



Redes Multimídia

Princípios da compressão em imagens

- **Redundância de codificação** (entrópica): acontece quando os níveis de cinza presentes em uma imagem são codificados com mais símbolos do que seria necessário.
- **Redundância espacial**: ocorre quando os pixels vizinhos em uma imagem são muito parecidos ou iguais.
- **Redundância espectral**: existente em imagens que possuem mais de uma faixa espectral. Correlação entre os valores espectrais em uma mesma posição da matriz de pixels de cada uma das bandas.
- **Redundância de psicovisual**: o olho humano não possui a mesma percepção para as diferentes informações visuais.

Métodos de codificação de imagem

- **Modos de cor de imagem**

- **RGB (24 bits):** sistema que utiliza três cores por pixel, comportando assim até 16,7 milhões de cores. Cada cor tem sua representação feita em 8 bits.
- **CMYK (32 bits):** usa quatro cores por pixel. Cada cor é representada por 8 bits. No CMYK tem como base a qualidade da luz absorvida por essas cores quando impressas em um papel.
- **Escala de cinza (8 bits):** usa um byte por nível de cinza.
- **Preto e branco (1 bit):** cada pixel é representado por um bit.
- **Cor indexada (1 a 8 bits):** conhecido como modo de 256 cores. Cada pixel assume um valor de cor em uma paleta de 256 cores.

Técnicas de compressão

- Seja X a imagem original (matriz de entrada), H a função de codificação e Y o resultado final do processo.

$$Y = H(X)$$

- **Codificadores sem perda de informação:** conhecidos como codificadores universais, possuem a característica de manter a integridade da informação codificada. $Y \rightarrow X$
- **Codificadores com perda de informação:** conhecidos como quantizadores, possuem taxas maiores de compressão ao custo da perda de fidelidade da informação.

Transformada Discreta de Cosseno (DCT)

- Converte um bloco de pixels em uma matriz de coeficientes, transformando a imagem do domínio espacial para o domínio da frequência.
- Blocos podem ser de 4x4, 8x8, 16x16, 32x32 ou 64x64.
- Padrão JPEG – *Joint Photographic Experts Group*.
 - (ISO/IEC 10981-1): 8x8.

Transformada Discreta de Cosseno (DCT)

- DCT é definida por:

$$C(i, j) = \alpha(i) \cdot \alpha(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cdot \cos\left(\frac{(2x+1) \cdot i\pi}{2N}\right) \cos\left(\frac{(2y+1) \cdot j\pi}{2N}\right)$$

$$\alpha(x) = \begin{cases} \sqrt{\frac{1}{N}}, & \text{se } x = 0 \\ \sqrt{\frac{2}{N}}, & \text{se } x = 1, 2, \dots, N-1 \end{cases}$$

- f representa os valores da imagem original, N a dimensão da imagem ou bloco.

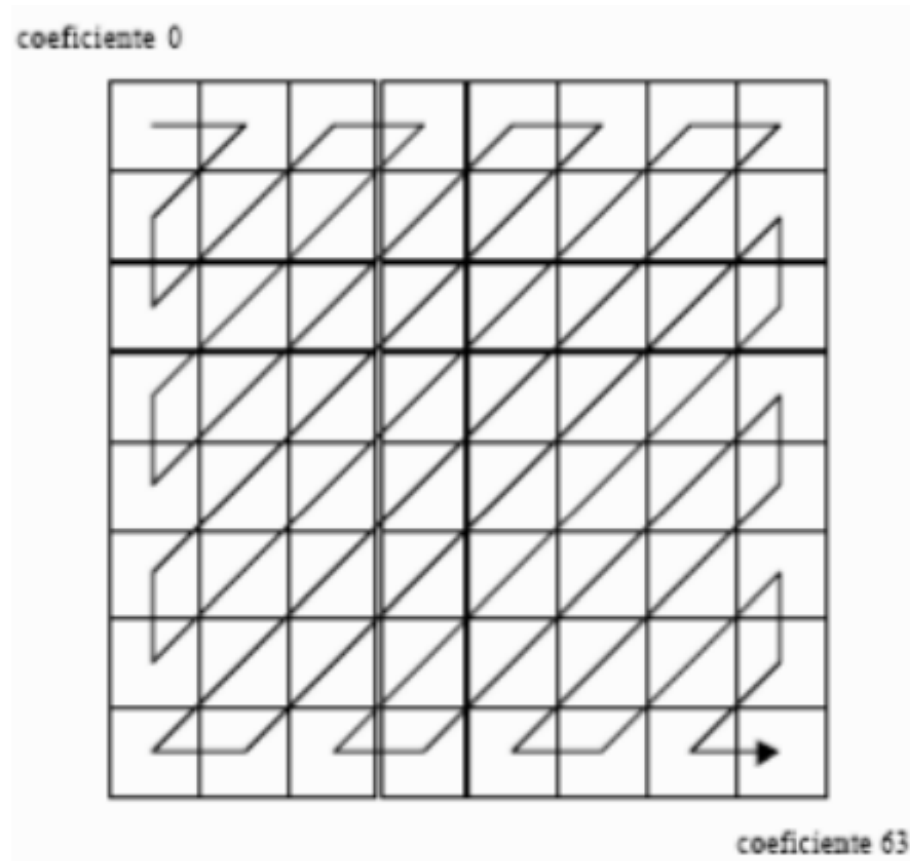
Transformada Discreta de Cosseno (DCT)

- DCT⁻¹

$$f(x, y) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \alpha(i) \cdot \alpha(j) \cdot C(i, j) \cos\left(\frac{(2x+1) \cdot i\pi}{2N}\right) \cos\left(\frac{(2y+1) \cdot j\pi}{2N}\right)$$

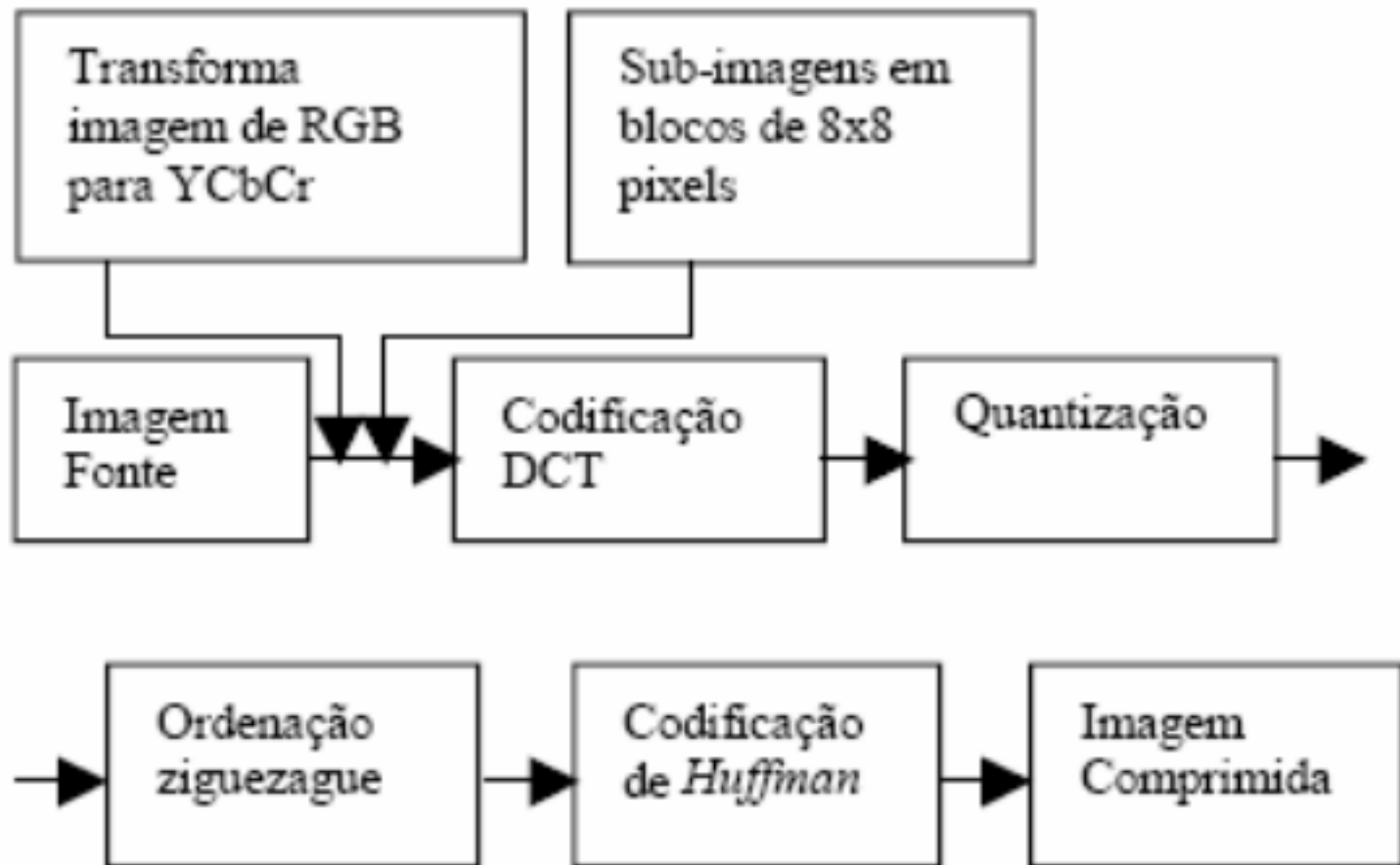
Processo de codificação JPEG

- Ordenação em zig-zag dos coeficientes DCT.



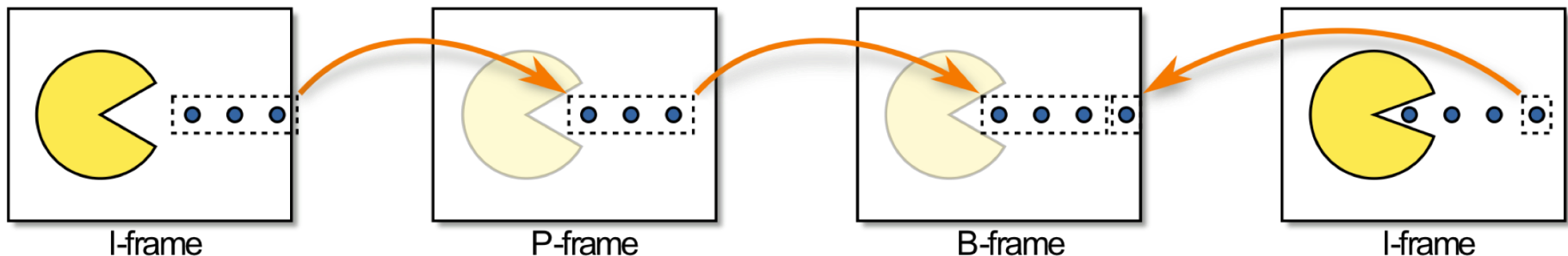
Processo de codificação JPEG

- Operação de compressão JPEG.



Princípios de compressão de vídeo

- **Redundância espacial (intra-quadro):** codificado com JPEG.
- **Redundância temporal (inter-quadro):** em um conjunto de quadros, pequenos movimentos acontecem de um quadro para outro.
 - Valores típicos para filmes: 20 quadros/s.



Propriedades de vídeo

- Talvez a característica mais destacada do vídeo seja sua alta taxa de bits.
- O vídeo distribuído pela Internet costuma variar de **100 kbps** para videoconferências de baixa qualidade até mais de **3 Mbps** para os filmes de fluxo de vídeo com alta definição.
- Outra característica importante do vídeo é que ele pode ser compactado, compensando assim a qualidade com a taxa de bits.
- Também podemos usar a compactação para criar múltiplas versões do mesmo vídeo.

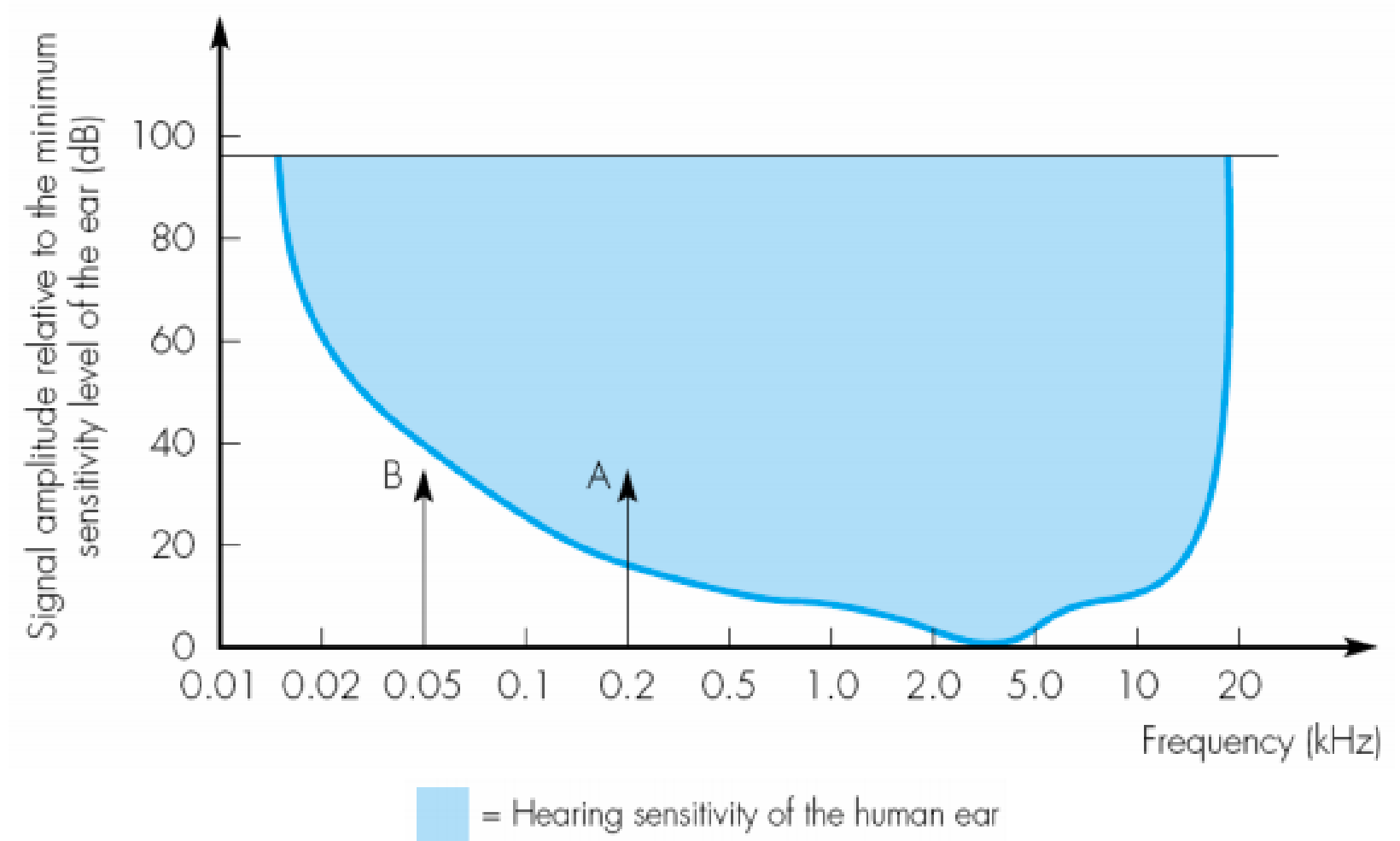
Áudio

- Sinal de voz: amostragem de 50Hz a 10kHz.
- Música: amostragem de 15Hz a 20kHz.
- Codificadores:
 - PCM: ITU-T G711, G721, G722, G726
- Modelos perceptual – modelo psicoacústico:
 - Mascaramento de frequências.
 - Mascaramento temporal.
- MPEG Layer 1, Layer 2, Layer 3 (MP3)
- Dolby Digital AC-3
- AAC (MPEG2-AAC e MPEG4-AAC)

