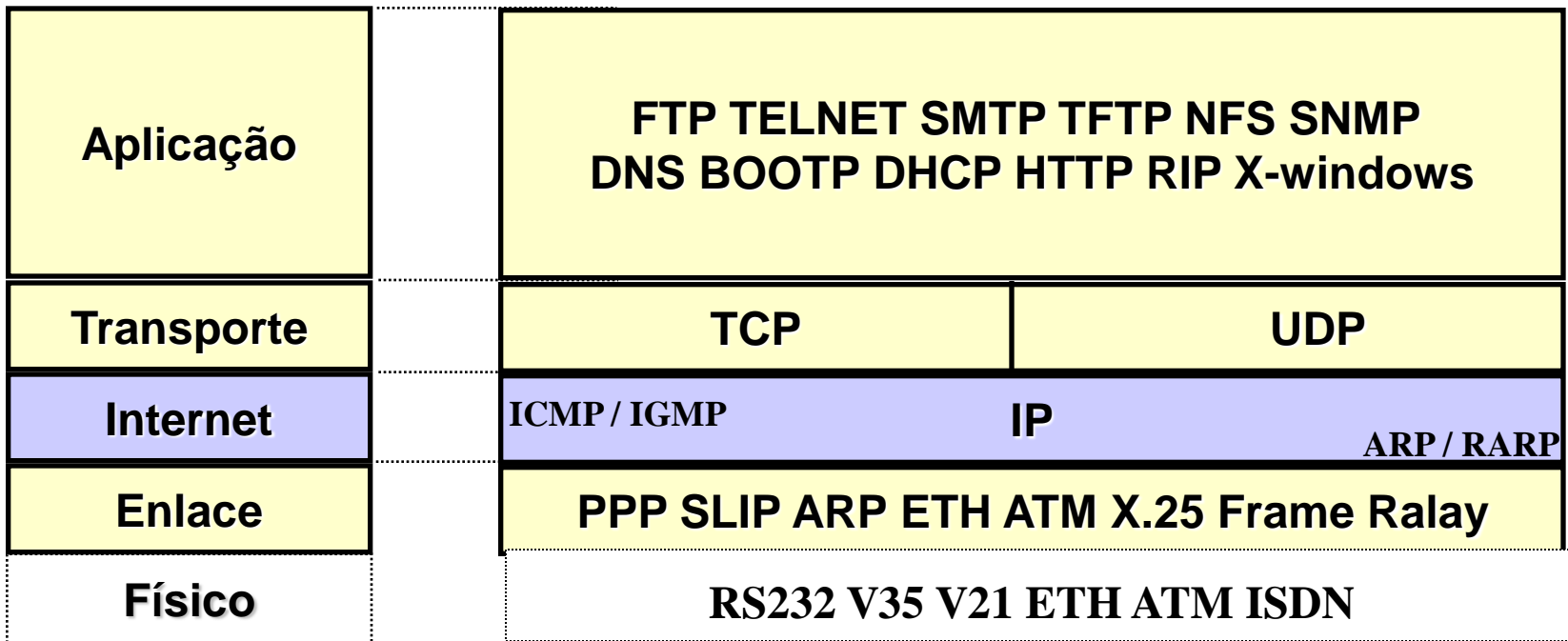


Camada INTERNET da Arquitetura TCP/IP

Alfredo Gama de Carvalho Júnior
agama@embratel.com.br

Natal - Abril 2008

Camada Internet



Camada INTERNET

- Proporciona a conectividade de redes: *Internetworking*
 - Define uma rede virtual única sobre distintas plataformas de hardware
- Protocolos
 - IP - Internet Protocol
 - Move dados entre máquinas da rede, descobrindo as rotas adequadas. Oferece serviço não orientado a conexão a camada de Transporte.
 - ICMP - Internet Control Message Protocol
 - Sinaliza erros do IP e controla seu comportamento
 - IGMP - Internet Group Management Protocol
 - Auxilia o IP gerenciando grupos de multicasting

Camada Internet - Funções

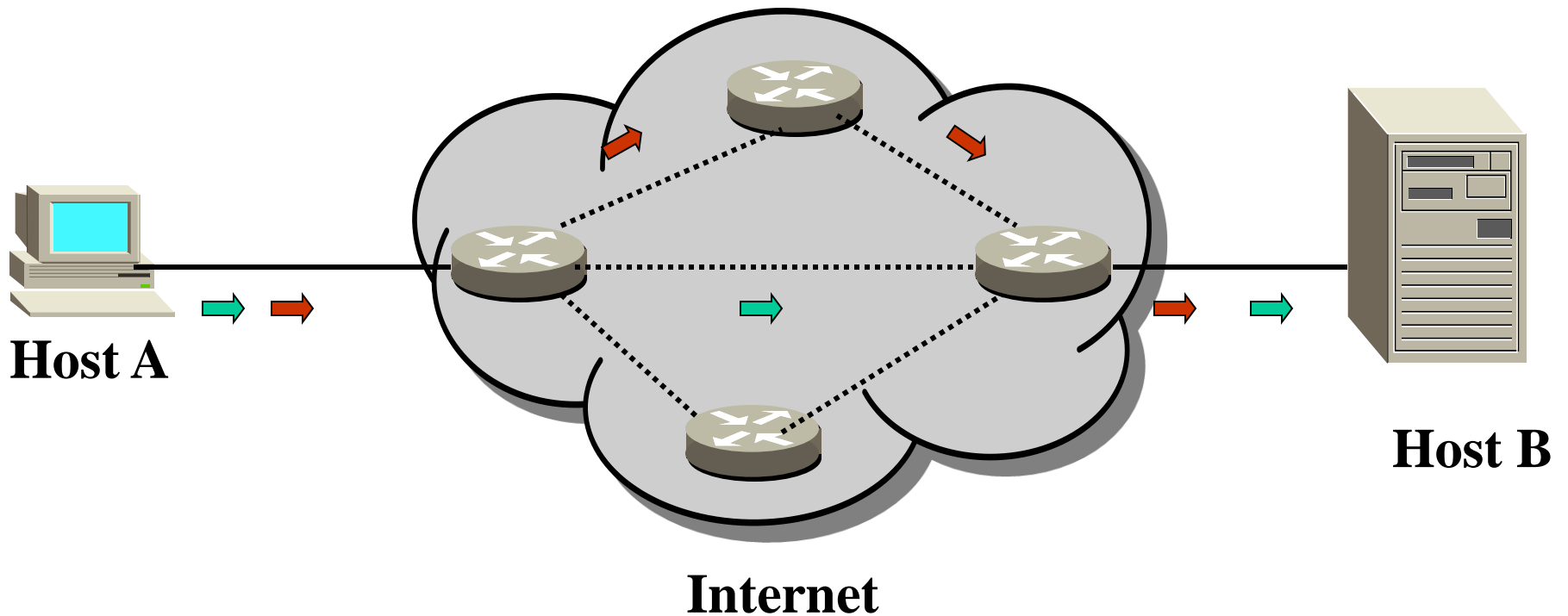
- Definir para que computador e por meio de qual interface, para computador ligado a mais de uma rede (Roteador), as informações oriundas da Camada de Transporte deverão ser encaminhadas.
- Definir o destino das informações recebidas por uma das interfaces de rede (**reenviar** para outra interface, **enviar** para uma das entidades de Transporte ou um simples **descartar**).
- Agrupar ou reagrupar informações em datagramas de tamanho compatível com a tecnologia de rede usada na transmissão
- **Sinalizar para camada de Transporte**, condições de erro verificadas na transmissão dos datagramas

