

Capítulo 8

Voz sobre IP (VoIP)

Este capítulo apresenta uma das mais importantes aplicações multimídia, a voz sobre IP, ou simplesmente telefonia IP. VoIP é um termo que engloba a transmissão de voz sobre redes IP (na rede corporativa da empresa ou na Internet). Quando o conceito de VoIP surgiu, este representou uma revolução na maneira em que chamadas telefônicas de longa distância poderiam ser conduzidas. Hoje em dia, entretanto, a VoIP engloba muito mais do que ligações mais baratas de longa distância para amigos e familiares. VoIP é um termo geral para tecnologias que utilizam protocolos Internet, permitindo a conexão entre rede de pacotes e redes comutadas para troca de voz, fax e outras formas de informação que tradicionalmente vinham sendo transportada através de conexões dedicadas de circuito comutado, disponibilizados pela *Public Switched Telephone Network* –PSTN. O desafio da VoIP é o de entregar um fluxo de pacotes de voz de forma confiável ao cliente.

O uso da VoIP possibilita o desenvolvimento de novas aplicações impossíveis de serem implementadas na telefonia tradicional. Além disso, ela permite a redução de custos pela convergência de duas redes tradicionalmente separadas, reduzindo custos de instalação e manutenção, e a economia em chamadas de longa distância. Entretanto, os investimentos para a implementação da VoIP são altos e existe o risco da mudança em uma área tão estratégica como a telefonia.

8.1 Serviço Telefônico PSTN

O serviço telefônico fornecido pela telefonia tradicional, a PSTN (*Public Switched Telephone Network*) usa conexões por comutação de circuito, o que significa que quando uma chamada é feita, é estabelecido um circuito dedicado de um telefone até o outro. Décadas de conhecimento e experiência permitiram ao serviço fornecido pela PSTN alcançar a alta qualidade e a disponibilidade que possui hoje. As redes a comutação a circuitos garante taxa de bits, atraso limitado e constante (sem variação). O nível de qualidade esperada da rede PSTN é denominado de “cinco-noves”, o que significa que a rede como um todo deve estar disponível e funcional para 99,999% do tempo [Walker e Hicks, 2002].

O serviço telefônico fornecido pela VoIP ocorre de maneira bem diferente daquele oferecido pelas PSTNs. A VoIP faz a transferência de voz como dados em uma rede IP. Ao invés da comutação a circuito, a conexão telefônica em VoIP é por comutação de pacotes, onde múltiplos dispositivos computacionais compartilham uma rede de dados. O serviço oferecido pela rede geralmente é do tipo melhor esforço, que, ao contrário da comutação a circuito, não fornece garantias de qualidade, gerando um problema na implementação da VoIP.

8.2 Benefícios da VoIP

O fator econômico é o de maior peso quando da definição de novos investimentos. Nenhum investimento é realizado sem que haja garantias de redução dos custos totais ou da maximização de retorno de investimentos, além de trazer ganhos para a empresa. No entanto, um expressivo número de executivos concordam em que a implantação de uma infra-estrutura de telefonia baseada na tecnologia e soluções IP não é mais uma questão de “se” mas de “quando” será implementada.

VoIP oferece muitos benefícios em potencial, incluindo custos reduzidos, novas funcionalidades e redes convergentes.

8.2.1 Redução dos Custos

Quando a primeira tecnologia VoIP surgiu, o principal estímulo foi as “chamadas telefônicas grátis”. Chamadas de longa distância podem ser um dos maiores itens no orçamento de uma organização. Na VoIP, não existem chamadas de longa distância. Se a rede IP já está disponível e se já se paga uma ISP (Provedora de Serviços Internet), então VoIP utiliza uma infra-estrutura que já é paga, assim as chamadas de VoIP poderiam ser consideradas grátis.

Atualmente as corporações necessitam estabelecer conversações de voz para contatos profissionais, entre matriz e filiais, entre fornecedores e clientes, espalhados em grandes áreas geográficas. Como ilustrado na Figura 59, não utilizando a rede PSTN e sim a rede IP, existe uma redução de custos. O tráfego de chamada agora é do PBX para o gateway em vez do PBX para o switch PSTN, assim evitando o custo da PSTN.

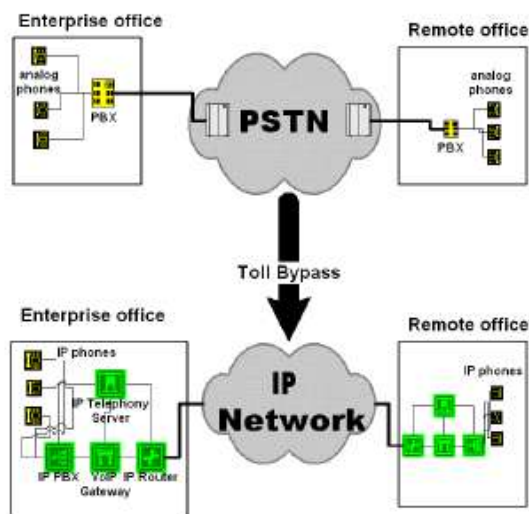


Figura 59. Redução dos custos das chamadas telefônicas

8.2.2 Redes Convergentes

Manter duas infra-estruturas de redes separadas de telefonia e de comunicação de dados não é simples nem barato. VoIP oferece uma infra-estrutura de rede única, construída sobre a rede IP. As vantagens disto são [Walker, 2002]:

- Uma única rede pode reduzir os custos de mantê-la. Em vez de comprar um PBX e uma infra-estrutura de rede para chamadas PSTN, pode-se investir na infra-estrutura de rede IP. Tanto a chamada de voz e tráfego de dados pode tomar vantagem desta melhora.
- VoIP pode prover um custo incremental reduzido da rede. Por exemplo, qual é o custo por usuário atual para serviços de telefonia? Qual é o custo de adicionar um novo usuário? Adicionando um usuário adicional ao sistema PBX tradicional pode requerer um upgrade para um novo PBX com maiores capacidades, assim incrementando o custo por usuário para o sistema. Ao contrário, muitas LANs campus têm uma boa escalabilidade, permitindo a adição de um novo usuário VoIP a custos reduzidos.
- Uma única rede oferece custos de cabeamento reduzidos, especialmente em novas construções. Em vez de implantar cabeamento para dados e voz, pode-se usar um único cabeamento. Esta convergência de voz e dados deve ser feita com cuidado. Você nunca deveria compartilhar um hub por uma chamada de voz IP e outro tráfego; se for solicitada uma consulta a uma base de dados enquanto uma chamada de voz for feita, a qualidade da chamada pode ser reduzida.
- Uma única rede pode facilmente incorporar infra-estruturas sem fio. Cabear uma casa ou escritório para uma rede de dados pode ser cara, assim muitas organizações usam redes sem fio usando a tecnologia 802.11. Estas redes sem fio suportam tanto aplicações IP como VoIP.
- Economia no suporte e gerenciamento. Em um sistema de PBX tradicional é necessário ter uma equipe para gerenciar o sistema de telefonia e outra para gerenciar a rede de dados. Em um sistema VoIP, estes funcionários normalmente são unidos. Com a convergência da infra-estrutura, pode-se reduzir a equipe interna necessária para suporte e gerenciamento. Mas esta economia pode requerer um mais alto custo inicial de treinamento.
- Manutenção e adições de novos usuários. VoIP torna a adição, movimentação ou trocas de usuários e serviços fáceis de serem realizados. VoIP usa protocolos IP tal como DHCP para permitir que telefones IP sejam automaticamente reconfigurados quando movidos de um local para outro.

8.2.3 Novas Funcionalidades

Existem várias novas aplicações e características que o VoIP possibilitam [Walker, 2002]:

- Sistema de mensagem Unificado. Muitos vendedores estão oferecendo a integração de correios de voz, e-mail e fax. A habilidade de se obter mensagens a qualquer hora, em qualquer lugar, e em qualquer modo tornam os sistemas unificados de mensagens interessantes para o aumento da produtividade.
- Roteamento de chamadas avançado. Um empregado que está trabalhando em casa pode ter sua chamada de negócios roteada para o seu telefone residencial. E roteamento de chamada pode incluir integração com o sistema CRM (Gerenciamento de Relacionamento com o Cliente) para verificar as informações do cliente e rotear a chamada para um grupo de suporte técnico apropriado.
- Integração em aplicações de negócio. Por exemplo, uma chamada de voz pode ser feita durante uma conversa via Chat, por exemplo, com o Windows Messenger da Microsoft. A possibilidade de fazer Chat com alguém e então clicando um botão para fazer uma chamada de voz é interessante.
- Mais fácil de adicionar novas características. Novas características podem ser adicionadas à implementação de VoIP muito mais rápido e fácil que em um PBX tradicional.

8.3 Protocolos de sinalização para VoIP

Programas de aplicações implementam os chamados protocolos de aplicação que utilizam protocolos de transporte (TCP ou UDP). Implementando uma chamada de voz em uma rede digital envolve duas fases (Figura 60): a configuração de chamada (*call setup*) – que é o equivalente VoIP da obtenção do tom de chamada da telefonia, digitação do número do telefone, soar o telefone ou um sinal de ocupado, e tirar o fone do gancho no outro lado; e então a conversação telefônica em si. Protocolos VoIP são necessários durante estas duas fases:

- Vários protocolos podem executar as configurações de chamada e encerramentos de chamada, incluindo H.323, SIP, MGCP, e Melco. Os programas que implementam o protocolo de configuração de chamada usam TCP e UDP para encapsular o dado trocado durante a configuração da chamada e seu encerramento.
- A troca de dados de voz codificados ocorre após o estabelecimento da chamada, usando dois fluxos de dados – um em cada direção – para deixar os dois participantes falarem ao mesmo tempo. Cada um destes fluxos de dados usa um protocolo de alto nível chamado RTP (*Real-time Transport Protocol*), que é encapsulado em datagramas UDP e IP para transferência.

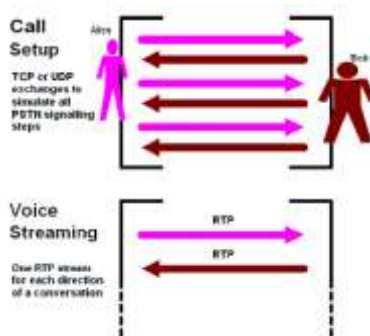


Figura 60. Fases envolvidas na chamada VoIP

Existem duas famílias de protocolos de configuração de chamadas:

- Os protocolos H.323 e MGCP (Media Gateway Control Protocol) oriundos da comunidade da telefonia. O H.323 é o protocolo mais usual e é na realidade uma família de padrões baseados em telefonia para multimídia, incluindo voz e videoconferência. MGCP é uma versão menos flexível, para uso em dispositivos mais baratos como telefones domésticos.
- SIP (Session Initiation Protocol) é um protocolo leve desenvolvido pela IETF (comunidade de redes). SIP é um protocolo bem mais leve (que o H.323) e praticamente faz a mesma coisa. Ele é

suportado pela Cisco e pela Nortel, e a Microsoft tem recentemente iniciado o desenvolvimento de interfaces clientes SIP com o sistema operacional Windows XP.