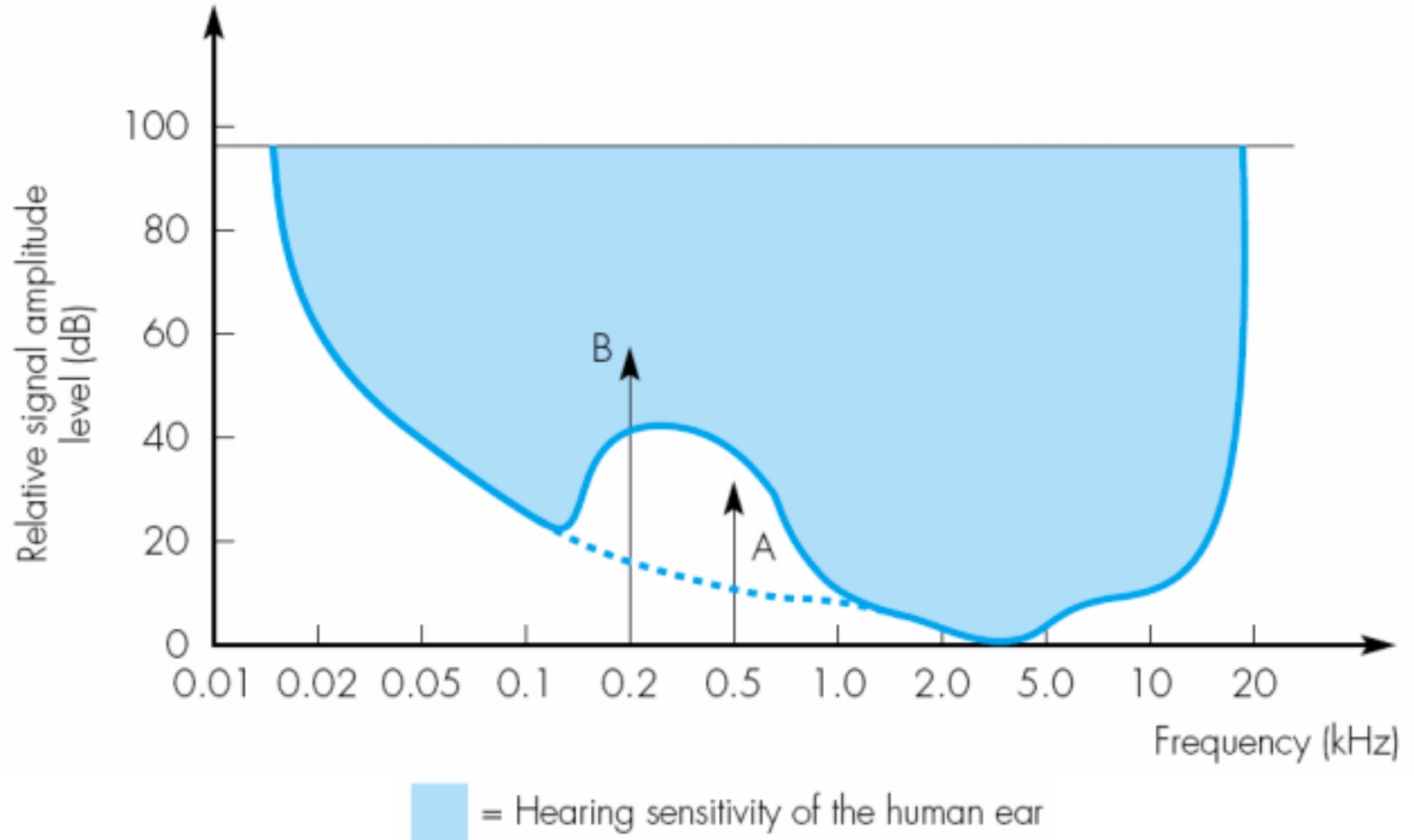
Codificação perceptual

- **Mascaramento de frequências:** um sinal mais forte pode reduzir a sensibilidade de ouvir outros sinais próximos na frequência.
- **Mascaramento temporal:** quando o ouvido escuta um som alto, ele leva um tempo curto para conseguir escutar um som mais baixo.

Codificação perceptual



Codificação perceptual



Propriedades de áudio

- O áudio digital tem requisitos de largura de banda muito menores do que o vídeo.
- Uma técnica de compactação popular para a música estéreo com qualidade quase de CD é **MPEG 1 layer 3**, mais conhecida como MP3.
- Embora as taxas de bit de áudio sejam em geral muito menores do que as de vídeo, os usuários costumam ser muito mais sensíveis a pequenas falha de áudio do que de vídeo.

Tipos de aplicações de redes multimídia

1. Áudio e vídeo de fluxo contínuo armazenados.
2. Voz e vídeo sobre IP interativos.
3. Áudio e vídeo de fluxo contínuo ao vivo.

Áudio e vídeo de fluxo contínuo armazenados

- Muitas empresas de Internet oferecem hoje vídeo de fluxo contínuo, incluindo YouTube (Google), Netflix e Hulu.
- Por algumas estimativas, ele contribui com mais de 50% do tráfego descendente nas redes de acesso à Internet atualmente.
- Ele tem três características distintas importantes:
 1. **Fluxo contínuo.**
 2. **Interatividade.**
 3. **Reprodução contínua.**

Voz e vídeo sobre IP interativos

- A voz interativa em tempo real pela Internet é chamada de telefonia da Internet.
- A maioria dos sistemas interativos por voz e vídeo permite que os usuários criem conferências com três ou mais participantes.
- Voz e vídeo interativos são muito usados na Internet hoje, com as empresas Skype, QQ e Google Talk.
- Aplicações de multimídia **interativas** são **tolerantes à perda**, e **sensíveis a atraso**.

Áudio e vídeo de fluxo contínuo ao vivo

- Aplicações ao vivo, do **tipo de difusão**, normalmente possuem muitos usuários que recebem o mesmo programa de áudio/vídeo ao mesmo tempo.
- A rede precisa oferecer a cada fluxo de multimídia ao vivo uma vazão média que seja maior que a taxa de consumo de vídeo.
- Como o evento é ao vivo, o atraso também pode ser um problema.
- Atrasos de até dez segundos ou mais desde o instante em que o usuário requisita a entrega/reprodução de uma transmissão ao vivo até o início da reprodução podem ser tolerados.

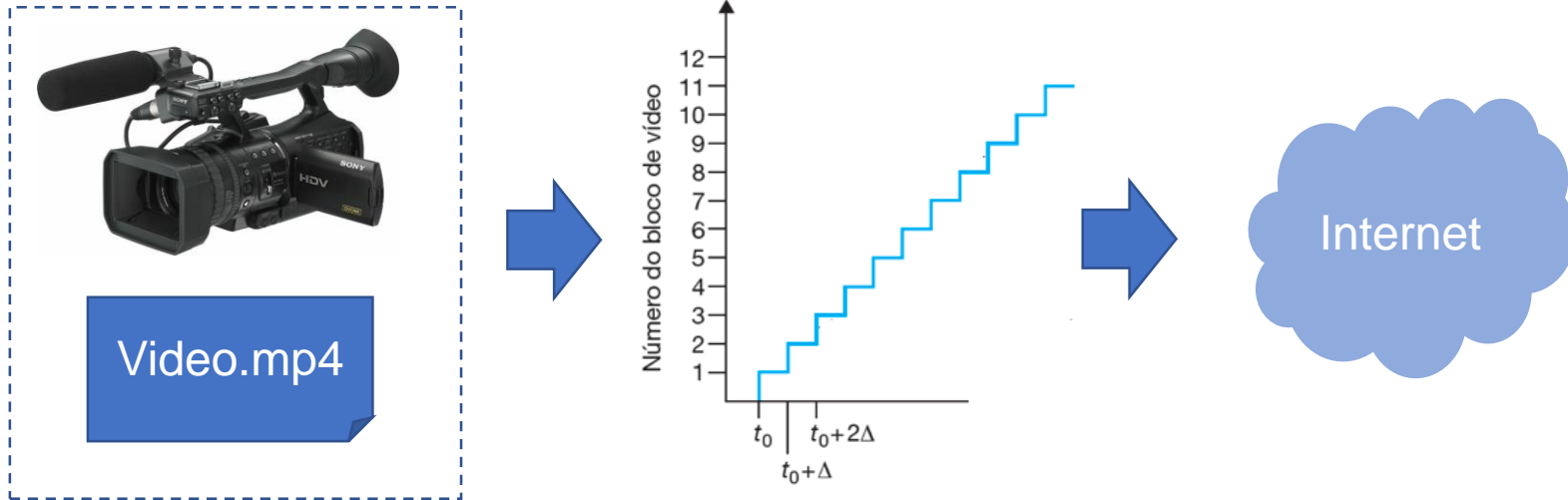
Vídeo de fluxo contínuo armazenado

Vídeo de fluxo contínuo armazenado

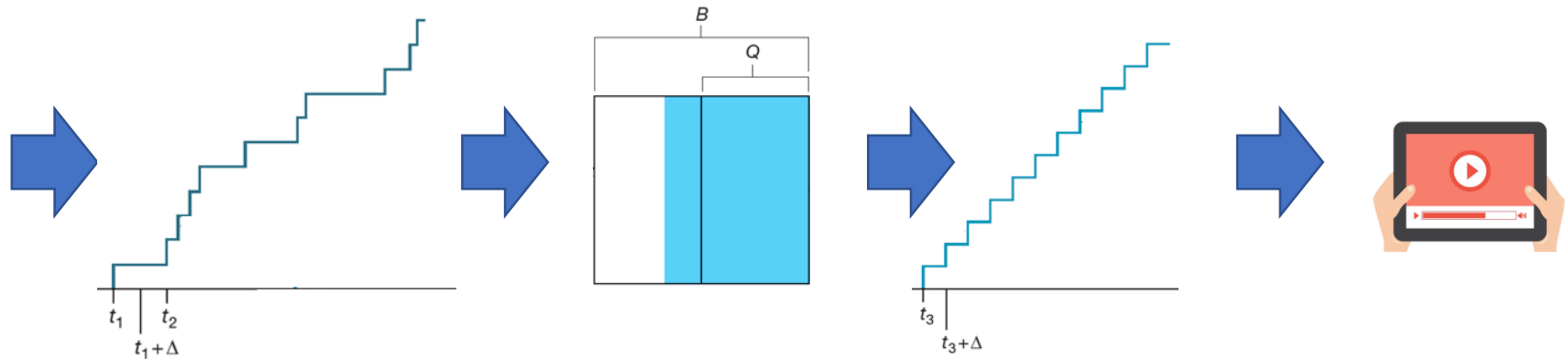
- Os sistemas de vídeo de fluxo contínuo podem ser classificados em três categorias:
 1. UDP de fluxo contínuo,
 2. HTTP de fluxo contínuo e
 3. **HTTP de fluxo contínuo adaptativo.**
- Uma característica comum de todas as três formas de vídeo de fluxo contínuo é o **uso extenso de buffer de aplicação no lado do cliente** para aliviar os efeitos de variar os atrasos de fim a fim e variar as quantidades de largura de banda disponível entre servidor e cliente.

