Camada de Rede

4

Conceitos Básicos

Roteamento

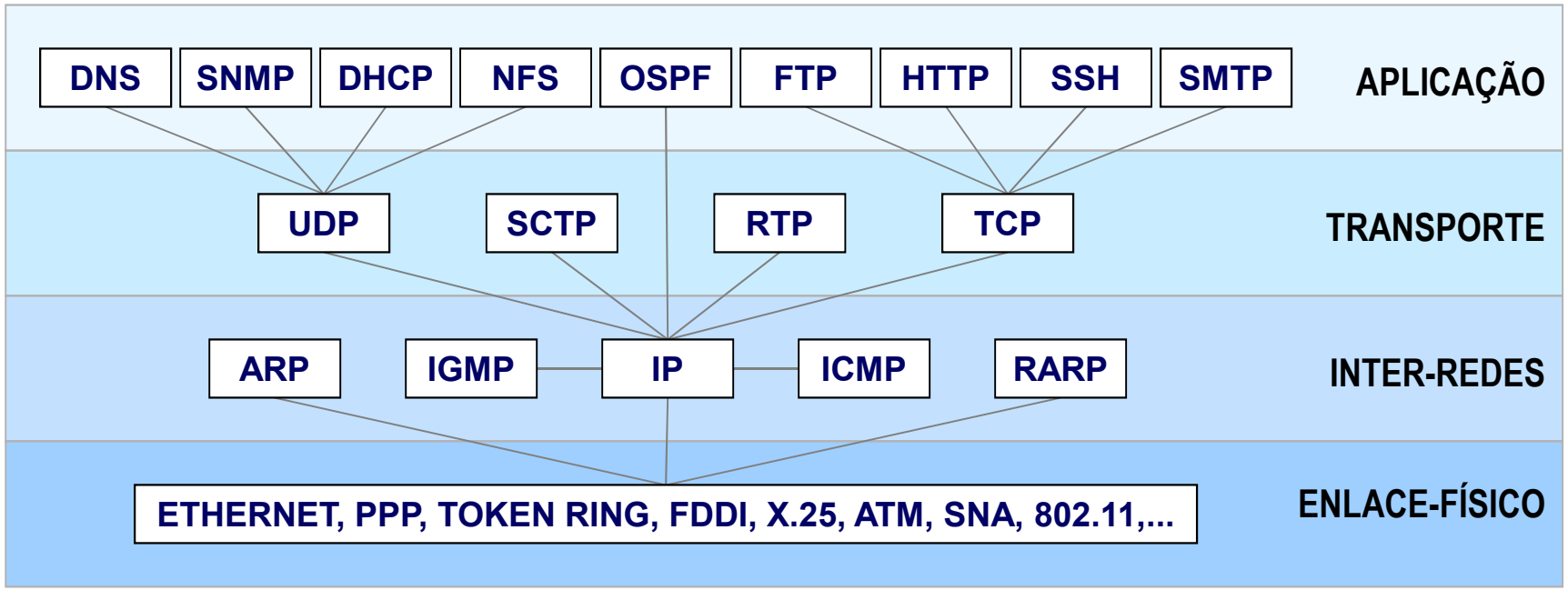
Reinaldo Gomes
reinaldo@dsc.ufcg.edu.br

Internet e TCP/IP

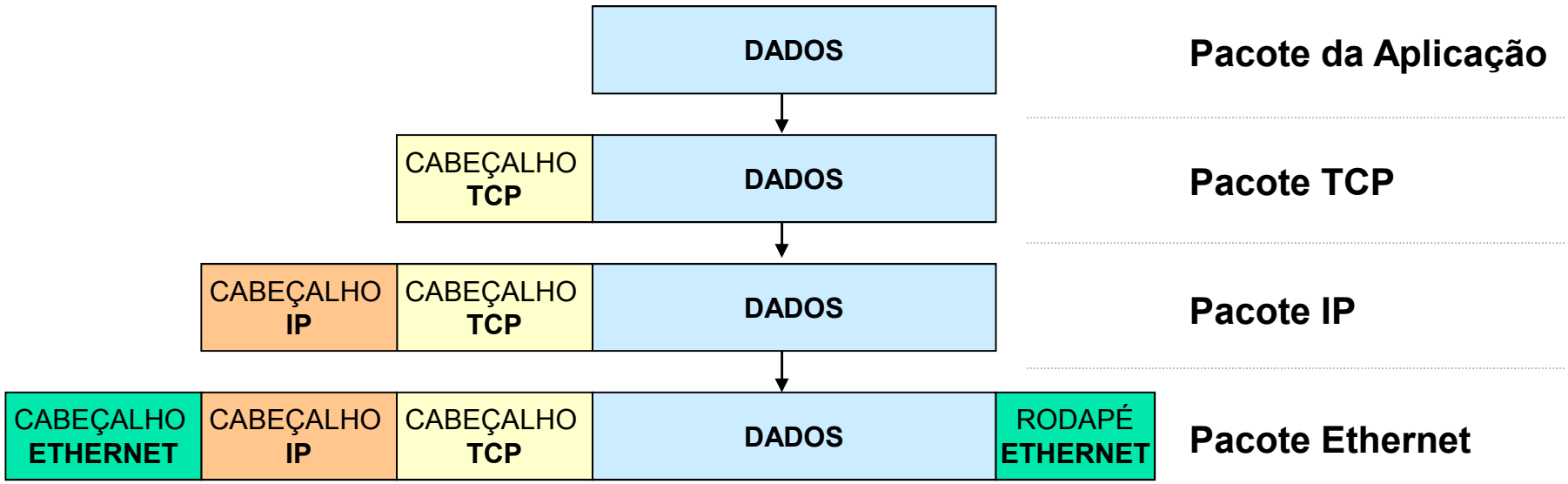
- Internet
  - Conjunto de redes de escala mundial, ligadas pelo protocolo IP
- TCP/IP
  - Família de protocolos de comunicação
  - Serviços e acesso independente de tecnologia
    - Permite a interconexão de redes físicas diferentes
    - Interconexão realizada por roteadores
- Protocolo IP
  - Não orientado a conexão, roteamento melhor esforço
    - Não confiável, sem controle de fluxo e de erros → simples
  - Roteamento baseado no endereço da rede de destino

Arquitetura TCP/IP

- TCP/IP = família de protocolos



Encapsulamento

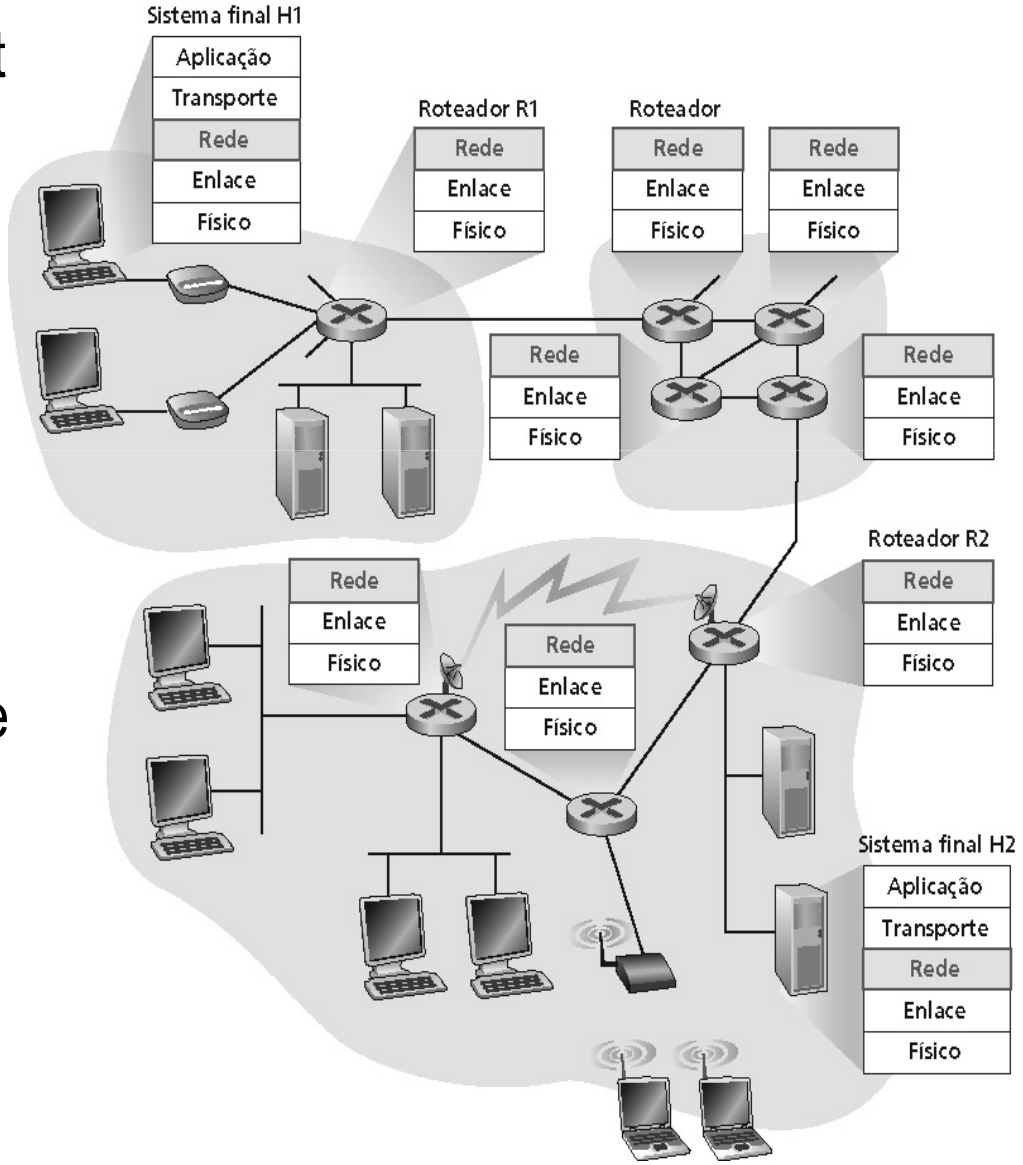


Camada de Rede: Objetivos

- Transferência de pacotes da origem para o destino
- Vários saltos (*hops*) intermediários no caminho
- Elementos de rede: roteadores ou comutadores
- Principal função: roteamento (encaminhamento)
- Outras funções
  - Controle de Congestionamento
  - Negociação de QoS
  - Interconexão de redes

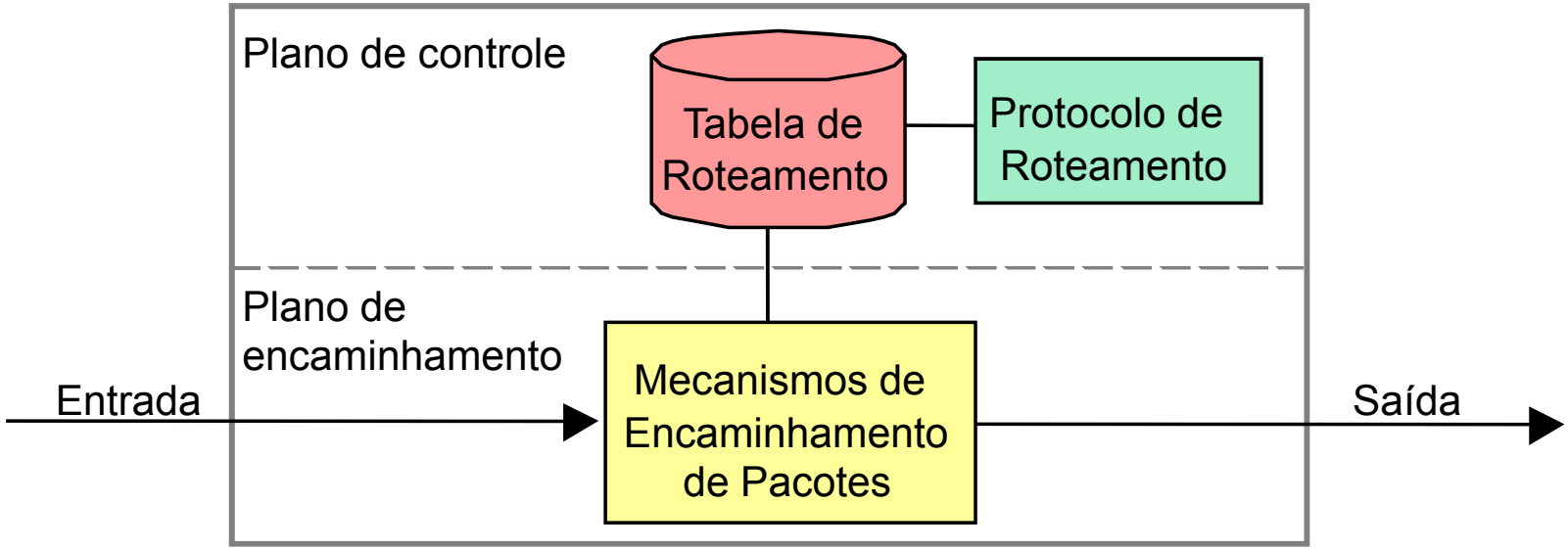
A camada de Rede

- Transporta segmentos do host transmissor para o receptor
- No lado transmissor, encapsula os segmentos em datagramas
- No lado receptor, entrega os segmentos à camada de transporte
- Protocolos da camada de rede em cada host roteador
- Roteador examina campos de cabeçalho em todos os datagramas IP que passam por ele



Roteamento e Encaminhamento

- **Roteamento**: estabelecimento dos melhores caminhos (rotas)
- **Encaminhamento**: processo de despachar cada pacote ao seu destino ou sistema intermediário



Questões de projeto

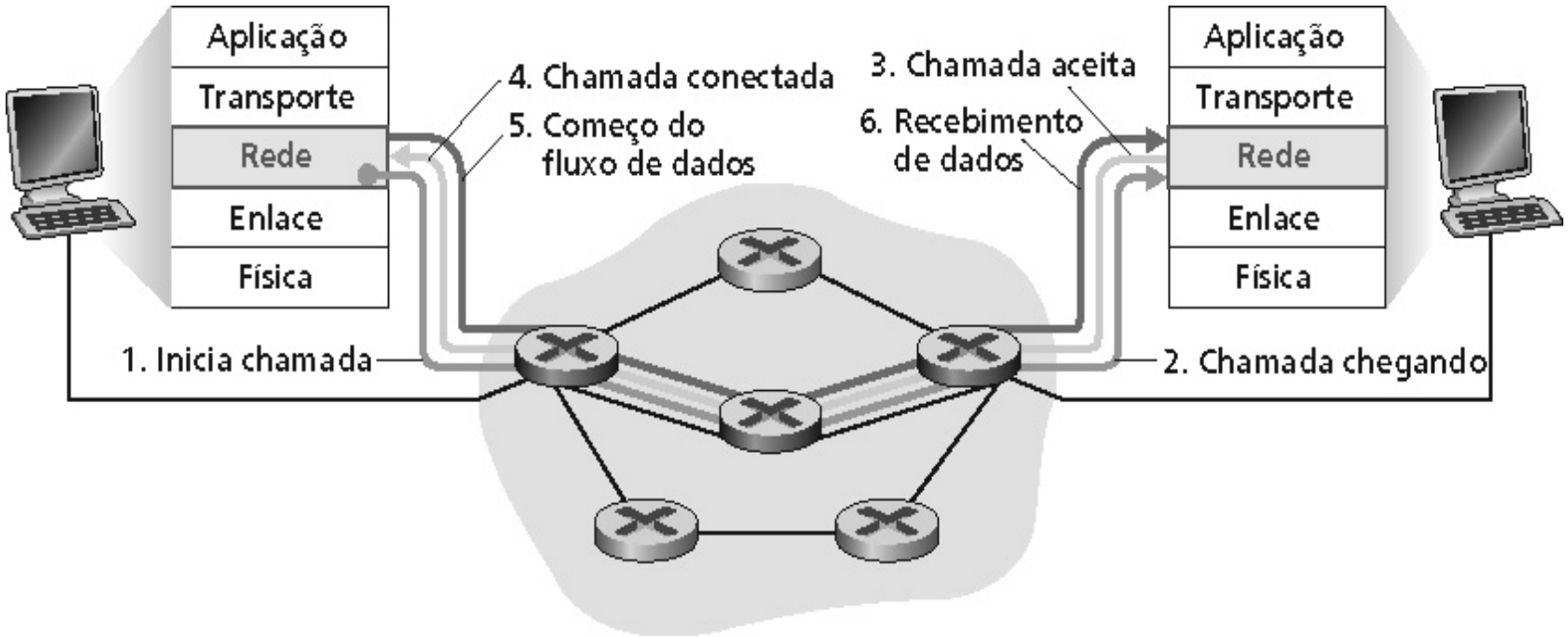
- Serviços oferecidos à camada de Transporte
  - Devem ser independentes da tecnologia da sub-rede
  - Proteção contra tipo, quantidade e topologia das sub-redes
  - Endereços devem ter plano de numeração único
- Tipos de Serviço
  - Orientado a conexão
  - Sem conexão