Embora a família H.323 de protocolos de configuração de chamada é predominante hoje, o é esperado que os quatro protocolos discutidos acima (H.323, MGCP, Megaco e SIP) deverão usados em iguais proporções nos próximos anos.

8.4 Padrão de Videoconferência H.323

O H.323 é uma recomendação da União Internacional de Telecomunicações (ITU) que fornece um framework (componentes, protocolos, e procedimentos) de comunicação para áudio, vídeo e dados através de redes a comutação de pacotes. Redes baseadas em pacotes incluem redes locais, corporativas, metropolitanas e mundiais baseadas em IP (incluindo a Internet) ou IPX (*Internet Packet Exchange*). Ele é considerado um protocolo de sinalização importante para VoIP, mas ele também abrange a videoconferência sobre redes a pacotes.

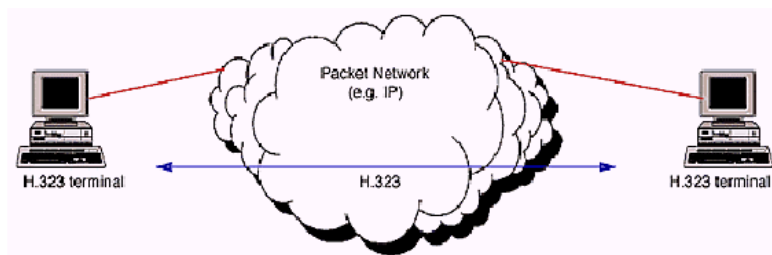


Figura 61. Terminais H.323 em redes a pacotes

O H.323 possibilita uma variada gama de aplicações interativas multimídia tais como:

- Internet fone;
- Videoconferência em desktop;
- Computação colaborativa;
- Conferência de negócios;
- Ensino a distância;
- Voz sobre IP (VoIP);
- Aplicações de suporte e Help Desk (localização em uma organização onde chamadas podem ser recebidas para resolução de problemas e/ou um conhecimento básico das informações acerca dos produtos/serviços de uma companhia).
- Compra/venda interativas;
- Outras.

O padrão H.323 é especificado pelo Grupo de Estudos 16 do ITU. A versão 1 da recomendação H.323 – sistemas de telefonia visual e equipamentos para LANs que não fornecem garantias de QoS – foi aprovada em 1996. Ela foi, como o nome sugere, duramente voltada para a comunicação multimídia em um ambiente de rede local e não fornece garantias de QoS.

A emergência das aplicações de voz-sobre-IP (VoIP) levou a uma revisão da especificação H.323. A ausência de um padrão para VoIP resultou em produtos que não eram compatíveis. Com o desenvolvimento do VoIP, novos requisitos surgiram, tal como fornecer comunicação entre telefones baseados em PC e telefones tradicionais em redes a comutação de circuito. Tais requisitos forçaram a necessidade de um padrão para a telefonia IP. Uma nova versão do H.323, chamada de versão 2 – sistemas de comunicação multimídia baseadas em pacotes – foi definida para acomodar estes requisitos adicionais e foi aprovada em janeiro de 1998. Novas características estão sendo adicionadas ao padrão H.323, que deve resultar na sua versão 3. Entre estas novas características estão incluídas fax sobre redes a pacote, comunicação gatekeeper-gatekeeper (visto mais adiante) e mecanismos de conexão rápida.

O H.323 é parte de uma série maior de padrões de comunicação que habilitam videoconferência através de uma gama de redes. Conhecido como H.32X, esta série inclui o H.320 e o H.324, que endereçam

comunicações ISDN e PSTN (*Public Switched Telephone Network* - Rede Comutada de Telefonia Pública), respectivamente. Os padrões associados são os seguintes:

- H.320 – o padrão de videoconferência ISDN original.
- H.322 – serviço de comunicação multimídia sobre LAN com suporte de QoS.
- H.323 – uma extensão do H.320 para videoconferência em redes locais
- H.324 – uma extensão do H.320 para videoconferência sobre linhas PSTN.
- Q.931 –serviço de sinalização de chamadas para estabelecimento e terminação de chamadas.
- T.120 – protocolo de conferência de dados tempo-real.

Por sua vez, o H.323 também é uma família de padrões:

- H.225 – protocolo de controle de chamada: RAS (Registration, Admission and Status). Especifica mensagens para controle de chamada incluindo sinalização, registro e admissão, empacotamento e sincronização de fluxos de medias.
- H.245 – protocolo de controle de mídia. Especifica mensagem para abertura e fechamento de canais para fluxos de mídias e outros comandos.
- H.261 – codificação de vídeo a taxa ≥ 64 kbps.
- H.263 - codificação de vídeo a taxa < 64 kbps.
- G.711 – codificação de áudio PCM 56/64 kbps
- G.722 - codificação de áudio para 7 Khz a 48/56/64 kbps
- G.723 - codificação de voz para 5.3 e 6.4 kbps
- G.728 - codificação de voz para 16 kbps
- G.729 - codificação de voz para 8/13 kbps

A adoção do padrão H.323 trás uma série de benefícios. Os principais são:

- Codificadores/Decodificadores padrões: ele estabelece padrões para compressão de descompressão de fluxos de áudio e vídeo, assegurando que equipamentos de diferentes fabricantes tenham alguma área de suporte comum.
- Interoperabilidade: ele permite aos usuários fazerem uma conferência sem se preocupar com a compatibilidade dos vários pontos finais. Além de assegurar que o receptor possa decodificar a informação, o H.323 estabelece métodos para clientes receptores comunicarem suas capacidades ao remetente.
- Independência de rede: ele foi projetado para executar sobre arquiteturas de rede comuns. Na medida em que evolui a tecnologia de rede e melhoram as técnicas de administração da largura de banda, soluções baseadas em H.323 poderão tirar proveito dessas capacidades aumentadas.
- Independência de Plataforma e Aplicação: ele não é dependente de hardware ou sistema operacional. Plataformas aderentes ao H.323 estão disponíveis em muitos tamanhos e formas, elas incluem computadores pessoais com capacidades de áudio e vídeo, plataformas dedicadas, telefones com IP habilitados e adaptadores para TV a cabo.
- Suporte multiponto: permite a realização de conferências multiponto.
- Administração de Largura de Banda: os tráfegos de áudio e vídeo geralmente geram uma grande taxas de bits e podem congestionar a rede corporativa. O H.323 trata este assunto provendo administração de largura de banda: os gerentes de rede podem limitar o número de conexões H.323 simultâneas dentro da rede ou a largura de banda disponível para aplicações H.323. Estes limites asseguram que o tráfego crítico não será prejudicado.
- Suporte Multicast: ele suporta transporte multicast em conferências multiponto, permitindo o uso mais eficiente da largura de banda uma vez que todas as estações no grupo multicast lêem um único fluxo de dados.

8.4.1 Arquitetura H.323

A recomendação H.323 cobre os requisitos técnicos para serviços de comunicação de áudio e vídeo em LANs que não fornece garantias de Qualidade de Serviço (QoS). H.323 referencia a especificação T.120 para conferência de dados. O escopo do H.323 não inclui a LAN em si ou a camada de transporte que

pode ser usada para conectar as várias LANs. Apenas os elementos necessários para a interconexão com a rede a comutação de circuitos (*Switched Circuit Network - SCN*) estão dentro do escopo do H.323.

O H.323 especifica quatro componentes principais: Terminais, Gateways, Gatekeepers e MCUs (*Multipoint Control Units*) como mostra a Figura 62. Terminais, Gateways e MCUs são classificados como dispositivos terminais (*endpoints*) – são dispositivos que podem iniciar e receber chamadas. Outras tecnologias importantes associadas ao H.323 são os CODECs usados para codificar e decodificar transmissões de vídeo e áudio.

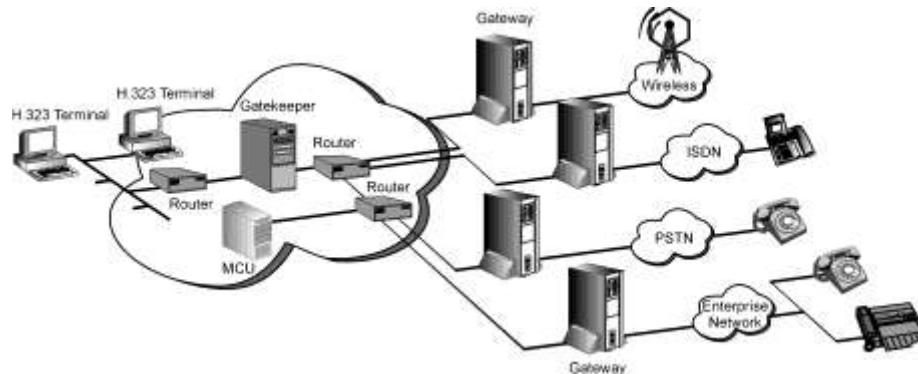


Figura 62. Componentes principais da arquitetura H.323

Os Gateway, Gatekeeper e MCU são componentes lógicos da recomendação H.323 e pode ser implementados conjuntamente em um mesmo hardware, podendo existir em um mesmo hardware as funções Gateway, Gatekeeper e MCU.

8.4.2 Terminais

Terminais H.323 são clientes (*endpoints*) que provem comunicação multimídia bidirecional em tempo real, executando a pilha H.323 e as aplicações multimídia. Ele suporta a comunicação de áudio e pode opcionalmente suportar a comunicação de vídeo e dados (prove funcionalidades tais como “chat” em texto, Quadro Branco - “white boarding - WB” compartilhado). Como o serviço básico fornecido por um terminal H.323 é a comunicação de áudio, um terminal H.323 é um elemento importante nos serviços de telefonia IP.

A principal meta do H.323 é a interoperabilidade com outros terminais multimídia. Terminais H.323 são compatíveis com: terminais H.324 das redes PSTN e redes sem fio; terminais H.310 na B-ISDN ou ISDN; terminais H.321 em B-ISDN; e terminais H.322 em redes locais com garantias de QoS.

Terminais podem ser implementados como softwares baseados em PCs ou em dispositivos stand-alone tais como vídeo fones, ou Internet fones. Atualmente, a grande maioria dos terminais é implementada como software de PCs. Como estes terminais não são rigidamente especificados pela recomendação H.323, terminais baseados em PCs normalmente usam uma placa de som, microfone e alto-falantes ou fones de ouvido. Existem muitos fabricantes que fornecem placas de PC onde pode ser conectado um telefone padrão.

Terminais H.323 devem suportar os seguintes componentes:

- H.245/H.248 para troca de capacidades do terminal e criação de canais de mídia.
- H.225 para sinalização de chamada e configuração de chamada.
- RAS para registro e outros controles de admissão com o gatekeeper.
- RTP/RTCP para sequenciamento de pacotes de áudio e vídeo.

