

Capítulo 7

Redes multimídia

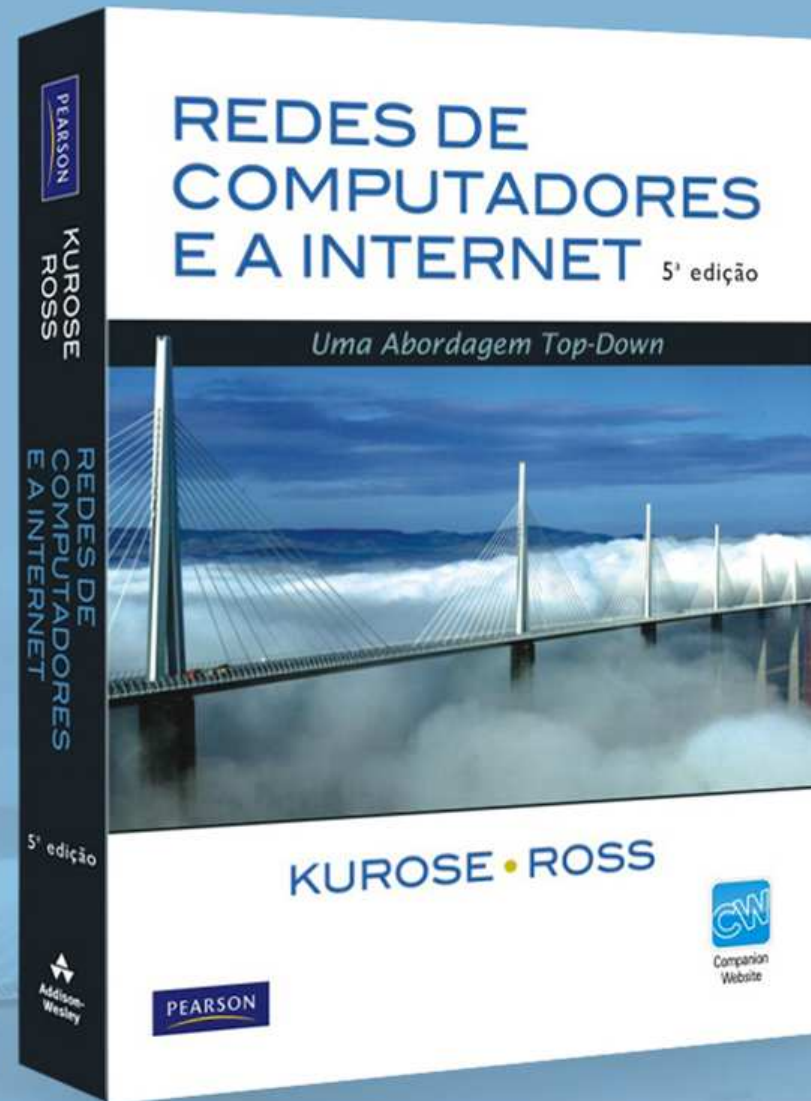
Nota sobre o uso destes slides ppt:

Estamos disponibilizando estes slides gratuitamente a todos (professores, alunos, leitores). Eles estão em formato do PowerPoint para que você possa incluir, modificar e excluir slides (incluindo este) e o conteúdo do slide, de acordo com suas necessidades. Eles obviamente representam *muito* trabalho da nossa parte. Em retorno pelo uso, pedimos apenas o seguinte:

- ❑ Se você usar estes slides (por exemplo, em sala de aula) sem muita alteração, que mencione sua fonte (afinal, gostamos que as pessoas usem nosso livro!).
- ❑ Se você postar quaisquer slides sem muita alteração em um site Web, que informe que eles foram adaptados dos (ou talvez idênticos aos) nossos slides, e inclua nossa nota de direito autoral desse material.

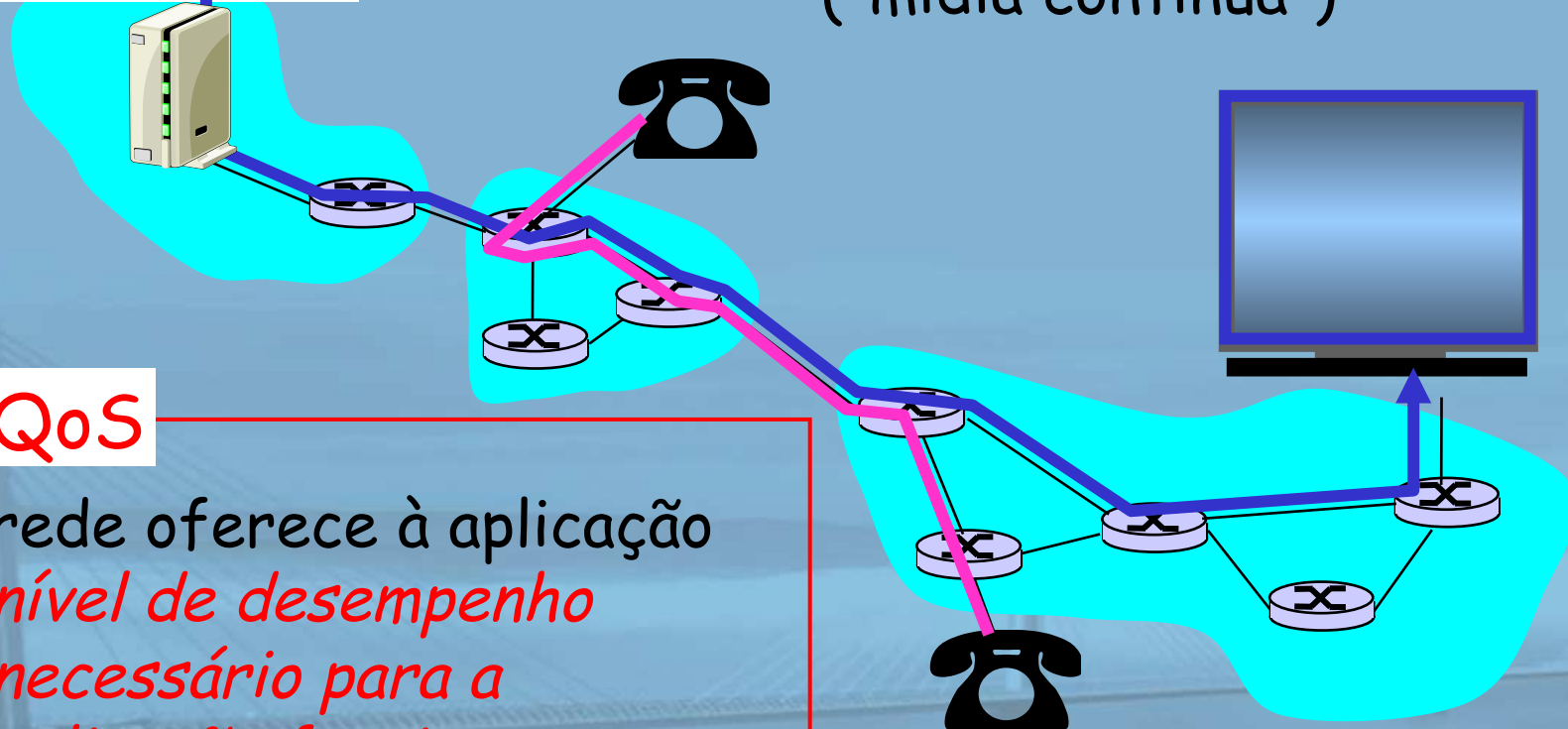
Obrigado e divirta-se! JFK/KWR

Todo o material copyright 1996-2009
J. F. Kurose e K. W. Ross, Todos os direitos reservados.



Multimídia e qualidade de serviços: o que é?

aplicações de multimídia:
áudio e vídeo de rede
("mídia contínua")



QoS

rede oferece à aplicação
nível de desempenho
necessário para a
aplicação funcionar.

Capítulo 7: Objetivos

Princípios

- ❑ classificar aplicações de multimídia
- ❑ identificar serviços de rede que as aplicações precisam usar
- ❑ fazer o melhor com o serviço de melhor esforço

Protocolos e arquiteturas

- ❑ protocolos específicos para melhor esforço
- ❑ mecanismos para fornecer QoS
- ❑ arquiteturas para QoS

Capítulo 7: Esboço

7.1 Aplicações de rede
multimídia

7.2 Áudio e vídeo de fluxo
contínuo armazenados

7.3 Fazendo o melhor
possível com o serviço de
melhor esforço

7.4 Protocolos para
aplicações interativas em
tempo real - RTP, RTCP,
SIP

7.5 Fornecendo classes
de serviço múltiplas

7.6 Fornecendo
garantias de
qualidade de serviços

Aplicações de rede multimídia (MM)

Classes de aplicações MM:

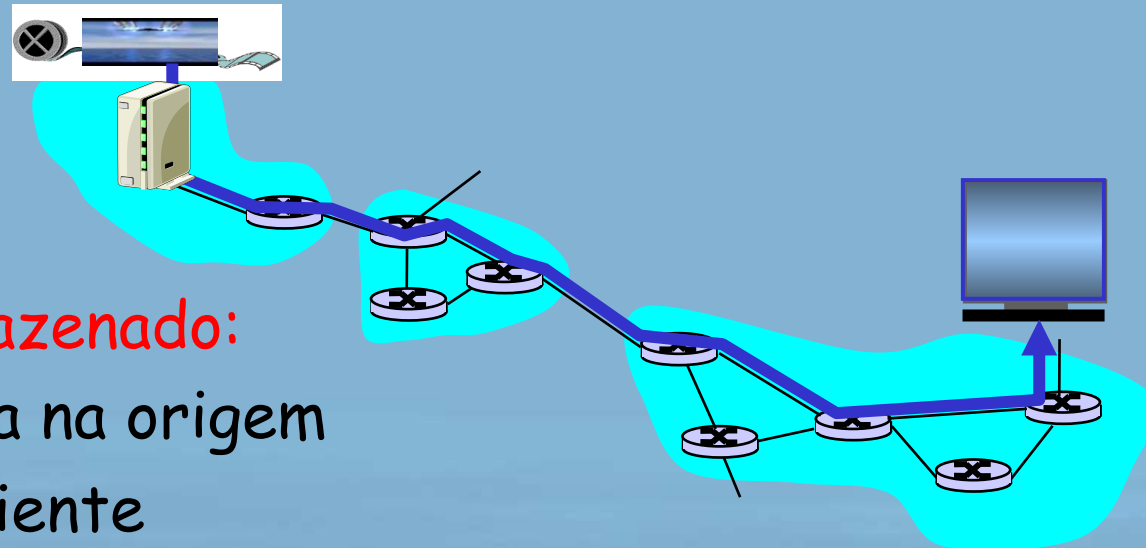
1. fluxo contínuo (*streaming*) armazenado
2. fluxo contínuo ao vivo
3. interativas, tempo real

Jitter é a variabilidade dos atrasos de pacote dentro do mesmo fluxo de pacotes

Características fundamentais:

- ❑ normalmente, **sensível ao atraso**
 - atraso fim a fim
 - jitter do atraso
- ❑ **tolerante a perdas:** perdas infrequentes causam pequenas falhas
- ❑ antítese de dados, que são *intolerantes* a falhas, mas *tolerantes* a atraso.

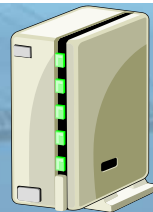
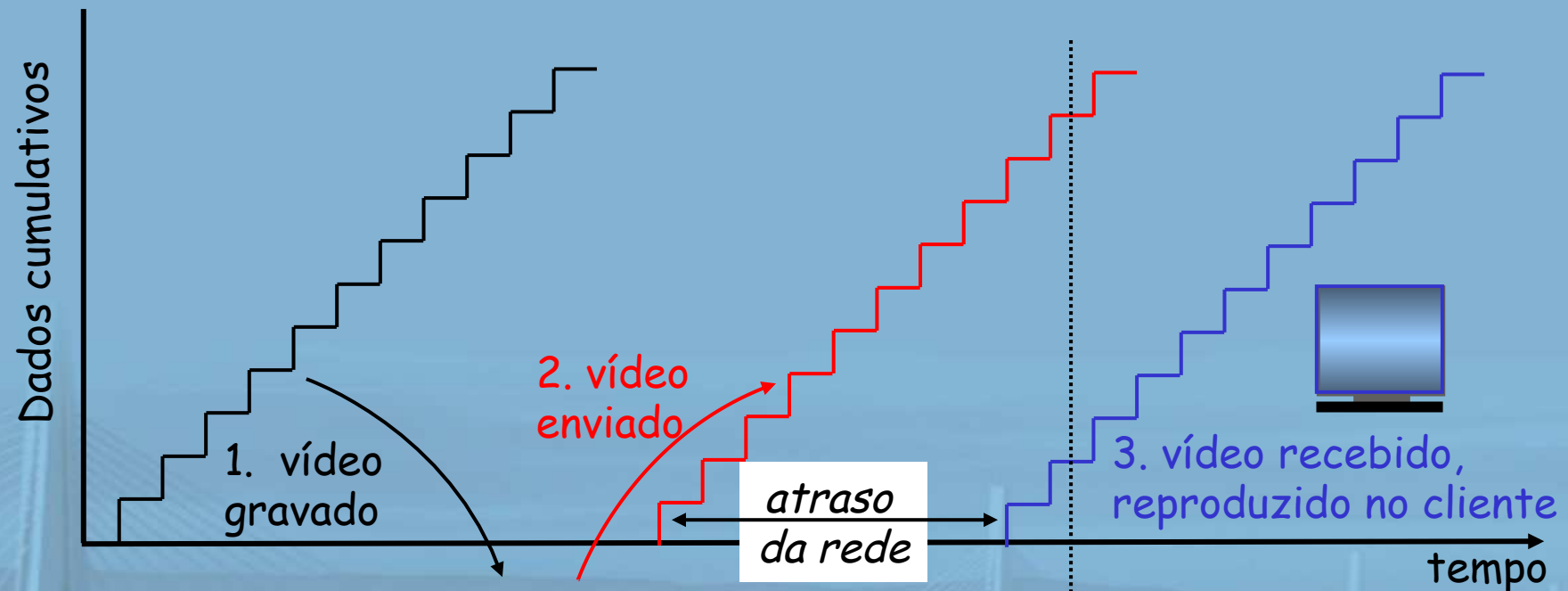
Multimídia armazenada de fluxo contínuo



Fluxo contínuo armazenado:

- ❑ mídia armazenada na origem
- ❑ transmitida ao cliente
- ❑ fluxo contínuo: reprodução do cliente começa *antes* que todos os dados tenham chegado
- ❑ restrição de tempo para dados ainda a serem transmitidos: a tempo para o reprodução

Multimídia armazenado de fluxo contínuo: o que é?



fluxo contínuo: neste momento, cliente reproduzindo parte inicial do vídeo, enquanto servidor ainda envia parte posterior do vídeo

Multimídia *Armazenado* de fluxo contínuo: interatividade



- ❑ *funcionalidade tipo VCR*: cliente pode dar pausa, voltar, avançar, pressionar barra deslizante
 - 10 seg de atraso inicial OK
 - 1-2 seg até efeito do comando OK
- ❑ restrição de tempo para dados ainda a serem transmitidos: em tempo para reprodução

Multimídia *ao vivo* em fluxo contínuo

Exemplos:

- ❑ programa de entrevistas por rádio da Internet
- ❑ evento esportivo ao vivo

Fluxo contínuo (como na multimídia *armazenada* em fluxo contínuo)

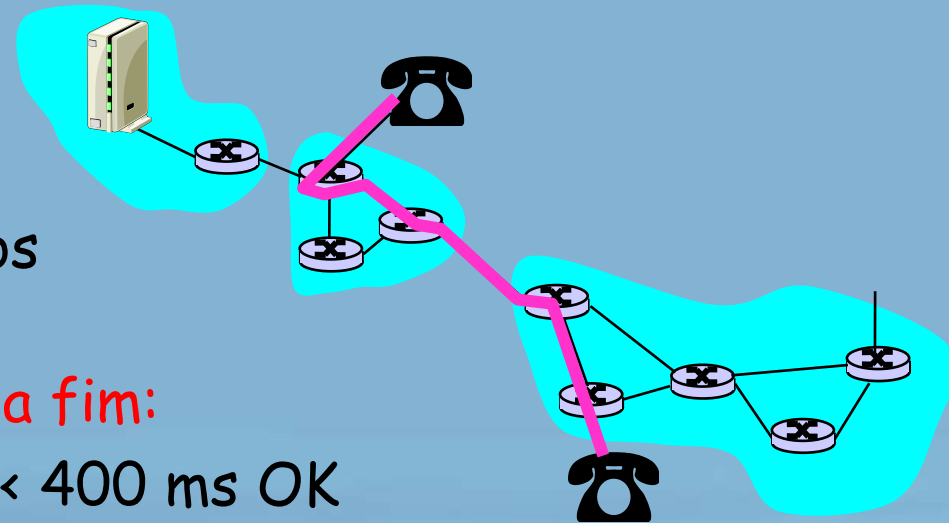
- ❑ buffer de reprodução
- ❑ reprodução pode atrasar dezenas de segundos após a transmissão
- ❑ ainda tem restrição de tempo

Interatividade

- ❑ avanço rápido impossível
- ❑ retornar, pausar possíveis!

