

Telefonia IP para Ambientes Móveis Compactos

1. [Introdução](#)
2. [Telefonia IP](#)
 1. [Voz](#)
 2. [Comutação](#)
 3. [Protocolos de Sinalização e Controle](#)
 1. [H.323](#)
 2. [SIP](#)
 4. [Aspectos de Tempo Real](#)
 5. [Aspectos Econômicos](#)
3. [Computação Móvel](#)
 1. [Qualidade de Serviço](#)
4. [Análise de Erros da Telefonia IP sobre Ambiente de Computação Móvel](#)
5. [Implementação](#)
 1. [Bibliotecas para Implementação de Telefonia IP](#)
6. Resultados, Conclusões e Trabalhos Futuros
7. [Referências](#)

Introdução

Este trabalho tem como objetivo construir um protótipo, baseado em duas tecnologias em franca ascensão, Telefonia IP e Computação Móvel, utilizando ainda de técnicas de sistemas embutidos. Este protótipo se propõe a ter as funcionalidades de um telefone celular, porém, não utiliza da sólida e bem conhecida tecnologia de telefonia celular. Ao contrário, nosso aparelho celular é construído sobre o protocolo IP, utilizando da técnica que ficou conhecida como *Wavelan*, para garantir mobilidade ao nosso sistema.

Este documento apresenta inicialmente a tecnologia de Voz sobre o Protocolo IP, que será tratada como Telefonia IP, mostrando seu estado atual de desenvolvimento, vantagens e problemas encontrados, padrões propostos e utilizados e expectativas de custo e utilização. Esta seção do documento tem ainda o objetivo de ser uma referência sobre este assunto.

Na terceira seção, apresentaremos o ambiente de Computação Móvel, apresentando suas dificuldades e vantagens, os padrões utilizados, especialmente aquele utilizado na construção do nosso protótipo. Vamos apresentar também os conceitos de qualidade de serviço e suas limitações sobre um ambiente de computação móvel.

Na quarta seção, apresentaremos os experimentos realizados sobre o ambiente de computação móvel com o objetivo de verificar a qualidade de voz transmitida sobre este ambiente, utilizando diversos *vocoders* e incluindo uma política de redundância de informações transmitidas. Estes testes permitiram definir aspectos importantes para o transmissão de voz no nosso protótipo.

Na quinta seção apresentaremos então nosso protótipo. O hardware e o software utilizados na sua implementação, os padrões utilizados, as pilhas de protocolos usadas na implementação. As limitações de energia, portabilidade e alcance do sinal.

Na última seção apresentamos então os resultados e as conclusões obtidas com este trabalho. As possibilidades de utilização do protótipo e as melhorias necessárias. Também estaremos apresentando as possibilidades vislumbradas de trabalhos futuros, utilizando os conhecimentos obtidos neste trabalho de forma a agregar valor e possivelmente transformá-lo em produto.

Telefonia IP

Uma revolução está para acontecer no ambiente de telecomunicação. Há décadas, desde a invenção do telefone, que a exigência básica para uma comunicação telefônica é o estabelecimento de um circuito entre os dois assinantes ou interlocutores. Isto se dá ainda hoje, na maioria das ligações telefônicas realizadas. É certo que o circuito é digitalizado e multiplexado, mas sua presença é indispensável na comunicação telefônica que nós chamaremos de convencional e que representa a telefonia pública hoje utilizada, conhecida como PSTN (*Public Switched Telephonic Network* ou Rede Telefônica Pública Comutada).

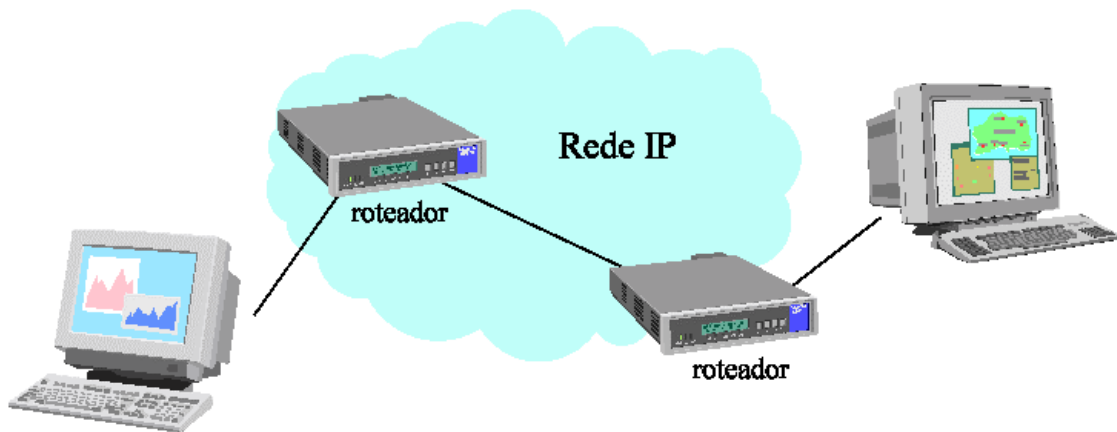
Mas esta necessidade está acabando, com a utilização de redes de pacotes para tráfego de voz. Dentro deste conceito, a voz é empacotada e transmitida em redes compartilhadas, juntamente com dados. E agora é hora da utilização de redes IP. O protocolo que conquistou o mundo de dados com a Internet agora vai entrar também no mundo de voz. Este é o objetivo da Telefonia IP.

O termo Telefonia IP, Telefonia Internet ou ainda Voz sobre IP (VoIP - *Voice over IP*) tem se aplicado a utilização de redes baseadas em protocolo IP na camada de rede para transporte de voz, em especial a Internet. Podemos destacar as vantagens da utilização deste serviço:

- Compartilhamento da rede para o tráfego de voz com o tráfego de dados (email, www, ftp, etc);
- Unificação de redes de transporte, sinalização e gerência sobre a mesma rede, com economia de infraestrutura e manutenção.
- Meio de transmissão de baixo custo, comparado ao sistema telefônico;
- Possibilidade de compactação e supressão de silêncio, reduzindo a largura de banda utilizada.
- Utilização de rede já instalada;
- Possibilidade de oferecer outros serviços como correio de voz, call centers via Internet, segunda linha virtual, etc.
- Possibilidade de disponibilização de diversos serviços como email, fax, voz, web, com o auxílio de tecnologias com reconhecimento e síntese de voz, através de interface telefônica.

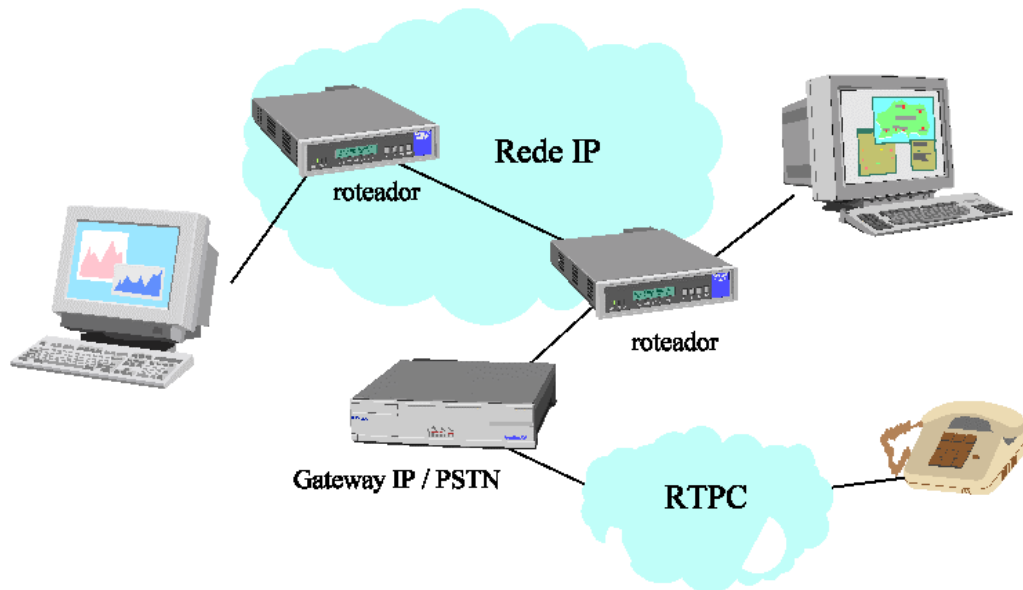
Um primeiro ambiente que podemos imaginar de Telefonia IP é para chamadas entre dois computadores, usando de seus recursos multimídia. Este cenário é apenas o mais simples. Na realidade, Telefonia IP refere-se ao fluxo de voz sobre pacotes transportado sobre redes que utilizam o Internet Protocol (IP) para terminais onde o fluxo é decodificado em voz novamente. Estes terminais podem ser computadores, ou da mesma forma telefones convencionais analógicos. Neste caso, porém, um *gateway* é requerido para converter os pacotes de áudio para o formato analógico que o telefone é capaz de reproduzir[1].

Computador a computador



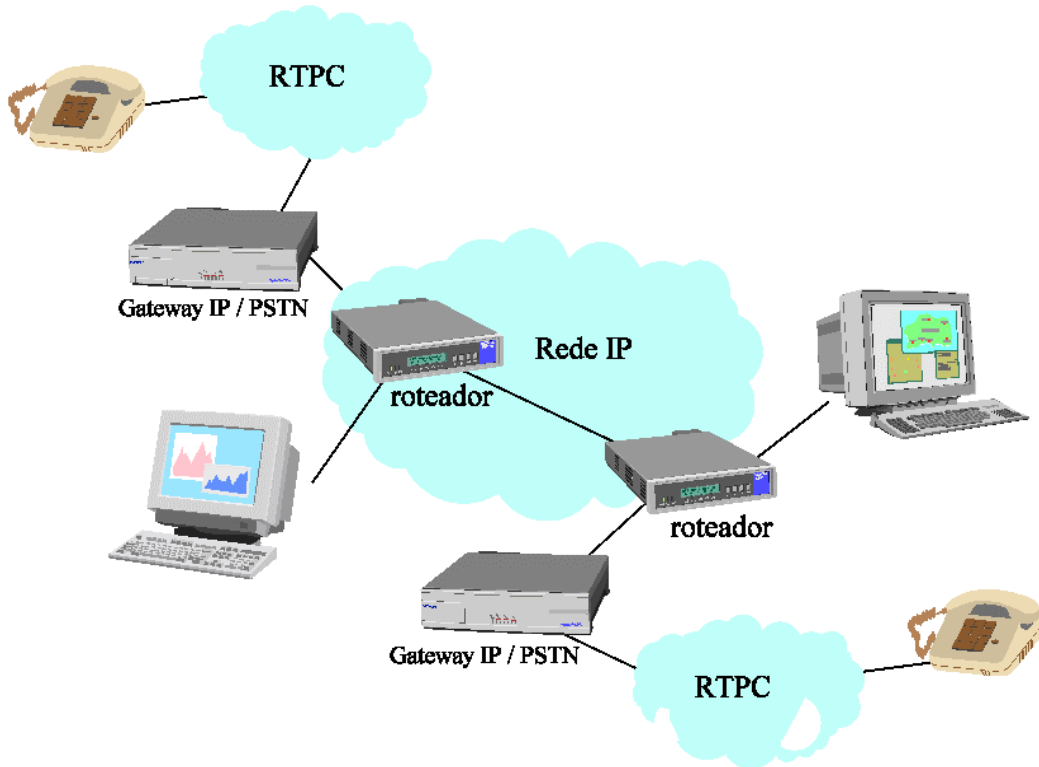
Este primeiro cenário é o mais simples e antigo da Telefonia IP. Dois usuários se comunicando sem que haja utilização da rede telefônica convencional. Vários softwares estão disponíveis para esta utilização, podendo utilizar protocolo proprietário ou padrão. Nesta última opção, podemos encontrar a interação de softwares de diferentes fabricantes. A codificação de voz é feita pelos computadores envolvidos e a transmissão é feita através da rede IP.

Computador a telefone



Neste ambiente necessitamos da utilização de um *gateway*, um equipamento que faz a interface entre a Telefonia IP e a telefonia pública convencional. Todos os protocolos envolvidos com a Telefonia IP terminam neste equipamento e são aí processados e resolvidos. A partir dele a rede telefônica convencional é responsável pela comunicação da voz. Ele converte voz, sinalização e controle da rede IP para a rede telefônica convencional, permitindo a comunicação entre usuários dos dois ambientes: rede IP e telefonia convencional. Posteriormente vamos discutir mais sobre este equipamento.

Telefone a telefone



Finalmente, temos a utilização da Telefonia IP para ligação de dois assinantes da telefonia convencional. Desta forma, a rede IP é utilizada apenas como uma forma de transporte da voz, por ser uma meio mais barato. Um exemplo desta utilização é para chamadas internacionais, para redução de custos. As operadoras públicas podem também ocasionalmente substituir parte de sua rede convencional por uma rede IP.

A utilização de uma rede de protocolo IP para a transmissão de voz reduz em muito os custos de telefonia. Uma rede IP é muito mais barata que uma rede telefônica convencional. Esta redução no custo se dá devido a uma série de fatores:

- Comutação da rede IP é feita através de software, por dispositivos como *roteadores* e *switches*, ao contrário da rede pública, que é toda comutada por hardware. Isto garante um desempenho melhor da rede pública, mas permite uma grande redução no custo da rede IP;
- A Rede IP não garante qualidade de serviço, ou seja, pode incluir atrasos na voz, perda de pacotes e ruídos na voz não previstos. A rede telefônica convencional garante diversos parâmetros de qualidade, que tornam seu custo mais elevado.
- Compartilhamento do canal. Em uma rede IP, o mesmo meio físico é compartilhado para várias ligações, com inclusão de alguns fatores como supressão de silêncio e compressão que podem reduzir a largura de banda utilizada por conversação em poucos kilobytes por segundo. Na telefonia convencional dois canais de 64 kbps, simplex, são reservados para cada conversação, independente da existência ou não de tráfego de voz nos circuitos, ou seja, mesmo nos momentos de silêncio. Além

disso, a mesma rede IP pode ser usada para a transmissão de dados, de forma a se reaproveitar uma infraestrutura existente.

Mas devemos observar que a maior vantagem da utilização de Telefonia IP não é o baixo custo. Enquanto ligações de longa distância baratas estão incentivando o uso, as razões pelas quais as companhias são atraídas para a Telefonia IP são a facilidade de criação de serviços e a consolidação de suas redes, unificando voz e dados.

A adição de novos serviços e funcionalidades atende a uma série de exigências novas no mundo de telecomunicações. Por exemplo pode-se criar uma vídeo-conferência sobre a rede IP somente pela definição de como o fluxo de vídeo é codificado e decodificado. Ao contrário, fazer vídeo-conferência sobre a RTPC pode ser um processo difícil e caro, por envolver a agregação de vários canais de 64 kbps usados para voz em um canal maior que suporte vídeo. Da mesma forma, como Telefonia IP é baseado em software é mais fácil adicionar serviços como correio de voz, siga-me e outros. Podemos ainda prever a integração de diversos serviços que utilizam de tecnologias agregadas como reconhecimento e síntese de voz. Por exemplo, é possível ouvir todos os emails recebidos por um telefone comum, utilizando a síntese de voz e transmissão sobre IP para que estes emails cheguem sobre um telefone comum. Desta mesma forma, serviços como fax, web e outros podem ser acessados através de telefones comuns.

Uma outra vantagem para a Telefonia IP é a consolidação das redes de dados. Atualmente, companhias telefônicas mantêm duas redes, uma para voz e outra para dados. Como redes de voz chegam a ser 10 vezes mais caras que as redes de dados, o uso de VoIP elimina a rede de voz, mantendo uma rede única. A manutenção, gerência, suporte e desenvolvimento são todos aplicáveis sobre uma única rede. Normalmente as operadoras precisam manter dois times diferentes para cada uma das redes, dadas suas características diversas.

Verificando a topologia de cada rede, podemos encontrar, nas operadoras de telecomunicações, três sub-redes distintas, para uma rede de voz:

- Rede transporte de voz - Formada pelos troncos de interligação de centrais telefônicas;
- Rede de sinalização - Responsável por tarefas como controle de chamada entre centrais e tarifação.
- Rede de gerência - Responsável pela gerência de desempenho, configuração, segurança, tarifação e falhas.

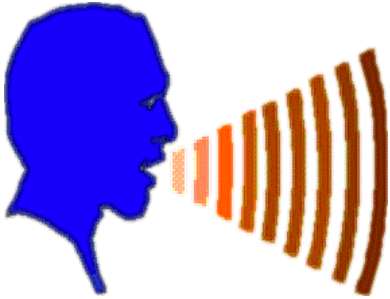
Sobre a telefonia convencional, é necessário manter uma estrutura de rede para cada uma destas sub-redes, com protocolos e canais ocupados distintos, triplicando o custo final. Com o uso de telefonia IP, unifica-se todas estas redes sobre a mesma infraestrutura e utilizando sempre o mesmo protocolo de rede: IP.

Falta agora questionar sobre a qualidade final da transmissão de voz sobre Telefonia IP. Se a rede estiver sobrecarregada, podem ser inseridos atrasos que podem atrapalhar a qualidade final da voz. Todavia, se Telefonia IP é usada em uma rede privada dedicada,

com a utilização de mecanismos de garantia de [qualidade de serviço](#), a qualidade final de voz será bem próxima do nível exigido pelos clientes. Espera-se que as operadoras de telecomunicações desenvolvam suas redes baseadas no protocolo IP para permitir então esta utilização. Podemos observar que novas operadoras já estão se desenvolvendo sobre estas redes e que as atuais operadoras também já estão consolidando suas redes IP.

Mesmo antes que as operadoras consolidem suas redes IP para a utilização de Telefonia Internet, sua utilização já é realidade em ambientes de telefonia privada, interno a grandes empresas. Outro uso é o oferecimento de Telefonia sobre a Internet, como está hoje, sem qualidade de serviço, porém a custos muito baixos, que justificam seu uso para clientes que não sejam muito exigentes.

Voz



A voz humana é uma forma de onda mecânica com frequências principais na faixa que vai de 300 a 3400 Hz, com alguns padrões de repetição definidos em função do timbre de voz e dos fonemas emitidos durante a conversação. O primeiro problema da telefonia em geral é a reprodução com qualidade da voz humana em um terminal à distância.

Em um ambiente de telefonia totalmente analógico isto é possível pela transmissão da forma de onda entre os interlocutores através de meio metálico, com possíveis ampliações analógicas. Isto porém representava um custo alto pela impossibilidade de se utilizar o meio físico para a transmissão de mais de um canal. Desde o advento da telefonia digital, existe a necessidade de se codificar a voz sobre um formato binário, digital que possa ser multiplexado de forma a compartilhar meios de transmissão.

Basicamente duas formas podem ser usadas para a codificação da voz humana:

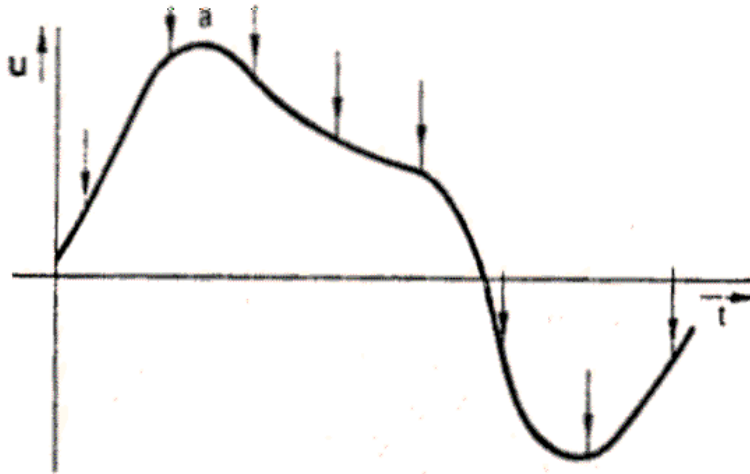
- Baseado na forma de onda
- Baseado nos padrões de voz

Vamos vislumbrar neste momento a utilização da codificação da voz baseada na forma de onda, por ser utilizada na telefonia convencional digital. Adiante voltaremos neste assunto.

PCM

A técnica PCM (*Pulse Code Modulation*) de codificação da voz é baseada na amostragem da voz humana. De acordo com a teoria de Nyquist de processamento de sinais, para representar fielmente um sinal, a taxa de amostragem precisa ser de, pelo menos, duas vezes a maior frequência presente neste sinal.

A fala gera um sinal que originalmente tem um espectro de até 7 kHz. Para efeitos de comunicação telefônica ficou definido que a parte do espectro até 3,4 kHz seria suficiente para uma boa conversação. Outro dado importante é a faixa dinâmica do sinal de voz: em torno de 60 dB. Isto exige no mínimo 10 bits para codificação. Para efeitos práticos, estabeleceu-se uma taxa de amostragem de 8 kHz com resolução de 11 ou 12 bits como parâmetros mínimos. Como porém a sensibilidade acústica é logarítmica e não linear, usando compressão logarítmica do tipo A-law ou μ -law, resulta em 8 bits finais por amostragem. Temos assim uma taxa de 64kbps.



Amostragem Periódica de um sinal telefônico analógico.

Economia de Largura de Banda

Existem alguns aspectos relacionados com a natureza do sinal de voz que permitem adicionar mecanismos de compressão. São eles:

- Distribuição não uniforme de amplitudes;
- Correlação entre amostras sucessivas;
- Correlação entre ciclos sucessivos;
- Fator de inatividade ou percentual de silêncio;
- Densidade espectral média não uniforme, confirmando a redundância de informações;
- Densidade espectral instantânea, ou presença de formantes que se mantêm inalterados durante 20 a 30 ms;

Uma importante é a supressão de silêncio. Em uma conversação telefônica, em apenas 40 % do tempo, o canal de voz está ativo, ou seja o assinante está falando. Com um mecanismo de supressão de silêncio podemos então reduzir a largura de banda utilizada em cerca de 60 %. Um mecanismo conhecido como VAD (*Voice Activity Detection*) é usado para perceber a presença do silêncio e removê-lo. Esta função pode também ser realizada por alguns *vocoders* dispensando a presença de um VAD PCM.

Fluxo PCM é analisado

1011010111010011 11001001 00100100 00111100 10010011 11100001 00100100 00111100 10010011 10110101 11010011 1100100100100100 00111100 10010011 1110000100100100

Eco é removido

1011010111010011 11001001 .0010011 11100001 00100100 00111100 10010011 10110101 11010011 1100100100100100 00111100 10010011 1110000100100100

Silêncio é suprimido pelo VAD

1011 11101 111001001 .0010011 .0010011 .0010010011 10110101 11010011 11001001 00111100 1110000100100100

Fluxo restante é enviado para o Vocoder

10110101 11010011 11001001 00100100 00111100 10010011 11100001 00100100 00111100 10010011 10110101 11010011

Compressão de sinais

Para reduzir a banda do canal necessária para a transmissão de voz digitalizada são utilizadas técnicas de compressão de voz. Isto deve acontecer em tempo real para possibilitar a comunicação e interação. Estes algoritmos permitem reduzir a banda necessária para transmissão de voz a fim de viabilizar sistemas de comunicação digital com taxas bastante reduzidas e tendo como objetivo uma qualidade semelhante ao sistema público de telefonia analógica.

A compressão de sinais é baseada em técnicas de processamento que retiram informações redundantes, previsíveis ou inúteis. A compressão pode acontecer com ou sem perda de informação. Tudo depende da degradação que se admite para o sinal e do fator de compressão que se deseja atingir.

Os dois principais enfoques com que o tema de codificação digital de voz é atualmente abordado são:

- Codificação da forma de onda (ou não-paramétrica)
Esta classe de métodos aborda o problema de codificação de voz de uma forma direta. A idéia é codificar o sinal de voz de forma a se obter um sinal reproduzido cuja forma de onda se assemelhe ao máximo à do sinal original. A codificação da forma de onda é utilizada quando uma qualidade elevada do sinal é exigida. Dentro desta categoria se encontram os métodos mais simples de codificação digital de voz.
- Codificação da fonte (ou paramétrica)
Todos os recursos utilizados na codificação da fonte[6], consideram quase que apenas as características do sinal de voz. Na verdade, no caso da codificação da forma de onda, não existe uma preocupação com a forma de produção do sinal amostrado. A abordagem agora consiste em um modelamento do processo de produção da voz [7]. Este modelamento é então simplificado de forma que possa ser implementado a um custo computacional apropriado [12]. Entretanto, o mecanismo de produção da voz humana não é simples e, portanto, modelagens, aproximações e simplificações acabam por tornar o processo de codificação da fonte inferior em qualidade em relação à codificação da forma de onda. Por outro lado, a codificação

da fonte permite uma redução considerável na taxa de transmissão, porém perde-se fatores essenciais como timbre de voz, tornando-a mais impessoal. Dessa forma, sistemas de codificação da fonte são utilizados quando é necessária uma baixa taxa de transmissão.