Terminais H.323 devem também suportar o CODEC de áudio G.711. Componentes opcionais em um terminal H.323 são CODECs de vídeo, protocolos de conferência de dados T.120 e MCU.

Um exemplo de um terminal é mostrado na Figura 63. O diagrama mostra a interface do equipamento do usuário, codec de vídeo e de áudio, equipamento de telemática, H.225.0 layer, funções de controle do sistema e a interface com a rede baseada em pacotes.

Um terminal H.323 deve implementar diversos dos protocolos especificados pelo H.323 (Figura 63):

- CODECs de áudio;
- CODECs de vídeo;
- H.225 registro, admissão e status (RAS);
- H.225 sinalização de chamada;
- H.245 controle de sinalização;
- RTP (Real-Time Transfer Protocol);
- RTCP (Real-Time Control Protocol);

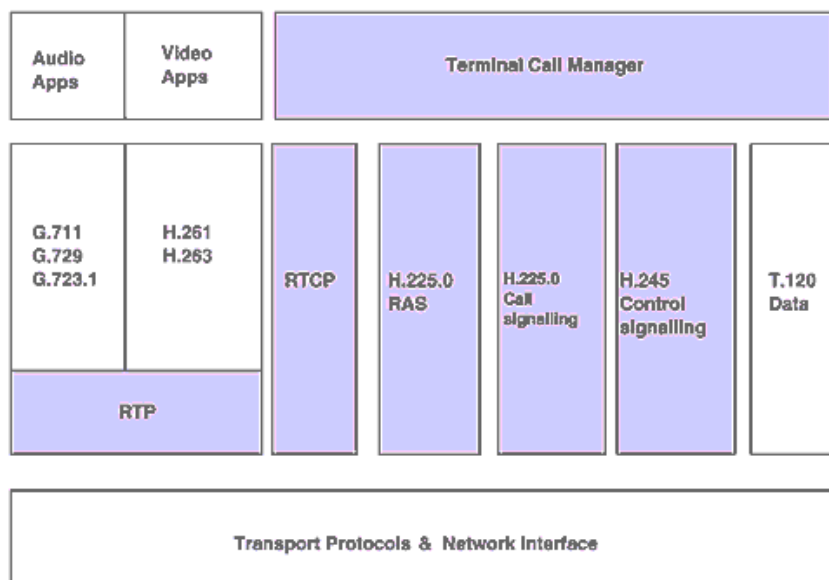


Figura 63. Protocolos H.323

A seguir são apresentados alguns dos elementos que compõem o terminal H.323. Os demais protocolos serão apresentados posteriormente.

CODEC de Áudio

Um CODEC de áudio codifica o sinal oriundo de um microfone para transmissão por parte de um terminal H.323 e decodifica o código de áudio recebido e o envia ao alto falante no terminal H.323 receptor.

Os algoritmos de compressão suportados pelo H.323 são todos padrões aprovados pelo ITU. Terminais H.323 devem suportar o padrão G.711 para compressão de voz. Suporte para os outros padrões de voz ITU é opcional.

As diferentes recomendações ITU para digitalização e compressão de fala refletem uma relação de compromisso entre qualidade da fala, taxa de bits, capacidade de computação e retardo do sinal. O G.711 geralmente transmite voz a 56 ou 64kbps, bem dentro da largura de banda disponível na LAN, porém ele foi projetado originalmente para rede de taxa contínua de bits. Pelo fato de o padrão G.723 operar em taxa de bits muito baixa (5.3 e 6.3 kbps), ele está sendo altamente considerado como um CODEC obrigatório e tende a ser o CODEC predominante em aplicações H.323.

CODEC de Vídeo

Um CODEC de vídeo codifica o sinal de vídeo oriundo de uma câmera para transmissão no lado do terminal H.323 emissor e decodifica o código de vídeo que é enviado para o display do terminal H.323 receptor.

Embora a capacidade de vídeo seja opcional, qualquer terminal H.323 que possibilite vídeo deve suportar o CODEC H.261; o suporte para H.263 é opcional. A informação de vídeo é transmitida a uma taxa não superior aquela selecionada durante a negociação de capacidades.

H.225 Layer

O H.225 layer descreve a camada mais baixa que encapsula e desencapsula os vários fluxos de dados para dentro de pacotes para transmissão sobre uma LAN. O H.225 também manipula o controle de chamadas para iniciar e finalizar chamadas entre terminais, gateways, MCUs e gatekeepers. A camada H.225 inclui o Q.931, RAS e o RTP/RTCP.

H.225 Sinalização de Chamada

O H.225 sinalização de chamada é usado para estabelecer um canal entre dois endpoints H.323, na qual dados em tempo-real podem ser transportados. Isto é obtido pela troca de mensagens H.225 no canal de sinalização de chamada confiável (Por exemplo, TCP em uma rede baseada em IP). O canal de sinalização de chamada é aberto entre dois pontos finais H.323 ou entre um ponto final e o gatekeeper. O Canal de Sinalização de Estabelecimento de Chamada usa o Q.931 para estabelecer uma conexão entre dois terminais. As mensagens H.225 são trocadas entre endpoints se não existam gatekeepers na rede H.323.

H.245 Sinalização de Controle

H.245 sinalização de controle é usado para trocar mensagens de controle fim-a-fim para governar a operação do ponto final H.323. Elas são transportadas sobre os canais de controle H.245. Este canal de controle é o canal lógico 0 e é permanentemente aberto, diferente dos canais de mídia. Estes mensagens transportam informações relacionadas a:

- Troca de potencialidades: para que os terminais possam enviar aos seus pares a capacidade de receber e processar fluxos de mídia.
- Abertura e fechamento dos canais lógicos usados para transportar fluxos de mídia;
- Mensagens de controle de fluxo;
- Comandos e indicações gerais.

RTP (Real-Time Transport Protocol)

Como visto anteriormente nesta apostila, o RTP fornece serviços de liberação de áudio e vídeo tempo-real fim a fim. Como já apresentado, o RTP fornece a identificação do tipo de dado transportado, numeração da sequência, marcas temporais, e monitoração.

RTCP (Real-Time Transport Control Protocol)

Como visto anteriormente nesta apostila, o RTCP é o protocolo parceiro do RTP e fornece serviços de controle. A principal função do RTCP é fornecer uma realimentação sobre a qualidade da distribuição dos dados. Outras funções RTCP incluem um identificador da fonte RTP.

Dados

Data-conferência é uma capacidade opcional. Quando suportada, possibilita a colaboração através de aplicações tais como: Quadro branco, compartilhamento de aplicações e transferência de arquivos.

O H.323 suporta data-conferência através da especificação T.120. O padrão ITU, T.120 endereça conferências ponto-a-ponto e multiponto. Ele provê interoperabilidade no nível de aplicação, rede e transporte. Um sistema H.323 pode suportar dados incorporando capacidades T.120 nos clientes e unidades de controle multiponto. A MCU pode controlar e multiplexar as informações de conferência de dados.

A recomendação para suporte multicast no T.120, conhecida como T.125 Anexo A ou Protocolo de Adaptação Multicast, foi aprovado pelo ITU em janeiro de 1998.

Terminais em Redes IP

O H.323 usa comunicações seguras/confiáveis e incertas/não confiáveis. Sinais de controle e dados requerem transporte seguro porque os sinais devem ser recebidos na ordem na qual foram enviados e não podem ser perdidos. Fluxos de áudio e vídeo perdem o valor com o tempo. Se um pacote de áudio chegar atrasado, pode não ter relevância ao usuário final. Sinais de áudio e vídeo usam transporte incerto/não confiável de modo mais eficiente.

A transmissão segura de mensagens usa o modo orientado a conexão para transmissão de dados. Na pilha IP, este tipo de transmissão é realizado com o TCP. A transmissão segura garante a seqüência, correção de erros, controle de fluxo na transmissão de pacotes, mas pode atrasar a transmissão e reduzir a vazão.

O H.323, cuja pilha de protocolos na rede IP é mostrada na Figura 64, usa os serviços seguros de conexão de fim-a-fim do (TCP) para o Canal de Controle H.245, para o Canais de Dados T.120, e para o Canal de Sinalização da Chamada.

Dentro da pilha de IP, serviços incertos/não confiáveis são providos pelo protocolo UDP. O H.323 usa UDP para áudio, vídeo, e para o Canal RAS (este último devido a necessidade de multicast).

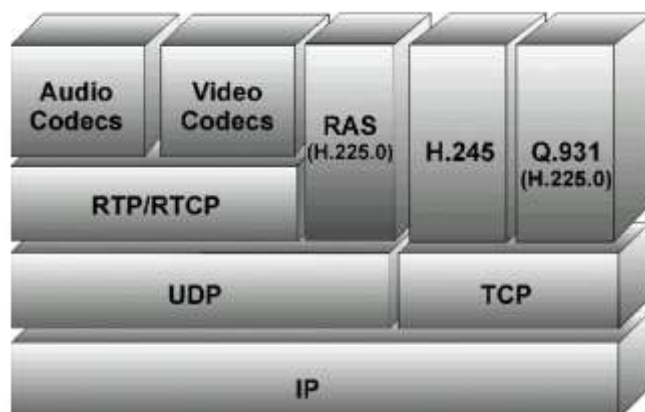


Figura 64. Pilha de Protocolos H.323

8.4.3 Gateways

Um gateway H.323 é um *endpoint* que provê comunicação bidirecional em tempo-real entre terminais H.323 e outros terminais padrão ITU (isto é telefones), ou outro Gateway H.323. Eles executam a tradução de controle de chamada e de conteúdo necessária nas chamadas para converter uma chamada da rede de comutação de pacotes formato H.323 para o formato das redes de comutação a circuito PSTN e vice-versa.

Gateways são componentes opcionais na rede. Eles são necessários quando se está conectando outros tipos de terminais tais como telefones comuns (redes PSTN e PBXs) ou terminais H.320 (Vídeo conferência ISDN).

Um uso comum para gateways é para transporte de tráfego de telefone de longa distância sobre uma rede IP. As empresas podem reduzir seus custos de chamadas de longa distância usando gateways para transporte de voz sobre as redes de dados existentes. Neste tipo de aplicação, o usuário discar um número de acesso local para conectar-se ao gateway, e discar o número de destino. O Gateway local faz a conexão IP para o outro gateway localizado na área de destino chamada. O Gateway remoto completa a chamada discando o número local. A Figura 65 ilustra uma chamada de gateway-para-gateway.



Figura 65. Chamada de Gateway-para-Gateway

Gateways são usados também para fornecer uma interface entre clientes H.323, tal como Microsoft NetMeeting e a rede PSTN, como visto na Figura 66. Esta configuração pode ser usada por um "call center" para permitir que usuários on-line possam contatar um atendente a partir do web site.

