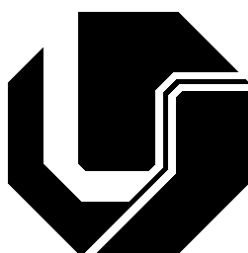


UNIVERSIDADE FEDERAL DE UBERLÂNDIA
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA



“ESTUDO DO IMPACTO DO
COMPRIMENTO DOS DADOS DE VOZ DE
PACOTES IP EM REDES VOIP EM UM DOMÍNIO
DE SERVIÇOS DIFERENCIADOS”

HERON FRANÇA DE OLIVEIRA

JUNHO

2005

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

“ESTUDO DO IMPACTO DO
COMPRIMENTO DOS DADOS DE VOZ
DE PACOTES IP EM REDES VOIP EM UM DOMÍNIO
DE SERVIÇOS DIFERENCIADOS.”

HERON FRANÇA DE OLIVEIRA

Junho

2005

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

“ESTUDO DO IMPACTO DO
COMPRIMENTO DOS DADOS DE VOZ
DE PACOTES IP EM REDES VOIP EM UM DOMÍNIO
DE SERVIÇOS DIFERENCIADOS.”

Dissertação apresentada por Heron França de Oliveira à
Universidade Federal de Uberlândia para obtenção do título
de Mestre em Engenharia Elétrica aprovada em 02/06/2005
pela Banca Examinadora:

Professor Dra. Edna Lúcia Flôres (orientadora)

Professor Dr. Paulo Roberto Guardieiro

Professor Dr. Rodrigo Pinto Lemos

**“ESTUDO DO IMPACTO DO COMPRIMENTO DOS DADOS DE VOZ DE
PACOTES IP EM REDES VOIP EM UM DOMÍNIO DE SERVIÇOS
DIFERENCIADOS.”**

HERON FRANÇA DE OLIVEIRA

Dissertação apresentada por Heron França de Oliveira à Universidade Federal de Uberlândia
como parte dos requisitos à obtenção do título de Mestre em Engenharia Elétrica.

Prof. Dra. Edna Lúcia Flôres

Orientadora

Professor Darizon Alves de Andrade

Coordenador do Curso de Pós Graduação

AGRADECIMENTOS

Agradeço primeiramente à minha orientadora Profa. Dra. Edna Lúcia Flôres que, desde o princípio, me incentivou a dar início e continuidade neste trabalho. Compartilhando conhecimentos e experiências, validando cada etapa conquistada na elaboração desta dissertação.

Aos Profs. Dr. Paulo Roberto Guardieiro, Dr. Rodrigo Pinto Lemos e Dr. Pedro Frosi pela honra de compartilharem comigo seus valiosos conhecimentos no direcionamento inicial deste trabalho.

À minha esposa Rosimar e meus filhos Yuri, Yantor e Taynah, pelo apoio, pelo amor e por compreenderem os momentos de ausência. Amo a todos vocês.

À minha Mãe Mary, meu Pai José, e meus Irmãos, Aldo, Caio e Saulo, juntamente com toda a minha família e todos os meus amigos, por compartilharem comigo esta conquista.

Aos meus colegas de trabalho, Dácio, Iraci, Isa'lice, Joneci, Júnio César, Maria Helena, Milton, Plínio e à amiga Solange que muito contribuiu para a realização deste trabalho. Obrigado a todos vocês por terem me apoiado sempre que foi necessário.

E agradeço a DEUS acima de tudo, que é o princípio de toda vitória.