Organização Interna

- Circuitos Virtuais (CVs)
  - Analogia aos circuitos físicos da rede telefônica
  - Rede complexa e segura
  - Geralmente orientada a conexões (conexões na camada de rede são geralmente chamadas de circuitos virtuais)
  - Ex.: Redes ATM
- Datagramas
  - Analogia com o serviço de “telegramas”
  - Rede simples e não confiável
  - Geralmente é não orientada a conexões (mas pode ser)
  - Ex.: Internet

Circuitos Virtuais

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



Circuitos Virtuais

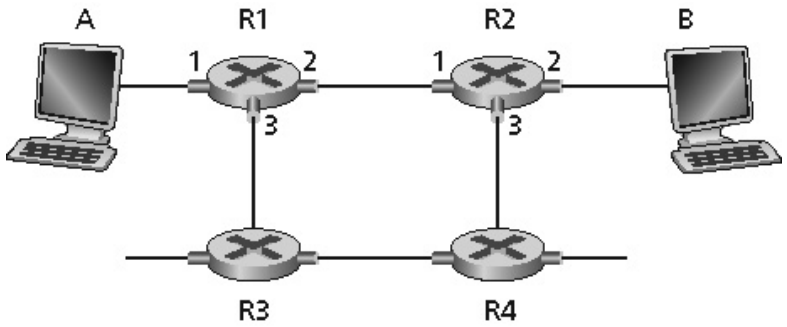


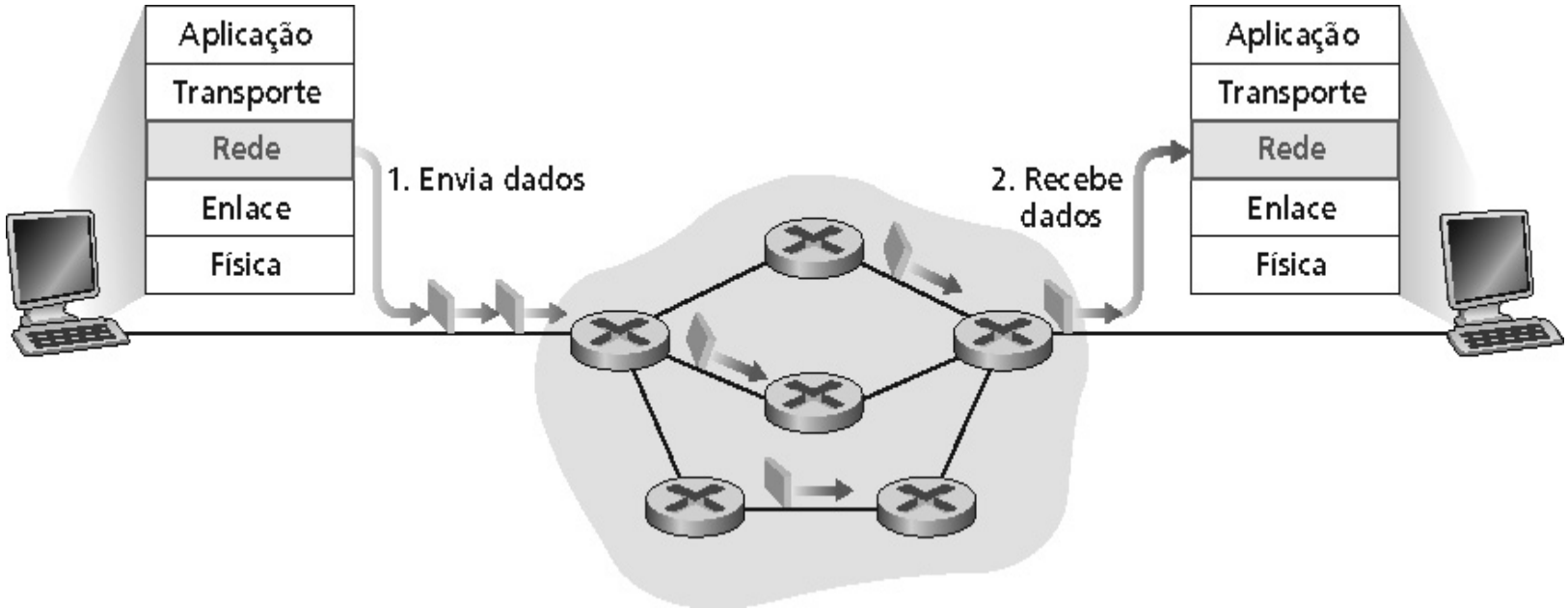
Tabela de comutação no roteador R1:

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

Roteadores mantêm informações de estado de conexão

Rede de Datagrama

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



Circuito Virtual X Datagrama

Questão	Subrede de Datagrama	Subrede de Circuito Virtual (VC)
Configuração de Circuito	Não necessária	Requerida
Endereçamento	Cada pacote contém endereços da fonte e do destino	Cada pacote contém um número de CV
Informação de Estado	Subrede não mantém informação de estado	Cada CV requer espaço na tabela da subrede
Roteamento	Cada pacote é roteado independentemente	Rota escolhida quando CV é configurado; todos os pacotes seguem esta rota
Efeito de falhas do roteador	Nenhuma, exceto a perda de pacotes durante a falha	Todos os CVs que passam pelo roteador com falha são Terminados
Controle de congestionamento	Difícil	Fácil se memória suficiente tiver sido alocada a priori para cada CV

Roteamento

- Algoritmo de roteamento
  - Parte da camada de rede responsável pela decisão sobre a linha de saída para a transmissão de um pacote
  - Pode ser implementado por um protocolo de roteamento, ou executado de maneira estática
- Protocolo de roteamento
  - Software utilizado pelos roteadores para que eles estabeleçam tabelas de roteamento consistentes
  - Qualquer protocolo de roteamento deve comunicar informação da topologia da rede para todos os outros roteadores, para tomar decisões de roteamento

Roteamento

- Datagramas
  - Decisão de roteamento deve ser tomada a cada pacote
- Circuitos Virtuais
  - Decisão de roteamento é tomada somente no estabelecimento da conexão
  - Também chamado de roteamento por sessão (*session*)

Requisitos dos Protocolos de Roteamento

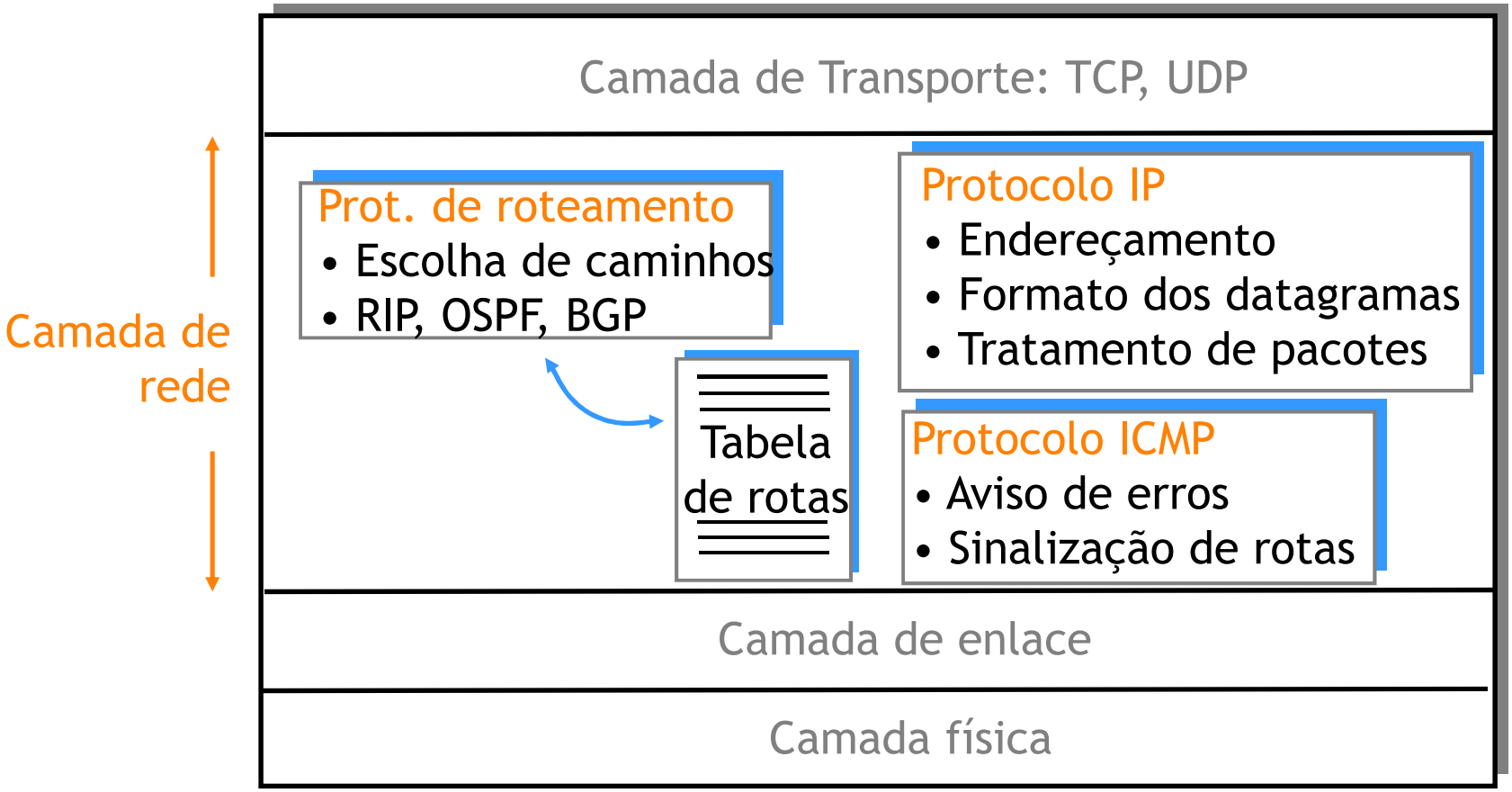
- Minimizar o tamanho da tabela de roteamento
  - Tamanho pode inviabilizar implementação eficiente
- Minimizar as mensagens de controle
  - Podem prejudicar desempenho da rede
- Robustez
  - Deve evitar laços e oscilações no funcionamento
- Usar caminhos ótimos
  - Escolher o "melhor" caminho

Classes de Algoritmos de Roteamento

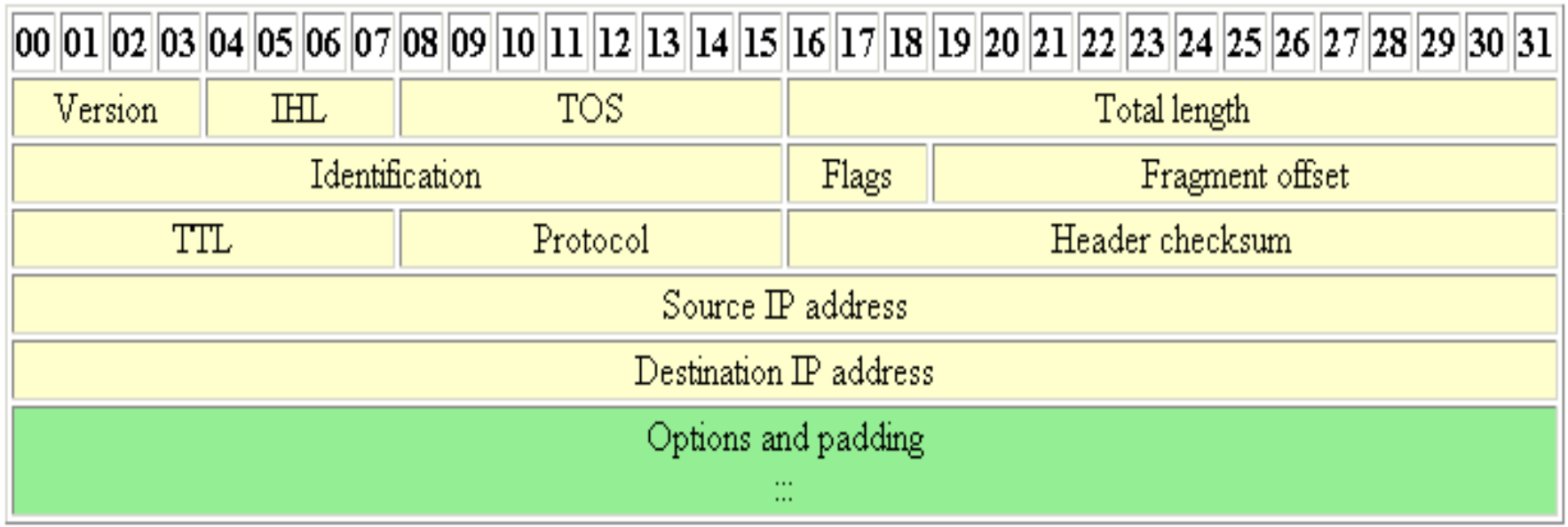
- Não adaptativo (estático)
  - Decisão do roteamento não é baseada em estimativas de tráfego atual e da topologia
- Adaptativo (dinâmico)
  - Muda decisões de roteamento para refletir mudanças na topologia, bem como, no tráfego
  - Métricas
    - Custo, caminho, carga, tamanho da fila