Vídeo de fluxo contínuo armazenado

Transmissão:



Recepção:



Vídeo de fluxo contínuo armazenado

- Atraso de reprodução do cliente no vídeo de fluxo contínuo



UDP de fluxo contínuo

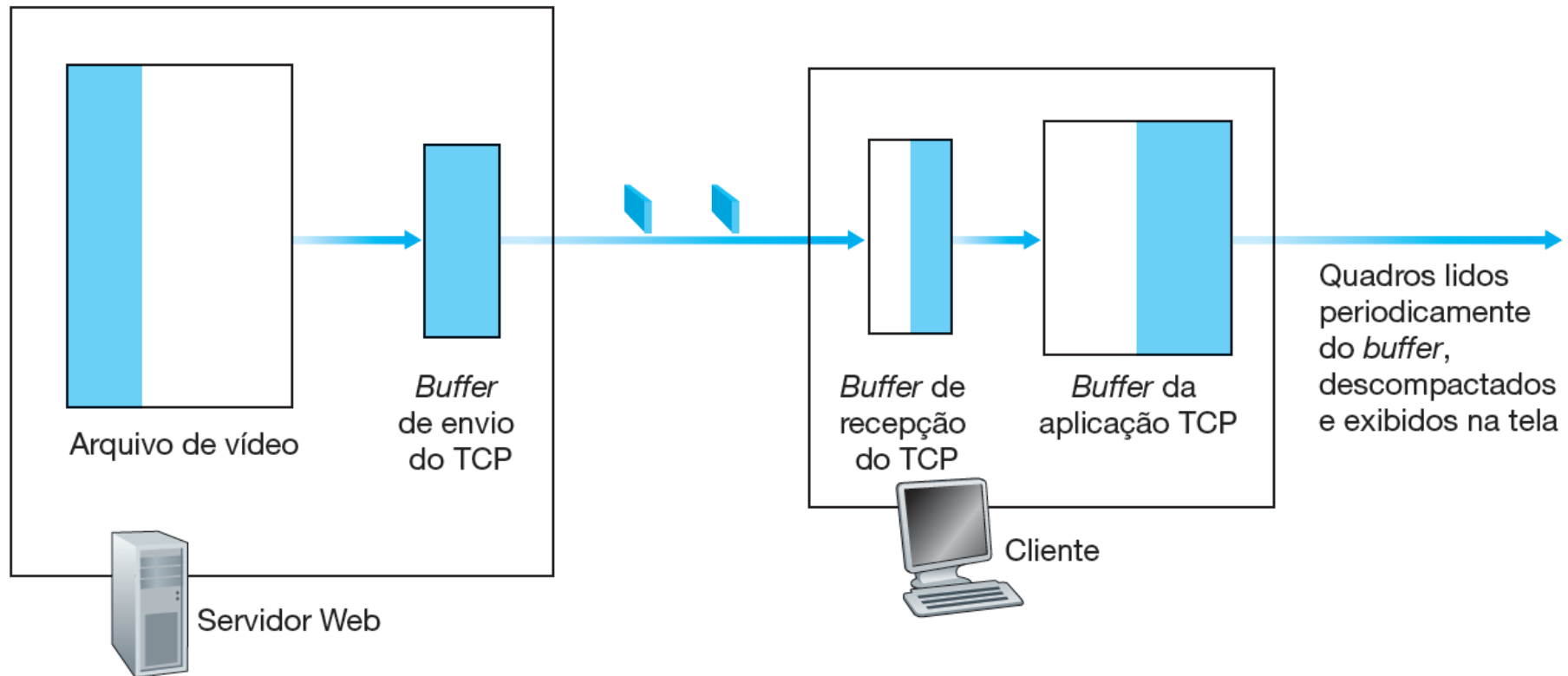
- Com o UDP de fluxo contínuo, o servidor transmite vídeo a uma taxa que corresponde à taxa de consumo de vídeo do cliente.
- Normalmente usa um pequeno buffer no lado do cliente.
- O UDP de fluxo contínuo com taxa constante pode deixar de oferecer reprodução contínua.
- Ele exige um servidor de controle de mídia.
- Muitos firewalls são configurados para bloquear o tráfego UDP.

HTTP de fluxo contínuo

- No HTTP de fluxo contínuo, o vídeo é apenas armazenado em um servidor HTTP como um arquivo comum com uma URL específica.
- O uso do HTTP sobre TCP também permite ao vídeo atravessar firewalls e NATs mais facilmente.
- Vídeos de fluxo contínuo sobre HTTP também deixam clara a necessidade de um servidor de controle de mídia, tal como um servidor RTSP.

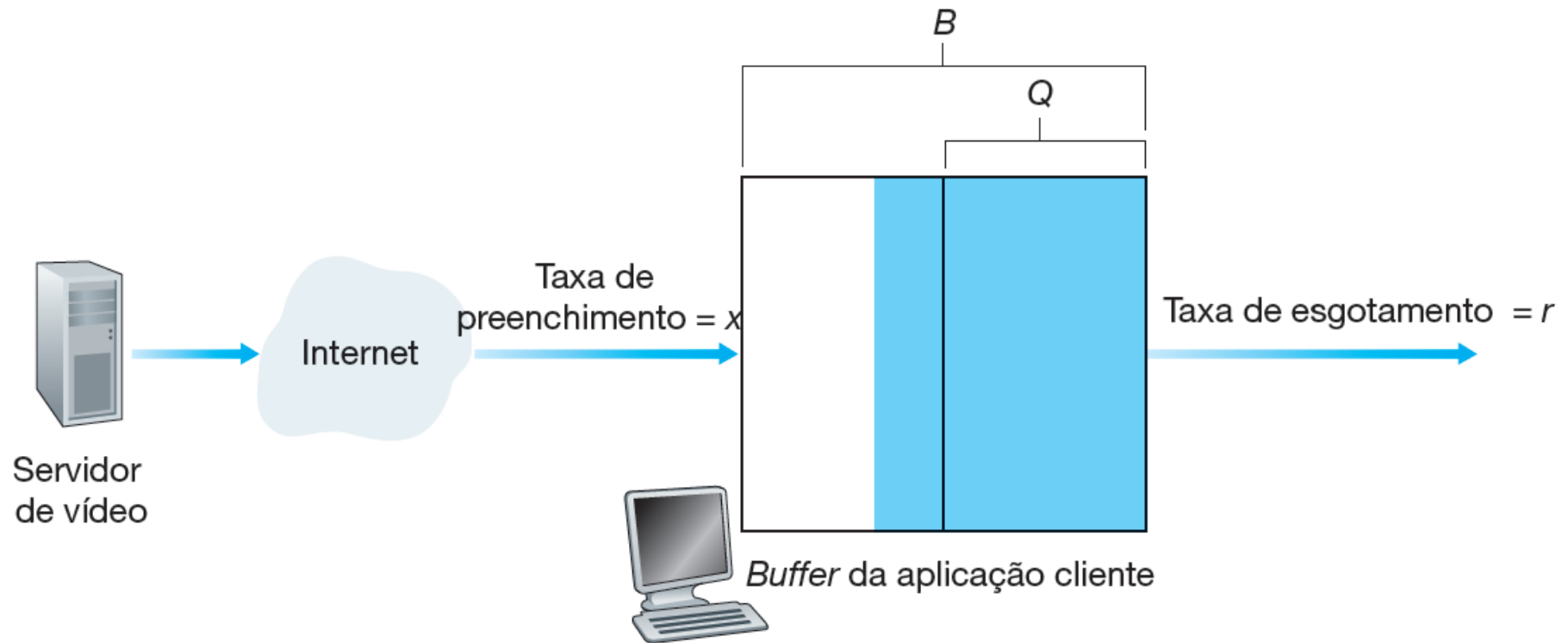
Buffer de aplicação do cliente e buffers TCP

A figura abaixo ilustra a interação entre o cliente e o servidor para HTTP de fluxo contínuo.



Análise do vídeo de fluxo contínuo

Análise do uso de buffer no lado do cliente, para vídeo de fluxo contínuo



Fluxo contínuo adaptativo e DASH

- Pelo DASH, o vídeo é codificado em muitas versões diferentes, cada qual com uma taxa de bits e um diferente nível de qualidade.
- O cliente seleciona diferentes trechos um de cada vez com mensagens de requisição HTTP GET.
- DASH permite aos clientes com diferentes taxas de acesso à Internet fluir em um vídeo por diferentes taxas codificadas.
- Com o DASH, cada versão do vídeo é armazenada em um servidor HTTP, cada um com uma diferente URL.

Redes de distribuição de conteúdo

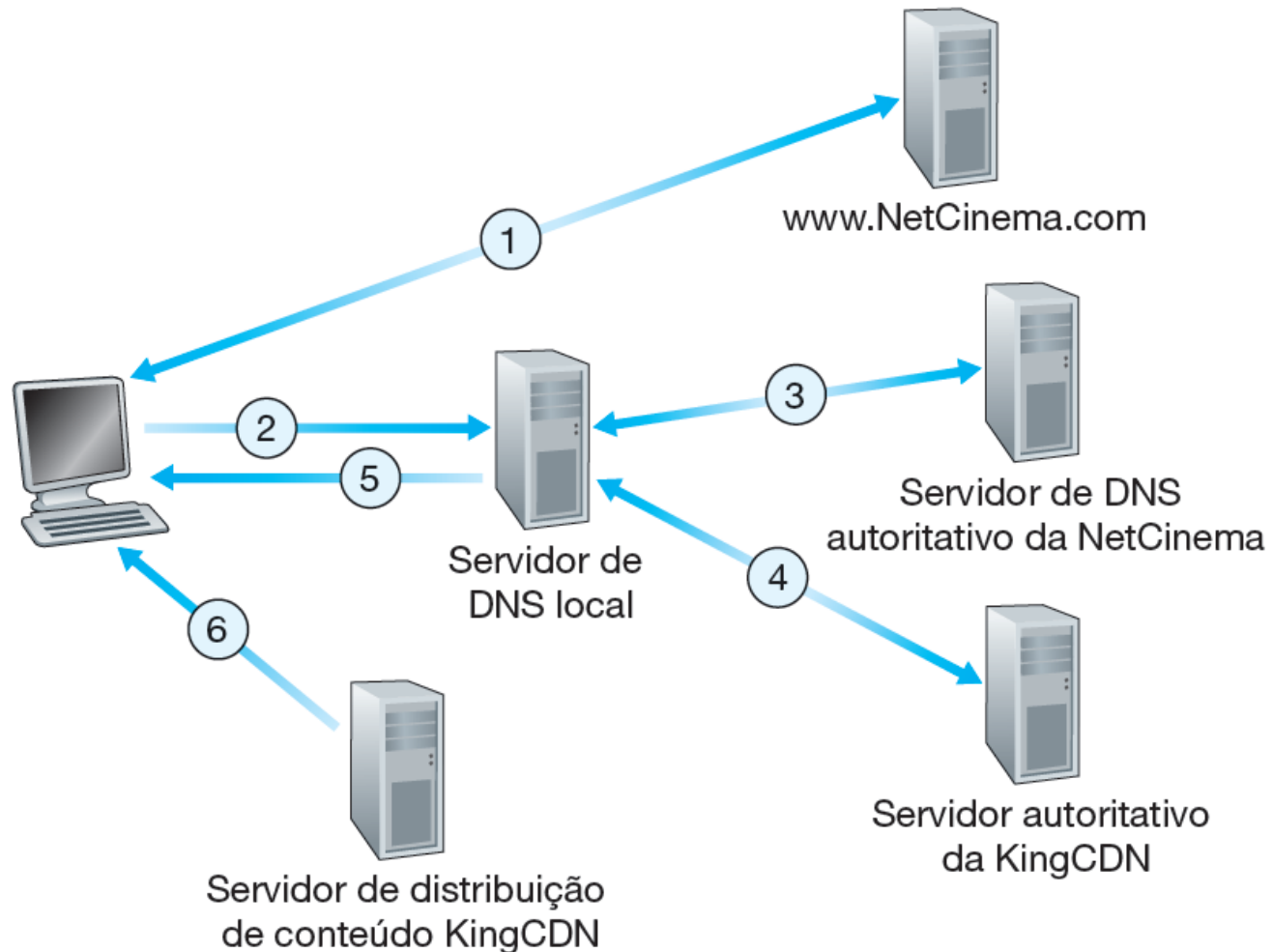
Uma CDN (Rede de Distribuição de Conteúdo):

- Gerencia servidores em múltiplas localidades distribuídas geograficamente,
- Armazena cópias dos vídeos em seus servidores, e
- Tenta direcionar cada requisição do usuário para uma localidade CDN que proporcionará a melhor experiência para o usuário.

Pode ser uma CDN privada ou uma CDN de terceiros.

Operação da CDN

A maioria das CDNs utiliza o DNS para interceptar e redirecionar requisições.



