

VOZ SOBRE IP: UMA VISÃO GERAL

Nelson Luiz Leal Fernandes

RESUMO

O uso de voz sobre IP (VoIP) vem sendo uma das grandes metas de investimentos por fornecedores de soluções e usuários de telecomunicações nos últimos anos. Esta tecnologia abre um novo horizonte para as possíveis aplicações integrando-se voz e dados num mesmo equipamento terminal de usuário, aproximando pessoas geograficamente distantes, aumentando a interatividade de aplicativos e diminuindo os custos de comunicação quando comparada às convencionais ligações telefônicas interurbanas.

Serão abordados aspectos técnicos desta tecnologia, como: a evolução das diversas técnicas de digitalização de voz empregadas ao longo do tempo; os pontos relevantes da rede de transporte para o tráfego de voz digitalizada; as duas principais propostas de sinalização na rede (H.323 e SIP); a principal forma de sinalização para reserva de recursos da rede com compromisso de qualidade na comunicação (RSVP); e alguns dos protocolos e implementações necessários para o transporte de voz sobre IP.

1. INTRODUÇÃO

A cada década temos presenciado mudanças significativas no serviço telefônico. A década de 50 foi marcada pela possibilidade de chamadas internacionais diretas, através da introdução dos cabos transatlânticos; nos anos 60, as centrais e transmissões digitais melhoraram em muito a qualidade do sinal de áudio; nos anos 70, as centrais programáveis viabilizaram serviços como as chamadas em espera e discagem por tons; na década de 80, o sistema de sinalização em canal comum *out-of-band* (como o *Signaling System 7 – SS7*) possibilitou serviços como os números 0800. Por fim, a década de 90, marca definitivamente a trajetória da transmissão e sinalização telefônica analógica, rumo a uma infra-estrutura baseada em redes de pacotes. Contudo, os primeiros artigos sobre o assunto datam do início de 70, com a primeira experiência de transmissão de pacotes IP com áudio, entre a *University of Southern California* e o *Massachusetts Institute of Technology*, em agosto de 1974. A primeira RFC (*Request for Comments*) sobre pacotes de voz, RFC 741, foi publicada em 1977.

É certo que a transição da infra-estrutura telefônica atual para um ambiente puro de voz sobre redes de pacotes será longa, haja avisto o enorme investimento feito para implantação do legado. Estamos então com a oportunidade de repensar toda a estrutura de serviço de telefonia, frente aos ambientes e serviços de redes disponíveis, e as novas aplicações viáveis com esta tecnologia.

2. VANTAGENS & DESVANTAGENS

Independente da tecnologia adotada, as corporações já vêm num movimento de integração entre voz e dados na mesma infra-estrutura de rede há alguns anos. As

vantagens são claras, pois os custos envolvidos na manutenção de equipes técnicas, infra-estruturas diferenciadas e ligações internacionais são reduzidos com a integração. Os aumentos do leque de novas aplicações, da disseminação de microcomputadores pessoais (para funcionamento como terminal multimídia), da capilaridade das redes IP e da banda de transmissão disponível para o usuário, contribuíram em muito para voz sobre IP (VoIP) tornar-se uma realidade.

Apesar das primeiras redes com o protocolo IP existirem desde o início da década de 70, na década de 90 houve uma verdadeira explosão no número de redes IP, em nível mundial, consolidado de vez este protocolo com um “padrão de *facto*” para interconexão de equipamentos e redes de comunicação de dados.

Contudo, a diferença de preço entre o terminal telefônico convencional e um equipamento para uso de VoIP (seja um equipamento próprio ou um microcomputador) ainda é um forte fator limitante para uso desta última solução em larga escala. Além disso, a alta disponibilidade da redes telefônicas convencionais aliada à falta de garantia de qualidade de serviço originalmente herdada do IP, são aspectos de peso na comparação entre os ambientes legado e VoIP.

3. ALTERAÇÃO DO CENÁRIO ATUAL

Não se pode precisar exatamente quando, mas é esperada a mudança do cenário de comunicação de voz e dados atual para uma realidade integrada em larga escala, onde os meios de transmissão deverão servir aos dois “mundos”, de forma transparente ao usuário, mas competente no tocante ao consumo de recursos. Neste “novo mundo”, que tecnicamente já é real, o mesmo ponto de rede local que serve a um microcomputador equipado com a capacidade de comunicação de voz, pode ser usado para ligação de um telefone IP.

Enquanto não se tenha a completa substituição do legado de telefonia, independente de estarmos tratando de um longo período de transição obrigatório, dado o volume de equipamentos envolvidos no mundo inteiro, ou por simples opção para maturação tecnológica de novas propostas, é possível a comunicação entre terminais de VoIP e aparelhos da rede telefônica convencional, RDSI (Rede Digital de Serviços Integrados) ou de telefonia móvel, através de *gateways* apropriados e já disponíveis no mercado.

4. CODIFICAÇÕES DO SINAL DE VOZ

A digitalização de um sinal de voz, permite que seu armazenamento e transmissão seja feito de forma mais eficiente. A primeira codificação digital da fala data de 1928 (por Homer Dudley), mas apenas na década de 70 teve uso fora da área militar.

As redes telefônicas adotaram inicialmente a técnica de codificação PCM (*Pulse Code Modulation* – Modulação por Codificação de Pulsos), que consiste em 8.000 amostragens do sinal de voz contínuo, por segundo, representando o valor discreto amostrado em 8 bits. Isto implica na necessidade de um canal digital de 64Kbps para transmissão de cada canal de voz. Este tipo de codificação procura reproduzir o

sinal amostra por amostra. Possui baixo atraso para o processo e pequena complexidade, mas requer um taxa de transmissão elevada.

Ao longo dos anos, novas técnicas de codificação foram desenvolvidas, explorando-se os modelos de produção da voz. Estas técnicas fazem a segmentação do sinal analógico em intervalos periódicos, para formação de quadros após a digitalização. Os quadros são compostos por informações do sinal de voz deste período, mais as de uma parcela do quadro subsequente. O tempo necessário para coletar as informações do próximo quadro, chamamos de *lookahead*. A taxa requerida por esta técnica de codificação é baixa, mas o atraso e a complexidade são elevados, em comparação com a técnica descrita no parágrafo anterior.

O ITU-T (*Telecommunication Standardization Sector of International Telecommunication Union*) padronizou várias codificações ao longo dos anos. A tabela 4.1 enumera as principais, contendo informações relevantes de cada uma delas.

| Recomendação | Codificação | Taxa (Kbps) | Quadro / <i>lookahead</i> (ms) | Ano |
|--------------|-------------|----------------|--------------------------------|------|
| G.711 | PCM | 64 | 0,125 / 0 | 1972 |
| G.726 | ADPCM | 40, 32, 24, 16 | 0,125 / 0 | 1990 |
| G.728 | LD-CELP | 16 | 0,625 / 0 | 1992 |
| G.729 | CS-ACELP | 8 | 10 / 5 | 1996 |
| G.723.1 | MP-MLQ | 6,3 | 30 / 7,5 | 1996 |
| G.723.1 | ACELP | 5,3 | 30 / 7,5 | 1996 |

Tabela 1 – Alguns padrões de codificação do sinal de voz (identificação da Recomendação ITU-T, tipo de codificação usada, taxa de transmissão requerida, tempos mínimos necessários para formação de um quadro e de *lookahead* e ano de publicação da Recomendação). As siglas das codificações significam: PCM - *Pulse Code Modulation*, ADPCM - *Adaptive Differential Pulse Code Modulation*, LD-CELP - *Low-Delay Code Excited Linear Prediction*, CS-ACELP - *Conjugate-Structure Algebraic-Code-Excited Linear-Prediction*, MP-MLQ - *Multipulse Maximum Likelihood Quantization* e ACELP - *Algebraic-Code-Excited Linear-Prediction*.