A camada de Rede

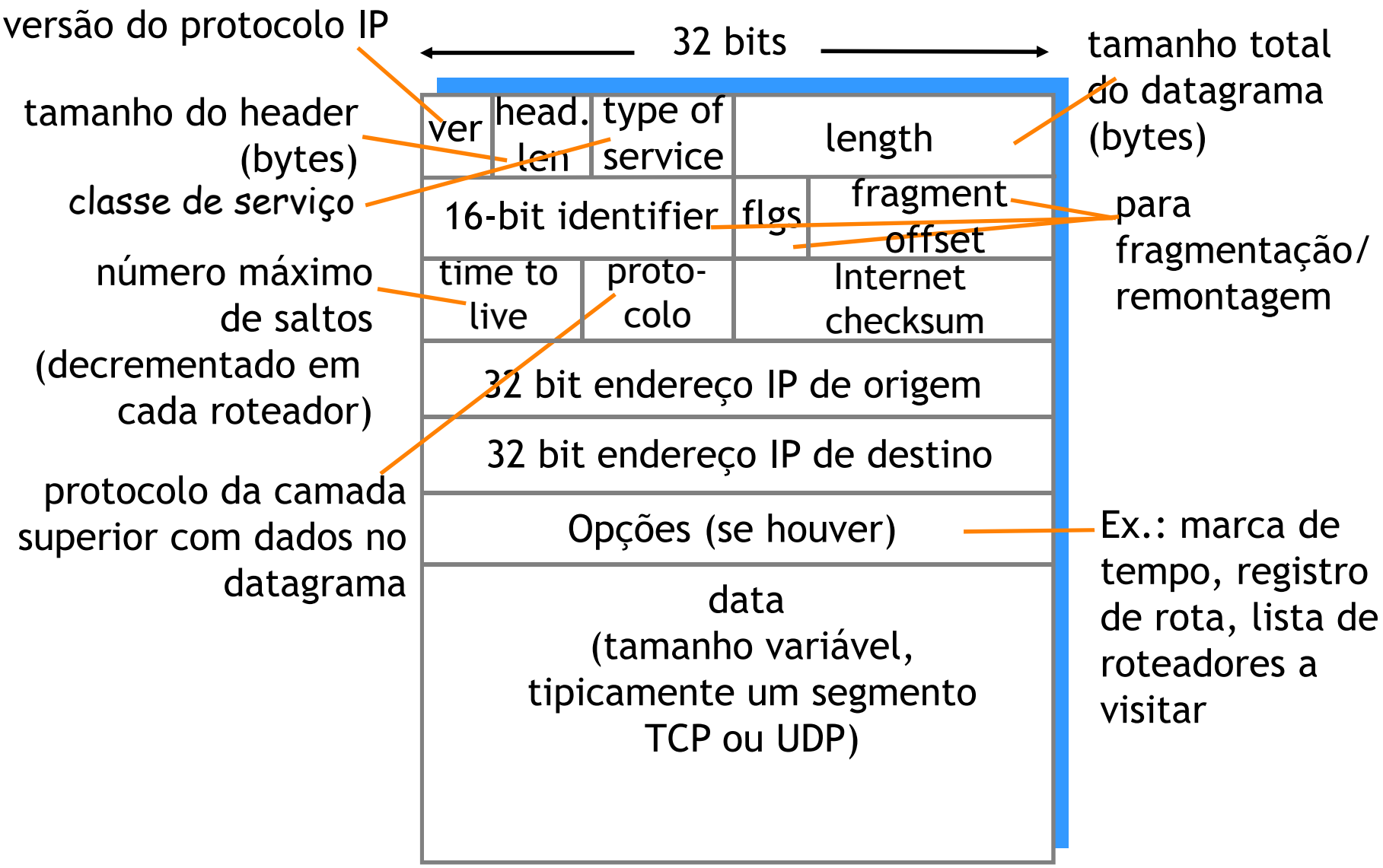
Entidade de rede em roteadores ou hospedeiros:



Cabeçalho IP



Cabeçalho IP



Campos do cabeçalho IP

Campo	Bits	Descrição
Version	**4**	**Identifica a versão do protocolo IP (IPv4, IPv6)**
Header Length	**4**	**Tamanho do cabeçalho (deve ser múltiplo de 32 bits). Se for 20 bytes, deve indicar 5 (5 × 32 = 160 bits, ou 20 bytes).**
Type-of-Service Flags	**8**	**Para permitir priorização de pacotes, dependendo da aplicação.**
Total Packet Length	**16**	**Tamanho total do pacote em bytes, incluindo cabeçalho e dados**
Fragment Identifier	**16**	**Identifica partes de um pacote fragmentado, para ajudar remontagem**
Fragmentation Flags	**3**	**Informações e controle sobre fragmentação (ex: instrução para roteadores não fragmentarem o pacote)**
Fragmentation Offset	**13**	**Posição dos dados deste pacote (fragmentado) em relação ao pacote original (não fragmentado). Múltiplo de 8 bytes.**

Campos do cabeçalho IP

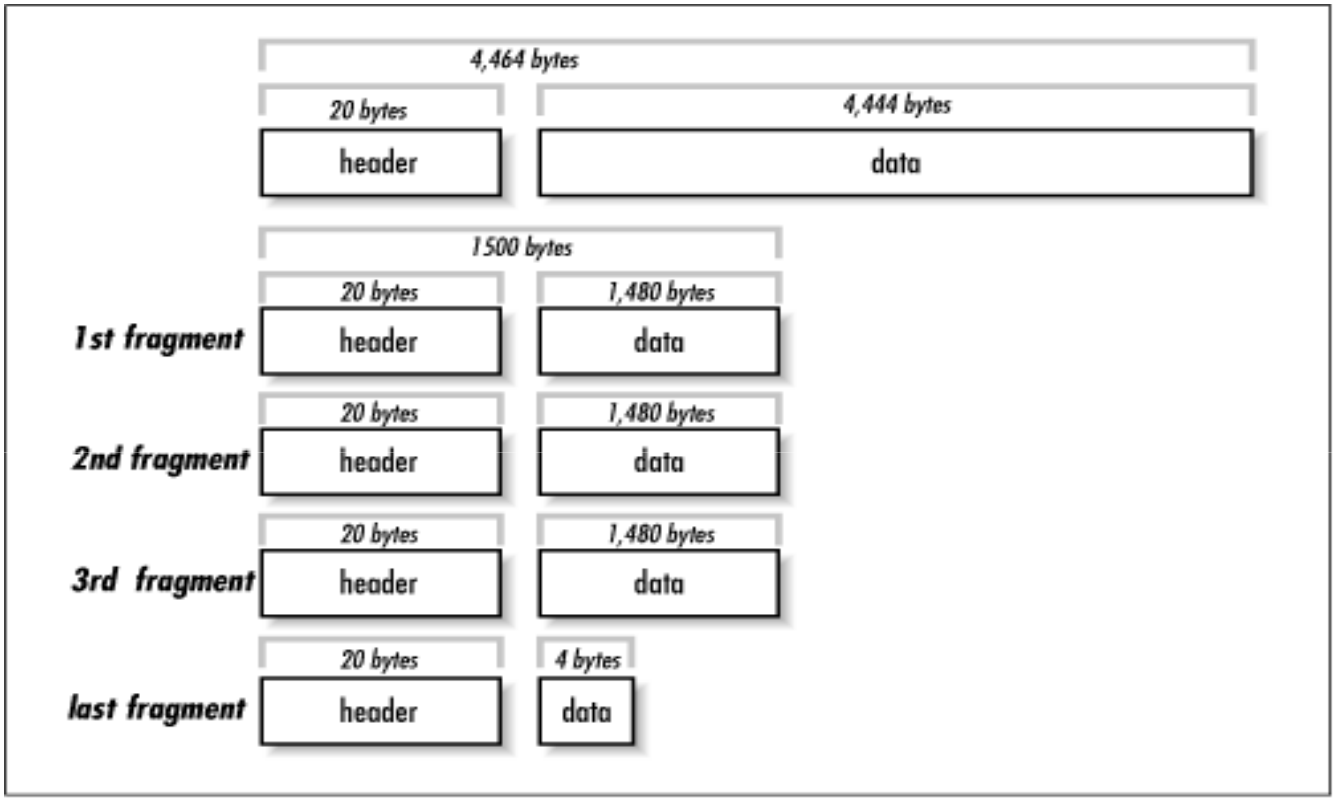
Campo	Bits	Descrição
Time-to-Live	8	Indica o número restante de "hops" que o pacote pode seguir antes de ser considerado inválido e descartado.
Protocol Identifier	8	Identifica o protocolo da camada superior, contido no corpo do pacote.
Header Checksum	16	Checksum do cabeçalho IP
Source IP Address	32	Endereço IP do remetente do pacote
Destination IP Address	32	Endereço IP do destinatário do pacote
Options (optional)	varia	Opções adicionais para type-of-service, Source Routing, Timestamp etc. Raramente usado.
Padding (if required)	varia	Se um pacote não for múltiplo de 32 bits, informações nulas são adicionadas para completar o tamanho.
Data	varia	Os dados que o IP deverá carregar (TCP, UDP, Fragmento, outros protocolos, dados puros ICMP, ...)

Fragmentação e remontagem

- MTU - *Maximum Transmit Unit* é o tamanho máximo que um pacote pode ter
- MTU é padronizado de acordo com a interface física
- Cada roteador deve fragmentar o pacote antes de encaminhá-lo para uma interface, no caso do tamanho original ser maior que o MTU

Topology	**MTU (in bytes)**	**Defined By**
Hyperchannel	65,535	RFC 1374
16 MB/s Token Ring	17,914	IBM
802.4 Token Bus	8,166	RFC 1042
4 MBs Token Ring	4,464	RFC 1042
FDDI	4,352	RFC 1390
DIX Ethernet	1,500	RFC 894
Point-to-Point Protocol (PPP)	1,500	RFC 1548
802.3 Ethernet	1,492	RFC 1042
Serial-Line IP (SLIP)	1,006	RFC 1055
X.25 & ISDN	576	RFC 1356
ARCnet	508	RFC 1051

Fragmentação e remontagem



- Exemplo: Pacote original tem tamanho total 4464 e será encaminhado por uma interface Ethernet
- Deverá ser fragmentado

Fragment	Fragment Identifier	Reserved Flag	May Fragment Flag	More Fragment Flags	Fragment Offset	Packet Length
1	321	0	0	1	0	1,500
2	321	0	0	1	185	1,500
3	321	0	0	1	370	1,500
4	321	0	0	0	555	24

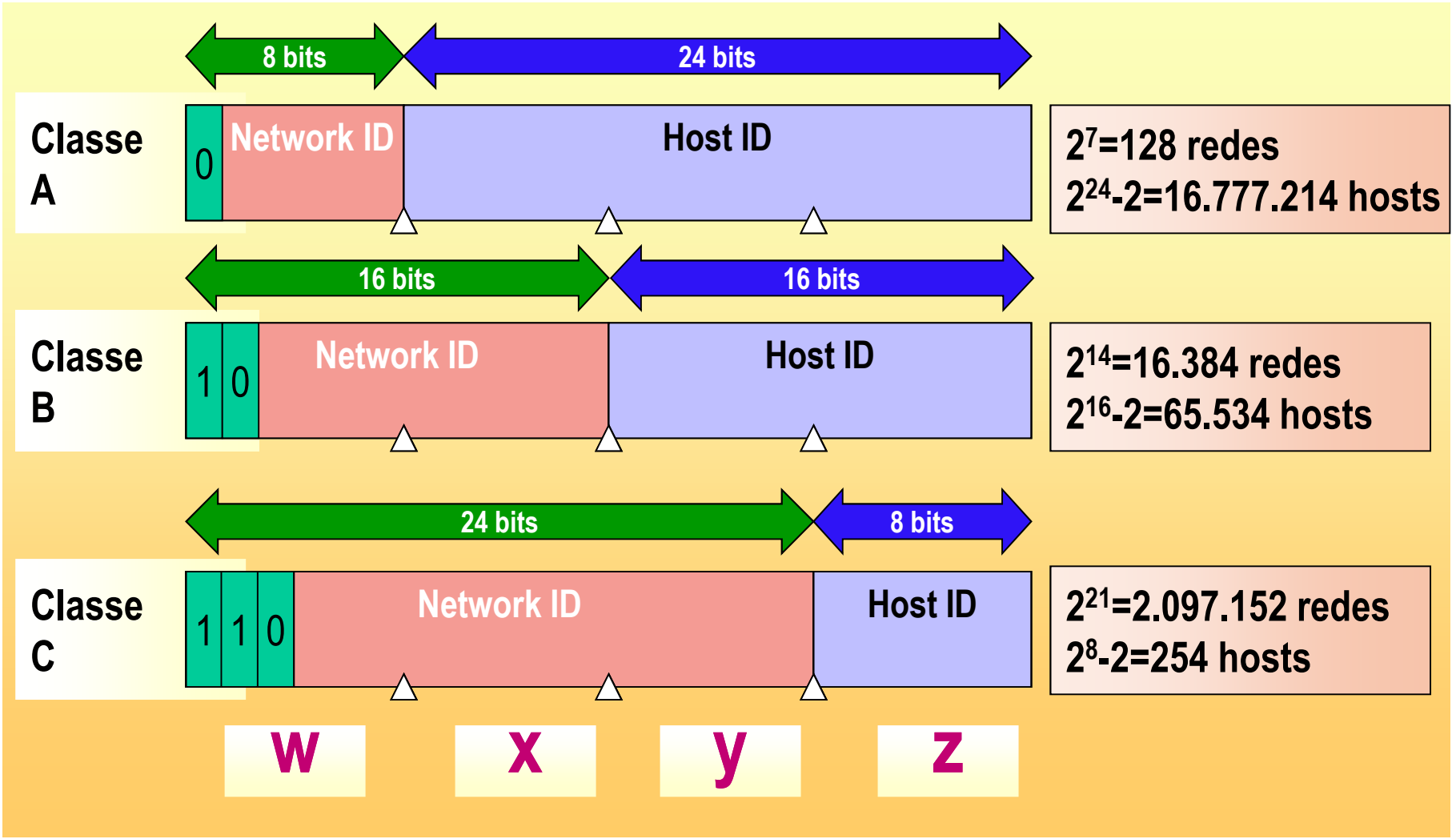
Fragmentação e remontagem

Fragment	Fragment Identifier	Reserved Flag	May Fragment Flag	More Fragment Flags	Fragment Offset	Packet Length
1	321	0	0	1	0	1,500
2	321	0	0	1	185	1,500
3	321	0	0	1	370	1,500
4	321	0	0	0	555	24

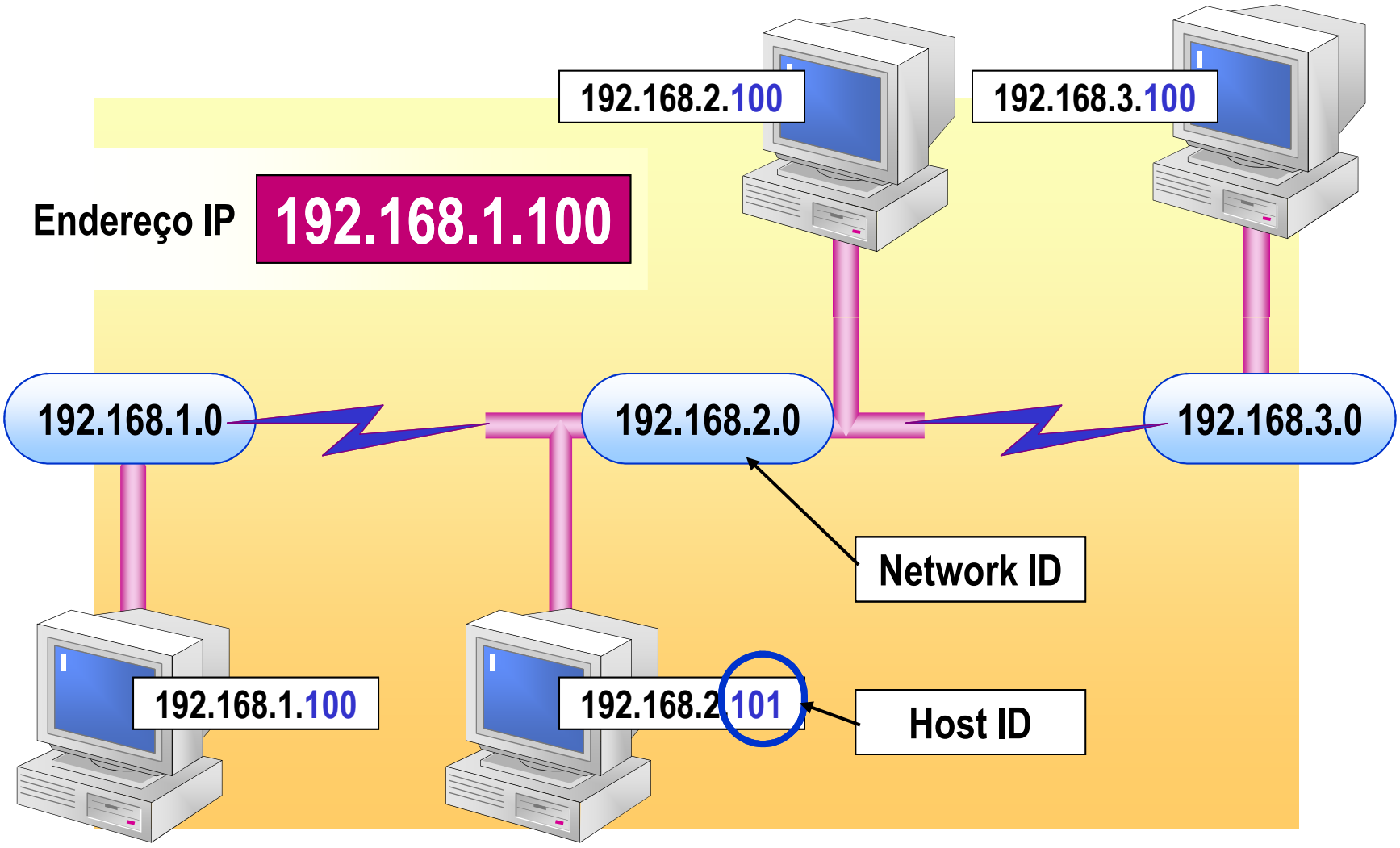
- Cada fragmento pertence a um mesmo pacote original (mesmo valor no campo Fragmentation Identifier)
- O primeiro bit do campo "Flags" deve ser 0 (reservado)
- Fragmentos não podem novamente ser fragmentados, logo "May Fragment" deve ser 0.
- O flag "More Fragments" indica se existe mais fragmentos adiante
- 1o. fragmento → "Fragment Offset" = 0, tam=1500 (20+1480) bytes
- 2o. fragmento → "Fragment Offset" = 185 (1480/8), tam=1500
- 3o. fragmento → "Fragment Offset" = 370 ((1480×2)/8), tam=1500
- 4o. fragmento → "Fragment Offset" = 555 ((1480×3)/8), tam=24 (20+4)