Estratégias de seleção de *cluster*

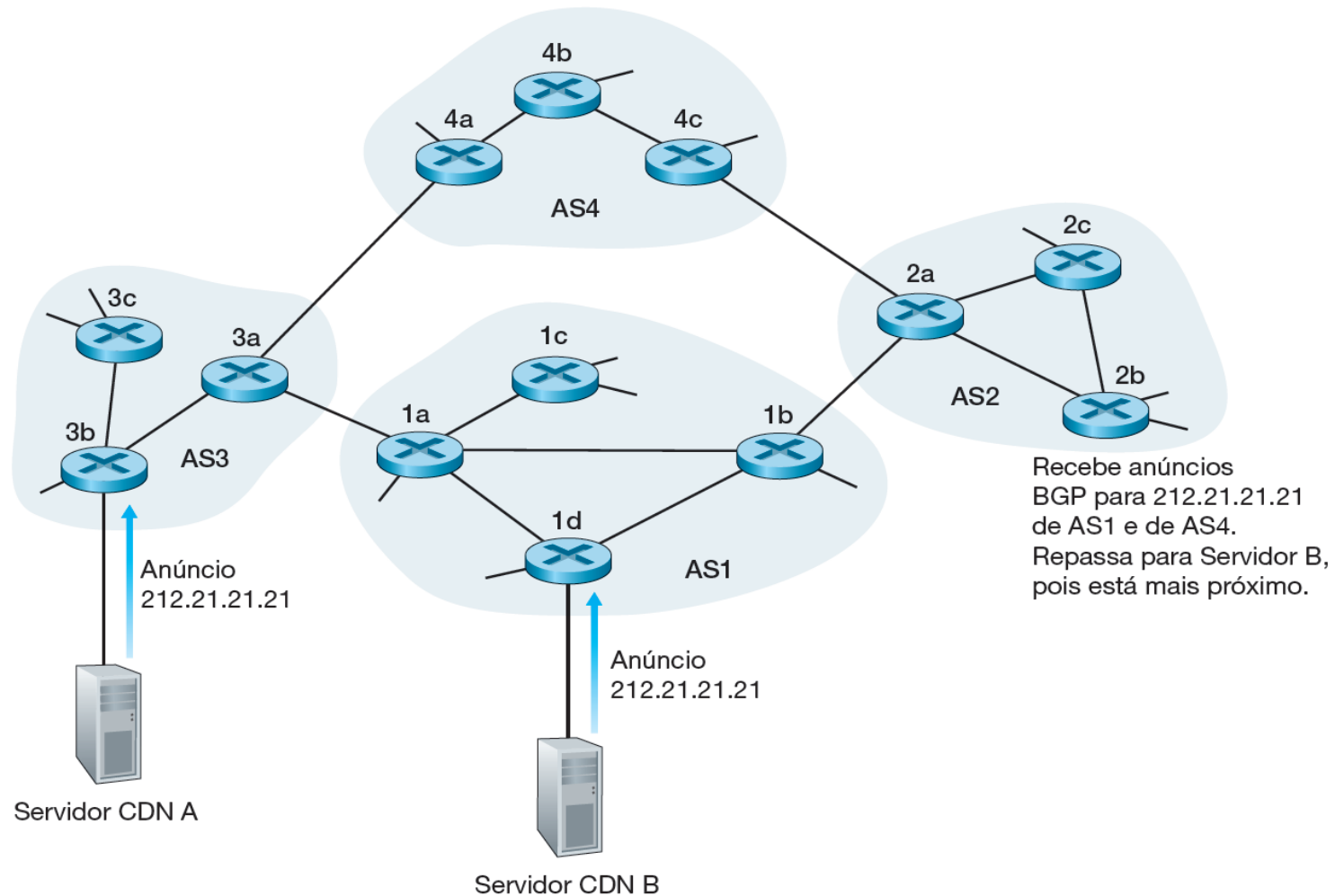
- A estratégia de seleção de *cluster* é um mecanismo para direcionamento dinâmico de clientes para um *cluster* de servidor ou uma central de dados dentro da CDN.
- Uma estratégia simples é associar o cliente ao cluster que está geograficamente mais próximo.
- Para determinar o melhor *cluster* para um cliente baseado nas atuais condições de tráfego, as CDNs podem realizar medições em tempo real do atraso e problemas de baixo desempenho entre seus clusters e clientes.

Estratégias de seleção de *cluster*

- Uma alternativa para medir as propriedades dos caminhos é usar as características do tráfego mais recente entre os clientes e os servidores da CDN.
- Uma abordagem muito diferente para combinar clientes com servidores CDN é utilizar o IP para qualquer membro do grupo.
- A ideia por trás do IP para qualquer membro do grupo é colocar os roteadores na rota da Internet dos pacotes do cliente para o cluster “mais próximo”, como determinado pelo BGP.

Estratégias de seleção de *cluster*

Usando o anycast IP para rotear clientes para o cluster da CDN mais próximo



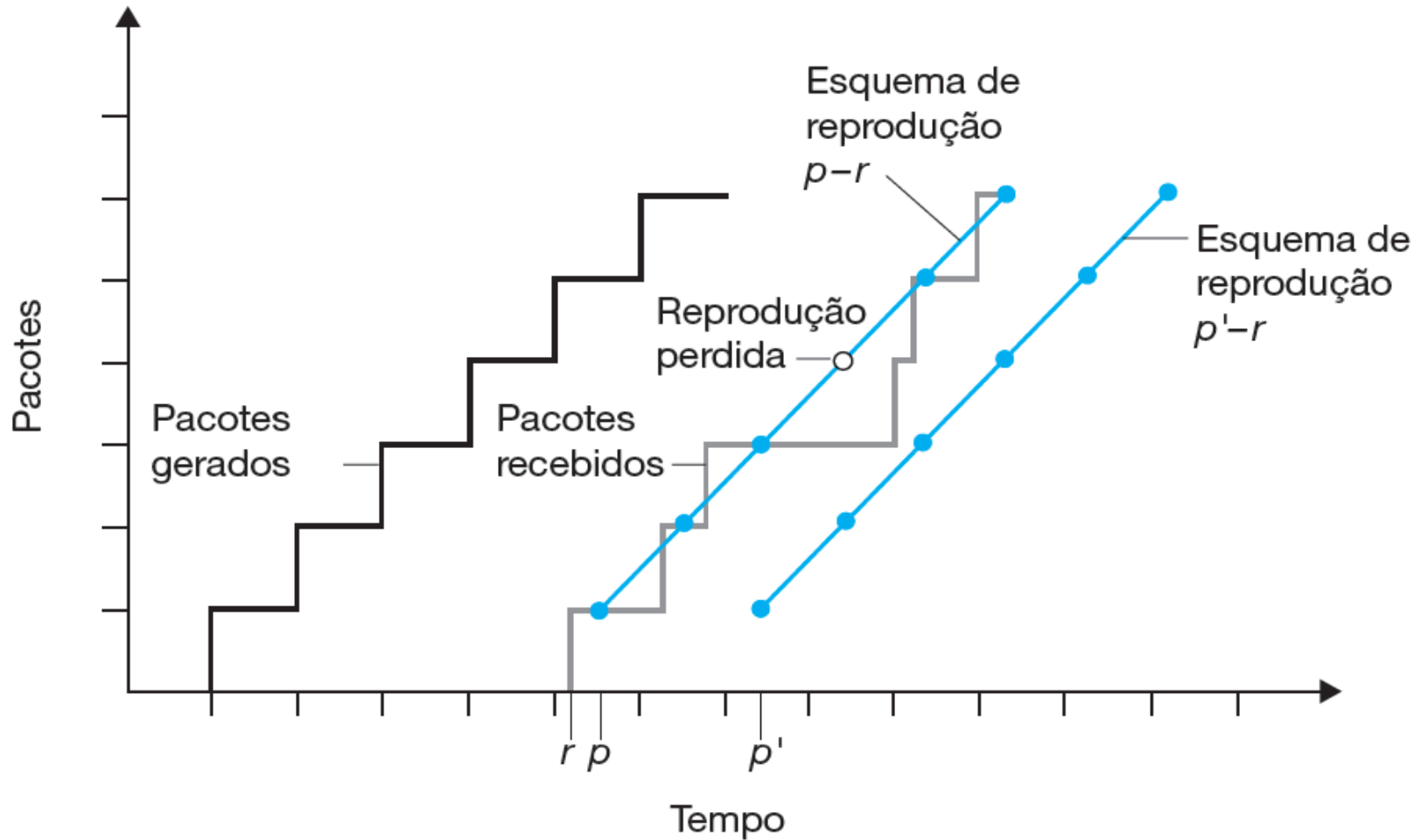
Voice-Over-IP

As limitações de um serviço IP de melhor esforço

- **Perda de pacotes** – A perda pode ser eliminada enviando os pacotes por TCP, em vez de por UDP.
- **Atraso de fim a fim** – é o acúmulo de atrasos de processamento, de transmissão e de formação de filas nos roteadores; atrasos de propagação nos enlaces e atrasos de processamento em sistemas finais.
- **Variação de atraso do pacote** – o tempo decorrido entre o momento em que um pacote é gerado na fonte e o momento em que é recebido no destinatário pode variar de pacote para pacote.

Eliminação da variação de atraso no receptor para áudio

Perda de pacotes para diferentes atrasos por reprodução fixa



Eliminação da variação de atraso no receptor para áudio

Atraso por reprodução adaptativa

- Segue um algoritmo genérico que o receptor pode usar para ajustar seus atrasos de reprodução adaptativamente.
- Para tal, consideremos:
 - t_i = marca de tempo do i -ésimo pacote = o instante em que o pacote foi gerado pelo remetente.
 - r_i = o instante em que o pacote i é recebido pelo receptor.
 - p_i = o instante em que o pacote i é reproduzido no receptor.

Eliminação da variação de atraso no receptor para áudio

- Estimativa do atraso médio para a recepção do i -ésimo pacote.

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

- Estimativa de desvio médio de atraso em relação ao atraso médio esperado:

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

Eliminação da variação de atraso no receptor para áudio

- Se o pacote i for o primeiro de uma rajada de voz, seu tempo de reprodução p_i será computado como:

$$p_i = t_i + d_i + k \cdot v_i$$

sendo k uma constante positiva.

- O ponto de reprodução de qualquer pacote será:

$$p_j = t_j + p_i - t_i$$

Recuperação de perda de pacotes

- Um pacote será considerado perdido se nunca chegar ao receptor ou se chegar após o tempo de reprodução programado.
- Aplicações de VoIP frequentemente usam algum tipo de esquema de prevenção de perda.

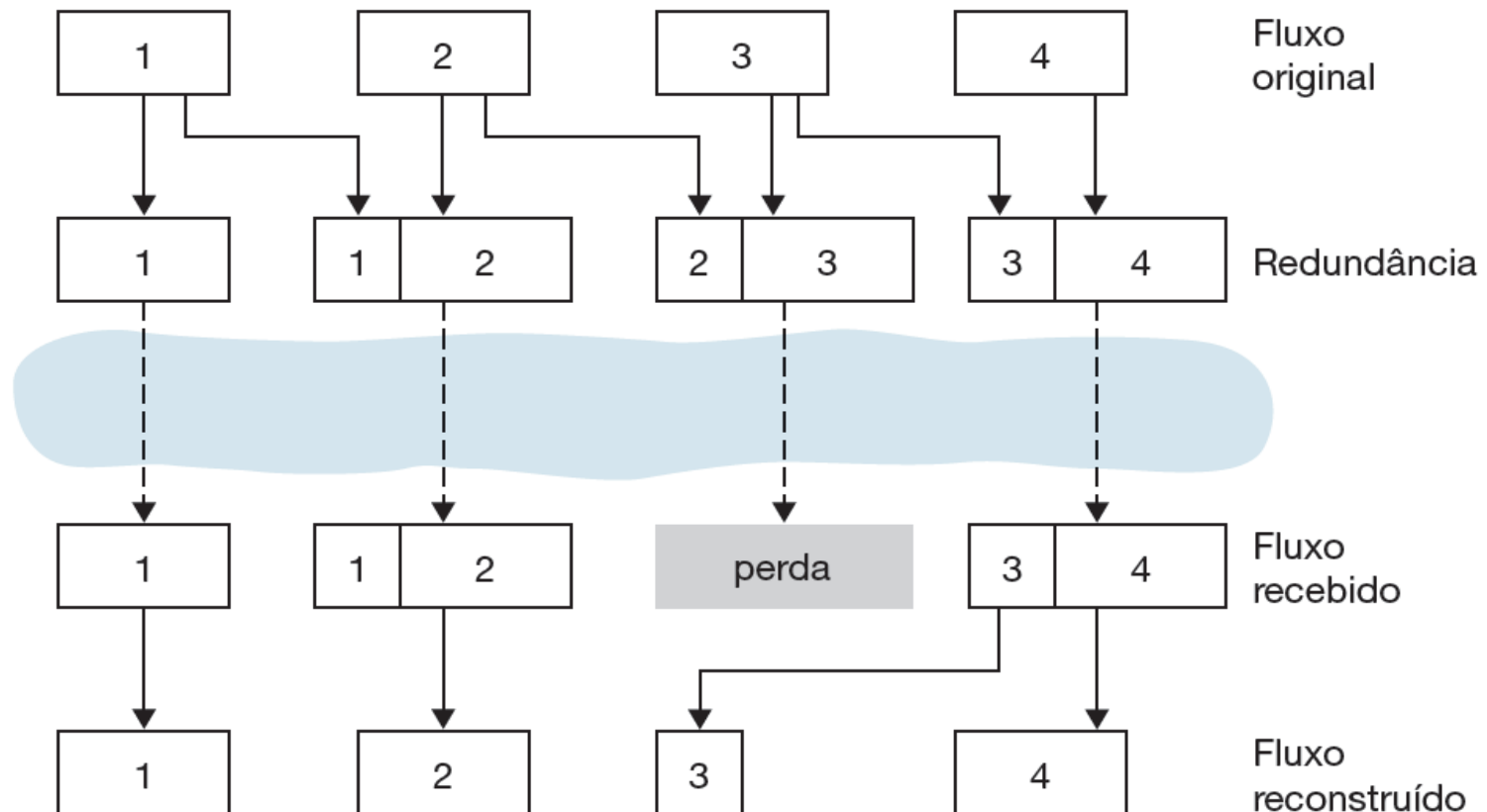
Forward Error Correction (FEC)

- A ideia básica da FEC é adicionar informações redundantes ao fluxo de pacotes original.