Protocolo IP - Características

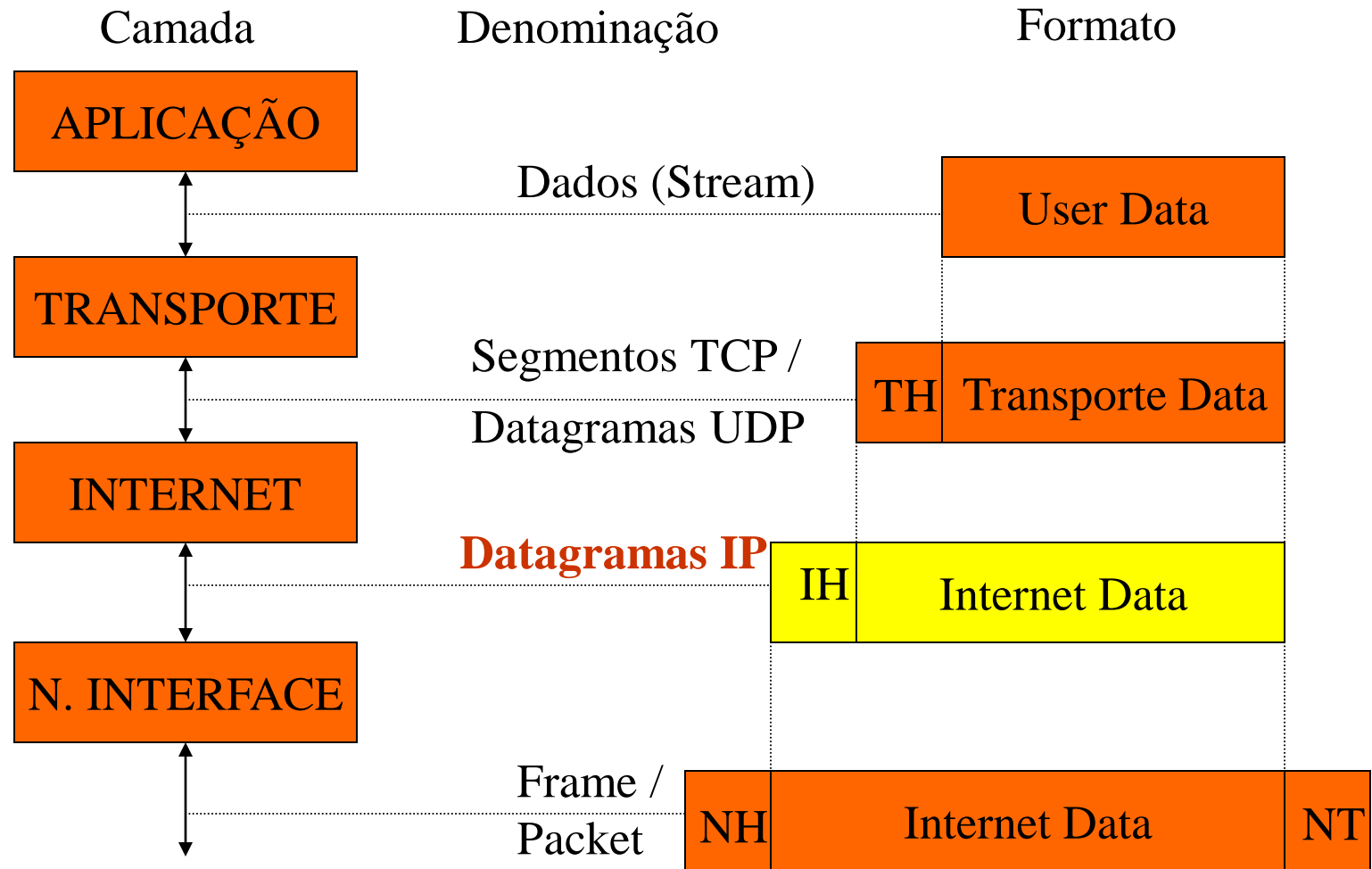
- **Serviço não orientado a conexão**
 - Simplificar Roteadores
 - Comunicação tolerante a falha.
- **Serviço sem reconhecimento de entrega**
- **Serviço sem verificação de integridade dos dados.**
 - Valida apenas o *header*

Pacotes podem se perder, chegar fora de seqüência ou duplicados.

Protocolo IP - Características



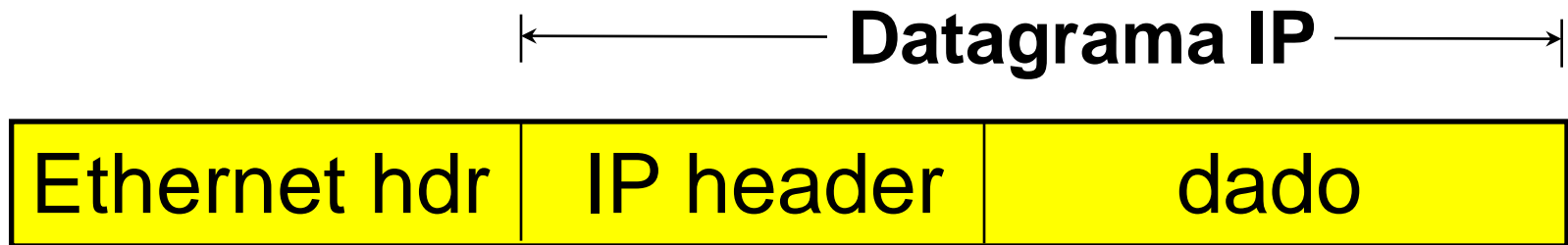
Unidade de Informação - Protocolo IP



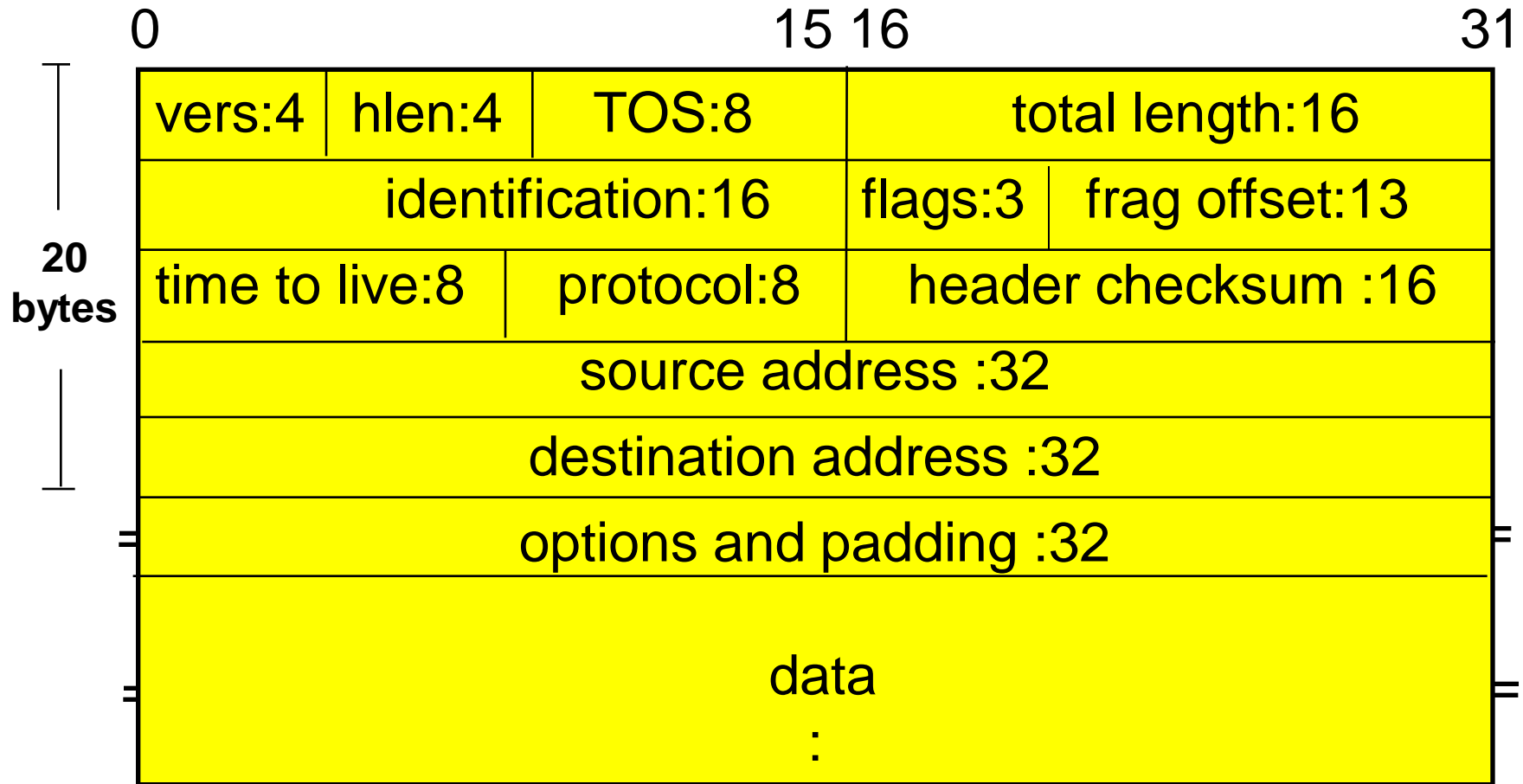
TH-TCP=20 Bytes, TH-UDP=8 Bytes, IH = 20 Bytes e NH varia com tecnologia