RESUMO

A qualidade de serviço (QoS) na rede IP é essencial em uma aplicação VoIP. Devido a inerente natureza de transmissão incerta da tecnologia de melhor-esforço utilizada e as demais características das redes IP, a QoS e a confiabilidade não podem ser garantidas nos serviços VoIP. Este trabalho estuda o funcionamento de uma rede de voz sobre o protocolo IP e propõe uma análise do impacto do comprimento dos dados de voz no pacote IP de uma rede VoIP em um domínio de serviços diferenciados, em relação aos seus efeitos nas métricas de medição da qualidade de serviço (largura de faixa, atraso e variação do atraso fim-a-fim). Os resultados obtidos nas simulações realizadas neste trabalho mostraram significativa melhoria na eficiência da transmissão, com a redução da largura de faixa utilizada para os dados de voz de maior comprimento. No entanto, os resultados obtidos para os pacotes com maior comprimento de dados da voz também apresentaram um sensível aumento do atraso médio fim-a-fim, provocando redução na qualidade da voz. Quanto à variação do atraso fim-a-fim, os resultados mostraram apenas uma pequena redução *jitter* média fim-a-fim em relação ao aumento do comprimento dos dados da voz por pacote IP para o segundo cenário do modelo simulado neste trabalho. A solução para o problema de redução da qualidade da voz em relação às métricas analisadas nesta dissertação que afetam essa qualidade e com aproveitamento da eficiência da transmissão, pode-se sugerir, pelos resultados obtidos, que a variação da quantidade de quadros por pacote seja feita de forma dinâmica na fonte de geração do tráfego pelos *gateways* de VoIP.

Palavras-chave: VoIP, QoS, DiffServ, Redes IP, Qualidade de voz.

ABSTRACT

The quality of service (QoS) in the IP network is essential in VoIP applications. Due to inherent nature of uncertain transmission of the best-effort used and much characteristics technology of IP networks too, the QoS and the reliability can not be guaranteed in VoIP's Services. This work studies the functioning of a voice over IP network and proposes a study of the impact of the length of the voice data in IP packets of a VoIP's network in a domain of differentiated services, regarding their effects in the measurement QoS's metrics (bandwidth, delay end-to-end and delay end-to-end variation). The results obtained in the accomplished simulations in this work showed significant improvement in the efficiency of the transmission, with the reduction of the bandwidth used to the voice data of larger length. However, the results also obtained for the packets with data larger length of the voice introduced a sensitive increase of the average delay end-to-end, provoking reduction in the voice quality. Regarding the variation of the delay end-to-end, the results showed only a small reduction of the average jitter end-to-end with the increase of the length of the voice data for IP packets, for the second scenery of the simulated model in this work. The solution for the reduction problem of the voice quality regarding the analyzed metrics in this dissertation that affect this quality and with utilization of the efficiency of the transmission of the bandwidth used, it can suggest, by the obtained results, that the quantity variation of the frames for packet be made of dynamic form in the generation source of the traffic by *VoIP's Gateways*.

Keywords: VoIP, QoS, DiffServ, IP Networks, Voice quality

Estudo do impacto do comprimento dos dados de voz de pacotes IP em redes VoIP em um Domínio de Serviços Diferenciados.

Sumário

CAPÍTULO I	1
INTRODUÇÃO	1
1.1 – Motivação.....	1
1.2 – Trabalhos Relacionados	6
1.3 – Proposta deste Trabalho.....	7
1.4 – Estrutura desta Dissertação	7
1.5 – Considerações Finais deste Capítulo	8
CAPÍTULO II.....	9
VOZ SOBRE IP.....	9
2.1 – Introdução.....	9
2.2 – Integração de Voz e de Dados Sobre IP	11
2.3 – Exemplos de Cenários de VoIP	11
2.4 – Padrões de VoIP	13
2.4.1 – O empacotamento em VoIP	15
2.4.2 – Sinais DTMF	15
2.5 – Dispositivos de VoIP	17
2.6 – Considerações Finais deste Capítulo	18
CAPÍTULO III	19
PROTOCOLOS DE SINALIZAÇÃO DE CHAMADA E DE CONTROLE DE GATEWAYS.....	19
3.1 – Introdução.....	19
3.2 – Arquitetura H.323	20