Recuperação de perda de pacotes

Forward Error Correction (FEC)

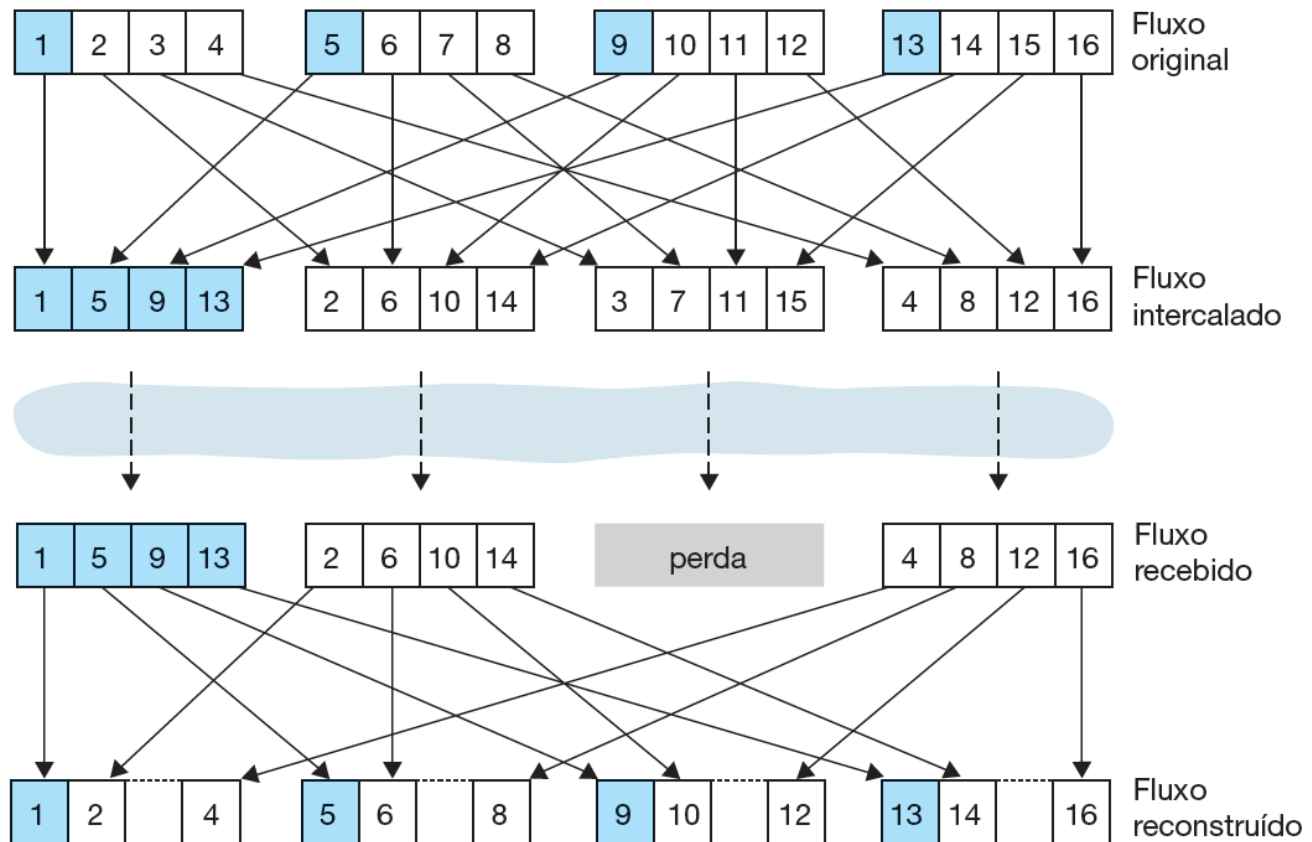
- Dando carona à informação redundante de qualidade mais baixa



Recuperação de perda de pacotes

Intercalação

- Como uma alternativa à transmissão redundante, uma aplicação de VoIP pode enviar áudio intercalado.



Protocolos para aplicações interativas em tempo real

Protocolo de tempo real (RTP)

- O RTP pode ser usado para transportar formatos comuns como PCM, ACC e MP3 para som e MPEG e H.263 para vídeo.
- O protocolo também pode ser usado para transportar formatos proprietários de som e de vídeo.
- Hoje, o RTP é amplamente implementado em muitos produtos e protótipos de pesquisa.
- Também é complementar a outros importantes protocolos interativos de tempo real, entre eles SIP.

Protocolo de tempo real (RTP)

- O RTP comumente roda sobre UDP.
- Se uma aplicação incorporar o RTP, ela interagirá mais facilmente com as outras aplicações de rede multimídia.
- O RTP não fornece nenhum mecanismo que assegure a entrega de dados a tempo nem fornece garantias de qualidade de serviço.
- O RTP permite que seja atribuído a cada origem seu próprio fluxo independente de pacotes RTP.
- Pacotes RTP não são limitados às aplicações individuais.

Protocolo de tempo real (RTP)

- Os quatro principais campos de cabeçalho do pacote RTP são:

Tipo de carga útil	Número de sequência	Marca de tempo	Identificador de sincronização da origem	Campos variados
--------------------	---------------------	----------------	--	-----------------

- Tipos de carga útil de áudio suportados pelo RTP:

Número do tipo de carga útil	Formato de áudio	Taxa de amostragem	Vazão
0	PCM μ -law	8 kHz	64 kbits/s
1	1016	8 kHz	4,8 kbits/s
3	GSM	8 kHz	13 kbits/s
7	LPC	8 kHz	2,4 kbits/s
9	G.722	16 kHz	48-64 kbits/s
14	Áudio MPEG	90 kHz	—
15	G.728	8 kHz	16 kbits/s

Protocolo de tempo real (RTP)

- Alguns tipos de carga útil de vídeo suportados pelo RTP:

Número do tipo de carga útil	Formato de vídeo
26	Motion JPEG
31	H.261
32	Vídeo MPEG 1
33	Vídeo MPEG 2

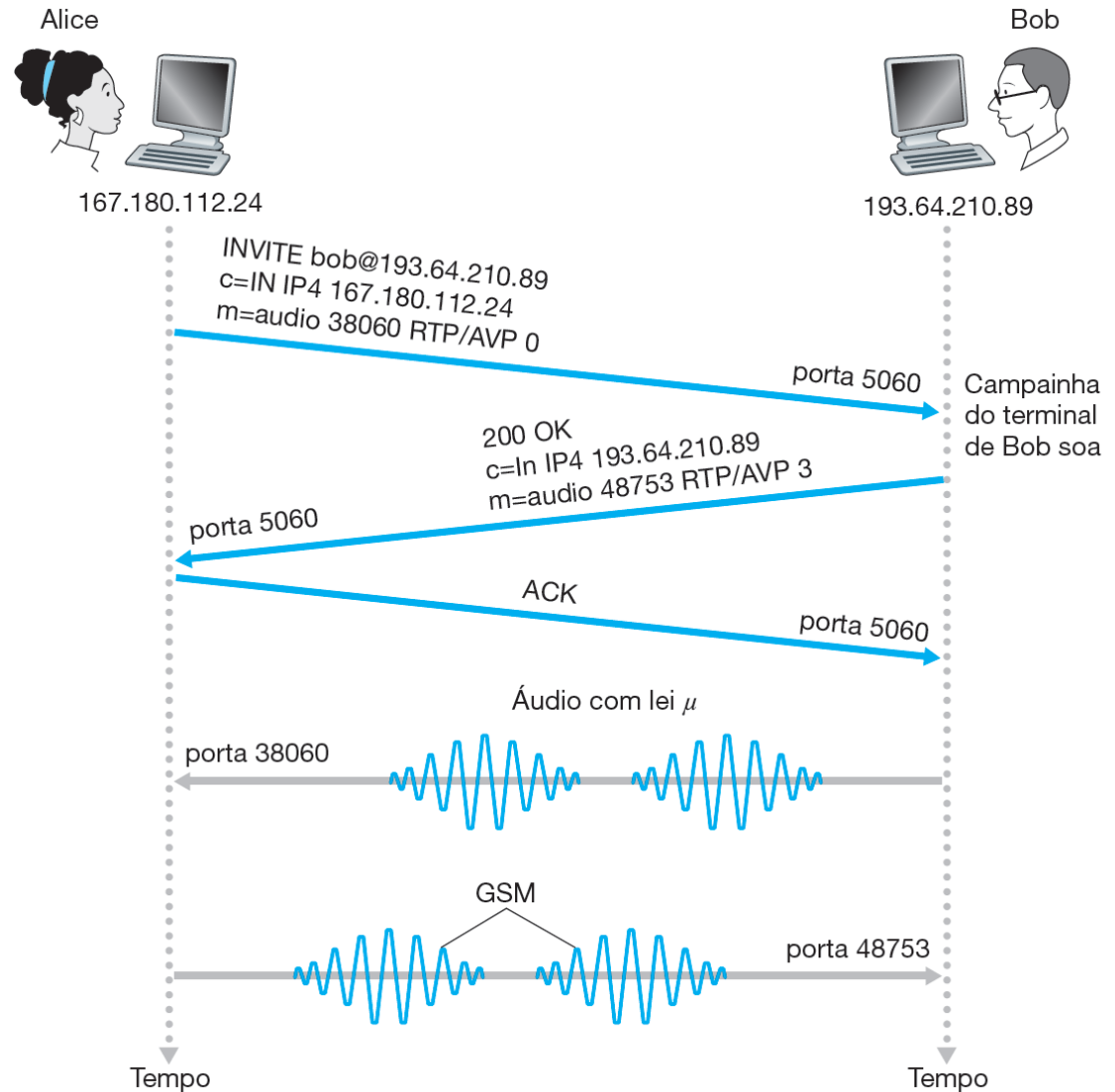
SIP

O Protocolo de Inicialização de Sessão (SIP) é um protocolo aberto e simples, que faz o seguinte:

- Provê mecanismos para **estabelecer chamadas** entre dois interlocutores por uma rede IP.
- Provê mecanismos que permitem a quem chama determinar o **endereço IP atual** de quem é chamado.
- Provê mecanismos para **gerenciamento de chamadas**.

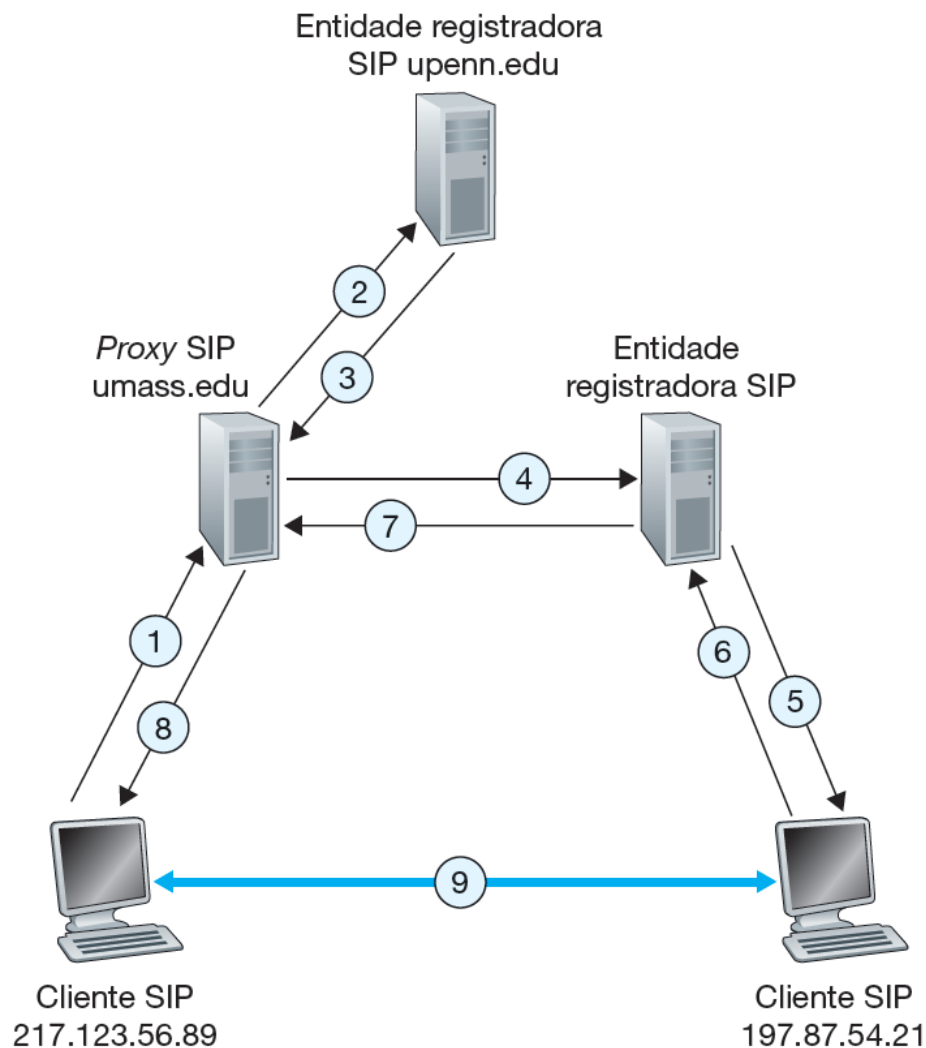
SIP

- Chamada SIP quando Alice conhece o endereço IP de Bob:



SIP

- Inicialização de sessão com proxies e entidades registradoras SIP:



Suporte de rede para multimídia

Suporte de rede para multimídia

- Três técnicas em nível de rede para dar suporte a aplicações de multimídia:

Técnica	Granularidade	Garantia	Mecanismos	Complexidade	Implementação no momento
Obtendo o melhor do serviço de melhor esforço	todo o tráfego tratado igualmente	nenhuma ou flexível	suporte da camada de aplicação, CDNs, sobreposições, provisão de recurso em nível de rede	mínima	em toda a parte
Serviço diferenciado	diferentes classes de tráfego tratadas de formas diferentes	nenhuma ou flexível	marcação, regulação e programação de pacotes	média	alguma
Garantias de qualidade de serviço (QoS) por conexão	cada fluxo de origem-destino tratado de forma diferente	flexível ou rígida, uma vez admitido o fluxo	marcação, regulação e programação de pacotes; admissão e sinalização de chamadas	leve	pouca