Encapsulamento IP

- Dentro do Frame Ethernet

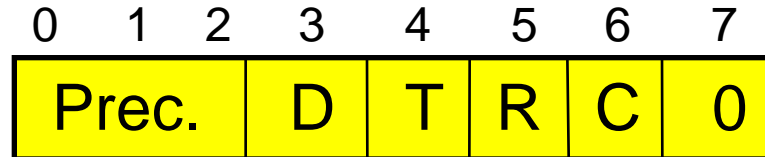


Datagrama IP - Formato



Datagrama IP - Campos

- **Vers** - versão = 4
- **Hlen** - tamanho do header, em unidades de 4 bytes. Normalmente tem valor 5 (20 bytes)
- **TOS** - tipo de serviço, QOS desejada. Sugere o tipo de tratamento que os roteadores devem dar ao datagrama



bits	se 0	se 1
0-2	Precedência (0 a 7 crescente)	
3	Normal delay	Baixo delay
4	Throughput Normal	Alto throughput
5	Confiabilidade Normal	Alta Confiabilidade
6	Custo Normal	Menor custo
7	Reservado	

Datagrama IP - Formato

- **Total length** - Tamanho do datagrama em bytes
 - incluindo header
 - máximo é 64Kbytes (16 bits). Limitado pela rede física
 - Raramente maiores que 1,5 Kbytes. Geralmente 576 bytes
- **Identification** - N° de seqüência **único** gerado no host origem ou roteadores, utilizado pelo destino para associar fragmentos.
- **Flags:**
 - **DF (Don't Frag):** Impede a fragmentação do datagrama
 - **MF(More Fragments):** Indica se datagrama é o 1° ou intermediário resultante de uma fragmentação.
 - **RES.:** Reservado

Datagrama IP - Formato

- **Frag offset** - Informa, em um datagrama que é resultado de uma fragmentação, o deslocamento da informação contida nesse datagrama em relação a informação contida no datagrama original, em unidades de 8 bytes.

Datagrama IP - Formato

- **TTL** - Time to live
 - Número máximo de roteadores que o datagrama pode passar. É decrementado em cada Roteador e quando atinge valor zero é descartado (mensagem *ICMP TTL exceed*).
Normalmente = 30
 - Evita que pacotes “velhos” fiquem, em loop, circulando na rede.
- **Protocol** - Indica qual o protocolo cujas informações estão encapsuladas no campo DATA. Usado no Host destino para entregar dados ao protocolo correspondente
 - 1 = ICMP
 - 6 = TCP
 - 17 = UDP
 - OSPF = 89
 - IGMP = 2

Datagrama IP - Formato

- **Header checksum** - Complemento de um do *header*
 - Não garante integridade dos dados
- **Source address** - Endereço IP Fonte (constante)
- **Destination address** - Endereço IP Destino (constante)
- **PAD** - Permite o enchimento, com zeros, do Header IP para um valor múltiplo de 4 bytes)
- **DATA** - Tam. variável. Irrelevante para o Protocolo IP.

Datagrama IP - Formato

- **Option** - Informações adicionais para testes que auxiliam no monitoramento e controle de interligação em redes (Roteamento e Segurança)
 - Cada opção consiste em um octeto de código que pode ser seguido por um comprimento de octeto único e um conjunto de octetos de dados para tal opção.
 - Ler item 7.8 Comer
 - Não implementado no IPv6 (otimização do processamento)

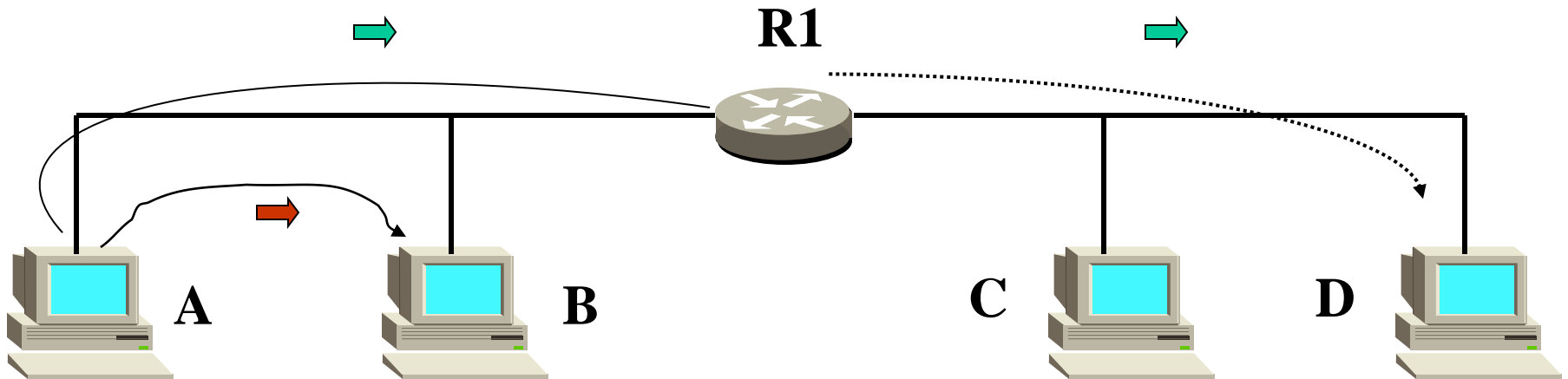
Roteamento IP

- Tarefa executada pelo software IP dos **Hosts** e **Roteadores** para determinar um caminho para enviar os datagramas através da Internet.
- Baseia-se em uma tabela interna ao protocolo IP, presente em **hosts** e **roteadores**, que relaciona endereços IP de destino com endereços IP de equipamentos conectados **diretamente** ao equipamento em que reside a **Routing Table**. Esse par de endereços IP é denominado **route**.