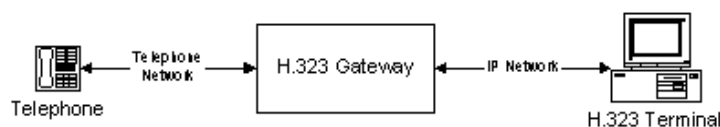


Figura 66. Chamada de Terminal-para-Gateway

Um gateway H.323 é formado por dois componentes: MGC (Media Gateway Control) e o MG (Media Gateway). O elemento MG é responsável em prover a conversão de mídia de um tipo de rede para um outro formato exigido pelo outro lado da rede. Por exemplo, o MG pode fazer a conversão de circuitos comutados para pacotes comutados em uma rede IP. Já o MGC é o elemento que gerencia todo o processo de criação, modificação, monitoração e exclusão de chamadas através do MG. A implementação do MG e do MGC pode ser realizada de duas formas [Soares, 2002]:

- Composite Gateway: onde o MG e MGC coexistam em um mesmo dispositivo físico;
- Decomposite Gateway: onde os elementos MG e MGC estão separados fisicamente, e um elemento MGC pode gerenciar múltiplos elementos MG's. Toda a comunicação entre o MGC e os MGs é realizada através do protocolo de sinalização e controle H.248 (chamada também de megaco). O controle dos vários MG por um único equipamento MGC torna o sistema mais escalável.

Como o gateway opera com os dois lados da rede, H.323 e PSTN, ele interage com os protocolos respectivos a cada tipo de rede. Do lado da LAN, o gateway implementa o H.245 para controle de sinalização quando da utilização de Composite Gateway, H.248 para controle de sinalização quando da utilização de Decomposed Gateway, H.225 para sinalização de chamada para estabelecimento e encerramento de chamadas e outros protocolos específicos da recomendação H.323. Já do lado da PSTN, o gateway implementa protocolos específicos para a telefonia convencional, como por exemplo ISDN e SS7.

8.4.4 Gatekeepers

Um gatekeeper pode ser considerado o cérebro da rede H.323. Ele é ponto de foco para todas as chamadas dentro da rede H.323. Embora eles não são necessários, os gatekeepers fornecem serviços importantes tal como endereçamento, autorização e autenticação de terminais e gateways; gerenciamento de largura de banda; contabilidade; faturamento; e debitação.

O gatekeeper (Figura 67) é um componente opcional do padrão H.323. Ele provê diversos serviços importantes e deverá fazer parte da maioria das redes H.323 no futuro.

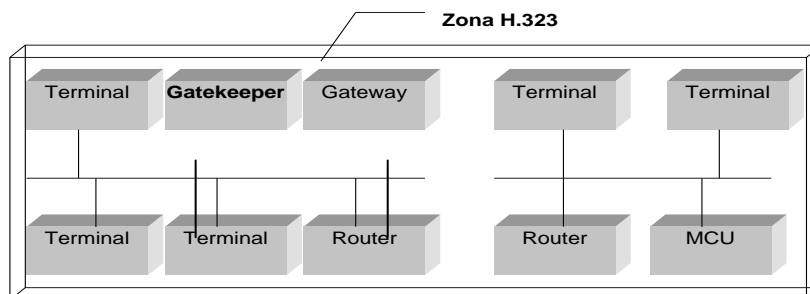


Figura 67. Uma rede H.323 com Gatekeeper

A seguir serão apresentados os serviços fornecidos por um gatekeeper.

- Gerenciamento de Zona: Gatekeepers introduzem o conceito de **zonas**. Uma zona é uma coleção de todos os terminais, gateways e MCUs gerenciadas por um gatekeeper. Uma zona deve incluir pelo menos um terminal, e pode ou não incluir gateways ou MCUs. Uma zona tem um e somente um gatekeeper. Ele é independente da topologia da rede e pode incluir múltiplos seguimentos de LAN.
- Controle de admissão: O gatekeeper concede ou nega admissão à zona para endpoints H.323. Isto pode estar baseado na disponibilidade de largura de banda ou algum outro critério.
- Controle de Largura de Banda: O gatekeeper também controla todos os pedidos para mudanças de largura de banda. Isto pode estar baseado na contenção de largura de banda na zona, em uma limitação de largura de banda específica estabelecida pelo administrador da zona, ou algum outro critério.
- Tradução de endereços: O gatekeeper mantém um banco de dados dos endpoints associados a sua zona. Em casos de pedido de endpoints, ele traduz pseudônimos de endpoint (por exemplo, hostname, endereço de e-mail) para endereços de rede (por exemplo, endereços IP) e traduz

endereços externos (número de telefone formato E.164) para endereços de rede. O endpoint usa então o endereço de rede para iniciar uma chamada a outro endpoint.

- **Autorização de Chamada:** O gatekeeper pode escolher rejeitar chamadas de um endpoint devido à falha de autorização (por exemplo, você escolhe não aceitar chamadas de H.323 de um concorrente).
- **Administração de Largura de Banda:** O gatekeeper pode limitar o número de endpoints H.323 que podem acessar a rede simultaneamente, rejeitando novas chamadas devido a limitações de largura de banda. Administração de largura de banda é permite controlar a banda de rede limitando o número de novas chamadas. Isto difere de controle de largura de banda, definido acima, o qual refere-se a pedidos de mudanças de largura de banda em chamadas existentes.
- **Administração de Chamadas:** O gatekeeper mantém um banco de dados de chamadas de H.323 em andamento. Estes dados são usados para localizar o estado de endpoints para uma variedade de propósitos, inclusive localização de chamada e administração de largura de banda.

H.225 Registro, Admissão e Status

Registro, admissão e status (RAS) é o protocolo entre os endpoints (terminais e gateways) e gatekeepers. O RAS é usado, entre outros, para:

- **Descoberta do gatekeeper (GRQ).** O processo de descoberta do gatekeeper é usado pelos endpoint H.323 para determinar o gatekeeper na qual o endpoint deve registrar. Esta descoberta pode ser feita estaticamente ou dinamicamente. Na descoberta estática, o endpoint conhece o endereço de transporte do seu gatekeeper a priori. No método dinâmico, o endpoint realiza um multicast da mensagem GRQ ("Who is my gatekeeper?") no endereço multicast de descoberta de gatekeepers. Um ou mais gatekeepers podem responder a com uma mensagem GCF ("I can be your gatekeeper").
- **Registro do endpoint.** Registro é o processo usado pelos endpoints para se juntar a uma zona e informar o gatekeeper dos endereços de alias e de transporte.
- **Localidade do endpoint.** É o processo pelo qual o endereço de transporte de um endpoint é determinado e dado seu nome alias ou endereço E.164.
- **Controle de admissão,** para restringir a entrada de um endpoint na zona.
- **Controle da largura de banda.**
- **Desengajamento,** onde o endpoint é desassociado de um gatekeeper e de sua zona.

As mensagens RAS são transportadas por um canal RAS que é não confiável (Por exemplo, UDP). A troca de mensagens RAS deve ser associada com timeouts e numeradas.

H.225 Sinalização de Chamada

Como visto anteriormente, o H.225 sinalização de chamada é usado para estabelecer um canal entre dois endpoints H.323, na qual dados em tempo-real podem ser transportados. As mensagens H.225 são trocadas entre endpoints se não existam gatekeepers na rede H.323. Quando um gatekeeper existir na rede, as mensagens H.225 são trocadas diretamente entre endpoints ou entre os endpoints após serem roteadas pelo gatekeeper:

- No modelo de chamada roteado, o endpoint originador da chamada (A) (Figura 68) envia um pedido admissão de zona ao gatekeeper via o canal RAS. O gatekeeper recebe as mensagens de sinalização de chamada sobre os canais de sinalização de chamada vindo de um endpoint e roteia elas para outro endpoint no canal de sinalização de chamada do outro endpoint. Isto permite ao gatekeeper monitorar o estado e controlar todas as chamadas dentro de sua zona.

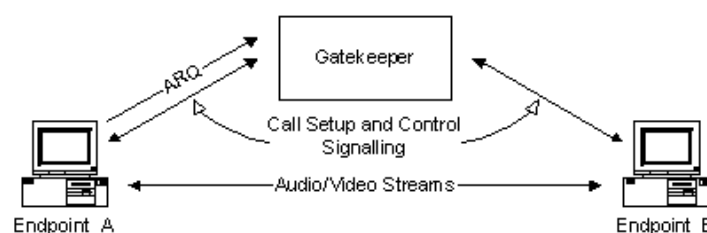


Figura 68. Modelo de Chamada Roteada

- No modelo de chamada direta, o endpoint originador da chamada (A) (Figura 69) envia um pedido de admissão de zona (ARQ) para o gatekeeper. Durante a confirmação da admissão, o gatekeeper indica que os endpoints podem trocar mensagens de sinalização de chamada diretamente. O estabelecimento da chamada e a sinalização de controle são feitos diretamente entre os endpoints.

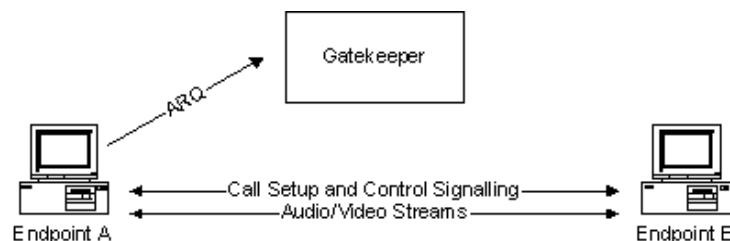


Figura 69. Modelo de Chamada Direta

8.4.5 Multipoint Control Units

MCUs fornecem suporte para conferências de 3 ou mais terminais H.323. Todos os terminais participantes na conferência estabelecem uma conexão com o MCU. O MCU gerencia recursos de conferência, negocia entre terminais para solicitar o CODEC (codificador/decodificador) de áudio e vídeo a ser usado, e pode também manipular o fluxo multimídia.

Os gatekeepers, gateways e MCUs são componentes logicamente separados do padrão H.323, mas podem ser implementados como um dispositivo físico único.

Uma MCU consiste de pelo menos um Controlador Multiponto (MC) e de Processador(es) Multiponto (MP) opcionais:

- O MC provê as funcionalidades de controle de chamada necessárias para conferência multiponto, incluindo negociação de facilidade comuns entre os endpoints.
- O MP, se existir, provê processamento (multiplexação ou chaveamento) de fluxos de diferentes mídias (áudio, vídeo, e/ou dados).

MCs e MPs funcionalmente podem ser incorporados em outras entidades do H.323 – terminais, gateways ou gatekeepers.

No H.323, conferências multiponto são controladas de diversos modos e configurações. A recomendação utiliza os conceitos de conferência centralizada e descentralizada, como ilustrado na Figura 70.