O processo de transmissão de voz através de uma rede digital envolve alguns atrasos, caracterizados a seguir:

- **atraso de algoritmo** – gerado pelo módulo de codificação do sinal, é causado pelo tempo necessário para a formação de um quadro, mais o tempo de *lookahead*;
- **atraso total de codec** – refere-se à soma do atraso de algoritmo ao tempo de processamento necessário para codificar o sinal na origem e decodificar no destino (ambos somados). É claro que o tempo para processamento depende do potencial do equipamento usado, mas é próximo ao tempo de codificação;
- **atraso total de sistema** – refere-se à soma do atraso total de codec aos tempos necessários à serialização e propagação do sinal digital, da origem ao destino, e ao tempo despendido para enfileiramento dos quadros (*buffer* para amenizar o efeito do *jitter*).

| | G.723.1 | G.729 |
|----------------------|----------------|--------------|
| Taxa (Kbps) | 5,3 e 6,3 | 8 |
| Atrasos quadro | 30 ms | 10 ms |
| <i>lookahead</i> | 7,5 ms | 5 ms |
| processamento | 30 ms | 10 ms |
| Total | 67,5 ms | 25 ms |
| MIPS | 16 / 14,6 | 10,5 |
| RAM (<i>words</i>) | 2.200 | 2.000 |

Tabela 2 – Comparativo entre as Recomendações ITU-T de maior uso no mercado, para pouco consumo de banda. Através das informações de MIPS (*Million Instructions Per Second* – milhões de instruções por segundo) e memória RAM (*Read Only Memory* – memória apenas de leitura), gastos no processamento de codificação do sinal, temos uma noção da complexidade de cada uma delas. As informações de MIPS foram as de menor valores dentre vários DSPs (*Digital Signal Processor Chip* – Integrado processador de sinal digital) conhecidos pelo autor da referência [7].

Como cada tipo de codificação trabalha com uma técnica e exigência de banda diferentes, a qualidade do sinal de voz analógico resultante apresenta diferentes níveis de qualidade, devendo ser classificado segundo algum critério. O método de classificação mais empregado usa uma pontuação entre 1 e 5, onde um grupo de pessoas escuta

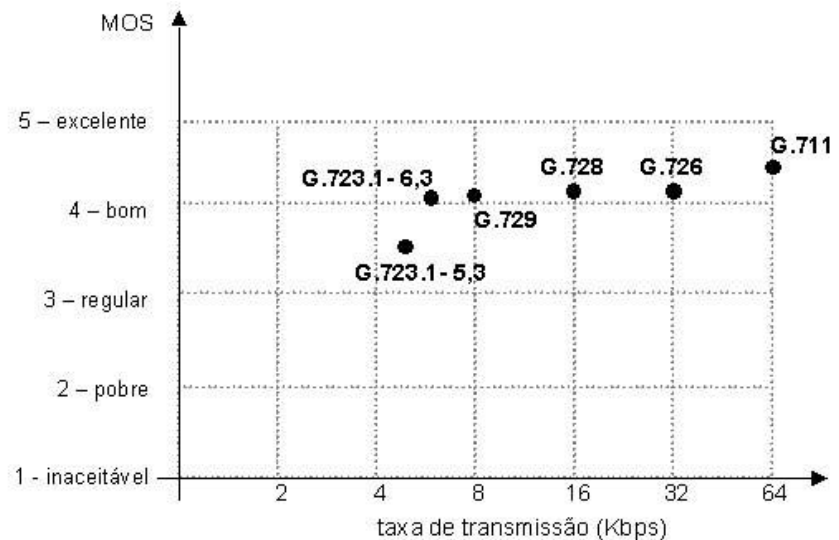


Figura 1 – MOS de algumas Recomendações ITU-T.

várias amostras de voz decodificada (sendo o mais variado possível e envolvendo vozes masculinas e femininas em igual quantidade). É tirada a média das pontuações, gerando um valor denominado MOS (*Mean Opinion Score*).

Na figura 1, podemos comparar a qualidade de algumas das principais Recomendações ITU-T. É importante lembrar que como este método de classificação é subjetivo, podemos encontrar pequenas variações de resultados, dependendo das fontes de pesquisa ou em medições práticas de laboratório, já que cada grupo de “jurados” possui sensibilidade auditiva própria, e o material de voz a ser julgado também não é idêntico para cada caso.

6. SERVIÇO DE VOZ SOBRE REDES IP

A transmissão de voz codificada em uma rede com o protocolo IP, possui características peculiares a este ambiente e importantes a se considerar. Como o IP, por si só, não oferece nenhuma garantia de Qualidade de Serviço, categorizado como tráfego de melhor resultado (*best-effort*), outros protocolos e soluções complementares devem ser agregados na formação da solução final, para permitirem um resultado comparável com o observado na rede de voz convencional.

Define-se como Qualidade de Serviço (QoS) o conjunto de requisitos necessários a uma aplicação, para qual exige-se que parâmetros como atrasos, banda, perdas, etc., estejam dentro de limites bem definidos.

6.1. Fatores que Influenciam na Qualidade de Serviço

Vários são os fatores que devem ser observados, para que a transmissão de voz em IP seja viável e atinja a qualidade desejável pelos projetistas e usuários de sistemas de áudio.

6.1.1. Banda

Cada tipo de codificador necessita de uma banda mínima para transmissão de um canal de voz. Já foi descrito anteriormente os requisitos de banda para cada uma das principais codificações, e deve-se estar alerta ao consumo de banda em um projeto onde vários canais de voz compartilharão o mesmo canal digital, pois os valores de canais de alta velocidade são bastante onerosos. Também, é claro, existe um compromisso entre a qualidade desejada para o sinal de voz recebido (item 5) e o custo da banda necessária. Podemos dividir a observação em dois ambientes: o corporativo e o residencial. O primeiro normalmente pode contar com meios de comunicação dedicados, sejam em ligações à Internet ou suas redes IP próprias, com velocidades de 64 Kbps à 2 Mbps. São meios de boa capacidade, contudo, compartilham aplicações de dados (emulação de terminal, transferências de arquivos, consulta a bases de dados, tráfego Web, etc.) e vários canais de voz.

O ambiente residencial é caracterizado por um acesso discado à Internet, que ao longo do tempo vem aumentando a velocidade de 28,8Kbps, 33,6Kbps, até os dias de hoje com a grande maioria a 56Kbps.

A fim de minimizar o requisito de banda, as técnicas mais empregadas são de supressão de silêncio na conversação e compressão de cabeçalhos dos pacotes IP.

Supressão de Silêncio

Ao longo de uma conversação existem vários períodos períodos de silêncio. Valendo-se disto, as implementações dos codificadores de voz podem reduzir a banda consumida por cada canal. O total de redução prático dependerá do tipo de codificação adotada, chegando a valores aproximadamente 25% menores, no caso da G.729.

Compressão de Cabeçalhos IP

As soluções de áudio sobre IP utilizam, além deste protocolo, o UDP (*User Datagram Protocol*) e o RTP (*Real Time Transport Protocol*) (será visto adiante), como protocolos de transporte. A soma dos cabeçalhos dos três protocolos resulta em 40 bytes (20 para o IP, 8 para o UDP e 12 para o RTP). Tomando como exemplo uma implementação de G.729 com um pacote formado por dois quadros de amostragem ou G.723.1 a 5,3Kbps com um pacote formado por um quadro de amostragem, em ambos os casos teremos 20 bytes de informação a ser transmitida (estas são implementações comuns no mercado). Fica evidente o despropósito na distribuição de bytes úteis e de controle. Assim, uma técnica adotada é descrita na RFC 2508 [25] onde a maioria dos pacotes terão seus cabeçalhos comprimidos para 2 ou 4 bytes, dependendo do uso de *checksum* pelo UDP ou não. A idéia básica é que após a transmissão do primeiro pacote descomprimido, vários campos dos pacotes seguintes podem ser suprimidos ou variam de forma conhecida, sendo “remontados” no destino.

