado pelo administrador da rede



Addr. Entry Size = 2 (Router Address + Preference)

Anúncios (*Router Advertisements*)

- São enviados ao endereço **224.0.0.1** (todas as máquinas) ou a **255.255.255.255**
- Informação sobre o roteador de saída
  - Deve ser volátil para evitar uso rotas em desuso
  - Tempo de vida (*Lifetime*)
    - 30 min.
- Anúncios (*router advertisements*) enviados a cada 7 min.
  - Evitar congestionamento da rede
  - Como o período é longo, estações podem enviar solicitações



Escolha do Roteador

Instituto de computação

- *Router solicitation*
  - Enviadas a **`224.0.0.2`** (*"todos os roteadores"*) ou **`255.255.255.255`**
- O roteador envia a resposta
  - À estação, ou
  - A todas as estações, se o momento do anúncio estiver próximo
- Estações podem receber várias respostas
  - Devem considerar apenas os roteadores na sua sub-rede
  - Devem selecionar o de maior valor de preferência
  - Devem enviar todo o tráfego para este roteador

Dynamic Host Configuration Protocol (DHCP)

Instituto de computação

- A premissa até o momento é que cada estação conhece o seu próprio endereço IP
  - Endereço pré-configurado
- Entretanto, isso pode nem sempre ser verdade...
  - Nesses casos, é necessário obter um endereço IP

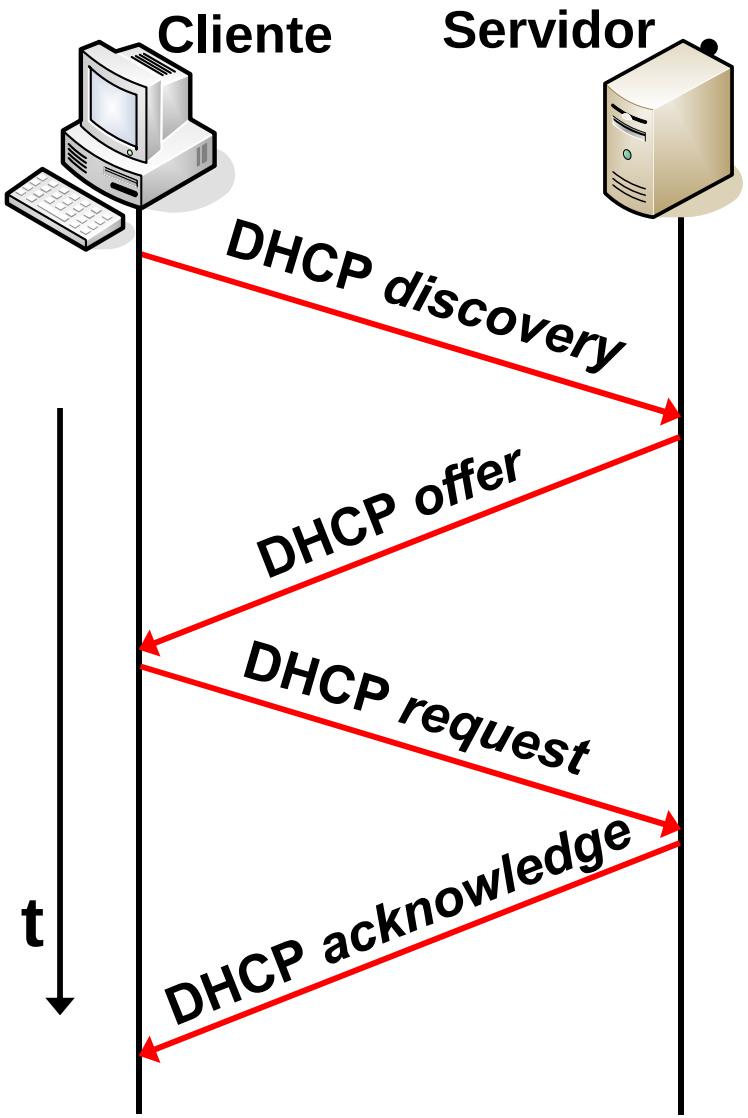
- Alguns protocolos com essa finalidade são
  - RARP: *Reverse Address Resolution Protocol*
  - BOOTP: *Bootstrap Protocol*
  - DHCP
    - Mais utilizado atualmente

Dynamic Host Configuration Protocol (DHCP)

Instituto de computação

- Aloca automaticamente endereços IP para estações em uma sub-rede
  - Os endereços podem ser reusados
- Passa outras informações adicionais
  - Ex. Rota *default*, máscara de sub-rede, servidor DNS
- Utiliza uma arquitetura cliente-servidor
  - Cliente DHCP
    - Estação que solicita parâmetros de configuração de rede
  - Servidor DHCP
    - Estação que responde as solicitações por parâmetros de configuração das estações clientes

Dynamic Host Configuration Protocol (DHCP)



Processo realizado em 4 etapas:

- DHCP *discovery*
  - Cliente envia mensagem em *broadcast* para descobrir os servidores disponíveis
- DHCP *offer*:
  - Servidores DHCP disponíveis respondem com um endereço IP disponível e outras configurações de rede
- DHCP *request*
  - Cliente escolhe uma das ofertas recebidas e solicita individualmente a um servidor as suas configurações
- DHCP *acknowledge*
  - Servidor envia endereço IP e as outras configurações de rede

Network Address Translation (NAT)

Instituto de computação

- Recurso utilizado inicialmente para contornar a possível escassez de endereços IP
  - Usado por mais da metade dos usuários domésticos nos EUA
- Endereço IP público X Endereço IP privado
  - Endereço IP público
    - Definido em escopo global → Internet
      - Endereço roteável
  - Endereço IP privado
    - Definido em escopo local → rede local
      - Endereço não-roteável
        - » Blocos de endereços definidos pelo IANA: Rede 10.0.0.0/8, 192.168.0.0/16 e 172.16.0.0/12

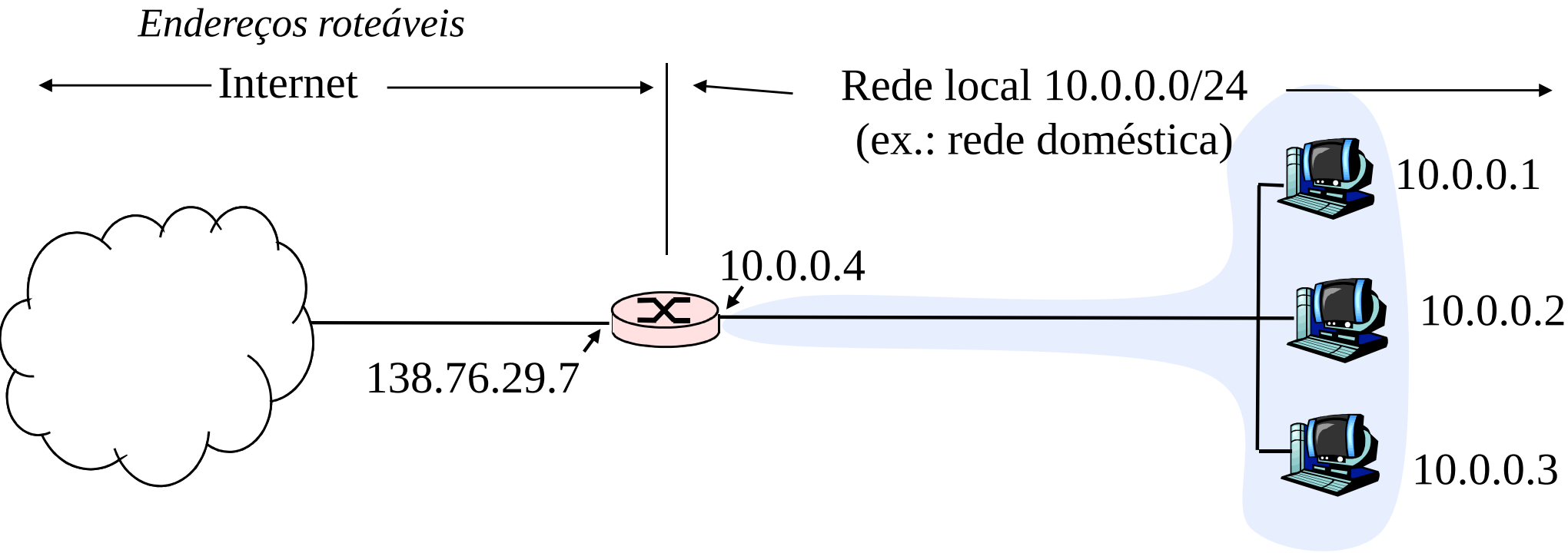
Network Address Translation (NAT)