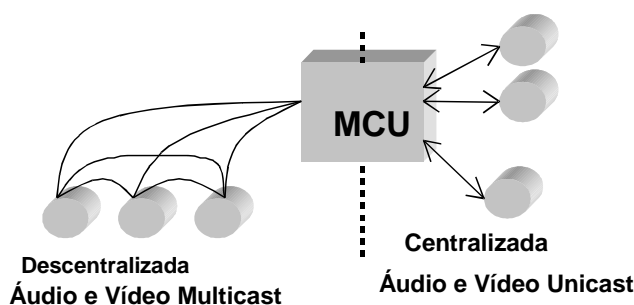


Figura 70. MCU – Conferência Centralizada, Conferência Descentralizada.

A conferência centralizada multiponto requer a existência de uma MCU. Todos os terminais enviam áudio, vídeo, dados e fluxos de controle para a MCU no modo ponto-a-ponto. O MC de forma centralizada gerencia a conferência utilizando funções de controle do H.245 que também define as capacidades de cada terminal. O MP faz a multiplexação de áudio, distribuição de dados, e funções de seleção de vídeo tipicamente executados em conferências multiponto e envia o fluxo resultante para os terminais participantes. O MP pode também fornecer conversão entre diferentes CODECs e taxas de bits e pode também usar multicast para distribuir o vídeo processado.

Conferências multiponto descentralizadas podem fazer uso da tecnologia multicast. Os terminais H.323 enviam áudio e vídeo multicast para outros participantes sem enviar dados para uma MCU. Note que o

controle de dados multiponto continua sendo processado centralizadamente pela MCU, e as informações de controle H.245 são transmitidas no modo ponto-a-ponto para um MC. Os terminais receptores são responsáveis pelo processamento de múltiplos fluxos de áudio e vídeo. Os terminais usam o canal de controle H.245 para indicar ao MC quantos fluxos simultâneos de áudio e vídeo eles podem decodificar. O MP pode ainda prover seleção de vídeo e multiplexação de áudio em conferência multiponto descentralizada.

Conferências multiponto híbridas usam uma combinação de características de conferências centralizadas e descentralizadas. Sinais H.245 e fluxos de áudio e vídeo são processados através de mensagens ponto-a-ponto para a MCU. E a partir desta, os sinais de áudio e vídeo são transmitidos para os terminais H.323 participantes através de multicast. Uma vantagem da conferência centralizada é que todos os terminais H.323 suportam comunicação ponto-a-ponto. A MCU pode fazer múltiplas transmissões unicast para os participantes e não são requeridas capacidades especiais na rede. Por outro lado, a MCU pode receber múltiplos fluxos, multiplexar áudio e vídeo e transmitir fluxos multicast, conservando largura de banda da rede.

O suporte H.323 para conferências multiponto híbridas permite que alguns terminais estejam em uma conferência centralizada e outros em uma conferência descentralizada. Neste caso o MCU provê uma ponte entre os dois tipos. O terminal não está atento à natureza mista da conferência, somente ao modo de conferência na qual ele envia e recebe. Suportando ambas as abordagens multicast e unicast, o H.323 é adequado tanto às tecnologias de rede atuais quanto às futuras gerações de redes. O suporte a multicast faz uso mais eficiente da largura de banda, mas exige maior poder computacional dos terminais que devem multiplexar e selecionar seu próprio fluxo de recepção de áudio e vídeo. Adicionalmente, suporte a multicast é requerido nos roteadores e switches da rede.

Considere um exemplo onde uma conferência multiponto é realizada entre três clientes (Figura 71). Um terminal cliente (Cliente B) executa uma função MC. Todos os terminais podem usar multicast para participar em uma conferencia descentralizada. Uma função MP em cada nó irá multiplexar e apresentar os sinais de áudio e vídeo destinados ao usuário. Esta abordagem minimiza a necessidade de recursos especializados na rede. Entretanto, a rede deve ser configurada para suportar multicast.

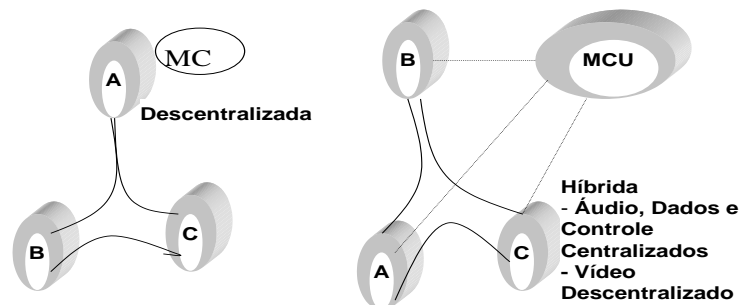


Figura 71. Conferência Descentralizada, Conferência Híbrida

Uma MCU pode ser usada para manipular somente áudio, dados e funções de controle. Nesta configuração o vídeo pode ser multicast conservando largura de banda da rede. O MCU pode ser um sistema dedicado ou um terminal com capacidades extras.

8.4.6 Ferramentas H.323

Esta seção apresenta um levantamento das aplicações (terminais clientes, MCUs, Gatekeepers e Gateways) que implementam o padrão H.323.

Servidores H.323

Chama-se servidor H.323 um endpoint com capacidade de fazer funções de MCU, de Gatekeeper ou de Gateway ou qualquer combinação entre estes três tipos. Este item lista as características de alguns desses servidores disponíveis no mercado.

Existem alguns servidores H.323 que implementam a funcionalidade de unidade de controle multiponto MCU:

- RADVision MCU-323 (<http://www.radvision.com>)

- Ezenia Encounter NetServer (<http://www.ezenia.com>)
- White Pine MeetingPoint Conference Server (<http://www.wpine.com>)

Segundo o padrão H.323, o gatekeeper é definido como um componente opcional, porém ele está se tornando elemento fundamental no centro de controle de redes convergentes para voz, dados e vídeo. A seguir são listados alguns produtos:

- RADVision NGK-100 (<http://www.radvision.com>);
- Ezenia Encounter Gatekeeper (<http://www.ezenia.com>)
- White Pine MeetingPoint Conference Server (<http://www.wpine.com>)

Abaixo estão alguns exemplos de produtos gateways H.323:

- RADVision H.323/H.320 Gateway – L2W-323P (<http://www.radvision.com>)
- Ezenia Encounter Netgate (<http://www.radvision.com>)

8.5 PROTOCOLO SIP

Os dois padrões mais importantes dentro da telefonia IP são o H.323, desenvolvido pelo ITU-T, e o SIP, da IETF. Até mesmo a existência do *gateway* H.323/SIP é essencial para a alavancagem de uma rede convergente de dados e telefonia IP.

Para suporte a telefonia na Internet, a IETF (*Internet Engineering Task Force*) propôs um mecanismo simples para sinalização telefônica na rede IP, criando um protocolo chamado SIP (*Session Initiation Protocol*) [HANDLEY,1998], com raízes no protocolo HTTP. O SIP é um protocolo que estabelece, modifica e termina sessões multimídia entre dois ou mais participantes. Estas sessões podem ser conferências multimídia, aulas pela Internet, telefonia sobre Internet, distribuição multimídia, entre outras.

O SIP é perfeito para a integração com a web, pois ele já nasceu fruto da mesma estrutura utilizada no HTTP. Sua aderência e escalabilidade o tornam um excelente protocolo para a utilização em larga escala da telefonia IP, e criação de complexos serviços telefônicos nesta rede convergente. Entretanto, o H.323, mais antigo, é base para muitos equipamentos e softwares amplamente utilizados na telefonia IP, como a solução de hardware e software da Cisco para telefonia IP, ou o software de telefonia IP Netmeeting, da Microsoft.

Em pouco tempo, o SIP mostrou-se tão abrangente para a telefonia e com tanto poder de integração com as tecnologias voltadas para a própria Web, que muitas empresas passaram a desenvolver implementações a partir dele [SINGH,2000]. Ele se tornou um grande concorrente na arena de Telefonia IP.

Os serviços do SIP para o estabelecimento e o encerramento de sessões multimídia incluem [Soares, 2002]:

- Localização de usuário: um usuário pode se movimentar por toda a rede, logo, ele precisa ser localizado antes de efetivamente iniciar uma comunicação com este usuário. Este procedimento determina a localização do usuário e se o mesmo pode ser usado para comunicação;
- Capacidades do usuário: este procedimento é utilizado para determinar as capacidades de mídia dos usuários envolvidos na comunicação e para determinar os parâmetros de mídia a serem usados;
- Disponibilidade do usuário: após um usuário ser localizado precisa-se saber se ele está disponível para uma nova comunicação. Este procedimento determina se o usuário possui recursos disponíveis para iniciar uma nova comunicação.
- Configuração da chamada: é o processo de definição dos parâmetros que serão usados para o estabelecimento da chamada;
- Controle de chamada: é o processo de gerenciamento da chamada, incluindo processos de transferência de ligações e encerramento da mesma.

8.5.1 Componentes SIP

A tabela abaixo apresenta os componentes básicos do SIP e formam o que é chamado de rede SIP. Um Servidor é uma aplicação que é responsável em receber as mensagens SIP do tipo request dos usuários e enviar mensagens do tipo response. Um Servidor pode integrar as funções de Proxy Server, redirect Server ou user agent Server (uma parte do user agent que processa solicitações request). Normalmente um Servidor é um hardware específico que pode implementar todas as funções citadas neste item. Inclusive, um Servidor pode conter uma porção User Agent Client.

Componente SIP	Função
SIP User Agent	Cliente da arquitetura, ou o ponto final da comunicação multimídia. Tem duas partes: User agent client que é responsável pelo início da comunicação entre cliente e servidor e iniciar solicitações request; User agent Server que processa as mensagens do tipo request.
SIP Proxy Server	Servidor de redirecionamento de requisições e respostas SIP. Passa a realizar a sinalização como se fosse o originador da chamada, e quando a resposta lhes é enviada, ela redirecionada para o originador real. Ele atua como cliente e servidor, sendo responsável em receber mensagens e encaminhar a outros servers, às vezes através de alguma tradução de endereços. Quando necessário, um servidor Proxy reescreve uma mensagem antes de encaminhá-la a frente. Os servidores Proxy podem também ser usados para rotear mensagens e aplicar regras de segurança.
SIP Redirect Server	Responsável em fornecer a um cliente a lista de endereços possíveis para se alcançar um cliente destino. Ele pode fazer o mapeamento de um endereço em nenhum ou em outros novos endereços, para alcançar um cliente destino.
SIP Registration Server	Servidor SIP que suporta requisições de registro, usadas para registrar as informações dos usuários em algum Servidor de Localização.
Servidor de Localização	Oferece funcionalidades de armazenamento e consulta de registros de usuários SIP. Para localização, são usadas bases de dados locais ou servidores LDAP (<i>Lightweight Directory Access Protocol</i>), onde é possível montar diretórios de usuários e seus perfis

8.5.2 Mensagens SIP

O protocolo SIP é baseado no HTTP e, assim como este, suporta o transporte de qualquer tipo de carga em seus pacotes, pelo uso de Mime-Types (*Multipurpose Internet Mail Extensions*). O SIP funciona numa arquitetura cliente/servidor, e suas operações envolvem apenas métodos de requisição e respostas, como observado também no HTTP. Os métodos de requisição do SIP são apresentados na tabela abaixo.