6.1.2. Atraso

O tempo necessário para que um pacote de áudio gerado uma origem chegue até seu destino não deve ultrapassar um patamar adequado, sob a pena de degradar a qualidade da aplicação. O patamar ideal depende fundamentalmente de três aspectos: o tipo de interatividade entre os usuários da aplicação, o nível de exigência dos usuários da aplicação e o quanto se está disposto em gastar para viabilizar uma solução que reflita pequenos atrasos. Dependendo da aplicação em questão, o grau de interação entre os usuários é grande ou não. Considerando uma conversa entre duas ou mais pessoas, o tempo entre a geração do pacote de voz e a entrega no destino deve estar entre 200 e 300 ms, conforme estudos [9,14]. Já em se tratando de uma aplicação de áudio *stream*, onde a informação é transmitida em um único sentido, estes valores podem ser relaxados consideravelmente. A Norma G.114 do ITU-T, coloca que atrasos totais no sistema entre 0 e 150 ms são aceitável para a maioria das aplicações, entre 150 e 400 ms deve ser avaliado o impacto na qualidade da aplicação (sendo os casos típicos onde é usado satélite como parte da solução de transmissão) e superior a 400 ms geralmente é inaceitável, salvo os casos onde, por exemplo, não existe outra forma de acesso senão com duplo salto de satélite ou a aplicação não exija forte interação entre os usuários.

As principais parcelas que compõem o atraso total de sistema são: o processo de codificação, o tempo de transmissão e o tempo de propagação na rede.

6.1.3. Jitter

Outro fator importante para garantia de qualidade do sinal de voz recebido é a variação do tempo entre chegadas de pacotes consecutivos. Até determinados limiares, um tempo entre chegadas maior, mas com uma variação menor, é observada como de melhor qualidade do que o contrário. A esta variação do tempo entre chegadas dá-se o nome de *jitter*.

Na transmissão de voz sobre IP os datagramas podem tomar caminhos diferentes na rede, resultando em diferentes tempos de propagação, ou podemos ter congestionamentos momentâneos que obrigam a maiores retardos. Para contornar este problema, são usados *buffers* nas entradas dos equipamentos decodificadores, afim de guardar alguns pacotes como “reserva” em uma fila, que é servida de forma mais constante possível. Mesmo que alguns pacotes sofram uma demora maior que a normal para chegada (guardando certos limites), os pacotes no *buffer* são enviados para o decodificador com uma cadência homogênea. A quantidade ótima de pacotes armazenados nestes *buffers* depende do tamanho dos pacotes de voz, taxa de transmissão, atraso médio da rede e como não poderia deixar de ser, exigência da qualidade de voz requerida.

6.1.4. Taxa de Perdas e Erros

Na transmissão de voz em redes de pacotes, não podemos nos dar ao luxo de solicitarmos retransmissões dos quadros com erros ou perdidos, usando protocolos do tipo ARQ (*Automatic Repeat Request*), sob pena de queda na qualidade do sinal resultante, devido à grande variação de tempo entre chegadas de quadros consecutivos (*jitter*). Uma alternativa estudada para as perdas e erros em rajadas de poucos pacotes é o uso de algoritmos de FEC (*Forward Error Correction*) [7,8], onde um mesmo pacote IP conteria vários quadros de voz, implicando em redundância de quadros quando tomados pacotes consecutivos. Esta alternativa deve ser considerada apenas para o caso de uso de técnicas que gerem pequenos atrasos de codificação (por exemplo: G.711, G.726 e G.728, conforme tabela 1), já que o tempo necessário para a formação de um pacote poderia tornar-se crítico ao ponto de inviabilizar a solução. Mas, como sempre, ganha-se por um lado, mas perde-se por outro, vale lembrar que para estes casos não poderíamos contar com bandas menores que 16Kbps.

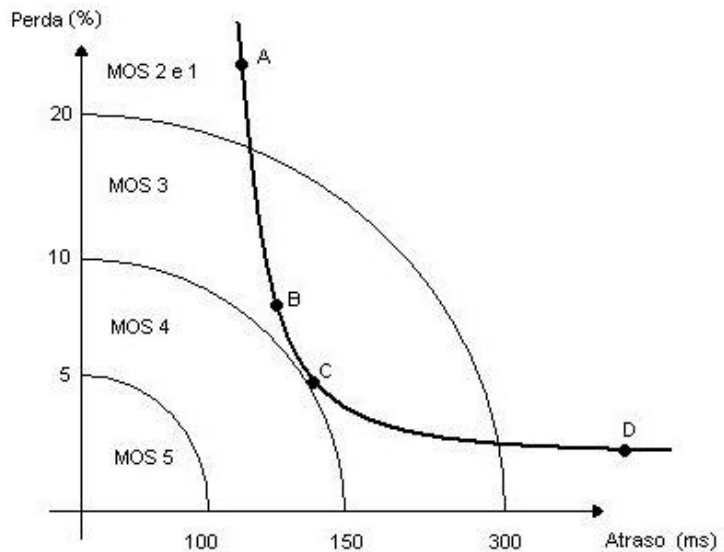


Figura 2 – Relação entre tamanho de *buffers*, perdas de pacotes e qualidade do sinal recebido.

É interessante observar, que influência a perda de pacotes causada por *overflow* dos *buffers* de entrada dos nós de rede pode causar na qualidade do sinal recebido. A figura 2 ilustra esse fenômeno para o caso da Recomendação G.723.1. No ponto C temos a melhor relação entre atrasos e perdas gerados pelos tamanhos dos *buffers*. Conforme aumentamos os tamanhos, caminhamos em direção ao ponto D, diminuindo as perdas, mas também aumentando o atraso total na rede, levando ao

distanciamento da boa qualidade. Por outro lado, se diminuirmos os *buffers*, o atraso diminui até um valor mínimo próximo do ponto A, mas as perdas aumentam e qualidade também fica prejudicada.

6.2. Implementação de Qualidade de Serviço

Para a aplicação em questão (voz sobre IP), quando se fala em garantia de Qualidade de Serviço, implicitamente estamos nos referindo à garantia de meio com banda disponível mínima suficiente para a transmissão do sinal codificado, atraso máximo de 300 ms, por exemplo, e *jitter* próximo de zero. Para isto, um conjunto de mecanismos específicos deve ser implementado, junto ao IP.

De forma mais abrangente, arquiteturas com a finalidade de proverem serviços diferenciados na Internet, tem sido o foco de extensivas pesquisas nos últimos anos. Existem duas abordagens diferentes do IETF (*Internet Engineering Task Force*), mas não mutuamente exclusivas, com esta finalidade: a de serviços integrados (IntServ) e a de serviços diferenciados (DiffServ).

IntServ – Focada no fluxo individual de pacotes, entre os mesmos pontos de origem e destino. Nesta abordagem cada fluxo pode requisitar níveis de serviços diferentes à rede, sendo tipicamente quantificados como banda mínima necessária ou tolerância a um atraso máximo definido. Os quatro componentes básicos do IntServ são: unidade de controle de admissão (que identifica se a rede pode suprir o serviço requisitado), unidade de classificação (inspeciona os campos dos pacotes para determinar suas classes e o nível de serviço acordado), unidade de *schedule* (aplica um ou mais mecanismos

de gerência de tráfego para garantir que o pacote seja transmitido à rede, a tempo de satisfazer a banda e atraso adequados ao tipo de fluxo) e o protocolo de reserva de recursos RSVP (que é responsável pela sinalização dos níveis de requisitos aos nós da rede, vide item 6.2.1).

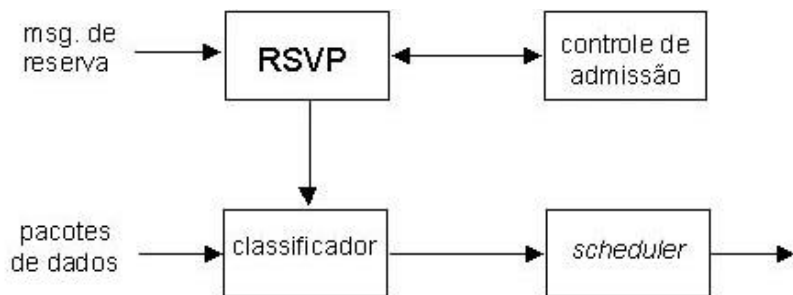


Figura 3 – Componentes básicos da arquitetura IntServ.