3.2.1 – Elementos da arquitetura H.323	21
3.2.2 – A sinalização da arquitetura H.323	24
3.2.3 – Sinalização RAS	26
3.2.4 – Sinalização de chamada H.225 (<i>Call Signaling</i>)	26
3.2.5 – Sinalização de controle H.245	27
3.3 – Protocolo SIP	28
3.3.1 – Entidades da rede SIP	29
3.3.2 – O estabelecimento de uma chamada SIP	30
3.3.3 – Resumo da sintaxe das mensagens do SIP	33
3.3.4 – Forma de endereçamento adotada no SIP	35
3.3.5 – Interoperando com outros sistemas de sinalização	36
3.3.5.1 – Interoperando com a rede de telefonia pública comutada	37
3.3.5.2 – Interoperação do protocolo SIP com a arquitetura H.323	37
3.4 – Protocolos de Sinalização e Controle de Gateway	38
3.4.1 – A arquitetura <i>softswitch</i>	40
3.4.2 – Protocolos para controle de <i>gateways</i> de mídia	41
3.4.3 – O MGCP	42
3.4.4 – O MEGACO/H.248	43
3.4.4.1 – Entidades lógicas do MEGACO	43
3.5 – Considerações Finais deste Capítulo	46
CAPÍTULO IV	47
TÉCNICAS DE CODIFICAÇÃO, QUALIDADE DA VOZ E QOS	47
4.1 – Introdução	47
4.2 – Codificação da voz	48
4.2.1 – O codificador G.711 ou PCM	48
4.2.2 – O codificador G.723.1 (ACELP e MP-MLQ)	49
4.2.3 – O codificador G.729 ou CS-ACELP	50
4.2.4 – Métodos de qualificação do sinal de voz	52
4.2.4.1 – <i>Mean opinion score</i> (MOS)	53
4.2.4.2 – <i>Perceptual speech quality measure</i> (PSQM)	53
4.2.4.3 – <i>Perceptual speech quality measurement plus</i> (PSQM+)	53
4.2.4.4 – <i>Perceptual analysis measurement system</i> (PAMS)	54
4.2.4.5 – <i>Perceptual evaluation of speech quality</i> (PESQ)	54
4.2.5 – Fatores que influenciam a qualidade da voz na rede VoIP	55
4.2.5.1 – Largura de faixa	55
4.2.5.2 – Atraso	58
4.2.5.3 – Variação de atraso (<i>jitter</i>)	60
4.2.5.4 – Fragmentação	61
4.2.5.5 – Vazão, precisão e perda de pacotes	61
4.3 – Quality of Service (QoS)	62
4.3.1 – QoS fim-a-fim	63
4.3.2 – Parâmetros de QoS para aplicações de VoIP	64
4.3.3 – Resumo das soluções de QoS	65

4.3.4 – Mecanismos de implementação de QoS	65
4.3.4.1 – <i>Dejitter buffer</i>	66
4.3.4.2 – Serviço de melhor-esforço	67
4.3.4.3 – Arquitetura <i>Intserv</i>	67
4.3.4.4 – Arquitetura <i>DiffServ</i>	70
4.3.4.5 – MPLS	71
4.3.4.6 – Compressão de cabeçalho	72
4.3.4.7 – Detector de atividade da voz	73
4.3.5 – Políticas de QoS	73
4.3.6 – Combinando as soluções de QoS	74
4.4 – Considerações Finais deste Capítulo	75
CAPÍTULO V	76
SIMULAÇÕES E RESULTADOS OBTIDOS	76
5.1 – Introdução	76
5.2 – O Simulador de Redes NS	76
5.2.1 – Suporte ao Diffserv no NS	77
5.2.2 – Geração de tráfego VoIP no NS	79
5.3 – Métricas	80
5.4 – Modelo da Simulação deste Trabalho	80
5.5 – Condução dos Experimentos e Resultados Obtidos	86
5.5.1 – Primeiro cenário – G.711/G.729 de 1 a 5 quadros por pacote	87
5.5.2 – Segundo Cenário – G.729 de 1 a 10 quadros por pacote	90
5.6 – Validação dos Resultados Obtidos	94
5.7 – Conclusões deste Capítulo	97
CAPÍTULO VI	100
CONCLUSÕES, CONTRIBUIÇÕES DESTE TRABALHO E SUGESTÕES PARA FUTUROS TRABALHOS	100
6.1 – Considerações Finais	100
6.2 – Contribuições deste Trabalho	101
6.3 – Sugestões para Futuros Trabalhos	101
REFERÊNCIAS BIBLIOGRÁFICAS	102

