

Multimedia Networking

Audio e Video

Em termos de **sensibilidade a atrasos e perdas**, é importante destacar que videos requerem uma entrega mais suave dos dados, ou seja, uma entrega com menos atrasos e perdas. Isso deve-se ao facto dos vídeos serem altamente sensíveis às mudanças da qualidade de transmissão, enquanto que a qualidade de audio é menos afetada.

Existem duas principais formas de redundâncias em vídeo:

- A **redudância temporal** é a repetição de padrões ao longo do tempo. Esta redudância pode ser explorada com técnicas de precisão que utilizam informação anterior para codificar a imagem atual, o que resulta numa redução do tamanho dos dados.

“em vez de enviar a frame completa em i+1, enviar apenas diferenças em relação à frame i”

- A **redudância espacial** é a repetição de padrões na mesma imagem, como a repetição de cores ou padrões de textura. Esta redundância pode ser explorada com técnicas de compactação baseadas em transformadas, que identificam e codificam os componentes de alta frequência de uma imagem.

“em vez de enviar N valores da mesma cor (tudo roxo), enviar apenas dois valores: valor da cor (roxo) e o número de valores repetidos (N)”

O processo de compactação pode ser eficiente em termos de redução do tamanho de dados, mas também pode ter impacto negativo na qualidade do vídeo devido à perda de informação durante o processo de compactação. Além disso, a complexidade do processo de compactação e decodificação também pode afetar o desempenho do sistema, especialmente em dispositivos de baixa potência.

A taxa de bits por amostra resultante de um sinal audio PCM é:

$$\frac{\text{vezes} * \underbrace{\text{niveis} * \log_2(\text{niveis})}_{\text{niveis} * \log_2(\text{niveis})}}{1024}$$

As três categorias principais de aplicações multimédia são:

- **Aplicações de conteúdo passivo**: são aplicações que apenas apresentam conteúdo multimédia, sem interação do utilizador. Exemplos: filme, vídeo, música, apresentação de slides.
- **Aplicações de conteúdo ativos**: são aplicações que permitem ao utilizador interagir com o conteúdo multimédia. Exemplos: jogos, aplicações de edição de vídeos, aplicações de realidade virtual.

- **Aplicações de comunicação multimédia:** são aplicações que permitem aos utilizadores comunicarem uns com os outros através de múltiplos meios, tais como voz, vídeo, mensagens de texto e compartilhamento de conteúdo. Exemplos: conferências, chamadas de vídeo e de voz.

Video Streaming

As duas soluções protocolares para suporte do serviço de video streaming sobre HTTP são:

- O **HLS (*HTTP Live Streaming*)** é um protocolo desenvolvido pela Apple e é amplamente usado para transmissão de vídeo em dispositivos Apple. O HLS divide o fluxo de vídeo em pequenas partes, chamadas de “segmentos”, que são transmitidas sucessivamente ao longo da sessão de streaming.
- O **DASH (*Dynamic Adaptive Streaming over HTTP*)** é uma solução de codificação e transmissão independente da plataforma, que permite que os dados sejam transmitidos de forma adaptativa ao longo da sessão de streaming, ajustando automaticamente a qualidade do vídeo à largura de banda disponível. O DASH é compatível com vários tipos de dispositivos, como smartphones, computadores e smartTVs.

Vantagens e desvantagens de cada solução protocolar:

- O HLS divide o fluxo do vídeo em segmentos de tempo fixo e armazena-os no servidor. O cliente vai descarregando os segmentos conforme forem necessários, permitindo que a qualidade do vídeo se adapte à largura de banda da rede. Além disso, o HLS é amplamente suportado por dispositivos móveis, o que o torna numa escolha popular para aplicações de streaming em dispositivos móveis. No entanto, o HLS tem uma maior latência devido ao tempo de buffer, o que pode afetar negativamente a experiência do utilizador.
- O DASH é baseado na fragmentação do conteúdo em pequenos segmentos, que são descarregados individualmente pelo cliente. Isso permite que o cliente se adapte à largura de banda disponível e ao estado da rede em tempo real, aumentando a qualidade do vídeo e minimizando o buffer. Além disso, o DASH suporta codecs como H.264 e H.265, o que aumenta a eficiência da compressão do vídeo. No entanto, o DASH é mais complexo que o HLS, o que pode levar a problemas de implementação.