Endereçamento IP

Classes de Endereços IP



Endereços IP



Endereços especiais

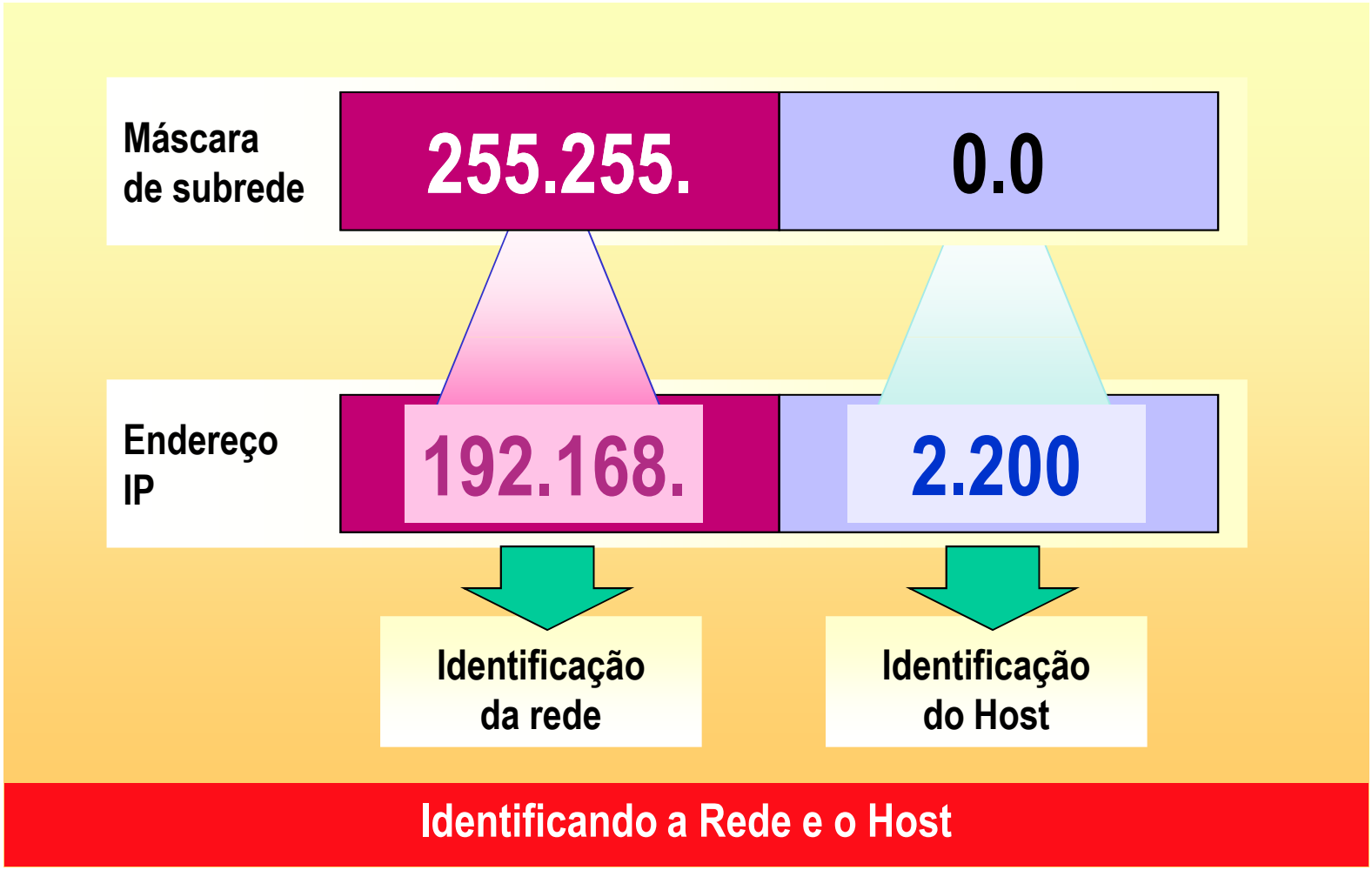
- Subredes privadas – Não podem ser publicados na Internet
  - 10.x.x.x (classe A)
  - 172.16.x.x - 172.31.x.x (classe B)
  - 192.168.x.x (classe C)
- Endereços especiais
  - 0.0.0.0 este host
  - 0.0.0.124 host 124 nesta rede
  - 255.255.255.255 todos os hosts desta rede
  - N.N.N.255 todos os hosts da rede N.N.N
  - 127.X.X.X Loopback
- Alguns endereços inválidos para hosts
  - 10.1.0.0 IP do host não pode ser 0
  - 10.1.0.255 IP do host não pode ser 255
  - 10.123.255.4 Subrede não pode ter valor 255
  - 0.12.16.89 Parte do endereço não pode ter valor 0
  - 255.9.56.45 Parte do endereço não pode ter valor 255
  - 10.34.255.1 Parte do endereço não pode ter valor 255

Máscara de subrede

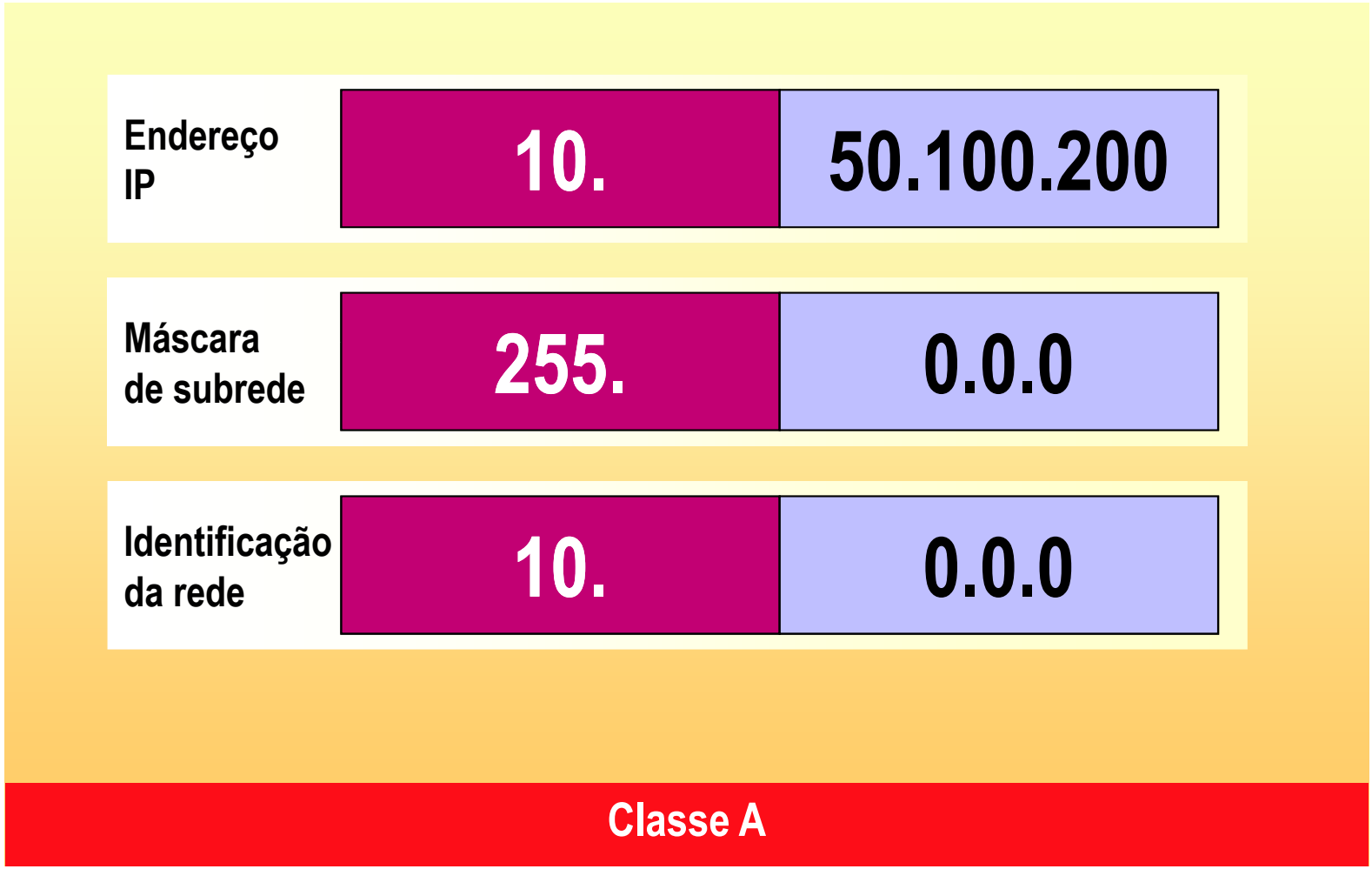
- A máscara serve para indicar qual parte do endereço IP identifica o endereço de rede e qual parte identifica o endereço do host
  - os bits 1 indicam a parte do endereço da rede
  - os bits 0 indicam a parte do endereço do host
- Máscaras default:
  - **Classe A** - 255.0.0.0 11111111.00000000.00000000.00000000
  - **Classe B** - 255.255.0.0 11111111.11111111.00000000.00000000
  - **Classe C** - 255.255.255.0 11111111.11111111.11111111.00000000
- Executa-se um AND lógico entre os bits da máscara e endereço IP e obtém o *Network Address*
- No endereço da rede todos os bits do host são 0
- No endereço de broadcast, todos os bits do host são 1
- Exemplo:
  - **10001100.10110011.11110000.11001000 140.179.240.200 Endereço IP**
    11111111.11111111.00000000.00000000 255.255.000.000 Máscara classe B

    10001100.10110011.00000000.00000000 140.179.000.000 Network Address
    10001100.10110011.11111111.11111111 140.179.255.255 Network Broadcast

Máscaras de Subrede



Máscaras de Subrede



Máscaras de Subrede

Endereço IP	10.50.	100.200
Máscara de subrede	255.255.	0.0
Identificação da rede	10.50.	0.0