Multimídia interativa em tempo real

- ❑ **aplicações:** telefonia IP, videoconferência, mundos interativos distribuídos
- ❑ **requisitos de atraso fim a fim:**
 - áudio: < 150 ms bom, < 400 ms OK
 - inclui atrasos em nível de aplicação (empacotamento) e de rede
 - atrasos maiores observáveis prejudicam interatividade
- ❑ **inicialização da sessão**
 - Como o destino anuncia seu endereço IP, número de porta, algoritmos de codificação?



Multimídia sobre a Internet de hoje

TCP/UDP/IP: "serviço de melhor esforço"

- *sem* garantia sobre atraso e perda



? ? ? ? ? ?
Mas você disse que as aplicações de multimídia ?
exigem que QoS e nível de desempenho
? sejam eficazes! ? ?



Aplicações de multimídia na Internet de hoje usam técnicas em nível de aplicação para aliviar (ao máximo) os efeitos de atraso e perda.

Como a Internet deverá evoluir para dar melhor suporte à multimídia?

Filosofia de serviços integrados:

- ❑ mudanças fundamentais na Internet para as aplicações reservarem largura de banda fim a fim
- ❑ requer software novo, complexo nos hospedeiros e roteadores

Laissez-faire

- ❑ sem mudanças importantes
- ❑ mais largura de banda quando necessário
- ❑ distribuição de conteúdo, multicast da camada de aplicação
 - camada de aplicação

Filosofia de serviços diferenciados:

- ❑ menos mudanças na infraestrutura da Internet, oferecendo serviço de 1ª e 2ª classes



Qual é a sua opinião?

Algumas palavras sobre compactação de áudio

- ❑ amostra de sinal analógico
 - telefone: 8.000 amostras/s
 - música de CD: 44.100 amostras/s
- ❑ *"Quantization is the process of converting a continuous analog audio signal to a digital signal"*
- ❑ cada amostra quantizada, ou seja, arredondada
 - p. e., $2^8 = 256$ valores quantizados possíveis
- ❑ cada valor quantizado representado por bits
 - 8 bits para 256 valores
- ❑ exemplo: 8.000 amostras/s, 256 valores quantizados --> 64.000 bps
- ❑ receptor converte bits para sinal analógico:
 - alguma redução de qualidade

Exemplos de taxas

- ❑ CD: 1.411 Mbps
- ❑ MP3: 96, 128, 160 kbps
- ❑ Telefonia da Internet: 5,3 kbps em diante

Algumas palavras sobre compactação de vídeo

- ❑ vídeo: sequência de imagens exibidas em taxa constante
 - p. e. 24 imagens/s
- ❑ imagem digital: array de pixels
 - cada pixel representado por bits
- ❑ redundância
 - espacial (dentro da imagem)
 - temporal (de uma imagem para a seguinte)

Exemplos:

- ❑ MPEG 1 (CD-ROM) 1,5 Mbps
- ❑ MPEG2 (DVD) 3-6 Mbps
- ❑ MPEG4 (normalmente usado na Internet, < 1 Mbps)

Pesquisa:

- ❑ vídeo em camadas (escalável)
 - adapta camadas à largura de banda disponível

Capítulo 7: Esboço

- ❑ 7.1 Aplicações de rede multimídia
- ❑ 7.2 **Áudio e vídeo de fluxo contínuo armazenados**
- ❑ 7.3 Fazendo o melhor possível com o serviço de melhor esforço
- ❑ 7.4 Protocolos para aplicações interativas em tempo real - RTP, RTCP, SIP
- ❑ 7.5 Fornecendo classes de serviço múltiplas
- ❑ 7.6 Fornecendo garantias de qualidade de serviços

Multimídia armazenada de fluxo contínuo

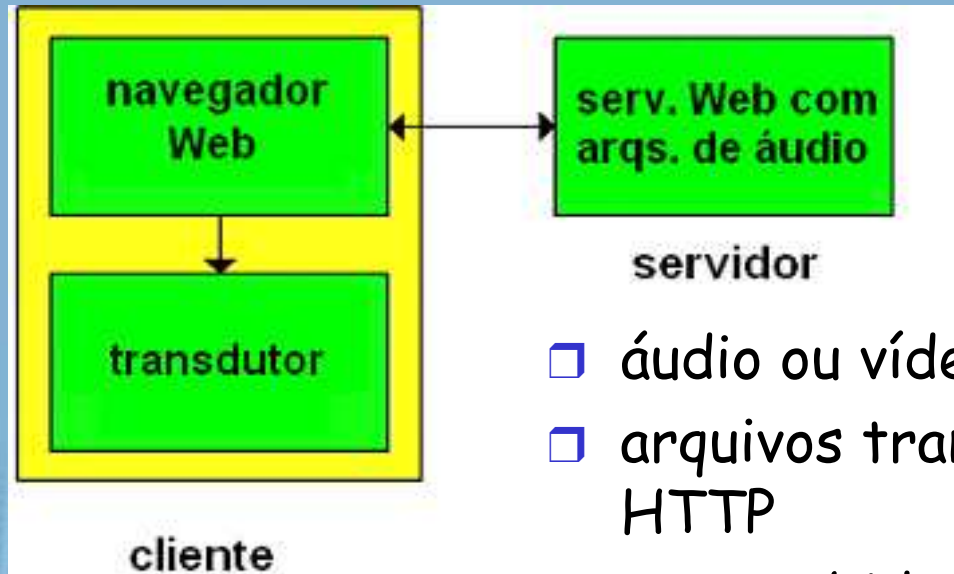
técnicas de fluxo contínuo em nível de aplicação para obter o máximo do serviço de melhor esforço:

- buffering no cliente
- uso de UDP *versus* TCP
- múltiplas codificações de multimídia

Media Player

- ❑ eliminação da variação de atraso (jitter)
- ❑ descompressão
- ❑ supressão de erro
- ❑ interface gráfica de usuário sem controles para interatividade

Multimídia na Internet: técnica mais simples

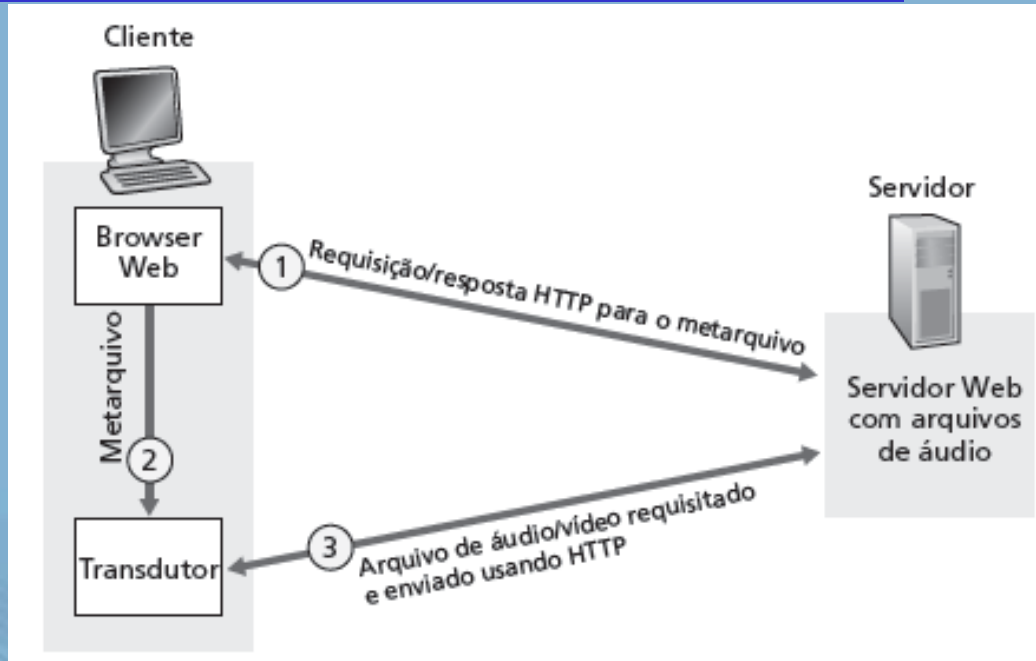


- ❑ áudio ou vídeo armazenados em arquivo
- ❑ arquivos transferidos como objetos HTTP
 - recebidos por inteiro no cliente
 - depois passados ao transdutor

áudio, vídeo sem fluxo contínuo:

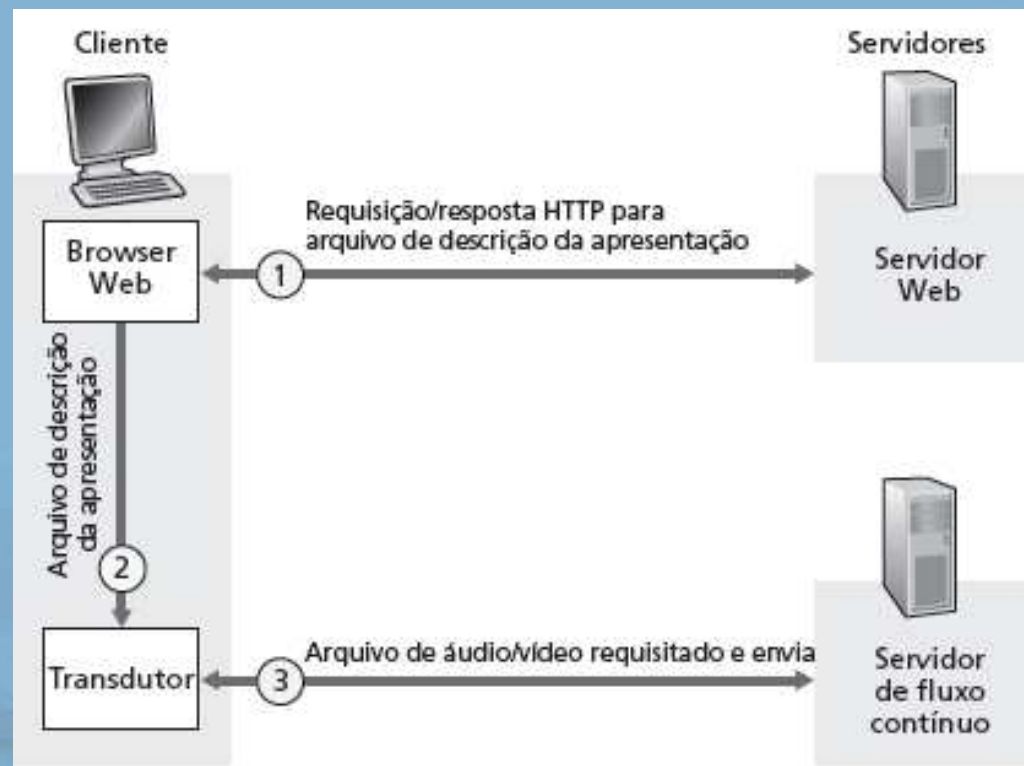
- ❑ sem "canalização", longos atrasos até reprodução!

Multimedia na Internet: técnica de fluxo contínuo



- ❑ navegador apanha (GET) **metarquivo**
- ❑ navegador dispara transdutor, passando metarquivo
- ❑ transdutor contata servidor
- ❑ servidor envia **fluxo contínuo** de áudio/vídeo ao transdutor

Fluxo contínuo de um servidor de fluxo contínuo

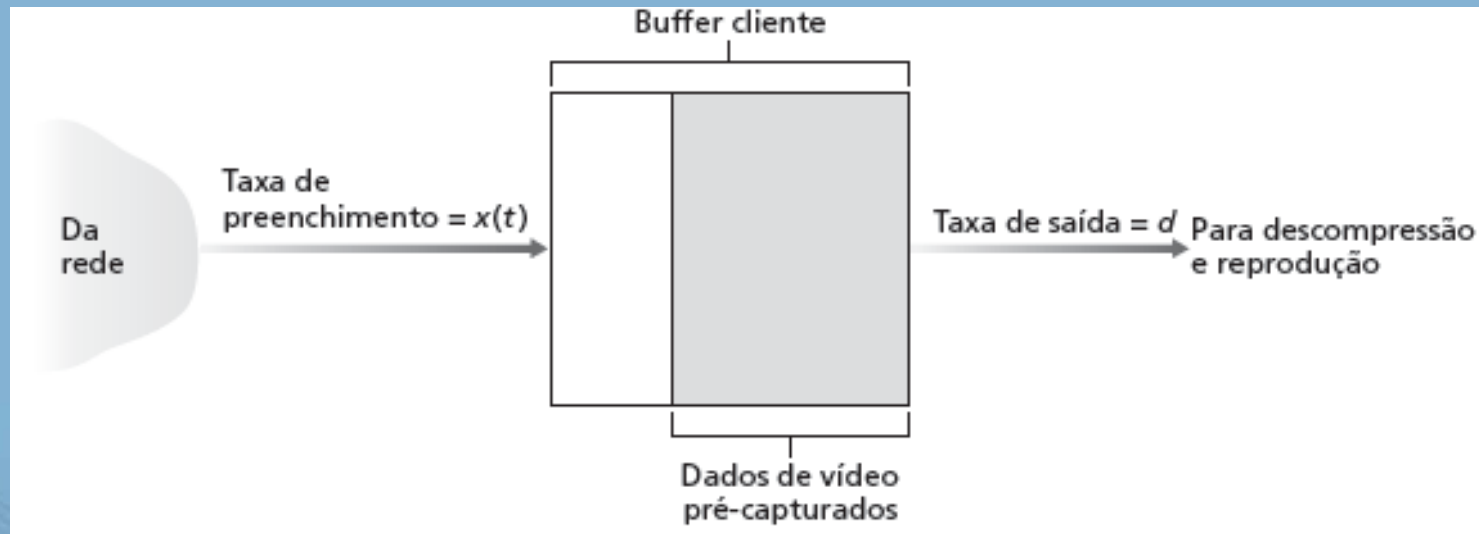


- ❑ permite protocolo não HTTP entre servidor e transdutor
- ❑ UDP ou TCP para etapa (3); veja mais adiante

Multimídia de fluxo contínuo: buffer no cliente



- buffer no cliente, atraso na reprodução compensa atraso adicional da rede, jitter



- buffer no cliente, atraso na reprodução compensa atraso adicional da rede, jitter

Multimídia de fluxo contínuo: UDP ou TCP?