DiffServ – Este tem como premissa básica o foco não no fluxo individual de pacotes, mas sim no tráfego agregado com requisitos de serviços similares. Para classificação dos diferentes tipos de tráfego, existe o seguinte grupo de classes: Serviço *Premium* – para aplicações que necessitem de atrasos e *jitter* pequenos; Serviço Garantido – para aplicações que necessitem de um serviço melhor que o *best-effort* padrão; e Serviço Olímpico – que é subdividido em serviços ouro, prata e bronze, com qualidades decrescentes. Na prática, esta classificação é implementada no IPv4 através do campo TOS (*Type of Service*) do cabeçalho IP.

Os três primeiros bits, dos oito contidos no “TOS”, formam o subcampo *PRECEDENCE*, usado para identificar a importância do pacote. Quanto maior o valor neste campo, maior é a prioridade dada ao pacote .

6.2.1. RSVP

Definido na RFC 2205, o protocolo RSVP (*Resource Reservation Protocol*) não é um protocolo de roteamento, mas sim trabalha em conjunto com ele, desempenhando a função de sinalizar aos nós a necessidade de reserva de recursos de rede, permitindo que as aplicações possam usufruir da Qualidade de Serviço requerida para seus fluxos de dados, num ambiente de serviços integrados em redes IPv4 ou IPv6.

Suas funcionalidades principais são: suporte a *unicast* e *multicast*; reserva de recursos da rede apenas no sentido exigido, transparência quanto ao tipo de dados, atualização periódica quanto à capacidade e topologia da rede.

Caso um trecho intermediário ou sub-rede não suporte RSVP, é possível o uso de “tunelamento” entre os trechos que suportem o protocolo. Contudo, não teremos a garantia de QoS, devendo ser dimensionada banda suficiente para o tráfego pretendido ou garantia do recurso por outro método.

Suporta três tipos de tráfegos diferentes:

- **Melhor Resultado** (*best-effort*) – tráfego IP tradicional, sem qualquer garantia de serviço, adequado para aplicações de transferência de arquivos ou correio eletrônico, por exemplo;
- **Sensível a Velocidade** (*rate-sensitive*) – são os que necessitam de garantia de banda mínima, como os gerados por videoconferência;
- **Sensível a Atraso** (*delay-sensitive*) – são os que necessitam de garantia de atraso mínimo na rede, como por exemplo o tráfego de imagem.

O RSVP trabalha com quatro tipos básicos de mensagens:

- **reservation-request messages** – enviada por cada receptor ao transmissor, percorrendo o caminho indicado pela *path message* e solicitando reserva dos recursos necessários;
- **path messages** – enviada pelo transmissor ao receptor, percorrendo todos os roteadores indicados como de melhor caminho pelo protocolo de roteamento. Esta mensagem contém informações sobre as características do caminho, afim de subsidiar o receptor na solicitação de recursos;
- **error and confirmation messages** – existem três tipos diferentes:
 - (a) **path-error messages** – quando ocorre um erro no registro de caminho de uma *path message*, esta mensagem é retornada ao transmissor;
 - (b) **reservation-request error messages** – informa ao receptor falha de admissão, banda indisponível, serviço não suportado ou caminho ambíguo, em retorno a uma *reservation-request message*;
 - (c) **reservation-request acknowledgment messages** – enviada ao receptor como confirmação de recebimento da *reservation-request message*;
- **teardown messages** – desfaz o caminho de reserva de recursos sem esperar por um *timeout*.

Reserva de recursos na rede

Imaginemos duas máquinas, uma transmissora e outra receptora, onde determinada aplicação necessita reservar recursos da rede para funcionamento adequado. A transmissora envia à receptora uma *path message* contendo informações dos recursos solicitados pela aplicação, características do caminho roteado entre elas e o caminho de volta à transmissora. Após manusear as informações contidas na *path message*, a receptora monta uma *reservation-request message* contendo efetivamente os recursos desejados para a aplicação em questão, segundo as possibilidades da rede. Esta mensagem é então enviada à transmissora, obedecendo o trajeto contrário do envio, e solicitando a reserva de recursos em cada roteador pelo caminho (figura 4).

Note que apesar da transmissora ter indicado os recursos necessários, quem solicita a reserva na prática é a receptora, pois é ela quem pode avaliar a capacidade da rede entre as duas máquinas.

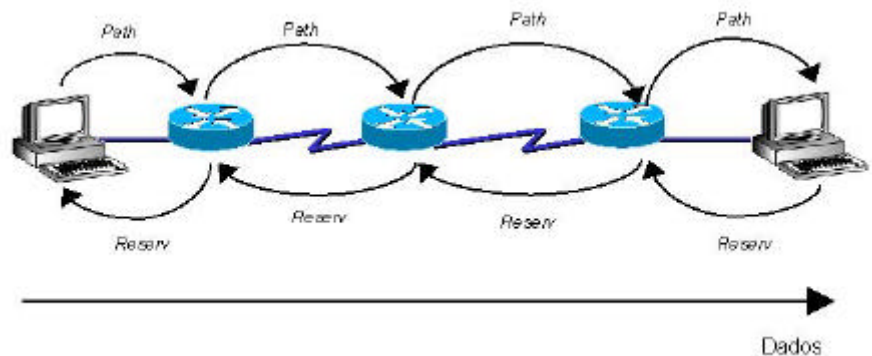


Figura 4 – Troca básica de mensagens para reserva de recursos.

Estado *soft* do RSVP

Este estado permite que a rede RSVP suporte dinamicamente as alterações do número de membros envolvidos e de topologia da rede. O estado *soft* RSVP é criado e periodicamente atualizado pelas mensagens *path* e *reservation-request*. O término deste estado ocorre quando as mensagens citadas anteriormente não são trocadas dentro de um determinado período de tempo, ocorrendo *timeout*, ou por solicitação explícita de uma mensagem *teardown*.

7. PROPOSTAS

A telefonia IP traz consigo a oportunidade de oferecimento não só dos serviços convencionais, mas também um amplo leque de criativas novas opções, integrando comunicação de voz, aplicações de dados e a potencialidade de terminais com poder de processamento. Devido ao grande parque instalado de telefonia convencional, a convivência dos dois ambientes será longa. Por conta disto, a arquitetura de sinalização deve suportar tanto a tradicional como a que permita os novos serviços. A arquitetura definida para a telefonia IP também deve se preocupar com a forma de autenticação e autorizações dos usuários, que para suporte às novas aplicações deve procurar maior semelhança com a abordagem empregada às aplicações de dados do que com a telefonia convencional.

Um bom exemplo da mudança de paradigma seria a alteração na forma de tarifação das chamadas. O modelo atual está voltado para o aparelho telefônico e não para o usuário do serviço, proporcionando pouca flexibilidade e graves problemas de segurança. Mantendo-se este mesmo modelo para o ambiente IP, os problemas de segurança são ainda piores, pois não é difícil um usuário alterar a identificação de seu terminal, usufruindo de facilidades concedidas a um grupo com maior privilégio (por exemplo a possibilidade de chamadas internacionais em um ambiente de escritório). Quantas pessoas já passaram por um leve constrangimento na hora de pedir para utilizar o aparelho telefônico de outro, sabendo que o custo da ligação será de responsabilidade do dono do aparelho. A proposta de tarifação pela identificação do usuário resolve todos estes problemas, além de outras vantagens: uso de aparelhos telefônicos públicos sem a necessidade de compra prévia de fichas ou cartões; facilidade na implementação de IP dinâmico para os terminais telefônicos, com o uso de DHCP (*Dynamic Host Configuration Protocol*); facilidade na implementação de tarifação reversa disponível apenas para um grupo seletivo de usuários (por exemplo: funcionários de uma empresa em deslocamento, poderiam acessar a rede interna por linha discada ou manter conversação sem ônus, em horário de expediente); etc.