Roteamento IP

- A tabela de um Roteador grava:
 - Que conexão pode ser utilizada para alcançar determinada rede.
 - Alguma indicação de custo e/ou performance
 - Número de HOP

Formas de Roteamento

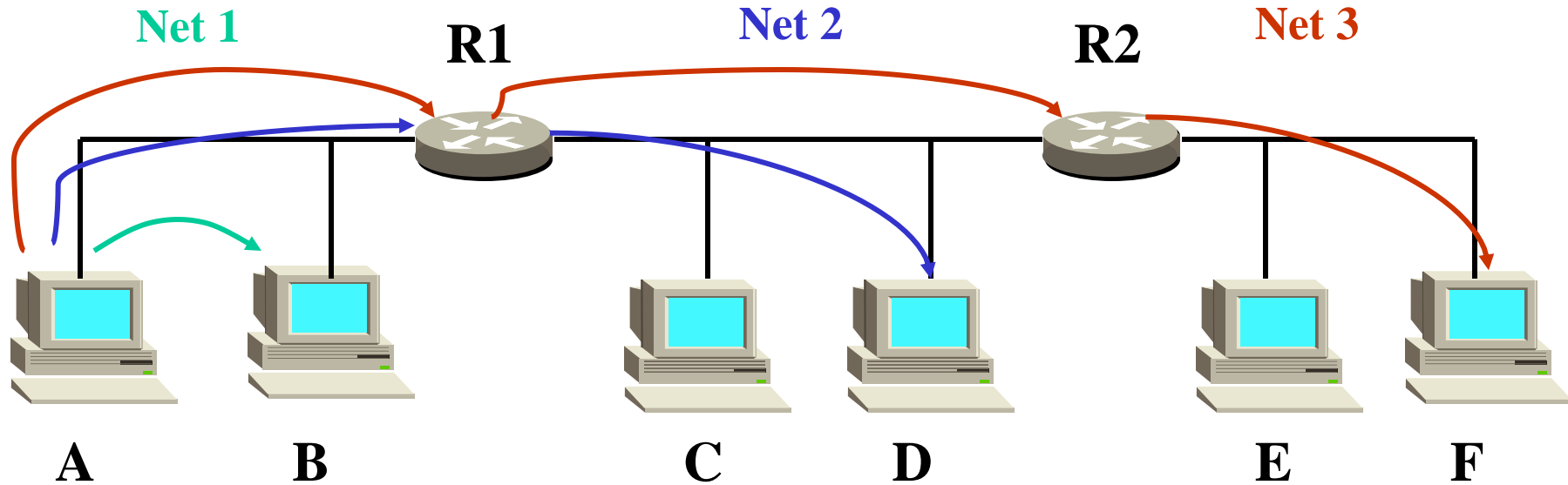


- **Roteamento Direto:** Mesma rede física
Fonte: **A**, Dest: **B**, Próx: **B** (Próx. = Dest)

- **Roteamento Indireto:**

Fonte: **A**, Dest: **D**, Próx: **R1** (Próx. \neq Dest)

Tabelas de Roteamento



HOST A	
Dest	Gateway
B	B *
C	R1
D	R1
E	R1
F	R1

ROT R1	
Dest	Gateway
A	A *
B	B *
C	C *
D	D *
E	R2
F	R2

HOST D	
Dest	Gateway
A	R1
B	R1
C	C *
E	R2
F	R2

ROT R2	
Dest	Gateway
A	R1
B	R1
C	C *
D	D *
E	E *
F	F *

HOST F	
Dest	Gateway
A	R2
B	R2
C	R2
D	R2
E	E *

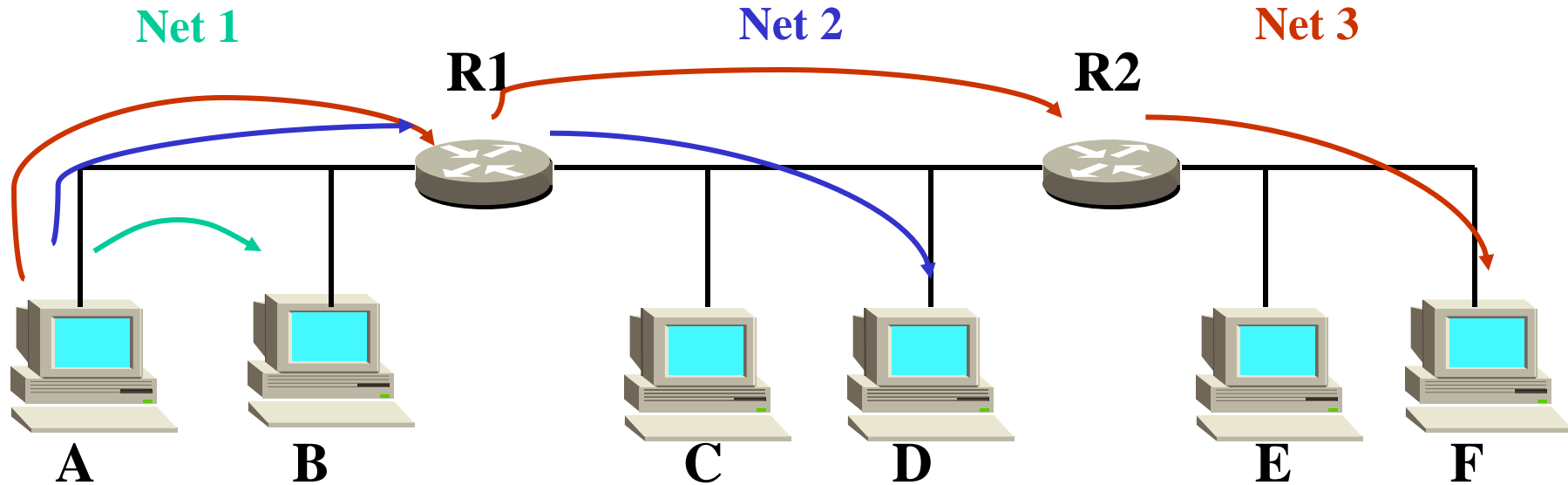
Tabelas de Roteamento

- A última etapa do processo de roteamento é sempre um roteamento direto.
- Na tabela de roteamento não consta a informação completa de como se chega a um destino, apenas o próximo passo para isso.
- O campo do datagrama *Destination IP Address* não se altera independentemente do endereço de entrega obtido na Tabela de Roteamento.

Tabelas de Roteamento

- O endereço de entrega é apenas um parâmetro (variável do SW IP) passado à camada Network Interface e aí traduzido em endereço físico do equip. destino sem alteração dos IPs originais.
- Para evitar enormes Tabelas, função do número de equipamentos de cada rede, todos os equipamentos de uma mesma rede foram agrupados em uma única rota apontada pelo **endereço IP de rede** correspondente.