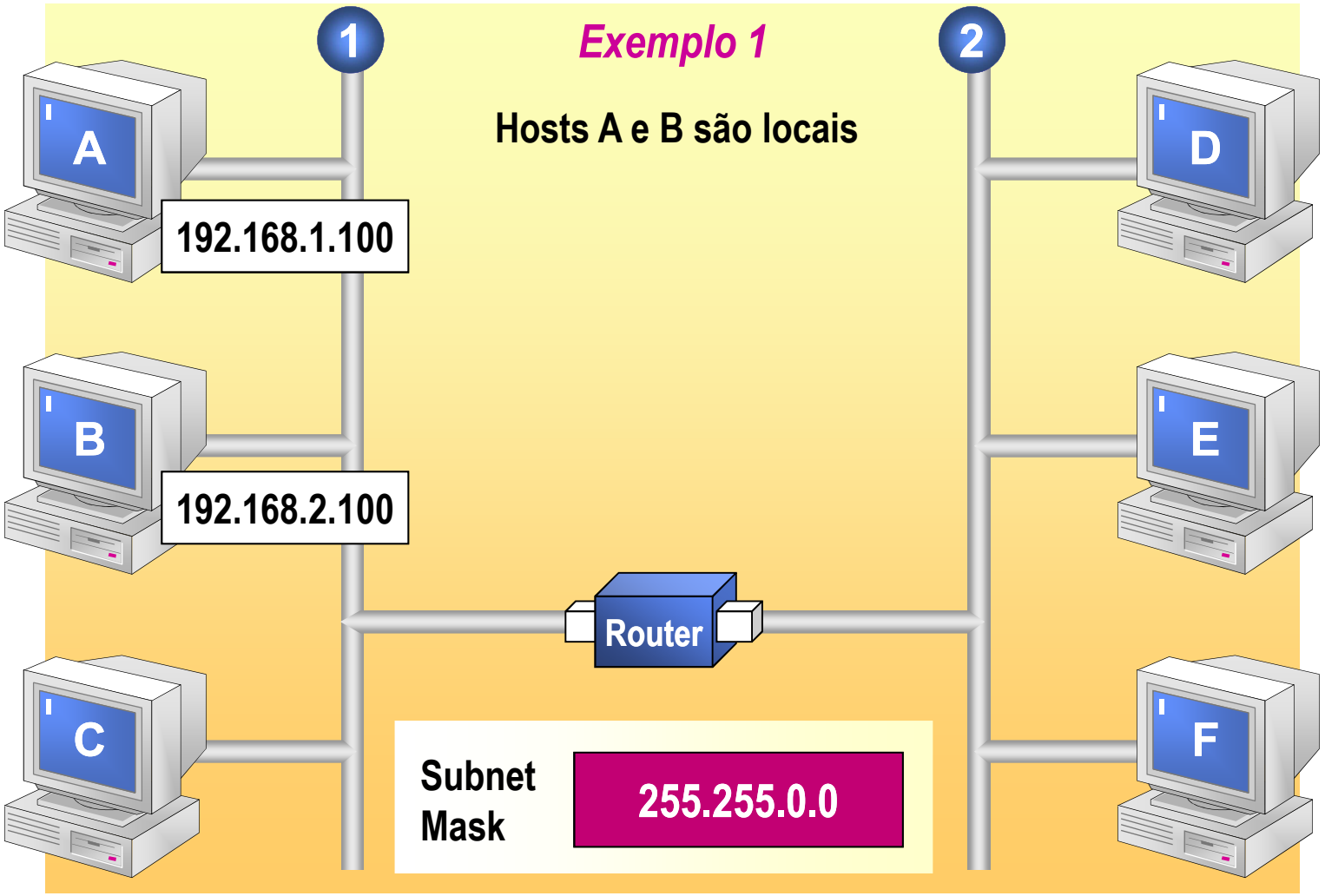
Classe B

Máscaras de Subrede

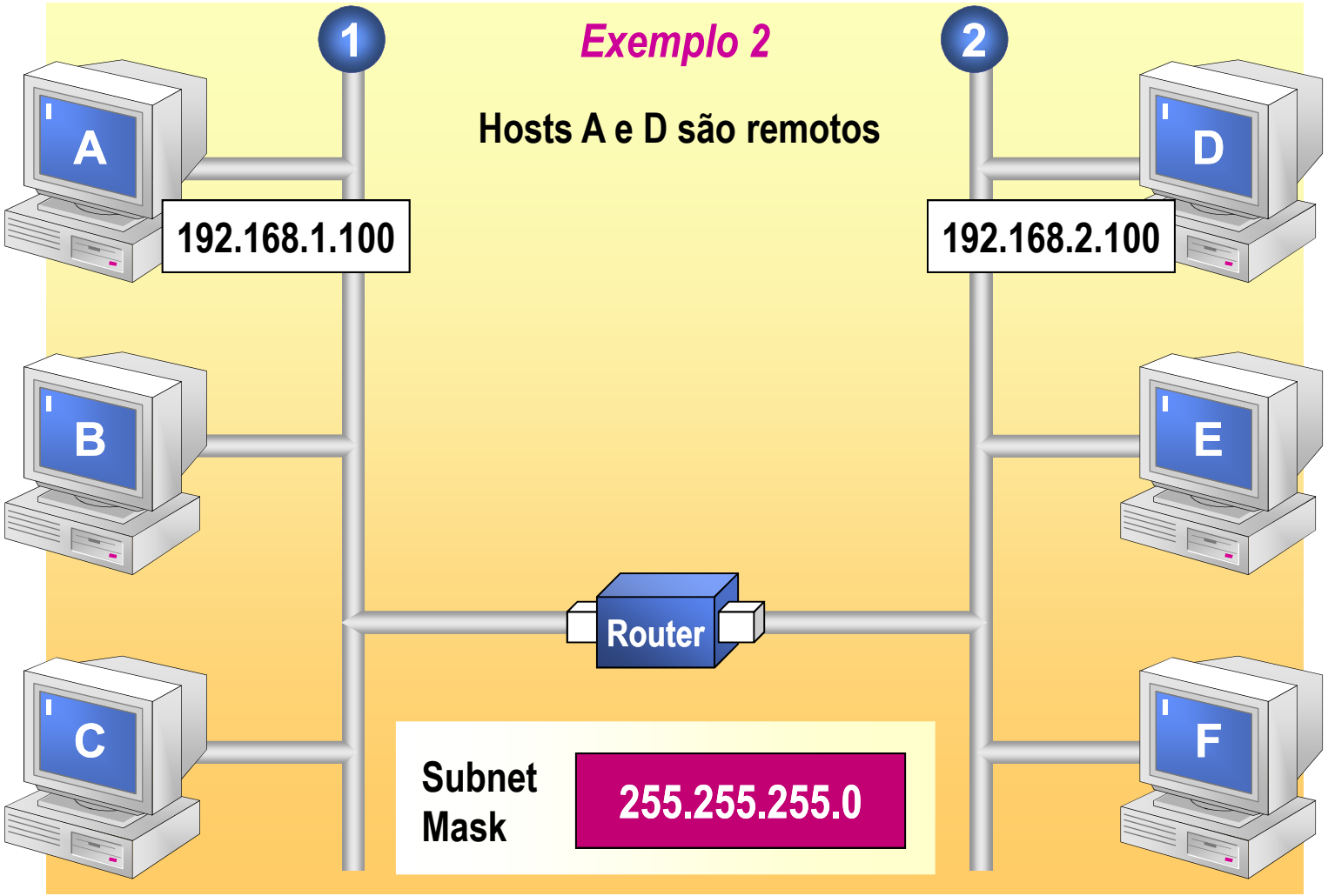
Endereço IP	10.50.100.	200
Máscara de subrede	255.255.255.	0
Identificação da rede	10.50.100.	0

Classe C

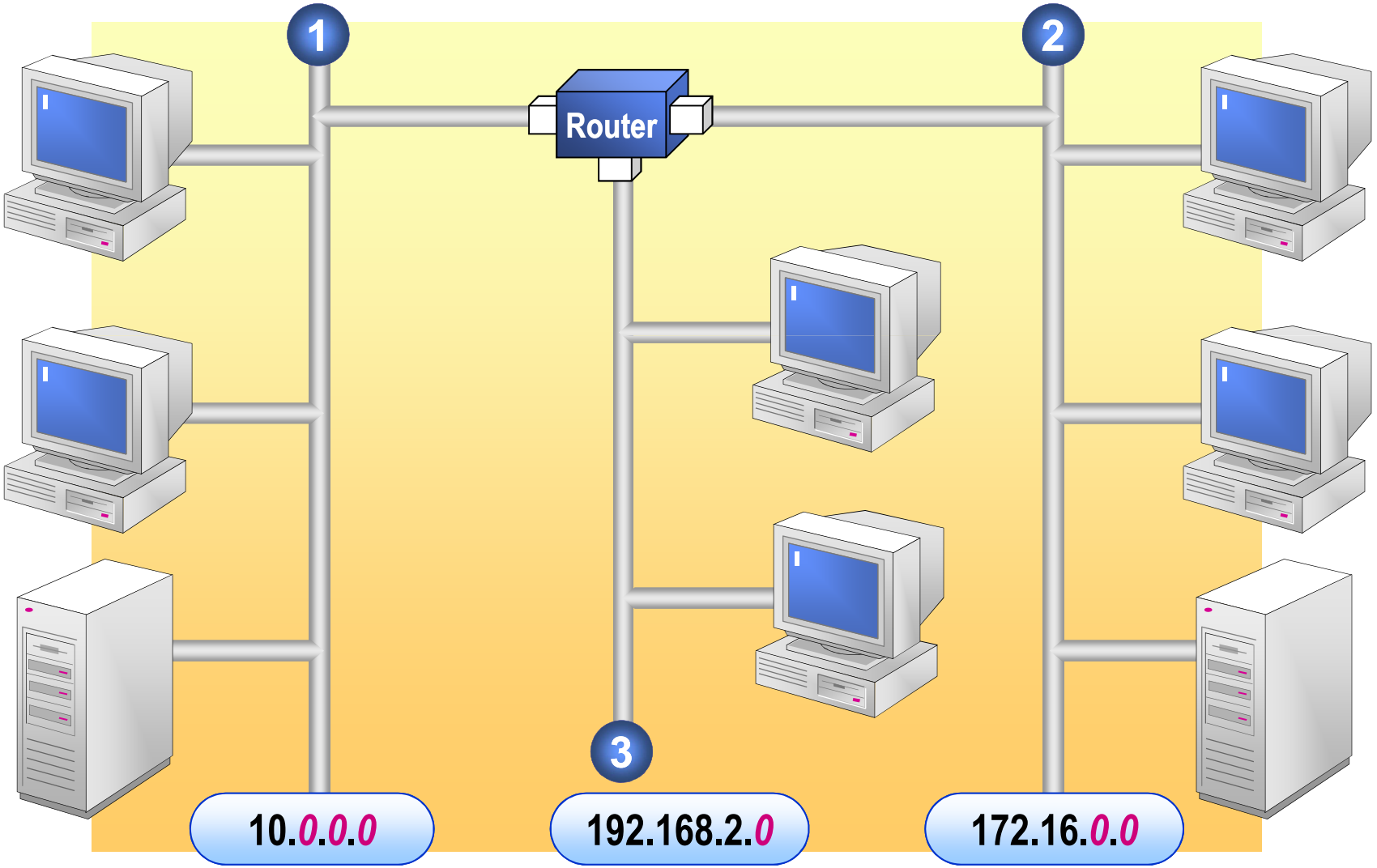
Determinando hosts locais e remotos



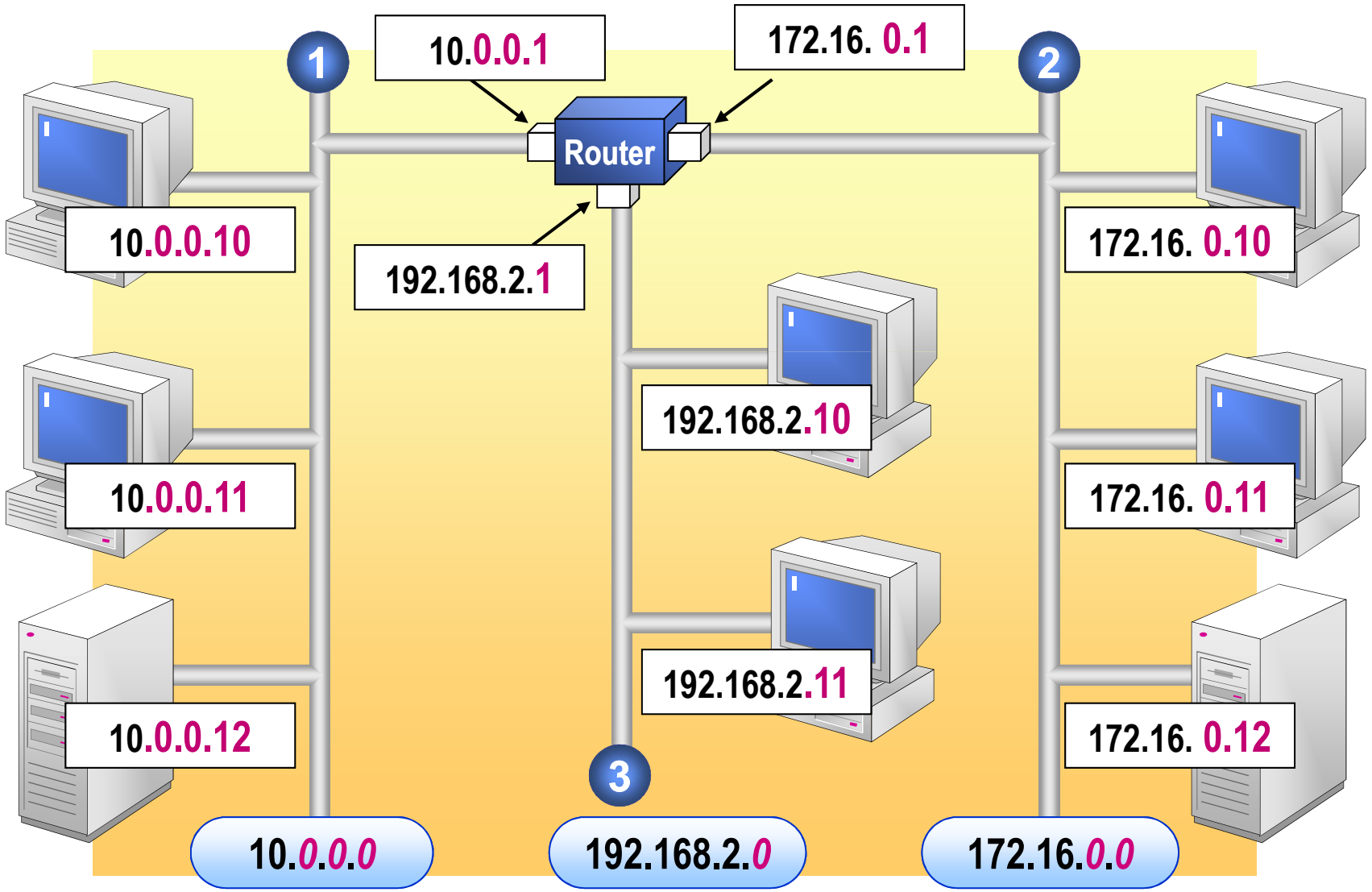
Determinando hosts locais e remotos



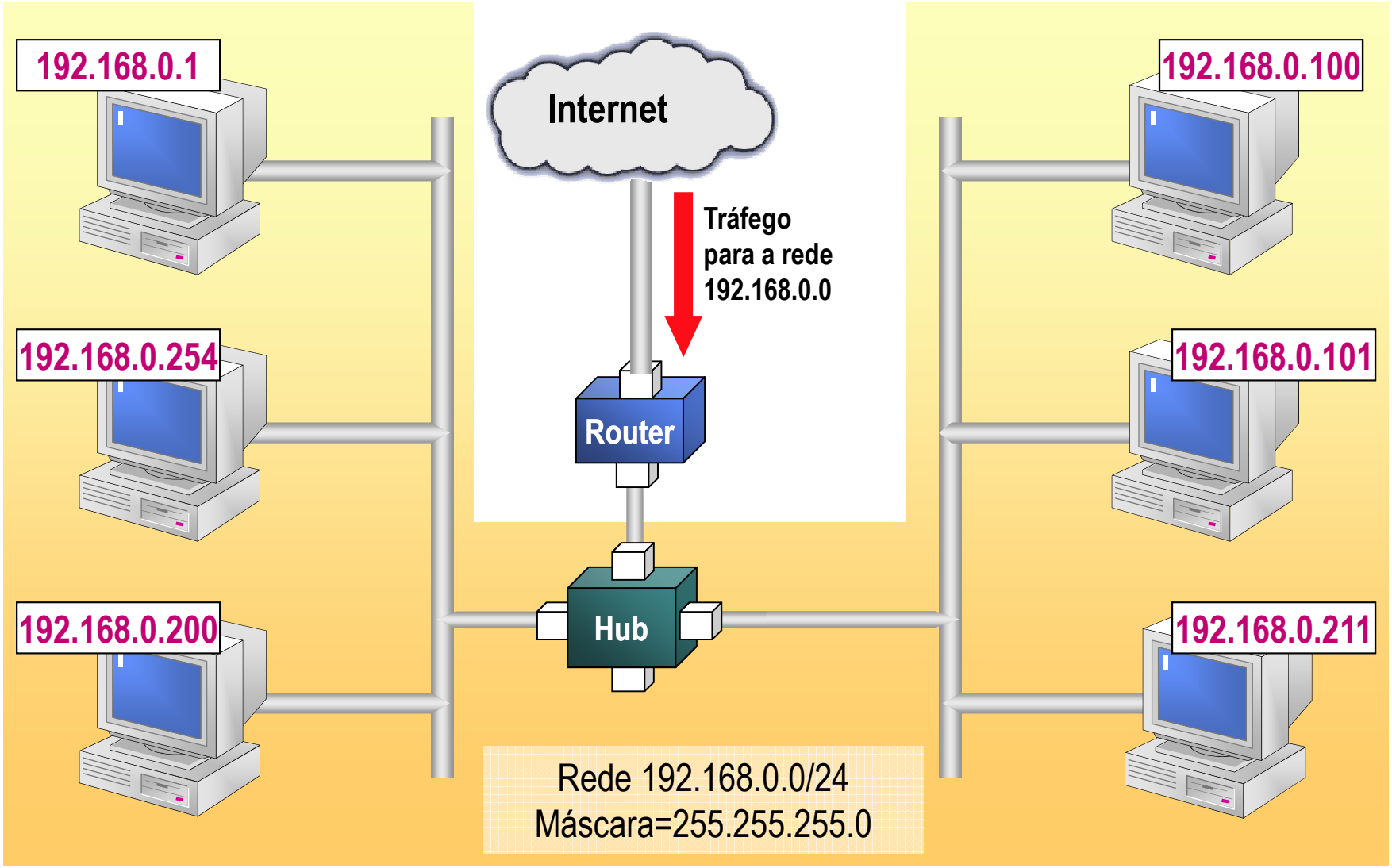
Atribuindo endereços das redes



Atribuindo endereços dos hosts



Subredes


192.168.0.1
192.168.0.254
192.168.0.200
Internet
Tráfego
para a rede
192.168.0.0
Router
Hub
192.168.0.100
192.168.0.101
192.168.0.211
Rede 192.168.0.0/24
Máscara=255.255.255.0

Dividindo um Classe C em 2 subredes

- Exemplo: Rede 192.168.0.0
- Com máscara de subrede 255.255.255.0
  - Mascara = 11111111.11111111.11111111.00000000
  - 1 subrede, notação 192.168.0.0/24

Network	Hosts		Broadcast Address
	from	to	
192.168.0.0	192.168.0.1	192.168.0.254	192.168.0.255

- Rede 192.168.0.0 com máscara de subrede 255.255.255.128
  - Mascara = 11111111.11111111.11111111.10000000
  - 2 subredes, notação 192.168.0.0/25