Instituto de computação

- IP *masquerading*
  - Processo de tradução dos endereços de uma rede local com endereços privados para endereços públicos
    - Consiste em "mascarar" um espaço de endereços privados para Internet
  - Roteador de manter estado dos fluxos que possuem pacotes traduzidos
    - Necessário para encaminhar respostas para a origem
  - Roteador responsável pela tradução pode converter
    - Endereço IP da origem para endereço IP próprio
    - Porta de origem para uma porta conhecida

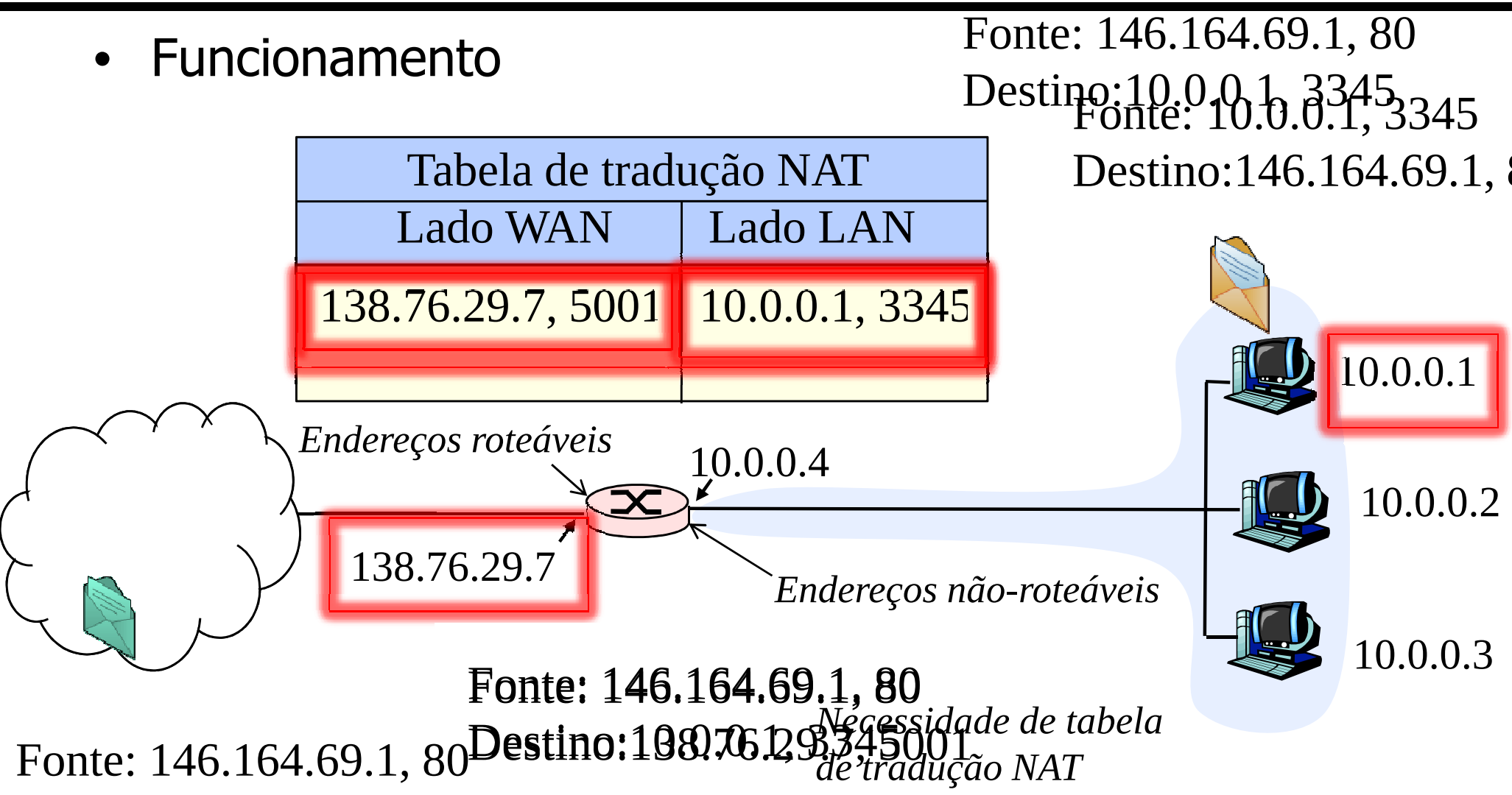
Network Address Translation (NAT)

- Estrutura



Network Address Translation (NAT)

- Funcionamento



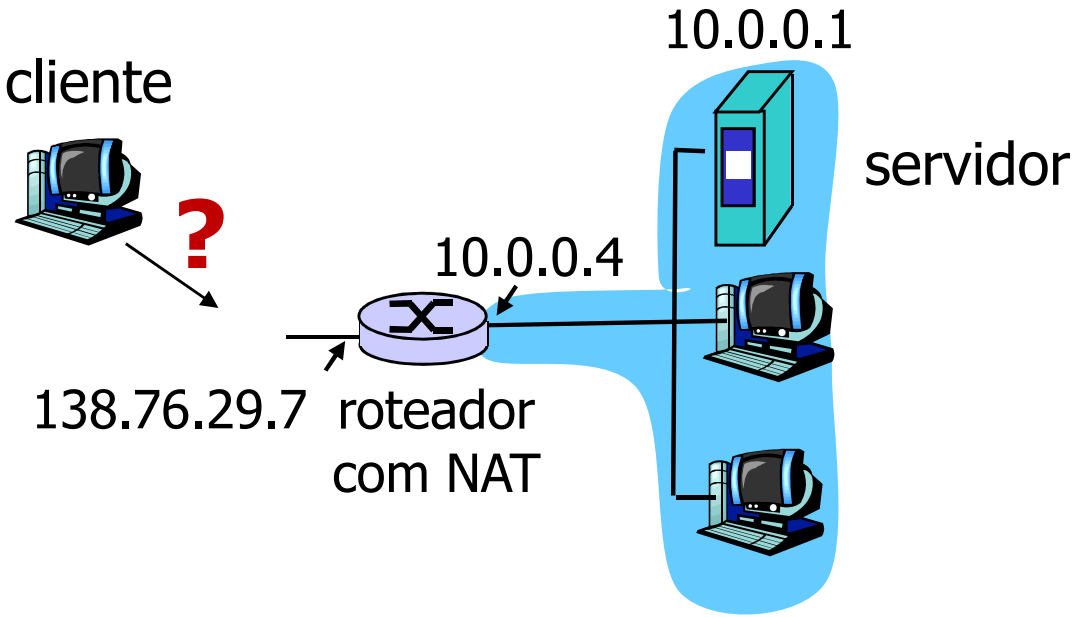
Network Address Translation (NAT)