Modelos para Vocoders

Os VOCODERS analisam o conteúdo espectral do sinal da fala e identificam os parâmetros que são entendidos pelo ouvido. Estes parâmetros são transmitidos e usados na recepção para sintetizar o padrão da voz. A forma de onda resultante pode não ser semelhante à original mas as diferenças não são percebidas ou, ainda que o sejam, são consideradas aceitáveis para a aplicação.

Basicamente utiliza-se um modelo de predição, ou seja, geração de códigos baseada em análise de amostras dos padrões repetitivos conhecidos como *pitches*:

- Codificação Preditiva Linear (LPC)

Em LPC o trato vocal é modelado como um sistema linear, pois utiliza-se um modelo de filtro digital tipo IIR (Resposta Impulsiva Infinita); o qual, excitado por um sinal de entrada apropriado (que no caso representa o ar fluindo a partir dos pulmões através das cordas vocais que podem estar contraídas ou distendidas) fornece o sinal de voz observado. O sinal de voz é amostrado em janelas de 20 a 40 ms, espaçadas de 10 a 30 ms (portanto com superposição de janelas). Neste, os próprios parâmetros do sistema linear são transmitidos juntamente com o sinal de excitação apropriado. Conseguem-se taxas a partir de 2,4 kbps, mas com qualidade baixa (voz “metálica”).

- Codificação com Excitação a Pulsos Regulares (RPE)

Neste algoritmo amostra-se uma janela de 20 ms da voz, que é dividida em 40 amostras. São feitas interações sucessivas até se obter expressões que minimizem o erro em relação ao sinal original. Ao final do processo obtém-se uma sequência de parâmetros numa taxa de 13kbps. É o algoritmo do padrão europeu GSM de telefonia móvel.

- Codificação Preditiva Linear com Excitação Multipulsos (MLPC)

Este algoritmo é uma evolução do LPC pois os pulsos de excitação não são de apenas dois tipos (vocálicos e não vocálicos) mas assumem vários valores em função das amostras de voz, independentemente do som ser vocálico ou não.

- Predição Linear com Excitação de Códigos (CELP e LD-CELP)

É uma outra variação do algoritmo LPC onde a codificação é baseada em tabelas (“codebook”) que contém os valores de códigos (vetores) que representam as possíveis amostras do sinal de voz. Uma vez identificado o melhor vetor que representa a amostra, o índice deste vetor na tabela é transmitido. Tabelas usuais são de 256 ou 512 vetores. Vale ressaltar que nestes algoritmos é exigida grande capacidade de processamento. O algoritmo CELP acarreta atraso da ordem de 105 ms na conversação.

Existe uma variação chamada LD-CELP (“Low Delay”) cujo atraso é da ordem de 5 ms. Seu “codebook” é de 1024 vetores.

- Predição Linear com Excitação de Soma Vetorial (VSELP)

É uma variação do algoritmo CELP onde a busca na tabela (“codebook”) é feita de forma ligeiramente diferente: são duas tabelas de 128 vetores.

- Predição Linear com Excitação de Códigos Algébricos (ACELP e CS-ACELP)
É outra variação do algoritmo CELP com um “codebook” especializado.

Algoritmos de Compressão Padronizados

Enumeram-se a seguir os principais algoritmos de compressão de áudio de voz padronizados por órgãos internacionais como ITU (International Telecommunication Union), TIA (Telecommunication Industries Association) e USFS (United States Federal Standards). A utilização de algoritmos padronizados é importante para a interoperabilidade do sistema.

Codificação na forma de onda

- ITU G-711 (PCM logarítmico lei A ou lei μ a 64 kbps): O sinal de voz na faixa de 4 kHz, ou seja, filtrado de 100 a 3400 Hz, é amostrado a 8 kHz com 14 (lei μ) ou 13 bits (lei A) e sofre uma compressão logarítmica para 8 bits. A compressão logarítmica é adequada por causa da natureza do sinal de voz: os sinais de baixa amplitude são mais frequentes e mais importantes.
- ITU G-726 (ADPCM a 40, 32, 24 ou 16 kbps): O sinal de voz é amostrado a 8 kHz, codificado em 8 bits (leis A ou μ) e são transmitidas diferenças entre amostras com 5,4, 3 ou 2 bits em quantização adaptativa. Outros métodos de compressão de áudio ADPCM incluem as Recomendações-padrão G-721 (taxa de 32 kbps) e a G-724 (taxa de 24 kbps); e o algoritmo de compressão de áudio interativo [15,21].
- ITU G-722 (SB-ADPCM a 64, 56 ou 48 kbps): O sinal de voz numa banda de 7 kHz é amostrado a 16 kHz e codificado em 14 bits, gerando assim uma taxa de 224 kbps. As amostras passam por um circuito codificador que as trata em duas sub-bandas e geram códigos numa taxa de 64, 56 ou 48 kbps. Nestes dois últimos casos, pode-se transmitir também dados a 8 ou 16 kbps para completar os 64 kbps. Este algoritmo é próprio para aplicações de videoconferência uma vez que telefones comuns não respondem na faixa de 7 kHz

Codificação na fonte

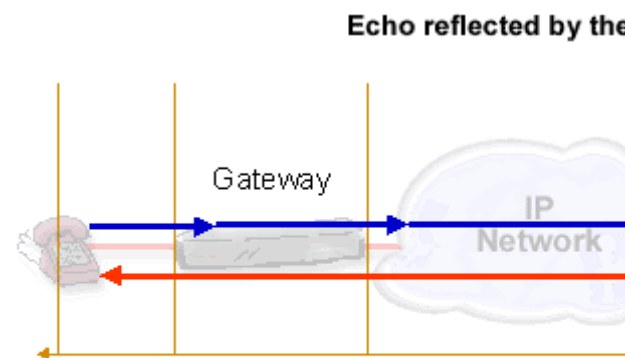
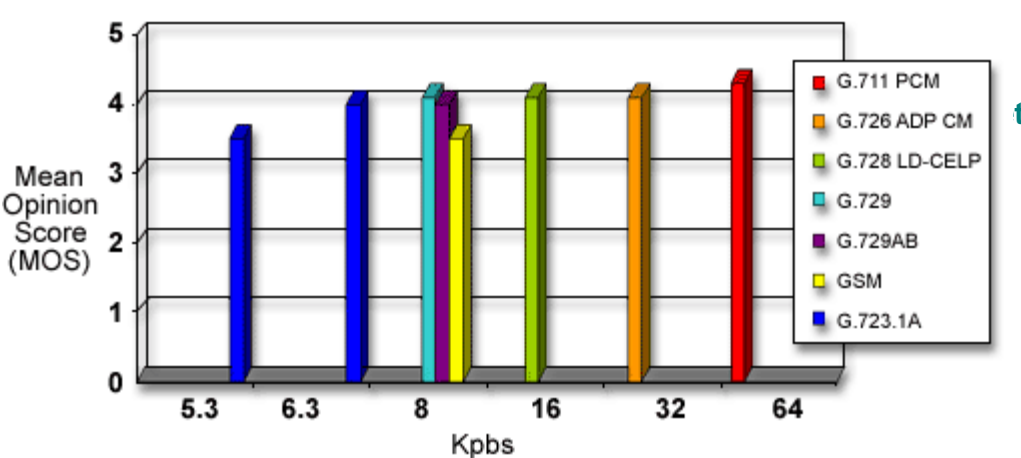
- ITU G-723 (ACELP a 5,3 ou 6,3 kbps): Em cada janela de 30 ms do sinal de voz, são analisadas 240 amostras de 16 bits do sinal de voz (tomadas a 8 kHz) para identificação de pitches e são gerados 12 ou 10 códigos de 16 bits, conforme o

algoritmo esteja configurado para uma taxa de 6,3 ou 5,3 kbps. É um VOCODER próprio para videoconferência, telefonia móvel digital e aplicações multimídia.

- ITU G-728 (LD-CELP a 16 kbps): Em cada janela de 0,625 ms do sinal de voz, são analisadas 5 amostras de 8 bits (que internamente voltam para PCM linear) e é gerado 1 código de 10 bits. O “codebook” é uma tabela de 1024 valores. Os pontos fortes deste VOCODER são o pequeno atraso e a qualidade do sinal de voz (muito parecida com a compressão não paramétrica). Por isso é adequado para a maioria das aplicações. Permite inclusive o reconhecimento de DTMF e Modem de até 2400 bps.
- ITU G-729 (CS-ACELP a 8 kbps): Em cada janela de 10 ms do sinal de voz, são analisadas 80 amostras de 8 bits para a geração de 10 códigos de 8 bits. Em [23] são apresentados os aspectos importantes deste algoritmo, que é referenciado como CS-ACELP. É um VOCODER bastante otimizado nos seus vários aspectos (qualidade, atraso, taxa de compressão) e assim tende a ser bastante difundido.
- TIA IS-54 (VSELP a 8 kbps): Em cada janela de 20 ms do sinal de voz, são analisadas 160 amostras de 8 bits para a geração de 10 códigos de 16 bits. Este VOCODER é patente da Motorola e é próprio para telefonia móvel digital.
- TIA IS-641 (ACELP a 7,4 kbps): Em cada janela de 20 ms do sinal de voz, são analisadas 160 amostras de 8 bits para a geração de um frame de 148 bits. Também é próprio para telefonia celular digital.

Remoção de Eco

Do usuário até a central telefônica a conexão é feita por dois fios, par de cobre trançado, com transmissão *full duplex*, ou seja, a voz compartilha o meio nos dois sentidos. A partir da central telefônica, ou do equipamento de acesso, são separados os canais de conversação de forma que a voz do assinante A vai trafegar por circuito diferente do assinante B. Chamamos de híbrido o equipamento que faz a conversão de dois fios para quatro fios, do modo *full duplex* para *simplex*. Remotamente, na outra ponta, um equipamento semelhante é usado para converter para dois fios que chegam ao usuário do outro lado. O equipamento híbrido remoto, que converte os quatro fios para dois fios de B gera um eco que será percebido pelo assinante A caso chegue com atraso superior à percepção humana, algo em torno de 10 ms. Na prática, o eco é um problema significativo quando este atraso é maior que 50 ms. A partir daí, faz-se necessário a presença de um mecanismo de cancelamento do eco.



Echo is detected in the near end gate from the gateway (10ms echo tail) and

Como o atraso da Telefonia IP com certeza é maior que este valor, haverá percepção de eco. Para evitar isto, o *gateway* ou o terminal de Telefonia IP deve realizar a função de remoção de eco.

Verificação da Qualidade - MOS

O método MOS (*Mean Opinion Score*) é derivado do método ACR (*Absolute Category Rating*) para estimar a qualidade de sistemas de transmissão de voz. O MOS requer que algumas pessoas avaliem a qualidade geral de exemplos de fala submetidos a *vocoders* para propósitos de comunicação telefônica.

Os avaliadores atribuem notas em escala de 5 categorias, como mostrado no quadro abaixo:

Nota	Significado
1	Ruim: ininteligível, o usuário não entende a mensagem decodificada. Possui interrupções horríveis devido às degradações
2	Pobre: o sinal possui interrupções devido às degradações e o usuário tem que fazer um esforço considerável para entender alguns trechos.
3	Moderado: a qualidade da voz é ruim, o usuário sente incomodado com as degradações, porém não tem interrupções e ainda consegue entender a mensagem. (requer esforço moderado)
4	Bom: a voz é agradável de se ouvir, ou seja, percebe degradações mas não se incomoda com elas, pois são mínimas. (nenhum esforço apreciável é requerido)
5	Excelente: o usuário não consegue diferenciar o trecho original com o corrompido, ou seja, não percebe degradações no sinal. (nenhum esforço é requerido)

A especificação completa e procedimentos para o ACR são descritos em [3]. O objetivo deste teste pessoal é apresentar uma figura da opinião pessoal dos jurados observando o sinal como transmitido pelo sistema de comunicação ou pelo algoritmo sendo testado. Como normalmente conduzidos, estudos pessoais de sistemas de comunicação requerem que os avaliadores escutem a fala processada por um período de tempo definido.

Comparação de Vocoders

O vocoder utilizado influi na qualidade do sistema de várias maneiras. Ele inclui por si uma distorção na amostra de voz. Esta distorção pode ser comparada entre vários *vocoders* utilizando a técnica MOS de avaliação de qualidade acústica da voz. Abaixo temos um gráfico de comparação

O segundo aspecto a ser observado é a sua largura de banda ocupada. Em uma rede saturada, com grande número de canais de voz pela rede, isto pode ser crítico. No gráfico acima vemos a informação de banda e de qualidade cruzadas para os principais *vocoders*.

Um outro aspecto é o atraso introduzido pela formação de um pacote. Isto depende do número de amostras que são tomadas pelo vocoder e da taxa de compressão. O vocoder G.723, cujas taxas são mais baixas tem tempo de formação de pacote de 30 ms, o que pode introduzir um atraso significativo. G.729 e G.728 introduzem atrasos menores.

Um último aspecto a ser analisado é a complexidade do algoritmo de compressão, para que possam ser dimensionados processadores que suportem o algoritmo sem imposição de atrasos adicionais.

Comutação

Comutação por Circuitos

Tradicionalmente, as redes de telefonia, tanto a nível público ou privado tem se desenvolvido utilizando-se de comutação baseada em circuitos. Estas redes são conhecidas por *Time Division Multiplexing* (TDM), devido ao método de multiplexação baseado na divisão do tempo. Através deste paradigma, um assinante A, presente em um ponto da rede de telefonia, ao se comunicar com um outro assinante B, é estabelecido um circuito exclusivo para garantir o canal de comunicação. Este circuito, de largura de banda determinada, normalmente 64kbps, é usado para o transporte de voz entre o assinante A e o assinante B, e é multiplexado para a transmissão em um meio físico compartilhado. Caso não haja comunicação entre os dois assinantes, como em momentos de silêncio, o canal é desperdiçado. Dizemos que a comutação é baseada em circuitos porque os elementos do sistema que garantem que a voz chegue ao destino são comutadores que estabelecem circuitos na fase de negociação da chamada e mantém estes circuitos até o término da chamada.

Vários canais podem ser multiplexados em um feixe de canais do tipo E1 (32 canais) ou T1 (24 canais) para ligação entre centrais telefônicas. Deve ser observado que um canal de comunicação é mantido durante todo o período de conversação entre dois assinantes. Embora possa ser multiplexado para economia de meio físico de transmissão, o canal deve existir e tem largura fixa de 64 kbps.

Comutação por Pacotes

Em oposição, podemos transmitir a voz convertendo-a para pacotes que são elementos mínimos em uma rede comutada por pacotes. Nesta abordagem, o conceito de circuito não existe, de forma que todos os pacotes compartilham o mesmo meio físico e lógico. Os comutadores deste sistema analisam cada pacote que chega e verifica qual seu destino, antes de enviá-lo em sua rota. Desta forma, são mantidas rotas que são respeitadas e verificadas a cada pacote.

A verificação de destino e conseqüente roteamento substitui a comutação da telefonia convencional, com algumas diferenças:

- É baseada em pacotes, de modo que cada pacote registra informações de origem e destino, e compartilha o mesmo circuito físico, estas informações adicionais inserem perdas de eficiência;
- É realizada na camada de rede, IP, baseada no endereçamento deste protocolo, no formato de 4 octetos;
- É realizada por software, ao contrário da comutação na telefonia fixa, de forma que se torna muito mais flexível, permitindo verificações de erros, priorização de pacotes, bloqueio de pacotes indesejáveis ou inseguros, etc. A desvantagem da

execução de software é que o desempenho do sistema se torna mais crítico, de forma que atrasos maiores podem ser inseridos neste processo.

Uma pacote representa um conjunto de bytes sequenciados e agrupados, marcados por cabeçalhos que determinam suas características, como protocolo utilizado, tamanho do pacote, sequenciamento, endereço de origem e endereço de destino.



Com esta abordagem, não existe desperdício de largura de banda, uma vez que somente a porção da comunicação que efetivamente é utilizada gera tráfego sobre a rede. Assim, podemos aplicar algumas técnicas como supressão de silêncio, que representa a retirada de todos os pacotes que não tenham conteúdo da rede, além da compressão dos pacotes de voz, de forma a reduzir a banda utilizada por uma comunicação entre dois assinantes. Algumas tecnologias tem se desenvolvido para a transmissão de voz sobre redes que utilizam a comutação baseada em pacotes. Os principais protocolos de rede utilizados para estes fim são: ATM, *Frame Relay*, e IP (*Internet Protocol*). A telefonia sobre o protocolo IP é o objeto de nosso estudo e tem como principais vantagens:

- Capilaridade de rede já instalada no mundo, a Internet, cujo alcance é indiscutível;
- Baixo custo, devido à simplicidade do protocolo utilizado. Em contrapartida, o protocolo é pobre em termos de recursos, como garantia de qualidade de serviço e disponibilidade de largura de banda;
- Grande número de aplicações já desenvolvidas utilizando este protocolo. Todas estas aplicações podem ser integradas no ambiente de telefonia.

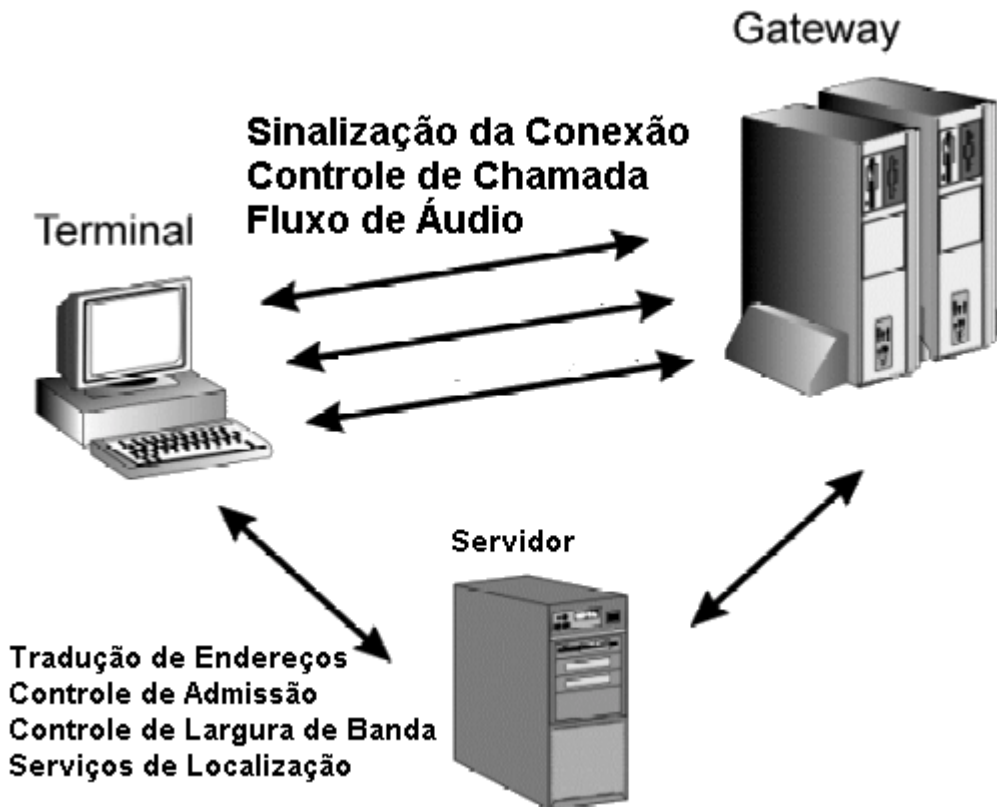
Outros modelos de rede, como ATM apresentam características que poderiam tornar interessante seu uso para transmissão de voz, como garantia de qualidade de serviço. Mas este tipo de rede não se difundiu como as rede com o protocolo IP, de forma que não são acessíveis a baixo custo e com tamanho número de aplicações.

Uma rede ATM não utiliza o mesmo princípio de comutação de uma rede IP. A rede ATM tem comutação baseada em células, que são pequenos pacotes (53 bytes), comutados por hardware, devidos à simplicidade dos circuitos de comutação. Possui o conceito de circuitos virtuais, com reserva de banda, que garante a qualidade de serviço para transporte de voz, porém de forma cara, o que inviabiliza sua utilização.

Padrões de controle para Telefonia IP

Para realizar uma chamada são necessários protocolos de controle e sinalização para execução de algumas tarefas como localização do usuário a ser chamado, notificação de chamada, notificação de aceite da chamada, início da transmissão de voz, finalização da transmissão de voz e desconexão. Isto sem considerar os serviços adicionais no ambiente de telefonia. Inicialmente estas tarefas eram executadas por protocolos proprietários, hoje, dois conjuntos de protocolos padrões dominam o cenário de telefonia.

A maioria dos protocolos de sinalização estão focadas na conexão e no controle da chamada, que estão separados da transmissão de conteúdo multimídia (voz) entre os terminais. Isto porque a conexão e o controle de chamada podem envolver terceiros elementos, como um servidor de localização. Desta forma, permite que os terminais transmitam a voz sem a intervenção deste terceiro elemento.



Podemos atribuir ao servidor as seguintes tarefas:

- Tradução de endereços - Usuários do serviço telefônico são identificados por um número e usuário da Internet ou outra rede IP podem ser identificados pelo endereço IP da máquina onde o usuário está localizado ou pelo endereço eletrônico do usuário. Em ambas as abordagens, o número telefônico deve ser traduzido, permitindo assim que o usuário receba corretamente suas chamadas;
- Controle de Admissão - Autorização ou não de utilização do sistema, tarifação, etc.

- Controle de largura de banda - Capaz de reduzir ou aumentar a carga na rede através da autorização de chamadas, transferências de arquivos, recursos de vídeo, etc.
- Serviços de Localização - Um determinado usuário, registrado através de um email, pode manter a informação de sua localização no Servidor, de forma a permitir que suas chamadas sejam corretamente direcionadas para ele.

Todos estes serviços envolvem comunicação entre o terminal e o servidor ou entre um gateway e o servidor. O tráfego de voz, por sua vez, não passa pelo servidor, passando diretamente entre os terminais e/ou gateways. Além disto, também os serviços de controle de chamada e sinalização da conexão são feitos diretamente entre os elementos finais.

Várias soluções proprietárias foram utilizadas para transmissão de voz em redes IP. Porém estas soluções não garantiam interoperabilidade entre os equipamentos. Podemos verificar então, o advento de dois padrões que podem ser utilizados para garantir a interoperabilidade entre sistemas de Telefonia IP. O padrão [H.323](#), proposto pela ITU-T[3] e o padrão [SIP](#) proposto pelo IETF [4].

Independente do padrão utilizado, o protocolo normalmente utilizado para transmissão do fluxo de voz é o *Real Time Protocol* (RTP)[5]. RTP por sua vez utiliza de *User Datagram Protocol* (UDP) para transmitir estes pacotes. RTP é feito sobre UDP porque a velocidade é necessária e fundamental no processo e a perda de pacotes pode ser compensada no processo de codificação e decodificação da voz.

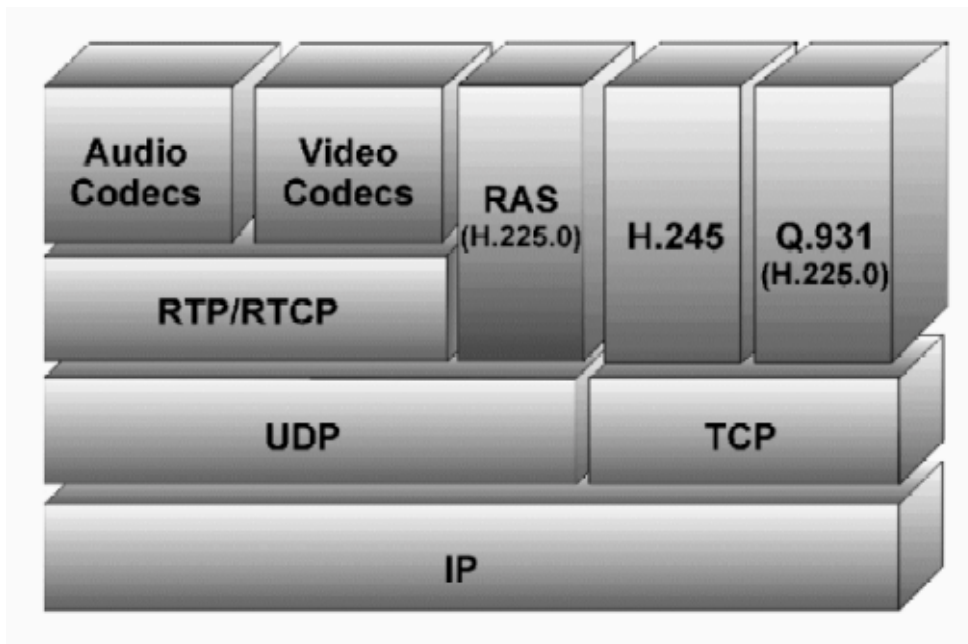
Para garantir a comunicação entre terminais na Internet e telefones analógicos presentes na rede telefônica convencional, os protocolos devem ser capazes de interagir com os equipamentos nesta rede. Esta interconexão é efetuada por um *gateway*. Um *gateway* é responsável por converter comandos do protocolo de controle baseados em IP para protocolos tradicionais de sinalização como SS7 e ISDN, além de converter o fluxo de voz para sinais de voz compatíveis com a rede telefônica convencional.