Dimensionando redes de melhor esforço

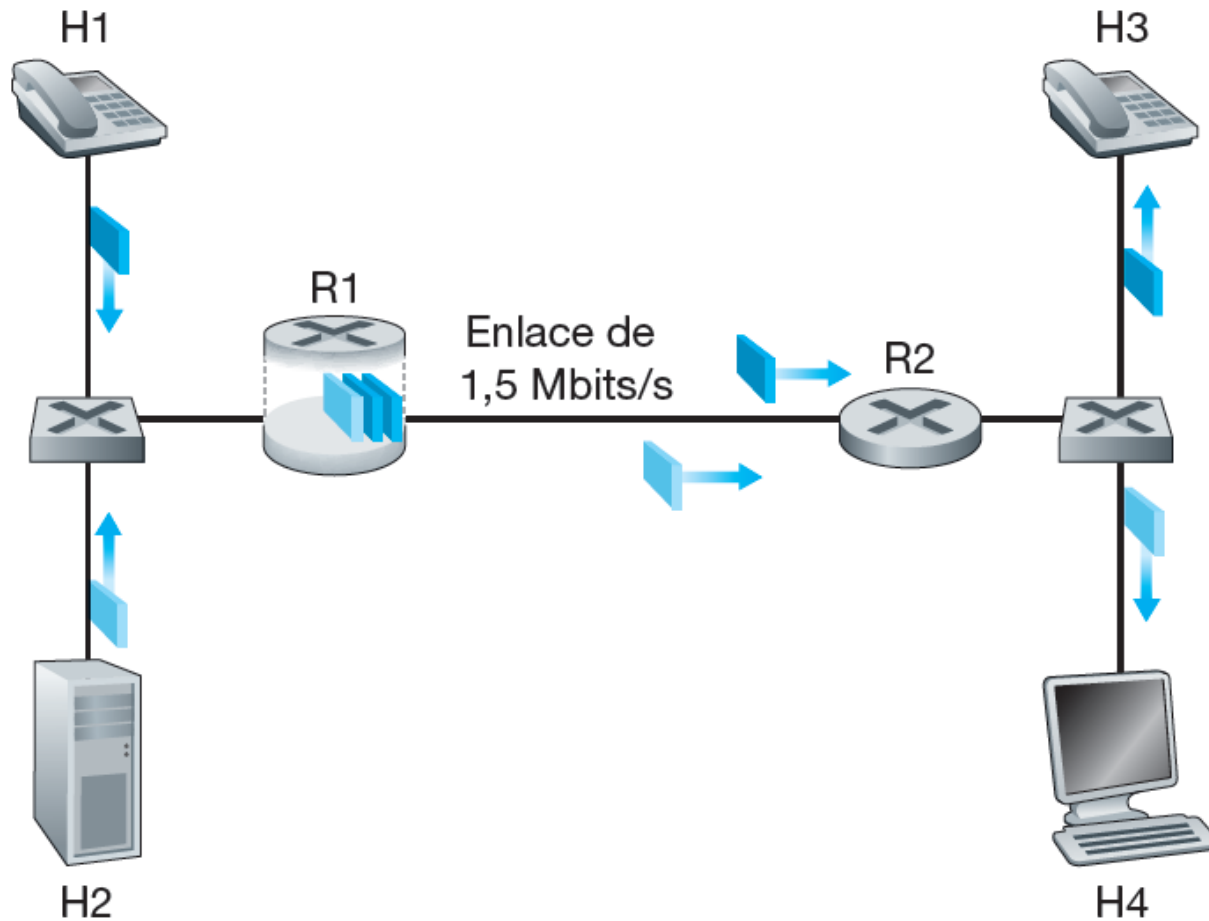
- A questão de como projetar uma topologia de rede para alcançar determinado nível de desempenho de fim a fim é um problema de projeto de redes que muitas vezes é chamado **de dimensionamento de redes**.
- Sabendo que a Internet de melhor esforço de hoje poderia dar suporte para o tráfego de multimídia em um nível de desempenho apropriado, se fosse dimensionada para fazer isso, a questão é por que a Internet de hoje não o faz.
- As respostas são principalmente econômicas e organizacionais.

Fornecendo múltiplas classes de serviço

- Já estamos acostumados com diferentes classes de serviço em nossas vidas diárias.
- É importante observar que esse serviço diferencial é fornecido entre agregações de tráfego, ou seja, entre classes de tráfego, e não entre conexões individuais.
- Até mesmo há três décadas, a visão de fornecer diferentes níveis de serviço a diferentes níveis de tráfego estava evidente.
- No entanto, levou um longo período para que pudéssemos perceber essa visão.

Cenários motivadores

- Competindo aplicações de áudio e HTTP

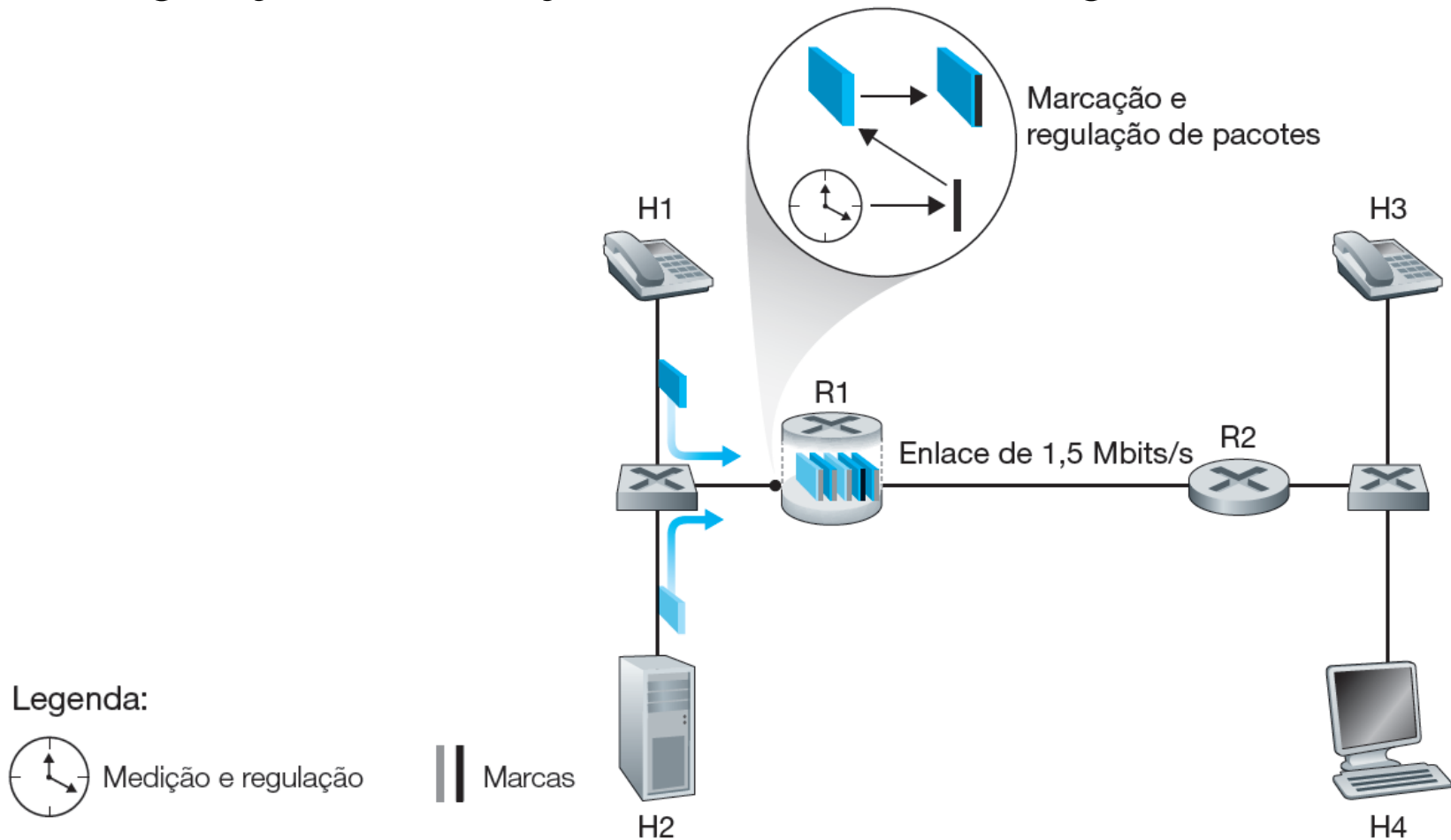


Cenários motivadores

- **Princípio 1:** A marcação de pacotes permite que um roteador faça a distinção de pacotes pertencentes a diferentes classes de tráfego.
- **Princípio 2:** É desejável fornecer algum grau de isolamento de tráfego entre as classes, para que uma classe não seja afetada adversamente por outra classe de comportamento inadequado.
- **Princípio 3:** Ao fornecer isolamento de classes ou fluxos, é desejável que se usem os recursos da maneira mais eficiente possível.