Lista de Figuras

Figura 2. 1 – Camadas RM-OSI e protocolos VoIP.....	11
Figura 2. 2 – Convergência das redes de voz e de dados.	11
Figura 2. 3 – Telefonia IP.....	12
Figura 2. 4 – Relação entre os padrões e protocolos de VoIP.....	14
Figura 2. 5 – Arquitetura TCP/IP e as camadas de VoIP.....	15
Figura 2. 6 – Pacote IP para transmissão.....	15
Figura 3. 1 – Elementos da arquitetura H.323.....	21
Figura 3. 2 – As três fases da conexão H.323.....	25
Figura 3. 3 – Operação do servidor <i>proxy</i> SIP.	30
Figura 3. 4 – Operação de redirecionamento SIP.....	31
Figura 3. 5 – Exemplo de estabelecimento de uma chamada SIP.....	33
Figura 3. 6 – Formato da mensagem SIP.....	34
Figura 3. 7 – Interoperação entre as arquiteturas H.323 e SIP.....	38
Figura 3. 8 – Tráfego de sinalização e mídia em um serviço interurbano.....	39
Figura 3. 9 – Interoperação entre a arquitetura <i>Softswitch</i> e a RTPC.....	41
Figura 3. 10 – Síntese de um protocolo controlador de <i>gateway</i> de mídia.....	42
Figura 4. 1 – Diagrama de blocos da codificação de um sistema de comunicação.....	48
Figura 4. 2 – Diagrama conceitual de blocos do modelo de síntese CELP.....	51
Figura 4. 3 – Relação da largura de faixa e da qualidade de voz dos codificadores.	52
Figura 4. 4 – Escala de atrasos e os efeitos que estes produzem na qualidade de voz.....	58
Figura 4. 5 – RSVP utilizado para reserva de recursos.	69
Figura 4. 6 – Estrutura do quadro VoIP na Ethernet.....	73
Figura 4. 7 – Pedido e resposta do COPS.....	74
Figura 5. 1 – Topologia da simulação utilizada neste trabalho.....	82
Figura 5. 2 – Vazão média obtida no primeiro cenário.....	88
Figura 5. 3 – Média do atraso fim-a-fim dos pacotes de voz do primeiro cenário.....	89
Figura 5. 4 – Média da variação do atraso do primeiro cenário.....	90
Figura 5. 5 – Vazão média obtida no segundo cenário.....	91
Figura 5. 6 – Média do atraso fim-a-fim dos pacotes de voz do segundo cenário.....	92

Figura 5. 7 – Média da variação do atraso do segundo cenário.....	93
Figura 5. 8 – Percentual de pacotes com latência maior que 300 ms do segundo cenário.....	94

Lista de Tabelas

Tabela 2. 1 – Comparação entre os serviços de voz por comutação de circuitos e VoIP	10
Tabela 3. 1 – Principais componentes do H.323 para VoIP.....	20
Tabela 3. 2 – Pilha de protocolos do H.323.....	20
Tabela 3. 3 – Mensagem dos tipos <i>request</i> e <i>response</i>	34
Tabela 4. 1 – Tipos de quadros existentes no G.723.1.....	50
Tabela 4. 2 – Duração do quadro e tamanho do <i>payload</i> por Codificador de Voz/Algoritmo.....	51
Tabela 4. 3 – Pontuação média de opinião.....	53
Tabela 4. 4 – Cálculo da vazão necessária.....	56
Tabela 4. 5 – Vazão necessária para o G.729 até 12 quadros por pacote.....	57
Tabela 4. 6 – Vazão necessária para o G.723 até 12 quadros por pacote.....	57
Tabela 4. 7 – Algoritmo de compressão e a classificação MOS.....	62
Tabela 5. 1 – Fontes de tráfego utilizadas no primeiro cenário.....	84
Tabela 5. 2 – Mapa dos valores utilizados na topologia do primeiro cenário.....	84
Tabela 5. 3 – Fontes de tráfego utilizadas no segundo cenário.....	85
Tabela 5. 4 – Mapa dos valores utilizados no segundo cenário.....	85
Tabela 5. 5 – Parâmetros utilizados no primeiro cenário da simulação.....	87
Tabela 5. 6 – Parâmetros utilizados no segundo cenário da simulação.....	91
Tabela 5. 7 – Intervalo de confiança da vazão média fim-a-fim do primeiro cenário.....	95
Tabela 5. 8 – Intervalo de confiança do atraso fim-a-fim do primeiro cenário.....	95
Tabela 5. 9 – Intervalo de confiança do jitter fim-a-fim do primeiro cenário.....	96
Tabela 5. 10 – Intervalo de confiança da vazão fim-a-fim do segundo cenário.....	96
Tabela 5. 11 – Intervalo de confiança do atraso fim-a-fim do segundo cenário.....	97
Tabela 5. 12 – Intervalo de confiança do jitter fim-a-fim do segundo cenário.....	97