O ITU-T e o IETF possuem propostas de ambientes, delineando este novo cenário de telefonia.

7.1. ITU-T

O ITU-T (*Telecommunication Standardization Sector of International Telecommunication Union*) tem como proposta para a telefonia IP a Recomendação H.323, que descreve uma arquitetura contendo terminais, equipamentos e serviços para comunicação multimídia sobre LANs (*Local Area Networks*), sem garantia de qualidade de serviço. Na verdade a H.323 é uma recomendação “guarda-chuva”, referenciando outras recomendações ITU-T e até o protocolo RTP (*Real-Time Transport Protocol*) do IETF. A primeira versão desta recomendação foi publicada em 1996 e a segunda em 1998. Esta última contém melhorias no tocante à segurança, diminuição do tempo necessário para liberação do canal após a chamada ter sido atendida pelo destino, incorporação de serviços adicionais como transferência, redirecionamento, etc. e maior integração com T.120 (protocolo de transmissão de dados multimídia).

Escopo

Basicamente, destina-se a transmissão em ambiente com uma ou mais redes locais, sem garantia de qualidade de serviço. Como por exemplo: Ethernet (IEEE 802.3), Fast Ethernet (IEEE 802.10), FDDI (no modo sem garantia de QoS) e Token Ring (IEEE 802.5).

Componentes da arquitetura

(a) Terminal

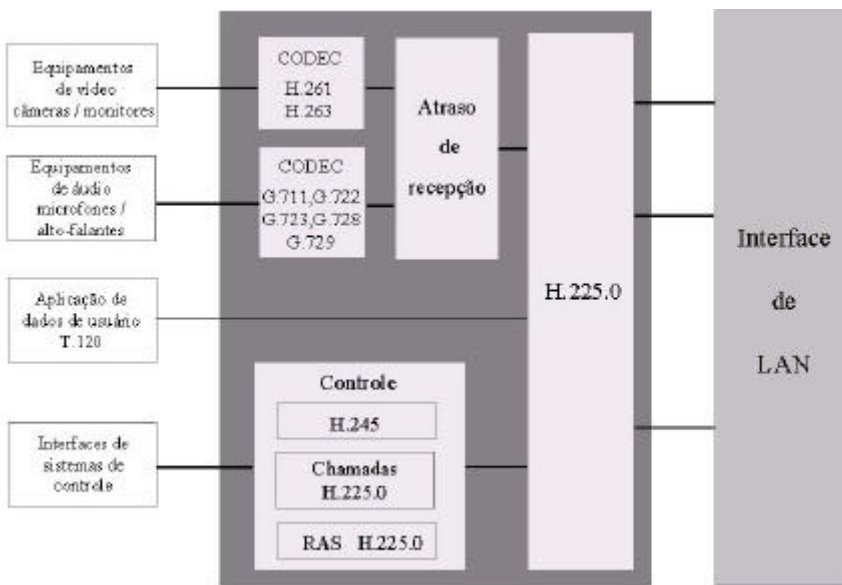


Figura 5 – Terminal H.323

Por definição é um ponto de rede que provê comunicação bidirecional, em tempo real, com outro equipamento terminal, *gateway* ou MCU. Pode ser um equipamento telefônico com a funcionalidade de comunicação através do protocolo IP, um microcomputador com *software* específico e *hardware* multimídia para comunicação de voz ou qualquer outro tipo de

equipamento que possa estar conectado em rede e permita a transferência apenas voz; voz e dados; voz e vídeo; ou voz, dados e vídeo, entre os usuários.

Na figura 5 estão ilustrados os elementos internos de um terminal H.323, sendo eles:

- **Codec (codificador/decodificador) de vídeo** – suportando obrigatoriamente a Recomendação H.261 e opcionalmente a H.263, referentes à codificação de sinal de vídeo. A implementação deste elemento é opcional;
- **Codec de áudio** – suportando obrigatoriamente a Recomendação G.711 e opcionalmente as Recomendações G.722 (sinal de 7KHz codificado em 64Kbps), G.728, G.729, áudio MPEG1 e G.723.1. A implementação deste elemento é obrigatória ;
- **Atraso de recepção** – responsável pelo controle do *jitter* nas recepções de áudio e/ou vídeo. A implementação deste elemento é obrigatória;
- **Canal de dados** – apesar de não estar explicitamente indicado na figura, suporta um ou mais canais de dados, uni ou bidirecionais. O T.120 é a base padrão para interoperabilidade entre terminais H.323 ou outro tipo suportado pelos ambientes externos ligados por *gateway*. A implementação deste elemento é opcional;
- **Controle H.245** – responsável pelo manuseio de mensagens de controle fim-a-fim, que governam a operação da entidade H.323, no tocante à troca do modo de operação de receptor para transmissor e vice-versa, abertura e fechamento de canais lógicos para troca de informações de um transmissor para um ou mais receptores, requisições de preferência de modos de operação (*unicast* ou *multicast*), mensagens de controle de fluxo e comandos gerais e indicações;

- **Controle RAS** – controle das mensagens de sinalização para registro, admissão, troca de banda, pedido de informação sobre a situação atual e desvinculação, com o *gatekeeper*;
- **Controle de Chamadas** – usado para o estabelecimento de conexão entre dois terminais H.323;
- **H.225.0** – estabelece o formato das mensagens trocadas pelos canais lógicos de áudio, vídeo dados ou controle.

(b) Gateway

Converte, apropriadamente, diferentes formatos de mensagens (ex.: H.225.0 / H.221 - estrutura de quadros em canais de 64 a 1920 Kbps em telesserviços audiovisuais) e procedimentos de comunicação (ex.: H.245 / H.242 - estabelecimento de comunicação entre terminais audiovisuais em canais digitais até 2 Mbps). A conversão entre diferentes formatos de áudio, vídeo ou dados também pode ser feita pelo *gateway*.

(c) Gatekeeper

É um elemento opcional no sistema H.323, provendo serviços de controle de chamadas para os terminais. Possui como funções obrigatórias a tradução de endereços (permite o uso de apelidos (*alias*) em lugar dos Endereços de Transporte), controle de admissão (autoriza o acesso de recursos usando por base critérios como: permissão de acesso em período predeterminado, banda disponível, etc.), controle de banda (rege os pedidos de troca de banda em uso) e a gerência de zona (prove as funções acima para terminais, *MCUs (Multipoint Control Units)* e *gateways* nele registrados).

Opcionalmente, o *gatekeeper* pode apresentar as funções de controle de sinalização de chamadas (pode controlar todo processo de sinalização de chamadas entre terminais, ou deixar que os próprios terminais tratem da sinalização), autorização de chamadas (através da sinalização prevista na H.225.0, o *gatekeeper* pode rejeitar chamadas de um terminal sem autorização adequada), gerência de banda (controla o número de acessos simultâneos de terminais H.323 à LAN, podendo rejeitar uma chamada por falta de banda disponível), gerência de chamadas (controla que terminais possuem chamadas estabelecidas, para o caso de uma nova chamada a um terminal ocupado ser informada aos módulos de direito).

Um conceito importante, introduzido pela segunda versão do H.323 é o de Zona. Por definição, Zona abrange a coleção de todos os terminais, *gateways*, e *MCUs*, gerenciados por um único *gatekeeper*. Cada Zona possui apenas um *gatekeeper* e não precisa estar confinada geograficamente a mesma LAN.

(d) Multipoint Control Unit – MCU (Unidade de Controle de Multiponto)

Elemento para suporte a conferência com três ou mais pontos, podendo funcionar em três modos distintos:

- **Modo Centralizado:** a comunicação entre a MCU e os terminais ou *gateway* é *unicast*. Dados, áudio, vídeo e controle passam obrigatoriamente pela MCU;
- **Modo Descentralizado:** os terminais trocam informações de controle, e opcionalmente de dados, de forma centralizada com a MCU, mas trocam áudio e vídeo entre si por *multicast*;
- **Modo Híbrido:** a comunicação de dados e controle sempre se dá de forma centralizada com a MCU. Contudo, podemos ter áudio também centralizado e *multicast* de vídeo ou vice-versa.