Instituto de computação

- Quebra do requisito fim-a-fim da Internet
  - Nós na Internet não conseguem se comunicar com nós "atrás" de dispositivos NAT
    - Prejudicam as aplicações par-a-par, por exemplo
- Soluções
  - Mapeamento estático de portas
  - UPnP (*Universal Plug-and-Play*)
    - Padrão que utiliza protocolos para realizar mapeamento automático de portas
  - Uso de nós intermediários
    - Ex.: Skype

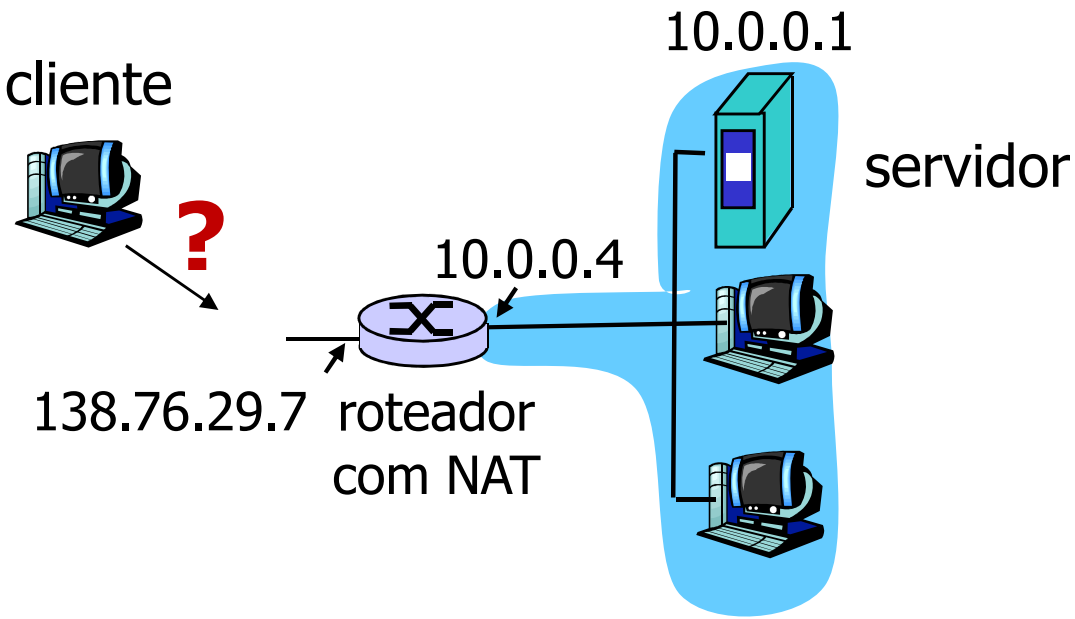
Como atravessar o NAT?

- Problema: cliente quer se conectar com servidor, cujo endereço é 10.0.0.1
  - Endereço do servidor é local à LAN que usa NAT
    - Cliente não pode usá-lo como endereço de destino
  - Único endereço válido externamente é o do roteador: 138.76.29.7



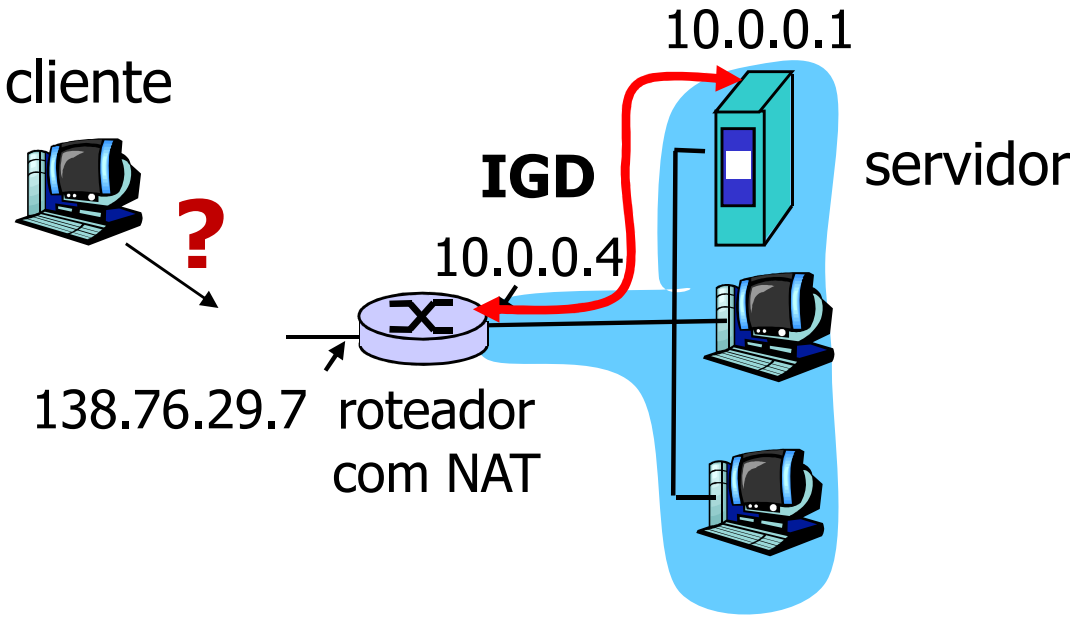
Como atravessar o NAT?

- Solução 1: Configurar o roteador NAT **estaticamente** para encaminhar pedidos de requisição de conexão em uma dada porta para o servidor
  - Ex.: (123.76.29.7, porta 2500) sempre é encaminhado para (10.0.0.1, porta 25000)



Como atravessar o NAT?

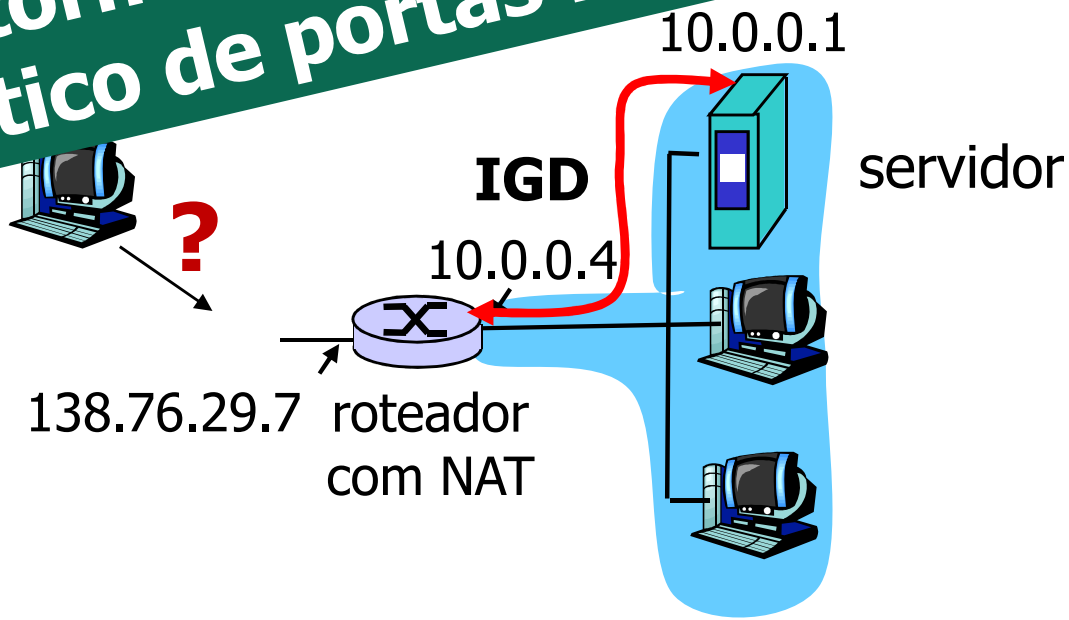
- Solução 2: *Universal Plug and Play* (UPnP) *Internet Gateway Device* (IGD) *Protocol*
  - Permite que estações "atrás" do roteador
    - Aprendam o endereço IP público (138.76.29.7)
    - Adicionem/removam mapeamentos de portas (com tempos de liberação)



Como atravessar o NAT?