Tabelas de Roteamento com Endereços de Redes



Dest IP Address AND Subnet mask = End Rede Dest. → Gateway

HOST A	ROT R1	HOST D	ROT R2	HOST F
Dest Gateway	Dest Gateway	Dest Gateway	Dest Gateway	Dest Gateway
N1 DIR	N1 DIR	N1 R1	N1 R1	N1 R2
N2 R1	N2 DIR	N2 DIR	N2 DIR	N2 R2
N3 R1	N3 R2	N3 R2	N3 DIR	N3 DIR

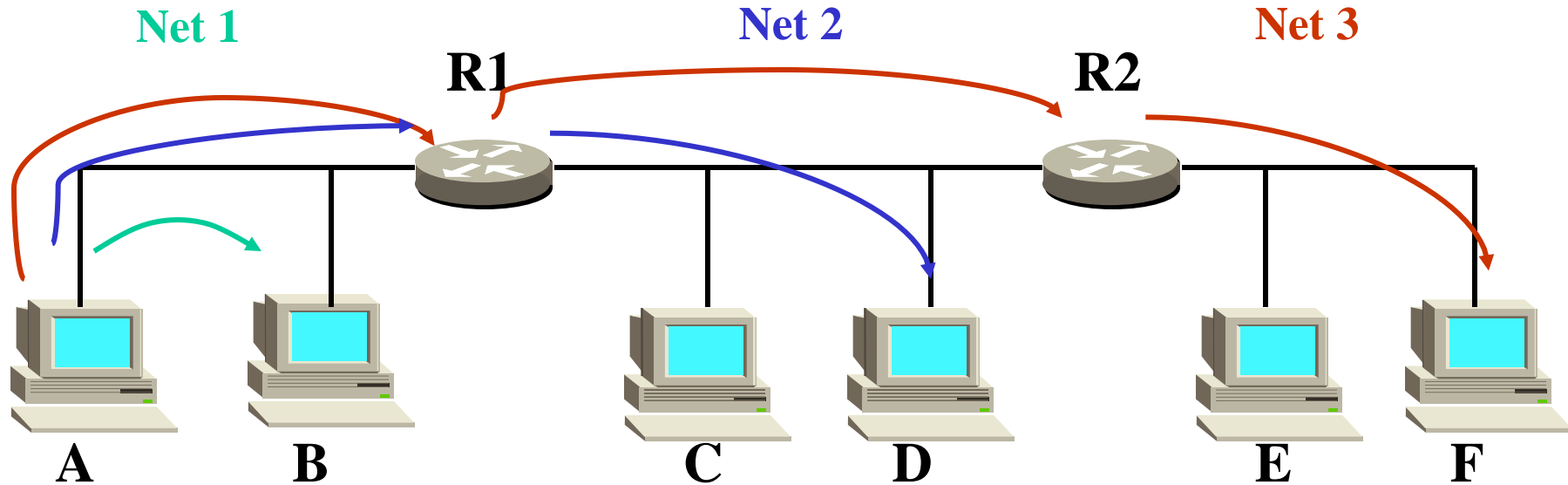
Tabelas de Roteamento com Endereços de Redes

- Host origem monta datagrama com informações recebidas da camada de Transporte e extrai seu endereço destino
- Consulta sua tabela de roteamento, executando AND entre esse endereço e o campo *mask* das entradas, descobrindo assim, o endereço *Destination*. Daí, do campo *Gateway* é retirado o endereço do roteador ou host destino.
- Na tabela de roteamento, a *mask* 255.255.255.255 permite especificação de *host routes*.

Tabelas de Roteamento com Endereços de Redes

- Nas tabelas de roteamento, além dos campos *Destination*, *Mask* e *Gateway* existe o campo *Interface* (roteadores).
- O SW IP pode indicar o roteamento direto de diversas formas (um campo flag ou informando o endereço da própria interface de rede no campo gateway).
- Mesmo agrupando os endereços de uma rede em um única rota não resolvemos o problema do tamanho das Tabelas de Roteamento.
- Será que cada roteador possui rota para todas as redes da Internet?

Tabelas de Roteamento com Default Gateway



HOST A	
Dest	Gateway
N1	DIR
D,E,F	R1

ROT R1	
Dest	Gateway
N1	DIR
N2	DIR
N3	R2

HOST D	
Dest	Gateway
N1	R1
N2	DIR
N3	R2

ROT R2	
Dest	Gateway
N1	R1
N2	DIR
N3	DIR

HOST F	
Dest	Gateway
N1,N2	R2
N3	DIR

Tabelas de Roteamento com Default Gateway

- *Default Gateway* ou *Default Route* é o endereço para o qual os datagramas que não possuam entradas na tabela de roteamento devem ser enviados.
- *Default Gateway* não devem existir para *roteadores de backbones* para evitar loops em caso de endereços inexistentes (Time to live).
- O problema da extensa tabela de rotas destes roteadores, passou a ser contornado a partir de 1992, após uma série de resoluções, cujo resultado é conhecido como *CIDR - Classless Inter-Domain Routing* (abolição do conceito de classes A, B e C).

CIDR - Classless Inter-Domain Routing

- Tornou obrigatório o uso de *bits contíguos* para formação de subnetting
 - Notação convencional: 142.1.0.0 / 255.255.0.0
 - Notação CIDR: 142.1.0.0 / 16
- Permitiu a agregação de redes - *Supernetting*
 - admite distribuição de endereços intermediário entre classe C e B (esgotamento de end. Classe B)
 - admite que redes com endereços contíguos sejam agregadas em uma única rota na tabela de roteamento. Isso implicou em distribuição de endereços observando critério geográfico (1995).

CIDR - Classless Inter-Domain Routing

- Geografia determina associação de endereços IP.
- Mundo dividido em 8 regiões: (RFC1466)

Multi-regional 192.0.0.0 - 193.255.255.255

Europe 194.0.0.0 - 195.255.255.255

Others 196.0.0.0 - 197.255.255.255

North America 198.0.0.0 - 199.255.255.255

Central/South America 200.0.0.0 - 201.255.255.255

Pacific Rim 202.0.0.0 - 203.255.255.255

Others 204.0.0.0 - 205.255.255.255

Others 206.0.0.0 - 207.255.255.255

IANA reserved 208.0.0.0 - 223.255.255.255

CIDR - Supernetting

- Uma organização recebeu um bloco (2^n) contíguo de endereços classe C
- Consegu-se usando bits que pertencem ao endereço de rede como bits de host
- Altera-se a máscara default para classe C, mudando-se alguns bits de 1 para 0.

255	255	252	0
11111111	11111111	11111100	00000000

CIDR - Exemplo de Supernetting

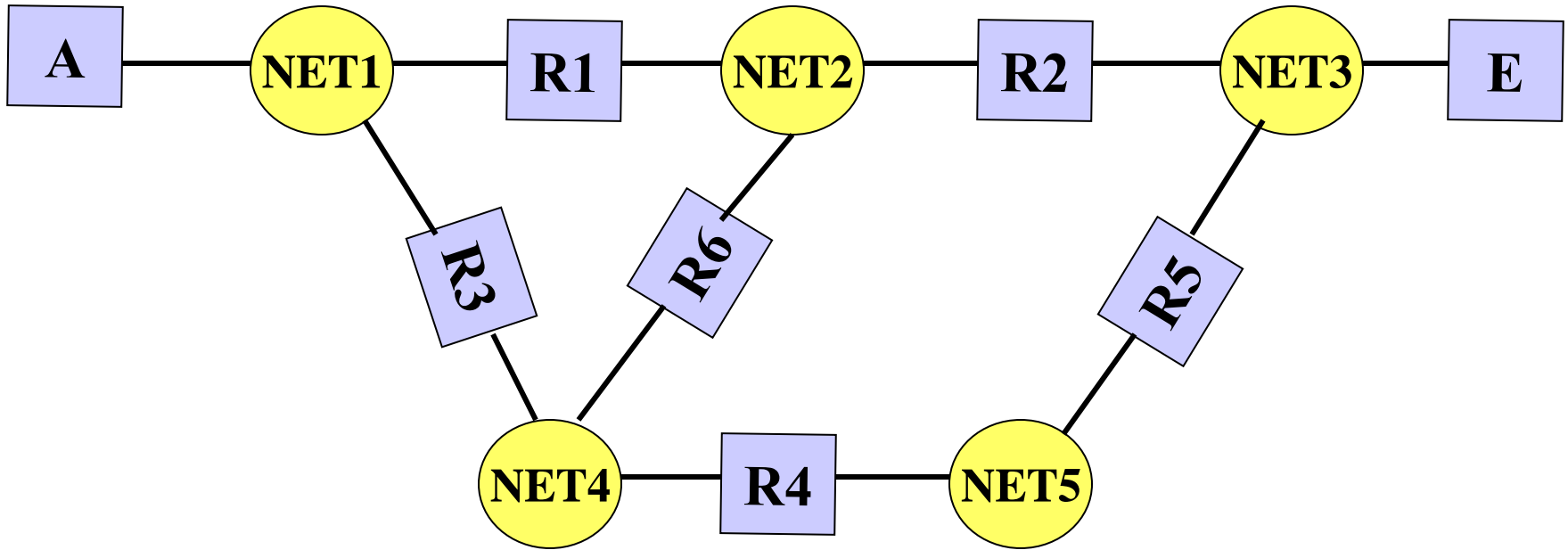
- Uma organização com 4 classes C
 - 194.0.32.0, 194.0.33.0, 194.0.34.0 194.0.35.0