H.323

A ITU-T propôs o padrão H.323[3], sendo mais difundido atualmente, especialmente por ser o precursor da Telefonia IP, primeiro padrão a ocupar esta lacuna. Este padrão prevê, entre outros:

- Algoritmos padrões de compressão que devem ser implementados de forma a garantir compatibilidade, conhecidos como áudio *codecs* ou *vocoders*;
- Protocolos utilizados para o controle da chamada, estabelecimento dos canais de comunicação e negociação de qualidade de serviço;
- Interoperabilidade com outros terminais de voz, como telefonia convencional, ISDN, voz sobre ATM, e outros; permitindo assim a construção de gateways.
- Elementos ativos do sistema e suas funções.

O protocolo H.323 utiliza em suas diversas funcionalidades de uma família de Recomendações ITU-T: H.225[4] para conexão, H.245[5] para controle, H.332 para grandes conferências, H.335 para segurança, H.246 para interoperabilidade com RTPC, e a série H.450.x para serviços suplementares. Todos estes padrões fazem parte da série H de Recomendações.



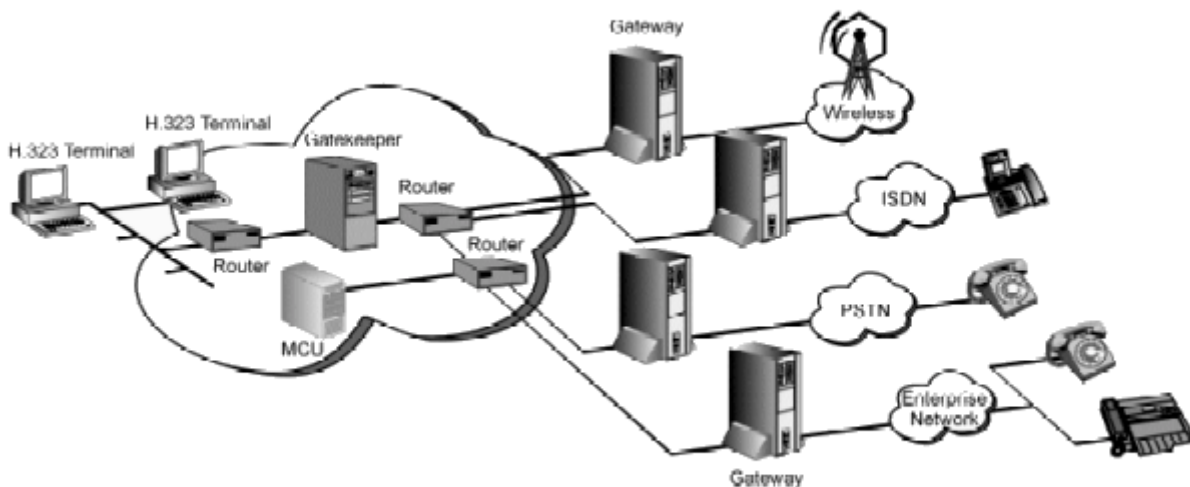
Podemos verificar a pilha de protocolos definidos na Recomendação H.323 na figura acima. Podemos verificar:

- *Audio Codecs* e *Video Codec*, utilizando do protocolo RTP para transmissão do fluxo de pacotes;

- RAS, *Registration, Admission and Status*, mensagens trocadas entre os terminais e unidades de gerência de conexão, executando tarefas como localização de usuários e reserva de largura de banda.
- RTP/RTCP, *Real Time protocol / Real Time Control Protocol*, usado para transporte do fluxo de pacotes multimídia, com características de tempo real, executando também funções de estatísticas de qualidade de serviço;
- H.245, Controle de canais de áudio e vídeo;
- Q.931, Mensagens trocadas entre os terminais para sinalização de chamada.

Devemos observar ainda que os protocolos RTP, RTCP e RAS utilizam na camada de transporte de protocolo não orientado a conexão (UDP), ao contrário da dupla H.245/Q.931, que utiliza de protocolo orientado a conexão (TCP).

Podemos verificar também na Recomendação H.323, os elementos que compõe uma rede de Telefonia IP. Estes elementos aparecem na figura abaixo e podem ser definidos como se segue:



Arquitetura de Rede e Componentes H.323

Os seguintes elementos são definidos neste ambiente:

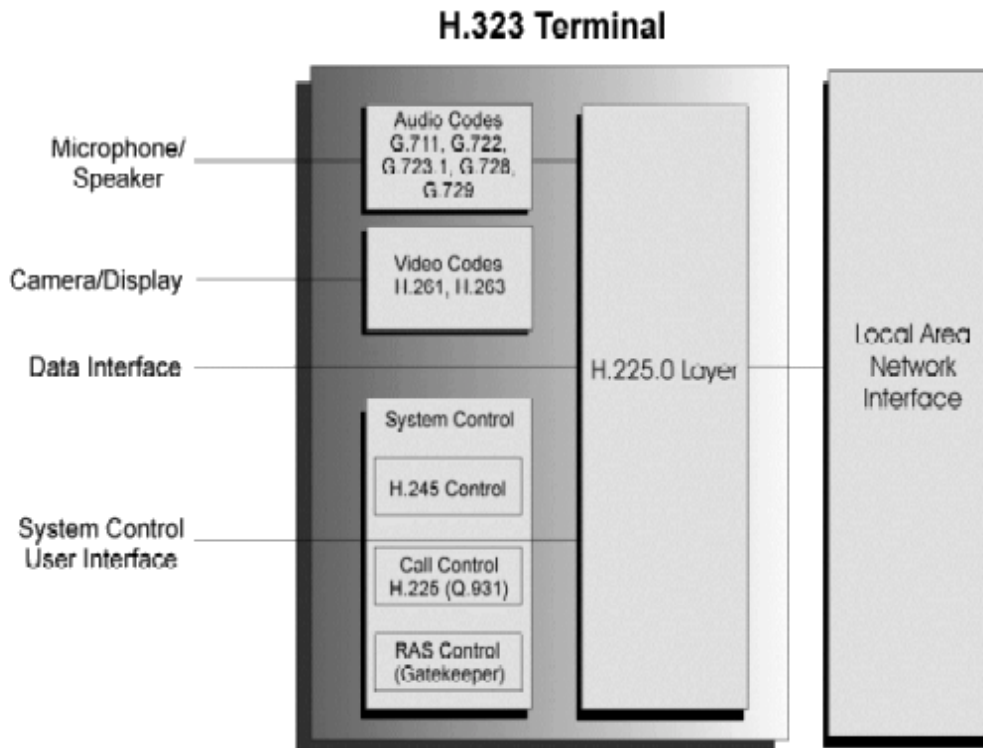
- Terminal H.323 – Computador onde está implementado o serviço de telefonia IP, atuando como terminal de voz, vídeo e dados, através de recursos multimídia;
- Gateway H.323 – Elemento situado entre uma rede IP e outra de telecomunicações, como o sistema telefônico convencional (PSTN), rede integrada de serviços digitais (ISDN), rede de telefonia celular; de forma a permitir a interoperabilidade entre as duas redes;

- *Gatekeeper* – Servidor de tradução, localização e admissão. Provê tradução de endereços e controle de acesso à rede H.323 para terminais, *gateways* e *MCUs*. Pode prover também outros serviços como gerenciamento de largura de banda e localização de *gateways*.
- Unidade de controle de Multiponto (MCU) – Provê facilidades para três ou mais terminais e *Gateways* participarem em uma conferência multiponto.

Não se faz necessária a implementação de todos estes elementos para efetuar uma comunicação em ambiente de Telefonia IP. Podemos, utilizando de apenas dois terminais H.323, estabelecer uma comunicação de voz, vídeo e dados. Os outros elementos permitem gerenciamento da chamada e interconexão com outras redes.

Um dos protocolos mais importantes do padrão H.323 é o H.225, responsável pelo controle da chamada. Além disto, descreve também as funcionalidades do RAS, *Registration, Admission and Status*, o qual faz uma aplicação cliente acessar recursos, como roteamento, serviços de diretórios, localização de usuários, providos por um *Gatekeeper*.

Podemos verificar na figura abaixo, as camadas utilizadas em um terminal H.323. Este terminal é normalmente implementado por software sobre computadores pessoais multimídia.



Vários softwares comerciais e freeware implementam terminais do padrão H.323 e garantem interoperabilidade. Os principais são:

-  [Microsoft NetMeeting](#)
-  Intel Internet VideoPhone,
-  [VocalTec Internet Phone](#)
-  [Voxware VoxPhone](#)
-  [Netscape Conference](#)
-  [NetSpeak WebPhone](#)

Também está disponível a pilha de protocolos que implementa o padrão H.323 sobre a forma de código-fonte C e C++, compilável na plataforma Linux e também na plataforma Windows. Esta pilha de protocolos foi desenvolvida por um consórcio de desenvolvedores e pode ser encontrada no [Projeto OpenH323](#) [8]. Como alternativa, outra pilha de protocolos também disponível em código fonte para o padrão H.323 é distribuída livremente pela [Vovida Networks](#)[9].

A complexidade do protocolo H.323 é muito grande. Sua composição incluindo todas estas subrecomendações resulta em mais de 700 páginas de padrões. Além disto, ao contrário da maioria dos protocolos utilizados na Internet que são baseados em texto, o H.323 é binário, baseado na *Abstract Syntax Notation On* (ASN.1) e *Packet Encoding Rules* (PER). Isto torna o H.323 computacionalmente exaustivo.

A versão 1 do padrão H.323 foi criticada por ter o desempenho pobre e muitos problemas de incompatibilidade entre diferentes fabricantes. Por não ter muitos recursos, diversos fabricantes acabaram por incluir suas próprias extensões para o padrão, de forma proprietária. Isto resultou em incompatibilidade em diversos pontos. ITU-T resolveu estes problemas com a versão 2 do padrão, o qual provê mecanismos de conexão rápida que eliminam problemas de desempenho e incluiu muitos novos recursos, de forma a eliminar a necessidade de extensões proprietárias.

Protocolos IETF para a Telefonia IP

SIP - Session Initiation Protocol

A alternativa do protocolo H.323 é SIP, o qual foi concebido na Universidade de Columbia e depois submetido para aprovação do IETF (*Internet Engineering Task Force*). Como outros protocolos dominantes na Internet (HTTP, FTP, SMTP), SIP é baseado em texto. SIP foi aprovado como RFC (*Request For Comment*), sendo publicado como RFC 2543, em março de 1999.

SIP preocupa-se com a sinalização e não tenta definir qualquer aspecto de comunicação multimídia, como H.323 faz. Consequentemente, ele pode ser documentado em menos que 130 páginas de leitura relativamente fácil. SIP reutiliza algumas características de outros protocolos, como por exemplo, os mesmos cabeçalhos, erros e regras de codificação do HTTP.

O protocolo de sinalização SIP inicia uma sessão entre usuários. É usado para serviços de localização de usuários, controle de chamada e gerência de participantes em uma conferência (usando uma extensão SIP[10]). SIP não define o tipo de sessão que é estabelecida, de forma que pode ser facilmente usado tanto para estabelecer sessões de jogos como para estabelecer conferências de vídeo/áudio[9].

Cada requisição SIP consiste de um conjunto de campos de cabeçalho que descrevem a chamada seguidos por uma mensagem que descreve uma sessão individual que está sendo realizada na chamada. Normalmente SDP (que iremos descrever) é usado, mas pode ser negociado outro protocolo. Além disto, como SIP não define todo o potencial do protocolo usado em comunicações multimídia, ele deve ser combinado com outro protocolo IETF para criar uma solução completa. Por exemplo, você pode usar as características de tocar e gravar do *Real Time Streaming Protocol* (RTSP) para prover funcionalidades de *voicemail* ou secretária eletrônica com SIP.

SIP é um protocolo cliente-servidor, parecido tanto em sintaxe como semântica ao protocolo HTTP[11]. Requisições são geradas por uma entidade cliente e enviadas para uma entidade receptora ou servidor. O servidor processa a mensagem e então envia uma resposta para o cliente. A requisição e a resposta que chegam a seguir são chamadas transações. Todo software SIP tem um sistema final que interage com um usuário conhecido como *User Agent*. Um *user agent* tem dois componentes, um ***user agent client(UAC)*** e um ***user agent server(UAS)***. O UAC é responsável por iniciar as chamadas, enviando requisições, e o UAS é responsável por responder às chamadas, enviando respostas. Uma aplicação de telefonia Internet contém ambos UAC e UAS. Vale destacar a diferença com um *Web Browser* que contém apenas um cliente.

Com funções específicas, espalhados pela rede, existem três tipos de servidores. **Servidores de registros**, os quais recebem atualizações sobre a localização corrente de cada usuário; **servidores proxy**, os quais recebem requisições e enviam-nas para outros servidores, conhecidos então como servidores ***next-hop***. O servidor ***next-hop*** pode ser outro servidor ***proxy***, um UAS ou um servidor de redirecionamento. Um **servidor de redirecionamento** também recebe requisições e determina um servidor ***next-hop***. Todavia,

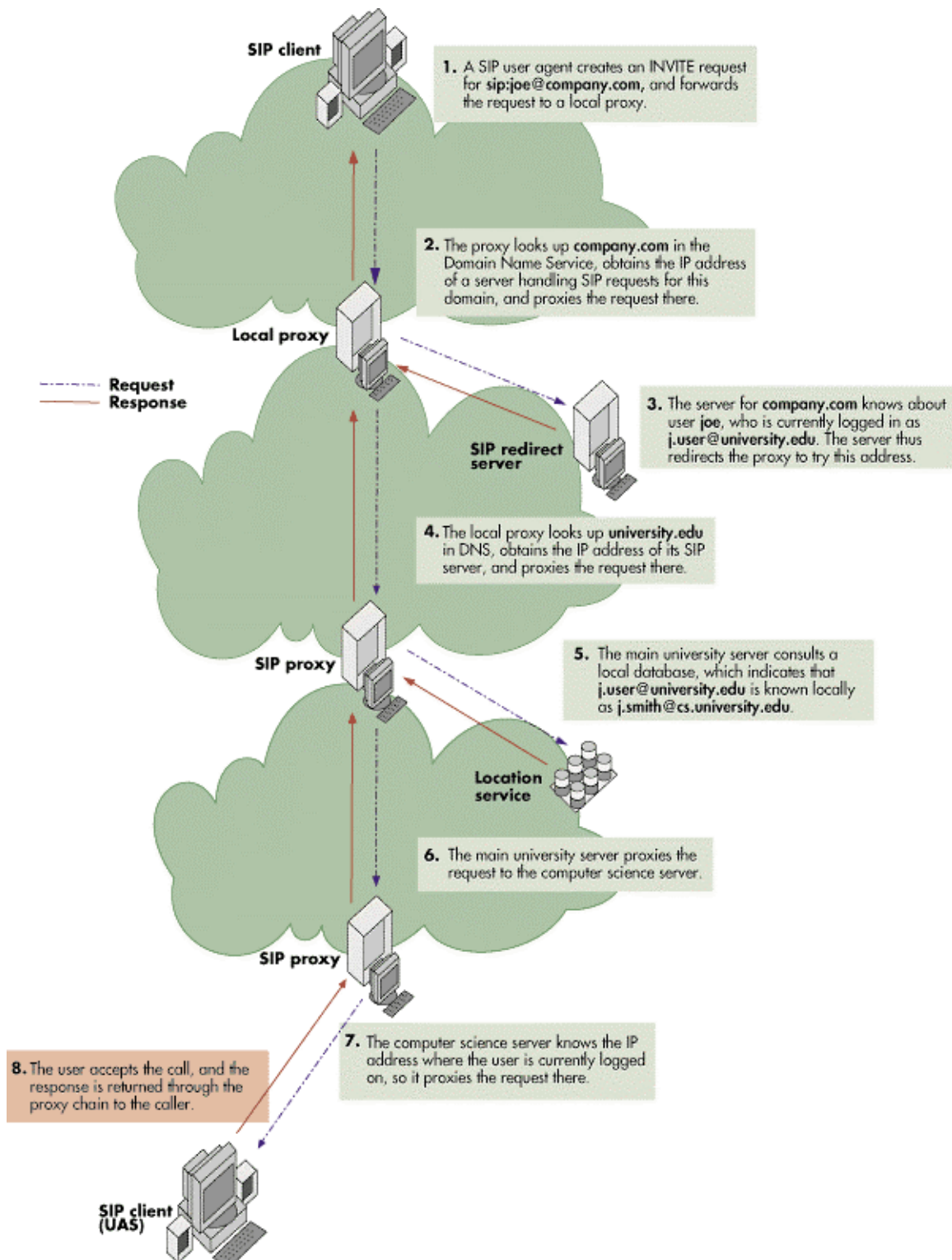
ao contrário de reenviar a requisição neste caso, ele retorna o endereço do servidor *next-hop* para o cliente. A função primária dos servidores *proxy* e de redirecionamento são roteamento de chamadas, ou seja a determinação do conjunto de servidores a serem usados no caminho para completar a chamada. Um servidor *proxy* ou um servidor de redirecionamento pode usar qualquer meio para determinar o servidor *next-hop*, incluindo executar programas e consultar bancos de dados. Um servidor *proxy* pode também duplicar uma requisição, enviando cópias para múltiplos servidores *next-hop* de uma vez. Isto permite que uma requisição de início de chamada tente diversas localizações diferentes ao mesmo tempo. A primeira localização que responder é conectada com o cliente chamador.

Como no HTTP, o cliente requisita a chamada de métodos ou comandos no servidor. SIP define muitos métodos. INVITE convida um usuário para uma chamada. BYE termina uma conexão entre dois usuários em uma chamada. OPTIONS solicita informações sobre as capacidades, mas não configura uma chamada. ACK é usado para confirmar mensagens trocadas para estabelecimento. CANCEL termina a procura por um usuário. Finalmente REGISTER adiciona informações sobre um usuário para um servidor de registro.

Um cliente inicia uma chamada enviando uma requisição INVITE. Esta requisição contém campos de cabeçalho usados para levar informações sobre a chamada. Os campos de cabeçalho mais importante são TO e FROM, os quais contém o endereço do chamador e do chamado. O campo SUBJECT identifica o assunto da chamada. O campo CALL-ID contém um identificador único da chamada e o campo CSEQ contém um número de sequência. O campo CONTACT lista endereços onde o usuário pode ser contactado. Este último é colocado em respostas de um servidor de redirecionamento, por exemplo. O campo REQUIRE é usado para a negociação do protocolo a ser utilizado como extensão ao SIP. Os campos CONTENT_LENGTH e CONTENT-TYPE são usados para levar informações sobre o conteúdo da mensagem. No conteúdo da mensagem vemos ainda a descrição da sessão que será estabelecida.

Extensões podem ser definidas com novos campos. Uma extensão[10], usada para controle de chamada, incluem novos cabeçalhos para tarefas como transferências de chamadas e conferências com várias pessoas[12, 13].

Vemos abaixo uma típica transação SIP:



SDP

SDP descreve sessões multimídia, para telefonia e aplicações de difusão como rádio ou TV via Internet. O protocolo inclui informações sobre:

- Fluxo de mídia: Uma sessão multimídia pode conter muitos fluxos de mídia, como fluxos de áudio, vídeo, etc. SDP leva informações sobre o número e o tipo de cada fluxo de mídia. São definidos áudio, vídeo, dados, controle e aplicações como tipos de fluxos, similares aos tipos MIME usados por email Internet.
- Endereços: Para cada fluxo, o endereço do destinatário (*unicast* ou *multicas*) é indicado, de forma a garantir independência. Assim, um usuário pode receber a voz através de um telefone de baixo atraso e o vídeo através de uma workstation, em uma videoconferência.
- Portas: Para cada fluxo, a porta UDP para recepção e/ou envio é indicada.
- Tipo de conteúdo: Os formatos de mídia que podem ser usados durante a sessão. Para sessões *unicast* (tradicional telefonia IP) este campo corresponde ao campo *capability set* do protocolo H.323.
- Horário de início e término: Para sessões do tipo *broadcast* como um programa de televisão, o horário de início, o horário de término e a duração são enviados. Assim, pode-se anunciar ou convidar usuários para um programa de TV semanal, por exemplo.
- Origem: Para sessões do tipo *broadcast* a sessão traz quem é o responsável e como pode ser contactado.

SDP leva esta informação em formato textual, assim como SIP. Quando uma chamada SIP é estabelecida, a mensagem INVITE contém informações SDP descrevendo os parâmetros aceitáveis para o chamador, incluindo suas capacidades. A resposta do chamado contém uma versão modificada desta descrição, incorporando suas capacidades.

Temos abaixo um exemplo de campos SDP:

```
v                               =                               0
o = sergiool 87728 8772 IN IP4 15.164.10.1
s                               = Ola, senhores!
u                               = http://www.dcc.ufmg.br/~sergiool
e                               = sergiool@dcc.ufmg.br
c                               = IN IP4 150.164.10.1
b                               = CT:64
t                               = 3086272736 0
k                               = clear:manhole cover
m                               = audio 3456 RTP/AVP 96
a                               = rtpmap:96 VDVI/8000/1
m                               = video 3458 RTP/AVP 31
m                               = application 32416 udp wb
a = orient:portrait
```

Podemos identificar "v" como indicador de versão da sessão. Linha "o" apresenta o conjunto de valores que formam um identificador único para a sessão, incluindo endereço IP, portas utilizadas e usuário. Linhas "u" e "e" indicam a URL e o endereço de e-mail para mais informações sobre a sessão. A linha "c" indica o endereço para a sessão, a linha "b" indica a largura de banda (64kps), e a linha "t" o tempo de início e o final, onde 0 indica que a sessão continua indefinidamente. A linha "k" traz a chave de criptografia para a sessão. Existem 3 linhas "m", cada qual identifica um tipo de fluxo de mídia (áudio, vídeo e aplicação de quadro branco para desenhos, etc), o número da porta para o fluxo, o protocolo e a lista de tipos de conteúdo. A linha "a" representa um conteúdo. Por exemplo, a linha seguindo a definição do fluxo de áudio define o parâmetro de *codec* para RTP tipo 96.

Real Time Protocol

RTP (*Real time transport protocol*) é um protocolo padrão para transporte de dados com características de tempo real, como vídeo e áudio, que pode ser usado em serviços tanto como mídia sob demanda quanto interativos, do tipo telefonia IP. É composto por uma parte de transmissão de dados e outra de controle, chamada RTCP (*Real Time Control Protocol*). A parte de dados consiste de um protocolo leve, que provê suporte para aplicações com características de tempo real, incluindo reconstrução temporal de mensagens, detecção de perdas, segurança, selo de tempo e identificação de conteúdo[1]. RTCP é usado para monitorar a qualidade de serviço da transmissão e para transmitir informações sobre os participantes de dada sessão[2]. Apesar do RTP utilizar serviços do UDP e IP, ele pode ser implementado em outros ambientes, já que necessita somente de serviço de transporte não orientado à conexão.

O processo de transporte envolve acompanhar o fluxo de bits gerados pelo codificador de mídia, normalmente o *vocoder* de Telefonia IP, quebrando em pacotes, enviando os pacotes pela rede e reproduzindo o fluxo de bits no receptor. O processo é complexo porque pacotes podem ser perdidos, receber atrasos variados e serem entregues fora de ordem. O protocolo de transporte deve permitir ao receptor detectar estas perdas. Ele deve também transportar informações de temporização de forma que o receptor possa fazer também compensações para o atraso.

Para ajudar nestas funções, RTP define o formato de pacote que trafega pela rede. O pacote contém informação sobre a mídia, chamada *RTP Payload*, junto ao cabeçalho RTP. Este cabeçalho provê informações para o receptor que permitirão a reconstrução da mídia. RTP também especifica como o fluxo de bits do *vocoder* estão agrupados nos pacotes. RTP contém também dispositivos para *multicast*, permitindo seu uso em aplicações de conferência e ambientes *broadcast* usados para distribuição de mídia. É importante notar que RTP não tenta reservar recursos na rede para evitar perdas de pacotes e atrasos, ele apenas permite ao receptor notar a presença desta perda e do atraso.