Procedimento para o estabelecimento de chamadas

Fase 1 (RAS): Um terminal quer fazer uma chamada. Ele envia ao *gatekeeper* (GK) uma mensagem "*Admission Request*". Esta msg. contém: apelido do destino, nome ou número telefônico do usuário chamador, etc. O GK pode aceitar o pedido, enviando em retorno uma mensagem. "*Admission Confirm*", contendo o endereço de transporte associado à chamada ou rejeitar, enviando uma mensagem "*Admission Reject*" contendo o motivo (ex.: falta de banda);

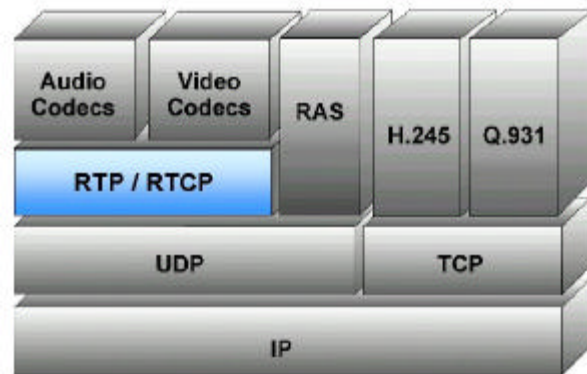


Figura 6 – Pilha de protocolos.

Fase 2 (Q.931): Esta fase é derivada da sinalização de estabelecimento de chamada fim-a-fim da RDSI (Rede Digital de Serviços Integrados). Provê a conexão lógica entre os pontos chamado e chamador. O Q.931 é um protocolo de sinalização usada em ISDN, implementado sobre a camada TCP.

Fase 3 (H.245): Tão logo a fase anterior tenha terminado, os pontos de origem e destino trocam informações sobre a natureza do tráfego (voz, dados e/ou vídeo) e seus formatos.

7.2. IETF

Esta Entidade possui vários protocolos que compõem boa parte da solução para o complexo problema de integração de serviços na Internet:

SIP (*Session Initiation Protocol* - RFC 2543) - Sinalização;
 SDP (*Session Description Protocol* - RFC 2327) - Definição de sessão multimídia;
 RTP/RTCP (RFC 1889) - Entrega fim-a-fim para dados interativos;
 RTSP (*Real Time Streaming Protocol* - RFC 2326) - Controle de servidores de mídia;
 CPL (*Call Processing Language - Draft*) - *Script* para o usuário final;
 GLP (*Gateway Location Protocol - Draft*) - Localização de usuários fora de Internet;
 MGCP (*Media Gateway Control Protocol - Draft*) - Controle de *gateways*;
 MDCP (*Media Device Control Protocol - Draft*) - Controle de dispositivos.

Neste item iremos abordar apenas o protocolo de sinalização SIP, usado para sinalização de seções multimídia.

O SIP é um protocolo de controle referente à camada de aplicações do Modelo de Referência OSI (*Open System Interconnection*), usado para iniciar, modificar ou terminar sessões ou chamadas multimídia entre usuários. Possui como funcionalidades a localização de usuários, estabelecimento de chamadas, suporte a *unicast* ou *multicast*, administração na participação de chamadas (transferências, conferência, ...) e possibilidade de participação de um usuário em terminal H.323, via *gateway*. É um protocolo cliente-servidor similar ao HTML no tocante à sintaxe e semântica das estruturas empregadas, com campos explicitamente descritos.

Durante seu funcionamento, requisições são geradas por um cliente e enviadas ao servidor (entidade receptora), que processa tal requisição e envia a resposta ao cliente. O cliente SIP, chamado “agente usuário”, é formado por dois módulos: o agente usuário cliente, responsável em gerar requisições e o agente usuário servidor, responsável em responder as requisições. Os dois módulos são sempre obrigatórios para todo cliente.

Existem três tipos diferentes de servidores SIP:

- **Servidor de Registro** - recebe atualizações da localização dos usuários na rede;
- **Servidor Proxy** - recebe requisições e as encaminha para outros servidores que contenha informação mais precisa sobre a localização do usuário chamado;
- **Servidor de Redirecionamento** - também recebe requisições, mas no lugar de encaminha-las para outros servidores, retorna o endereço do servidor para o qual a requisição deverá ser direcionada.

O SIP adota como forma de endereçamento uma URL (*Uniform Resource Locators*), similar a um endereço de correio SMTP (*Simple Mail Transfer Protocol*), formada por uma parte referente ao nome do usuário ou número telefônico (*user*) e outra ao nome do domínio ou endereço numérico da rede (*host*), isto é, [user@host](#).

Comandos usados pelo protocolo:

- *INVITE* - “convida” um usuário para uma chamada;
- *BYE* - termina a conexão entre dois usuários;
- *OPTIONS* - solicita informações de capacidade;
- *ACK* - resposta a um *INVITE*;
- *CANCEL* - interrompe a busca por um usuário;
- *REGISTER* - Informa a localização de um usuário ao servidor de registro.

Como exemplo do formato das mensagens do protocolo, suponhamos que a usuária Maria queira saber os tipos de codificações suportados pelo usuário Lucas, antes de fazer uma chamada. Através de um *OPTIONS request*, a usuária solicita a informação desejada ,utilizando o campo *Accept*. Esta mensagem é enviada do módulo cliente do agente usuário de Maria, para o módulo servidor do agente usuário de Lucas:

```

OPTIONS sip:lucas@example.com SIP/2.0
From: Maria <sip:maria@anywhere.org>
To: Lucas <sip:lucas@example.com>
Call-ID: 6378@host.anywhere.org
Cseq: 1 OPTIONS
Accept: application/sdp

```

Em resposta, o módulo cliente do agente usuário de Lucas envia um *OPTIONS response* indicando ser possível receber áudio no padrões: PCM Ulaw (*payload* tipo 0), 1016 (*payload* tipo 1), GSM (*payload* tipo 3), e SX7300/8000 (*payload* tipo 99); e vídeo H.261 (*payload* tipo 31) e H.263 (*payload* tipo 34):

```

SIP/2.0 200 OK
From: Maria <sip:maria@anywhere.org>
To: Lucas <sip:lucas@example.com> ;tag=376364382
Call-ID: 6378@host.anywhere.org
Content-Length: 81
Content-Type: application/sdp
v=0
m=audio 0 RTP/AVP 0 1 3 99
m=video 0 RTP/AVP 31 34
a=rtpmap:99 SX7300/8000

```

Processo de operação do SIP

A figura 7 ilustra um processo de estabelecimento de chamada. Um usuário encaminha um *INVITE request* para um *proxy* local, tendo como destino nelson@empresa.com (1). O *proxy* procura empresa.com no DNS (*Domain Name System*) e obtém o endereço IP do servidor que trata requisições SIP deste domínio. O *proxy* então encaminha o *request* para esse servidor (2). O servidor de empresa.com conhece o usuário nelson, mas ele está “logado” no momento como n.aluno@universidade.edu. Então, o servidor redireciona o *proxy* (3) para tentar este