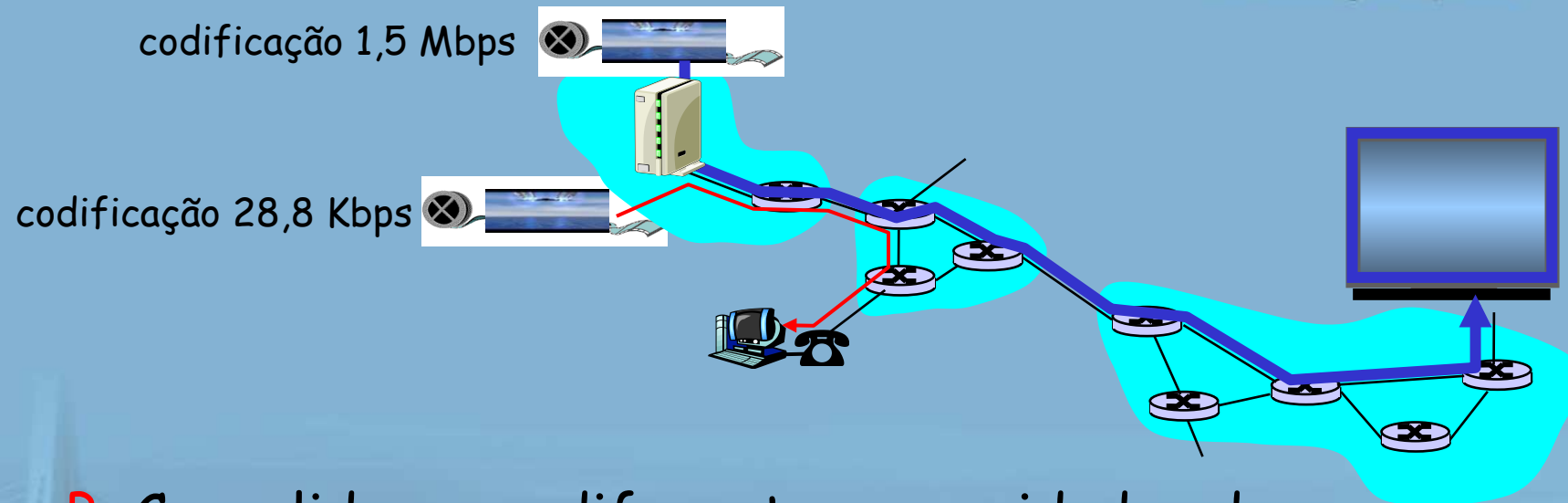
UDP

- ❑ servidor envia na taxa apropriada ao cliente (desatento ao congestionamento na rede!)
 - normalmente, taxa envio = taxa codif. = taxa constante
 - depois, taxa de preenchimento = taxa constante - perda de pacote
- ❑ pequeno atraso na reprodução (2-5 s) para remover jitter da rede
- ❑ recuperação de erro: se o tempo permitir

TCP

- ❑ envio na maior taxa possível sob TCP
- ❑ taxa de preenchimento flutua devido ao controle de congestionamento TCP
- ❑ maior atraso na reprodução: taxa de envio TCP suave
- ❑ HTTP/TCP passa mais facilmente pelos firewalls

Multimídia de fluxo contínuo: taxa(s) do cliente



P: Como lidar com diferentes capacidades de taxa de recepção do cliente?

- rede discada a 28,8 Kbps
- rede Ethernet a 100 Mbps

R: Servidor armazena e transmite várias cópias do vídeo, codificadas em diferentes taxas

Controle do usuário da mídia de fluxo contínuo: RTSP

HTTP

- ❑ não visa conteúdo de multimídia
- ❑ sem comandos para avanço rápido etc.

RTSP: RFC 2326

- ❑ protocolo da camada de aplicação cliente-servidor
- ❑ controle do usuário: retrocesso, avanço rápido, pause, reinício, reposicionamento etc....

O que ele não faz:

- ❑ não define como áudio, e vídeo são encapsulados para fluxo contínuo pela rede
- ❑ não restringe como a mídia de fluxo contínuo é transportada (UDP ou TCP possível)
- ❑ não especifica como transdutor mantém áudio/vídeo em buffer

RTSP: controle fora da banda

FTP usa canal de controle "fora da banda" :

- ❑ arquivo transferido por uma conexão TCP
- ❑ informação de controle (mudanças de diretório, exclusão de arquivo, renomeação) enviadas por conexão TCP separada
- ❑ canais "fora de banda", "na banda" usam números de porta diferentes

Mensagens RTSP também enviadas fora da banda:

- ❑ Mensagens de controle RTSP usam diferentes números de porta do fluxo contínuo de mídia: fora da banda
 - porta 554
- ❑ fluxo contínuo de mídia é considerado "na banda"

Exemplo do RTSP

Cenário:

- ❑ metarquivo comunicado ao navegador Web
- ❑ navegador inicia transdutor
- ❑ transdutor configura conexão de controle RTSP, conexão de dados ao servidor de fluxo contínuo

Exemplo de metarquivo

```
<title>Twister</title>
```

```
<session>
```

```
  <group language = en lipsync>
```

```
    <switch>
```

```
      <track type = audio
```

```
        e = "PCMU/8000/1"
```

```
        src = "rtsp://audio.example.com/twister/audio.en/lofi">
```

```
      <track type = audio
```

```
        e = "DVI4/16000/2" pt = "90 DVI4/8000/1"
```

```
        src = "rtsp://audio.example.com/twister/audio.en/hifi">
```

```
    </switch>
```

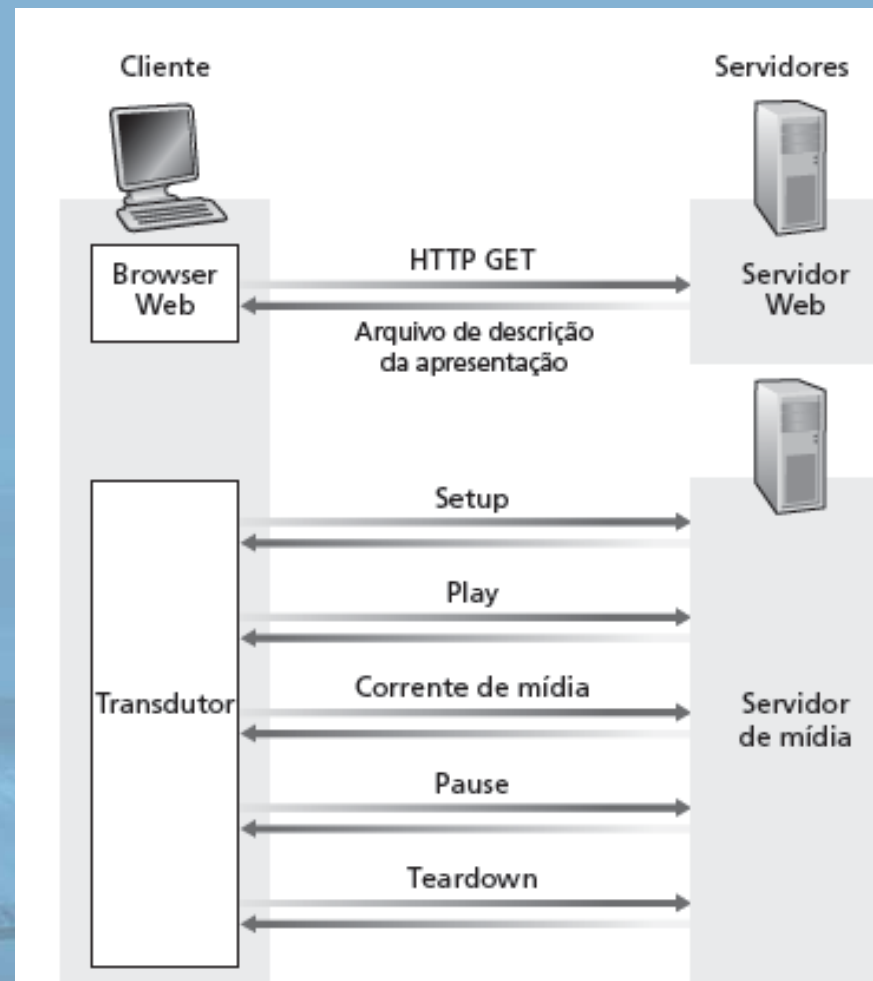
```
  <track type = "video/jpeg"
```

```
    src = "rtsp://video.example.com/twister/video">
```

```
</group>
```

```
</session>
```

Operação do RTSP



Exemplo de sessão RTSP

C: SETUP rtsp://audio.example.com/twister/audio RTSP/1.0
Transport: rtp/udp; compression; port = 3056; mode = PLAY

S: RTSP/1.0 200 1 OK
Session 4231

C: PLAY rtsp://audio.example.com/twister/audio.en/lofi RTSP/1.0
Session: 4231
Range: npt = 0-

C: PAUSE rtsp://audio.example.com/twister/audio.en/lofi RTSP/1.0
Session: 4231
Range: npt = 37

C: TEARDOWN rtsp://audio.example.com/twister/audio.en/lofi RTSP/1.0
Session: 4231

S: 200 3 OK

Capítulo 7: Esboço

- ❑ 7.1 Aplicações de rede multimídia
- ❑ 7.2 Áudio e vídeo de fluxo contínuo armazenados
- ❑ 7.3 Fazendo o melhor possível com o serviço de melhor esforço
- ❑ 7.4 Protocolos para aplicações interativas em tempo real - RTP, RTCP, SIP
- ❑ 7.5 Fornecendo classes de serviço múltiplas
- ❑ 7.6 Fornecendo garantias de qualidade de serviços

Aplicações interativas em tempo real

- ❑ telefone PC-a-PC
 - Skype
- ❑ PC-para-telefone
 - discado
 - Net2phone
 - Skype
- ❑ videoconferência com webcams
 - Skype
 - Polycom

Vamos agora
examinar um
exemplo de telefone
PC-a-PC na Internet
com detalhes

Multimedia interativa: Internet Phone

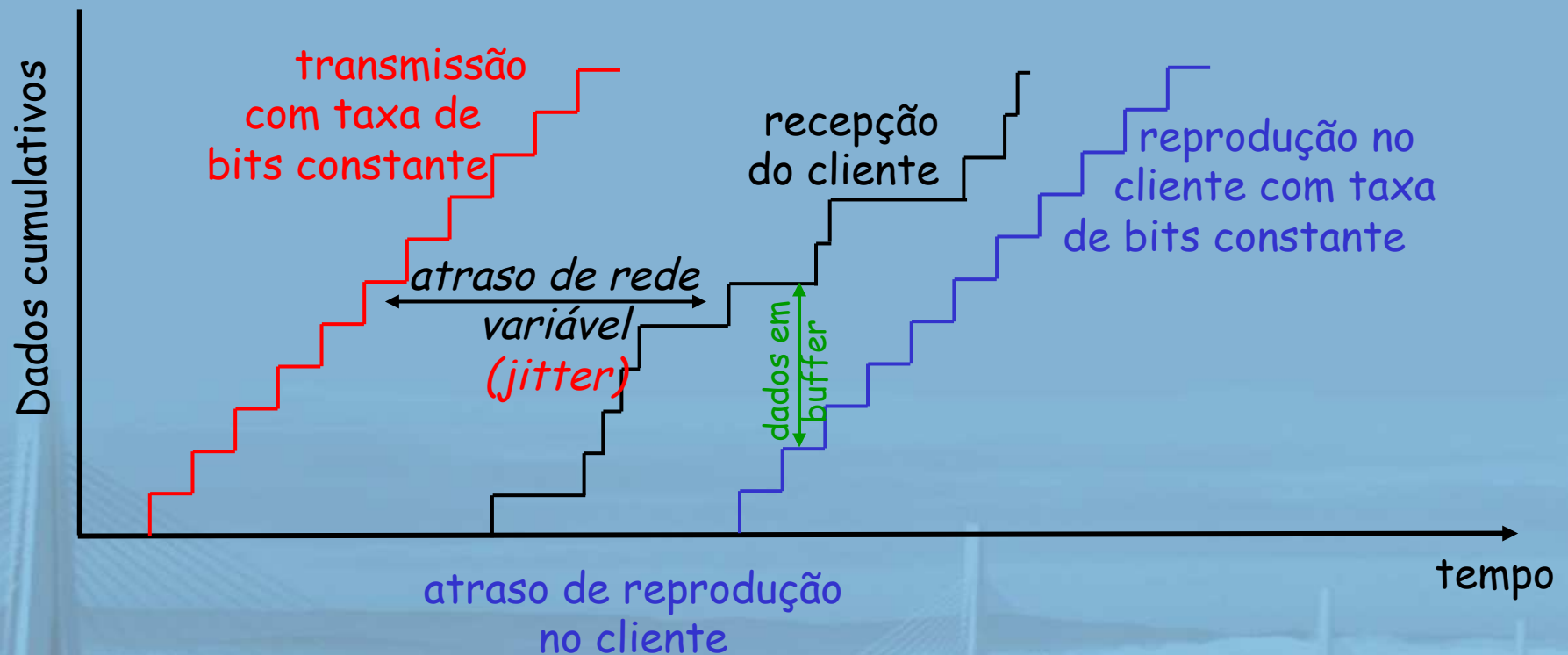
Apresento Internet Phone por meio de um exemplo

- ❑ áudio do locutor: alternando rajadas de voz e silêncio
 - 64 kbps durante a rajada de voz
 - pacotes gerados apenas durante as rajadas de voz
 - porções de 20 ms a 8 Kbytes/s: 160 bytes de dados
- ❑ cabeçalho da camada de aplicação acrescentado a cada porção
- ❑ porção + cabeçalho encapsulados no segmento UDP
- ❑ aplicação envia segmento UDP para socket a cada 20 ms durante a rajada de voz

Internet Phone: perda de pacote e atraso

- ❑ **perda na rede:** perda de datagrama IP devido a congestionamento na rede (estouro de buffer do roteador)
- ❑ **perda por atraso:** datagrama IP chega muito tarde para reprodução no receptor
 - atrasos: atrasos de processamento, filas na rede; sistema final (remetente, receptor)
 - atraso típico máximo tolerável: 400 ms
- ❑ tolerância a perda: dependendo da codificação de voz, perdas ocultas e taxas de perda de pacotes entre 1% e 10% podem ser toleradas

Variação de atraso



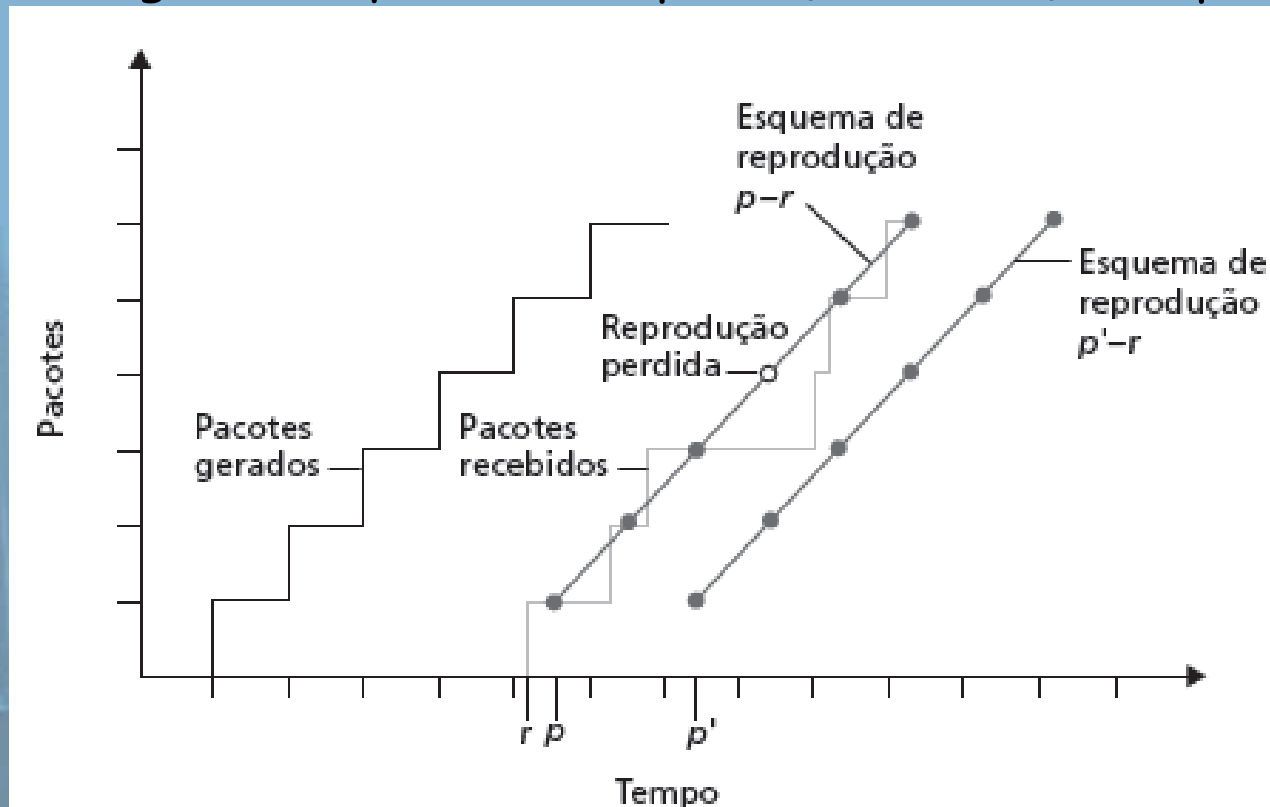
- considere atrasos de fim a fim de dois pacotes consecutivos: diferença pode ser mais ou menos 20 ms (diferença no tempo de transmissão)

Internet Phone: atraso de reprodução fixo

- receptor tenta reproduzir cada porção exatamente q ms após a porção ter sido gerada
 - porção tem marca de tempo t : reproduz porção em $t + q$.
 - porção chega após $t + q$: dados chegam muito tarde para reprodução e se “perdem”
- dilema na escolha de q :
 - *q grande*: menos perda de pacote
 - *q pequeno*: melhor experiência interativa