RTP é componente chave dentro de sistemas de Telefonia IP. Podemos verificar que além do SIP, o próprio padrão H.323 faz uso do RTP, de forma que podemos considerá-lo universal para uso de Telefonia IP.

Podemos verificar algumas funcionalidades do RTP:

- Sequenciamento: Atribui número de ordem aos pacotes. Isto pode ser usado para verificação de perdas e possível reordenamento de pacotes;
- Sincronismo intramídia: Pacotes de um mesmo fluxo podem sofrer diferentes atrasos. A variação deste atraso é prejudicial à reprodução da mídia. Buffers adicionais podem então ser usados para eliminar a diferença entre os atrasos (*jitter*). Estes mecanismos precisam de informações de tempo de cada pacote. RTP provê esta informação[16-18].
- Identificação de conteúdo: Em redes sem garantia de QoS, como a Internet e redes de computação móvel, condições de perda de pacotes e atraso podem variar ao longo do tempo. Podemos observar que diferentes *vocoders* apresentam diferentes características de qualidade final em função das perdas e atrasos. Este campo permite a alteração dinâmica do *vocoder* utilizado em função das perdas e do atraso a fim de melhorar a qualidade final acústica.
- Identificação de quadro: Os quadros são unidades básicas de transmissão dos *vocoders*. Um pacote pode conter vários quadros. RTP marca o início e o fim de cada pacote, necessário para identificação deste quadro.
- Identificação de origem: Necessária para indicar quem enviou o pacote. Isto porque em uma conferência *multicast* um mesmo fluxo pode ter várias origens.

Já o RTCP, que acompanha o RTP, que provê informações adicionais sobre seus participantes:

- Retorno de informações de Qualidade de Serviço: Receptores podem retornar informações sobre atraso, *jitter* e perdas, que pode ser usadas para adaptar a aplicação [19, 20, 21], como por exemplo alterar o *vocoder* que está sendo utilizado.
- Sincronismo Intermídia - Necessário para sincronizar diferentes fluxos, como áudio e vídeo, caso sua origem seja de servidores diferentes.
- Identificação do usuário: Email, Nome, Organização, número telefônico, etc.

RTCP necessita que todos os participantes enviem estas informações periodicamente. RTCP usa o mesmo endereço do RTP, porém usa uma porta diferente. Nem todas as aplicações RTP utilizam RTCP, que não é indispensável para as aplicações.

Real-Time Streaming Protocol

RTSP[23] é usado para controlar servidores de armazenamento de mídia. Estes dispositivos são capazes de reproduzir conteúdos de mídia armazenados, como filmes ou músicas, atuando como servidores sob demanda. Os comandos do RTSP são semelhantes a de um vídeo cassete, como *play*, *fast forward*, *rewind*, *record* e *pause*. Também existem comandos para configurar o servidor quanto ao endereço e porta utilizada, *vocoder* para voz e vídeo.

RTSP pode apresentar um grande número de aplicações em Telefonia IP:

- Parte de um conversa  o pode ser gravada;

- Pode ser reproduzido algum trecho de áudio ou vídeo durante uma sessão de Telefonia IP;
- Utilização de e-mail de voz;
- Serviços de secretária eletrônica para Telefonia IP.

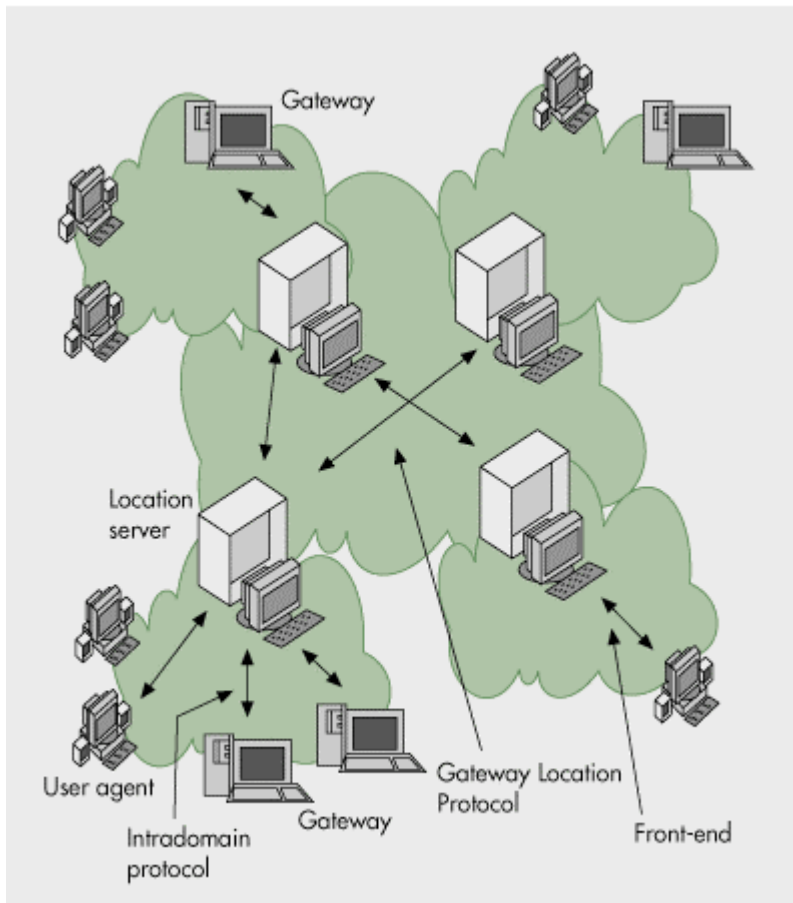
Call Processing Language

CPL[24, 25] é uma linguagem de scripts que permite que os usuários finais especifiquem o comportamento dos agentes de chamada, que executam em seu software de telefonia IP. O agente de chamada é chamado quando o servidor SIP recebe um pedido de chamada. Este agente executa as instruções geradas nesta linguagem, como por exemplo sempre tentar contactar determinada pessoa no seu celular, depois na sua casa e em seguida no escritório pelo menos uma vez.

Gateway Location Protocol

Para completar uma chamada para um telefone na rede pública tradicional, o cliente SIP deve enviar uma mensagem de INVITE para um gateway, para fazer a interconexão destas duas redes. Espera-se um grande número de gateways futuro, nos principais centros urbanos. Deve existir uma forma de encontrar o número IP do Gateway com base no número telefônico requisitado de forma que se otimize o custo da ligação ou permita-se a utilização de uma sub-rede mais confiável. O protocolo GLP é usado com esta finalidade.

A figura a seguir ilustra o protocolo GLP:



Um ambiente de Telefonia IP é dividido em domínios. Cada domínio mantém um número de gateways, usuários IP e servidores de localização. Os servidores de localização de um domínio conhecem sobre os gateways em seu domínio, por meio de um protocolo intradomínio como o *Service Location Protocol* (SLP). Este protocolo propaga informações sobre o gateway.

Não é possível que um simples domínio administrativo tenha gateways que acessem todos os número telefônicos existentes. Por isto é necessário que haja um acordos comerciais entre os vários domínios. Uma vez que este acordo é feito, é necessário que os Servidores de Localização de um domínio tenham como comunicar com outro. Isto é feito através do protocolo GLP. GLP permite que um Servidor de Localização atualize os registros de seu banco de dados com os gateways de outro domínio. Ele mantém informações do endereço IP do gateway, a faixa de número telefônicos que ele atende além de informações do próprio gateway. O princípio de funcionamento é semelhante ao serviço de nomes da Internet (DNS).

Quando um cliente deseja estabelecer uma ligação de telefonia IP com alguém na rede telefônica, isto pode ser feito das seguintes maneiras:

- LDAP - *Lightweight Directory Access Protocol* - Protocolo que permite acesso direto à base de dados de um servidor de localização.
- SIP - Ao invés de mandar um INVITE diretamente para o gateway, o cliente envia para o servidor de localização. Esta mensagem já contém o número telefônico a ser discado e o servidor de localização após consultar seu banco de dados, reenvia para o destino, atuando como servidor de redirecionamento.
- Páginas Web - O servidor de localização pode deixar disponível na rede IP, através do protocolo HTTP, todos os números telefônicos e a forma de acessá-los de forma que o usuário possa consultá-los e já disparar a requisição diretamente para o gateway.

Aspectos de tráfego de voz em tempo real

Atraso na Telefonia IP

Em um sistema de telefonia IP, vários fatores afetam a qualidade acústica final ao usuário. Entre eles já discutimos a escolha do algoritmo de codificação de voz e supressão de silêncio. A qualidade do meio de transmissão é também fundamental, pois uma grande quantidade de erros podem ser inseridos no meio, além de perdas ocasionais de pacotes.

Além disto, outro grande fator característico da qualidade é o atraso da voz, desde o momento em que é pronunciada até o momento em que é reproduzida. Muitos fatores atuam no sentido de aumentar este atraso, os quais vamos então discutir.

O atraso tem influência direta na qualidade de serviço da telefonia sobre IP. O usuário do sistema percebe um intervalo entre suas interlocuções igual a duas vezes ao atraso. Isto se dá porque o seu interlocutor só perceberá o fim da sua fala depois de um tempo igual ao atraso, e só depois começará a falar. Como sua fala também terá um atraso igual ao anterior, o usuário terá a impressão de atraso em dobro. A este tempo dá-se o nome de *round-trip*, que corresponde a duas vezes o atraso.

A ITU-T define da seguinte forma os valores em milissegundos para o atraso:



A telefonia convencional tipicamente não inclui atrasos maiores do que 150 milissegundos, o que é imperceptível para o homem. Ligações internacionais, principalmente quando envolvem satélites, podem ter atrasos de até 1 segundo, o que é considerado inaceitável para manter uma boa interlocução.

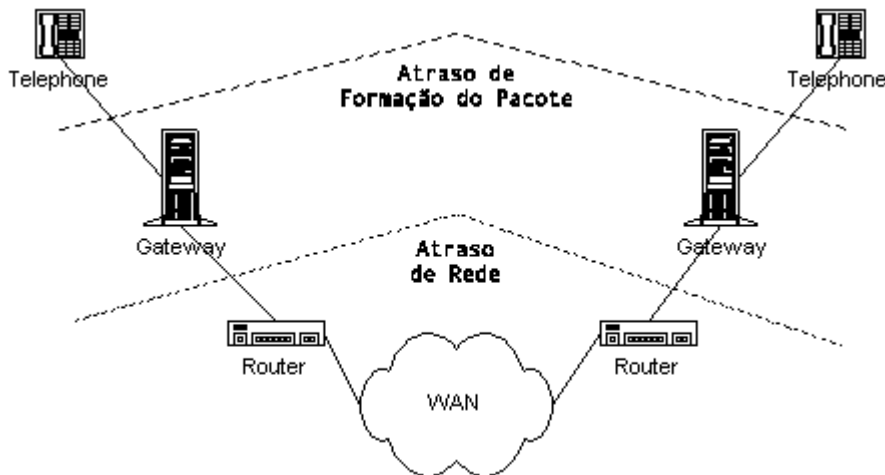
Entendendo o atraso

O atraso percebido num ambiente de telefonia IP, é gerado por uma série de pequenos atrasos:

- Atraso de formação do pacote: Tempo necessário para o preenchimento do pacote de voz a ser enviado. Os pacotes de voz são determinados pelo tamanho do vetor ou célula gerada pelo *vocoder* multiplicado pelo número de células que compõe este pacote. Atrasos médios são da ordem de 20 a 30 ms para a formação de uma célula. A maioria dos sistemas utiliza apenas uma célula por pacote, o que pode reduzir a eficiência do sistema, visto que o conteúdo útil do pacote é reduzido em relação ao tamanho total do pacote.
- Atraso de rede: Tempo necessário para o transporte pela rede do pacote do terminal origem até o terminal destino. É um tempo variável e que pode comprometer o sistema. Este tempo é a soma do tempo gasto em todos os roteadores da rede, além dos sistemas de transmissão dos próprios terminais.

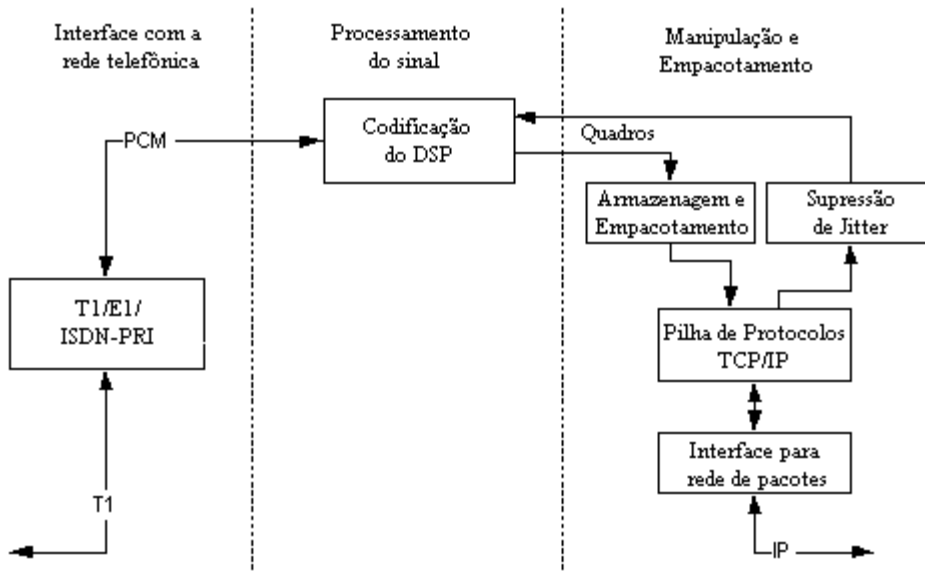
Sempre existe um processo de formação do pacote e outro correspondente para reprodução da voz. Estes dois podem acontecer no computador de origem e/ou destino ou então num gateway correspondente que faz a interface com uma rede de telefonia convencional.

A figura a seguir ilustra esta situação e mostra os responsáveis pelos atrasos.



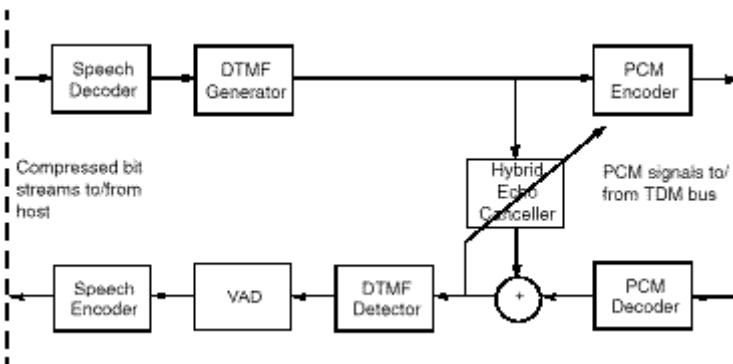
Atraso de formação do pacote

A próxima figura vai indicar o diagrama de blocos que permitirá entender o atraso de formação do pacote que se passa internamente os gateways. Poderemos analisar assim melhor este elemento.

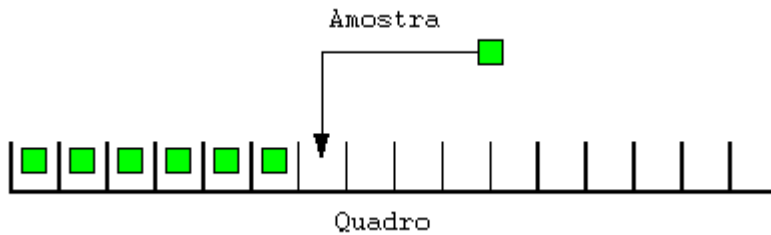


- **Interface com a rede telefônica:** A Rede telefônica utiliza de interfaces T1/E1 ou PRI para seus feixes PCM que devem ser convertidos para uma interface interna do gateway. Tipicamente esta conversão não é maior que 1 ms.

- **Atraso de processamento do sinal:** Um processador interno (DSP) é normalmente dedicado para esta função dada sua grande complexidade. Algumas funções são executadas por esta etapa: compressão e descompressão da voz, detecção de silêncio, detecção e geração de tom, cancelamento de eco, etc. Este processamento é chamado de codificação de voz ou *vocoder*. A próxima figura ilustra estas funções:



- **Atraso de empacotamento:** Os *vocoders* mais eficientes permitem que um quadro de voz seja tratado de uma única vez. Isto permite que instruções especiais possam ser usadas, garantindo a eficiência de soluções de telefonia IP de alta densidade. Assim, várias amostras de voz são tratadas simultaneamente.



Isto porém tem uma desvantagem. Nenhum dos dados pode ser processado até que o quadro esteja completamente cheio. Como a taxa de geração de amostras é normalmente 8000 vezes por segundo, o tamanho do quadro vai então interferir diretamente no atraso final. Um quadro de 100 amostras vai demorar então 12,5 ms, por exemplo. A tabela mostra então os principais *vocoders* e os tempos inseridos para formação do quadro.

<i>Vocoder</i>	Banda (em bits/s)	Duração do quadro (em milissegundos)	Tamanho do quadro (em bytes)
G.711*	64000	15	120
G.723.1	5300-6300	30	24
G.729a	8000	10	10
SX7300	7300	15	14
SX9600	9600	15	18

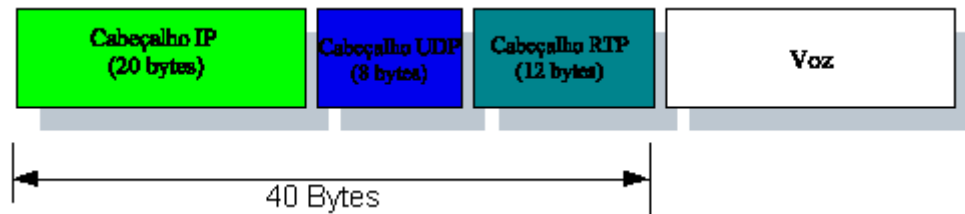
- **Tempo de processamento:** Depois que o quadro foi preenchido, o DSP deve processar o sinal de forma a resultar então de um sinal já codificado. Existe um tempo de processamento variável em função do poder de processamento do DSP. Se o elemento que estiver processando não for um DSP, como em computadores pessoais usados para telefonia IP, o tempo é gasto pelo processador central do sistema.

Sistemas de alta densidade podem ainda usar o mesmo processador para vários canais, de forma a aumentar o tempo de processamento final. Todavia este tempo normalmente é menor que o tempo de quadro, de forma que o processador já esteja ocioso na chegada do próximo quadro.

- **Manipulação do Pacote:** Processo efetuado sobre o pacote até que ele seja enviado para a rede, como passagem por buffers de saída, transmissão do DSP para a CPU, etc. Na recepção ainda podemos identificar um elemento chamado buffer de supressão de *jitter*. Vamos discutir sobre isto mais à frente. Entre o DSP, responsável pela criação do pacote a partir do fluxo de voz, e a CPU, responsável pelo envio para a rede, podem também haver buffers possivelmente implementados por hardware que podem também incluir atrasos em função do tamanho do pacote. Por exemplo um quadro do *vocoder* G.723.1 é maior que o

quadro do G.729 quase 3 vezes. Assim, se um mesmo *gateway* dois canais diferentes usarem estes *vocoders* simultaneamente, pode ser incluídos 3 pacotes G.729 para cada G.723.1 para facilitar a manipulação dos *buffers*. Consequentemente isto ocasionará um atraso adicional aos quadros do canal G.729.

- **Empacotamento:** Na CPU, os quadros são então montados em pacotes, para serem enviados pela rede. Este processo é feito pela pilha de protocolos TCP/IP. Os protocolos RTP, UDP e IP atuam sobre os quadros, de forma a termos como resultados pacotes com tratamento específico para a aplicação. Neste caso, a preocupação é a redução do tempo perdido com confirmação e retransmissão, ausentes nestes protocolos.



Considerando os três cabeçalhos, temos 40 bytes para informações como endereço, portas de serviços utilizadas, identificação do fluxo de voz, número de sequência, etc.

Para um pacote contendo voz no formato G.723.1, temos então, 40 bytes de cabeçalho para 24 bytes de voz, para transmissão de um único quadro, gerados a cada 30 ms. Isto representa uma eficiência de apenas 37,5 %. Uma opção é a inclusão de vários quadros no mesmo pacote, de forma a aumentar esta eficiência. Isto porém inclui um atraso igual ao número de quadros adicionais vezes o tempo de formação do pacote. Esta prática pode ser utilizada, por exemplo junto ao uso do *vocoder* G.729, pois o tempo de formação do pacote é baixo.

Uma outra alternativa para aumentar a eficiência é a inclusão de outro canal no mesmo pacote. Isto porém só é possível se o endereço de destino é o mesmo, como entre dois *gateways*. Esta alternativa não está prevista no padrão H.323, mas pode ser usado em aplicações proprietárias.

Atraso de Rede

Uma vez que os pacotes estão prontos para serem enviados, uma série de atrasos ainda podem ser incluídos de forma a aumentar em muito o atraso final. Eles estão relacionados com o processo de transmissão do pacote pela rede:

- **Atraso de acesso ao meio** - Este atraso é incluído todas as vezes que um pacote passa por um meio físico diferente. Por exemplo, se um pacote utiliza um modem analógico de 28800 bps, por exemplo, cada byte tem um tempo de transmissão de 0,35 segundos. Assim, a um pacote de 100 bytes será agregado um atraso de 35 ms. Acada roteador no percurso do pacote um atraso deste pode ser incluído. Se uma interface de 2Mbps é usada, como o atraso por pacotes cai para 0,4 ms.

- **Atraso de roteamento** - A rede IP é uma rede comutada por pacotes, como já discutimos. Isto implica que existam roteadores que determinem o caminho do pacote pela rede, durante seu trajeto, de forma que sua rota não está definida a priori e pode mudar dinamicamente durante o trajeto do pacote.

A política normalmente utilizada pelos roteadores é a política do melhor esforço, conforme discutido na seção de [qualidade de serviço](#). Desta forma, os pacotes recebem atrasos adicionais nas filas dos roteadores e podem até ser descartados, caso o roteador não tenha capacidade suficiente para atender à demanda de tráfego.

- **Atraso em *firewalls* e *proxy*** - Equipamentos do tipo *firewall* e *proxy* podem ser encontrados ao longo da rede IP, com a finalidade de adicionar segurança ao sistema. Estes equipamentos, além de adicionarem outros *buffers* ao sistema, têm de verificar o conteúdo do pacote, que pode necessitar de um tempo razoável, caso o equipamento não tenha boa resposta de processamento.

Jitter

É interessante observar no atraso de rede, que seu valor varia muito em função do ambiente e ao longo do tempo, sendo impossível dimensioná-lo se estiver baseado nas redes IP convencionais. Se todos os pacotes recebessem o mesmo atraso, este atraso seria propagado integralmente sobre a voz e seria o atraso final da voz.

O problema é a variação deste atraso. Como os pacotes chegam com diferentes atrasos, estes pacotes não podem reproduzir a voz diretamente ao usuário, pois, desta forma a voz sofreria cortes em função de atrasos maiores. A variação do atraso é conhecida como *jitter*.

Supressão de *jitter*

O *jitter* deve ser tratado para evitar efeitos indesejáveis sobre a voz. Se a voz fosse reproduzida de forma instantânea à sua recepção pelo terminal de telefonia IP, ela ficaria extremamente dependente do *jitter*. Por exemplo, se cada pacote representa 35 ms do fluxo de voz, eles deveriam chegar ao terminal com intervalo de 35 ms. Porém, graças ao efeito da rede, este intervalo varia, alguns pacotes podem chegar atrasados, resultando na

interrupção da voz. Outros pacotes podem ainda chegar num intervalo menor, que gera uma sobreposição da voz.

Para a manutenção do fluxo de voz, é necessário um armazenamento de pacotes por um tempo superior ao maior *jitter* observado. Como não existe previsão do valor do *jitter* em função inclusive da possível perda de pacotes, estima-se um valor médio para criação deste armazenamento. Este processo é chamado de supressão de *jitter* e como já foi observado, este armazenamento gera um atraso chamado atraso de supressão de *jitter*.