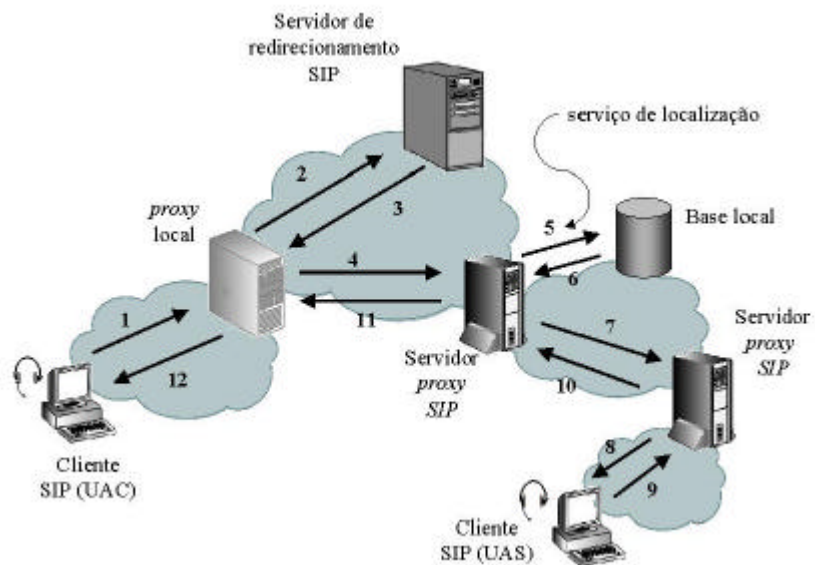


Figura 7 – Operação básica do SIP.

endereço. Também com o auxílio do DNS, o *proxy* obtém o endereço IP do servidor responsável por universidade.edu e envia o *request* para esse servidor (4). O servidor consulta uma base local (5) e indica que n.aluno@universidade.edu é localmente identificado como nelson.fernandes@eng.universidade.edu (6). O servidor principal de universidade.edu envia o *request* para o servidor da engenharia (7), que por sua vez, sabendo o endereço IP do usuário, envia o *request* para ele (8). O usuário aceita a chamada e responde retornando pela cadeia de *proxy* (9), (10), (11) e (12).

5.2.1. Comparação das Propostas

De forma resumida, podemos considerar que:

H.323

- Fortemente baseado em protocolos do ITU-T já existentes;
- Abordagem voltada para os equipamentos terminais.

SIP

- Similar ao HTTP;
- Abordagem voltada para os usuários de serviços integrados na Internet.

Ambos usam o RTP para troca de dados multimídia.

Baseado em [22], será apresentada uma comparação entre H.323 e SIP, considerando os seguintes aspectos:

- ① Complexidade ⇒ complexidade de implementação e/ou funcionamento;
- ② Expansão Funcional ⇒ facilidade na inclusão de novas funcionalidades;
- ③ Escalabilidade ⇒ facilidade para aumento da quantidade de elementos interligados;
- ④ Serviços ⇒ serviços oferecidos.

Complexidade

- O total de páginas na especificação do H.323 soma 736, em comparação com 128 páginas no SIP. Isto reflete maior complexidade para implementação do H.323, devido ao maior esforço do desenvolvedor para entendimento da especificação por completo;
- O H.323 faz uso de centenas de elementos diferentes para comunicação. Já o SIP trabalha com apenas 37 tipos de cabeçalhos diferentes;

- As mensagens do H.323 possuem representação binária para cada campo, enquanto o SIP utiliza texto para representação, facilitando em muito o entendimento visual rápido do protocolo;
- O H.323 usa vários protocolos sem uma separação clara, isto é, muitos serviços usam componentes dos vários protocolos, dificultando inclusive a configuração de *Firewalls*. O SIP coloca todas as informações necessárias em uma mesma requisição;
- A completa interoperabilidade com o H.323 exige a definição de perfis funcionais. O SIP trabalha de forma única para tratamento de uma mesma tarefa.

Expansão Funcional

- A estrutura textual dos campos no SIP, permite que novas características sejam incluídas de forma fácil e compatível com as versões anteriores. Os novos campos ou parâmetros podem ser colocados em qualquer parte da mensagem. Já no H.323, existem alguns locais predefinidos para inclusões futuras;
- O suporte a novos padrões de codificadores de áudio e vídeo é livre no SIP, basta que este novo “codec” seja registrado com uma identificação em órgão competente. No caso do H.323, os “codecs” devem ser padronizados pelo ITU, dificultando o processo de inclusões dos “codecs” de terceiros.

Escalabilidade

- Os *gateways* ou servidores SIP podem trabalhar nos modos *stateful* ou *stateless*. No segundo caso, o servidor recebe uma requisição de chamada, executa algumas operações, encaminha a requisição e depois não se preocupa mais com este processo. As mensagens contêm informações de estado suficientes para a resposta ser encaminhada corretamente. Os *gateways* e *gatekeepers* H.323 são *stateful*, mantendo controle do estado da chamada durante toda a sua duração. Quando tratamos de um ambiente com uma grande quantidade de chamadas simultâneas, isto pode implicar em um sério problema de performance;
- Conferências com H.323 obrigatoriamente necessitam da MCU para realização. Que centraliza toda sinalização, por menor que seja o número de participante. No caso de conferências com muitos participantes, este equipamento pode ser um “gargalo”. O SIP trabalha com o controle da conferência de forma distribuída pelos participantes, sem a necessidade de um equipamento centralizador.

Serviços

- O conjunto de serviços oferecidos pelos dois protocolos é bastante similar. Entende-se como serviço as facilidades de transferências, conferências, encaminhamento de chamadas, etc.

Em [35] é apresentada uma comparação entre os tempos necessários para o estabelecimento de uma chamada entre o SIP e o H.323.

A figura 9 mostra a troca de mensagens para os dois casos.

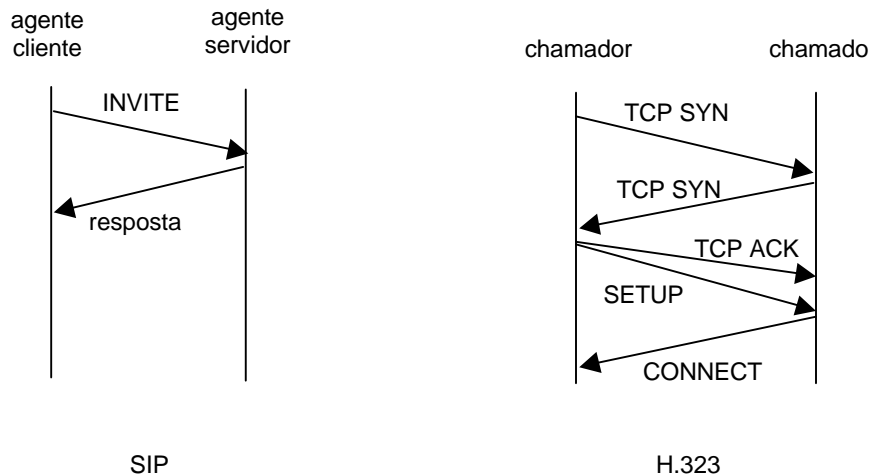


Figura 8 – Procedimentos para estabelecimento de chamada SIP (porção inicial) e H.323 (Fast Connect).

Foram feitas medidas de simulações dos dois processos entre alguns locais do Estados Unidos da América, e indicados os percentuais de dias onde ocorreu mais de 1% de rejeição de chamadas, estabelecido o tempo máximo de espera em 2 segundos.

| | | Boston | Chicago | West Coast | Washington | Colorado |
|------------|-------|--------|---------|------------|------------|----------|
| New York | SIP | 20,3 | 77,2 | 32,3 | 9,1 | 15,4 |
| | H.323 | 28,2 | 94,7 | 40,0 | 20,0 | 18,5 |
| Boston | SIP | | 1,6 | 31,5 | 0,0 | 5,4 |
| | H.323 | | 1,6 | 31,5 | 0,0 | 10,8 |
| Chicago | SIP | | | 34,3 | 5,2 | 28,6 |
| | H.323 | | | 34,3 | 6,9 | 61,4 |
| West Coast | SIP | | | | 33,3 | 45,3 |
| | H.323 | | | | 36,7 | 57,3 |
| Washington | SIP | | | | | 6,6 |
| | H.323 | | | | | 6,6 |

Tabela 3 – Percentual de dias onde mais de 1% das chamadas não foram completadas.