Nome do Método	Comportamento
INVITE	Indica que o usuário está sendo convidado a participar de uma sessão multimídia. O corpo da mensagem pode conter uma descrição da sessão, utilizando-se o protocolo de descrição de sessão SDP (<i>Session Description Protocol</i>).
ACK	Mensagem recebida como resposta a um INVITE. A requisição ACK pode conter o SDP de descrição da sessão negociada entre ambos os clientes. Se não contiver o SDP, o usuário chamado pode assumir a descrição dada pelo primeiro INVITE, se houver.
OPTIONS	Faz uma pergunta sobre quais métodos e extensões são suportados pelo servidor e pelo usuário descrito no campo de cabeçalho <To:>. O servidor pode responder a esta pergunta com o conjunto de métodos e extensões suportado pelo usuário e por ele mesmo.
BYE	Usado para liberar os recursos associados a uma ligação e forçar a desconexão da mesma.
CANCEL	Cancela uma requisição que ainda esteja pendente, ou seja, em andamento. Uma requisição é considerada pendente, se e somente se, ela não foi atendida com uma resposta final.
REGISTER	Um cliente usa este método para registrar o "alias" (apelido) do seu endereço em algum servidor SIP (registration Server).

Uma request-URI é um SIP URL, sendo este último usado dentro de uma mensagem SIP para identificar quem originou a mensagem (From), qual o destino corrente da mensagem (request-URI) e qual o destino final da mensagem (TO). Um SIP URL possui vários componentes, onde se pode estacar: USER (identifica o nome do usuário endereçado) e HOST (identifica o nome do domínio ou um endereço IP específico). Parâmetros genéricos podem ser usados em uma SIP URL para identificar dentre outros: protocolo de transporte (UDP ou TCP), número de porta específica para o protocolo de transporte que está sendo usado. Veja abaixo alguns exemplos de SIP URL:

- sip:usuário@dominio.com
- sip:usuário@10.1.1.1

■ sip:willrich@gateway.com;transport=udp

Quando requisições são atendidas, as respostas enviadas são identificadas por números, que significam a classe da resposta. Pode-se enviar diversas mensagens provisórias antes de se enviar uma resposta definitiva. Existem seis classes possíveis de resposta: Classe 1XX, respostas temporárias ou informativas; Classe 2XX, resposta final de sucesso; Classe 3XX, redirecionamento da requisição; Classe 4XX, erros no cliente; Classe 5XX, erros do servidor; e Classe 6XX, erros globais na rede.

Na Figura 72, tem-se o exemplo de um fluxo de convite para um usuário na “rede” SIP, mostrando características de mobilidade do usuário, mensagens de requisição, e mensagens de resposta finais. Acompanha-se na Figura 72 as descrições numeradas a seguir:

- 1) Usuário bruno pede para ser criada uma sessão entre ele e o usuário de "alias" cesar@land.ufrj.br. [Requisição SIP INVITE]
- 2) O servidor proxy então pergunta ao servidor de localização de usuários (*Location Server Database*) onde está o usuário com esse "alias" [usando o Protocolo LDAP].
- 3) A resposta deste servidor é a atual localização do usuário (esta é a característica de mobilidade na rede SIP. Seu último REGISTER partiu de ipanema.land.ufrj.br).
- 4) A requisição de abertura de sessão é então redirecionada pelo proxy para o endereço correto [Requisição SIP INVITE]. Então, o usuário cesar na máquina ipanema.land.ufrj.br será alertado, recebendo o toque de chamada [RING-RING].
- 5) Cesar decide se juntar à sessão e o seu cliente SIP responde para o servidor proxy que a sessão pode ser aberta [Resposta de Sucesso 200 OK para o Servidor Proxy].
- 6) O servidor proxy redireciona essa resposta ao cliente chamador [Resposta de Sucesso 200 OK redirecionada para bruno].
- 7) O cliente chamador bruno indica para o servidor que a negociação da sessão acabou e a sessão está aberta [Requisição ACK contendo a negociação final de mídia].
- 8) Enfim, o servidor proxy indica para o cliente chamado que a negociação da sessão acabou e a sessão está aberta [Requisição ACK contendo a negociação final de mídia].

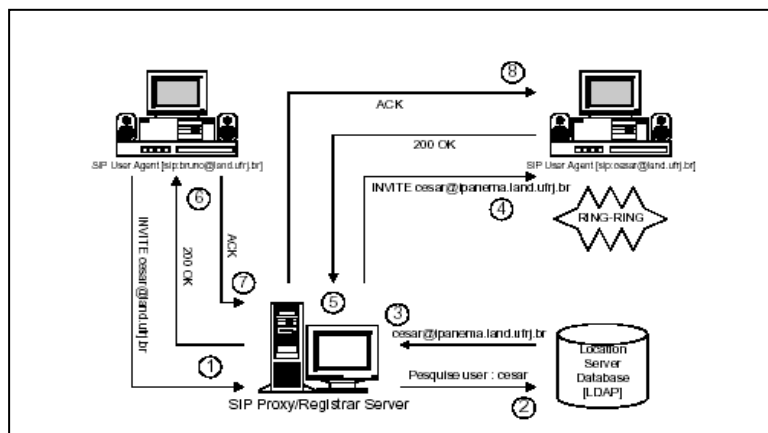


Figura 72. Estabelecimento de chamada SIP

8.5.3 H.323 vs. SIP

H.323 e SIP competem pela sinalização da telefonia IP. Existem muitos debates na indústria para saber qual protocolo é o melhor, H.323, SIP, ou talvez outro protocolo que possa estar em estágio inicial de desenvolvimento. Atualmente, não há um vencedor claro, mas os padrões aparecem estar evoluindo de tal modo que as melhores qualidades de um estão sendo inseridas no outro [Mehta, 2001]. Por exemplo, a evolução do H.323 da versão 1 para 4 tem focado na redução do tempo de configuração de chamada de 6 ou 7 idas e voltas para 1,5 como no SIP. Obviamente, esta convergência é altamente desejável para a interoperabilidade entre os dois protocolos e reduz a sobrecarga de sinalização.

H.323 e SIP suportam, de maneira semelhante, a maioria das funções requeridas pelo usuário final, tal como estabelecimento e encerramento de chamada, transferência de chamada, chamada em espera e conferência. Mas existem algumas diferenças funcionais, tal como o suporte do H.323 para indicação de

mensagem em espera e suporte para controle de moderador do SIP. Além disso, o H.323 v3 fornece um mecanismo mais robusto para troca de capacidades – o processo pelo qual é determinado se uma característica particular é suportado pelos dois participantes.

A tabela a seguir apresenta uma comparação pontual entre H.323 e SIP, embora não se pretenda com esta comparação apontar qual a melhor implementação, mas apenas descrever como são implementadas funções em ambas as especificações [Soares, 2002].

	H.323	SIP
Filosofia	Desenvolvido para gerenciar chamadas de voz, multimídia e serviços suplementares através de recomendações específicas para cada tipo de serviço	Desenvolvido para estabelecer uma sessão entre dois pontos finais, sem nenhum relacionamento específico com algum tipo de mídia.
Complexidade	Mais complexo	Relativamente mais simples
Formato das mensagens	Representação binária ASN (Abstract Syntax Notation)	Representação textual http
Transporte de mídia	RTP/RTCP	RTP/RTCP
Protocolo de transporte	TCP ou UDP	TCP ou UDP
Endereçamento	Entende URLs e H.164	Só URL
Escalabilidade	Não é muito escalável	Altamente escalável
Suporte a Firewall	Pode ser provido por um gateway H.323	Pode ser provido por um SIP Proxy
Autenticação	Através de H.235	Via http, ssl, pgp e outros
Encriptação	Via H.235	Via ssl, PGP e outros

8.6 Protocolo RTP

Mais recentemente, vários grupos de trabalho Internet têm sido formados para estudar os requisitos da comunicação multimídia e tempo-real. O Grupo de Trabalho Transporte de Áudio e Vídeo está desenvolvendo um protocolo que forneça um serviço de transporte fim-a-fim para dados com características tempo-real como áudios e vídeo, o protocolo RTP (*Real-Time Transport Protocol*) [Schulzrinne, 97]. O objetivo principal do RTP é satisfazer as necessidades da conferência multimídia multiusuários, mas seu uso não é limitado a isto.

RTP não contém praticamente nenhuma suposição específica sobre as capacidades das camadas inferiores. Ele não contém nenhum endereço da camada de rede de forma que RTP não é afetado por mudanças de endereçamento. Além disso, qualquer capacidade adicional da camada inferior como segurança ou garantias de qualidade de serviço pode ser usada obviamente por aplicações que empregam RTP. Inclusive RTP suporta transferência de dados para destinos múltiplos usando distribuição multicast se provido pelas camadas inferiores.

Os algoritmos RTP não são usualmente implementados como uma camada separada mas são parte do código da aplicação. É esperado que a nova geração de navegadores WWW também usarão RTP para fluxos de áudio e vídeo.

A Figura 73 mostra as relações entre vários protocolos. Nesta estrutura, o objetivo do RTP é fornecer uma comunicação fim-a-fim, tempo-real sobre a Internet. Note que ele fornece uma função de transporte de dados e é chamado protocolo de transporte, mas ele é realmente usado com frequência no topo do UDP, que é um protocolo de transporte sem conexão.

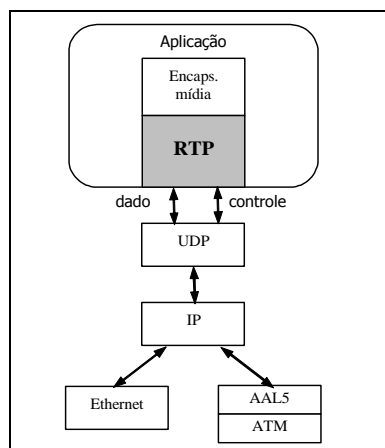


Figura 73. Pilha de protocolo típica para mídias contínuas usando RTP

RTP consiste de duas partes, uma parte de dados e outra de controle. O último é chamado de RTCP (*RTP Control Protocol*).