Controlando o atraso

Controlar o atraso é o ponto chave na tecnologia de Telefonia IP para permitir seu uso em larga escala. A solução prevê sua utilização em redes cujo tráfego seja conhecido e controlado. Algumas medidas podem ser tomadas no sentido de controlar a rede e assim o atraso:

- Uso de equipamentos de rede que possibilitem um controle de qualidade de serviço, seja ele através da priorização de pacotes ou através da reserva de recursos na rede;
- Garantir que a rede tem largura de banda suficiente para o tráfego observado. Algumas ferramentas de medição de tráfego podem ser usadas neste sentido;
- Verificar a disponibilidade de banda a cada instante, bem como a taxa de perdas e de erros e permitir a mudança dinâmica de *vocoder* utilizado, de forma a se adaptar às condições da rede. Alguns equipamentos já permitem esta mudança dinâmica;
- Verificar o atraso nos equipamentos que estão fora de controle da rede, como a Internet pública, por exemplo e se necessário alterar a velocidade dos enlaces para a Internet e a empresa fornecedora do serviço.
- Aumentar a eficiência dos pacotes, utilizando, se possível dois ou mais canais por pacote. Isto pode reduzir o tráfego da rede em até 50%.

Conversão de pacotes em fluxo de voz

Os pacotes são transportados sobre uma rede IP através do protocolo de transporte RTP (*Real Time Protocol*). Embora o RTP seja um protocolo da camada de transporte, ele utiliza de *sockets* UDP como mecanismo de transporte.

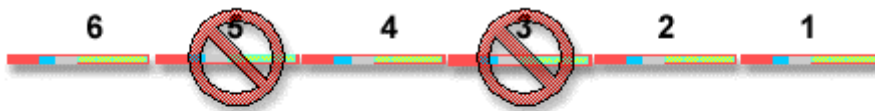
Como protocolo para tráfego em tempo real algumas características são mantidas:

- Não é orientado a conexão;
- Sem retransmissão de pacotes com erros ou pacotes que não chegaram ao destino;
- Descarte de pacotes que chegaram ao destino com atraso maior que o atraso máximo

Cabe à aplicação montar os pacotes recebidos de forma a manter o fluxo de voz para o usuário. Este fluxo de voz deve ser contínuo, independente do recebimento ou não dos pacotes correspondentes.

Já verificamos a necessidade de armazenamento dos pacotes antes deles serem reproduzidos como fluxo de voz, para eliminação de *jitter*. Assim, uma fila é criada de forma que são reproduzidos os pacotes do início da fila, com taxa correspondente àquela gerada pelo *vocoder* quando do processo de geração de pacotes, à medida que a fila recebe mais pacotes ao final através da rede. Os pacotes então são entregues ao *vocoder* para que o processo inverso ocorra, ou seja, o processo de decodificação dos pacotes na voz original.

Para garantir a continuidade da voz quando o pacote correspondente é perdido, seja por atraso, erros ou perda real do pacote, o espaço deve ser preenchido. Como alguns fonemas podem durar um tempo maior que o tempo de cada pacote, a repetição do pacote anterior por reduzir o impacto sobre a qualidade da voz. A política utilizada então é de substituição pelo último pacote válido, com redução do volume, para que o usuário perceba o mínimo



Também é importante garantir a sequência dos pacotes. Para isto o RTP provê um número básico de sequência que possibilita à aplicação obter informações em relação à ordem dos pacotes. Este número de sequência é então verificado para garantir a integridade do fluxo de voz.

Podemos verificar então a presença de pacotes atrasados, com número de sequência menor que o número esperado, que devem ser descartados, e pacotes adiantados, com número de sequência maior que o esperado. A fila de pacotes deve então ser ordenada de forma que os pacotes possam ser reproduzidos em ordem.





Aspectos Econômicos

O tráfego de pacotes hoje já supera o tráfego de voz em algumas operadoras de telecomunicações do mundo. Isto deve se firmar como verdade nos próximos anos, de forma que seja necessário a instalação de grandes backbones para o fluxo deste tráfego. A utilização de Telefonia IP permite a utilização deste mesmo backbone para o fluxo de tráfego de voz. Esta unificação em si já representa grande economia, pois exige-se a manutenção de apenas um backbone.

Outro argumento econômico a favor da Telefonia IP é o custo de transmissão e comutação entre centrais telefônicas. Vamos definir e analisar quatro tipos de centrais de comutação e suas interfaces, com base em pesquisa da Cisco Systems[4].

1. Centrais de comutação por circuito e interfaces por circuito, normalmente utilizadas na rede telefônica convencional, não utilizam *vocoders* para conversão de voz para pacotes. Vamos definir seu custo como 1 para comparação com as outras centrais que vamos descrever. O custo de ci;
2. Centrais de comutação por pacotes e interfaces por pacotes. Pode ser simplesmente um roteador ou *switch* ATM. Não utiliza de *vocoders* porque trabalha somente com pacotes. Baseado na central anterior, esta central tem custo menor que 0,1.
3. Centrais de comutação por pacote, mas com interface por circuito em uma das extremidades e pacotes em outra. Esta central pode ser interface entre uma rede de pacotes e uma rede de circuitos, atuando como *gateway*. Utiliza um *vocoder* para cada comunicação que passa por ela. Seu custo é de 0,8 em relação à primeira central apresentada.
4. Centrais com comutação por circuito e todas interfaces de pacotes. Utiliza dois níveis de *vocoders* para cada comunicação que passa por ela. Estes dois *vocoders* ainda não apresentam baixas de qualidade significativas, de forma a ser ainda aceitável. O custo desta central é 1,5 em relação à primeira central apresentada.

Na figura seguinte vemos estes quatro tipos de comutadores:

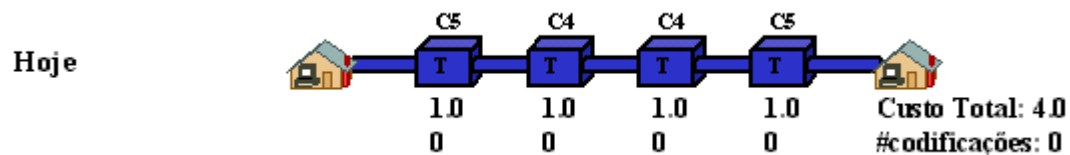
	Interface	Tecnologia	Interface	Custo	#Vocoders
	TDM	TDM	TDM	1.0	0
	IP	IP	IP	<0.1	0
	TDM	Ambos	IP	0.8	1
	IP	TDM	IP	1.5	2

Confirmamos estes números analisando o custo de equipamentos de telecomunicações. O custo de um assinante em um sistema convencional varia entre 100 e

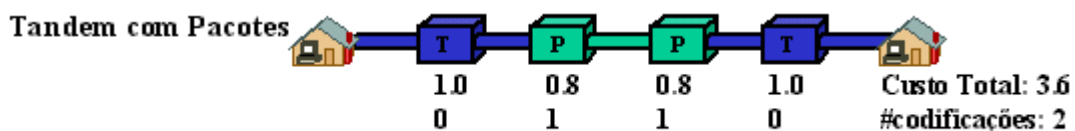
200 dólares, o que implica em um custo de 200 a 400 mil dólares por circuito OC3 de interface. O mesmo circuito OC3, em um comutador ATM custa cerca de 5 mil dólares.

Utilizando então destes modelos de centrais telefônicas, levando em consideração a comunicação entre dois assinantes A e B, a qual deve transpor duas centrais *Tandem*, que tem funções de interligação de centrais locais, além das respectivas centrais locais, podemos imaginar os seguintes ambientes:

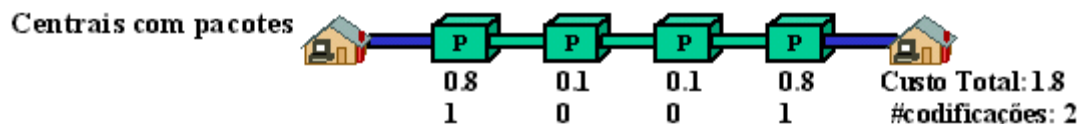
1. Hoje: Todas as centrais utilizam de telefonia convencional para realizar a comunicação entre os assinante A e B. Em cada uma aloca-se recursos com custo unitário, seguindo a definição que fizemos anteriormente de custos, com custo total de 4.0.



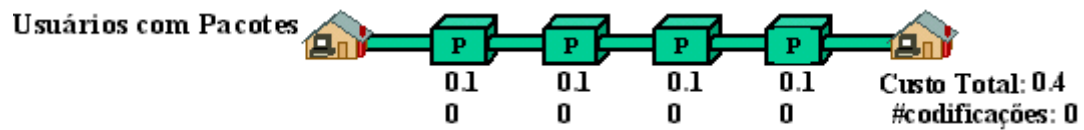
2. *Tandem com pacotes*: Substituindo as centrais intermediárias por centrais com tecnologia baseada em pacotes, como a telefonia IP, chegamos a um custo final de 3.6, inserindo dois processos de codificação/decodificação. É importante observar que este processo pode inserir perdas na qualidade acústica final da voz. Até 2 codificações são aceitáveis porém dentro dos limites da percepção humana.



3. Centrais com pacotes: Substituindo todas as centrais por tecnologia de comutação baseada em pacotes, reduzimos o custo final para 1.8. Isto porque as centrais *tandem* não necessitarão mais realizar a função de codificação/decodificação que será realizado nas centrais locais.



4. Usuários com pacotes: Neste último ambiente, todos os usuários já utilizariam da tecnologia de VoIP a partir de suas residências, através de aparelhos telefônicos especiais ou através de seus computadores pessoais. A interligação à central seria através de uma rede metropolitana, possivelmente interligando pequenas redes locais. Toda a tecnologia convencional seria substituída por tecnologia IP, e o custo final seria da ordem de 0.4. O processo de codificação/decodificação da voz seria realizado na própria residência do usuário, sem apresentar perdas significativas.



Assim podemos verificar o impacto da telefonia IP sobre as redes de telecomunicações, reduzindo drasticamente o custo de comutação. Outra abordagem deveria ser dada ao custo de transmissão que é também reduzido sensivelmente com o uso de VoIP. Isto porque com técnicas de codificação de voz e supressão de silêncio adequadas, chegamos a taxas de 4kbps para cada canal de voz em oposição aos 64kbps da telefonia convencional, com uma redução de 16 vezes o tráfego final.

Computação Móvel

Computação Móvel é o último estágio do desenvolvimento da computação pessoal. Em 1946, Illinois Bell Telephone Company introduziu um serviço de telefonia móvel, que permitia usuários na direção de veículos, comunicar-se com o sistema telefônico. Foi a primeira iniciativa de permitir comunicação bidirecional sem fio. Nos nossos dias temos a popularização do telefone celular, que garante mobilidade dos usuários do sistema telefônico[1].

São muitos os argumentos que nos levam a desenvolver um ambiente de computação móvel. Devemos pensar em primeiro lugar que nem todas as pessoas que necessitam do computador estão presentes no escritório. Vendedores são exemplos de trabalhadores com alta mobilidade que necessitam acessar bases de dados remotas e executarem operações diversas como emissão de pedidos, requisição de mercadorias, etc. Mesmo os usuais usuários de computadores no escritório passam por momentos de difícil acesso ao escritório, como em viagens, cursos, congressos, etc.

Podemos enumerar diversas vantagens de um sistema móvel como conforto, para utilização em qualquer ambiente; flexibilidade, para utilização em diversas aplicações que exijam movimento; disponibilidade, independente da localização do usuário; e algumas exigências, como portabilidade, facilitando o transporte; autonomia de energia, para garantir o funcionamento onde não existe disponibilidade de energia; e desempenho comparável a estações fixas.

A exemplo dos computadores pessoas fixos, deseja-se garantir capacidades de comunicação aos computadores móveis, de uma forma *wireless*, ou seja, sem fio. Esta comunicação deve ser feita através de sistema de ondas de rádio, utilizando de antenas para transmissão e recepção. Pela portabilidade do sistema, espera-se que as antenas não sejam grandes e nem mesmo muito potentes, para não ferir a autonomia de energia do sistema.

Um ambiente de **computação móvel** envolve computadores portáteis interligados em rede através de sistema de ondas de rádio. Vamos chamar de Unidade Móvel (UM) ao elemento de rede, computador portátil, interligado à rede de computação móvel. As Unidades Móveis são agrupadas ligadas em rede, divididas em subredes, cada qual mantida por um grupo de antenas, às quais vamos chamar de Ponto de Acesso. Não é estritamente necessária a presença de Pontos de Acesso. Um ambiente de computação móvel pode ser desenvolvido de forma independente de infraestrutura fixa, o que chamamos de redes Ad Hoc, onde eles comunicam entre si diretamente através das antenas.

A mobilidade porém, sempre implica em algumas condições, típicas do ambiente, que devem ser consideradas, independente do sistema de acesso:

- Capacidade de comunicação limitada com largura de banda variável e alta taxa de erros;

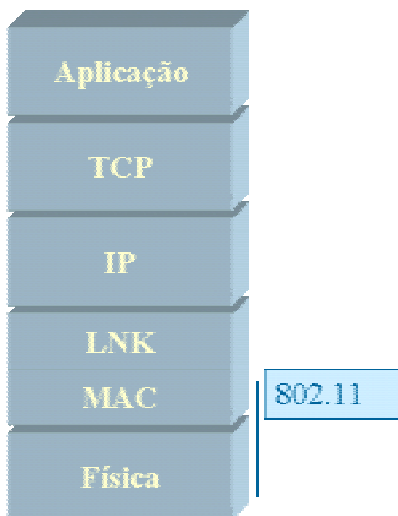
- Autonomia de energia limitada por baterias com limite de consumo, de forma que deve-se dispendir o mínimo de energia com processamento e dispositivos de apoio ao sistema.
- Limites físicos de hardware para garantia de portabilidade, limitando também o poder de processamento e dispositivos;
- Problemas de roteamento de pacotes quando variação da subrede onde está presente a UM.
- Perda temporária de comunicação quando deslocamento entre áreas mantidas por diferentes estações de rádio e renegociação de características de acesso.

Para contornar estes problemas, diversas soluções tem sido propostas, algumas das quais já em operação, outras em pesquisa. É o que vamos discutir em seção futura.

A princípio podemos utilizar o próprio sistema de telefonia celular para fazer a interconexão destes computadores em rede, porém, algumas limitações estão presentes neste tipo de acesso, especialmente no que diz respeito à baixa velocidade de acesso obtida neste sistema. Para estes sistema, a Unidade Móvel é um computador portátil utilizando de um modem e um telefone celular para se conectar à um servidor de acesso remoto. As estações de rádio são as estações do sistemas de telefonia celular e todas as operações de mobilidade são gerenciadas pela camada física deste sistema, provida pelo sistema de telefonia celular. Este sistema não é de nosso interesse, devido à baixa taxa de transmissão obtida, na faixa de 9.6 Kbps.

Outras formas de comunicação podem também ser usadas como *Wavelan* que permite a utilização de taxas de transmissão mais satisfatórias, entre 1 e 2 Mbps. Este sistema deve resolver porém problemas de roteamento, *handoff* que são resolvidos pelo sistema celular na abordagem anterior. Na próxima seção vamos abordar mais sobre este sistema, padronizado pelo IEEE 802.11.

IEEE 802.11



Este padrão, definido pelo IEEE em 1997, prevê uma rede local sem fio, conhecida por *LAN wireless*, em atendimento à padronização dos equipamentos já em uso para este tipo de rede. Estão definidos neste padrão a camada física, subcamada de acesso ao meio (MAC), como qualquer padrão 802.X e infraestrutura de rede.

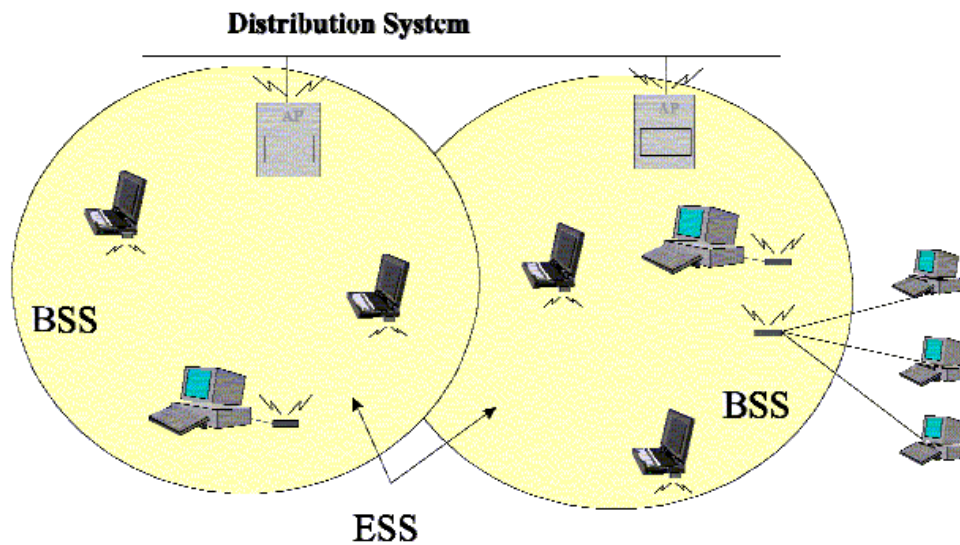
Uma rede IEEE 802.11 é baseada na arquitetura celular, onde o sistema é subdividido em células. Cada célula, chamada Conjunto de Serviços Básicos (*BSS*) é controlada por uma estação base, chamada Ponto de Acesso.

Embora uma rede local sem fio possa ser formada por uma simples célula, com um único ponto de acesso, e

até mesmo sem nenhum ponto de acesso, a maioria das instalações irão ser formadas por muitas células, onde os Pontos de Acesso são conectados para o mesmo tipo de backbone, conhecido como Sistema de Distribuição. A interconectividade da rede local sem fio é garantida por uma rede 802 simples nas camadas superiores do modelo OSI. A esta interconectividade damos o nome de Conjunto de Serviços Estendidos ou ESS.

Quando uma estação é ligada ou entra na área de cobertura de um ponto de acesso, temos o processo de **associação**. AO se deslocar de uma região coberta de um ponto de acesso para outro ponto de acesso, o processo realizado é de **reassociação**. Uma única célula pode cobrir a área de 5000 metros quadrados.

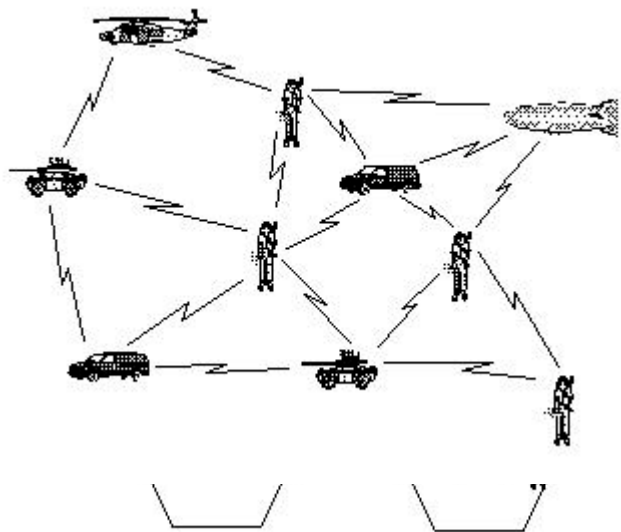
O diagrama a seguir mostra uma rede típica 802.11:



Na camada física temos três possibilidades de acesso:

- Infravermelho;
- Direct Sequenc Spread Spectrum (DSSS) a 2,4GHz com acesso direto, e
- Frequency Hopping Spread Spectrum a 2,4GHz com chaveamento de frequência entre 79 canais de 1 MHz com objetivo de restringir acesso aos dados.

Na camada MAC, o acesso ao meio é feito através do método CSMA/CA (*Carrier-sense Medium Access / Collision avoidance*). CSMA é bem conhecido no mercado por ser usado em Ethernet. Porém nesta rede o método é conhecido por CSMA/CD (CD para *Collision Detect*).



CSMA trabalha da seguinte maneira: Uma estação solicitada para transmitir escuta o meio. Se o meio está ocupado com a transmissão de outra estação, então a estação atrasa sua transmissão por um determinado período de tempo. Se o meio está livre, a estação transmite. No 802.11, através do uso de pacotes do tipo RTS (*Request to Send*), CTS(*Clear to Send*) e ACK(Confirmação), as estações são capazes de solicitar o uso de meio antes de transmitir, evitando assim colisões.

Não há limites teóricos para o número de usuários de uma mesmo ponto de acesso. Na prática, cerca de

120 usuários dividem o uso de um ponto de acesso sem problemas.

Infraestrutura Fixa X Ad-Hoc

Ainda é previsto dentro do IEEE 802.11 duas formas de conexão entre redes móveis. As redes de infraestrutura fixa e as redes independentes ou ad hoc.

- Redes de infraestrutura fixa: Neste tipo de rede o ponto de acesso é utilizado para comunicação entre as unidades móveis, de forma que uma unidade móvel sempre se comunica com outra somente através de um ponto de acesso. Assim, a rede fixa dá suporte à mobilidade e auxilia em tarefas como roteamento, processamento distribuído, redução de tráfego, adaptabilidade, etc.

- Redes Independentes ou Ad Hoc - Nestas redes, não é prevista a existência de qualquer infraestrutura fixa, e, quando existe, é conhecida pela rede como qualquer outra unidade móvel. As unidades móveis comunicam-se diretamente através do meio dispensando infraestruturas fixas, que podem encarecer o projeto. Se uma unidade deseja se comunicar com outra que não está dentro de seu alcance, ela o faz através de outras unidades móveis, que retransmitem os pacotes até que ele alcance seu destino. A maior dificuldade neste caso é o roteamento dos pacotes até o destino.

Segurança

Para manutenção da segurança dos dados, está prevista a utilização de algoritmo de chave compartilhada de 40 bits, conhecido como RC4. Este esquema é conhecido como WEP - *Wireless Equivalent Privacy*.

Internet Móvel

Após estabelecimento de mecanismos nas camadas inferiores que possibilitam o uso de computação móvel, devemos adaptar também as camadas superiores para permitir o seu uso de forma mais ampla. O primeiro passo é na camada de rede para permitir o uso de computação móvel na Internet.

Espera-se que uma unidade móvel seja reconhecida na Internet independente da subrede que se encontre. Sabemos porém que o endereço de uma estação depende de sua subrede. Esta subrede pode ser modificada com muita facilidade, de forma que a configuração de subrede presente não é cômodo que o endereço seja constantemente modificado. Um mecanismo transparente ao usuário deve ser utilizado com esta finalidade.

Para servidores móveis o problema é mais crítico ainda, pois deve ser reconhecido na Internet com um endereço fixo, independente da subrede e que se encontra. Espera-se também que a estação fixa em comunicação com a unidade móvel não veja diferenças entre esta unidade e outra estação qualquer. Para resolver estes problemas, foi proposto pelo IETF o IP Móvel, que é uma adaptação do protocolo IP convencional para redes móveis.

Roteamento

Seja uma unidade móvel presente em uma rede, do tipo internet, com seu endereço IP único capaz de garantir a capacidade de entrega correta de pacotes a ela destinada. Esta UM utiliza de serviços de um roteador presente na sua sub-rede para que possa comunicar-se que elementos presentes em outras subredes.

Em um ambiente de computação móvel, uma subrede é mantida por um ou mais pontos de acesso que mantém comunicação com a unidade móvel. Para garantir total mobilidade, devemos prever a possibilidade de locomoção da unidade móvel entre áreas mantidas por diferentes pontos de acesso, possivelmente mantidas em diferentes subredes.

Esta possibilidade de locomoção entre diferentes subredes deve ser transparente para o usuário, de forma a não lhe causar incômodo. Sob o ponto de vista do protocolo de rede, isto implica na mudança do endereço de rede da unidade móvel, pois seus pacotes passam a ser enviados e recebidos por outra rota. Para tratamento desta abordagem para o protocolo IP, temos a opção do IP Móvel, proposto pelo IETF[3].

Uma unidade móvel presente em uma subrede diferente, chamada estrangeira, da sua subrede de origem é chamada unidade móvel visitante. O IP Móvel prevê basicamente a associação de um número IP dinâmico e provisório para a unidade móvel visitante, a partir de um servidor de suporte à mobilidade presente na subrede estrangeira.