Atraso de reprodução fixo

- remetente gera pacotes a cada 20 ms durante rajada de voz
- primeiro pacote recebido no instante r
- primeiro esquema de reprodução: começa em p
- segundo esquema de reprodução: começa em p'



Atraso de reprodução adaptativo

- ❑ **Objetivo:** minimizar atraso de reprodução, mantendo a taxa de perda baixa
- ❑ **Técnica:** ajuste do atraso de reprodução adaptativo:
 - estime atraso da rede, ajuste atraso de reprodução no início de cada rajada de voz
 - períodos de silêncio compactados e alongados
 - porções ainda reproduzidas a cada 20 ms durante a rajada de voz

t_i = marca de tempo do i° pacote

r_i = o momento em que o pacote i é recebido pelo receptor

p_i = o momento em que o pacote i é reproduzido no receptor

$r_i - t_i$ = atraso na rede para i° pacote

d_i = estimativa do atraso médio da rede após receber i° pacote

estimativa dinâmica do atraso médio no receptor:

$$d_i = (1 - u)d_{i-1} + u(r_i - t_i)$$

onde u é uma constante fixa (p. e., $u = 0,01$)

- também útil para estimar desvio médio do atraso, v_i :

$$v_i = (1 - u)v_{i-1} + u|r_i - t_i - d_i|$$

- estima d_i , v_i calculado para cada pacote recebido (mas usado apenas no início da rajada de voz)
- para primeiro pacote na rajada de voz, tempo de reprodução é:

$$p_i = t_i + d_i + Kv_i$$

onde K é uma constante positiva

- pacotes restantes na rajada de voz são reproduzidos periodicamente

- P: Como o receptor determina se o pacote é o primeiro em uma rajada de voz?
- se não há perda, receptor examina marcas de tempo sucessivas
 - diferença de marcas de tempo sucessivas > 20 ms --> rajada de voz começa
 - com possível perda, receptor deve examinar marcas de tempo e números de sequência.
 - diferença de marcas sucessivas > 20 ms e números de sequência sem lacunas --> rajada de voz começa

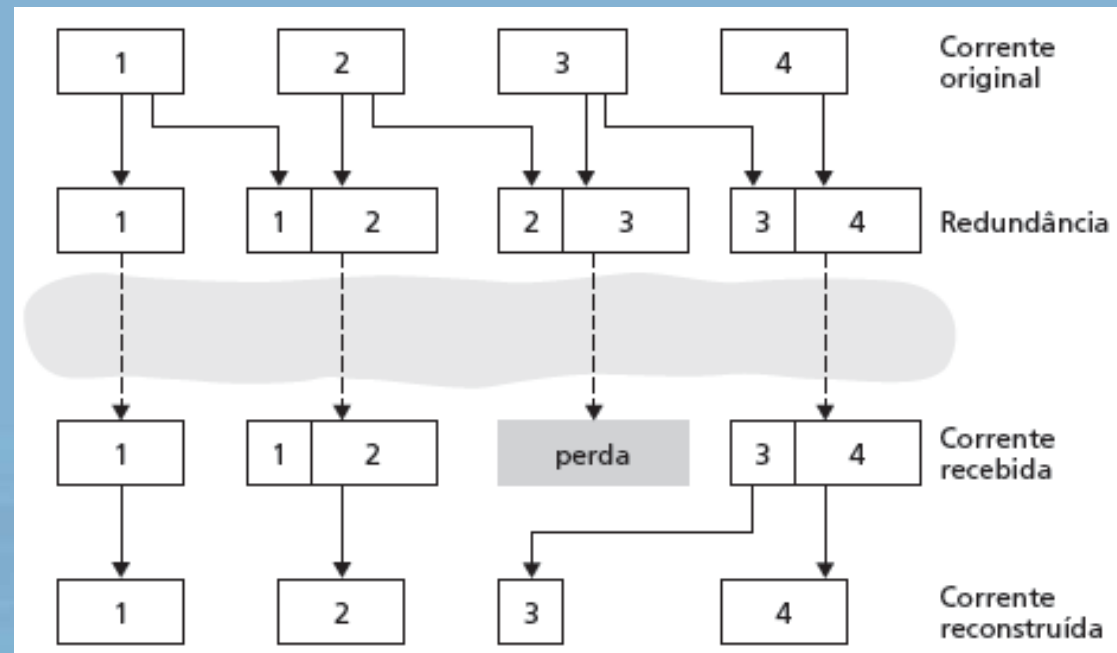
Recuperação de perda de pacotes

Forward Error Correction (FEC): mecanismo simples

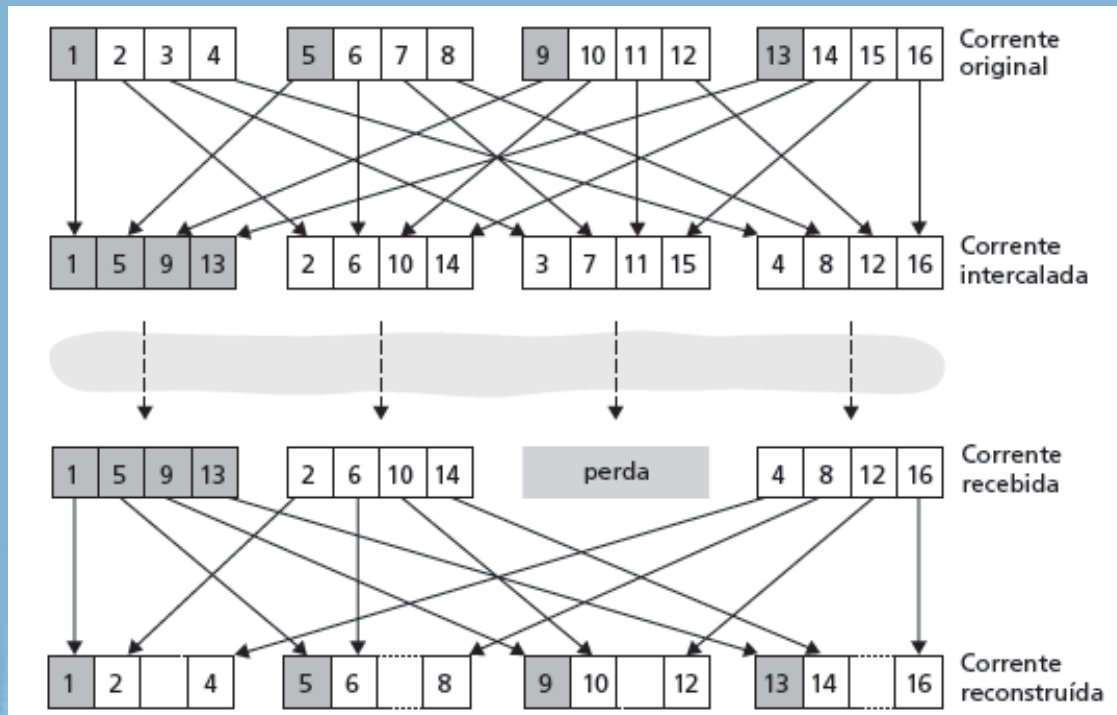
- ❑ para cada grupo de n porções, crie porção redundante com OR exclusivo de n porções originais
- ❑ envie $n + 1$ porções, aumentando largura de banda pelo fator $1/n$.
- ❑ pode reconstruir n porções originais se no máximo uma porção perdida dentre $n + 1$ porções
- ❑ atraso de reprodução: tempo suficiente para receber todos $n + 1$ pacotes
- ❑ dilema:
 - aumente n , menos desperdício de largura de banda
 - aumente n , maior atraso de reprodução
 - aumente n , maior probabilidade de que 2 ou mais porções se percam

2º mecanismo FEC

- ❑ "fluxo contínuo de menor qualidade"
- ❑ envia fluxo com resolução de áudio inferior como informação redundante
- ❑ p. e., PCM de fluxo nominal a 64 kbps e GSM de fluxo redundante a 13 kbps.



- ❑ sempre que há perda não consecutiva, receptor pode ocultar a perda
- ❑ também pode anexar $(n-1)^a$ e $(n-2)^a$ porção com baixa taxa de bits



Intercalação

- ❑ porções divididas em unidades menores
- ❑ por exemplo, quatro unidades de 5 ms por porção
- ❑ pacote contém pequenas unidades de porções diferentes
- ❑ se pacote perdido, ainda tem a maioria de cada porção
- ❑ sem overhead de redundância, mas aumenta atraso de reprodução

Content Distribution Networks (CDNs)

Replicação de conteúdo

- ❑ difícil enviar grandes arquivos (p. e., vídeo) de único servidor de origem em tempo real
- ❑ *solução*: replicar conteúdo em centenas de servidores pela Internet
 - conteúdo baixado para servidores CDN antes da hora
 - *conteúdo "perto" do usuário evita dados (perda, atraso) do envio por longos caminhos*
 - servidor CDN normalmente na rede da borda/acesso

servidor de origem na América do Norte



nó de distribuição de CDN



servidor CDN na América do Sul



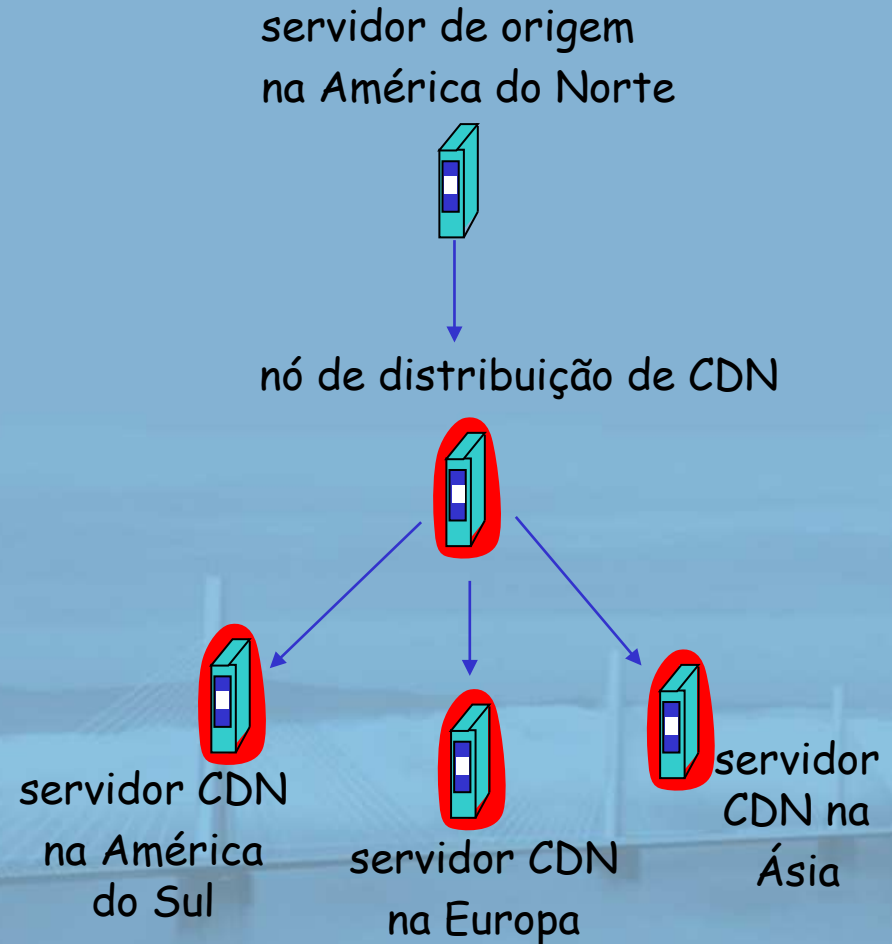
servidor CDN na Europa



servidor CDN na Ásia

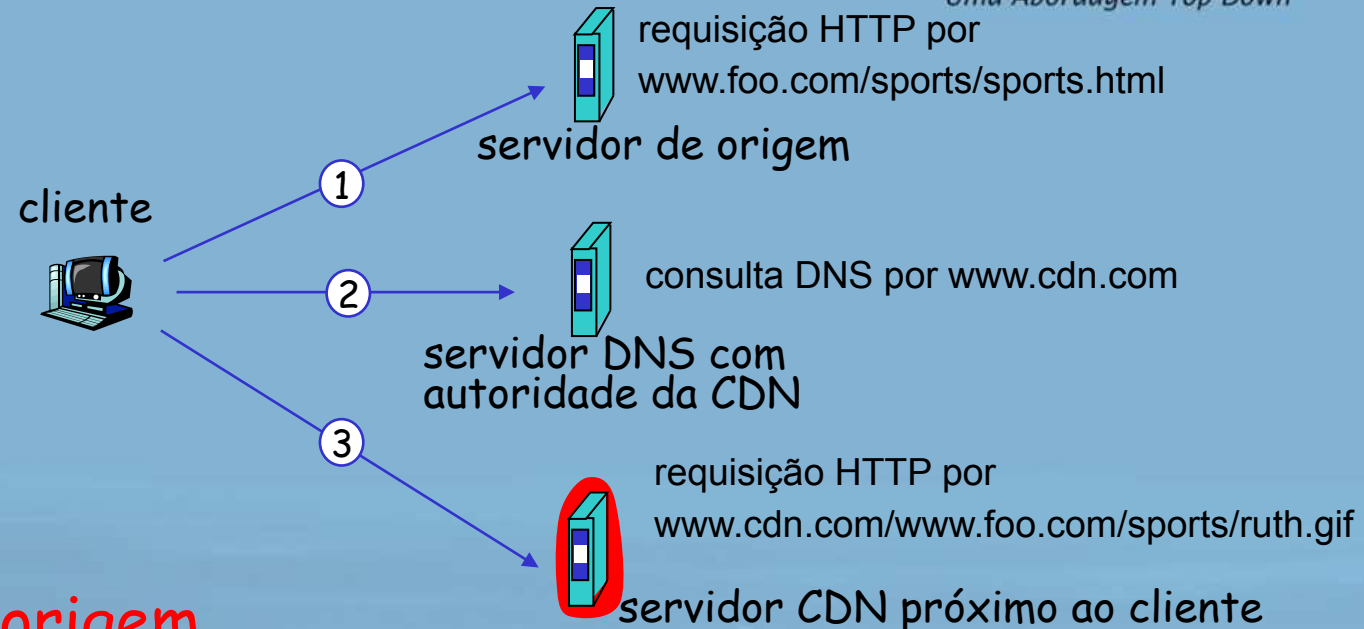
Replicação de conteúdo

- ❑ cliente CDN (p. e., Akamai) é provedor de conteúdo (p. e., CNN)
- ❑ CDN replica conteúdo do cliente nos servidores CDN
- ❑ quando provedor atualiza conteúdo, CDN atualiza servidores



Exemplo de CDN

Uma Abordagem Top-Down



servidor de origem
(www.foo.com)

❑ distribui HTML

❑ substitui:

http://www.foo.com/sports.ruth.gif

por
sports/ruth.gif

empresa de CDN (cdn.com)

❑ distribui arquivos GIF

❑ usa seu servidor DNS com
autoridade para rotear
requisições

http://www.cdn.com/www.foo.com/

Mais sobre CDNs

requisições de roteamento

- ❑ CDN cria um "mapa", indicando distâncias de ISPs de folha e nós CDN
- ❑ quando consulta chega no servidor DNS com autoridade:
 - servidor determina ISP do qual a consulta origina
 - usa "mapa" para determinar melhor servidor CDN
- ❑ nós CDN criam rede de sobreposição da camada de aplicação

Resumo: multimídia da Internet: sacola de truques

- ❑ **use UDP** para evitar controle de congestionamento TCP (atrasos) para tráfego sensível ao tempo
- ❑ **atraso de reprodução adaptativo** no cliente: para compensar o atraso
- ❑ lado servidor **combina largura de banda da corrente** com largura de banda do caminho disponível entre cliente e servidor
 - escolha entre taxas de corrente pré-codificadas
 - taxa dinâmica de codificação de servidor
- ❑ recuperação de erro (em cima do UDP)
 - FEC, intercalação, ocultação de erro
 - retransmissões, se o tempo permitir
- ❑ CDN: leva conteúdo mais para perto dos clientes

Capítulo 7: Esboço

- ❑ 7.1 Aplicações de rede multimídia
- ❑ 7.2 Áudio e vídeo de fluxo contínuo armazenados
- ❑ 7.3 Fazendo o melhor possível com o serviço de melhor esforço
- ❑ 7.4 Protocolos para aplicações interativas em tempo real - RTP, RTCP, SIP
- ❑ 7.5 Fornecendo classes de serviço múltiplas
- ❑ 7.6 Fornecendo garantias de qualidade de serviços