CAPÍTULO I

INTRODUÇÃO

1.1 – Motivação

Pode-se dizer que voz sobre IP é a sensação do momento em telefonia e redes convergentes, mas muito sensacionalismo tem sido criado em torno desse assunto. Voz sobre IP, ou VoIP, nada mais é do que o transporte da voz sobre redes que utilizam o protocolo IP (*Internet Protocol*).

O IP é um protocolo de pacotes, ou seja, a informação é “quebrada” em pequenos pedaços e transmitida ao destino. O VoIP é um tipo de voz sobre redes de pacotes e, como esperado, não é o único. Existem outros tipos tais como: Voz sobre *Frame Relay* (*Voice over Frame Relay* (VoFR)) e Voz sobre ATM (*Voice over ATM* (VoATM)).

Voz sobre *Frame Relay* já é uma realidade e vem sendo utilizada por alguns anos, tanto nas redes corporativas privadas como nas redes das grandes operadoras. No ambiente corporativo a forma de ligação mais comumente encontrada é um sistema *Private Automatic Branch eXchange* (PABX), ou Centrais Privadas de Comutação Telefônica (CPCT), interligada por uma linha de pacotes de protocolo *Frame Relay* através de um *gateway* denominado *Voz-Frad* e, ao mesmo tempo, sendo interligado a uma Rede Telefônica Pública Comutada (RTPC). A central telefônica deve ser capaz de realizar o gerenciamento do tráfego de transbordo, isto é, quando a rede privada *Frame Relay* (FR) estiver congestionada o PABX

deve redirecionar a ligação para a rede pública [1, 2]. VoATM é uma alternativa poderosa, todavia é uma solução bastante cara, talvez pela utilização de enlaces de altas taxas de transmissão e comutadores sofisticados de alto poder de processamento.

Entretanto, o IP é o protocolo mais amplamente utilizado, além de ser uma tecnologia barata, são as suas vantagens em relação às soluções de VoATM e VoFR, e devido a sua onipresença, uma vez que o IP é encontrado na maioria das redes de computadores.

A primeira vista, a telefonia de VoIP pode ser entendida como um paradoxo, certo de que o IP foi originalmente concebido para redes de dados e não possui garantia de entrega. Baseado na tecnologia de aplicações do “melhor esforço” (*best-effort*), ele está sujeito a erros, atrasos e variações de atraso (*jitter*).

Considerando a mesma linha de pensamento, as redes convencionais de comutação de circuitos têm atendido perfeitamente os usuários, com raras exceções, com excelente qualidade e confiabilidade extremamente alta em relação à disponibilidade do sistema, mas com alta vulnerabilidade quanto à privacidade das ligações, sujeita aos conhecidos “grampos” telefônicos de fácil acesso.