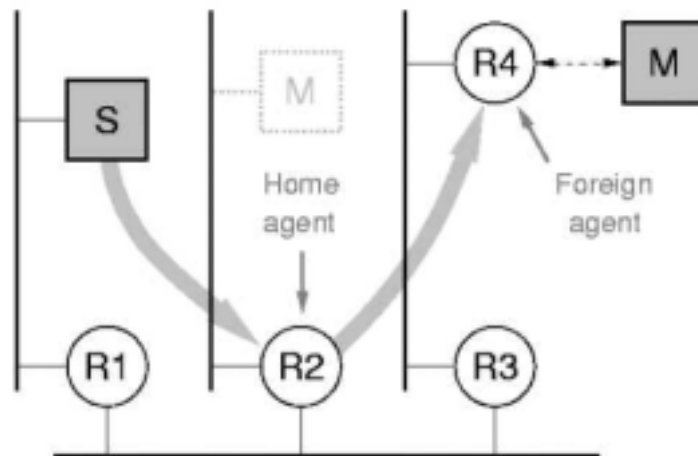
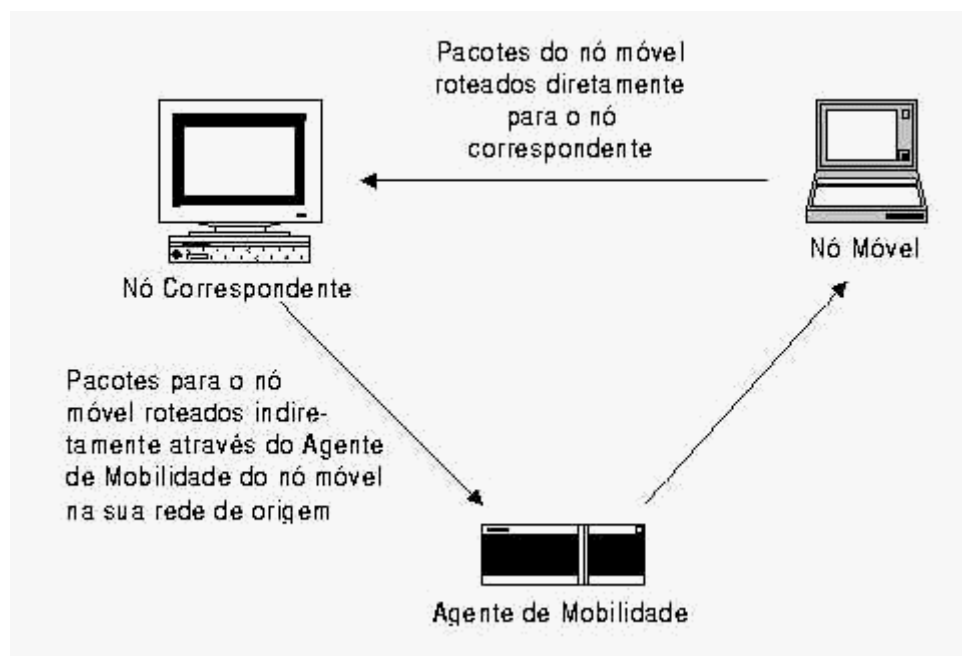


Figure 1 Basic architecture of the IETF Mobile IP protocol

Na subrede de origem da unidade móvel, deve estar presente um agente de mobilidade, configurado junto ao roteador da subrede de origem. Ao entrar em uma nova subrede, o unidade móvel recebe um novo endereço e o envia para o seu agente de mobilidade. Este agente de mobilidade é responsável por escutar todos os pacotes que são enviados para a unidade móvel quando esta não estiver presente na subrede de origem, encapsular estes pacotes e enviá-los para a subrede onde está presente a unidade móvel neste momento, endereçados ao novo endereço da unidade móvel. Este processo de encapsulamento é baseado em IP-in-IP[4].



Os pacotes a partir da unidade móvel para outro endereço qualquer de rede seguem diretamente para o destino, mas os pacotes com destino para a unidade móvel sempre

devem passar pelo agente de mobilidade, por serem endereçados para a subrede de origem da unidade móvel.

O IP Móvel prevê ainda, caso a rede de comunicação suporte mobilidade, que ela seja capaz de identificar, através de uma primeira comunicação com o agente de mobilidade, o endereço IP utilizado pela unidade móvel na subrede estrangeira, e permita a comunicação do nó correspondente diretamente com a unidade móvel na sua nova subrede. Isto porém, implica na manutenção de tabelas de conversão de endereços espalhadas pela rede, de forma a permitir o roteamento correto destes pacotes.

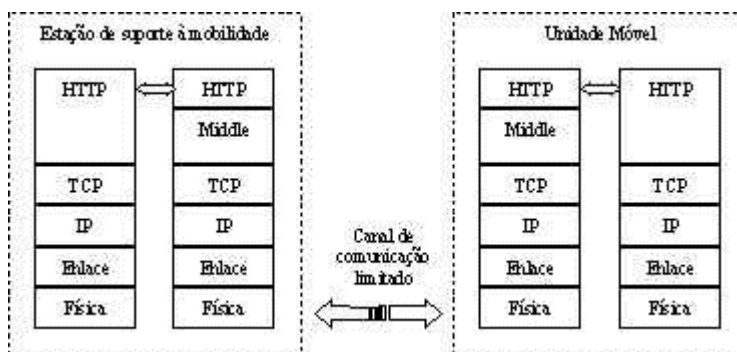
Adaptabilidade

Espera-se que uma unidade móvel seja capaz de acessar a todos os recursos disponíveis em uma estação ligada a uma rede através de cabeamento, contornando todos os problemas deste ambiente. Lança-se mão, então, de técnicas para maior aproveitamento do canal de comunicação, da energia do sistema e do processamento, de forma a limitar as tarefas executadas na UM. Podemos verificar algumas abordagens interessantes:

- Alguns elementos como estações fixas de suporte à mobilidade podem aumentar consideravelmente a capacidade do sistema em termos de comunicação e processamento[2]. A estas estações são reservadas determinadas tarefas como pre-processamento do tráfego de comunicação com o objetivo de reduzir a largura de banda necessária e adaptação dos serviços disponíveis para as unidades móveis.
- Protocolos nas diversas camadas podem ser adaptados para resolver problemas específicos. Podemos citar IP móvel para garantir capacidades de roteamento em ambientes móveis, e outras adaptações, especialmente na camada de aplicação, com o objetivo de reduzir a largura de banda necessária ou aumentar a confiabilidade.

Devemos adaptar a aplicação, o protocolo utilizado ou mesmo a pilha de protocolos utilizados no ambiente de computação móvel para garantir os mesmos parâmetros de QoS que seriam garantidos em uma aplicação semelhante em uma rede fixa.

Podemos alterar a pilha de protocolos entre a UM e estação fixa, de forma a criar uma camada intermediária que agregue as características necessárias. A figura seguinte apresenta um exemplo desta abordagem em relação ao protocolo HTTP.



No exemplo, a camada *Middle* é responsável por fazer uma adaptação do protocolo HTTP para o meio, de forma a economizar a banda e possivelmente o processamento. O resultado pode ser transparente para o usuário ou não.

Uma possibilidade é o uso de compactação na camada *Middle*. Isto reduz a banda de transmissão de forma transparente para o usuário. Esta função está prevista em camadas da pilha de protocolos do modelo OSI. Além da compactação, serviços de criptografia podem ser adicionados, de forma garantir restrições de segurança à rede móvel, que está mais sujeita a invasões devido à ausência de privacidade do meio.

Outra abordagem poderia ser de realizar um tratamento prévio sobre serviços oferecidos ao usuário. Objetos HTTP poderiam receber um tratamento prévio com o objetivo de reduzir o tamanho na transmissão. Por exemplo, as figuras poderiam ser inicialmente transmitidas e mostradas em preto-e-branco e, somente as figuras necessárias e requisitadas sob a ação do usuário seriam então retransmitidas coloridas. Este tratamento por sua vez não é transparente para o usuário.

Ao vislumbrar agora o serviço de Telefonia IP, desejamos adaptá-lo de forma a possibilitar seu uso em um ambiente de computação móvel. Vamos verificar as características deste serviço, para propor algumas adaptações:

- Tráfego de tempo real, de forma que um pacote só será útil se for recebido no tempo certo. Retransmissão é desnecessária. Esta característica é garantida pelo protocolo de transporte RTP.
- Exigência de largura de banda constante. Uma queda de comunicação significa fim da conversação. Esta largura de banda é determinada pelo algoritmo de codificação da voz, compressão e supressão de silêncio.
- Serviço muito sensível a erros. A perda de um pacote pode implicar em ruídos na saída de voz.
- Altas exigências de QoS. Especialmente no que diz respeito ao tempo de atraso dos pacotes e *jitter*.

O uso de diversos algoritmos de codificação de voz influi na largura de banda necessária para o serviço e na qualidade da reprodução da voz no outro extremo da conversação, de forma a ter possivelmente uma reprodução de voz mais mecanizada, com perda de alguns sons da fala.

A adaptabilidade a ser trabalhada para telefonia IP pode ser feita sobre estes algoritmos de codificação e decodificação, definidos pela ITU-T, além da possibilidade de compactação e supressão de silêncio. Assim, a estação fixa pode alterar dinamicamente estes métodos, atuando diretamente na voz, de acordo com a disponibilidade de largura de banda.

Outras propostas estão sendo estudadas para garantir esta adaptação:

- Transmissão de pacotes maiores, contendo não só o conteúdo de voz atual, mas também frações dos pacotes anteriores, permitindo assim uma maior recuperação em caso de erros.

Esta abordagem já foi utilizada pela *Lucent* em redes fixas e aumenta a largura de banda utilizada por este serviço;

- Retransmissão constante de pacotes, com o objetivo de substituir pacotes perdidos. Podemos observar que esta abordagem deve aumentar a utilização de largura de banda para este serviço, em tantas vezes quanto for feita a retransmissão.

Bibliografia

[1] Bates, Regis J; Gregory, Donald; *Voice and Data Communications Handbook* – McGraw-Hill Series on Computer Communications, 1997

[2] Mateus, Geraldo R.; Loureiro, Antônio A. F.; *Introdução à Computação Móvel* – 11^a Escola de Computação - SBC, 1998

[3] Charles Perkins, [IP Mobility Support version 2 - INTERNET DRAFT](http://www.ietf.org)- IETF - <http://www.ietf.org>, 1997

[4] Charles Perkins, [IP Encapsulation within IP](http://www.ietf.org). RFC 2003, IETF - <http://www.ietf.org>, 1996

[5] S. Deering, [Internet Protocol version 6](http://www.ietf.org) - RFC 1883 - IETF - <http://www.ietf.org>, 1997

[6] Charles Perkins, [Mobile Networking Through Mobile IP](http://www.ietf.org), IEEE Internet Computing Online, 1997

Qualidade de Serviço

Graças à dimensão que a Internet atingiu no mundo, capilaridade, alcance e presença nas mais diversas partes do mundo, toda ordem de aplicações tem sido desenvolvidas utilizando sua infra-estrutura e capacidade de comunicação. Algumas aplicações novas como Telefonia IP ou Vídeo sob Demanda podem apresentar dados multimídia em tempo real. Outros, que usam baixa largura de banda e baseadas em texto podem apresentar um rótulo de "alta-prioridade" porque eles processam informações comerciais de missão crítica.[1]

Em ambos os casos, a rede deve manipular os pacotes da aplicação de uma forma especial, diferente dos dados que originados de outras aplicações. Mas a Internet e, generalizando, redes IP não oferecem caminhos fáceis para identificação destes pacotes ou consequentemente seu tratamento especial.

Esta situação está mudando com a introdução do conceito de Qualidade de Serviço (QoS), que é a capacidade da rede de prover serviços não convencionais para um subconjunto de tráfegos agregados. Para este subconjunto, podemos definir parâmetros a serem garantidos. O protocolo da camada de rede é o responsável pelo controle de QoS.

Podemos lançar mão de alguns parâmetros para medir a Qualidade de Serviço oferecido aos usuários de unidades móveis, vislumbrando as principais aplicações que queremos desenvolver neste ambiente. Podemos destacar entre estes parâmetros:

- *Bandwidth* - Representa a velocidade do meio. Taxa de transmissão máxima disponível em um determinado momento para a comunicação entre dois nós da rede;
- *Delay* - Atraso médio. Tempo médio de transmissão de um pacote entre dois nós da rede.
- *Jitter* - Variação do *delay*. Medida da variação média do tempo de transmissão de um pacote.

Uma rede garante QoS quando consegue estabelecer parâmetros mínimos para determinado fluxo de informações que serão sempre conseguidos sob quaisquer circunstâncias da rede. Estes parâmetros podem ser associados por aplicação, de acordo com as exigências de cada aplicação.

Ferguson and Huston em seu livro *Quality of Service*[3], classificam aplicações de rede em três categorias: Elásticas, Tolerantes de Tempo Real e Intolerantes de Tempo Real.

Aplicações Elásticas, como HTTP, FTP e a maioria dos processos comerciais, são escritas para executarem tão rápidas quanto a rede permite. O contetor usual para estas aplicações é a rede por si própria. Congestionamentos de rede implicam que o fluxo de dados seja retardado, de forma a aumentar em muito o delay.

Aplicações tolerantes de tempo real são tipicamente fluxos de áudio e vídeo, os quais são executados com algum atraso, por causa dos buffers de recepção que compensam as variações da rede.

Aplicações de telefonia são classificadas na categoria de aplicações intolerantes de tempo real. Pessoas não toleram atrasos na voz de mais de que 200 ms segundos em uma interlocução. Isto cria grandes dificuldades para as aplicações gerenciarem. Buffers devem ser mantidos, mas o tempo a ser armazenado é muito pequeno e pode não ser suficiente.

O problema é que, enquanto aplicações elásticas podem tolerar um atraso maior, elas tentam consumir muito mais recursos de rede que elas podem. Em contraste, aplicações de voz necessitam de poucos recursos da rede, mas estes recursos devem estar disponíveis imediatamente.

Até recentemente, redes IP suportavam apenas uma classe de serviço: a do melhor esforço (*best effort*). A rede desenvolvia o melhor procedimento para entregar pacotes para seu destino mas sem garantias e sem recursos especiais reservados para qualquer dos pacotes.

Primeiro porque o protocolo TCP/IP original foi baseado na idéia de igualdade no acesso e sem tratamento especial para qualquer nó ou serviço. Com exceção dos nós fins da comunicação, nenhuma informação de estado da conexão era mantida em qualquer lugar da rede.

Depois, o trabalho interno dos roteadores usavam um política de transmissão de pacotes do tipo FIFO (*First In, First Out*). Se o roteador recebesse mais pacotes do que pode manipular os pacotes mais novos são descartados.

Podemos verificar algumas iniciativas de modificação do protocolo IP de forma a dotá-lo de garantia de QoS, ou pelo menos otimizar certos parâmetros.

Uma primeira abordagem é desenvolvida pelo grupo de trabalho *Integrated Services* do IETF e consiste basicamente definição de uma arquitetura de integração de serviços, chamada arquitetura IIS (*Internet Integrated Services*). O ponto mais importante deste trabalho é a reserva de largura de banda para atender aos requisitos de QoS de um determinado fluxo. Esta reserva é feita nos nós da rede, roteadores, gateways e outros elementos de rede, utilizando de protocolo específico para isto. Esta solução padrão existente hoje para a QoS é o *Resource Reservation Protocol* (RSVP), presente na maioria dos roteadores atuais. Um cliente RSVP pode reservar uma quantidade adequada de largura de banda para tráfego de voz e prover um fluxo de baixo atraso para os pacotes de voz. Enquanto isto somente é aplicável para os roteadores controlados diretamente, RSVP é também provido da habilidade de comunicar a reserva para outros roteadores.

Esta abordagem implica em manter o estado de cada fluxo pelos roteadores, o que significa uma grande quantidade de estados nos roteadores do "núcleo" (*core*) da rede que transportam um número muito elevado de fluxos. Estes estados e o processamento a eles associado acarretam uma sobrecarga destes roteadores, que passam a ser um gargalo do desempenho. Os problemas de escalabilidade encontrados na arquitetura IIS levaram ao estudo de soluções mais simples e eficientes para a implantação de QoS na Internet.

A segunda abordagem, desenvolvida pelo grupo de trabalho *Differentiated Services* do IETF, busca prover a diferenciação de serviços na Internet, sem a necessidade de um estado para cada fluxo e de sinalização a cada nó, aumentando assim a sua escalabilidade. Assim, esta solução pode ser usada nos roteadores do núcleo da rede[5]. A idéia básica desta proposta é renomear o campo ToS (*Type of Service*) do cabeçalho IP, denominando-o de DS (*Differentiated Service*) e a partir da análise deste campo, prover diferentes tratamentos aos pacotes de cada fluxo transportados pela rede [2]. Este novo campo é utilizado como referência para a prioridade de transporte nas filas de transmissão. Políticas de escalonamento com prioridades são utilizadas visando a manutenção da qualidade dos serviços que necessitam de certos parâmetros de acordo com o conteúdo do campo DS.



A figura acima compara as três abordagens. A primeira é convencional política do melhor esforço que não tem garantia de QoS. Como política intermediária, temos a política de serviços diferenciados, com suas características e finalmente a política de serviços integrados, com suas características.

Qualidade de Serviço para Telefonia IP

Aplicações de telefonia são aplicações intolerantes de tempo real. Para a manutenção de interatividade em uma conversação, assume-se que o retardo fim-a-fim do tráfego de voz deve estar abaixo de um certo valor reduzido. Deve-se considerar ainda que o atraso fim-a-fim implica em um intervalo entre interlocuções igual ao dobro do seu valor. Numa diálogo entre dois usuários A e B, um trecho da fala de A chega a B com o atraso imposto pela rede. O usuário B só pode iniciar sua fala após o fim da fala de A, que já está atrasada. O usuário A, por sua vez, além de perceber o atraso do início da fala de B, ainda percebe o atraso de rede sobre a fala de B, de forma a existir sempre um tempo igual a duas vezes o atraso entre as falas de A e B.

A ITU-T define [6] os valores do retardo fim-a-fim, de acordo com a aplicação:

0 a 150 ms - Aceitável para a maioria das aplicações;

150 a 400 ms - Aceitável desde que se tenha cuidado com o impacto do atraso sobre a qualidade da aplicação para o usuário. Por exemplo, comunicação via satélite podem tolerar atrasos até 400 ms.

Acima de 400 ms - Inaceitável para a maioria das aplicações de rede. Seria prejudicial à interatividade, necessitando de informações a respeito da posse da fala. Torna-se necessário a utilização da palavra "câmbio" ou semelhante ao final de cada fala.

O retardo fim-a-fim não pode contudo ser uma medida para o atraso médio da rede. Isto porque dado as condições da rede, um *jitter* é imposto à comunicação, causando uma variação no intervalo de recepção de pacotes. O atraso imposto à aplicação deve então ser maior que a soma do atraso médio mais o *jitter* observado.

Quanto à largura de banda, as aplicações de telefonia IP necessitam de uma largura de banda fixa, em dependência do *vocoder* utilizado no processo de codificação de voz. Estas taxas variam entre 5.6Kbps e 64 Kbps. Pode ainda ser utilizado um algoritmo para supressão de silêncio, capaz de reduzir a taxa de transmissão em mais de 50 %.

A maioria das atenções tem sido voltadas para a implementação de Telefonia sobre IP como uma solução para redes de longa distância (WAN). Isto faz sentido tendo em vista o quanto se gasta em um WAN, mas uma solução para LAN também deve ser observada pois, sem outras razões, a própria economia de cabeamento justifica. Deve ser observado que, embora o problema de Qualidade de Serviço é o mesmo tanto para WAN quanto para LAN, as soluções são um pouco diferentes[4].

- WAN: Considerando a transmissão de tráfego de voz em redes de longa distância, IP QoS ou outra solução de gerência de largura de banda podem limitar o pico de transmissão do tráfego de dados para garantir banda para voz. Note que a família de padrões H.323 foi criada para LANs sem garantia de QoS, e assim tem uso limitado sobre configurações WAN.
- LAN: A necessidade existente em ambientes LAN é ter um caminho para controlar o fluxo de dados de cada dispositivo agregado à rede. Seria ideal que cada aplicação controlasse sua taxa de envio para a rede, mas isto não é praticável, devido ao grande número de aplicações envolvidas que desconhecem a largura de banda disponível a cada momento. Outra abordagem seria a utilização do sistema operacional, na camada de rede, para controlar o tráfego que é injetado na rede a cada momento. Isto, pelo menos, é possível; de fato, Microsoft já está falando sobre esta capacidade como parte de seu ADSI (*Active Directory Service Interface*).

- [1] Metz, Chris; [IP QoS:Traveling in First Class on the Internet](#) - IEEE Internet Computing OnLine, Maio 1999
- [2] K. Nichols, [Definition of the Differentiated Services Field \(DS Field\) in the IPv4 and IPv6 Headers \(RFC 2474\)](#) - IETF, Dezembro 1998
- [3] Paul Ferguson, Geoff Huston, Carol Long, Quality Of Service: Delivery QoS On The Internet And In Corporate Networks, John Wiley Computer, 1998
- [4] Daniel J. McCullough and John Q. Walker, [Interested in VOIP? How to Proceed](#), from Voice 2000, a supplement to Business Communications Review, Abril 1999, pp. 16–22
- [5] Artur Ziviani, José F. de Rezende, Otto C. M. B. Duarte - Avaliação de Desempenho de Mecanismos de Diferenciação de Serviços na Internet, SEMISH, SBC, 1999
- [6] One-Way Transmission Time, ITU-T Recommendation G.114, 1993

Análise de Erros da Telefonia IP sobre Ambiente de Computação Móvel

Autores:

Sérgio de Oliveira, Soraia Lúcia da Silva, Paulo Ferreira de Moura Júnior, Antônio Alfredo Loureiro
{sergiool, soraia, mourajr, loureiro}@dcc.ufmg.br

Resumo:

Estamos observando a convergência das redes de voz e dados, unificando o sistema de telecomunicações em rede única, cuja comutação será baseada em pacotes. O fluxo de voz gera um tráfego de tempo real, com restrições para as quais nossas redes não estão preparadas. Muitas pesquisas tem sido desenvolvidas no intuito de adaptar o fluxo de voz para as redes existentes com manutenção da qualidade final ao usuário.

Sobre um ambiente de computação móvel, esta questão se torna ainda mais crítica, devida a alta taxa de erros deste ambiente. Procuramos observar a influência da taxa de erros de um fluxo de voz sobre a qualidade acústica final ao usuário, simulando um ambiente de telefonia sobre redes baseadas no protocolo IP. Foram considerados os principais codificadores de voz utilizados hoje no ambiente de Telefonia IP. Os experimentos foram realizados em redes fixas e em um ambiente móvel, resultando também de comparações de taxas de erros nestes ambientes.

Introdução

A migração das redes de voz, hoje baseadas em redes com comutação baseada em circuito, para redes cuja comutação é baseada em pacotes deve se tornar uma realidade nos próximos anos devido ao desenvolvimento da tecnologia de voz sobre protocolo IP, bem como a introdução de conceitos importantes sobre as redes IP para que elas suportem este serviço.

A convergência destas redes permitirá grandes avanços, em especial a redução do custo de ambos os serviços. A rede de telefonia hoje existente é cara e complexa devido ao modelo em que foi desenvolvida para garantir a comunicação respeitando parâmetros de qualidade rigorosos. Outro problema com esta rede diz respeito à sinalização e gerência. A rede de telefonia requer a existência de duas outras redes, uma para sinalização e outra para gerência, com infra-estruturas próprias para garantir o funcionamento do sistema. Utilizando-se de uma rede de pacotes para o tráfego da voz, pode-se aproveitar esta mesma rede para a completa operação da telefonia, incluindo sobre a mesma infra-estrutura as redes de sinalização e gerência.

Para permitir a utilização de telefonia sobre redes de pacotes devemos garantir que o serviço tenha requisitos mínimos de qualidade que garantam a utilização satisfatória do serviço. Veja que isto não implica necessariamente no atendimento a todos os requisitos existentes sobre a rede convencional, pois são extremamente rigorosos. Os principais requisitos dizem respeito à qualidade acústica da reprodução e atraso imposto pela transmissão.

Para a garantia da qualidade ao nível de aplicação, são exigidos alguns parâmetros de qualidade de serviço da camada de rede, conhecidos por QoS. Os parâmetros necessários para uma comunicação de telefonia sobre rede de pacotes são especialmente *jitter* e *latency*. O protocolo IP por sua vez não garante estes parâmetros, de forma que estão sendo desenvolvidas algumas abordagens para permitir a garantia destes parâmetros através de técnicas de garantia de QoS.

Sobre ambientes de computação móvel, o problema ainda é maior devido às características do meio que impõe uma rede instável, com largura de banda e atraso variáveis, além períodos de desconexão durante o processo de *handover*.

O objetivo deste trabalho é avaliar a influência dos parâmetros de qualidade sobre a taxa de erros e sobre a qualidade final acústica de um sistema de telefonia IP. Isto foi feito sobre redes locais fixas e também em redes móveis, sempre em ambiente onde não existe a garantia de QoS.

Com base nestes resultados, esperamos propor técnicas alternativas para adaptar um sistema de telefonia IP para um ambiente de computação móvel de forma a garantir a qualidade final acústica ao usuário.