- Solução 2: *Universal Plug and Play* (UPnP) *Internet Gateway Device* (IGD) *Protocol*
  - Permite que estações "atrás" do roteador
    - Aprendam o endereço IP público (1
    - Adicionem/removam
      liberação)
      mpos de

UPnP: torna automático o mapeamento estático de portas no roteador NAT



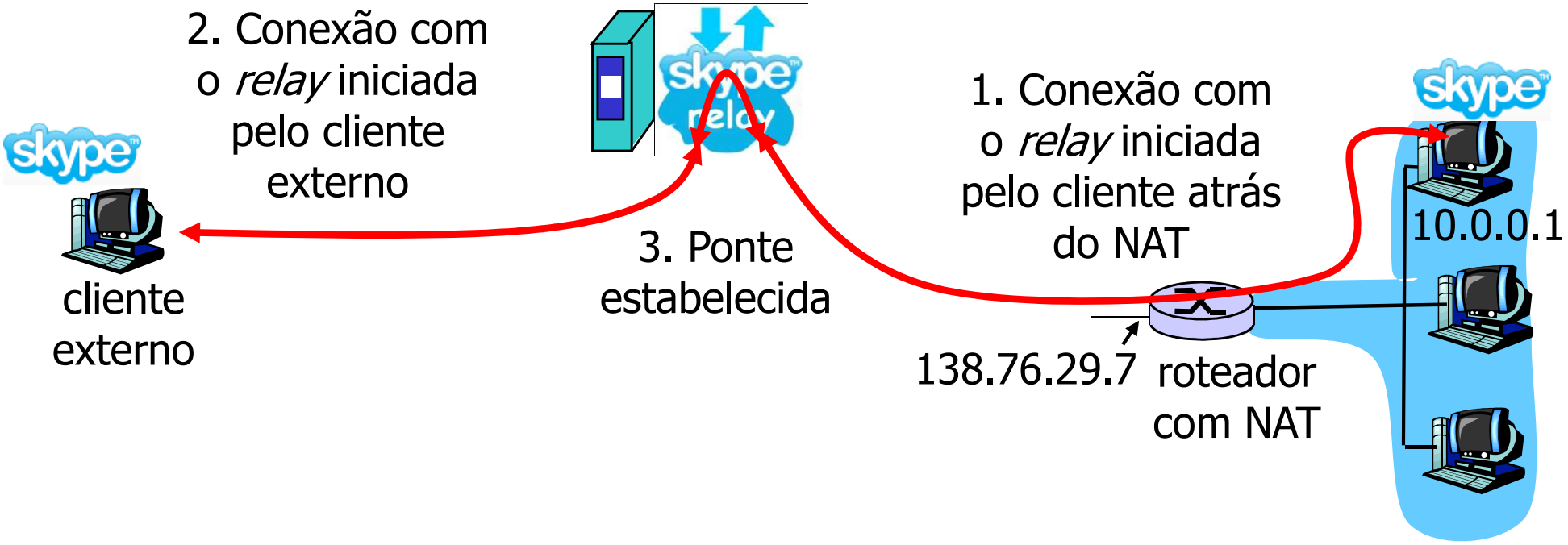
UPnP – *Universal Plug and Play*

Instituto de computação

- Funcionamento
  - Aplicação executando em uma estação pode solicitar um mapeamento NAT entre (*IP privado, #porta privado*) e (*IP público, #porta público*) para algum número de porta público
  - Se o NAT aceitar o pedido, estações remotas podem iniciar conexões TCP para a porta pública e as aplicações podem anunciar seu par IP e porta públicos para a Internet

Como atravessar o NAT?

- Solução 3: Uso de nós intermediários (Ex. Skype)
  - Cliente atrás do NAT estabelece uma conexão com o nó intermediário (*relay*)
  - Cliente externo se conecta ao nó intermediário
  - Nó intermediário faz uma ponte entre as duas conexões



Instituto de Computação

IPv6

IPv6

Instituto de computação

- No início dos anos 90
  - Previsões pessimistas
    - Todos os endereços IPv4 (32 bits) vão se esgotar até 2002
    - CIDR apenas contorna o problema temporariamente
- Em julho de 1994
  - É proposto o protocolo "*Simple Internet Protocol Plus*"
    - O principal autor foi Steve Deering
    - O *Simple Internet Protocol Plus* foi utilizado como base do IPv6
  - A diferença principal entre o IPv4 e o IPv6 é a quantidade de bits no valor do endereço
    - De 32 bits passa para 128 bits

IPv6

Instituto de computação

- O IPv6 pode ser considerado uma versão com endereços maiores do IPv4
  - No IPv6 também...
    - Cada endereço identifica uma interface
    - Cada estação pode possuir múltiplas interfaces (*multihomed*)
  - Entretanto...
    - O IPv6 os endereços em três categorias
      - *Unicast*
      - *Multicast*
      - *Anycast*

Lembrete: originalmente, o IPv4 dividia os endereços em cinco diferentes classes: de A até E

Unicast, *Multicast* e *Anycast*

Instituto de computação

- Endereços *Unicast*
  - Identificam apenas uma interface
    - Utilizado para comunicações ponto-a-ponto
- Endereços *Multicast*
  - Identificam um grupo de interfaces
    - Utilizado para comunicações de grupo
- Endereços *Anycast*
  - Identificam um grupo de interfaces
    - Apesar de identificar um grupo de estações, a mensagem é entregue apenas para uma delas sem uma definição *a priori*
      - Utiliza critérios de escolha como proximidade da origem

Notação dos Endereços

Instituto de computação

- Endereços de 128 bits divididos em oito inteiros de 16 bits cada
  - Inteiros são separados por dois pontos (“:”)
  - Utilizam notação hexadecimal
    - Ex.: FEDC:BA98:7654:3210:FEDC:BA98:7654:3210
- O tamanho do endereço dificulta a manipulação
  - Alguns administradores acham uma boa idéia, pois usuários tem que manipular nomes e não endereços
  - Entretanto, eles não são maioria
    - Os endereços IPv6 podem ser escritos com algumas simplificações

Notação dos Endereços

Instituto de computação

- Inteiros com 0's à esquerda podem ser simplificados
  - Suprime-se os 0's
    - Ex.: FEDC:**0008**:7654:**0010**:FEDC:BA98:0000:3210 →
      FEDC:**8**:7654:**10**:FEDC:BA98:0000:3210
- Inteiros compostos somente por 0's podem ser simplificados
  - Suprime-se todos os 0's
    - Ex.: FEDC:8:7654:10:FEDC:BA98:**0000**:3210 →
      FEDC:8:7654:10:FEDC:BA98**::**3210
  - Se houver uma sequência de inteiros zerados
    - Ex.: FEDC:BA98:7654:**0000:0000**:BA98:7654:3210 →
      FEDC:BA98:7654**::**BA98:7654:3210

Notação dos Endereços

Instituto de computação

- E se fosse assim?

  FEDC:BA98:7654:**0000:0000**:BA98:**0000**:3210

  Como ficaria?