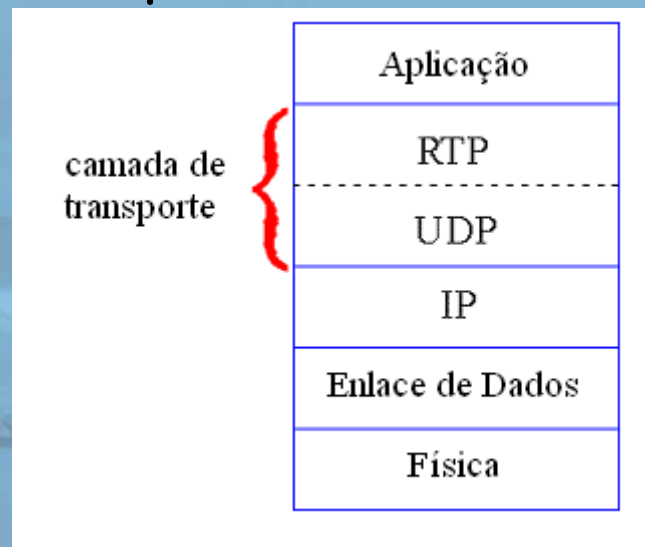
Real-Time Protocol (RTP)

- ❑ RTP especifica estrutura de pacote para transportar dados de áudio e vídeo
- ❑ RFC 3550
- ❑ pacote RTP oferece
 - identificação de tipo de carga útil
 - numeração de sequência de pacote
 - marca de tempo
- ❑ RTP roda em sistemas finais
- ❑ pacotes RTP encapsulados em segmentos UDP
- ❑ interoperabilidade: se duas aplicações de telefone da Internet rodam RTP, então elas podem ser capazes de trabalhar juntas

RTP roda sobre UDP

bibliotecas RTP oferecem interface da camada de transporte que estende UDP:

- números de porta, endereços IP
- identificação de tipo de carga útil
- numeração de sequência de pacote
- marca de tempo



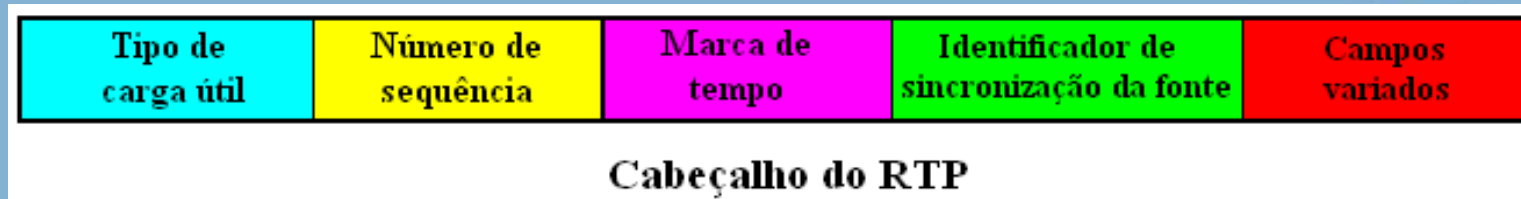
Exemplo de RTP

- ❑ considere o envio de voz codificada por PCM a 64 kbps por RTP
- ❑ aplicação coleta dados codificados em porções, p. e., cada 20 ms = 160 bytes em uma porção
- ❑ porção de áudio + cabeçalho RTP formam pacote RTP, que é encapsulado no segmento UDP
- ❑ cabeçalho RTP indica tipo de codificação de áudio em cada pacote
 - remetente pode alterar codificação durante conferência
- ❑ cabeçalho RTP também contém números de sequência, marcas de tempo

RTP e QoS

- ❑ RTP não oferece qualquer mecanismo para garantir entrega de dados a tempo ou outras garantias de QoS
- ❑ encapsulamento RTP só é visto nos sistemas finais (não) por roteadores intermediários
 - roteadores fornecendo serviço do melhor esforço, não fazendo esforço especial para garantir que os pacotes RTP chegam ao destino em tempo

Cabeçalho do RTP



tipo de carga útil (7 bits): indica tipo de codificação sendo usada atualmente. Se o remetente mudar a codificação no meio da conferência, ele informa ao receptor por meio do campo de tipo de carga útil.

- Tipo de carga útil 0: PCM lei μ , 64 kbps
- Tipo de carga útil 3, GSM, 13 kbps
- Tipo de carga útil 7, LPC, 2,4 kbps
- Tipo de carga útil 26, Motion JPEG
- Tipo de carga útil 31. H.261
- Tipo de carga útil 33, vídeo MPEG2

número de sequência (16 bits): incrementa para cada pacote RTP enviado e pode ser usado para detectar perda de pacote e restaurar sequência de pacote.

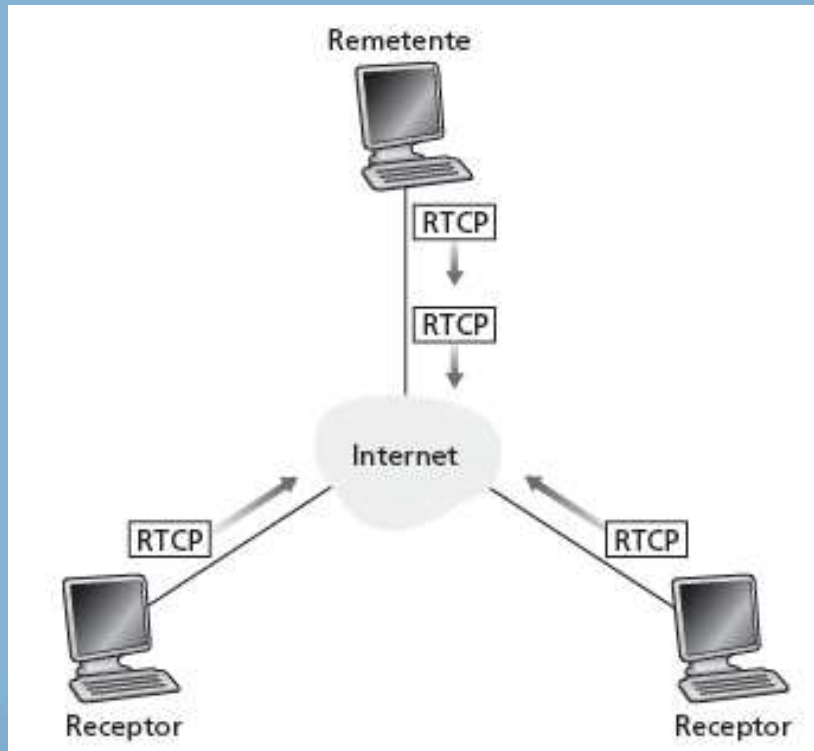
- campo de marca de tempo (32 bytes): instante de amostragem do primeiro byte neste pacote de dados RTP
 - para áudio, o clock da marca de tempo incrementa para cada período de amostragem (p. e., a cada 125 μ s para clock de amostragem de 8 KHz)
 - se a aplicação gera porções de 160 amostras codificadas, então marca de tempo aumenta em 160 para cada pacote RTP quando a origem está ativa. Clock da marca de tempo continua a aumentar em taxa constante quando a origem está inativa.
- campo SSRC (32 bits): identifica origem da corrente de RTP t. Cada corrente na sessão RTP deverá ter SSRC distinto.

Tarefa de programação de RTSP/RTP

- ❑ crie um servidor que encapsule quadros de vídeo armazenados em pacotes RTP
 - apanhe quadro de vídeo, inclua cabeçalhos RTP, crie segmentos UDP, envie segmentos para socket UDP
 - inclua números de sequência e marcas de tempo
 - cliente RTP fornecido para você
- ❑ escreva também lado cliente do RTSP
 - emita comandos de reprodução/pausa
 - servidor RTSP fornecido para você

Real-Time Control Protocol (RTCP)

- ❑ funciona em conjunto com RTP.
- ❑ cada participante na sessão RTP transmite periodicamente pacotes de controle RTCP a todos os outros participantes
- ❑ cada pacote RTCP contém relatórios de remetente e/ou receptor
 - estatísticas de relatório úteis à aplicação: # pacotes enviados, # pacotes perdidos, jitter entre chegadas etc.
- ❑ informações de retorno podem ser usadas para controlar desempenho
 - remetente pode modificar suas transmissões com base nessas informações



- ❑ cada sessão RTP: normalmente, um único endereço multicast; todos os pacotes RTP/RTCP pertencentes à sessão utilizam endereço multicast.
- ❑ pacotes RTP, RTCP distinguidos um do outro por números de porta distintos.
- ❑ para limitar o tráfego, cada participante reduz o tráfego RTCP à medida que o número de participantes da conferência aumenta

Pacotes RTCP

Pacotes de relatório do receptor:

- ❑ fração de pacotes perdidos, último número de sequência, jitter médio entre chegadas

Pacotes de relatório do remetente:

- ❑ SSRC da corrente RTP, hora atual, número de pacotes enviados, número de bytes enviados

Pacotes de descrição da fonte:

- ❑ endereço de e-mail do remetente, nome do remetente, SSRC da corrente RTP associada
- ❑ oferecem mapeamento entre o SSRC e o nome do usuário/hospedeiro

Sincronização de correntes

- ❑ RTCP pode sincronizar diferentes correntes de mídia dentro de uma sessão RTP
- ❑ considere aplicação de videoconferência para a qual cada remetente gera uma corrente RTP para vídeo, uma para áudio.
- ❑ marcas de tempo em pacotes RTP ligadas aos clocks de amostragem de vídeo e áudio
 - não ligada à hora de um relógio comum
- ❑ cada pacote de relatório do remetente RTCP contém (para pacote gerado mais recentemente na corrente RTP associada):
 - marca de tempo do pacote RTP
 - horário em que o pacote foi criado
- ❑ receptores usam a associação para sincronizar a reprodução do áudio, vídeo

Escalabilidade da largura de banda do RTCP

- ❑ RTCP tenta limitar seu tráfego a 5% da largura de banda da sessão.

Exemplo

- ❑ Considere um remetente, enviando vídeo a 2 Mbps. Então, RTCP tenta limitar seu tráfego a 100 Kbps.
- ❑ RTCP oferece 75% de taxa aos receptores; 25% restantes ao remetente
- ❑ 75 kbps é igualmente compartilhado entre receptores:
 - com R receptores, cada receptor consegue enviar tráfego RTCP a $75/R$ kbps.
- ❑ remetente consegue enviar tráfego RTCP a 25 kbps.
- ❑ participante determina período de transmissão de pacote RTCP calculando tamanho médio do pacote RTCP (pela sessão inteira) e dividindo pela taxa alocada

SIP: Session Initiation Protocol [RFC 3261]

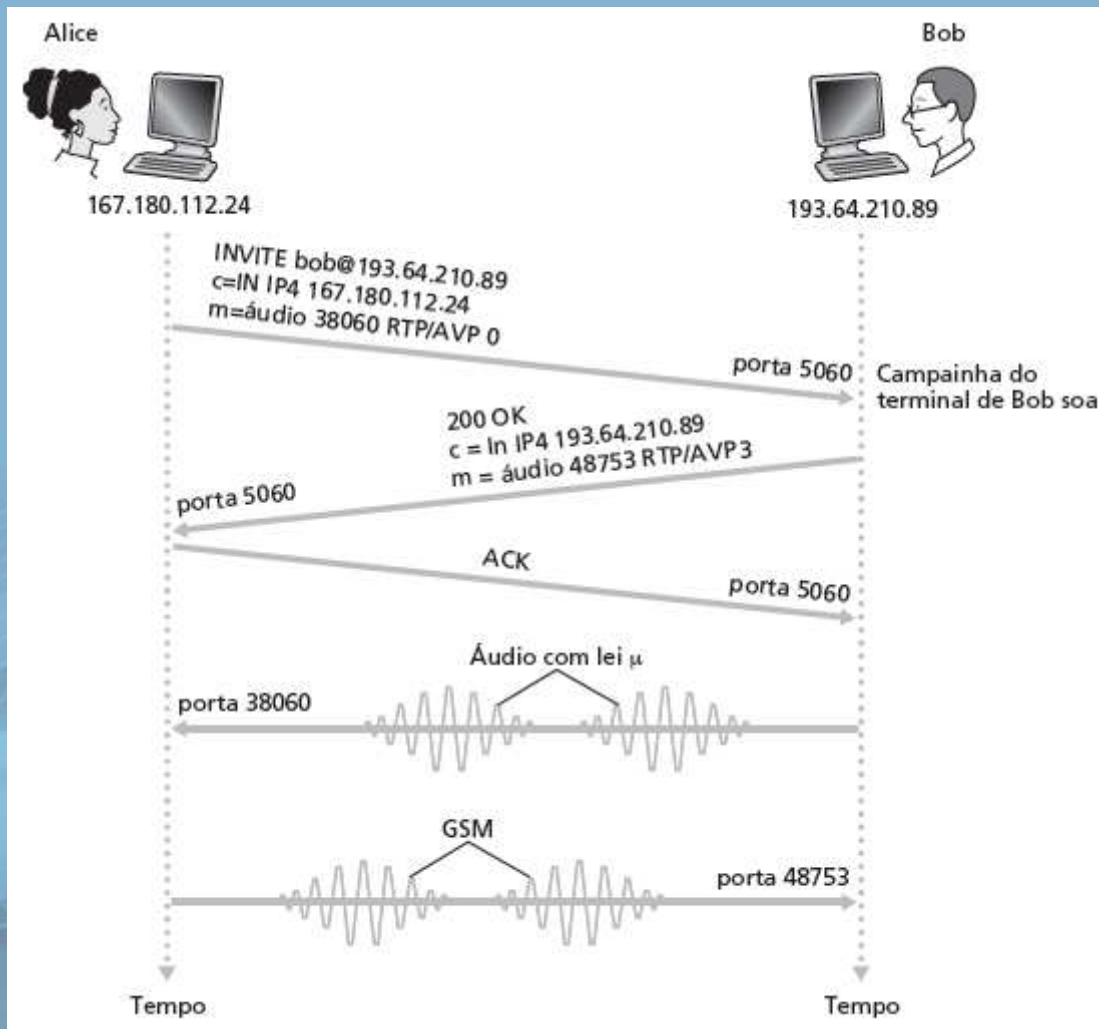
Visão a longo prazo do SIP:

- ❑ todas as ligações telefônicas e de videoconferência ocorrem pela Internet
- ❑ pessoas são identificadas por nomes ou endereços de e-mail, em vez de números de telefone
- ❑ você pode alcançar um receptor, não importa onde ele esteja ou o endereço IP que ele esteja usando atualmente

Serviços do SIP

- ❑ Estabelecendo uma chamada, o SIP oferece mecanismos ..
 - para o remetente permitir que o receptor saiba que ele deseja estabelecer uma chamada
 - assim, quem chama e quem é chamado podem combinar sobre tipo de mídia e codificação
 - encerrar chamada
- ❑ determine endereço IP atual de quem é chamado:
 - relacione identificador mnemônico ao endereço IP atual
- ❑ gerenciamento de chamada:
 - inclua novas correntes de mídia durante chamada
 - mude codificação durante chamada
 - convide outros
 - transfira e retenha chamada

Estabelecendo chamada para endereço IP conhecido



- Mensagem de convite SIP de Alice indica seu número de porta, endereço IP, codificação que ela prefere receber (PCM lei μ)
- Mensagem 200 OK de Bob indica seu número de porta, endereço IP, codificação preferida (GSM)
- Mensagens SIP podem ser enviadas por TCP ou UDP; aqui, enviada por RTP/UDP.
- número de porta default do SIP é 5060.