Este documento está organizado da seguinte maneira: Na seção 1 foi apresentado o contexto deste artigo bem como seu objetivo. Na seção 2 apresentaremos os conceitos importantes para este trabalho bem como suas definições e informações importantes,

discute-se telefonia IP, qualidade de serviço e ambientes de computação móvel. Na seção 3 apresentaremos o ambiente de teste e a metodologia utilizada. Na seção 4 mostraremos os resultados obtidos e encerraremos com uma conclusão e discussão sobre os trabalhos futuros na seção 5.

Conceitos e Definições

Telefonia IP

O termo Telefonia IP, Telefonia Internet ou ainda Voz sobre IP (VoIP - *Voice over IP*) tem se aplicado a utilização de redes baseadas em protocolo IP na camada de rede para transporte de voz, em especial a Internet. Podemos destacar as vantagens da utilização deste serviço:

- Compartilhamento da rede para o tráfego de voz com o tráfego de dados
- Meio de transmissão de baixo custo, comparado ao sistema telefônico;
- Possibilidade de compactação e supressão de silêncio;
- Possibilidade de oferecer outros serviços adicionais;
- Unificação do transporte, sinalização e gerência sobre a mesma rede;
- Possibilidade de unificação de diversos serviços como email, fax, voz, web, com o auxílio de tecnologias com reconhecimento e síntese de voz;

Um engano sobre Telefonia IP é pensar que seu maior benefício seja para chamadas de longa distância de baixo custo. Enquanto ligações de longa distância baratas estão criando boas áreas de trabalho, as razões pelas quais as companhias são atraídas para a Telefonia Internet são a facilidade de criação de serviços e a consolidação de suas redes.

Um outro engano sobre VoIP é sobre sua qualidade final de transmissão de voz. Se a rede estiver sobrecarregada, pode ser inseridos atrasos que podem atrapalhar a qualidade final da voz. Todavia, se Telefonia IP é usada em uma rede privada dedicada, com garantia de QoS, a qualidade final de voz está bem próxima do nível da telefonia convencional.[2]

Computação Móvel

Podemos enumerar diversas vantagens de um sistema móvel como conforto, para utilização em qualquer ambiente; flexibilidade, para utilização em diversas aplicações que exijam movimento; disponibilidade, independente da localização do usuário; e algumas exigências, como portabilidade, facilitando o transporte; autonomia de energia, para garantir o funcionamento onde não existe disponibilidade de energia; e desempenho comparável a estações fixas.

A mobilidade porém, sempre implica em algumas condições, típicas do ambiente, que devem ser consideradas, independente do sistema de acesso:

- Capacidade de comunicação limitada com largura de banda variável e alta taxa de erros;

- Autonomia de energia limitada por baterias com limite de consumo, de forma que deve-se despendar o mínimo de energia com processamento e dispositivos de apoio ao sistema.
- Limites físicos de hardware para garantia de portabilidade, limitando também o poder de processamento e dispositivos;
- Problemas de roteamento de pacotes quando variação da subrede onde está presente a UM.
- Perda temporária de comunicação quando deslocamento entre áreas mantidas por diferentes estações de rádio e renegociação de características de acesso.

A princípio podemos utilizar o próprio sistema de telefonia celular para fazer a interconexão destes computadores em rede, porém, algumas limitações estão presentes neste tipo de acesso, especialmente no que diz respeito à baixa velocidade de acesso obtida neste sistema. Para estes sistema, a Unidade Móvel é um computador portátil utilizando de um modem e um telefone celular para se conectar à um servidor de acesso remoto. As estações de rádio são as estações do sistemas de telefonia celular e todas as operações de mobilidade são gerenciadas pela camada física deste sistema, provida pelo sistema de telefonia celular. Este sistema não é de nosso interesse, devido à baixa taxa de transmissão obtida, na faixa de 9.6 Kbps.

Outras formas de comunicação podem também ser usadas como *Wavelan* que permite a utilização de taxas de transmissão mais satisfatórias, entre 1 e 2 Mbps. Este sistema deve resolver porém problemas de roteamento, *handoff* que são resolvidos pelo sistema celular na abordagem anterior.

Wavelan foi padronizado pelo comitê IEEE 802.11, definindo:

- Camada física: Com acesso infravermelho, acesso direto à frequência de 2,4 GHz, e acesso por chaveamento de canais, para garantia de segurança de acesso, com 79 canais de 1MHz, também na faixa de 2,4GHz.
- Subcamada de Acesso ao Meio, utilizando o mecanismo de CSMA/CA, *Carrier-sense Medium Access / Collision avoidance*, semelhante ao princípio da Ethernet, com a inclusão de um mecanismo para evitar a ocorrência de colisões.
- Infra-estrutura de rede, em duas configurações: redes com infra-estrutura fixa e redes *ad hoc*

Nosso experimentos sobre redes móveis se deram sobre *Wavelan* padrão IEEE 802.11 de fabricação da *Lucent Technologies* com taxas de transmissão de 2Mbps.

Espera-se que uma unidade móvel seja capaz de acessar a todos os recursos disponíveis em uma estação ligada a uma rede através de cabeamento, contornando todos os problemas deste ambiente. Lança-se mão, então, de técnicas para maior aproveitamento do canal de comunicação, da energia do sistema e do processamento, de forma a limitar as tarefas executadas na UM.

Devemos adaptar a aplicação, o protocolo utilizado ou mesmo a pilha de protocolos utilizados no ambiente de computação móvel para garantir os mesmos parâmetros de QoS que seriam garantidos em uma aplicação semelhante em uma rede fixa.

Estamos avaliando o serviço de telefonia IP sobre redes móveis com a finalidade de propor modificações neste serviço de forma a adaptá-lo ao ambiente de computação móvel.

Qualidade de Serviço

Qualidade de Serviço (QoS) é a capacidade da rede de prover parâmetros específicos de rede para um subconjunto de tráfegos agregados. Para este subconjunto, podemos definir parâmetros a serem garantidos. O protocolo da camada de rede é o responsável pelo controle de QoS.

Podemos lançar mão de alguns parâmetros para medir a Qualidade de Serviço oferecido aos usuários de unidades móveis, vislumbrando as principais aplicações que queremos desenvolver neste ambiente. Podemos destacar entre estes parâmetros:

- *Bandwidth* - Representa a velocidade do meio.
- *Delay* - Atraso médio.
- *Jitter* - Variação do *delay*.
- *Tempo de silêncio durante um handoff*

Uma rede garante QoS quando consegue estabelecer parâmetros mínimos para determinado fluxo de informações que serão sempre conseguidos sob quaisquer circunstâncias da rede. Estes parâmetros podem ser associados por aplicação, de acordo com as exigências de cada serviço.

Qualidade de Serviço para Telefonia IP

Aplicações de telefonia são aplicações intolerantes de tempo real. Para a manutenção de interatividade em uma conversação, assume-se que o retardo fim-a-fim do tráfego de voz deve estar abaixo de um certo valor reduzido. A ITU-T define [6] os valores do retardo fim-a-fim, de acordo com a aplicação:

- 0 a 150 ms - Aceitável para a maioria das aplicações;
- 150 a 400 ms - Aceitável desde que se tenha cuidado com o impacto do atraso sobre a qualidade da aplicação para o usuário.
- Acima de 400 ms - Inaceitável para a maioria das aplicações de voz.

O retardo fim-a-fim não pode contudo ser uma medida para o atraso médio da rede. Isto porque dado as condições da rede, um *jitter* é imposto à comunicação, causando uma variação no intervalo de recepção de pacotes. O atraso imposto à aplicação deve então ser maior que a soma do atraso médio mais o maior *jitter* observado.

Quanto à largura de banda, as aplicações de telefonia IP necessitam de uma largura de banda fixa, em dependência do *vocoder* utilizado no processo de codificação de voz. Estas taxas variam entre 5.6Kbps e 64 Kbps. Pode ainda ser utilizado um algoritmo para supressão de silêncio, capaz de reduzir a taxa de transmissão em mais de 50 %. No nosso experimento não estamos levando em consideração supressão de silêncio.

Características da Voz e da Conversação Telefônica

A fala gera um sinal que originalmente tem um espectro de 7 kHz. Para efeitos de comunicação telefônica ficou definido que a parte do espectro até 3,4 kHz seria suficiente para uma boa conversação. Outro dado importante é a faixa dinâmica do sinal de voz: em torno de 60 dB. Isto exige no mínimo 10 bits para codificação. Para efeitos práticos, estabeleceu-se uma taxa de amostragem de 8 kHz com resolução de 11 ou 12 bits como parâmetros mínimos.

Existem alguns aspectos relacionados com a natureza do sinal de voz que permitem adicionar mecanismos de compressão. São eles:

- Distribuição não uniforme de amplitudes;
- Correlação entre amostras sucessivas;
- Correlação entre ciclos sucessivos;
- Fator de inatividade ou percentual de silêncio;
- Densidade espectral média não uniforme, confirmando a redundância de informações;
- Densidade espectral instantânea, ou presença de formantes que se mantêm inalterados durante 20 a 30 ms;

Compressão de sinais

Para reduzir a banda do canal necessária para a transmissão de voz digitalizada são utilizadas técnicas de compressão de voz. Isto deve acontecer em tempo real para possibilitar a comunicação e interação. Estes algoritmos permitem reduzir a banda necessária para transmissão de voz a fim de viabilizar sistemas de comunicação digital com taxas bastante reduzidas e tendo como objetivo uma qualidade semelhante ao sistema público de telefonia analógica [22].

A compressão de sinais é baseada em técnicas de processamento que retiram informações redundantes, previsíveis ou inúteis. A compressão pode acontecer com ou sem perda de informação. Tudo depende da degradação que se admite para o sinal e do fator de compressão que se deseja atingir.

Os dois principais enfoques com que o tema de codificação digital de voz é atualmente abordado são:

- Codificação da forma de onda (ou não-paramétrica)
- Codificação da fonte (ou paramétrica), baseada em uma modelagem do processo de produção da voz [7].

A codificação da fonte permite uma redução considerável na taxa de transmissão, porém perde-se fatores essenciais como timbre de voz, tornando-a mais impessoal. Dessa forma, sistemas de codificação da fonte são utilizados quando é necessária uma baixa taxa de transmissão.

Os VOCODERS analisam o conteúdo espectral do sinal da fala e identificam os parâmetros que são entendidos pelo ouvido, através de codificação de fonte. Estes parâmetros são transmitidos e usados na recepção para sintetizar o padrão da voz. A forma de onda resultante pode não ser semelhante à original mas as diferenças não são percebidas ou, ainda que o sejam, são consideradas aceitáveis para a aplicação.

Algoritmos de Compressão Padronizados

Vamos descrever os algoritmos de compressão utilizados em nosso experimento. Os resultados são apresentados em função do algoritmo utilizado.

Codificação na forma de onda

- ITU G-711 (PCM logarítmico lei A ou lei μ a 64 kbps): O sinal de voz na faixa de 4 kHz, ou seja, filtrado de 100 a 3400 Hz, é amostrado a 8 kHz com 14 (lei μ) ou 13 bits (lei A) e sofre uma compressão logarítmica para 8 bits. A compressão logarítmica é adequada por causa da natureza do sinal de voz: os sinais de baixa amplitude são mais frequentes e mais importantes.
- ITU G-726 (ADPCM a 40, 32, 24 ou 16 kbps): O sinal de voz é amostrado a 8 kHz, codificado em 8 bits (leis A ou μ) e são transmitidas diferenças entre amostras com 5,4, 3 ou 2 bits em quantização adaptativa. Outros métodos de compressão de áudio ADPCM incluem as Recomendações-padrão G-721 (taxa de 32 kbps) e a G-723 (taxa de 24 kbps); e o algoritmo de compressão de áudio interativo [15,21].

Codificação na fonte

- ITU G-723 (ACELP a 5,3 ou 6,3 kbps): Em cada janela de 30 ms do sinal de voz, são analisadas 240 amostras de 16 bits do sinal de voz (tomadas a 8 kHz) para identificação de pitches e são gerados 12 ou 10 códigos de 16 bits, conforme o algoritmo esteja configurado para uma taxa de 6,3 ou 5,3 kbps. É um VOCODER próprio para videoconferência, telefonia móvel digital e aplicações multimídia.

RTP

RTP (*Real time transport protocol*) é um protocolo padrão para transporte de dados com características de tempo real, como vídeo e áudio, que pode ser usado em serviços tanto como mídia sob demanda quanto interativos, do tipo telefonia IP. É composto por uma parte de transmissão de dados e outra de controle, chamada RTCP (*RTP control*

protocol). A parte de dados consiste de um protocolo leve, que provê suporte para aplicações com características de tempo real, incluindo reconstrução temporal de mensagens, detecção de perdas, segurança, selo de tempo e identificação de conteúdo[1]. RTCP é usado para monitorar a qualidade de serviço da transmissão e para transmitir informações sobre os participantes de dada sessão[2]. Apesar do RTP utilizar serviços do UDP e IP, ele pode ser implementado em outros ambientes, já que necessita somente de serviço não orientado à conexão.

A telefonia IP utiliza sempre do protocolo RTP. Em nosso trabalho tivemos de implementar este protocolo por não ser disponível de forma aberta nos sistemas operacionais. Assim foi possível analisar todas as estatísticas a nível de transporte de tempo real.

Métricas coletadas

Para o nosso experimento as métricas coletadas foram o número de pacotes com erro, pacotes recebidos fora de ordem, pacotes descartados por chegarem "atrasados" e o *jitter* estimado. Todas elas são disponibilizadas pelo RTP: as três primeiras são contadores, incrementados à medida que pacotes nas condições citadas são encontrados; já o *jitter* envolve cálculo. O *jitter* pode ser definido como a variância estatística da diferença (D) no espaçamento do tempo de chegada dos pacotes no receptor comparado com o emissor para cada par de pacotes. A diferença no espaçamento é calculada através da diferença no "tempo de trânsito relativo" de dois pacotes. O tempo de trânsito relativo é a diferença entre o tempo de saída de um dado pacote, registrado através do selo de tempo contido no próprio, e o tempo de chegada no receptor medido na mesma unidade. Se S_i é o selo de tempo do pacote i e R_i o tempo de chegada do pacote i , para dois pacotes (i e j) tal diferença pode ser expressa como:

$$D(i, j) = (R_j - R_i) - (S_j - S_i) = (R_j - S_j) - (R_i - S_i)$$

O *jitter* é calculado continuamente para cada novo pacote usando a diferença no espaçamento para esse pacote e o anterior em ordem de chegada (não necessariamente em sequência) de acordo com a seguinte fórmula apresentada no RFC 1889:

$$J_i = J_{i-1} + (|D(i-1; i)| - J_{i-1}) / 16$$

Segundo Schulzrinne [2], o parâmetro de ganho 16 dá uma boa razão de redução de ruído e mantém uma taxa de convergência razoável.

Ambientes de Testes

Foi desenvolvido um aplicativo capaz de simular um ambiente de telefonia IP para transmissão de arquivos de voz. Assim, simula-se a telefonia para arquivos que são mais facilmente avaliados.

Este aplicativo inclui biblioteca de RTP *Real Time Protocol* capaz de gerar as mesmas condições sobre as quais a telefonia IP acontece sobre as camadas de transporte e de rede.

Os arquivos codificados sobre os *vocoders* especificados em seção anterior foram transmitidos respeitando a taxa de transmissão de cada *vocoder*, o tempo de duração da voz presente nos arquivos e o tamanho dos pacotes gerados por cada *vocoder*.

Registrarmos dados importantes na transmissão destes arquivos, como número de pacotes perdidos, atraso médio imposto pela rede *jitter* em cada pacote. Utilizamos de um método para avaliação da qualidade acústica de cada arquivo transmitidos e comparamos a qualidade final acústica com a taxa de erros, para verificar a influência da taxa de erros sobre os diversos *vocoders*.

Aspectos de tráfego de voz em tempo real

Delay e jitter

Cada pacote responsável pelo transporte de voz é submetido a um atraso variável, de acordo com o tráfego presente na rede e do processamento realizado sobre ele. Se todos os pacotes recebessem o mesmo atraso, este atraso seria propagado integralmente sobre a voz e seria o atraso final da voz.

O problema é a variação deste atraso. Como os pacotes chegam com diferentes atrasos, estes pacotes não podem reproduzir a voz diretamente ao usuário, pois, desta forma a voz sofreria cortes em função de atrasos maiores. A variação do atraso é conhecida como *jitter*.

Quanto maior o *jitter*, maior será o reflexo deste atraso sobre a voz, de forma a incluir cortes e tornar a voz inteligível. Desta forma, é necessário haver um armazenamento dos pacotes durante um certo período de tempo, evitando cortes durante a execução da voz transmitida, suprimindo as variações do *jitter*. Este armazenamento resulta em um atraso inicial da reprodução da voz.

Os valores máximos para este atraso estão definidos como parâmetros de qualidade de serviço para telefonia IP. Vamos chamá-lo de *delay*.

O dimensionamento do *delay* é ponto importante neste trabalho. Uma série de valores deve ser avaliada com vista a obter o menor *delay* inicial de forma a compensar o *jitter*. Podemos ainda prever a adaptação do *delay* de acordo com as condições da rede. O *jitter* deve ser analisado sobre tráfego gerado por diversos *vocoders* sobre ambientes de redes reais, especialmente redes não-determinísticas.

Em nossos trabalhos, utilizamos de armazenamento de pacotes de 25 ms, para supressão do *jitter*, que é acrescido ao atraso final. *Jitter* maior que 25 ms gera perda de pacote.

Conversão de pacotes em fluxo de voz

Os pacotes são transportados sobre uma rede IP através do protocolo de transporte RTP (*Real Time Protocol*). Embora o RTP seja um protocolo da camada de transporte, ele utiliza de *sockets* UDP como mecanismo de transporte.

Cabe à aplicação montar os pacotes recebidos de forma a manter o fluxo de voz para o usuário. Este fluxo de voz deve ser contínuo, independente do recebimento ou não dos pacotes correspondentes.

Para garantir a continuidade da voz quando o pacote correspondente é perdido, seja por atraso, erros ou perda real do pacote, o espaço deve ser preenchido. Como alguns fonemas podem durar um tempo maior que o tempo de cada pacote, a repetição de um pacote pode suprir de forma imperceptível a ausência do pacote anterior. A política utilizada então é de substituição pelo último pacote válido.

O tamanho dos pacotes varia de acordo com os métodos de compressão utilizados. Em geral, um pacote representa uma amostra de voz de 20 ms ou 30 ms. Vale observar que este tamanho varia em função da taxa de transmissão do *vocoder* utilizado.

Também é importante garantir a sequência dos pacotes. Para isto o RTP provê um número básico de sequência que possibilita à aplicação obter informações em relação à ordem dos pacotes. Este número de sequência é então verificado para garantir a integridade do fluxo de voz.

Podemos verificar então a presença de pacotes atrasados, com número de sequência menor que o número esperado, que devem ser descartados, e pacotes adiantados, com número de sequência maior que o esperado. Estes últimos normalmente representam a perda dos pacotes anteriores. Neste caso, devemos preencher o espaço dos pacotes perdidos e atualizar o número de sequência corrente.

Mean Opinion Score (MOS)

O método MOS é derivado do método ACR (*Absolute Category Rating*) para estimar a qualidade de sistemas de transmissão de voz. O ACR requer que algumas pessoas avaliem a qualidade geral dos exemplos de fala para propósitos de comunicação telefônica.

Os avaliadores atribuem notas em escala de 5 categorias, como mostrado no quadro abaixo:

Nota	Significado
1	Ruim: ininteligível, o usuário não entende a mensagem decodificada. Possui interrupções horríveis devido às degradações
2	Pobre: o sinal possui interrupções devido às degradações e o usuário tem que fazer um esforço considerável para entender alguns trechos.
3	Moderado: a qualidade da voz é ruim, o usuário sente incomodado com as degradações, porém não tem interrupções e ainda consegue entender a mensagem. (requer esforço moderado)
4	Bom: a voz é agradável de se ouvir, ou seja, percebe degradações mas não se incomoda com elas, pois são mínimas. (nenhum esforço apreciável é requerido)
5	Excelente: o usuário não consegue diferenciar o trecho original com o corrompido, ou seja, não percebe degradações no sinal. (nenhum esforço é requerido)

A especificação completa e procedimentos para o ACR são descritos em [3]. O objetivo deste teste pessoal é apresentar uma figura da opinião pessoal dos jurados

observando o sinal como transmitido pelo sistema de comunicação ou pelo algoritmo sendo testado. Como normalmente conduzidos, estudos pessoais de sistemas de comunicação requerem que os avaliadores escutem a fala processada por um período de tempo definido.

Redes Utilizadas

Os testes se deram em três ambientes distintos de rede:

- Rede local ethernet segmentada por *switch* com baixo tráfego;
- Rede local ethernet do departamento, com alto tráfego, através de três subredes distintas;
- Rede local *wavelan* baseada nos produtos da *Lucent* a 2 Mbps, padrão 802.11, sem tráfego local;

Resultados Obtidos

Duas amostras de voz, de prefixos "rosa" e "so", foram criadas gravadas, uma com voz masculina e outra feminina e utilizadas em nosso experimento utilizando os três ambientes de rede mencionados e os seguintes *vocoders*, já descritos anteriormente:

- G726 a 24, 32 e 40 Kbps
- G723 a 5,3 e 6,3 Kbps
- G711 a 64kbps;

Os *vocoders* acrescentaram um sufixo ao arquivo, de acordo com a taxa de transmissão do *vocoders* utilizado.

Devemos ainda observar um último sufixo "10" ou "20". Este sufixo não traz nenhuma informação importante, apenas está presente para diferenciar os arquivos idênticos que foram transmitidos simultaneamente simulando a capacidade *full duplex* da telefonia IP, que ocupa dois canais caso não seja utilizada a supressão de silêncio. Foram avaliados ambos os sentidos da transmissão, embora não haja diferenciações dos resultados.

Segue o sumário dos resultados, com os seguintes dados:

- Coluna arquivos especifica o arquivo transmitido, e *vocoder* utilizado;
- Cabeçalho "Rede do Laboratório", "Rede do Departamento" e "Rede Sem Fio", representam respectivamente uma rede local de baixo tráfego, rede local de alto tráfego e rede local sem fio;
- Os dados obtidos são:
 - Ti - Número de pacotes perdidos por *timeout*, ou seja não chegaram a tempo de serem reproduzidos;
 - Ma - *Jitter* máximo da transmissão, conforme descrito anteriormente;

- DP - Desvio Padrão do *jitter* durante a transmissão;
- Me - Média do *jitter* durante a transmissão;
- N - Médias das notas atribuídas por 8 avaliadores pela técnica MOS.

	Rede do Laboratório					Rede do Departamento					Rede Sem Fio				
Estatísticas Arquivos	Ti	Ma	DP	Me	N	Ti	Ma	DP	Me	N	Ti	Ma	DP	Me	N
Rosa2410	0	49	6,68	9,43	4.5	0	170	17,3	16,3	4.4	0	777	83,5	27,2	4.4
Rosa2420	0	23	1,35	7,28	4.3	2	253	16,2	4,3	4.3	3	721	71,3	23,4	4.1
Rosa3210	0	47	4,8	8,17	4.8	24	1328	232	210	2.8	254	2325	225	915	1.3
Rosa3220	0	19	1,08	7,05	4.6	0	318	50,5	39,1	4.8	0	1167	126,	122	4.3
Rosa4010	0	14	0,97	7,22	4.9	273	116	17,7	22,9	1.5	3	338	44,1	81,9	4.4
Rosa4020	0	45	5,27	8,54	5.0	0	164	18,0	20,3	5.0	0	46	6,8	16,4	4.9
Rosa5310	382	3004	342	2586	1.4	6	2329	198	1336	4.1	0	1647	154	1302	4.0
Rosa5320	0	1404	154	1285	3.3	83	2278	269	1632,	2.9	156	2538	313	206	2.0
Rosa6310	0	1391	145	1287	4.1	9	1941	169	1371	4.3	0	1441	147	1304	4.3
Rosa6320	0	1397	145	1290	4.3	33	2208	220	1462	3.4	1	1640	156	1324	2.5
Rosa6410	0	45	5,25	8,45	4.9	12	775	137	219	3.9	4	521	86,6	37,4	5.0
Rosa6420	0	17	1,53	7,0	5.0	6	1089	126	55,1	4.1	0	656	94,3	40,3	5.0
So2410	62	1320	387	163	3.3	8	547	87	76,3	4.1	0	737	120	239	4.9
So2420	0	99	8,84	9,74	4.6	27	2888	345	115	3.0	2	597	62,7	37,3	4.4
So3210	0	82	7,25	8,80	5.0	8	911	125	118	3.3	11	760	120	95,4	3.4
So3220	0	50	4,40	8,17	5.0	0	322	36,7	24,1	4.3	61	1379	307	287	2.3
So4010	0	14	1,19	6,98	2.4	63	1717	255	226	2.3	10	967	116	106	1.9
So4020	0	20	1,32	7,03	2.0	72	2060	341	227	2.1	1	293	28,7	14,1	2.1
So5310	0	1397	164	1278	3.6	347	3107	381	2618	1.0	82	3116	616	1890	3.3
So5320	0	1406	172	1278	3.8	7	1916	184	1354	3.3	0	1398	131	1300	3.8
So6310	343	2960	388	2409	1.0	319	2676	339	2219	1.0	151	2595	337	2052	2.6
So6320	296	4009	463	3109	1.6	536	1858	1912	1440	1.0	1	1640	156	1324	3.9
So6410	0	44	5,14	8,38	5.0	3	622	96,8	71,3	4.4	47	3003	512	404	4.5
So6420	0	309	28,3	13,8	4.9	46	737	120	239	4.1	0	2577	388	222	5.0

Podemos observar a partir dos resultados:

- Os *vocoders* de taxa de transmissão mais baixa são mais suscetíveis a erros. Poucos pacotes perdidos representam uma grande perda de qualidade.
- Dois fatores são responsáveis por queda de qualidade: perda de pacotes por *timeout* e *jitter*. A perda de pacotes é esperada, visto que acaba por adulterar o arquivo. Em relação ao *jitter*, isto também é verdade, uma vez que a variação do atraso ocasiona a necessidade de descarte de pacotes que não chegam no momento certo.
- Quase todas as notas acima de 4 dizem respeito a arquivos que não tiveram erros de transmissão. Esta queda de qualidade é ocasionada pelo processo de codificação de decodificação da voz que é amenizado nos arquivos de maior taxa de transmissão (G711).
- Todos os ambientes de rede são não-determinísticos e sujeitos a rajadas de alto tráfego alto que comprometem a qualidade da telefonia, por não haver garantia de qualidade de serviço. Isto é observado nos arquivos de baixa qualidade, notas 1 a 3.