O **uso de UDP como transporte para serviços de vídeo streaming** pode resultar em:

- Perdas de pacotes durante a transmissão já que o UDP não oferece garantias de entrega, o que pode causar interrupções na qualidade do vídeo.
- Problemas de sincronização, pois o UDP não garante a ordem dos pacotes durante a transmissão.

- Latência elevada e atrasos na entrega dos pacotes porque o UDP não tem mecanismos para lidar com a congestão na rede.
- Problemas de segurança uma vez que o UDP não oferece criptografia ou outros mecanismos de segurança, tornando-o vulnerável a ataques.
- Problemas de escalabilidade, pois o UDP não possui mecanismos de gestão de congestão, comprometendo o seu desempenho em redes congestionadas ou sob elevada carga.

O atraso inicial expectável (buffering delay) é calculado através da fórmula:

$$\text{Espera (Mbits)} / \text{Taxa (Mbps)}$$

Serviço de voz sobre IP

O serviço de voz sobre IP (VoIP) permite a transmissão de conversas telefónicas através da rede IP. Em comparação com as redes telefónicas tradicionais, o VoIP é mais sensível às variações do desempenho da rede IP subjacente. Algumas dessas variações incluem perdas de pacotes, atrasos e variações na largura de banda disponível.

A qualidade do serviço VoIP depende diretamente da qualidade da rede IP subjacente. Ocorrendo perdas de pacotes, o audio pode ficar truncado ou ser interrompido, o que reduz a clareza da comunicação. Atrasos na entrega dos pacotes, causados por variações na largura de banda disponível, podem criar eco na comunicação.

No entanto, as tecnologias de compressão de audio, controlo da qualidade de serviço e deteção de erros podem ser implementadas para minimizar o impacto destas variações na rede e melhorar a qualidade do serviço. A utilização de redes IP de alta velocidade também podem melhorar a qualidade do serviço VoIP.

Um dos métodos estudados para recuperar perdas de pacotes de voz é o uso de pacotes de redundância. Este método consiste na inclusão de informações adicionais em cada pacote de voz, permitindo ao recetor reconstruir pacotes perdidos. O custo/benefício desse método é geralmente considerado alto, pois requer a inclusão de informações adicionais em cada pacote, o que resulta numa maior largura de banda utilizada.

SIP

O protocolo SIP (Session Initiation Protocol) é um protocolo de sinalização usado na comunicação de voz e vídeo sobre IP. O objetivo principal do SIP é permitir a iniciação, manutenção e terminação de comunicações de múltiplos meios, incluindo chamadas de voz,

vídeo, mensagens instantâneas e presença. As pessoas são identificadas pelo nome ou email, ao invés de pelo número de telefone.

Além disso, o SIP visa fornecer uma solução escalável e flexível para a gestão de comunicações em rede, possibilitando a integração de diferentes serviços e aplicações de comunicação, bem como a sua utilização em diferentes tipos de redes, tanto fixas como móveis.

O SIP também pretende fornecer uma plataforma aberta e interoperável para a comunicação, permitindo aos utilizadores estabelecer comunicações com qualquer outro utilizador de IP independentemente das redes ou provedores de serviço envolvidos.

O SIP é um protocolo independente que pode ser utilizado com outros protocolos IETF, como o RTP, o RTSP e o SDP para construir uma arquitetura de multimédia completa.

Entidades principais do SIP

O protocolo SIP é composto por quatro entidades principais que sustentam a sua operação:

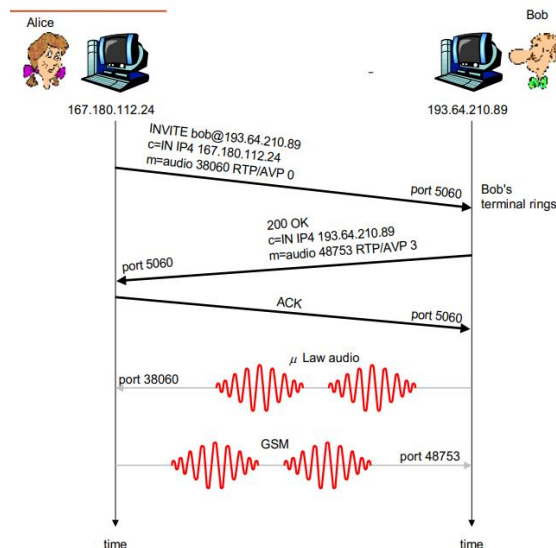
- **Agente de Utilizador** (User Agent): é uma aplicação que permite ao utilizador iniciar, manter e terminar sessões SIP. O agente de utilizador pode ser implementado tanto no lado do cliente como no lado do servidor. O agente de utilizador do lado do cliente envia solicitações SIP ao servidor, enquanto o agente de utilizador do lado do servidor responde a essas solicitações.
 - ✚ **User Agent Client** (UAC): É o componente SIP no lado do cliente que inicia a chamada SIP. É responsável por enviar mensagens SIP ao servidor de chamadas para iniciar, manter ou terminar uma sessão de comunicação.
 - ✚ **User Agent Server** (UAS): É o componente SIP no lado do servidor que é responsável por lidar com as mensagens SIP recebidas do UAC. Ele responde às mensagens do UAC para iniciar, manter ou terminar uma sessão de comunicação.
- **Servidor de Redirecionamento** (*Redirect Server*): é um servidor que, quando recebe uma solicitação SIP, responde indicando ao agente de utilizador o endereço do próximo destino a ser contactado. O servidor de redirecionamento é usado para redirecionar a chamada para outro destino, evitando a necessidade de uma busca adicional pelo destino da chamada.
- **Servidor Proxy** (*Proxy Server*): é um servidor que age como intermediário entre os agentes de utilizador, fornecendo a camada de transporte, roteamento e autenticação de solicitações SIP. O servidor de proxy verifica a autorização do utilizador para iniciar uma sessão e, em seguida, redireciona a solicitação para o destino apropriado.
- **Servidor de Registo** (*Registrar Server*): é um servidor que mantém o registo dos endereços de contato dos agentes de utilizador registados. Quando um agente de utilizador inicia uma sessão, ele envia uma solicitação de registo ao servidor de

registro, que, em seguida, atualiza sua base de dados com o endereço de contato do agente de utilizador.

Estabelecimento de uma sessão entre dois utilizadores SIP localizados...

No mesmo domínio SIP:

1. Alice envia uma mensagem SIP INVITE a Bob, indicando o seu número de porta, endereço IP e codificação preferencial (PCM ulaw).
2. Bob responde com uma mensagem 200 OK, indicando o seu número de porta, endereço IP e codificação preferencial (GSM).
3. As mensagens SIP podem ser enviadas por TCP, UDP ou outro protocolo de transporte; neste exemplo, são enviadas por RTP/UDP.
4. O número de porta SIP padrão é 5060.
5. Negociação de codificação: se Bob não tiver um codificador PCM ulaw, responderá com 606 Not Acceptable Reply, listando os seus codificadores. Alice pode então enviar uma nova mensagem INVITE, anunciando um codificador diferente.
6. Rejeição de chamada: Bob pode rejeitar a chamada com respostas como "ocupado", "partido", "pagamento necessário" ou "proibido".
7. O meio de transmissão pode ser enviado por RTP ou outro protocolo.



Em domínios SIP diferentes:

1. Alice envia uma mensagem SIP INVITE, mas não conhece o endereço IP de Bob.
2. São necessários servidores intermediários SIP para estabelecer a chamada.
3. Alice envia e recebe mensagens SIP usando o número de porta padrão 5060.

4. Alice especifica no cabeçalho Via: que o cliente SIP envia e recebe mensagens SIP por meio de UDP.
5. A mensagem INVITE de Alice é encaminhada através de um servidor proxy SIP que conhece o endereço IP do Bob.
6. O servidor proxy envia a mensagem INVITE para Bob.
7. Bob responde com um 200 OK, indicando que aceita a chamada.
8. Bob também indica o seu endereço IP, número de porta e a codificação preferida (GSM).
9. O servidor proxy encaminha a resposta 200 OK de Bob para Alice.
10. Alice responde com ACK, confirmando a estabilização da chamada.
11. A comunicação de mídia é iniciada através de RTP ou outro protocolo, entre os endereços IP e portas indicados por Alice e Bob na resposta 200 OK.

Comparação entre SIP e H.323

O H.323 é um protocolo de sinalização completo e integrado para conferências multimídia que inclui sinalização, registo, controle de admissão, transporte e codecs. Já o SIP é apenas uma componente que pode ser combinada com outros protocolos e serviços, trabalhando com RTP, mas não o obrigando. O H.323 é originário da ITU (telefonía) enquanto que o SIP vem da IETF e tem um sabor mais semelhante ao da Web, sendo globalmente mais simples e amplamente adotado, incluindo nas tecnologias 3GPP, 3GPP2 e 4G.