11111111	11111111	11111100	00000000	mask = 255.255.252.0
11000010	00000000	00100000	00000000	net = 194.0.32.0
11000010	00000000	00100001	00000000	net = 194.0.33.0
11000010	00000000	00100010	00000000	net = 194.0.34.0
11000010	00000000	00100011	00000000	net = 194.0.35.0

E bit-a-bit resulta 194.0.32.0

- *A rede desta organização mudou de 4 redes para uma única com 1022 hosts*
- *Qual o endereço de Rede e Broadcast?*

Rotas Múltiplas



A → R1 → R2 → E

A → R1 → R6 → R4 → R5 → E

A → R3 → R4 → R5 → E

A → R3 → R6 → R2 → E

Rotas Múltiplas

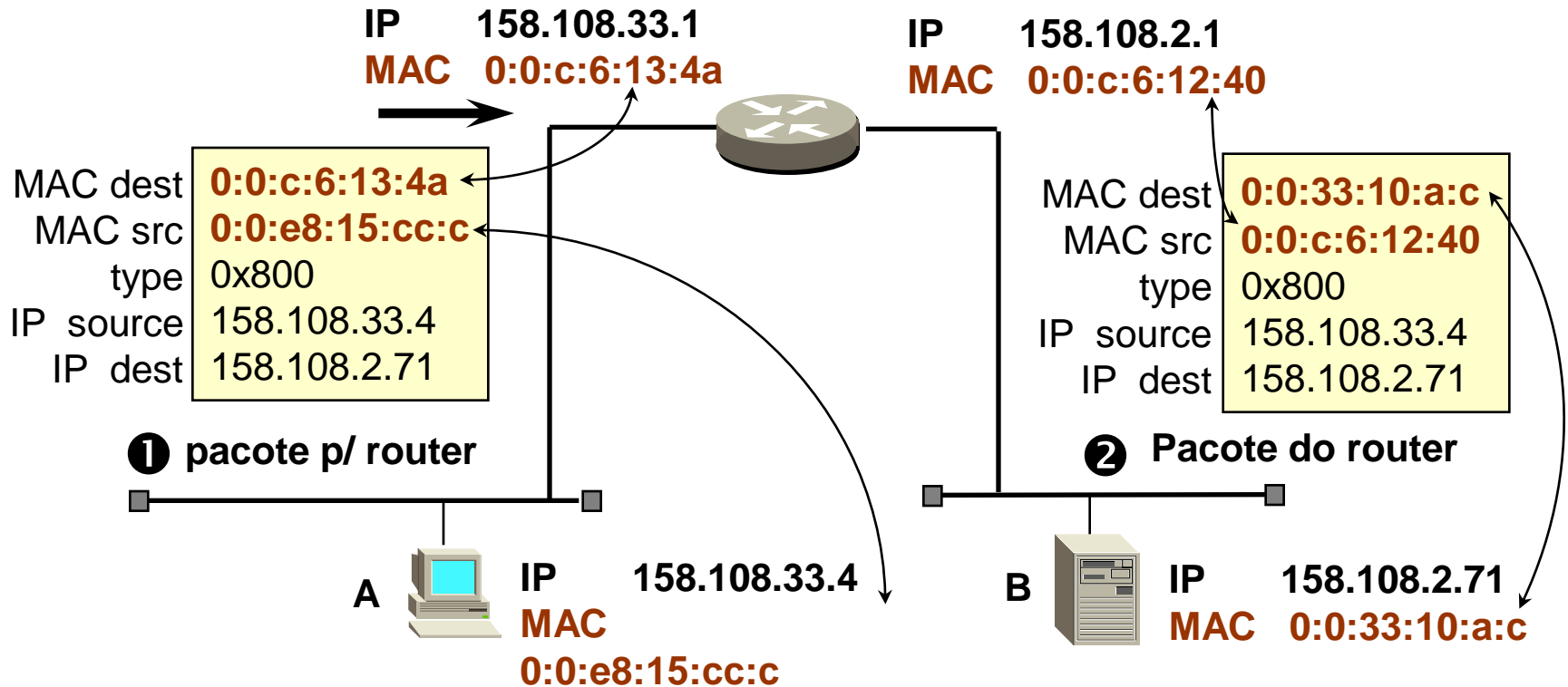
- Critérios para escolha do caminho:

Exigem mais de uma entrada para cada destino.

- Campo *Metric* da tabela de roteamento indica a *distância* (nº de roteadores)
- Função do tráfego de cada rota - *load balanced* (evita sobrecarga de rede ou roteador).
- Função do QoS indicado no campo TOS (Low Delay, High throuput, priorid., confiab. e custo).

Reframing de IP

IP irá reenquadrar o pacote de A enviado a B



Muda MAC Adress, IP permanece o mesmo

Forma da Tabela de Rotas

- **C:\>netstat -r**
- **Route Table**
- **Active Routes:**

Network Address	Netmask	Gateway Address	Interface	Metric
0.0.0.0	0.0.0.0	10.8.13.1	10.8.13.37	1
10.8.13.0	255.255.255.0	10.8.13.37	10.8.13.37	1
10.8.13.37	255.255.255.255	127.0.0.1	127.0.0.1	1
10.255.255.255	255.255.255.255	10.8.13.37	10.8.13.37	1
127.0.0.0	255.0.0.0	127.0.0.1	127.0.0.1	1
224.0.0.0	224.0.0.0	10.8.13.37	10.8.13.37	1
255.255.255.255	255.255.255.255	10.8.13.37	10.8.13.37	1

- **Active Connections**

Proto	Local Address	Foreign Address	State
TCP	pcntl032:1077	PCF616:nbsession	TIME_WAIT

Roteamento IP

- Alguns campos da tabela de roteamento:
 - Destino: Endereço destino (host ou rede)
 - Gateway: Endereço de entrega para atingir esse host.
 - Interface: Interface pela qual se dará a entrega
 - Metric: Custo da rota
 - Flags (opcional): Indicadores de tipo de roteamento (direto, TOS, etc).

Manipulação das Tabelas de Roteamento

- Existem três maneiras:
 - Estática - manualmente em cada equipamento

```
% route add 158.108.20.0 158.108.33.1
```
 - Dinâmica, via *ICMP Redirect* -
 - Roteador destino percebe que o datagrama seria melhor encaminhado por outro roteador da mesma rede e informa ao host que originou o datagrama.
 - Dinâmica, via protocolo de roteamento
 - Protocolos responsáveis pela troca de informações de roteamento entre os equipamentos de uma internet (RIP, OSPF, BGP, EGP, HELLO).