Estabelecendo uma chamada (mais)

- ❑ negociação codec:
 - suponha que Bob não tenha codificador PCM lei μ
 - Bob responderá com 606 Not Acceptable Reply, listando seus codificadores. Alice pode então enviar nova mensagem INVITE, anunciando codificador diferente
- ❑ rejeitando uma chamada
 - Bob pode rejeitar com respostas "busy," "gone," "payment required," "forbidden"
- ❑ mídia pode ser enviada por RTP ou algum outro protocolo

Exemplo de mensagem SIP

```
INVITE sip:bob@domain.com SIP/2.0
Via: SIP/2.0/UDP 167.180.112.24
From: sip:alice@hereway.com
To: sip:bob@domain.com
Call-ID: a2e3a@pigeon.hereway.com
Content-Type: application/sdp
Content-Length: 885

c = IN IP4 167.180.112.24
m = audio 38060 RTP/AVP 0
```

Notas:

- ❑ Sintaxe de mensagem HTTP
- ❑ sdp = protocolo de descrição de sessão
- ❑ Call-ID exclusivo para cada chamada

- ❑ Aqui, não sabemos o endereço IP de BOB. Servidores SIP intermediários são necessários.
- ❑ Alice envia e recebe mensagens SIP usando porta default do SIP, 5060
- ❑ Alice especifica na Via: cabeçalho que cliente SIP envia, recebe mensagens SIP por UDP

Tradução de nome e localização de usuário

- ❑ quem chama receptor só tem nome ou endereço de e-mail de quem ele quer chamar
- ❑ precisa obter endereço IP do hospedeiro atual de quem é chamado:
 - usuário se movimenta
 - protocolo DHCP
 - usuário tem dispositivos IP diferentes (PC, PDA, dispositivo de carro)
- ❑ resultado pode ser baseado em:
 - hora do dia (trabalho, casa)
 - quem chama (não quer que o chefe ligue para sua casa)
 - estado de quem é chamado (chamadas enviadas ao correio de voz quando já estiver falando com alguém)

Serviço fornecido por servidores SIP:

- ❑ entidade registradora SIP
- ❑ servidor proxy SIP

Entidade registra a SIP

- quando Bob inicia cliente SIP, cliente envia mensagem REGISTER do SIP ao servidor de registro de Bob (semelhante à função necessária no Instant Messaging)

Mensagem REGISTER:

```
REGISTER sip:domain.com SIP/2.0  
Via: SIP/2.0/UDP 193.64.210.89  
From: sip:bob@domain.com  
To: sip:bob@domain.com  
Expires: 3600
```

Proxy SIP

- ❑ Alice envia mensagem de convite ao seu servidor proxy
 - contém endereço sip:bob@domain.com
- ❑ proxy responsável por rotear mensagens SIP a quem é chamado
 - possivelmente através de vários proxies
- ❑ quem é chamado envia resposta de volta pelo mesmo conjunto de proxies
- ❑ proxy retorna mensagem de resposta SIP a Alice
 - contém endereço IP de Bob
- ❑ proxy semelhante ao servidor DNS local

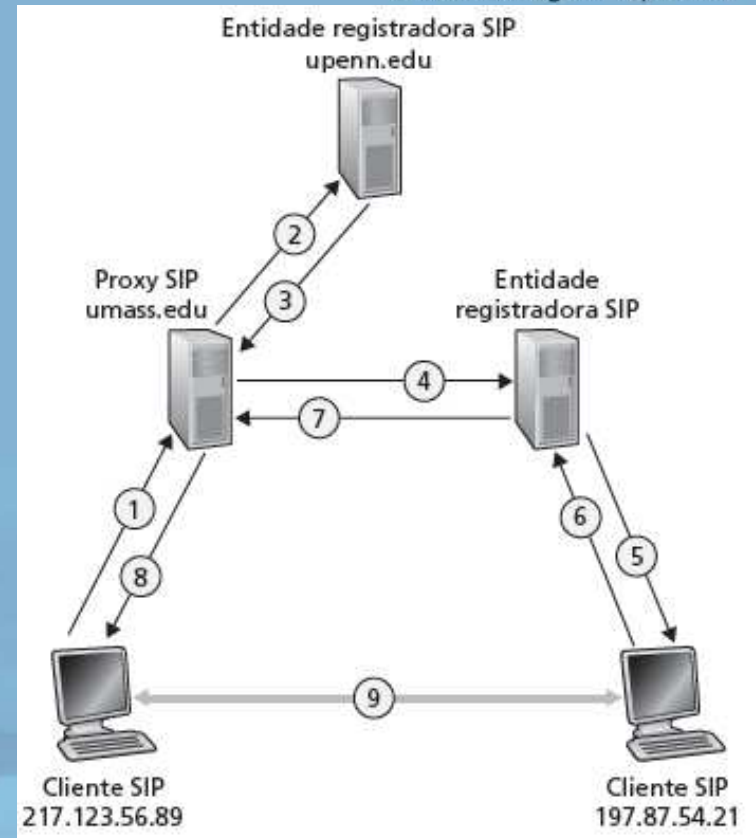
Exemplo

Remetente `jim@umass.edu`
faz chamada a `keith@upenn.edu`

- (1) Jim envia mensagem INVITE para proxy SIP umass SIP.
- (2) Proxy repassa pedido ao servidor registrador upenn.
- (3) servidor upenn retorna resposta de redirecionamento, indicando que deve tentar `keith@eurecom.fr`
- (4) proxy umass envia INVITE à registradora eurecom.

- (5) registradora eurecom repassa INVITE a 197.87.54.21, que está rodando cliente SIP de keith.
- (6-8) Resposta SIP enviada de volta
- (9) mídia enviada diretamente entre os clientes.

Nota: também há uma mensagem ack do SIP, que não aparece na figura.



Comparação com H.323

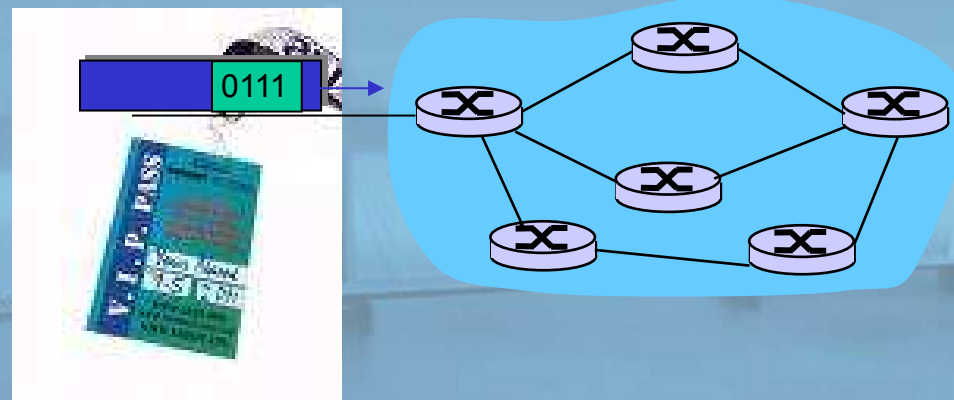
- ❑ H.323 é outro protocolo de sinalização para tempo real, interativo
- ❑ H.323 é um pacote completo, verticalmente integrado, de protocolos para conferência em multimídia: sinalização, registro, controle de admissão, transporte, codecs
- ❑ SIP é um componente isolado. Funciona com RTP, mas não o exige. Pode ser combinado com outros protocolos e serviços
- ❑ H.323 vem do ITU (telefonia).
- ❑ SIP vem do IETF: Apanha muitos de seus conceitos do HTTP
 - SIP tem forma de Web, enquanto H.323 tem forma de telefonia
- ❑ SIP usa o princípio KISS: Keep It Simple, Stupid (mantenha a simplicidade, seu ignorante!)

Capítulo 7: Esboço

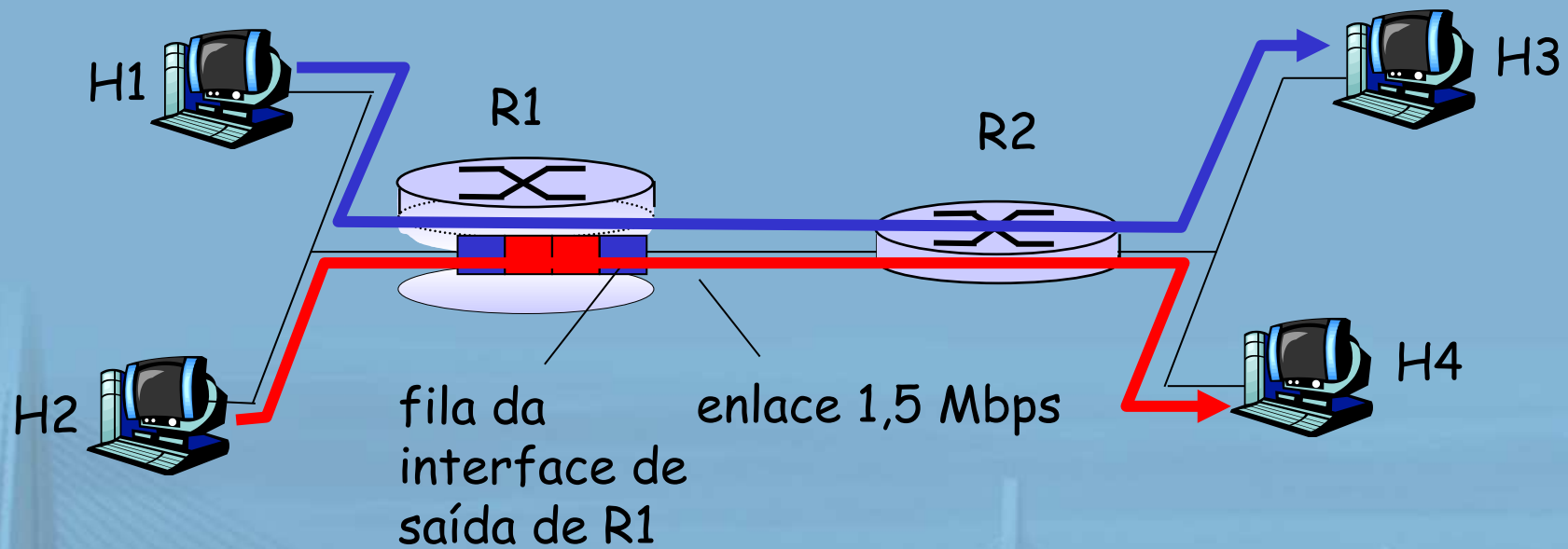
- ❑ 7.1 Aplicações de rede multimídia
- ❑ 7.2 Áudio e vídeo de fluxo contínuo armazenados
- ❑ 7.3 Fazendo o melhor possível com o serviço de melhor esforço
- ❑ 7.4 Protocolos para aplicações interativas em tempo real - RTP, RTCP, SIP
- ❑ 7.5 Fornecendo classes de serviço múltiplas
- ❑ 7.6 Fornecendo garantias de qualidade de serviços

Fornecendo múltiplas classes de serviço

- ❑ até aqui: fazer o melhor com serviço de melhor esforço
 - todo o modelo de serviço em um tamanho
- ❑ alternativa: múltiplas classes de serviço
 - particionar tráfego em classes
 - rede trata diferentes classes de tráfego de formas diferentes (analogia: serviço VIP X serviço normal)
- ❑ granularidade: serviço diferencial entre múltiplas classes, não entre conexões individuais
- ❑ história: bits de ToS

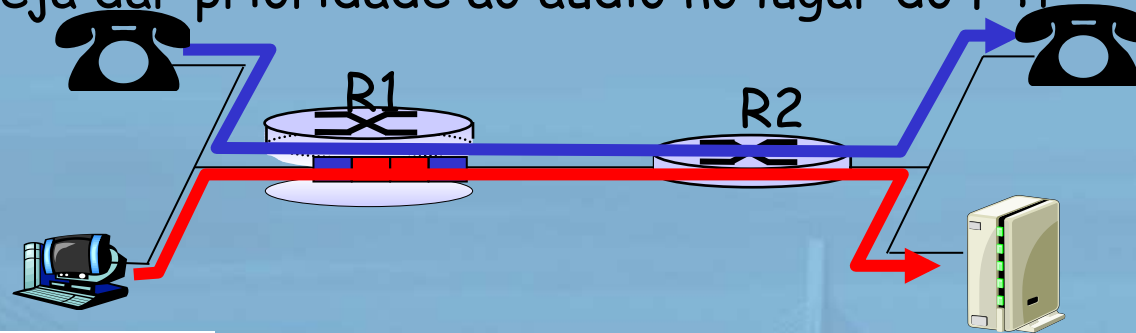


Múltiplas classes de serviço: cenário



Cenário 1: FTP e áudio misturados

- Exemplo: telefone IP a 1Mbps, FTP compartilha enlace de 1,5 Mbps.
 - rajadas de FTP podem congestionar roteador e causar perda de áudio
 - deseja dar prioridade ao áudio no lugar do FTP



Princípio 1

Marcação de pacote necessária para roteador distinguir entre diferentes classes; e nova política de roteamento para tratar pacotes de acordo.

Princípios de garantias de QOS (mais)

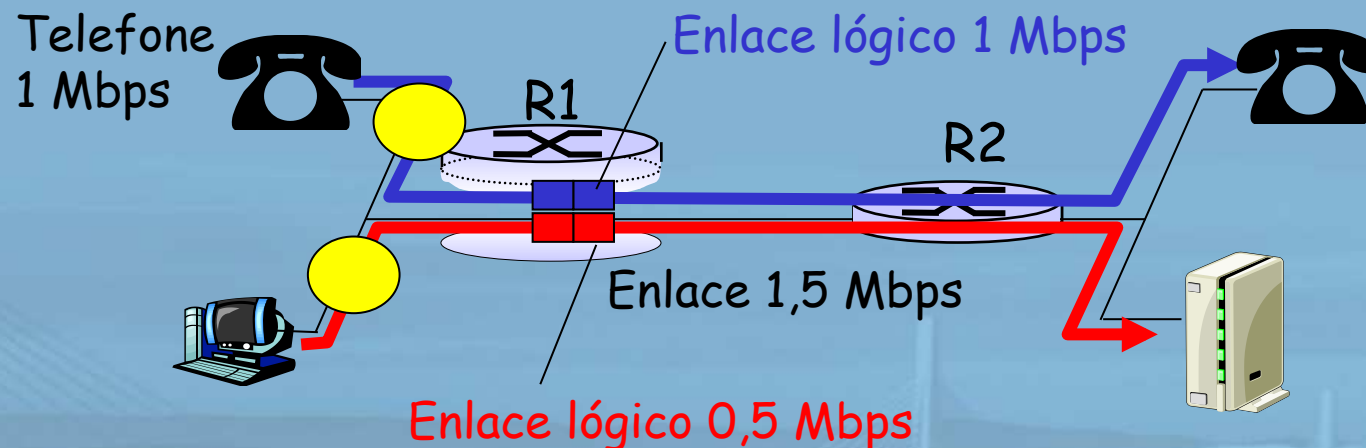
- ❑ e se as aplicações se comportarem mal (áudio envia mais do que a taxa declarada)
 - regulação: força de aderência dá origem às alocações de larg. banda
- ❑ marcação e regulação na borda da rede:
 - semelhante a ATM UNI (User Network Interface)



Princípio 2

forneça proteção (*isolamento*) de uma classe para outras

- ❑ Alocar largura de banda *fixa* (não compartilhável) ao fluxo: uso *ineficaz* da largura de banda se os fluxos não usarem sua alocação

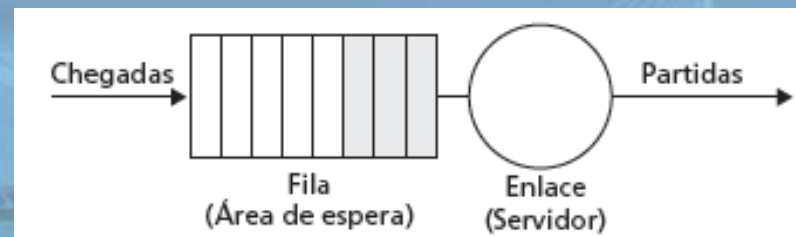


Princípio 3

Ao fornecer isolamento, é desejável usar recursos da forma mais eficiente possível

Mecanismos de escalonamento e regulação

- ❑ **escalonamento**: escolher próximo pacote a enviar no enlace
- ❑ **escalonamento FIFO (First In First Out)** : enviar na ordem de chegada à fila
 - exemplo do mundo real?
 - **política de descarte**: se pacotes chegam à fila cheia, descartar quem?
 - descarte do fim: descartar pacote que chega
 - prioridade: descartar/remover com base na prioridade
 - aleatório: descartar/remover aleatoriamente