Pode-se pensar, portanto, ser uma perda de tempo a utilização da telefonia VoIP, uma vez que o serviço de transporte de voz já é oferecido pelas operadoras com ótima qualidade. A rede de telefonia de comutação de circuitos não permite a alteração do codificador de voz (*voice coder*) sob pena de ter de alterar toda a estrutura da rede telefônica mundial, uma vez definido o codificador de voz na origem, a rede de destino deverá ter a capacidade de decodificar o sinal. Uma vantagem muito importante do VoIP é a integração de várias aplicações de voz e dados, visto ser cada vez maior o interesse dos usuários nesse tipo de serviço como utilizado no *web browser*: clique para falar (“*click to talk*”), geralmente encontrado nas redes *Intranet* e *Internet* [3].

Algumas pessoas confundem voz sobre IP com voz sobre a *Internet*. A *Internet* é um conjunto de redes interconectadas pelo protocolo IP. Qualquer pessoa, em qualquer lugar,

pode acessar a *Internet* para uma enorme variedade de aplicações, fazendo com que o usuário fique a mercê da forma de utilização da rede por outros usuários, uma vez que a *Internet* não oferece garantia de utilização de recursos da rede. Até que as políticas de qualidade de serviço da *Internet* sejam alteradas de modo a oferecer garantia de entrega dos pacotes com controle do atraso e acessos com maior largura de faixa, a VoIP pela *Internet* não será competitiva com a telefonia convencional. Enquanto isso, a VoIP está direcionada às redes corporativas, nas quais a largura de faixa e o acesso são melhores administrados. Mesmo assim, muitos provedores de serviços de *Internet* (*Internet Service Provider* (ISP)) têm agregado mais esse valor aos seus serviços, impulsionado pela economia para os usuários finais em utilizar uma rede integrada, como pode ser confirmado por diversos serviços oferecidos por provedores de Internet. Conforme divulgado pela Teleco [4], já são dezessete (17) os prestadores de serviços de VoIP via Internet no Brasil.

Todavia, mesmo nas redes corporativas, nem tudo está definido. Existem grandes desafios para que o VoIP seja aceitável na corrida com a telefonia tradicional de comutação de circuitos, de forma que possa oferecer a mesma qualidade de voz e confiabilidade.

Qualidade de voz é o fator principal, tendo como ponto de partida a compreensão da fala pelo usuário e a comparação com o sistema telefônico tradicional. O protocolo IP não oferece nenhuma garantia e os pacotes podem ser entregues fora da seqüência que foi transmitida originalmente. Para cobrir as deficiências do IP outros protocolos foram desenvolvidos, tal como o protocolo de controle de transmissão (*Transmission Control Protocol* (TCP)), no caso de falhas ou quando na confirmação do recebimento do pacote não fosse confirmada.

Para que uma comunicação de voz sobre redes de pacotes tenha qualidade da fala satisfatória, deve ter alguns parâmetros bem definidos com a finalidade de evitar diversos efeitos indesejáveis aos interlocutores, que de alguma forma prejudicam a comunicação. Esses parâmetros são: em termos de qualidade de serviço (*Quality of Service* (QoS)) na rede IP

(como a utilização da largura de faixa, atraso, *jitter* (ou variação de atraso) e perda de pacotes), técnicas de codificação de voz (como a compressão, controle e supressão de silêncio, e controle de eco) e a interoperação com a rede pública (como o suporte às centrais telefônicas tradicionais e sinalização) [5] .

Na codificação da voz são analisados apenas os aspectos do som da voz e as frequências perceptíveis ao ouvido humano, enquanto que na codificação da fala são analisados diversos aspectos de interlocução quando dois ou mais interlocutores estão falando. Por exemplo, quando duas ou mais pessoas falam ao mesmo tempo, quando apenas uma pessoa fala ou quando ninguém fala. Nesses casos é necessário que o decodificador crie um ruído de ambiente (criando uma sensação de conforto ao ouvinte) e faça o casamento de impedâncias de forma a eliminar o eco. É necessário para uma boa qualidade da fala que o codificador da fala (*speech coder*) suprima o eco, minimize a perda de pacotes, otimize a utilização da largura de faixa e tenha boa qualidade de voz e que a voz não soe como “metálica” ou sintética (robotizada). Portanto, é fundamental uma análise criteriosa na escolha do codificador da fala [3].