Manipulação das Tabelas de Roteamento

- Na prática, o processo é misto.
 - No boot, os Hosts e Roteadores montam suas entradas nas tabelas para as redes conectadas diretamente.
 - Após isso, pode ser ativada uma Aplicação de roteamento (um ou mais protocolos) que permita sua complementação e manutenção (muitos equipamentos)
 - Para pequenas redes, utiliza-se rotas estáticas
 - O processo através da mensagem ICMP Redirect é pouco usado face problema de segurança quanto a facilidade de roteamento não autorizado

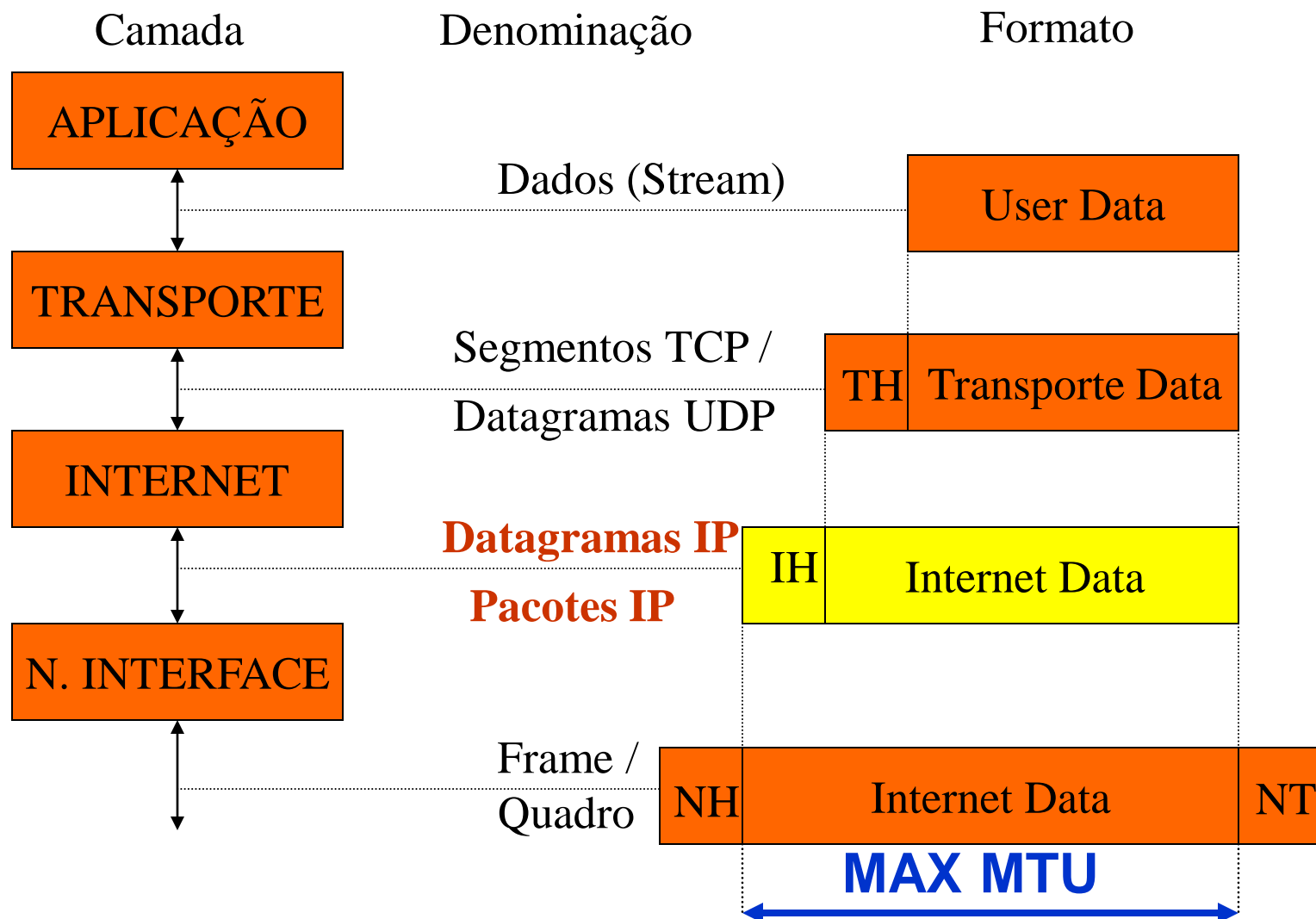
Manipulação das Tabelas de Roteamento

- Considerações:
 - Os hosts só executam roteamento de datagramas por eles gerados (não gera tráfego p/ datagramas recebidos não destinados a ele)
 - O Sw IP utiliza mensagens ICMP para informar ao emissor do datagrama condições de anormalidade.
 - Um datagrama broadcast proveniente de um site fora do ambiente da sub-rede só tem sentido se os roteadores que conectam estas sub-redes concordarem em difundi-lo.

Fragmentação e Remontagem

- O limite superior de dados que podem ser enviado em um tecnologia de enlace (Network Interface) é definido como **MTU** (*Maximum Transfer Unit*)
- MTU (em bytes):
 - FDDI : 4470
 - Ethernet : 1500
 - 802.3 : 1492
 - PPP: 296
- Caso datagrama IP $>$ MTU
 - IP quebra o datagrama em blocos menores - *Fragmentação*
 - Processo inverso - *Remontagem*

MTU - Maximum Transmission Unit



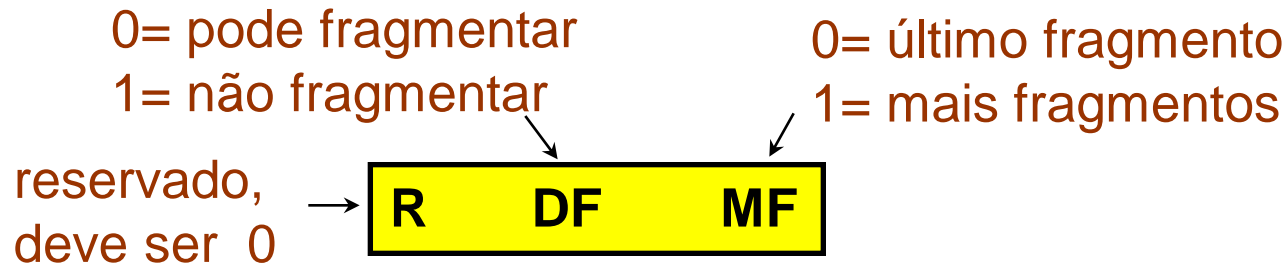
TH-TCP=20 Bytes, TH-UDP=8 Bytes, IH = 20 Bytes e NH varia com tecnologia

Fragmentação e Remontagem

- Fragmentação
 - Respeita limite da interface (MTU)
 - Gera datagramas normais, independentes
- Remontagem
 - No equipamento destino (custo alto)
 - Cada fragmento tem seu próprio *header*
 - Cada fragmento tem um identificador comum (16-bits)
 - Fragmentação acontece em blocos múltiplos de 8 bytes