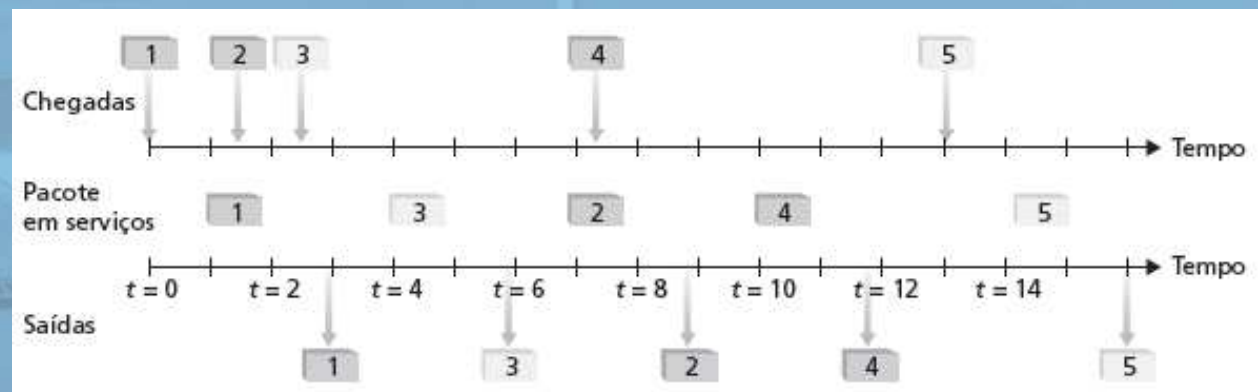
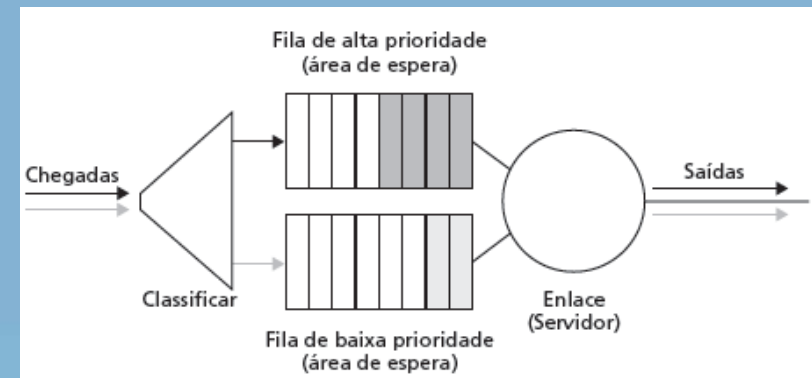


Políticas de escalonamento: mais

Escalonamento prioritário:

transmite pacote da fila com
prioridade mais alta

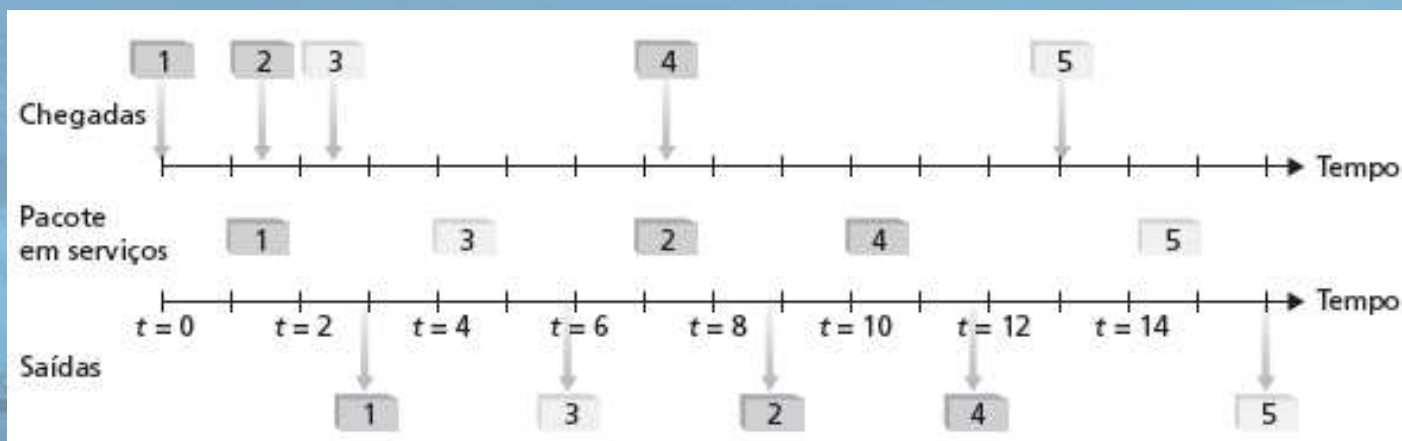
- múltiplas *classes*, com diferentes prioridades
 - classe pode depender da marcação ou outras informações de cabeçalho, p. e., origem/destino IP, números de porta etc.
 - exemplo do mundo real?



Políticas de escalonamento: ainda mais

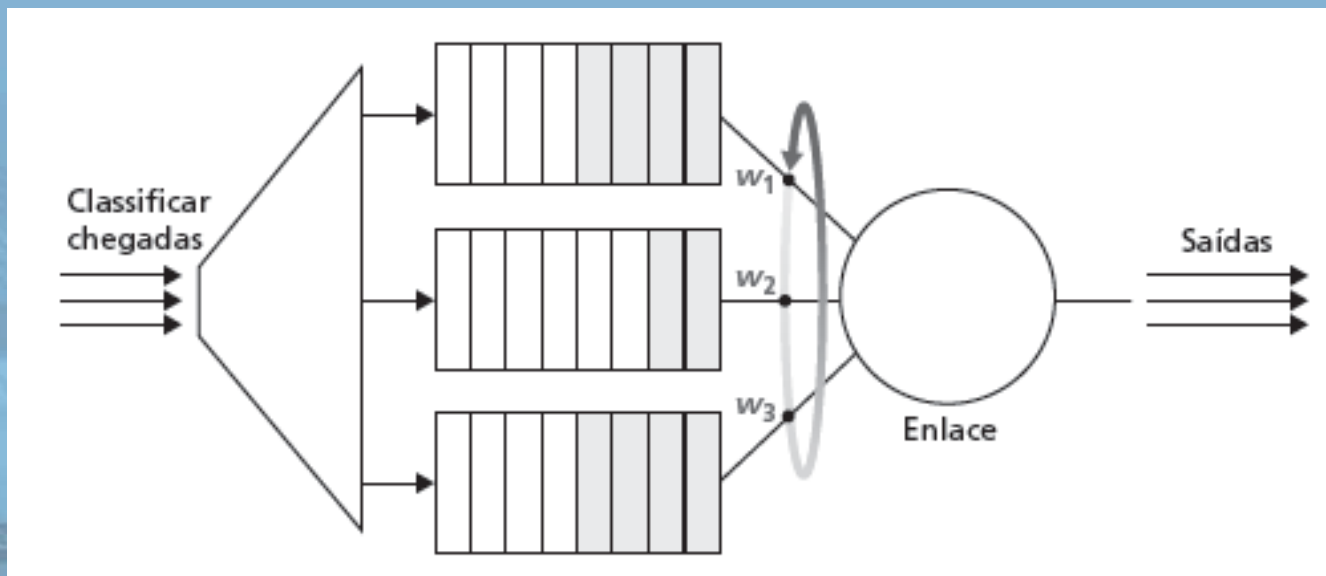
varredura cíclica:

- ❑ múltiplas classes
- ❑ varre ciclicamente as filas de classes, atendendo a uma de cada classe (se disponível)
- ❑ exemplo do mundo real?



Enfileiramento justo ponderado:

- ❑ varredura cíclica generalizada
- ❑ cada classe recebe quantidade ponderada de serviço em cada ciclo
- ❑ exemplo do mundo real?



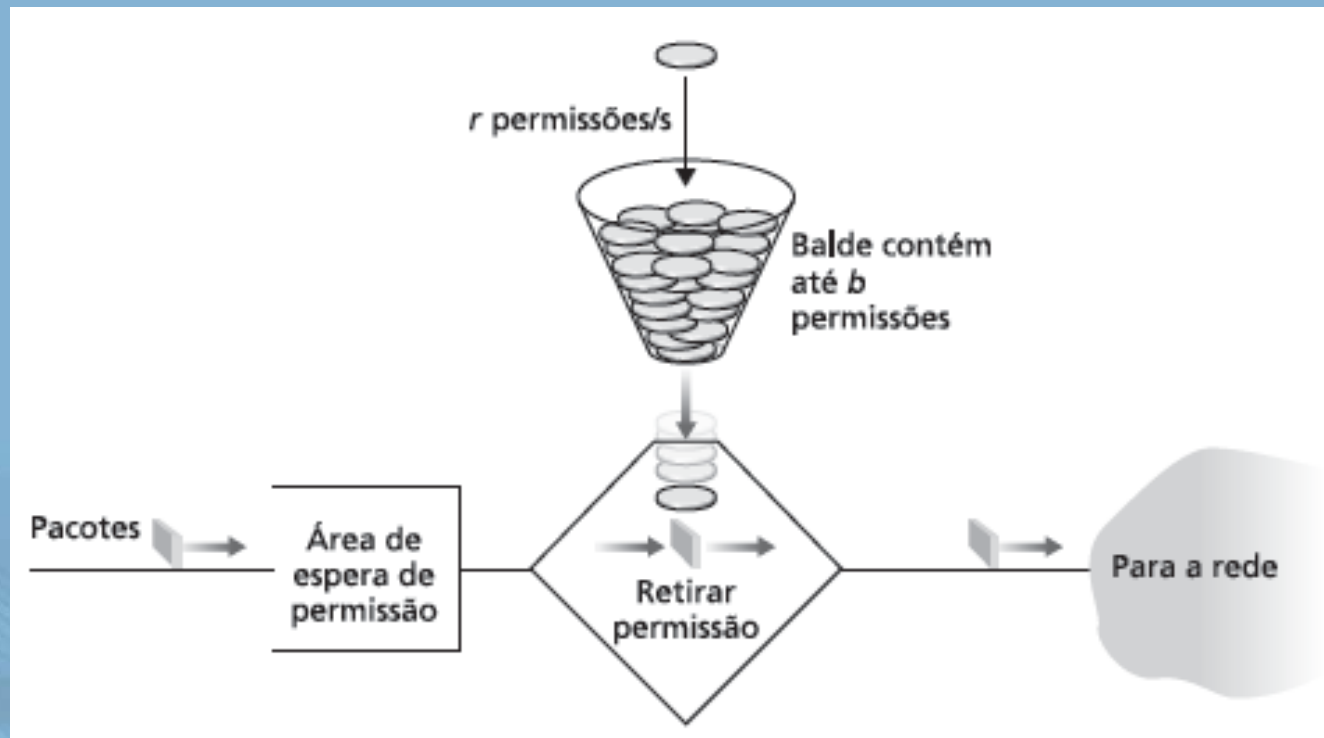
Mecanismos de regulação

Objetivo: limita tráfego para não exceder parâmetros declarados

Três critérios usados comumente:

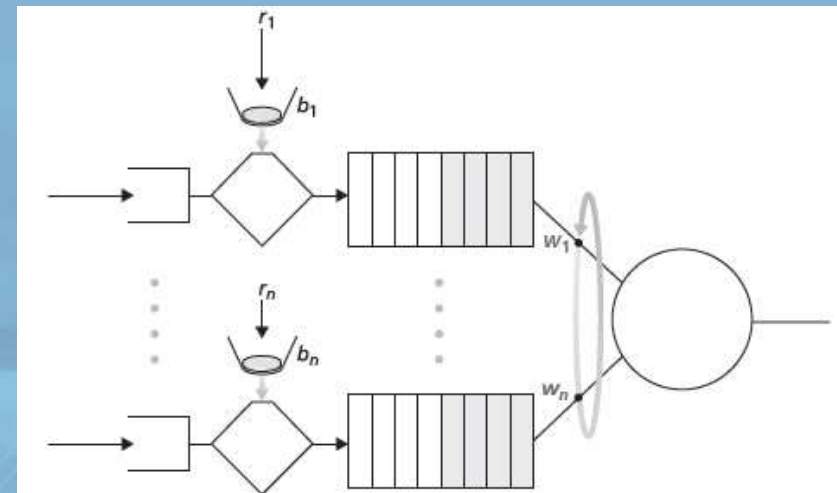
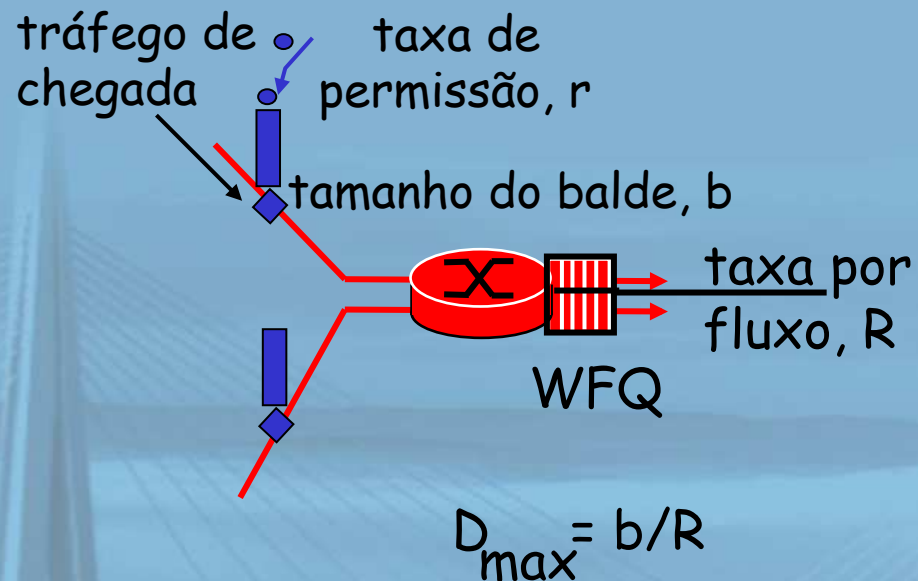
- ❑ *taxa média (longo prazo):* quantos pacotes podem ser enviados por unidade de tempo (no final das contas)
 - pergunta crucial: qual é o tamanho do intervalo: 100 pacotes por segundo ou 6000 pacotes por minuto têm mesma média!
- ❑ *taxa de pico:* p. e., 6000 pacotes por min. (ppm) média; 1500 ppm taxa de pico
- ❑ *tamanho da rajada (máximo):* número máximo de pacotes enviados consecutivamente (sem inatividade interveniente)

Balde de permissões: limita entrada a
Tamanho da Rajada e Taxa Média especificados.