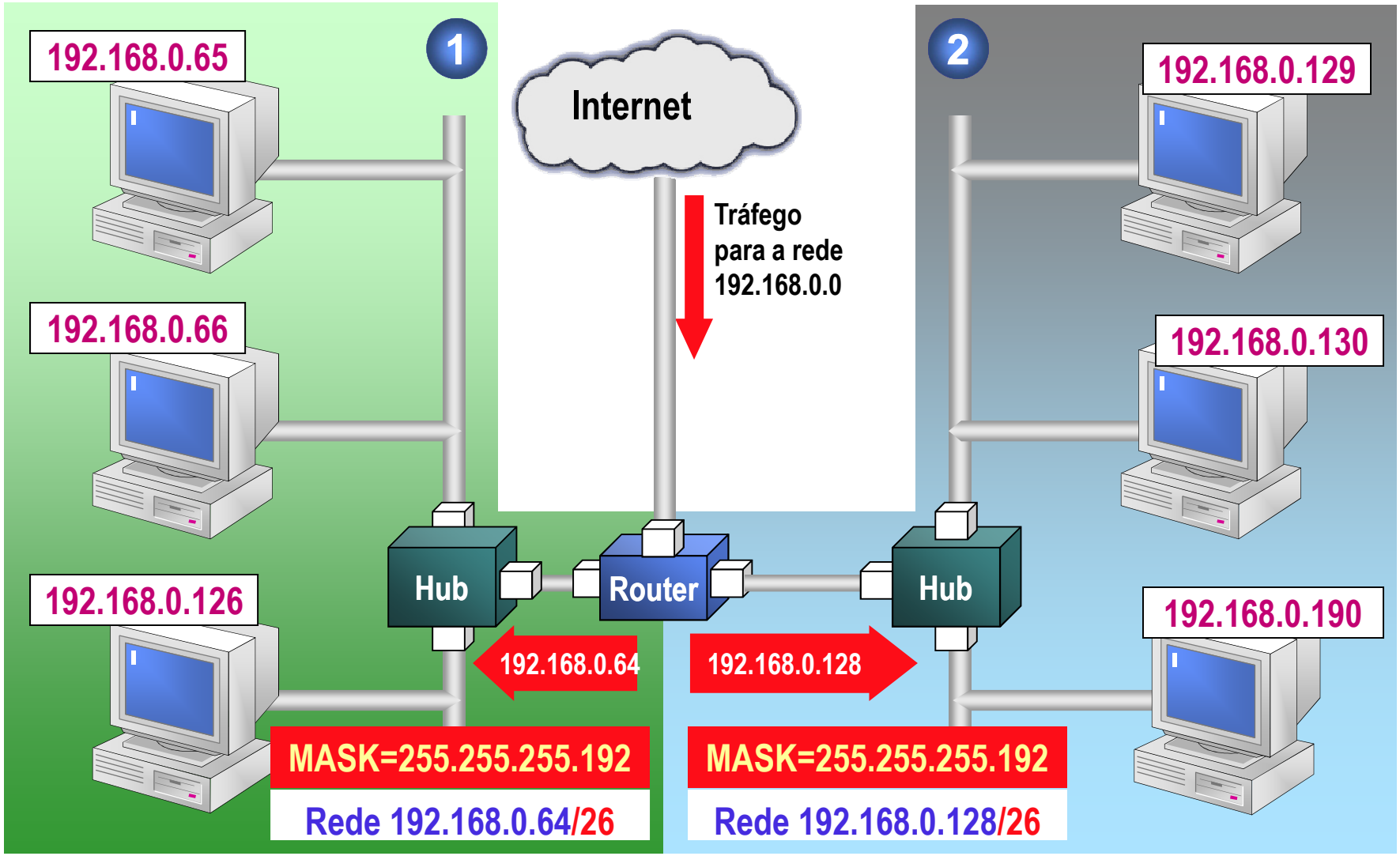
Network	Hosts		Broadcast Address
	from	to	
192.168.0.0	192.168.0.1	192.168.0.126	192.168.0.127
192.168.0.128	192.168.0.129	192.168.0.254	192.168.0.255

Dividindo um Classe C em 4 subredes

- Rede 192.168.0.0 com máscara de subrede 255.255.255.192
  - Mascara = 11111111.11111111.11111111.11000000
  - 4 subredes, notação 192.168.0.0/26

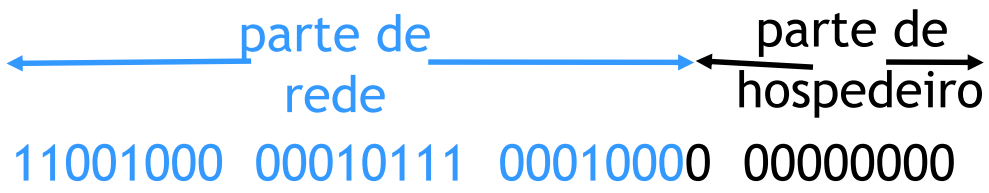
Network	Hosts		Broadcast Address
	from	to	
192.168.0.0	192.168.0.1	192.168.0.62	192.168.0.63
192.168.0.64	192.168.0.65	192.168.0.126	192.168.0.127
192.168.0.128	192.168.0.129	192.168.0.190	192.168.0.191
192.168.0.192	192.168.0.193	192.168.0.254	192.168.0.255

Subredes


1
2
192.168.0.65
192.168.0.66
192.168.0.126
Internet
Tráfego
para a rede
192.168.0.0
Hub
Router
Hub
192.168.0.64
192.168.0.128
MASK=255.255.255.192
Rede 192.168.0.64/26
MASK=255.255.255.192
Rede 192.168.0.128/26
192.168.0.129
192.168.0.130
192.168.0.190

Endereçamento IP atualmente: CIDR

- CIDR: Classless InterDomain Routing (pronuncia-se como a palavra cider)
- A porção de endereço de rede tem tamanho arbitrário
- Formato do endereço: a.B.C.D/x, em que x é o número de bits na parte de rede do endereço



200.23.16.0/23

CIDR: endereços reservados

CIDR Bloco de Endereços	Descrição	Referência
0.0.0.0/8	Rede corrente (só funciona como endereço de origem)	RFC 1700
10.0.0.0/8	Rede Privada	RFC 1918
14.0.0.0/8	Rede Pública	RFC 1700
39.0.0.0/8	Reservado	RFC 1797
127.0.0.0/8	Localhost	RFC 3330
128.0.0.0/16	Reservado (IANA)	RFC 3330
169.254.0.0/16	Zeroconf	RFC 3927
172.16.0.0/12	Rede Privada	RFC 1918
191.255.0.0/16	Reservado (IANA)	RFC 3330
192.0.0.0/24		
192.0.2.0/24	Documentação	RFC 3330
192.88.99.0/24	IPv6 para IPv4	RFC 3068
192.168.0.0/16	Rede Privada	RFC 1918
198.18.0.0/15	Teste de benchmark de redes	RFC 2544
223.255.255.0/24	Reservado	RFC 3330
224.0.0.0/4	Multicasts (antiga rede Classe D)	RFC 3171
240.0.0.0/4	Reservado (antiga rede Classe E)	RFC 1700
255.255.255.255	Broadcast	

Endereçamento IP: CIDR

- CIDR: Classless InterDomain Routing
  - Inicialmente somente grupos de endereços Classe C foram utilizados
  - Como o mesmo procedimento já foi também aplicado às antigas classes A e B, pode-se dizer que de fato o endereçamento em classes está descaracterizado e completamente substituído pelo CIDR (obs.: vários endereços de antigas classes A e B foram progressivamente realocados)

200.23.16.0/23

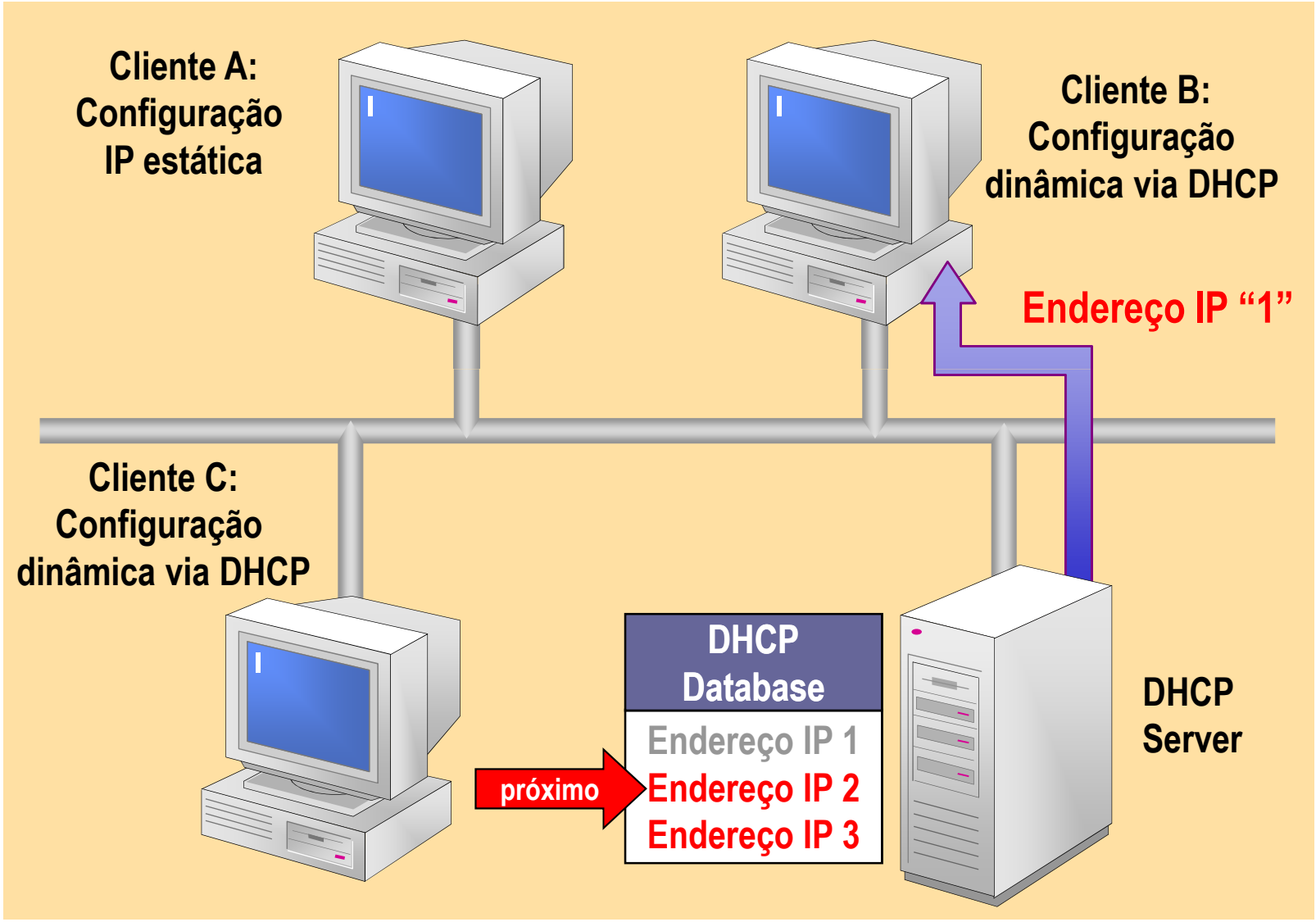
Como obter um endereço IP

- Como um sistema final pode obter um endereço IP?
  - Configuração estática
  - DHCP: Dynamic Host Configuration Protocol: obtém dinamicamente endereços IP de um servidor
    - “plug-and-play”

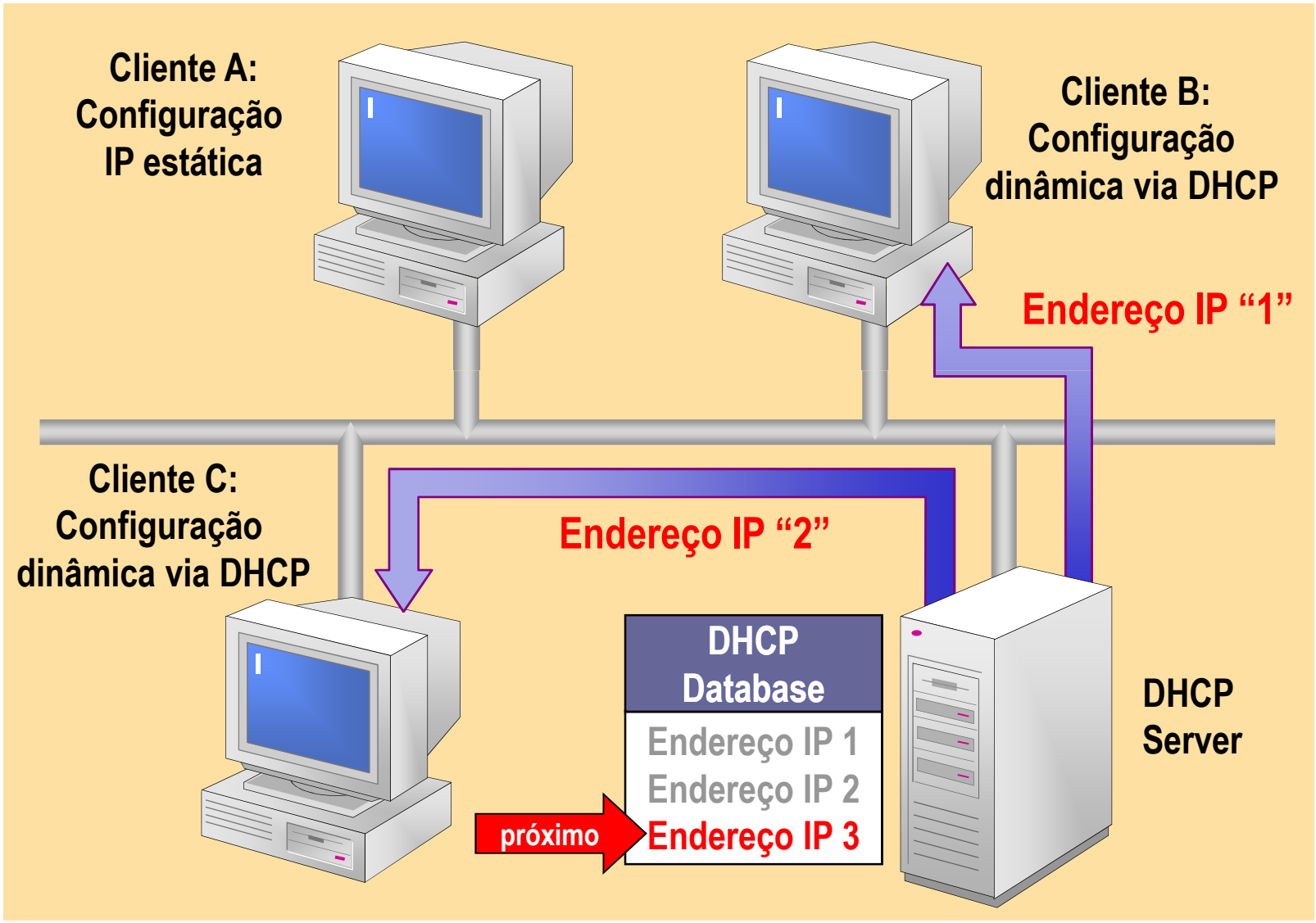
- Como uma rede obtém a parte de sub-rede do endereço IP ?
  - Obtém a porção alocada no espaço de endereço do seu provedor ISP (ICANN: internet corporation for assigned names and numbers)
    - Aloca endereços, gerencia DNS, atribui nomes de domínios, ...

bloco do ISP	11001000 00010111 00010000	00000000	200.23.16.0/20
Organização 0	11001000 00010111 00010000	00000000	200.23.16.0/23
Organização 1	11001000 00010111 00010010	00000000	200.23.18.0/23
Organização 2	11001000 00010111 00010100	00000000	200.23.20.0/23
...	…	…	…