  FEDC:BA98:7654**::**BA98:**0**:3210

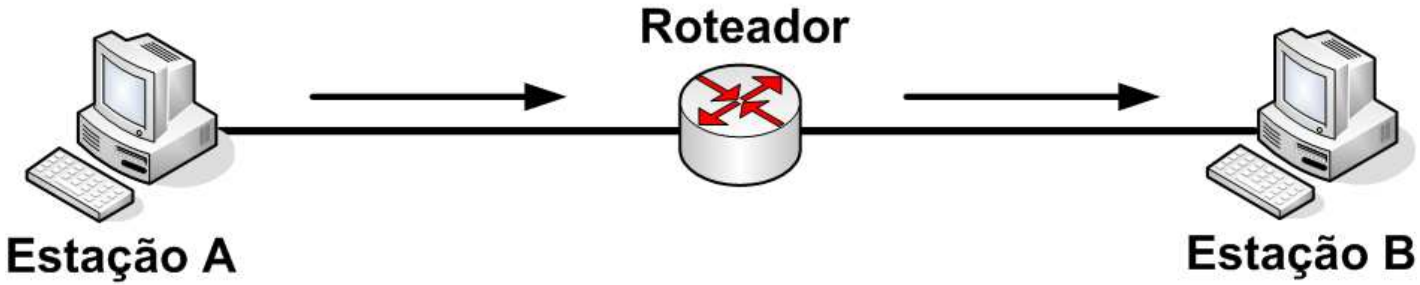
- Endereços IPv6 são obtidos prefixando 96 bits ao endereço IPv4
  - Endereço IPv4: 10.0.0.1
    - FEDC:BA98:7654:3210:FEDC:BA98:10.0.0.1

prefixo

Endereços *Unicast*

- Identificados pelo prefixo: 001
- Segue uma estrutura hierárquica
  - TLA (*Top Level Aggregator*)
    - Porção do endereço para identificar grandes provedores (registradores regionais)
  - Reservado
    - Reserva para possível aumento do número de TLAs

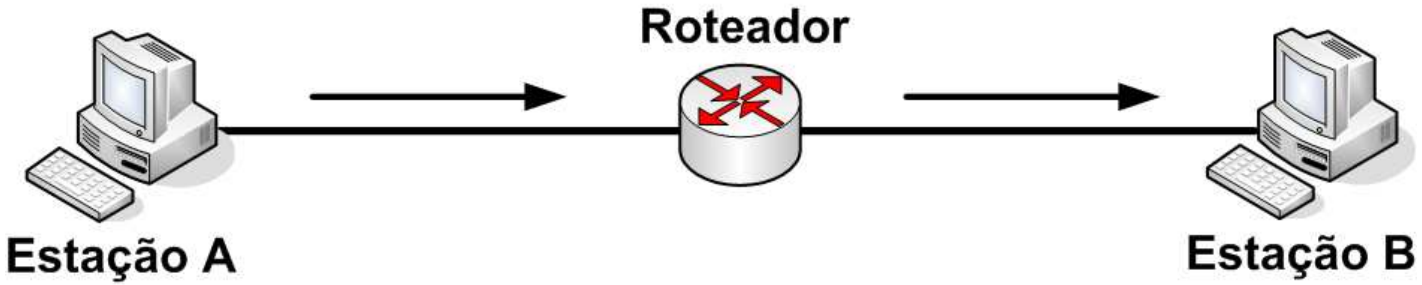
Comprimento (bits)	3	13	8	24	16	64
Parte do endereço	001	TLA	Reserv.	NLA	SLA	ID da interface



Endereços *Unicast*

- Segue uma estrutura hierárquica
  - NLA (*Next Level Aggregator*)
    - Identificador de um provedor menor que o TLA
  - SLA (*Site Local Aggregator*)
    - Identificador de uma organização que pode sub-dividir seu espaço de endereçamento em sub-redes
  - ID da interface
    - Identificador da interface de rede

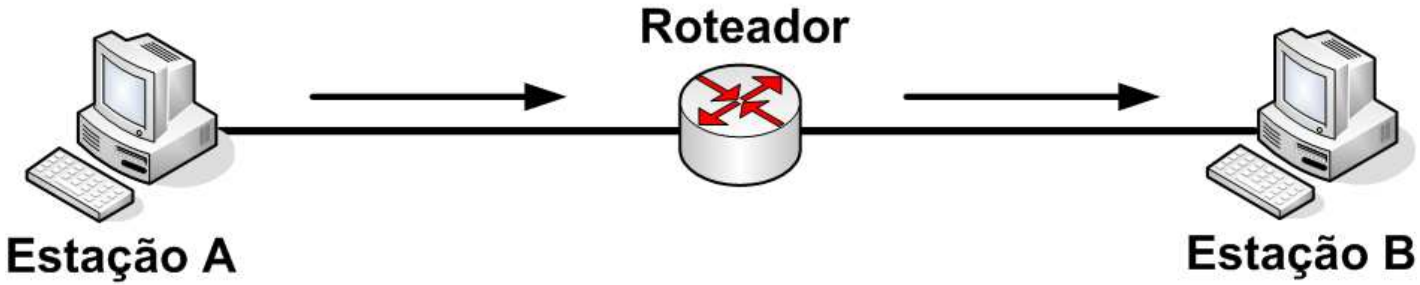
Comprimento (bits)	3	13	8	24	16	64
Parte do endereço	001	TLA	Reserv.	NLA	SLA	ID da interface



Endereços *Unicast*

- TLAs e NLAs
  - São utilizados por provedores que oferecem serviços de trânsito público de dados da Internet
- SLAs
  - São utilizados por organizações que não oferecem trânsito público na Internet

Comprimento (bits)	3	13	8	24	16	64
Parte do endereço	001	TLA	Reserv.	NLA	SLA	ID da interface