- ❑ balde pode manter permissões
- ❑ permissões geradas na taxa r permissões/seg, a menos que balde esteja cheio
- ❑ *sobre intervalo de tamanho t : número de pacotes admitidos menor ou igual a $(r t + b)$.*

- balde de permissões e WFQ se combinam para fornecer limite superior garantido no atraso, ou seja, *garantia de QoS!*



Serviços diferenciados da IETF

- ❑ querem classes de serviço “qualitativas”
 - “comporta-se como um fio”
 - distinção de serviço relativa: Platinum, Gold, Silver
- ❑ **escalabilidade**: funções simples no núcleo da rede, funções relativamente complexas nos roteadores (ou hospedeiros) da borda
 - sinalização, mantendo estado do roteador por fluxo difícil com grande número de fluxos
- ❑ não defina classes de serviço, fornece componentes funcionais para criar classes de serviço

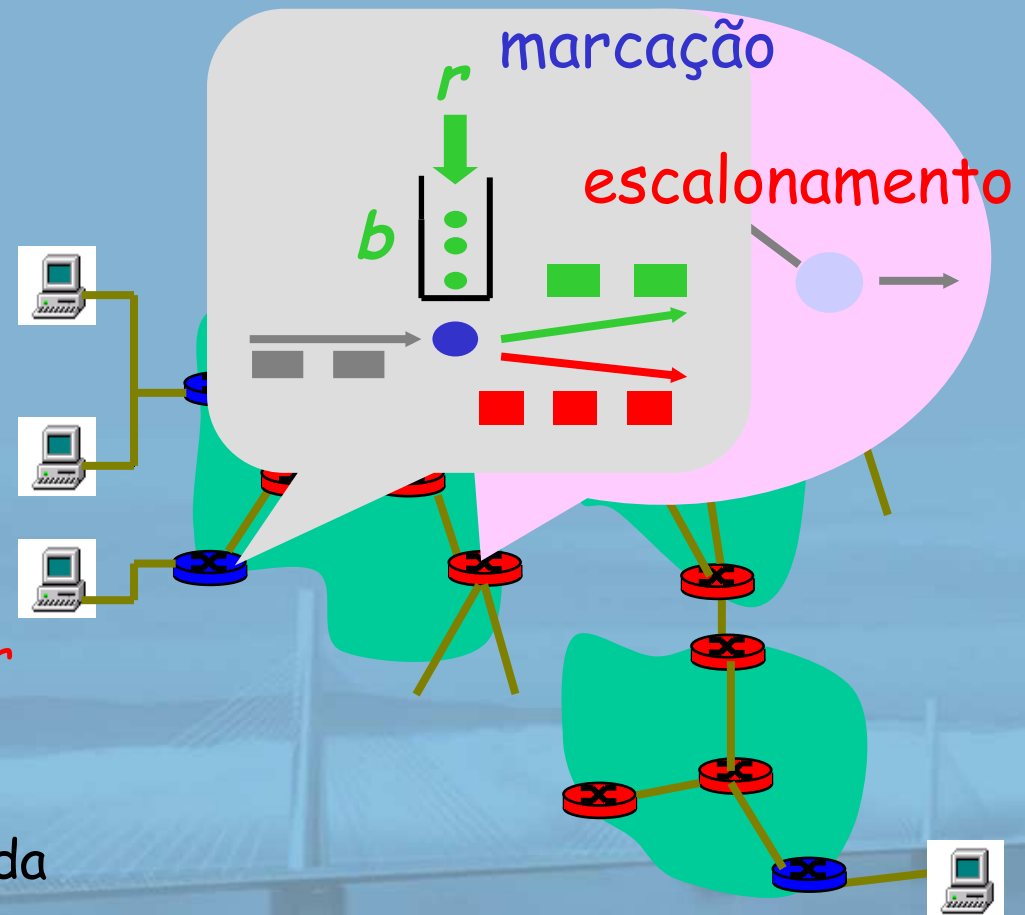
Arquitetura Diffserv

Roteador de borda:

- ❑ gerenciamento de tráfego por fluxo
- ❑ marca pacotes como **no perfil** e **fora do perfil**

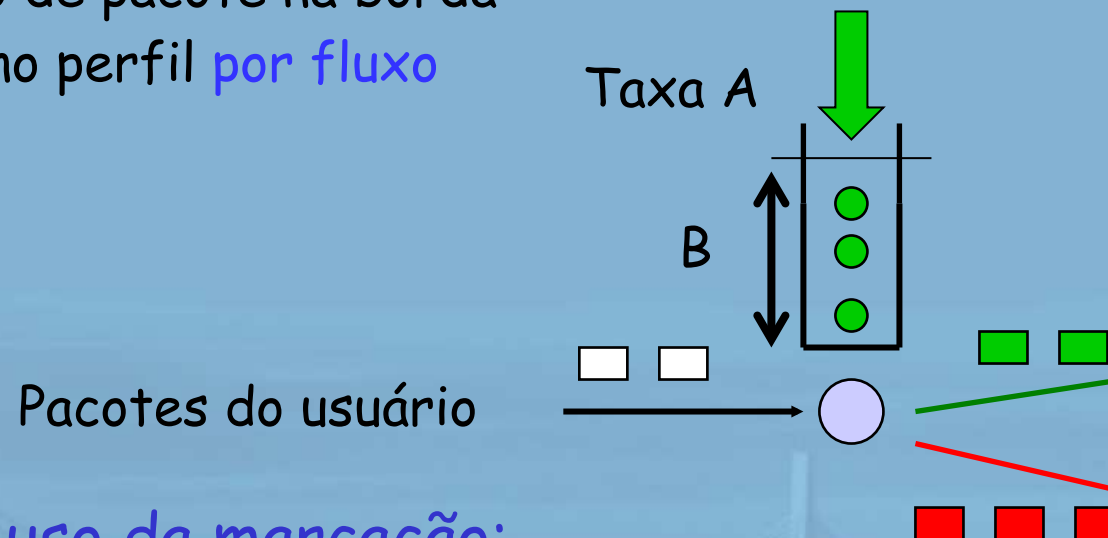
Roteador de núcleo:

- ❑ gerenciamento de tráfego **por classe**
- ❑ buffering e escalonamento baseados na **marcação** na borda
- ❑ preferência dada a pacotes **no perfil**



Marcação de pacote do roteador de borda

- perfil: taxa pré-negociada A , tamanho do balde B
- marcação de pacote na borda baseada no perfil **por fluxo**



Possível uso da marcação:

- marcação baseada em classe: pacotes de diferentes classes marcados de formas diferentes
- marcação intraclasses: parte do fluxo em conformidade marcada diferentemente da parte não em conformidade

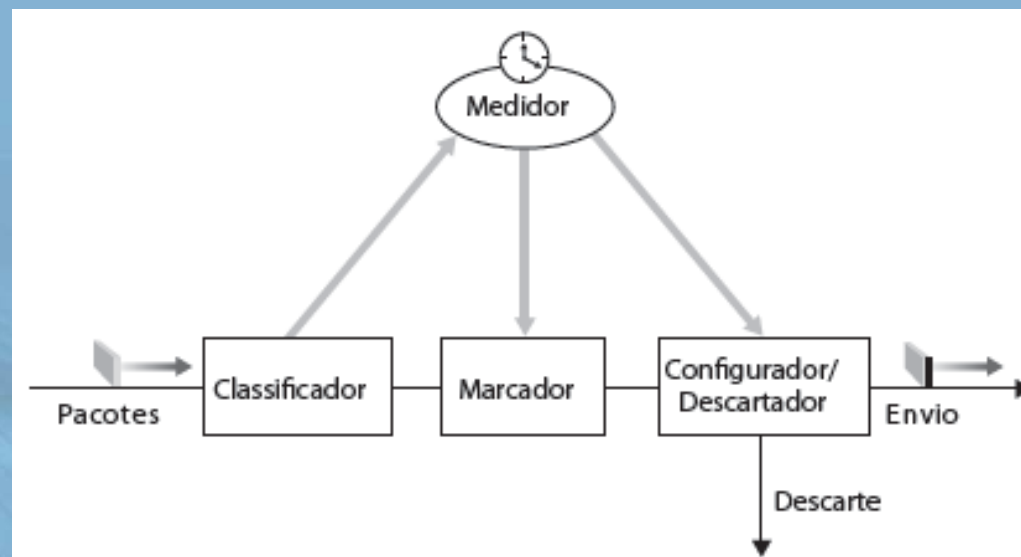
Classificação e condicionamento

- ❑ Pacote marcado no tipo de serviço (TOS) no IPv4, e classe de tráfego no IPv6
- ❑ 6 bits usados para Differentiated Service Code Point (DSCP) e determinar PHB que o pacote receberá
- ❑ 2 bits atualmente não são usados



pode ser desejável limitar a taxa de injeção de tráfego de alguma classe:

- ❑ usuário declara perfil de tráfego (p. e., taxa, tamanho de rajada)
- ❑ tráfego medido, modelado se não estiver em conformidade



Repasse (PHB)

- ❑ PHB resulta em comportamento de desempenho de repasse observável (mensurável) diferente
- ❑ PHB não especifica quais mecanismos usar para garantir comportamento de desempenho de PHB exigido
- ❑ Exemplos:
 - Classe A recebe $x\%$ da largura de banda do enlace de saída por intervalos de tempo de tamanho especificado
 - Pacotes de classe A saem antes dos pacotes de classe B

PHBs sendo desenvolvidos:

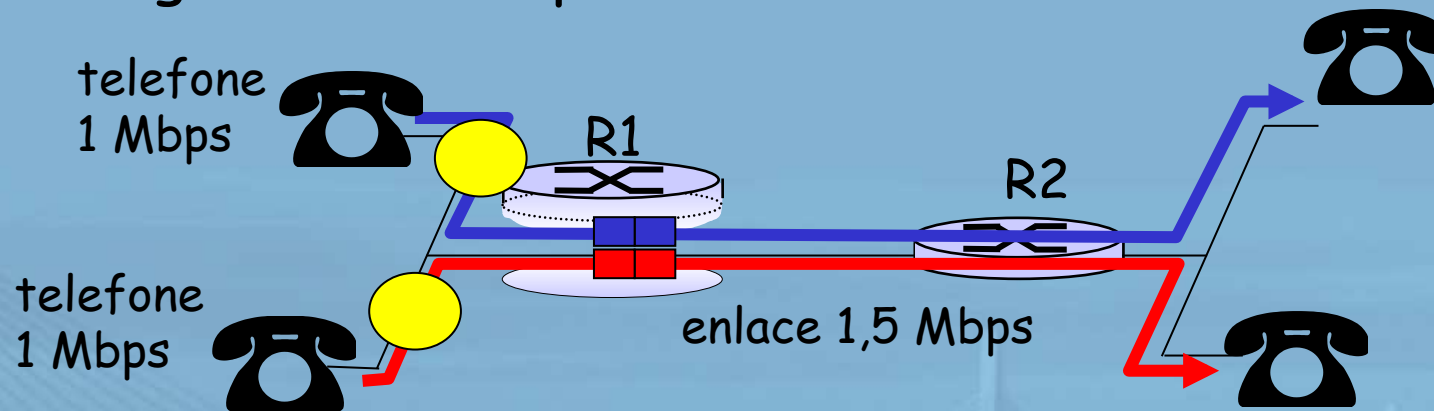
- ❑ **Repasse acelerado:** taxa de saída do pacote de uma classe igual ou superior à taxa especificada
 - enlace lógico com uma taxa mínima garantida
- ❑ **Repasse assegurado:** 4 classes de tráfego
 - cada uma com quantidade mínima de largura de banda garantida
 - cada uma com três partições de preferência de descarte

Capítulo 7: Esboço

- ❑ 7.1 Aplicações de rede multimídia
- ❑ 7.2 Áudio e vídeo de fluxo contínuo armazenados
- ❑ 7.3 Fazendo o melhor possível com o serviço de melhor esforço
- ❑ 7.4 Protocolos para aplicações interativas em tempo real - RTP, RTCP, SIP
- ❑ 7.5 Fornecendo classes de serviço múltiplas
- ❑ 7.6 Fornecendo garantias de qualidade de serviços

Princípios para garantias de QOS (mais)

- ❑ *Fato básico da vida: não pode admitir demandas de tráfego além da capacidade do enlace*



Princípio 4

Admissão de chamada: fluxo declara suas necessidades, rede pode bloquear chamada (p. e., sinal ocupado) se não puder atender as necessidades

Serviços integrados da IETF

- ❑ arquitetura para fornecer garantias de QoS em redes IP para sessões de aplicação individual
- ❑ reserva de recursos: roteadores mantêm informações de estado (tipo VC) de recursos alocados, requisições de QoS
- ❑ admitir/negar novas requisições de estabelecimento de chamada:

Pergunta: O fluxo recém-chegado pode ser admitido com garantias de desempenho enquanto não violar garantias de QoS feitas aos fluxos já admitidos?

Admissão de chamada

Sessão que chega precisa:

- ❑ declarar seu requisito de QoS
 - **Rspec**: define a QoS sendo requisitada
- ❑ caracterizar tráfego que enviará para rede
 - **Tspec**: define características de tráfego
- ❑ protocolo de sinalização: necessário para executar Rspec e Tspec aos roteadores (onde a reserva é exigida)
 - **RSVP** (Resource Reservation Protocol) - Protocolo para sinalizar explicitamente as necessidades de QoS de uma aplicação na rede

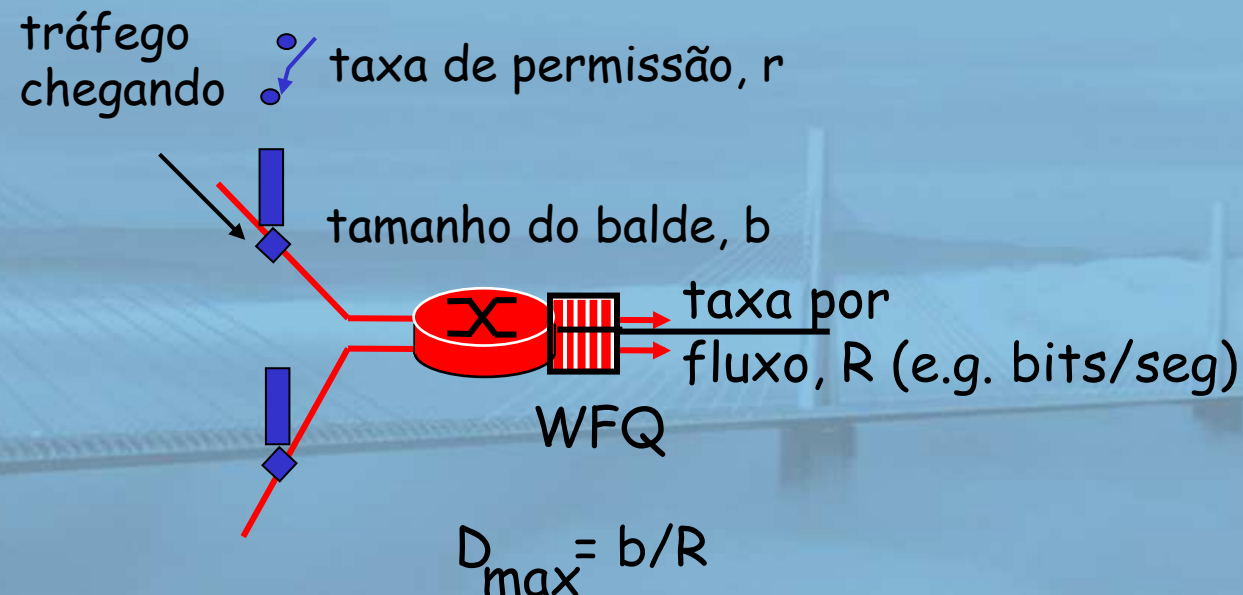
Intserv da QoS: Modelos de serviço [rfc2211, rfc 2212]

Serviço garantido:

- ❑ chegada de tráfego no pior caso: origem regulada por "leaky-bucket"
- ❑ *limite* simples matematicamente comprovável) sobre os atrasos [Parekh 1992, Cruz 1988]

Carga de serviço controlada:

- ❑ "uma qualidade de serviço próxima da QoS, que algum fluxo receberia de um elemento de rede não carregado."



Sinalização da Internet

encaminhamento sem
conexão (sem estado)
por roteadores IP

+

serviço de
melhor
esforço

=

nenhum protocolo
de sinalização de
rede no projeto
inicial do IP

- ❑ **Novo requisito:** reservar recursos ao longo do caminho fim a fim (sistema final, roteadores) para QoS nas aplicações de multimídia
- ❑ **RSVP:** Resource Reservation Protocol [RFC 2205]
 - “ ... permitir que usuários comuniquem requisitos à rede de modo robusto e eficaz.” ou seja, sinalização!
- ❑ antigo protocolo Internet Signaling: ST-II [RFC 1819]

Objetivos de projeto do RSVP

1. acomodar **receptores heterogêneos** (largura de banda diferente ao longo dos caminhos)
2. acomodar aplicações diferentes **com diferentes requisitos de recursos**
3. tornar o **multicast um serviço de primeira classe**, com adaptação para inclusão como membro de grupo multicast
4. **aproveitar roteamento multicast/unicast existente**, com adaptação a mudanças nas rotas do unicast/multicast subjacente
5. **controlar overhead de protocolo** para crescimento (no pior caso) linear no número de receptores
6. **projeto modular** para tecnologias subjacentes heterogêneas

RSVP: não...

- ❑ especifica como os recursos devem ser reservados
 - ❑ Em vez disso: um mecanismo para comunicar necessidades
- ❑ determina rotas que os pacotes tomarão
 - ❑ essa é a tarefa dos protocolos de roteamento
 - ❑ sinalização desacoplada do roteamento
- ❑ interage com repasse de pacotes
 - ❑ separação de planos de controle (sinalização) e dados (repasse)

RSVP: visão geral da operação

- ❑ remetentes e receptor se unem a um grupo de multicast
 - feito fora do RSVP
 - remetentes não precisam se unir ao grupo
- ❑ sinalização do remetente à rede
 - *mensagem de caminho*: torna a presença do remetente conhecida aos roteadores
 - *remoção do caminho*: exclui dos roteadores o estado do caminho do remetente
- ❑ sinalização do receptor à rede
 - *mensagem de reserva*: reserva recursos do(s) remetente(s) ao receptor
 - *remoção do caminho*: remove reservas do receptor
- ❑ sinalização da rede ao sistema final
 - erro de caminho
 - erro de reserva

Capítulo 7: Resumo

Princípios

- ❑ classificar aplicações de multimídia
- ❑ identificar serviços de rede que as aplicações precisam
- ❑ fazer o melhor com o serviço de melhor esforço

Protocolos e arquiteturas

- ❑ especificar protocolos para melhor esforço
- ❑ mecanismos para oferecer QoS
- ❑ arquiteturas para QoS
 - múltiplas classes de serviço
 - garantias de QoS, controle de admissão