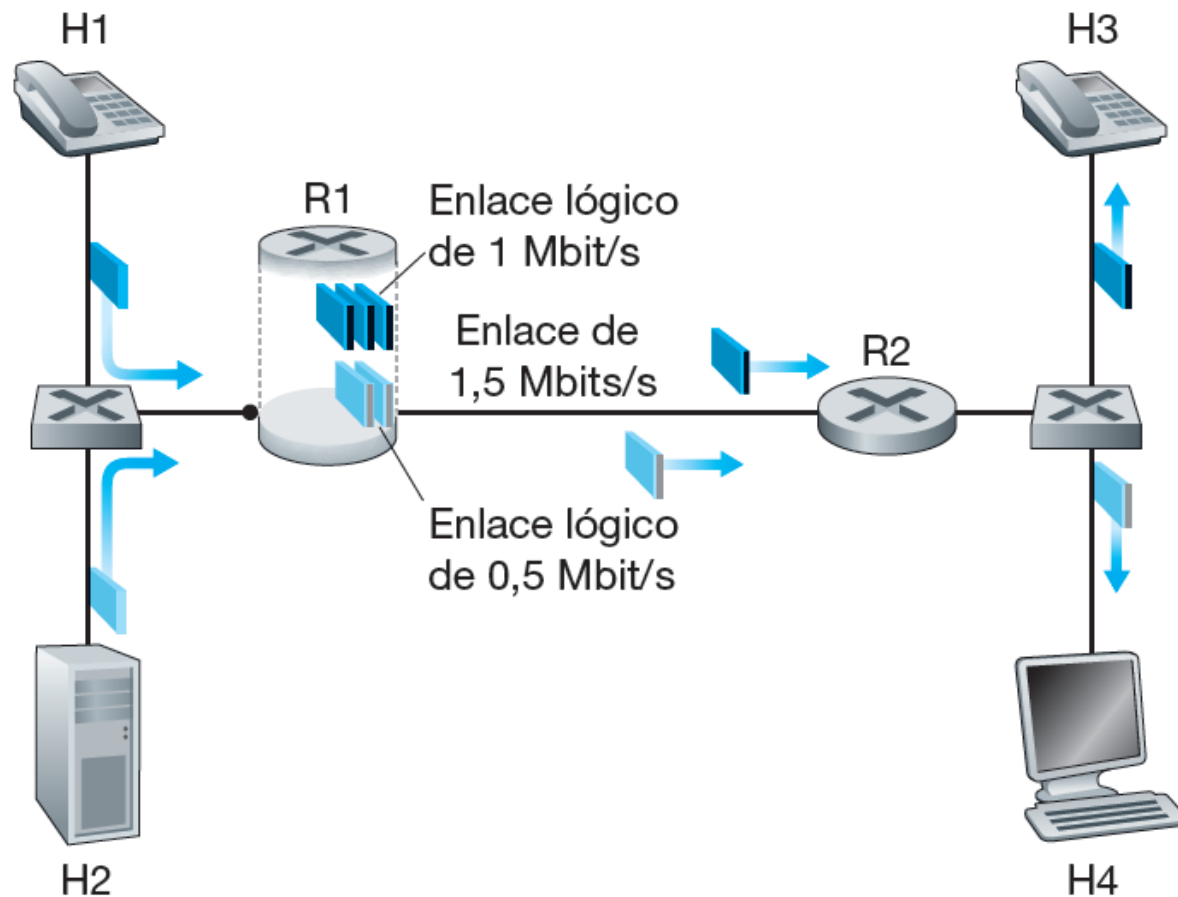
Cenários motivadores

- Regulação (e marcação) das classes de tráfego de áudio e HTTP



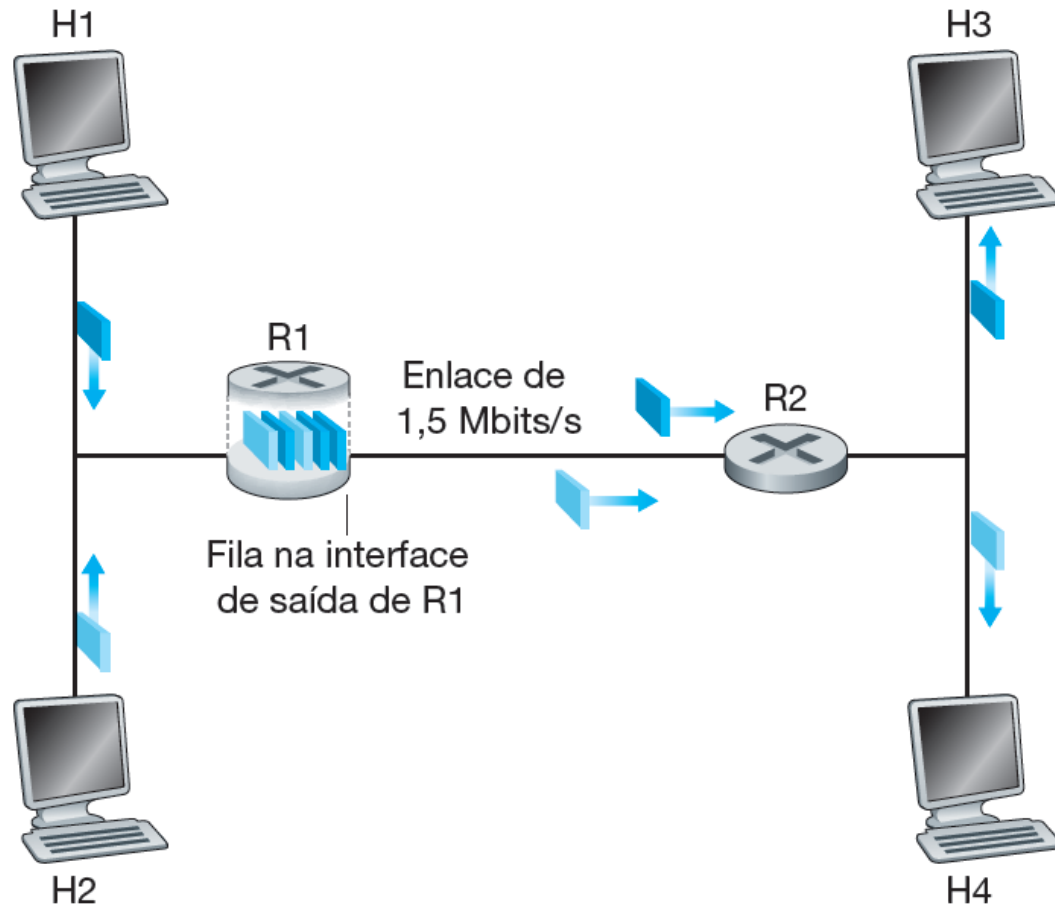
Cenários motivadores

- Isolamento lógico das classes de tráfego de áudio e HTTP



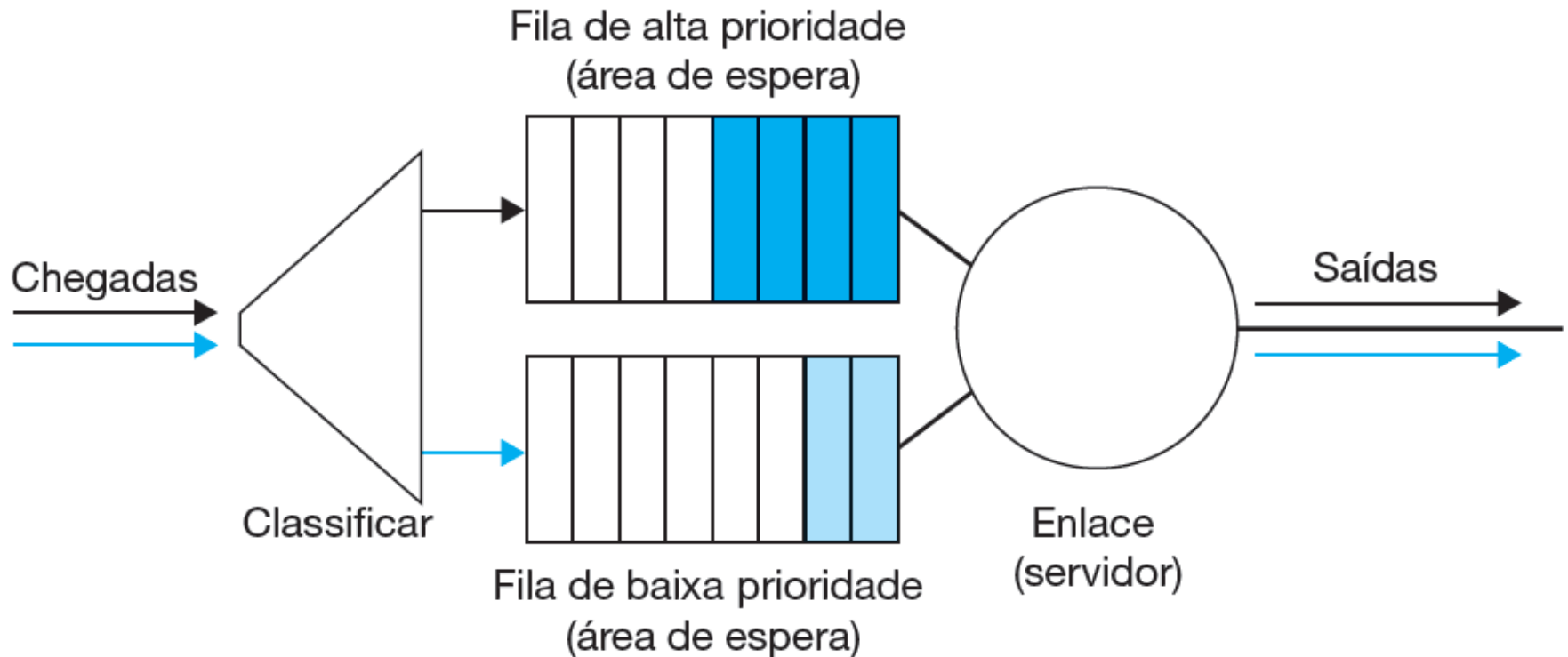
Mecanismos de escalonamento

- Primeiro a entrar/primeiro a sair (FIFO)



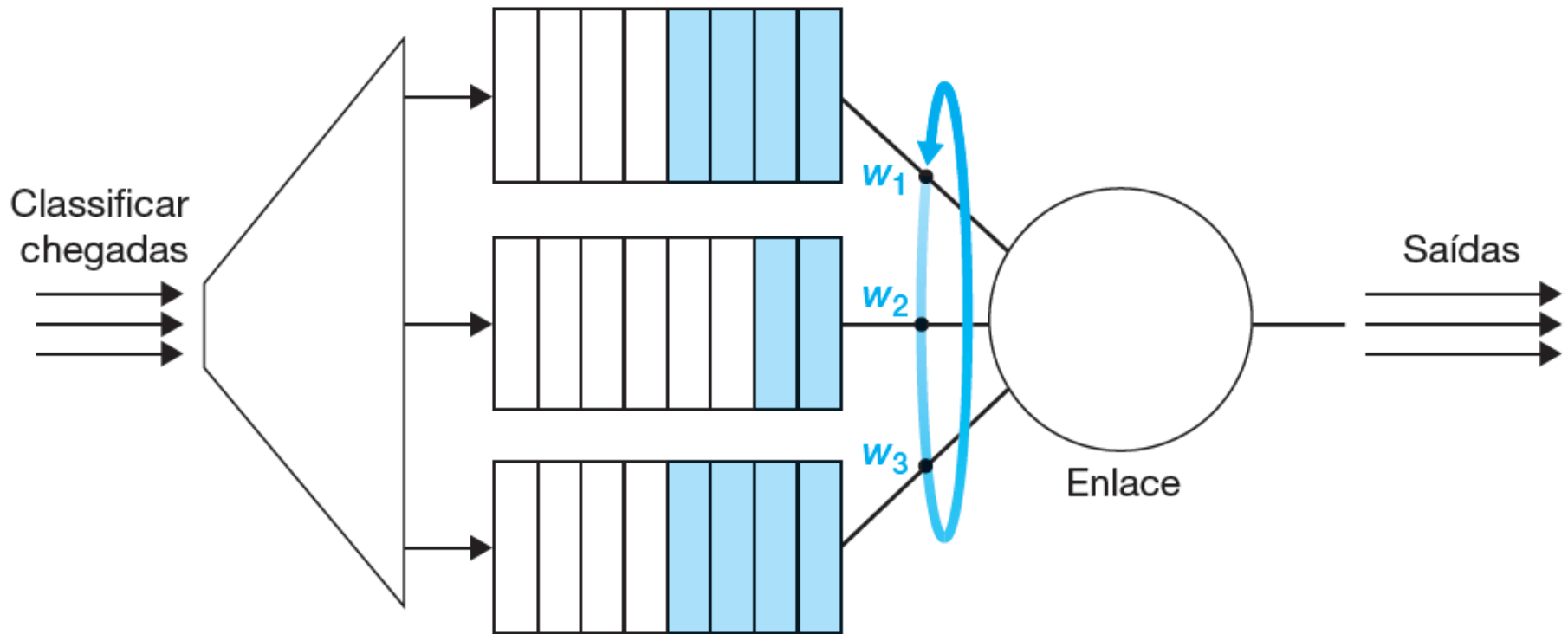
Mecanismos de escalonamento

- Enfileiramento prioritário



Mecanismos de escalonamento

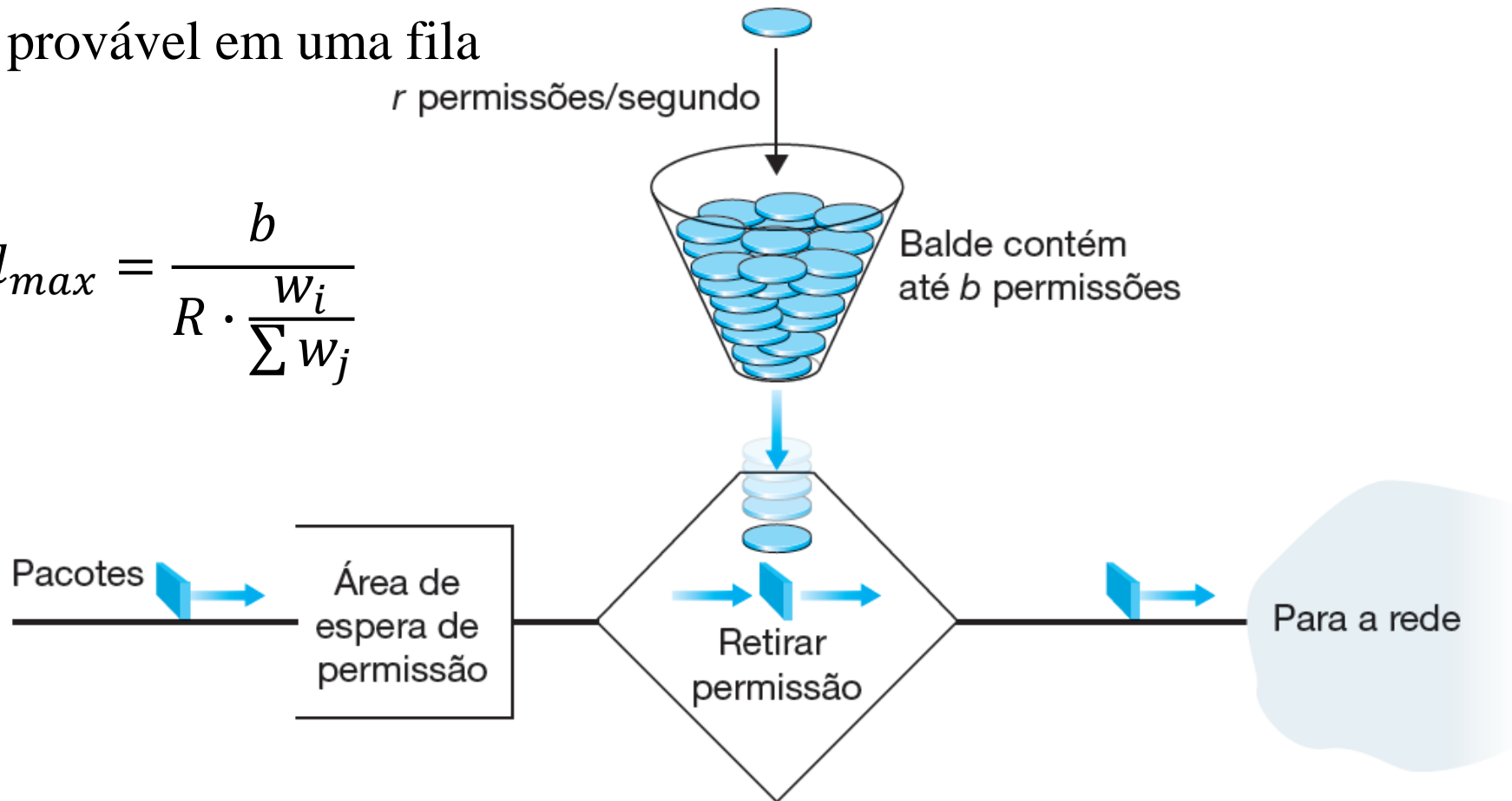
- Varredura cíclica e WQF (enfileiramento justo ponderado)



Mecanismos de escalonamento

- Balde furado + enfileiramento justo ponderado = máximo atraso provável em uma fila

$$d_{max} = \frac{b}{R \cdot \frac{w_i}{\sum w_j}}$$

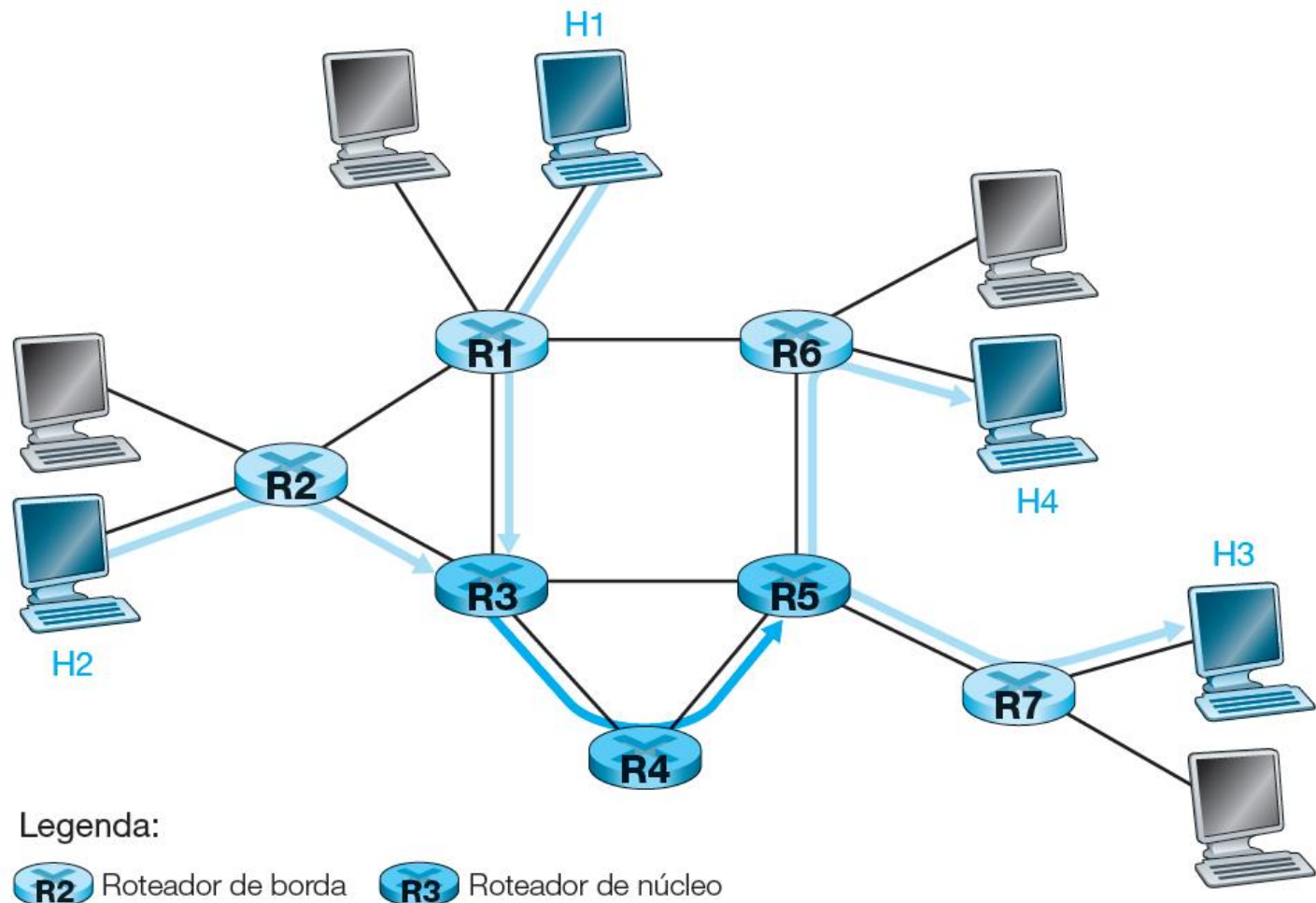


Diffserv

- Diffserv oferece diferenciação de serviço.
- A arquitetura Diffserv consiste em dois conjuntos de elementos funcionais:
 1. *Funções de borda*: classificação de pacotes e condicionamento de tráfego.
 2. *Função central*: envio.