Endereços *Unicast*

Instituto de computação

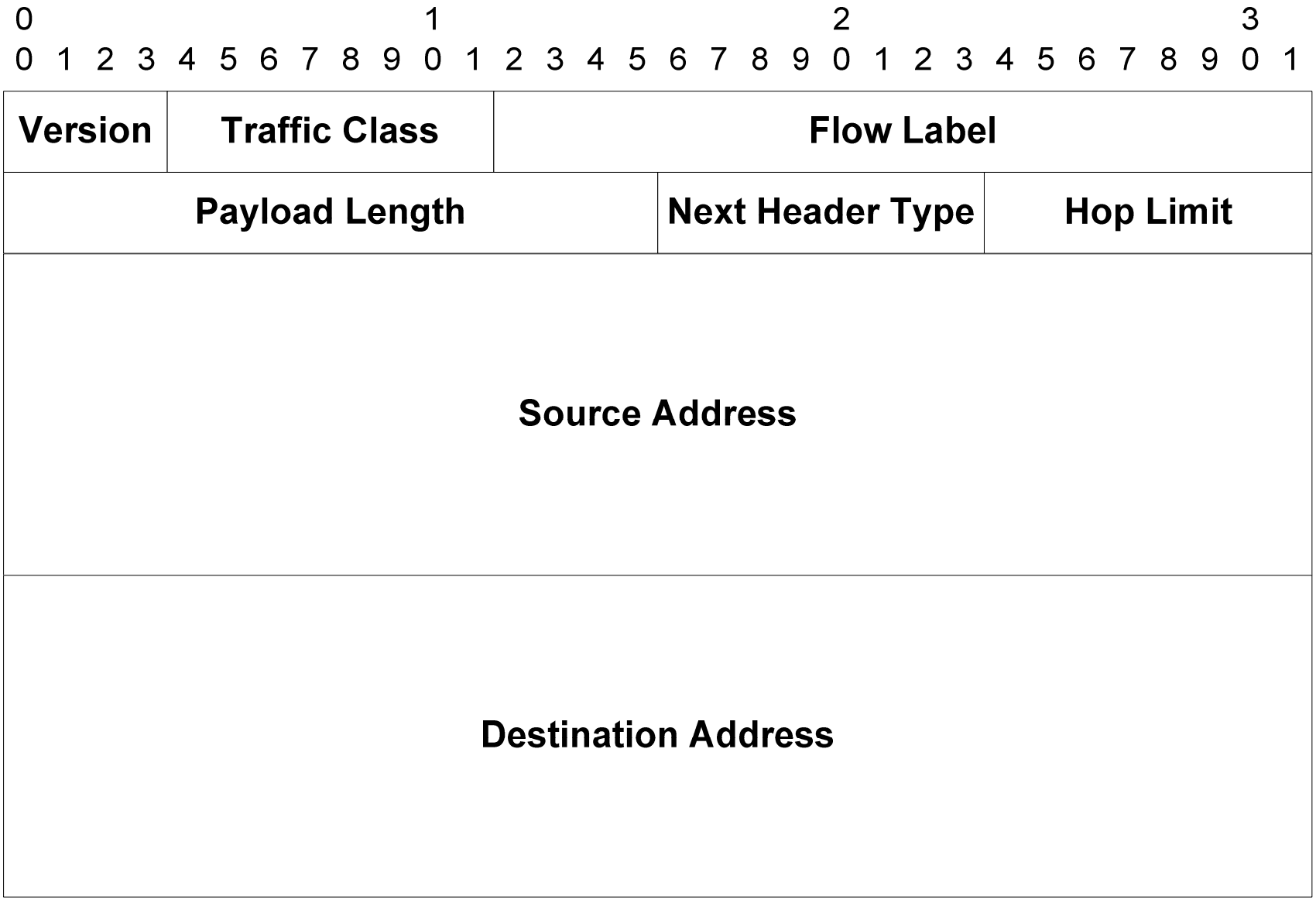
- Formatos Especiais
  - *Unspecified*
    - Endereço utilizado quando a estação não possui endereço configurado
    - Endereço todo em 0's
      - 0:0:0:0:0:0:0:0 ou simplesmente ::
  - *Loopback*
    - Endereço para testes da interface
      - 0:0:0:0:0:0:0:1
  - IPv4-*based*
    - Endereço IPv6 prefixado por um endereço IPv4

Endereços *Unicast*

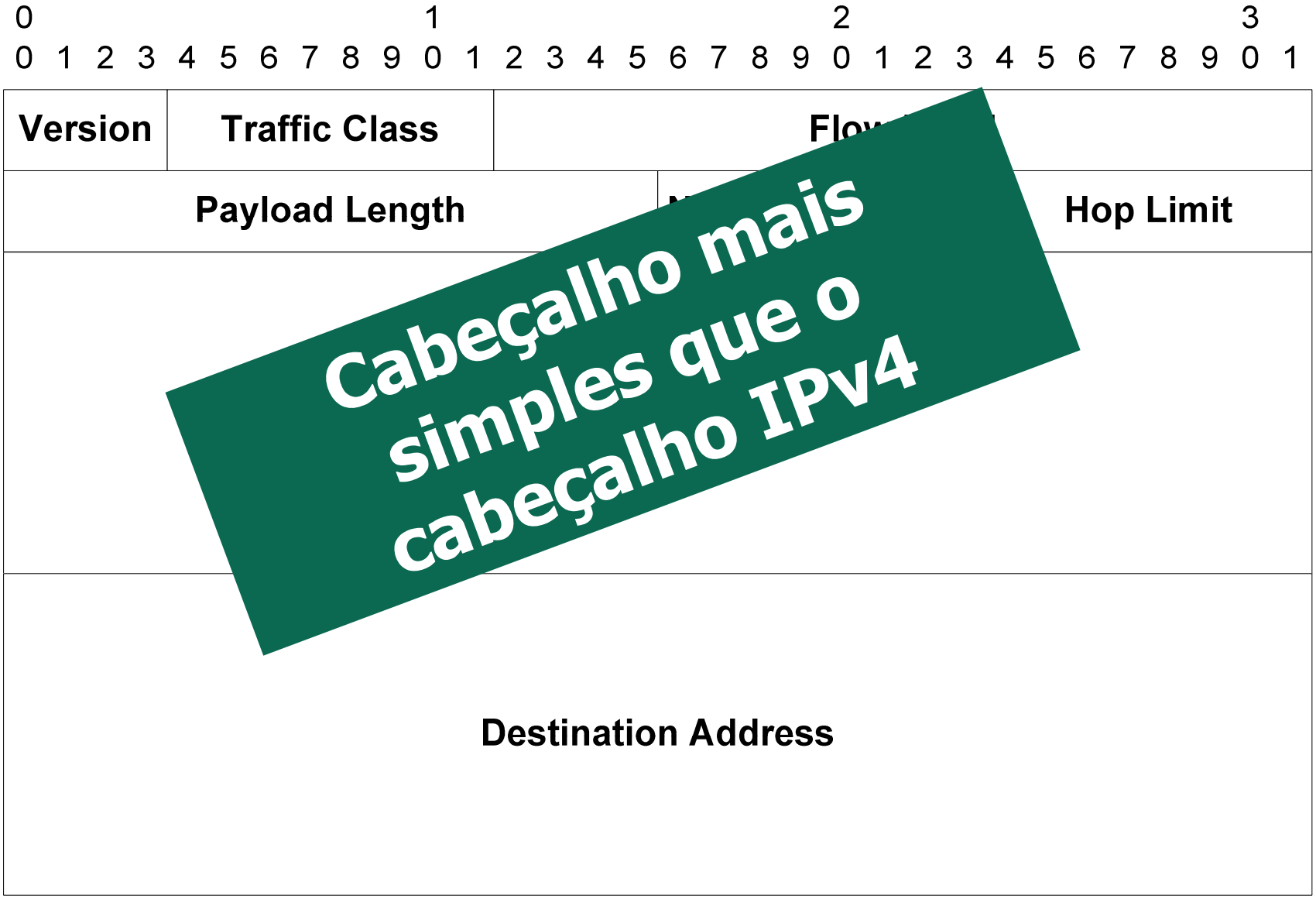
Instituto de computação

- Mais formatos especiais
  - *Site local*
    - Endereços privados
      - FEC0:0:0:<ID da sub-rede (16 bits)>:<ID da estação (64 bits)>
  - *Link local*
    - Endereços que ainda não conhecem o endereço de rede
      - FE80:0:0:<ID da sub-rede (16 bits)>:<ID da estação (64 bits)>

Cabeçalho IPv6



Cabeçalho IPv6



Cabeçalho IPv6

Instituto de computação

- Versão (4 bits)
  - Identifica a versão 6 do IP
    - Mesma função exercida no cabeçalho IPv4
- Classe de tráfico (8 bits)
  - Mesma função exercida no cabeçalho IPv4 pelo campo tipo de serviço
    - Campo utilizado para determinar prioridades de tráfego em função de requisitos da aplicação
- Rótulo do fluxo (20 bits)
  - Identifica pacotes do mesmo fluxo
    - Campo utilizado para acelerar a identificação de requisitos previamente utilizados

Cabeçalho IPv6

Instituto de computação

- Comprimento dos dados (16 bits)
  - No IPv4, o comprimento também considerava o tamanho do cabeçalho
    - No IPv6 não há o campo opções
- Próximo tipo de cabeçalho (8 bits)
  - Identificador do protocolo da camada superior
    - Diferente do IPv4, o IPv6 pode inserir extensões antes do cabeçalho do protocolo da camada superior
- Limite de saltos (8 bits)
  - Número de saltos
    - Semelhante ao TTL do IPv4
      - O IPv6 definiu que o limite de saltos é sempre decrementado de 1, o que já era feito na prática pelo IPv4