8.6.1 Parte de dados do RTP

A parte de dados do RTP é um protocolo fornecendo suporte para aplicações com propriedades tempo-real tal como mídias contínuas. Aplicações enviam dados em unidades de dado de protocolo RTP. A Figura 74 apresenta o cabeçalho de uma RTP PDU. Os significados dos campos são os seguintes:

- Version (V, 2 bits): identifica a versão utilizada do padrão RTP. A versão atual é identificada por 2, o valor 1 é usado para a primeira versão do draft e o valor 0 é usado pelo protocolo inicialmente implementado pela ferramenta de áudio vat.
- Padding (P, 1 bit): se este bit estiver a um, o pacote contém um ou mais bytes adicionais no final (Por exemplo, necessários por exemplo para criptografia).
- Extension (X, 1 bit): quando a um, uma extensão do cabeçalho segue o cabeçalho (contém informações dependentes de aplicação).
- CSRC count (CC, 4 bits): contém o número de identificadores CSRC que seguem o cabeçalho fixo.
- Marker (M, 1 bit): a interpretação deste bit é definido pelo perfil de mídia que define as características de diferentes mídias. Este campo é usado por exemplo para definir eventos como fronteiras de um quadro de vídeo.
- Payload type (PT, 7 bits): identifica o formato do dado transportado no pacote, por exemplo, vídeo MPEG, H.261 ou Motion-JPEG. O mapeamento do código ao formato do fluxo é normalizado.
- Sequence Number (16 bits): contém o contador de número de sequência dos pacotes RTP que é incrementado por um para cada pacote enviado.
- Timestamp (32 bits): esta marca temporal é incrementada com a frequência de relógio e determinada pelo formato do dado transportado no pacote. Por exemplo, para um áudio a taxa fixa, a marca temporal poderia ser incrementada por um para cada amostra. Vários pacotes RTP consecutivos podem ter marcas temporais iguais se eles são gerados logicamente no mesmo instante (Por exemplo, pertença ao mesmo quadro de vídeo).
- SSRC (32 bits): identifica a fonte do fluxo de pacotes RTP (fonte de sincronização). Todos os pacotes de uma fonte de sincronização fazem parte um mesmo espaço de número de sequência e de tempo de maneira que o receptor pode montar o dado de uma mesma fonte para a apresentação. Este identificador é escolhido aleatoriamente (existe detecção de conflito a este nível). Exemplos de fontes de sincronização incluem o emissor de um fluxo de pacotes derivados de uma fonte de sinal tal como microfone ou câmera, ou um mixer RTP8.

⁸ Um sistema intermediário que recebe pacotes RTP de uma ou mais fontes (possivelmente modificando o formato do dado) e combina os estes pacotes e os envia em um novo pacote RTP. Desde que o tempo entre fontes de entradas não serão geralmente sincronizadas, o mixer gerará seu próprio tempo para o fluxo combinado. Assim, todos os pacotes de dados originários do mixer serão identificados como tendo o mixer como fonte de sincronização.

- CSRC list (até 15 itens, 32 bits cada): identifica as fontes que contribuíram para o fluxo combinado produzido por um mixer RTP. O número de CSRCs é dado pelo CC. Identificadores CSRC são inseridos por mixers, usando os identificadores SSRC das fontes contribuintes. Uma aplicação exemplo é a áudio-conferência onde um mixer indica todos os interlocutores cuja fala foi combinada para produzir o pacote, permitindo que o receptor indique o interlocutor atual, mesmo que todos os pacotes de áudio contêm o mesmo identificador SSRC (que é do mixer).

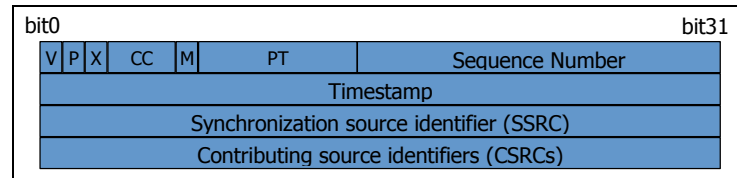


Figura 74. Cabeçalho de uma PDU RTP

As seguintes propriedades do RTP são interessantes para multimídia:

- Reconstrução temporal: o protocolo RTP não fornece nenhum mecanismo para garantir tempo de entrega, nem efetua reordenação de pacotes. Porém, com as marcas temporais e número de sequência incluídas no cabeçalho da mensagem, a aplicação receptora dos datagramas pode reconstruir a temporização da origem e a sequência de pacotes enviados. Para mídias a taxa de bits constante (CBR), como áudio PCM, os fluxos podem ser sincronizados pelos números de sequência, mas para mídias a taxa de bits variável (VBR), tal como vídeo codificado MPEG, são necessárias marcas temporais.
- Detecção de perdas pode ser feita a partir da numeração dos pacotes com um número de sequência definido no cabeçalho da mensagem. Com a numeração de sequência, é possível também que seja determinada a correta localização de um pacote sem necessariamente decodificar os pacotes em sequência, como em decodificações de vídeo. Além disso, um checksum pode ser adicionado em extensões do cabeçalho dependentes de aplicação.
- Identificação de conteúdo, graças a um código no cabeçalho da mensagem (por exemplo, se é um vídeo JPEG ou áudio GSM).
- A marca temporal e um identificador da fonte de sincronização (um microfone, um mixer, ou uma câmera) podem ser utilizados para realizar sincronizações intermídia e intramídia para aplicações multimídia. O receptor é então capaz de identificar a fonte de cada pacote podendo colocar no contexto apropriado e apresentar em um tempo apropriado para cada fonte.
- Mídias múltiplas podem ser multiplexadas sobre uma conexão e demultiplexadas no destino, isto graças a identificação das fontes contribuintes contida nos pacotes.

Dados de mídias contínuas são transportados em pacotes de dados RTP. Se pacotes RTP são transportados em datagramas UDP, pacotes de dados e de controle usam portas separadas. Se outros protocolos servem o RTP (como RTP diretamente sobre ATM AAL5), é possível transportar os dois uma unidade de dado de protocolo única, com controle seguido do dado.

8.1.1 Parte de controle do RTP

O protocolo de controle do RTP, chamado de RTCP, fornece suporte para conferência tempo real de grupos de qualquer tamanho dentro da Internet. RTCP é baseado na transmissão periódica de pacotes de controle para todos os participantes da seção. O protocolo de transporte deve fornecer a multiplexação dos pacotes de dados e de controle, por exemplo, usando números de portas UDP separadas. Para o RTP é atribuído um número de porta par e ao RTCP é atribuído o próximo número (ímpar).

O RTCP inclui:

- Monitorização de QoS e Controle de Congestionamento: os membros da sessão emitem periodicamente relatórios (aproximadamente a cada 5s) para todos os emissores contendo o número de sequência do último dado recebido, número de pacotes perdidos, e uma medida da variação temporal entre recepções e marcas temporais (necessários à estimação do atraso de transmissão entre emissores e receptores).

- Sincronização intermídia: aplicações que enviaram dados recentemente emitem um relatório que contém informações úteis para a sincronização intermídia: o tempo real e uma marca temporal correspondente. Estes dois valores permitem a sincronização de mídias diferentes, como a sincronização labial.
- Identificação da fonte: pacotes de dado RTP não identificam a fonte, mensagem RTCP contém o e-mail do usuário emissor e opcionalmente outras informações (nome, endereço, telefone).

8.6.2 RTP e a comunicação multimídia

O protocolo RTP tem algumas das características desejáveis para a comunicação multimídia. Algumas delas são:

- Ele é um protocolo leve, assim uma alta vazão pode ser obtida. Não há mecanismos de controle de fluxo e de erro, embora o emissor pode ajustar a taxa de transmissão baseado na informação de realimentação provida pelos pacotes de controle.
- Ele transporta informações usadas para realizar sincronizações intramídia e intermídia para aplicações multimídia.
- Multicast é possível se as camadas baixas fornecem roteadores multicast.
- RTP especifica ligações com QoS através de um perfil de tráfego. Ele assume que existe um número predefinido de perfis de tráfego que definem requisitos de QoS para aplicações típicas. Por exemplo, se o perfil especificado é “áudio telefone”, os parâmetros armazenados no perfil são: taxa de bit = 64kbps, máximo atraso = 100ms e máxima taxa de erro = 0,01. A gestão de recursos e garantias da QoS são de responsabilidade das camadas mais baixas. Por exemplo, ST-II pode ser usado para fornecer garantias de QoS (para reservar recursos ao longo da rota de pacotes de dados).
- RTP, apesar do nome, não assegura a comunicação tempo-real. Nenhum protocolo fim-a-fim, incluindo RTP, pode assegurar a comunicação em tempo. Isto requer suporte das camadas inferiores que realmente tem controle sobre a rede nos comutadores e roteadores.
- Outro problema é que o cabeçalho do pacote RTP é ligeiramente grande (12 bytes no mínimo) na qual outros cabeçalhos podem ser adicionados, tal como o cabeçalho UDP/IP/Ethernet. Isto resulta num baixo uso da largura de banda da rede se pacotes menores são usados. Esta sobrecarga de protocolo é notável particularmente na voz com seus tamanhos de pacote menores.

8.7 Componentes Essenciais da Telefonia IP

Telefonia sobre a Internet é uma técnica de transmissão de voz sobre rede de dados, que permite a sistemas de telefonia comerciais serem usados para chamadas de pessoas via redes IP. Os componentes essenciais da telefonia IP, além daqueles apresentados na seção 8.4, são (Figura 75): servidores de telefonia IP e PBX, Gateways VoIP, e fones IP.

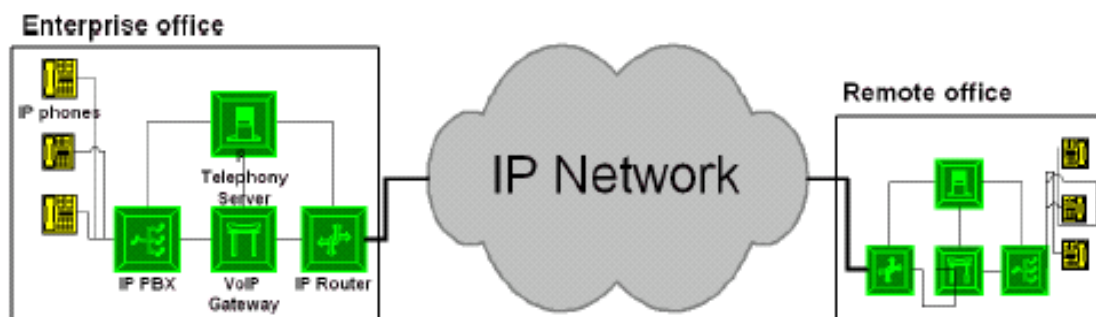


Figura 75. Elementos básicos da telefonia IP [Walker, 2002]

8.7.1 Servidores de telefonia IP e PBX

Muitas transações em redes de dados são baseadas no modelo cliente-servidor. Computadores clientes fazem pedidos de serviços a computadores servidores, que desempenham o serviço e retorna o

resultado (p.e. servidores web, email, banco de dados). Com a inclusão de dados de voz nas redes IP adiciona outro conjunto de servidores projetados para prover serviços de voz.

Na PSTN, o PBX é normalmente um sistema “Caixa Fechada” que fornece todas as funções de voz e funcionalidades necessárias de maneira proprietária. O gerenciamento da plataforma fechada é específico do vendedor do PBX. Com a implantação de VoIP, um PBX IP serve como um servidor principal da telefonia IP, fornecendo funções e características semelhantes aos providos por um PBX tradicional. Além destas, os PBXs IP estão provendo novas funções (visto mais adiante). Um PBX IP pode ser construído com um PC rodando um sistema operacional, como o Microsoft Windows, Linux ou Solaris.