A QoS é limitada pelos protocolos TCP/IP. Sendo necessários protocolos adicionais e/ou métodos suplementares para um controle mais eficiente da QoS. Além da recomendação H.323, outros protocolos estão sendo considerados no contexto, tais como o Protocolo de Reserva de Recursos (*Resource reSerVation Protocol* (RSVP)), a versão 6 do IP (IPv6), o Protocolo de Transporte de Tempo-Real (*Real-time Transport Protocol* (RTP)), a Arquitetura de Serviços Integrados (*Integrated Services Architecture* (ISA)), a Arquitetura *Softswitch*, o Protocolo de Inicialização de Sessão (*Session Initiation Protocol* (SIP)), os Serviços Diferenciados (*Diffserv*), etc. [5, 6].

Os padrões de arquitetura para o funcionamento do VoIP estão direcionados para o H.323, o SIP, o Protocolo de Controle de *Gateways* de Mídia (*Media Gateway Controller Protocol* (MGCP)) e o MEGACO/H.248. A opinião geral é que o H.323 é o mais complexo

deles. O SIP, desenvolvido pela Força Tarefa de Engenharia da Internet (*Internet Engineering Task Force* (IETF)), é a “estrela” do momento, por ser mais flexível e mais fácil, muitos dizem que logo estará substituindo o H.323, que é o líder de utilização no mercado de VoIP. O MEGACO/H.248, ou MEGACO como é conhecido dentro da IETF ou H.248 [7] dentro do ITU-T, é a convergência entre os padrões desenvolvida em conjunto pelo IETF e o Grupo de Estudos 16 do ITU-T. Como esperado, tendo sido desenvolvidos pelo IETF, o MGCP e o MEGACO são muito parecidos [6].

Certamente, como tais padrões tratam das sinalizações das chamadas, devem fazer interoperação com os padrões de sinalização ISDN Q.931 e SS7 da telefonia tradicional. Conforme citado anteriormente neste capítulo, a VoIP deve interoperar com a telefonia tradicional, realizando todos os serviços atuais e agregando outros serviços, integrando dados, para que desperte o interesse e seja atrativa ao usuário final [6, 8].

O protocolo SIP é, geralmente, utilizado em conjunto com o protocolo MGCP ou protocolo MEGACO. Neste conjunto, o SIP (utilizado como protocolo de sinalização) com o protocolo MGCP ou o protocolo MEGACO, pode ser chamado de *Softswitch*. Este termo é porque muitas funções manipuladas pela telefonia tradicional de comutação de circuitos são emuladas por sistemas de *software*. Nessa arquitetura, o SIP pode ser utilizado com outros protocolos, como o protocolo de fluxo em tempo real (*Real-Time Streaming Protocol* (RTSP)) para controle dos servidores de mídia, o protocolo de descrição de sessão (*Session Description Protocol* (SDP)) para definição da seção de multimídia, os protocolos RTP/RTCP para o transporte fim-a-fim dos dados de voz e outros protocolos de QoS.

A VoIP não é tão trivial quanto parece. É necessário que o especialista procure pesquisar quatro temas importantes: Redes de Pacotes IP (com ou sem fio), técnicas de codificação da voz, técnicas disponíveis de QoS em redes IP e o Sistema de Telefonia Fixa Comutada (STFC). Nesse sentido, o pesquisador estará abrangendo grande parte do tema de VoIP.

Este capítulo apresenta um levantamento bibliográfico dos trabalhos desenvolvidos em qualidade de serviço em VoIP, a proposta deste trabalho e a estrutura desta dissertação. Finalmente, são realizadas as considerações finais deste capítulo.

1.2 – Trabalhos Relacionados

Oouchi e outros [9] desenvolveram um estudo do comprimento apropriado dos dados de voz em pacotes IP para ajuste de sistemas de uma rede VoIP. Este estudo foi baseado em medidas de um ambiente real e examinou a qualidade da voz enquanto variava o comprimento dos dados de voz dos pacotes IP, sob várias condições da rede. A análise da qualidade da voz foi implementada utilizando o PSQM+ entre dois *gateways* de VoIP. O tráfego de voz foi gerado por um banco de dados de oito arquivos de amostras de voz, contendo amostras de voz de 4 homens e 4 mulheres, que falaram em japonês.