Diferenças para o Cabeçalho IPv4

Instituto de computação

- Não há o campo de opções
  - Cabeçalho de tamanho **fixo igual a 40 bytes**
    - Aumenta a velocidade de processamento
    - Caso haja necessidade, o IPv6 define extensões que "encapsulam" o cabeçalho IPv6
    - Não há a necessidade do campo de comprimento do cabeçalho (IHL) do IPv4
- Não há o campo *checksum*
  - Reduz processamento do cabeçalho
  - Entretanto, considera o risco de alterações no cabeçalho e consequente falhas no roteamento
    - Assume que o risco é minimizado já que a detecção de erros é realizada em protocolos de outras camadas

Diferenças para o Cabeçalho IPv4

Instituto de computação

- Não há campos para fragmentação
  - O IPv6 descobre a unidade máxima de transferência (MTU) do caminho
    - Processo chamado *path MTU discovery*
  - Estações que não desejam realizar o processo de descobrimento da MTU
    - Devem enviar pacotes com tamanho máximo reduzido
  - Caso a MTU utilizada for maior que o máximo
    - Roteadores mandam mensagem de erro à fonte quando não puderem enviar um datagrama "grande"
- Formato de endereços expandido
  - Endereços de 128 bits

Extensões do Cabeçalho IPv6

Instituto de computação

- Opções no IPv4
  - Roteadores são otimizados para o caso comum que é a **ausência** do campo opções
    - Sempre que o campo opções está presente, o pacote segue um caminho no roteador de desempenho inferior
- No IPv6
  - As possíveis opções são tratadas como extensões do cabeçalho IPv6
    - Portanto, entre o cabeçalho do IPv6 e o cabeçalho do protocolo de camada superior há a inserção das extensões do IPv6

Extensões do Cabeçalho IPv6

Instituto de computação

- Há seis extensões definidas
  1. Roteamento pela fonte
     - Extensão possui lista de endereços IPv6
  2. Fragmentação
     - A fragmentação é realizada pela fonte e não por elementos intermediários
  3. Autenticação
     - Extensão utilizada para autenticar a fonte e garantir integridade
  4. Criptografia
     - Extensão utilizada para autenticar a fonte, garantir integridade e criptografar os dados

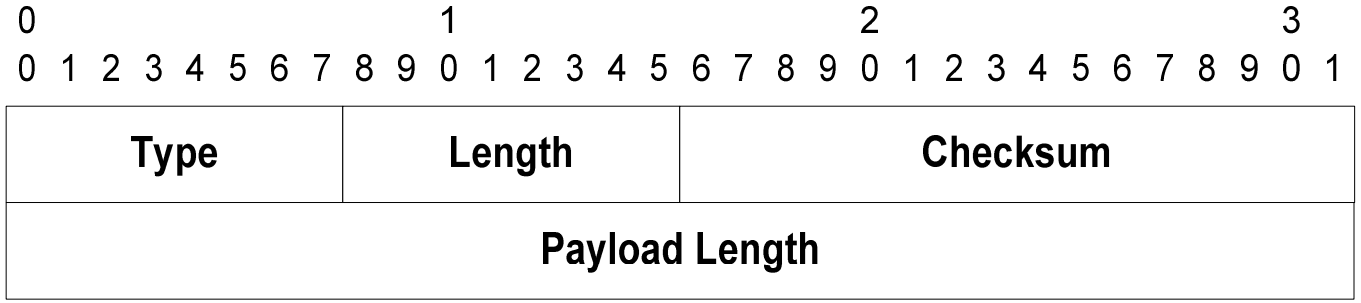
Extensões do Cabeçalho IPv6

Instituto de computação

- Há seis extensões definidas
  5. Opção de destino
     - Extensão examinada apenas pelo destino do pacote
  6. Opção salto a salto
     - Extensão é examinada por todos os nós do caminho até o destino
       - Oposta à opção de destino

ICMPv6

- Modificações com o mesmo objetivo que o do protocolo IP
  - Simplificações de funcionalidades que não eram mais utilizadas
- Incorporou funções de controle de grupos *multicast*
- Inclui novos tipos e códigos de erro
  - Pacote muito grande
  - Opções IPv6 não reconhecidas



1	Dest. unreachable
2	Packet too big
3	Time Exceeded
4	Parameter problem
128	Echo request
129	Echo reply

Descoberta do Próximo Salto

Instituto de computação

- Procedimento semelhante ao ARP no IPv4
  - Se o endereço físico do vizinho não estiver em *cache*
    - Estação envia uma mensagem de solicitação de vizinhança e coloca o *status* do processo como incompleto
  - Se já houver uma entrada para o vizinho, mas o *status* do processo for incompleto
    - Estação deve esperar o término do processo de descobrimento para conhecer o endereço físico do vizinho
  - Se a entrada existir e o processo estiver completo
    - Estação está pronta para enviar mensagem
  - Se a entrada existir, mas não tiver sido utilizada por um longo tempo
    - Estação utiliza o endereço físico conhecido, mas deve enviar uma solicitação de endereço

Descoberta do Próximo Salto

Instituto de computação

- Mensagem de solicitação de vizinhança
  - Mensagem ICMP do tipo 135
  - Contém o endereço IP do vizinho solicitado
    - Limite de saltos é sempre 1
  - Mensagem destinada a um grupo *multicast* composto por nós da mesma sub-rede
- Mensagem de anúncio de vizinhança
  - Mensagem ICMP do tipo 136
  - Contém o endereço físico solicitado

Descoberta do Roteador

Instituto de computação

- Assim como no IPv4
  - O destino pode não estar na mesma sub-rede
  - Portanto, é preciso descobrir qual roteador conhece caminhos para o destino
    - Descoberta do próximo salto

- Diferente do IPv4, no IPv6 os roteadores **sempre** enviam anúncios de rotas
  - Envios realizados periodicamente