Outros tipos de servidores de telefonia IP fornecem novos e interessantes serviços. A possibilidade de serviço de mensagem unificado (*Unified messaging*) – a convergência de mail de voz e e-mail – pode ser considerado um benefício da implementação de VoIP. Servidores de serviço de mensagem unificado também executam sobre plataformas PC e se comunicam com servidores de e-mail e PBXs IP para fornecer acesso a mensagens de vários modos.

Outro novo conceito introduzido com os servidores de telefonia IP é o *clustering*, em que vários destes servidores são agrupados em um *cluster* para oferecer aumento da escalabilidade, confiabilidade e redundância. Servidores em *cluster* funcionam em conjunto e podem ser visto como uma unidade, provendo poder de processamento combinado que logicamente aparece como um único servidor. *Clustering* não é disponível atualmente com PBXs tradicionais em uma PSTN.

Gatekeeper (porteiro) é outro tipo de servidor usado para fornecer características de controle de admissão e outras funções de gerenciamento para serviços multimídia. Estas características são apresentadas mais adiante.

8.7.2 Gateways VoIP

Gateways VoIP e roteadores IP transmitem datagramas de voz RTP por uma rede IP. Gateways VoIP fornecem uma conexão entre a rede VoIP e a PSTN ou a um PBX. Estes dispositivos realizam um papel importante na migração para a VoIP.

É necessário conectar a rede IP à PSTN para colocar chamadas para/de usuários PSTN. Para tanto, Gateways VoIP devem suportar o protocolo SS7 (protocolo de sinalização do PSTN), que é usado pelo gateway VoIP para sinalizar comutações no PSTN quando uma chamada telefônica é originária da rede VoIP com o chamado na PSTN. O gateway executa a digitalização (no caso do sinal de entrada ser um sinal analógico), compressão, demodulação (no caso do sinal de entrada se tratar de um sinal de fac-símile) e funções de empacotamento IP sobre o sinal de voz recebido.

Gateways VoIP podem também fornecer conversação entre diferentes codec, que é chamado *transcoding*. Se um codec diferente do G.711 (por exemplo, G.729), for usado em uma rede VoIP, os dados de voz devem ser convertidos para G.711 antes de serem transferidos para a PSTN.

Em um ambiente corporativo, os gateways VoIP podem interconectar PBXs tradicional para fornecer um caminho de migração e permitir uma migração gradual ao VoIP. Um exemplo desta utilização de gateways pode ser visto na Figura 76.

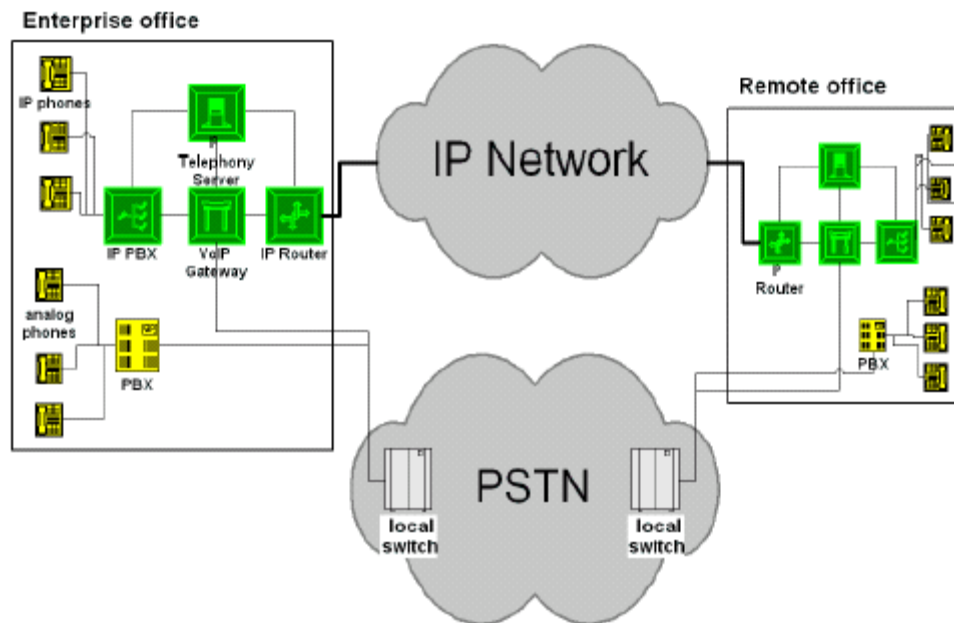


Figura 76. Uma rede VoIP com seus gateways conectados na PSTN

Na realidade, existem várias opções de empregar telefonia IP [Walker, 2002]:

- Toll bypass usando gateways VoIP: os gateways VoIP são usados para conectar PBXs tradicionais. O tráfego PSTN pode passar através do PBX, para o gateway VoIP, e sobre um backbone de rede IP. Esta opção pode prover redução de custos para chamadas entre sítios das corporações e oferece um primeiro passo para a migração de um emprego de VoIP maior. Alguns cenários adicionais deste tipo de implementação de VoIP podem ser encontrados em [Soares, 2002].
- VoIP com PBXs IP: é um passo a mais para a rede convergente. Integradores oferecem serviços para instalar, configurar, e testar VoIP usando PBX IP em conjunto ou substituindo os PBXs tradicionais.

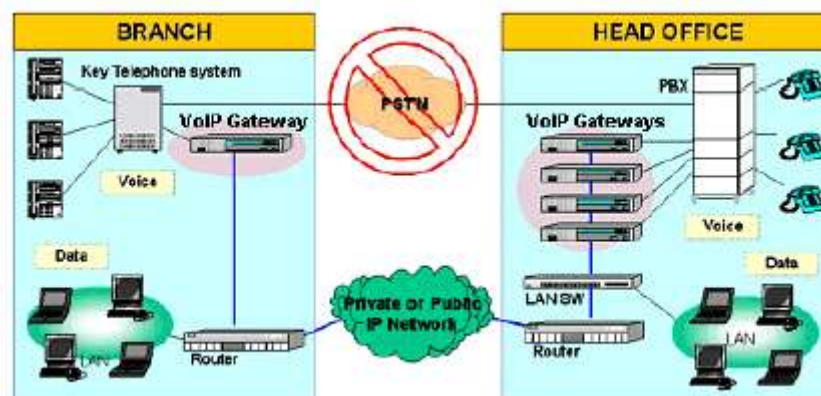


Figura 77. Toll bypass usando gateways VoIP. Fonte: OKI Network Technologies.

8.7.3 Roteador VoIP

Examinando cabeçalhos de pacotes IP, roteadores IP tomam a decisão necessária para mover pacotes para o próximo roteador ou hop ao longo do caminho para o destino, traçando a rota do pacote de voz através da rede.

Um roteador pode ter integrado as funcionalidades de Gateway VoIP, por exemplo um SwitchRouter Gateway IP, as quais a partir da adição de módulos de voz (placas) são capazes de prover a interface de conexão entre o dispositivo de voz (PBX) e a rede de dados, habilitando o transporte de chamadas VoIP.

Endereçamento VoIP

Uma Internet corporativa existente deveria ter um plano de endereçamento IP. Para o esquema de numeração IP, as interfaces de voz aparecem como hosts IP adicionais, ou como uma extensão do esquema existente ou como novos endereços IP.

A translação do número telefônico do PBX para um endereço IP é desempenhada por um DPM (Dial Plan Mapper), onde o número do telefone destino é mapeado para o endereço IP de destino. Quando o número é recebido do PBX, o roteador compara o número com aquele mapeado na tabela de roteamento. Se um casamento é encontrado, a chamada é roteada para o host IP. Após a conexão ser estabelecida, a conexão da Intranet corporativa é transparente para o usuário.

Fones IP geralmente utilizam o DHCP. O servidor DHCP fornece um endereço IP quando um host de rede, neste caso um fone IP, é ativo na rede. Usando o serviço DHCP, fones IP podem ser movidos com relativa facilidade. Quando você realoca um fone IP, movendo ele para outra subrede, por exemplo, o servidor DNS para aquela subrede pode ser capaz de encontrá-lo, amenos que você desabilite o DHCP do fone e coloque um endereço de rede estático. Neste caso, você terá um problema de configuração antes de mover. Por razões similares, se seu servidor DNS não funcionar, você poderia perder o serviço de telefonia. A disponibilidade do servidor de DNS e DHCP necessita ser monitorado. Inclusive, é interessante instalar servidores redundantes aumentar a disponibilidade deste serviço.

Roteamento VoIP

Um dos pontos fortes das redes IP é a maturidade e sofisticação dos seus protocolos de roteamento. O uso de um protocolo de roteamento moderno, tal como Enhanced Interior Gateway Routing Protocol (EIGRP), permite considerar o atraso quando do cálculo do melhor caminho. Além disso, arquiteturas de Qualidade de Serviços (QoS), como Serviços Diferenciados e Serviços Integrados/RSVP (Resource Reservation Protocol) podem ser usados para reservas de recursos para garantir a qualidade de voz.

O uso de tecnologias tag switching e MPLS (Multiprotocol Label Switching) também são interessantes para VoIP, permitindo estender o roteamento IP e facilitam o provimento de QoS, além de permitir a realização da Engenharia de Tráfego.

8.7.4 Equipamentos do Usuário

O equipamento de usuário é o dispositivo usado pelo usuário para solicitar e fazer uso dos serviços de voz. Este equipamento pode ser tanto um aparelho telefônico, sendo ligado a uma central PBX, direto ao gateway ou direto à rede local (telefone IP), ou um telefone baseado em PC, onde o telefone é ligado direto a um microcomputador via uma placa especial ou simplesmente um microcomputador com placa de som e software para telefonia IP (tipo netmeeting).

Para possibilitar a VoIP, o áudio analógico deve primeiro ser convertido em datagramas digitais, realizados pelos codecs. Os codecs podem estar situados no PBX IP ou no próprio telefone. Caso forem utilizados telefones convencionais, os codecs estão localizados no PBX IP. Chamadas que chegam são digitalizadas no PBX IP antes de serem retransmitidos para a rede IP. Outra alternativa é posicionar os codecs no próprio telefone. Estes novos telefones digitais são chamados fones IP. Em vez de ter conectores de telefone de 4 linhas, eles usualmente têm uma conexão LAN Ethernet. Um fone IP faz conexões de dados para um servidor de telefonia IP, que fazem o processamento de configuração de chamada.

Um computador pessoal também pode servir como telefone IP, bastando uma placa de áudio, microfone, fone de ouvidos e uma conexão à LAN. Este computador deve executar um software que realiza a codificação. Como um fone IP, o computador provavelmente se conecta a um servidor de telefonia IP para fazer o processamento da configuração de chamadas.