Conclusões e Trabalhos Futuros

A maior contribuição deste trabalho é a criação de um ambiente de teste e um metodologia para avaliação de telefonia IP.

Estamos partindo para a adaptação da telefonia IP para ambientes móveis de forma a eliminar as causas de baixa qualidade acústica final ao usuário e repetiremos nosso processo de avaliação diversas vezes sobre as propostas de adaptação com a finalidade de observar seu efeito sobre a qualidade final.

O objetivo final é permitir ainda a utilização de um terminal de telefonia IP em um elemento compacto, com restrições de desempenho impostas pelo tamanho do equipamento, de forma que a análise do codificador de voz utilizado é também essencial.

Bibliografia consultada

1. Schulzrinne, H., "RTP Overview", <http://www.cs.columbia.edu/~hgs/rtp/>
2. Schulzrinne, H., Casner, S., Frederick, R. e Jacobson, V., "RTP: A transport protocol for real-time applications", RFC 1889, Internet Engineering Task Force, Fevereiro 1996.
3. Telephone Transmission Quality Subjective Opinion Tests - Methods for Subjective Determination of Transmission Quality - International Telecommunication Union Recommendations

Implementação

Decisões de Implementação

Este projeto tem por objetivo implementar um sistema compacto que permita a utilização de recursos de Telefonia IP. Para isto, os seguintes princípios devem ser seguidos:

- Um agente de Telefonia IP deve executar na plataforma. A preferência é sobre utilizar um agente com código fonte aberto, para que seja possível incluir modificações para adaptar a aplicação e avaliar resultados. Este tipo de agente somente está disponível para sistemas Linux ou Windows;
- A plataforma deve ter recursos de rede móvel para atender aos princípios de mobilidade do sistema. Para aplicações deste tipo, o padrão 802.11 tem se destacado como solução de rede compatível com a Internet;
- Para sistemas compactos, o padrão PC104 tem se destacado. Um sistema com processador 486, 16 MB de memória RAM e 16 MB de memória FLASH pode atender às necessidades básicas de um sistema Linux preparado para esta função desde que não sejam executadas nesta CPU funções de compressão de voz.
- Suporte a PCMCIA é requerido para utilização de *Wavelan*, a placa de rede padrão 802.11
- Utilização de um DSP para realização de funções de compressão de voz.

Para atender a estes requisitos, escolhemos uma série de componentes para o nosso sistema, os quais iremos descrever a seguir.

Base Móvel

Linux

O Linux é um sistema operacional da família Unix, desenvolvido sobre a plataforma de aplicativos GNU[Referência], utilizando o núcleo desenvolvido por Linus Torvalds com ajuda de programadores em todo o mundo, disponibilizado sobre a forma de código aberto.

Linux é uma implementação POSIX independente, que inclui núcleo multitarefa, memória virtual, compartilhamento de bibliotecas com carga dinâmica, propriedades de gerência de memória, inclui protocolos de rede TCP/IP e outras características consistentes com sistemas Unix.[www.linux.org]

Após o desenvolvimento do núcleo por parte de Linus Torvalds, ele encontrou no Projeto GNU todos os aplicativos e bibliotecas que ele necessitava para completar o sistema. Assim foi desenvolvido um sistema livre, baseado na licença de uso de software conhecida como GNU *General Public License*[<http://www.linux.org/info/gnu.html>].

Nossa aplicação utiliza o kernel Linux na versão 2.2.13 [<ftp://sunsite.unc.edu/pub/linux/kernel/v2.2/linux-2.2.13.tar.gz>] além de um conjunto de aplicativos distribuídos na configuração conhecida por *Coyote*, especial para sistemas embutidos.

Sistema de arquivos em Memória RAM

O núcleo do Linux, como foi concebido, exige a presença de um sistema de arquivos raiz (*root filesystem*) para execução do sistema. Isto poderia ser impeditivo para utilização em sistemas embutidos que não utilizam dispositivos de armazenamento de massa, como disco rígido.

Especialmente na base móvel, onde o único dispositivo de armazenamento é uma memória FLASH, o uso de um sistema de arquivos neste dispositivo poderia ser impeditivo devido ao tempo longo de escrita. Como alternativa, o Linux permite a utilização de uma determinada região de memória RAM como sistema de arquivos raiz. Desta forma, todos os arquivos necessários ao funcionamento do sistema são mantidos nesta região de memória, garantindo a ausência de um dispositivo de armazenamento de massa e ainda verificando uma alta velocidade de acesso a estes arquivos.

A desvantagem desta abordagem é que a região de memória reservada para esta função não pode ser usada para outras tarefas do sistema. Como a base móvel tem apenas 16 MB (Megabytes) de memória RAM, e vamos necessitar 4 MB desta memória para manter todos os arquivos do sistema, vamos ter disponíveis apenas 12 MB para o processamento do sistema.

Além disto devemos observar que o sistema não utiliza de memória virtual, por não ter um dispositivo capaz de realizar esta tarefa.

Além de utilizar desta região de memória para funcionamento normal do sistema, o Linux permite ainda utilizar a memória RAM para a inicialização, sendo carregado para a memória, antes mesmo que o núcleo, pelo *boot loader* do Linux. Ele é carregado a partir de um pacote compactado especialmente criado para esta tarefa que contém todos os arquivos e *scripts* necessários para a inicialização do sistema. Podemos descrever então os passos na inicialização do sistema:

1. O *boot loader* é o primeiro a ser executado, chamado diretamente no processo de inicialização, por estar armazenado no início do dispositivo de armazenamento;
2. O *boot loader* carrega para a memória o pacote compactado que resultará no sistema de arquivos raiz, e logo em seguida o núcleo do sistema, também armazenado de forma compactada.
3. O núcleo é descompactado ainda pelo *boot loader* e recebe então o controle de processador.
4. O núcleo faz todas as inicializações necessárias, como reconhecimento dos dispositivos presentes no sistema, inicialização dos drivers e preparação da região de memória a ser utilizada como sistema de arquivos.

5. Neste momento o arquivo que contém o sistema de arquivos raiz que já está presente na memória é descompactado para a região preparada para isto. Só a partir deste ponto, é que o Linux está preparado para ler a partir da FLASH.
6. Em seguida são executados uma série de *scripts* que podem adicionar outros arquivos ao sistema, a partir de outros pacotes carregados então a partir da FLASH.
7. São inicializados todos os serviços disponíveis no sistema: servidores, gerenciadores de memória, gerenciadores de alarmes, gerenciadores de dispositivos, etc.
8. O sistema está então em execução pronto para disparar a aplicação de Telefonia IP.

Distribuição Coyote

Coyote Linux [www.coyotelinux.com] é uma distribuição do Linux desenvolvida para sistemas embutidos, capaz de ser armazenado em apenas um disquete com capacidade de 1,44 MB. A distribuição utiliza dos conceitos de sistema de arquivos em memória RAM para eliminar a exigência de disco rígido no sistema.

Vários pacotes são usados para armazenar arquivos que serão mantidos no sistema de arquivos raiz. Estes pacotes são descompactados e extraídos após a extração do pacote principal do sistema de arquivos de memória. Neste momento, o sistema é capaz de reconhecer o dispositivo de armazenamento que contém estes pacotes, que representa, neste sistema a memória FLASH.

Este princípio de funcionamento foi utilizado pela primeira vez no projeto *Linux Router*, e aprimorado pela distribuição *Coyote*. As melhorias dizem respeito à versão do kernel utilizada (2.2.13), e método de configuração.

Estes dois projetos porém, consideram a sua utilização para a construção de roteadores. Desta forma, uma terceira distribuição, baseada na distribuição *Coyote* pode ser considerada como desenvolvida, neste trabalho, para suportar algumas características não previstas na distribuição original. Entre elas:

- Suporte a dispositivos PCMCIA, em especial *Wavelan* e *Internet PhoneCard*;
- Configuração do sistema para apenas um dispositivo de rede, *Wavelan*;
- Limpeza dos pacotes desnecessários ao sistema que estão presentes na distribuição padrão;

A modificação da distribuição se dá sobre os pacotes que a compõe, que são armazenados na mídia de armazenamento, agrupadas no formato *tar* e compactadas no formato *gzip*. Para modificar um pacote, basta decompactar e desagrupar (`tar xfvz nome_do_pacote.lrp`), atualizar os arquivos do pacote, prevendo seu comportamento dentro do sistema resultante.

Podemos incluir na nossa distribuição, pacotes para fins especiais. Neste projeto, estamos incluindo dois pacotes adicionais:

1. pcmcia : Pacote que inclui o suporte a dispositivos PCMCIA, ou *PC Card*, como estão sendo chamados, incluindo os drivers utilizados na nossa aplicação, para a *wavelan* e *quicknet*;
2. iptel: Pacote que inclui todas as bibliotecas e aplicações para Telefonia IP, incluindo o terminal modificado.

Boot Loader

O *boot loader* é responsável pela carga do sistema operacional, juntamente com todo sistema necessário para partida. Para o Linux, são comuns o LILO[13], com certeza o mais utilizado, e o syslinux[], alternativa utilizada pelo *Coyote* para a carga do sistema.

As funções do *boot loader* dizem respeito à leitura do sistema do dispositivo de armazenamento, seja ele disco, memória FLASH, ou outro; carga na memória, incluindo possíveis operações de descompactação, passagem de parâmetros, etc; até que o sistema possa assumir o controle.

O **LILO** (*Linux Loader*) utiliza do princípio de manutenção de uma tabela de blocos de armazenamento do núcleo do sistema, a ser carregado na inicialização. Através dos endereços físicos dos blocos ele é capaz de carregar estes blocos para a memória, descompactar o núcleo do sistema, carregar um pacote para utilização como sistema de arquivos raiz na memória, passar todos os parâmetros necessários para o núcleo e acionar sua execução.

Já o **syslinux** carrega um driver para sistema de arquivos FAT, localiza neste sistema o núcleo do linux, faz sua carga, descompacta e executa, executando possivelmente também a carga de um arquivo que pode vir a se transformar no sistema de arquivos raiz. Este é mais simples e facilmente configurável. Por executar sistema de arquivos FAT, pode ser configurado a partir de um sistema como o MS-DOS que seja capaz de reconhecer o sistema de arquivos onde está instalado o syslinux.

O Coyote utiliza o syslinux como *boot loader*. Todos os pacotes do sistema são mantidos em uma partição com formatação FAT, que pode ser criada inclusive utilizando um sistema MS-DOS, sendo assim facilmente instalado e configurado. O arquivo de configuração *syslinux.cfg* inclui os parâmetros a serem utilizados, como:

- Arquivo que contém o núcleo do Linux;
- Utilização de Memória RAM para sistema de arquivos raiz e montante reservado para isto;
- Arquivo que contém o pacote para o sistema de arquivos e tipo de sistema de arquivos utilizado em RAM, normalmente MINIX[14];
- Lista de pacotes a serem adicionados ao sistema de arquivos raiz após a inicialização do sistema, incluindo pacotes de configuração e nossos pacotes;
- Dispositivo de armazenamento a partir do qual serão lidos estes pacotes e sua formatação;

Implementação de Vocoder

Para implementação do *vocoder*, é necessário eleger uma placa que cumpra esta função e que possibilite uma interface fácil ao sistema. Foi escolhido uma placa da Quicknet[17], conhecida como *Internet PhoneCard*, que tem as seguintes características:

- Interfaces de microfone e fones-de-ouvido;
- Compressão de voz nos padrões de Telefonia IP, G.723.1, G.711, G.729a e *TrueSpeech*;
- Interface através de barramento PCMCIA tipo II;
- Disponibilidade de drivers para Linux, facilitando a integração desta placa junto ao sistema;

Bibliotecas para Implementação de Telefonia IP

Estão disponíveis algumas implementações de pilhas de protocolos que podem ser usadas para a construção de agentes de Telefonia IP, bem como os próprios agentes em algumas implementações. As fontes destas implementações são: consórcio OpenH323[8] e a empresa Vovida Networks[9].

OpenH323

O Projeto OpenH323 tem como objetivo desenvolver um código aberto, com todas as características e interoperável, da pilha de protocolos descritos pelo padrão H.323 com licença de uso público, do tipo MPL (Mozilla Public License)[15].

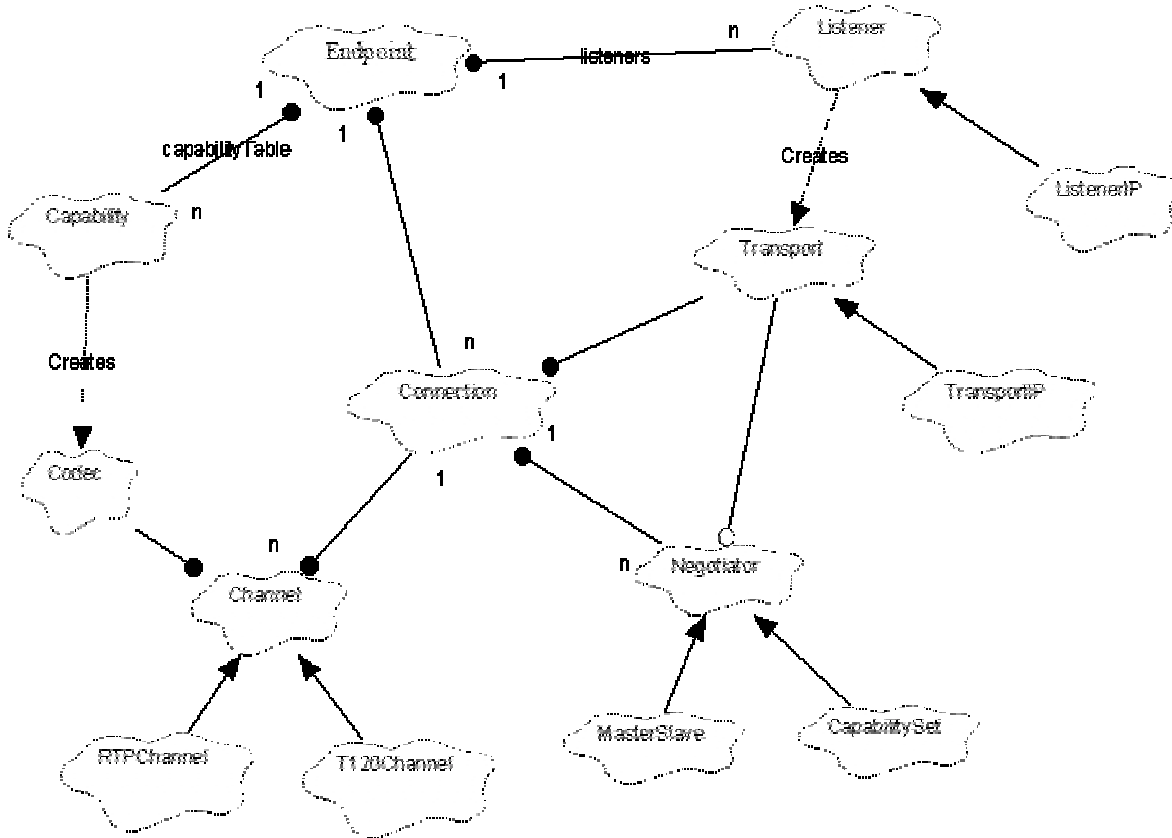
Este projeto é coordenado pela empresa australiana [Equivalence Pty Ltd](#) [16], e conta com a participação de diversos interessados pelo mundo.

A implementação utiliza uma biblioteca de classes C++, conhecida como PWLib, portátil para Windows, Linux, FreeBSD e outros sistemas Unix, de forma a tornar o código H.323 também portátil para estas plataformas. PWLib é uma biblioteca, também desenvolvida pela [Equivalence Pty Ltd](#) multi-plataforma, cujo objetivo é encapsular as diferenças entre os diversos sistemas em classes específicas que possam garantir compatibilidade entre eles. Por exemplo, características *multithreads* são encapsuladas em uma classe específica que pode assim esconder os detalhes de implementação para cada sistema. Desta forma, a linguagem C++ ganha uma característica de portabilidade muito interessante.

Além da pilha de protocolos, está disponível também um agente, implementado com base nesta pilha. Este agente é completo em termos de funcionalidades para uma aplicação de Telefonia IP, interoperável com aplicações como NetMeeting e Conference. Esta aplicação é conhecida como **voxilla**.

A aplicação voxilla suporta as placas de DSPs para tarefas de *vocoder* para a aplicação, fabricadas pela QuickNet[17], ou pode executar estas tarefas na própria CPU do sistema, utilizando dos algoritmos de *vocoder* G.711 e GSM. Com as placas Quicknet, é capaz de implementar os *vocoders* G.723.1, G.729.

A figura abaixo mostra o diagrama de classes da biblioteca OpenH323.



A classe base desta hierarquia é a classe *H323EndPoint*. Uma aplicação deve ter uma instância de uma classe descendente desta, definindo os principais parâmetros H.323, como tabela de capacidades, que inclui os *vocoders* suportados e tipos de canais que a aplicação suporta, como áudio, vídeo e dados.

A aplicação também deve criar instâncias de uma ou mais classes descendentes de *H323Listener*. Existe uma classe descendente desta para cada protocolo suportado. Por exemplo, *H323ListenerIP* deve ser usado para a Internet. Cada uma destas classes dispara uma *thread* que monitora o protocolo e, quando recebe uma nova chamada, cria uma instância de uma classe *H323Transport*. Assim como a classe *H323Listener*, existe uma classe descendente para cada protocolo suportado, como *H323TransportIP*.

O primeiro pacote recebido pelo *H323Transport* usando os protocolos Q.931 e H.225 sinaliza a presença de uma conexão. A classe *H323Connection* encapsula todos os parâmetros desta conexão entre dois agentes H.323. A instância da classe *H323EndPoint* mantém informações sobre as conexões ativas. Se nenhuma conexão entre os dois agentes existe, é criada uma nova instância da classe *H323Connection* e as negociações de sinalização se iniciam.

Uma aplicação normalmente cria uma classe descendente da classe *H323Connection*. Esta classe tem métodos virtuais que devem ser sobrepostos. Alguns

destes métodos são chamados pela própria pilha de protocolos para acionar métodos desta classe.

Implementação Vovida Networks

Como alternativa ao OpenH323, é possível utilizar das pilhas de protocolos desenvolvidas pela Vovida Networks [9]. Além de implementar a pilha de protocolos requerida pelo padrão H.323, também estão disponíveis os protocolos IETF para telefonia IP, incluindo SIP, SDP e RTP.

Embora os protocolos disponíveis ofereçam inclusive a opção do uso de SIP, a análise do código e da aplicação disponível para teste revela que a documentação é fraca e código perde em nível de clareza e maturidade para o projeto OpenH323, motivo pelo qual fizemos a escolha de trabalhar com o código deste projeto.

Referências e Bibliografia

BookMark:

[Entendendo o atraso da Telefonia IP](#)

[IP Telephony - Requirements and Possibilities](#)

[Quality Of Service Testing In The VoIP Environment: A Primer](#)

[Internet Telephony Gateways](#)

[Beyond Dial Tone: Opportunities for Value in IP Telephony](#)

[Voice Over IP Networks](#)

Referências:

- [1] Linden deCarmo, *Internet Telephony Protocols*, Dr. Dobb's Journal, Julho 1999
- [2] Telephone Transmission Quality Subjective Opinion Tests - Methods for Subjective Determination of Transmission Quality - International Telecommunication Union Recommendations
- [3] *Visual telephone systems and equipment for local area networks which provide a non-guaranteed quality of service*, ITU-T Recommendation H.323 (1996)
- [4] M. Handley, SIP: Session Initiation Protocol, IETF, RFC 2543, Março 1999
- [5] RTP
- [6] H.225
- [7] H.245
- [8] <http://www.openh323.org>
- [9] <http://www.vovida.com>
- [13] Lilo mini-Howto - Cameron Spitzer (cls@truffula.sj.ca.us), Alessandro Rubin (rubini@linux.it)
- [14] Sistemas Operacionais: Conceitos e Implementação - Andrew Tanenbaum
- [15] [MPL \(Mozilla Public license\) - http://www.mozilla.org/NPL/MPL-1.0.html](http://www.mozilla.org/NPL/MPL-1.0.html)
- [16] <http://www.equival.com/>
- [17] <http://www.quicknet.net/>

- [4] José Furst Jr., *Telefonia IP, O Novo Mundo das Telecomunicações*, Cisco Systems, Inc. 1999
- [9] Henning Schulzrinne - The IETF Internet Telephony Architecture and Protocols - available at <http://computer.org/internet/telephony/w3schrosen.htm>, 1999
- [10] H. Schulzrinne and J. Rosenberg, "SIP Call Control Services," Internet draft, IETF, Feb. 1998; work in progress.
- [11] R. Fielding et al., "Hypertext Transfer Protocol -- HTTP/1.1," RFC 2068, IETF, Jan. 1997; available at <http://www.ietf.org/rfc/rfc2068.txt>.
- [12] H. Schulzrinne and J. Rosenberg, "Signaling for Internet Telephony," Tech. Report CU-CS-005-98, Columbia Univ., New York, N.Y., Feb. 1998.
- [13] H. Schulzrinne and J. Rosenberg, "Signaling for Internet Telephony," Int'l Conf. Network Protocols, Austin, Tex., Oct. 1998.
- [15] H. Schulzrinne, "RTP Profile for Audio and Video Conferences with Minimal Control," RFC 1890, IETF, Jan. 1996; available at <http://www.ietf.org/rfc/rfc1890.txt>.
- [16] R. Ramjee et al., "Adaptive Playout Mechanisms for Packetized Audio Applications in Wide-Area Networks," *Proc. IEEE INFOCOM*, Computer Society Press, Los Alamitos, Calif., June 1994, pp. 680-688.
- [17] W.A. Montgomery, "Techniques for Packet Voice Synchronization," *IEEE JSAC*, Vol. SAC-1, Dec. 1983, pp. 1,022-1,028.
- [18] S.B. Moon, J. Kurose, and D. Towsley, "Packet Audio Playout Delay Adjustment: Performance Bounds and Algorithms," *ACM/Springer Multimedia Systems*, Vol. 5, Jan. 1998, pp. 17-28.
- [19] J.-C. Bolot and A.V. Garcia, "Control Mechanisms for Packet Audio in the Internet," *Proc. IEEE INFOCOM*, San Francisco, Calif., Mar. 1996.
- [20] I. Busse, B. Deffner, and H. Schulzrinne, "Dynamic QoS Control of Multimedia Applications Based on RTP," *Computer Comm.*, Vol. 19, Jan. 1996, pp. 49-58.
- [21] C. Perkins and O. Hodson, "Options for Repair of Streaming Media," RFC 2354, IETF, June 1998; available at <http://www.ietf.org/rfc/rfc2354.txt>.
- [26] J. Veizades et al., "Service Location Protocol," RFC 2165, IETF, June 1997; available at <http://www.ietf.org/rfc/rfc2165.txt>.