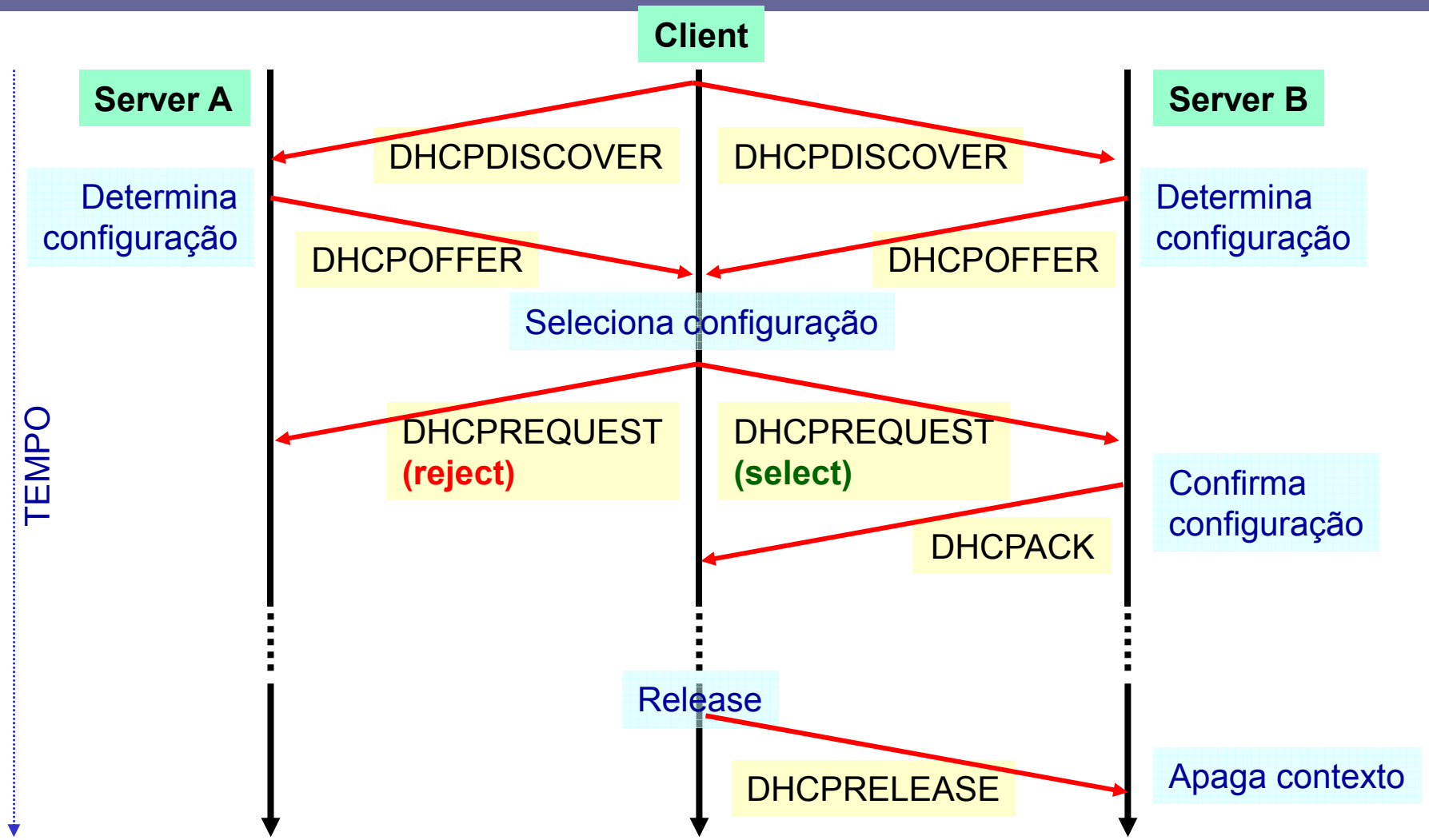
Como funciona o DHCP



Como funciona o DHCP



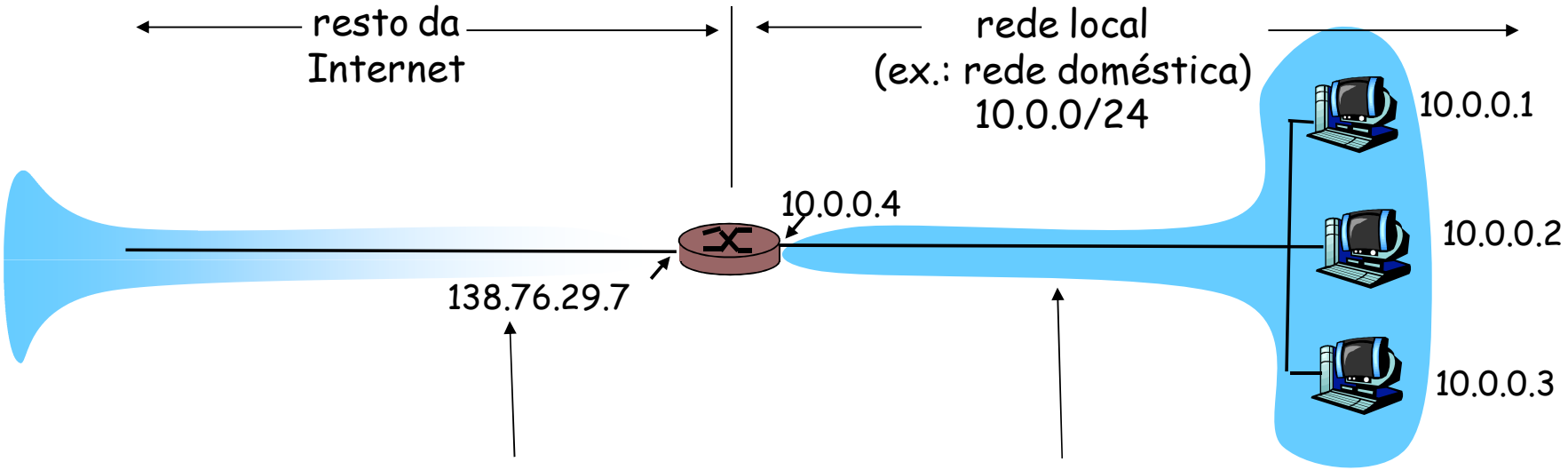
DHCP – o que acontece na rede



DHCP: Alguns Problemas

- Segurança:
  - Mensagens não são autenticadas
    - Alguém pode forjar um DHCP server ou um cliente
  - Cliente não pode confiar no servidor e vice-versa
- Configuração:
  - Para redes com mais de um servidor
    - Servidores na rede não podem trocar informações
    - Não existe um protocolo server-server
    - Devem ter espaços de endereçamento disjuntos para evitar distribuição de IP duplicados
  - Servidores são configurados manualmente

NAT: Network Address Translation



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

datagramas com origem ou destino nesta rede possuem endereço 10.0.0/24 para origem, destino (usualmente)

NAT: Network Address Translation

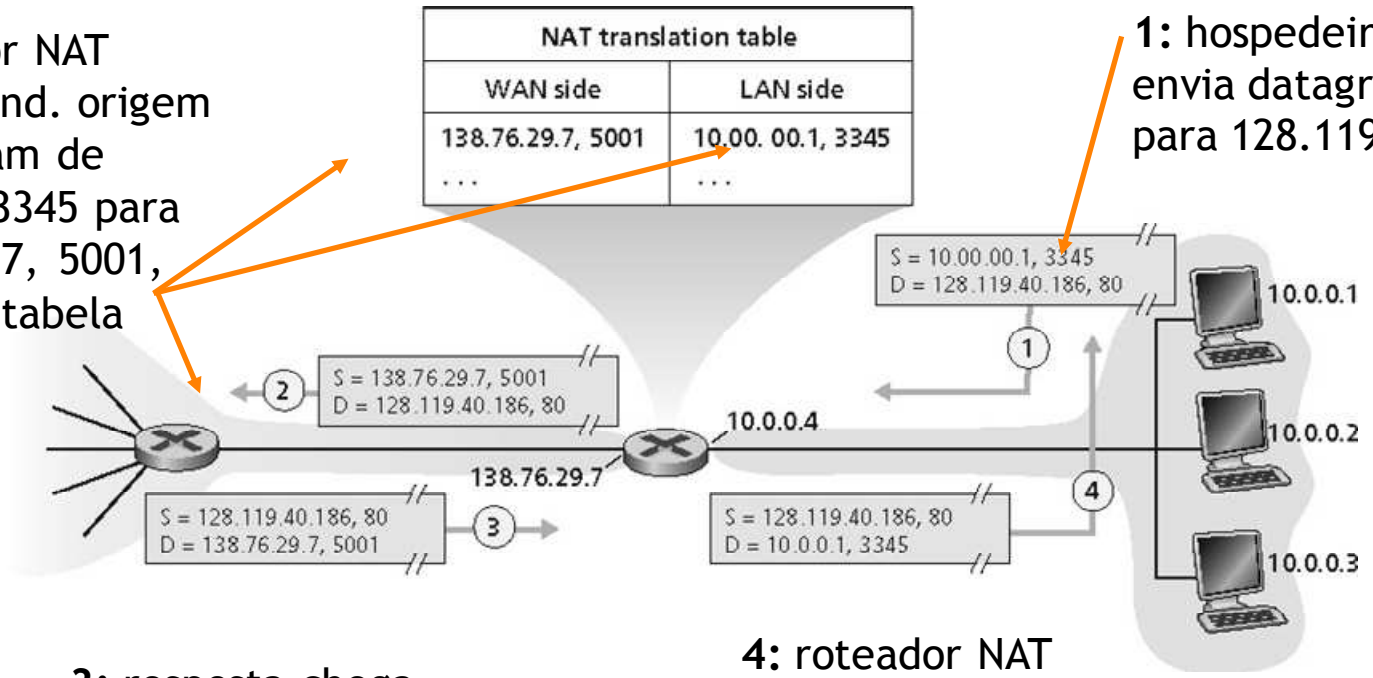
- Motivação: redes locais podem utilizar apenas um endereço IP:
  - Não é preciso alocar uma gama de endereços do ISP, apenas um endereço IP é usado para todos os dispositivos
  - Podem-se alterar os endereços dos dispositivos na rede local sem precisar notificar o mundo exterior
  - Pode-se mudar de ISP sem alterar os endereços dos dispositivos na rede local
  - Dispositivos da rede local não são explicitamente endereçáveis ou visíveis pelo mundo exterior (um adicional de segurança).

NAT: Network Address Translation

- Implementação: o roteador NAT deve:
  - Tratar datagramas que saem: substituir (endereço IP de origem, porta #) de cada datagrama para (endereço IP do NAT, nova porta #)
    - . . . clientes/servidores remotos responderão usando (endereço IP do NAT, nova porta #) como endereço de destino
  - Lembrar (na tabela de tradução do NAT) cada (endereço IP de origem, porta #) para o par de tradução (endereço IP do NAT, nova porta #).
  - Tratar datagramas que chegam: substituir (endereço IP do NAT, nova porta #) nos campos de destino de cada datagrama pelos correspondentes (endereço IP de origem, porta #) armazenados da tabela NAT

NAT: Network Address Translation

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



1: hospedeiro 10.0.0.1 envia datagrama para 128.119.40, 80

3: resposta chega endereço de destino: 138.76.29.7, 5001

4: roteador NAT substitui o endereço de destino do datagrama de 138.76.29.7, 5001 para 10.0.0.1, 3345

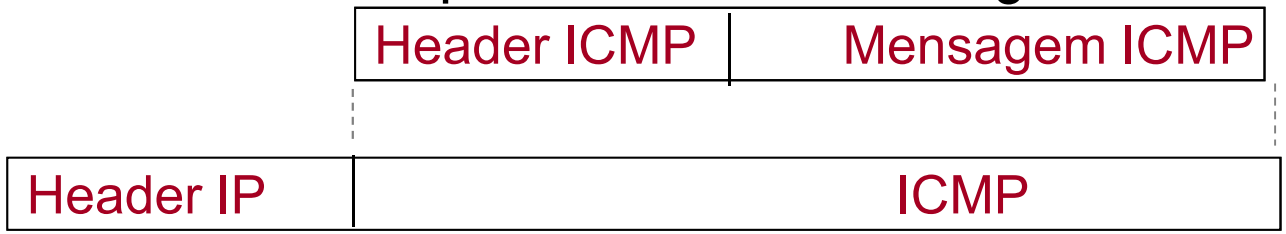
NAT: Network Address Translation

- Campo número de porta com 16 bits:
  - Possibilidade de 65536 conexões simultâneas com um único endereço de LAN!!!
- NAT é controverso:
  - Roteadores deveriam processar somente até a camada de rede (*Layer* 3): Violação do argumento fim-a-fim
  - A possibilidade de NAT deve ser levada em conta pelos desenvolvedores de aplicações: ex.: aplicações P2P
  - A escassez de endereços deveria ser resolvida pelo IPv6

ICMP

ICMP

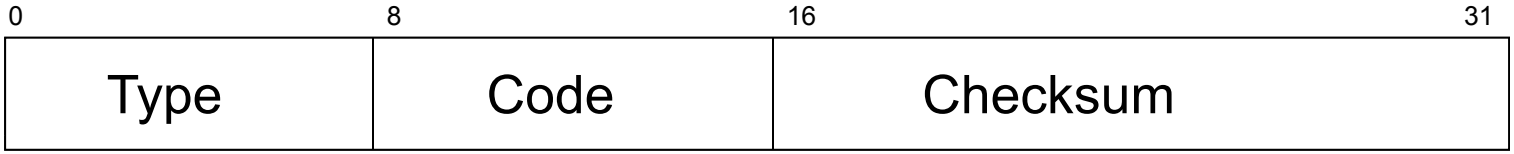
- Internet Control Message Protocol (ICMP)
  - Utilizado para enviar mensagens de erro e de controle
  - Protocolo de Nível 3, encapsulado em um datagrama IP



- Não é enviada uma mensagem ICMP para mensagens ICMP de erro
  - ICMP não relata erros que ocorram em mensagens ICMP
- Os erros são relatados somente sobre o primeiro fragmento (offset = 0)

Frame ICMP

- O ICMP tem um frame básico, comum a todos os tipo de mensagem



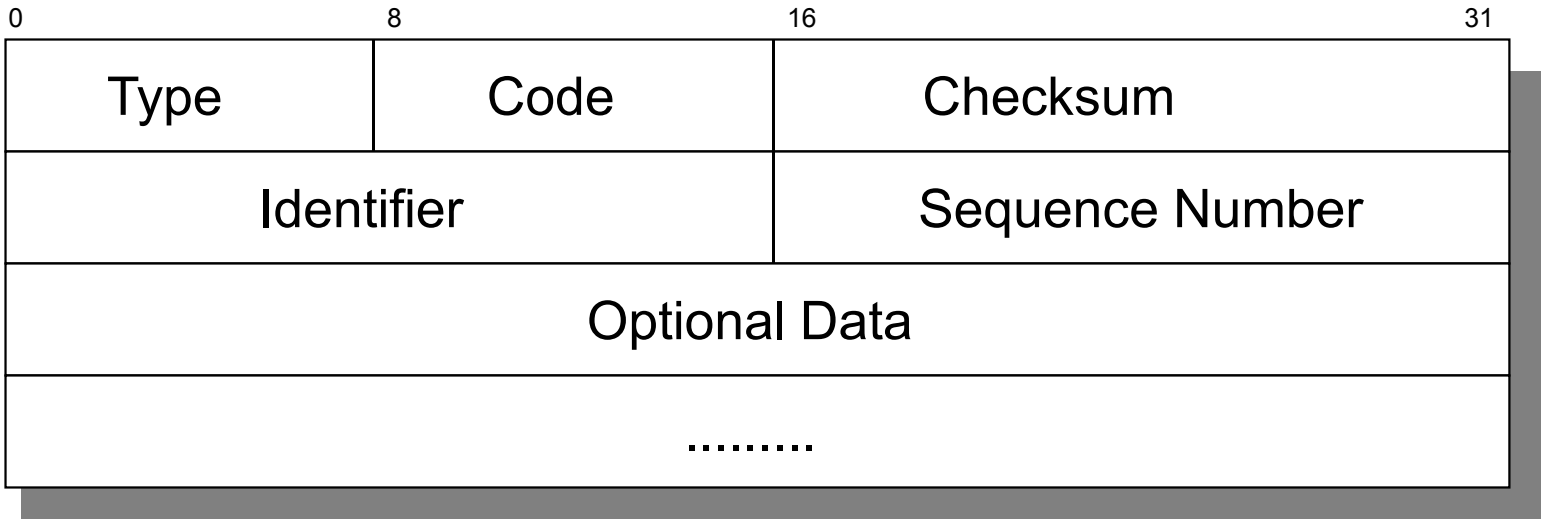
  - Type: tipo da mensagem
    - Echo, Timestamp, Destination Unreachable, ...
  - Code: tipo específico
  - Checksum: da mensagem ICMP