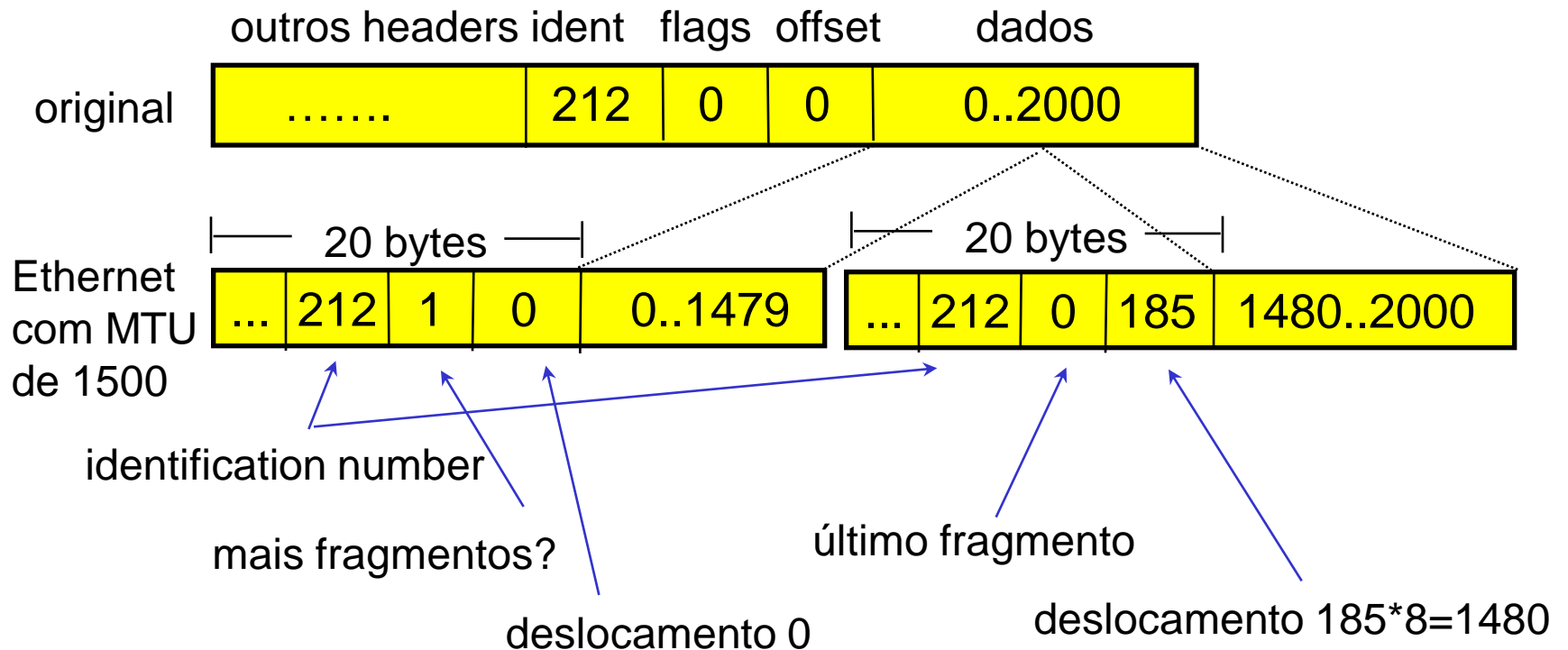
Implementação da Fragmentação

- Número de identificação
 - 16 bits comuns a todos os fragmentos
 - Não é um número sequencial!
- flags - 3 bits de controle



- **fragment offset** - Deslocamento do fragmento em relação ao início do datagrama original (em blocos múltiplos de 8 BYTES)

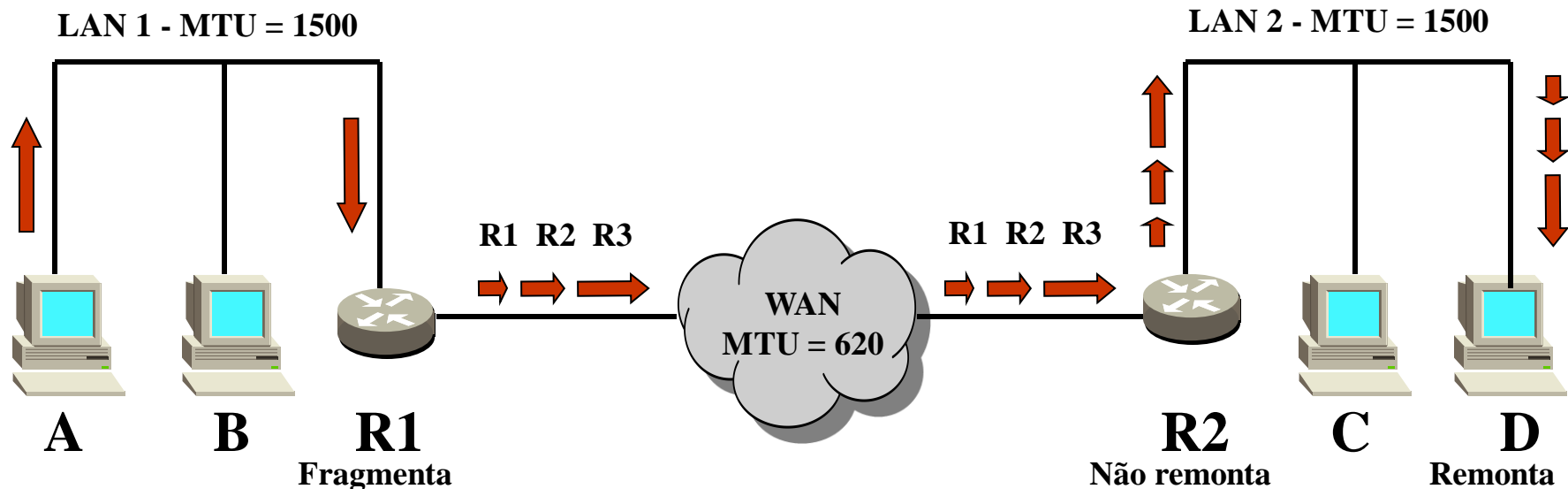
Amostra de Fragmentação



Problemas da Fragmentação

- O nó final não sabe o número de fragmentos. Gasto de memória não previsível na remontagem.
- Se um fragmento se perde, os demais são descartados.
- O destino inicia um *timer* quando recebe o primeiro fragmento, se há muita demora (30 seg), todos os demais são descartados
- IP não recupera erros ...
 - Mas faz notificação ICMP

Fragmentação e Remontagem



Campo	D0	D1	D2	D3
Total Length	1420	620	620	220
Length Identification	20 (ex)	20	20	20
Frag Offset	0	0	75	150
MF	0	1	1	0
DF	0	0	0	0

Processamento de dados recebidos da camada de Transporte

```
envia dados (dados, destino) {  
    consulta roteamento para obter endereço de entrega;  
    if (endereço destino é roteável) {  
        if (datagrama é maior do que o MTU da interface) {  
            monta datagramas-fragmento contendo dados;  
        } else {  
            monta datagrama único contendo dados;  
        }  
        envia o(s) datagrama(s) e o endereço de entrega para  
        interfaces definidas pelo roteamento;  
        sinaliza OK ao software da camada Transport;  
    } else {  
        sinaliza erro ao software da camada Transport;  
    }  
}
```

Processamento de datagramas recebidos da camada Interface de Rede

```
Recebe_datagrama (datagrama) {  
    if (checksum não confere) { descarta datagrama e encerra processamento;}  
    obtém endereço destino da datagrama;  
    if (datagrama é destinado a este equipamento) {  
        if (é fragmento de um datagrama maior) { aguarda demais fragmentos;}  
        envia datagrama completo para protocolo correspondente (PROTO);  
        encerra processo;  
    }  
    if (equipamento é um host) { descarta datagrama e encerra processamento }  
    consulta roteamento para obter endereço de entrega;  
    if (destino do datagrama não é roteável) { descarta datagrama e envia  
        mensagem de erro (ICMP) ao emissor; encerra processamento;}  
    if (tempo de vida do datagrama expirou (TTL -1 = 0) { descarta datagrama e envia mensagem  
        de erro (ICMP) ao emissor; encerra processamento;}  
    if (nova interface requer fragmentação) { fragmenta o datagrama;}  
    envia o(s) datagrama(s) e o endereço de entrega para a interface definida pelo roteamento;  
    encerra o processamento;  
}
```