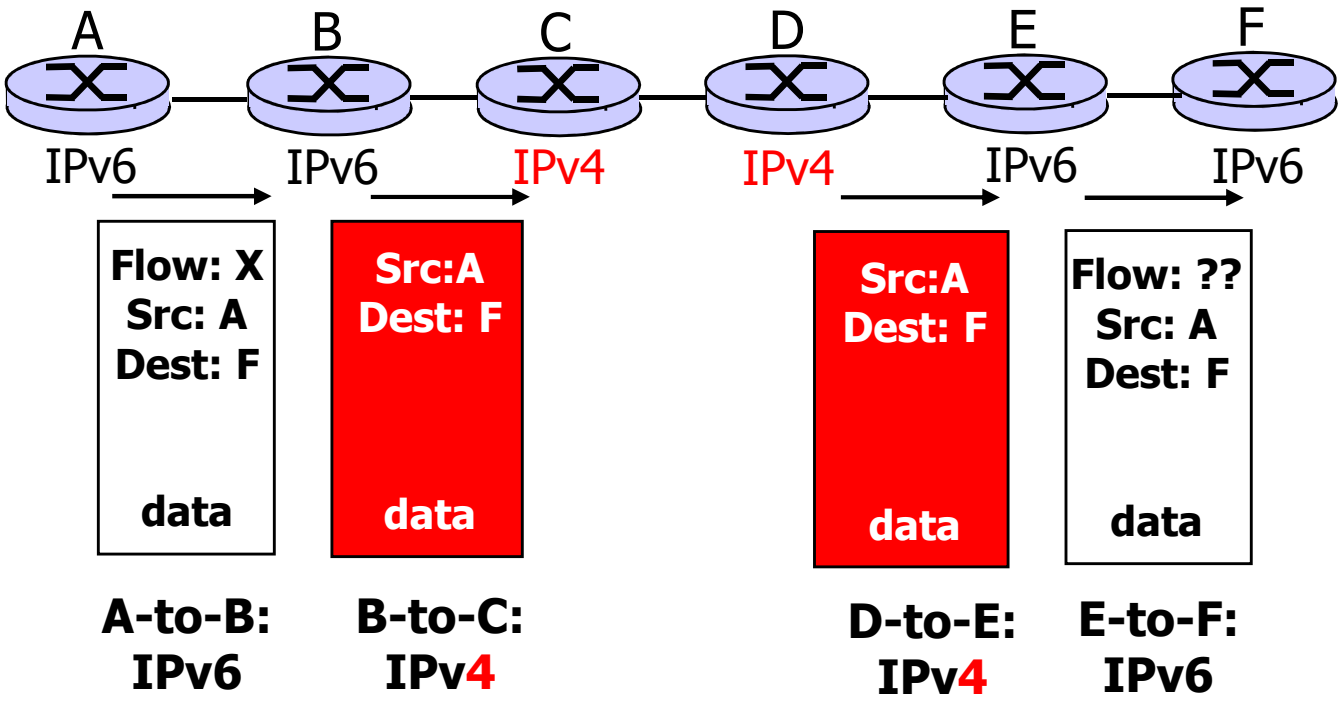
Transição IPv4 para IPv6

Instituto de computação

- Existem mecanismos para a transição do IPv4 para IPv6
  - Importantes durante a fase na qual as duas versões do protocolo operarão concomitantemente
    1. Pilha dupla
    2. Tunelamento IPv4 sobre IPv6

Pilha Dupla

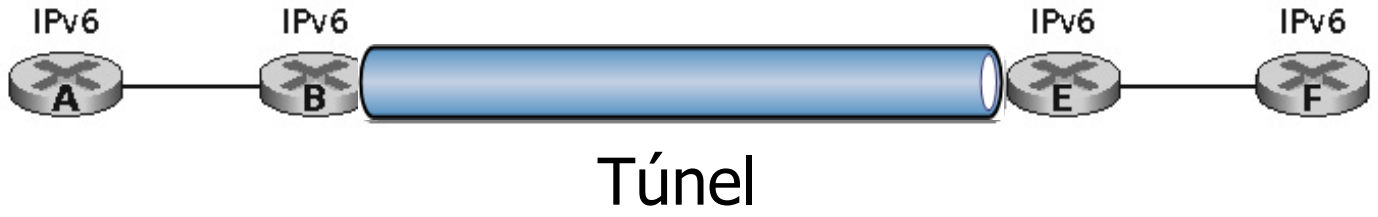
- Alguns roteadores implementam os protocolos IPv4 e IPv6



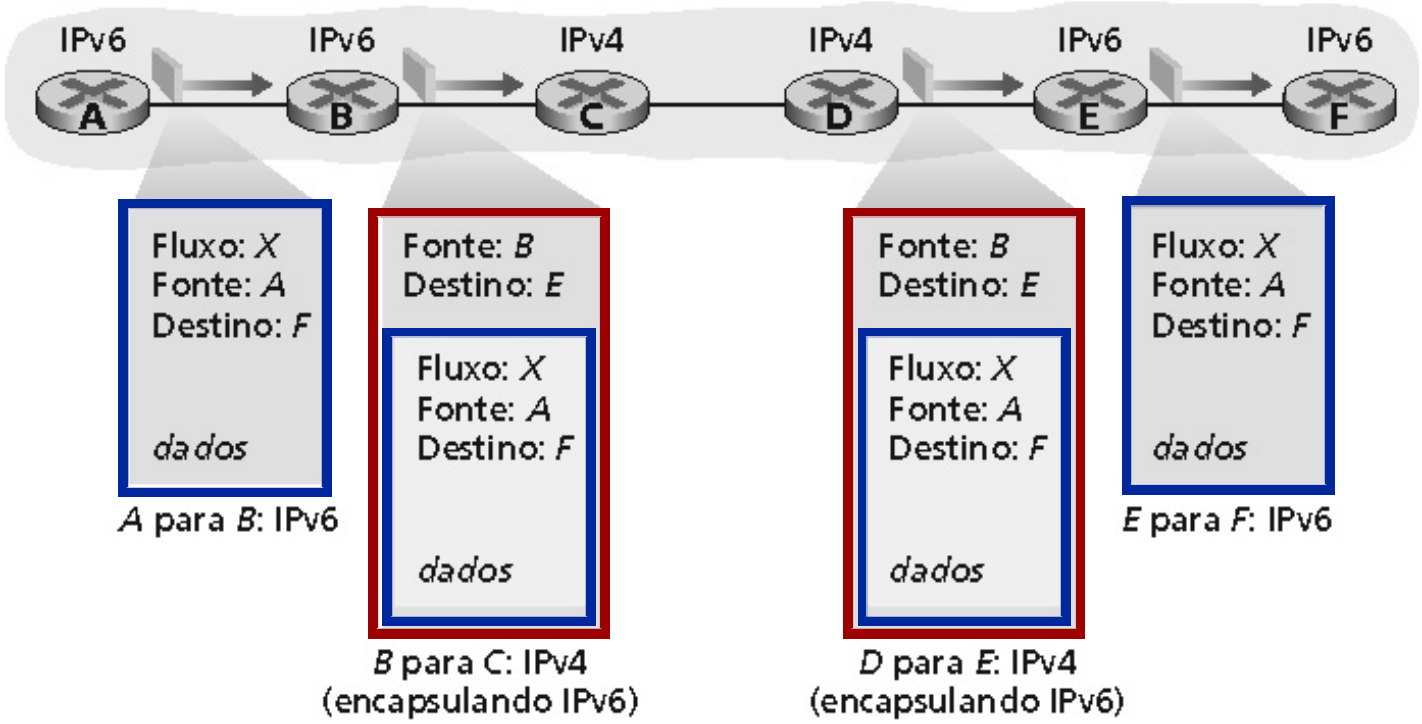
Desvantagem: Perda de informações contidas em campos do IPv6 que não existem no IPv4. Ex.: Rótulo do fluxo.

Tunelamento IPv4 sobre IPv6

Visão lógica



Visão física



Por que o IPv6 ainda não foi implementado na Internet?

Instituto de computação

- Apesar da existência de técnicas para a transição do IPv4 para o IPv6, ela ainda **não** ocorreu

- Inicialmente, o motivo do atraso era por falta de
  - Padrões específicos
  - Implementações e produtos
  - Infraestrutura de gerenciamento

- Entretanto, todos esses motivos já não existem mais
  - Surgiram outros problemas

Por que o IPv6 ainda não foi implementado na Internet?

Instituto de computação

- Problema de demanda
  - Os fabricantes só vão começar a fazer equipamentos IPv6 quando tiver demanda de mercado suficiente
  - Por outro lado, só haverá demanda de mercado quando houver equipamentos prontos
- Problema da base instalada
  - Tornar a Internet totalmente IPv6 requer mudanças de configuração e troca de roteadores em operação
  - Portanto, pode ser necessário estímulo para que tais alterações sejam realizadas

Instituto de computação

Programa de Pós-Graduação
Instituto de Computação
Universidade Federal Fluminense

Aulas 14, 15 e 16

Camada de Rede

Conceitos, modelos de serviço e IP

Igor Monteiro Moraes

Redes de Computadores