8. RTP / RTCP (*Real-Time Transport Protocol / RTP Control Protocol*)

Como a transmissão de voz sobre IP usa o UDP como protocolo de transporte, os datagramas podem tomar caminhos diferentes na rede, resultando em diferentes tempos de propagação, sendo possível inclusive que um pacote transmitido posteriormente, chegue ao destino primeiro que seu antecessor. Esse é apenas um dos problemas que o RTP se propõe a contornar.

O RTP provê um serviço de entrega fim-a-fim, para dados com características de tempo real, tais como áudio e vídeo interativos (RFC 1889). Suporta transferência de dados para múltiplos destinos, usando distribuição *multicast*. Seus serviços são: identificação do tipo de *payload*, numeração de seqüência, indicação do tempo da amostragem (*timestamping*) e monitorização de entrega. Ele não implementa as funcionalidades de entrega em prazo máximo, garantia de QoS e garantia de entrega.

O RTCP é baseado na transmissão periódica de pacotes de controle por todos os participantes de uma sessão, a fim de monitorizar a qualidade de serviço e transportar informações destes participantes (RFC 1889). Possui como funções básicas o retorno sobre a congestionamento e fluxo dos dados, transporte do "CNAME" de um elemento originador de chamada, para evitar conflitos de identificação em casos anômalos e controle da taxa de transmissão dos pacotes, para o caso de muitos participantes.

BIBLIOGRAFIA

- [1] ITU-T Rec. G.114, "*General Recommendations on the Transmission Quality for an Entire International Telephone Connection*", março de 1993;
- [2] L. Zhang *et al.*, "*RSVP: A New Resource ReSerVation Protocol*", IEEE Network, setembro de 1993;
- [3] D. E. Comer, *Internetworking with TCP/IP*, Prentice Hall, Vol. 1, 3ª Edição, 1995, páginas 89-107;
- [4] H. Schulzrinne *et al.*, "*RTP: A Transport Protocol for Real-Time Applications*", IETF RFC 1889, janeiro de 1996;
- [5] H. Schulzrinne *et al.*, "*SIP: Session Initiation Protocol*", IETF RFC 2543, janeiro de 1996;
- [6] H. Schulzrinne, "*RTP Profile for Audio and Video Conferences with Minimal Control*", IETF RFC 1890, janeiro de 1996;
- [7] J. Bolot e A. Vega-Garcia, "*Control Mechanisms for Packet Audio in the Internet*", Proc. NOSSDAV, Cambridge U.K., abril de 1996;
- [8] J. Bolot e A. García, "*The Case for FEC-Based Error Control for Packet Audio in the Internet*", Proc. NOSSDAV, Cambridge U.K., abril de 1996;
- [9] R. V. Cox and P. Kroon, "*Low Bit-Rate Speech Coders for Multimedia Communication*", IEEE Communication Mag. , dezembro de 1996;
- [10] G. Thom, "*H.323: The Multimedia Communications Standard for Local Area Networks*", IEEE Communication Mag. , dezembro de 1996;
- [11] A. S. Tanenbaum, *Redes de Computadores*, Campus, 3ª Edição, 1997, páginas 115-141;

- [12] Draft ITU-T Rec. H.323v2, “*Visual Telephone Systems and Equipment for Local Area Networks which Provide a Non-Guaranteed Quality of Service*”, fevereiro de 1997;
- [13] R. Braden *et al.*, “*Resource ReSerVation Protocol (RSVP)*”, IETF RFC 2205, setembro de 1997;
- [14] R. V. Cox, “*Three New Speech Coders from the ITU Cover a Range of Application*”, IEEE *Communication Mag.*, setembro de 1997;
- [15] G. Schröder and M. H. Sherif, “*The Road to G.729: ITU 8-kb/s Speech Coding Algorithm with Wireline Quality*”, IEEE *Communication Mag.*, setembro de 1997;
- [16] R. Salami *et al.*, “*ITU-T G.729 Annex A: Reduced Complexity 8 kb/s CS-ACELP Codec for Digital Simultaneous Voice and Data*”, IEEE *Communication Mag.*, setembro de 1997;
- [17] Georgia Tech – *Graphics, Visualization & Usability Center’s, Eight www User Survey*, http://www.gvu.gatech.edu/user_surveys/survey-1997-10/graphs/technology/Connection_Speed.html, outubro de 1997;
- [18] T. J. Kostas *et al.*, “*Real-Time Voice Over Packet-Switched Networks*”, IEEE *Network*, janeiro/fevereiro de 1998;
- [19] R. Gareiss, “*Voice Over IP Services: The Sound Decision*”, *Data Communication*, março de 1998;
- [20] A. Cray, “*Voice Over IP – Hear’s How*”, *Data Communication*, abril de 1998.
- [21] N. Anerousis *et al.*, “*The TOPS Architecture for Signaling, Directory Services and Transport for Packet Telephony*”, arquivo “*nossdav98-021.ps.gz*” em <http://www.nossdav.org/1998/papers/>, Proc. NOSSDAV, Cambridge U.K., julho de 1998;
- [22] H. Schulzrinne e J. Rosenberg, “*A Comparison of SIP and H.323 for Internet Telephony*”, arquivo “*nossdav98-043.ps.gz*” em <http://www.nossdav.org/1998/papers/>, Proc. NOSSDAV, Cambridge U.K., julho de 1998;
- [23] Georgia Tech – *Graphics, Visualization & Usability Center’s, Eight www User Survey*, http://www.gvu.gatech.edu/gvu/user_surveys/survey-1998-10/graphs/technology/q01.htm, outubro de 1998;
- [24] Cisco Systems Inc., “*Empowering the Internet Generation*”, palestra ministrada na Dataprev, 1998;
- [25] S. Casner e V. Jacobson., “*Compressing IP/UDP/RTP Headers for Low-Speed Serial Links*”, IETF RFC 2508, fevereiro de 1999;
- [26] X. Xiao e L. Ni, “*Internet QoS: A Big Picture*”, IEEE *Network*, março/abril de 1999;
- [27] C. Metz, “*RSVP: General-Purpose Signaling for IP*”, IEEE *Internet Computing*, maio/junho de 1999;
- [28] H. Schulzrinne, *Guest Editorial – “Internet Telephony”*, IEEE *Network*, maio/junho de 1999;
- [29] D. Rizzetto e C. Catania, “*A Voice over IP Service Architecture for Integrated Communications*”, IEEE *Network*, maio/junho de 1999;
- [30] H. Schulzrinne e J. Rosenberg, “*The IETF Internet Telephony Architecture and Protocols*”, IEEE *Network*, maio/junho de 1999;
- [31] B. Goodman, “*Internet Telephony and Modem Delay*”, IEEE *Network*, maio/junho de 1999;
- [32] C. Huitema *et al.*, “*An Architecture for Residential Internet Telephony Service*”, IEEE *Network*, maio/junho de 1999;
- [33] C. A. Polyzois, “*From POTS to PANS: A Commentary on the Evolution to Internet Telephony*”, IEEE *Network*, maio/junho de 1999;

- [34] Cisco systems Inc., “*Resource ReSerVation Protocol (RSVP)*”, http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/rsvp.html, julho de 1999;
- [35] H. Schulzrinne e T. Evers, “*Predicting Internet Telephony Call Setup Delay*”, <http://www.cs.columbia.edu/~hgs/papers/Eyer99:Predicting.ps>, setembro de 1999;
- [36] J. Martins, “Qualidade de Serviço (QoS) em Redes IP – Princípios Básicos, Parâmetros e Mecanismos”, http://www.jsmnet.com/Downloads/ReviewsJSMNet1_1.zip, setembro de 1999;
- [37] C. Dovrolis e P. Ramanathan, “*A case for Relative Differentiated Services and the Proportional Differentiation Model*”, *IEEE Network*, setembro/outubro de 1999;
- [38] DataBeam Corp, “*A Primer on the H.323 Series Standard*”, <http://www.databeam.com/h323/h323.primer.html>;
- [39] F. Aleixo, “RSVP – Uma Implementação de Qualidade de Serviço sobre a Tecnologia da Internet”, http://www.jsmnet.com/Downloads/RSVP_Fellipe.zip.