Tipos de Mensagens ICMP

Tipo	Descrição
0	Echo reply
3	Destination unreachable
4	Source quench
5	Redirect
8	Echo request
11	Time exceeded
12	Parameter unintelligible
13	Time-stamp request
14	Time-stamp reply

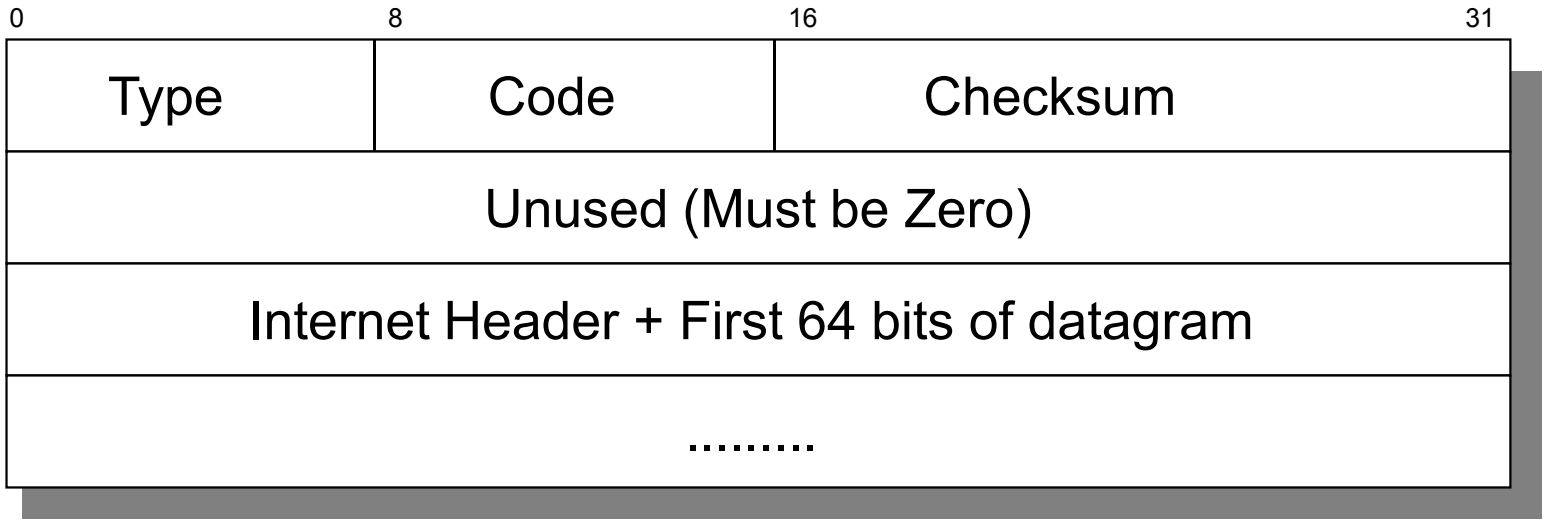
Tipo	Descrição
15	Information request
16	Information reply
17	Address mask request
18	Address mask reply

Echo Request e Echo Reply



- Type:
  - 8 – Echo Request
  - 0 – Echo Reply
- Code: sempre zero (0)

Destination Unreachable

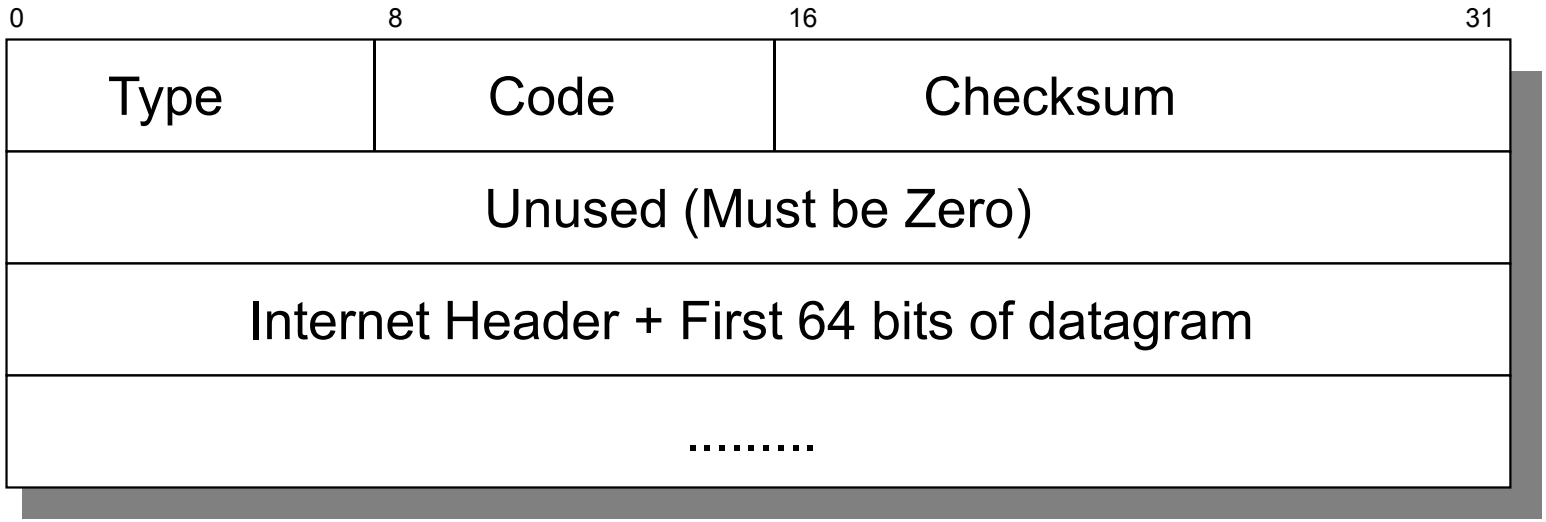


- Type:
  - 3 - Destination Unreachable
- Code:
  - Vários códigos específicos

Destination Unreachable - Códigos

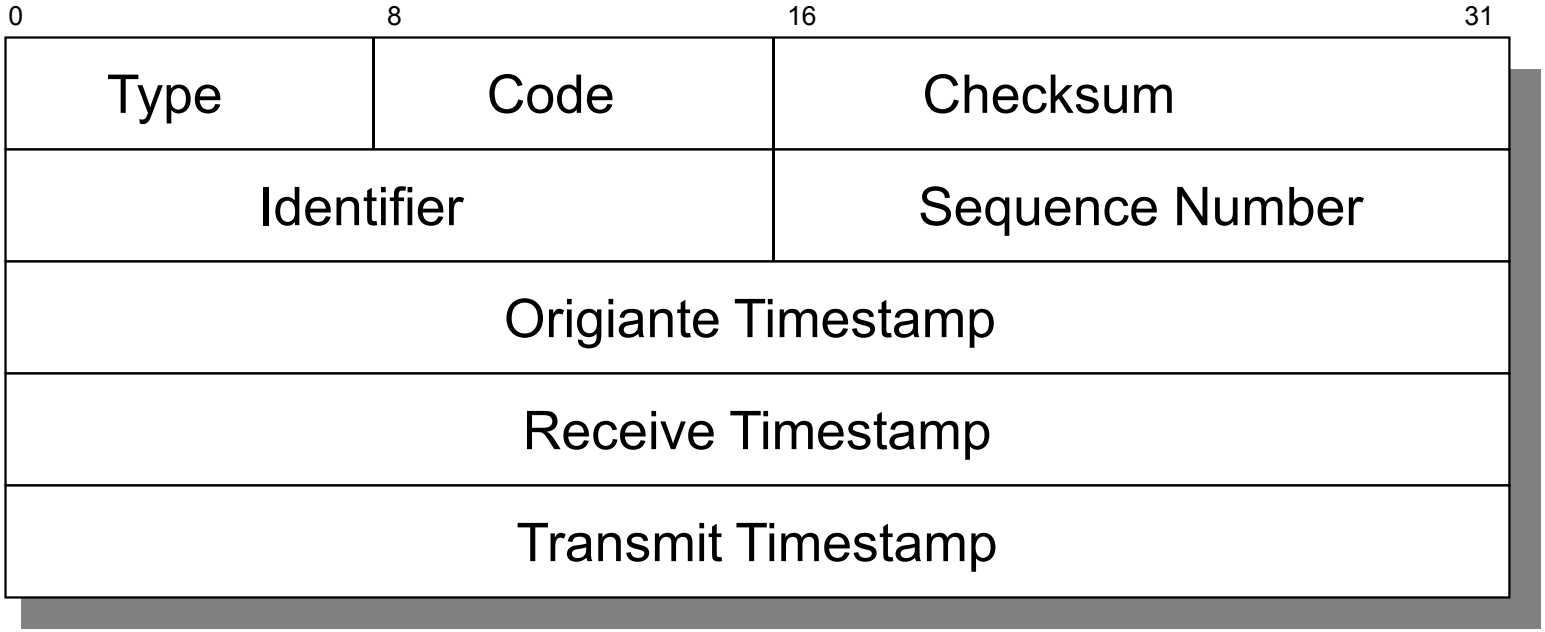
Código	Descrição
0	Network unreachable
1	Host unreachable
2	Protocol unreachable
3	Port unreachable
4	Fragmentation need and don’t fragment bit set
5	Source route failed
6	Destination network unknon
7	Destination host unknown
8	Source host isolated
9	Communication with dest net administratively prohibited
10	Communication with dest host administratively prohibited
11	Network unreachable for type of service
12	Host unreachable for type of service

Time Exceeded



- Type
  - 11 – Time Exceeded
- Código
  - 0 – Time-to_live count exceeded
  - 1 – Fragment reassembly time exceeded

Timestamp



- Type:
  - 13 ou 14 (request/reply)

Outras Mensagens

- Redirect
  - Gerada para correção de rotas
- Address Maks Request/Reply
  - Gerada para solicitação de máscara de subrede
- Source Quench
  - Gerada para controle de fluxo pelo nível 3

IPv6

Cabeçalho IPv6

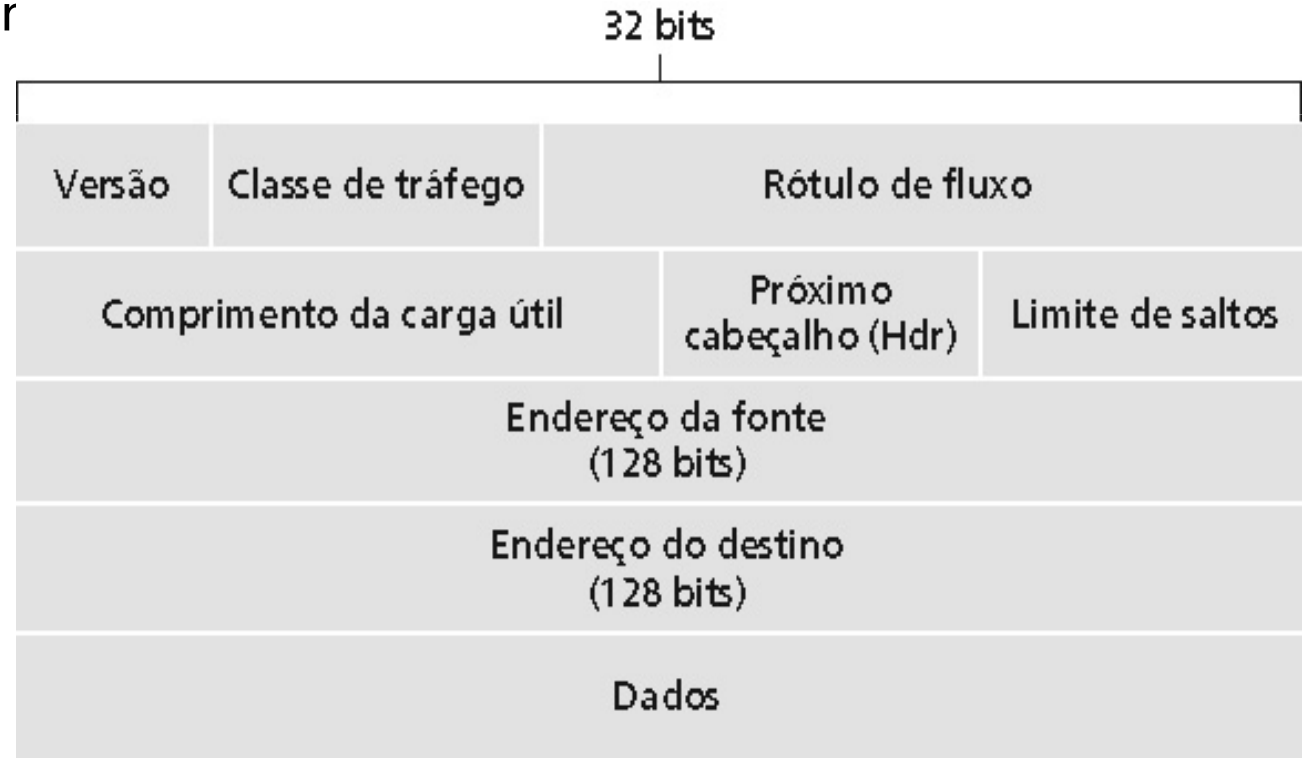
- Motivação inicial: o espaço de endereços de 32 bits está próximo de ser completamente alocado.
- Motivação adicional:
  - Melhorar o formato do cabeçalho para permitir maior velocidade de processamento e de transmissão
  - Mudanças no cabeçalho para incorporar mecanismos de controle de serviço (i.e., *Quality-of-Service*, QoS)
  - Formato do datagrama IPv6:
    - Cabeçalho fixo de 40 Bytes
    - Não é permitida fragmentação (a fragmentação e remontagem tomam muito tempo, retirando essa funcionalidade dos roteadores acelera o repasse de datagramas IP)

Cabeçalho IPv6

Priority: permitir definir prioridades diferenciadas para vários fluxos de informação

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

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



Outras mudanças do IPv4

- Endereços de 128 bits (i.e., 16 Bytes)
- Checksum: removido inteiramente para reduzir o tempo de processamento em cada salto
- Options: são permitidas, mas são alocadas em cabeçalhos suplementares, indicados pelo campo "Next header"
- ICMPv6: nova versão de ICMP
- Tipos de mensagens adicionais , ex.: "Packet Too Big"
- Funções de gerenciamento de grupos multicast

Transição do IPv4 para IPv6

- Nem todos os roteadores poderão ser atualizados simultaneamente
  - Não haverá um dia da vacinação!!!!!
  - Como a rede irá operar com roteadores mistos de IPV4 e IPV6?
  - Tunelamento: IPv6 transportado dentro de pacotes IPv4 entre roteadores IPv4

Tunelamento