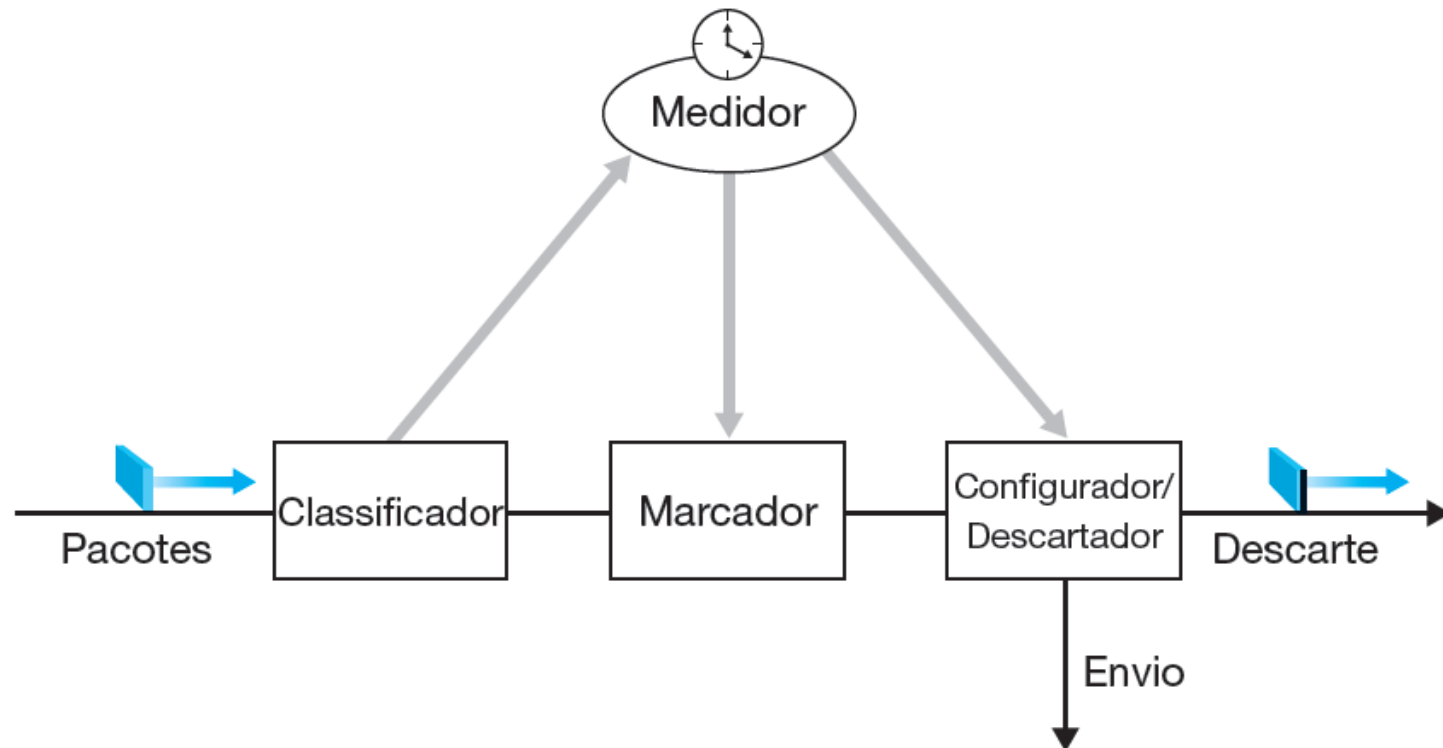
Diffserv

- Exemplo simples de rede Diffserv



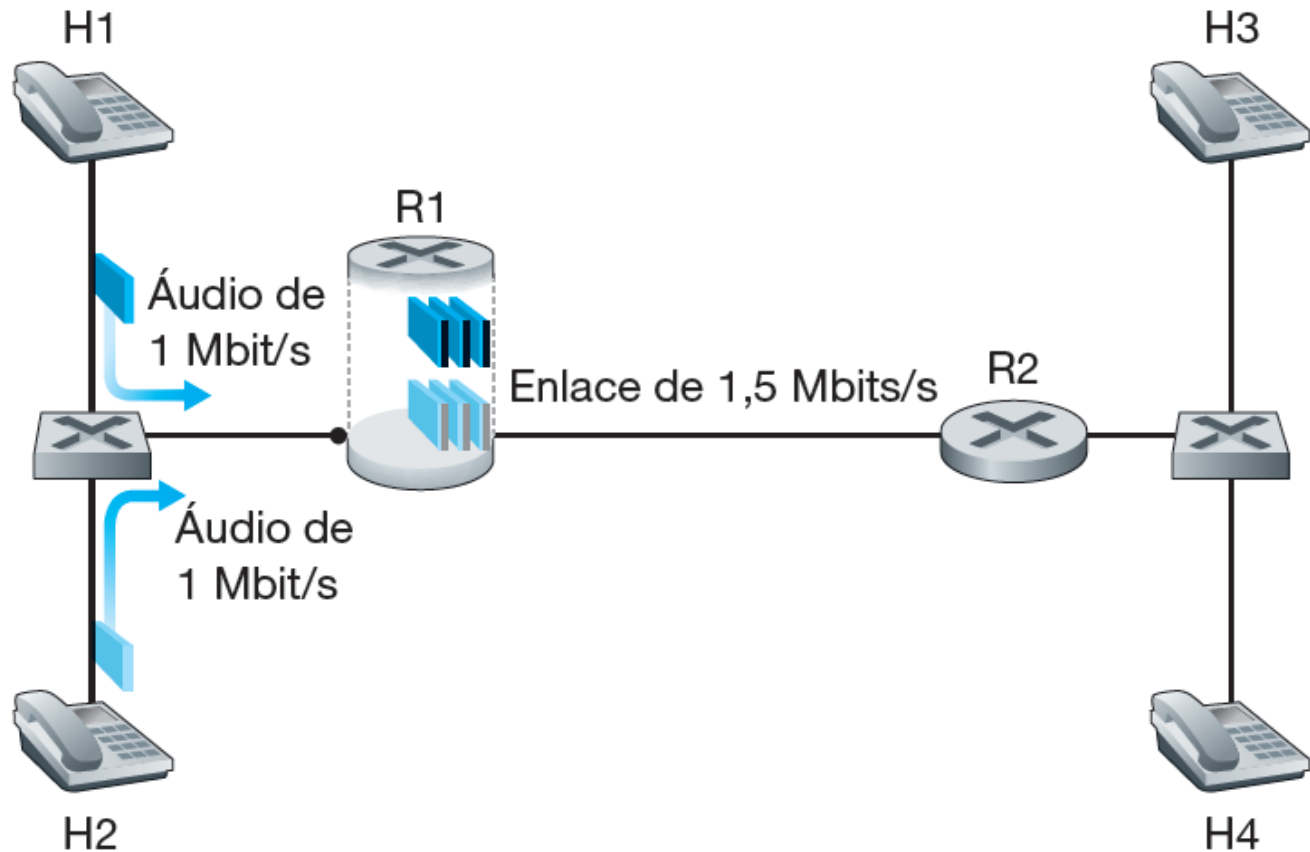
Diffserv

- Exemplo de uma rede Diffserv simples



Garantias de QoS por conexão: reserva de recurso e admissão de chamada

- Duas aplicações de áudio concorrentes sobrecarregando o enlace de R1 a R2



Garantias de QoS por conexão: reserva de recurso e admissão de chamada

- Sabendo que essas duas aplicações não podem ser atendidas simultaneamente, o que a rede deve fazer?
- Um dos fluxos de aplicação deve ser, enquanto o outro deve prosseguir, usando o 1 Mbit/s inteiro necessário pela aplicação.
- O processo de um fluxo declarar seu requisito de QoS e a rede aceitar o fluxo (na QoS solicitada) ou bloqueá-lo é denominado processo de admissão de chamada.

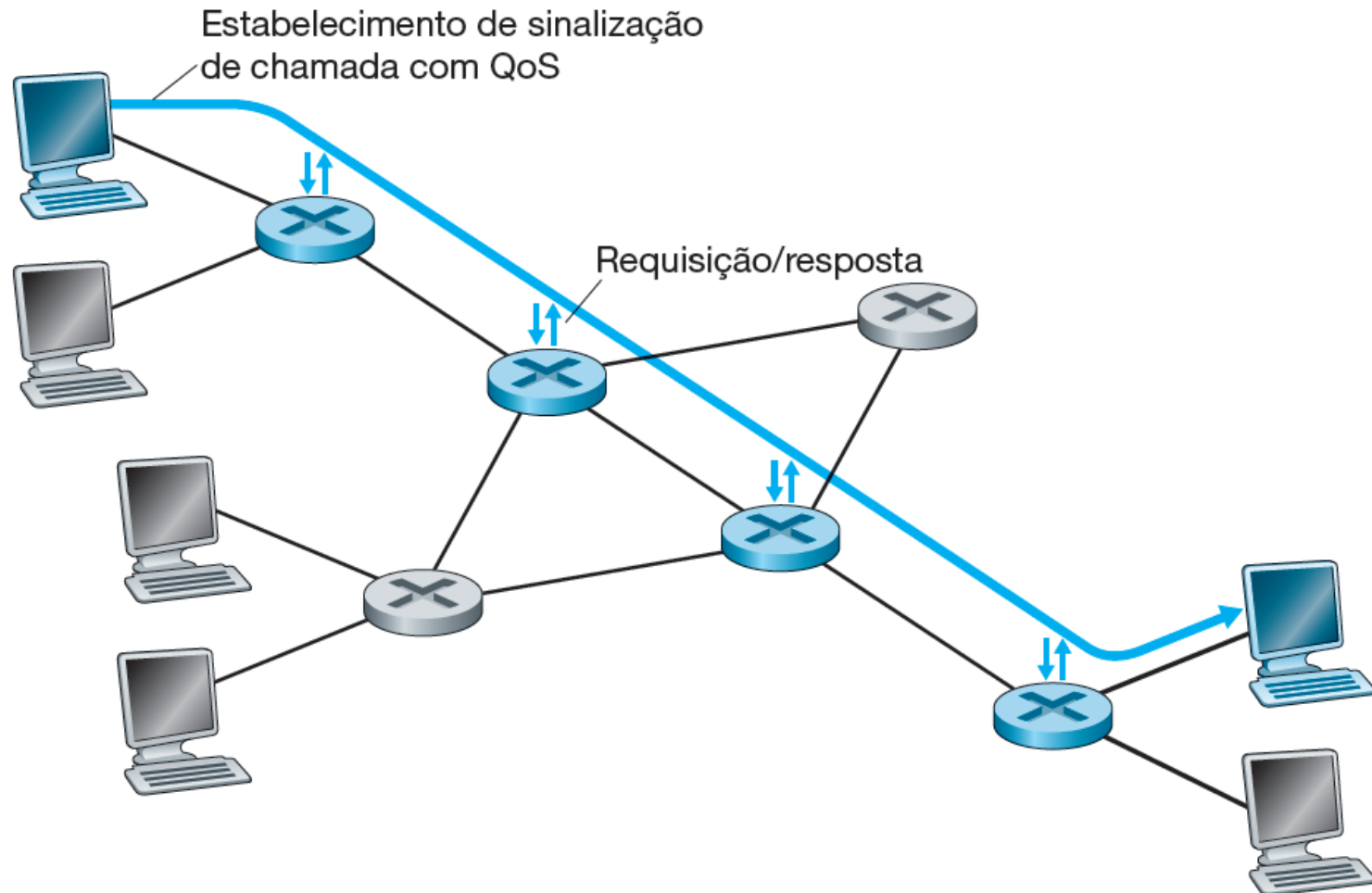
Garantias de QoS por conexão: reserva de recurso e admissão de chamada

- **Princípio 4:** Se recursos suficientes nem sempre estiverem disponíveis e a QoS tiver de ser garantida, é necessário um processo de admissão de chamada no qual os fluxos declaram seus requisitos de QoS e, então, são admitidos à rede ou bloqueados da rede.
- Nosso exemplo motivador na figura anteriormente apresentada enfatiza a necessidade de diversos novos mecanismos e protocolos de rede, se uma chamada tiver de garantir determinada qualidade de serviço uma vez iniciada:

Garantias de QoS por conexão: reserva de recurso e admissão de chamada

- Reserva de recurso.
- Admissão de chamada.
- Sinalização do estabelecimento de chamada.

Garantias de QoS por conexão: reserva de recurso e admissão de chamada



Próximo assunto:

Segurança em Redes de Computadores