Zheng e outros [10] analisaram as características do atraso da rede e o *jitter* de atraso e seus efeitos em voz sobre IP (VoIP) utilizando a solução de enfileiramento transiente. O principal objetivo da análise foi centralizado no tráfego de voz por um roteador com tráfego de fundo em rajadas na rede, ignorando os atrasos de codificação, decodificação e de empacotamento da voz. Nesse trabalho foi demonstrado que quando aumenta a carga de tráfego, a “explosividade” do tráfego e os comprimentos dos estados **on/off**, o comportamento do *jitter* piora significativamente, reduzindo a qualidade de voz.

Sun e outros [11] analisaram o impacto da posição da perda de pacotes na qualidade de voz em três (3) codificadores diferentes: G.729, G.723.1 e AMR, usando métodos objetivos da qualidade de voz percebida utilizando PSQM+, MNB e EMBSD. A posição da perda de pacotes refere-se a localização da perda, se no início, no meio ou no final de um segmento sonoro ou rajada de voz. Os resultados mostraram que a posição da perda de pacotes em relação à posição de voz teve um efeito severo na qualidade de voz percebida e este efeito causou maior impacto no princípio dos segmentos sonoros.

Outros trabalhos desenvolvidos relacionados ao tema de QoS em VoIP podem ser encontrados em Melo e outros [12], Fernandes e outros [13] e Neto e outros [14].

1.3 – Proposta deste Trabalho

O principal objetivo deste trabalho é a realização do estudo do impacto do comprimento dos dados de voz do pacote IP em um domínio de serviços diferenciados (*Diffserv*), tendo como finalidade a *qualidade da voz*, medida pela largura de faixa utilizada, atraso e variação do atraso dos pacotes. Esta análise será realizada através dos resultados de uma simulação de uma topologia de rede com fontes de VoIP com suporte de *Diffserv*.

Após os resultados obtidos, é também objetivo deste trabalho propor uma solução que possa minimizar os efeitos do comprimento dos dados da voz em uma rede VoIP conforme as condições de QoS da rede.

1.4 – Estrutura desta Dissertação

O Capítulo II descreve os principais conceitos básicos do transporte de voz sobre redes de pacotes IP, seus elementos e dispositivos que compõem a rede de VoIP, alguns exemplos de cenários de VoIP, e uma breve introdução à sinalização de voz sobre IP. Finalmente, são realizadas considerações finais desse capítulo.

O Capítulo III descreve os conceitos básicos das arquiteturas H.323, SIP e Softswitch e seus elementos, bem como os dois protocolos de sinalização entre *gateways* e controladores de mídia mais utilizados, o MGCP e o MEGACO/H.248, e suas principais características. Finalmente, são realizadas as considerações finais desse capítulo.

O Capítulo IV apresenta os codificadores de voz mais utilizados na aplicação de VoIP, as técnicas de medição, os fatores que influenciam na qualidade da voz, os conceitos fundamentais de qualidade de serviço (*Quality of Service* (QoS)) em uma rede VoIP, assim

como apresenta os principais mecanismos disponíveis atualmente para a sua implementação. Finalmente, são realizadas considerações finais desse capítulo.

No Capítulo V é apresentado o simulador NS, são descritos o conjunto de considerações e a topologia adotada na simulação realizada neste trabalho, e são mostrados os resultados obtidos com os testes realizados. Finalmente, são realizadas conclusões sobre os resultados obtidos nesses testes.

E finalmente, no Capítulo VI são apresentadas as conclusões, as contribuições desta dissertação e as sugestões para futuros trabalhos.

1.5 – Considerações Finais deste Capítulo

Este capítulo apresentou uma introdução aos temas relacionados ao transporte da voz sobre redes IP, um levantamento bibliográfico dos trabalhos desenvolvidos em qualidade de serviço em VoIP e a proposta deste trabalho.

CAPÍTULO II

VOZ SOBRE IP

2.1 – Introdução

Em síntese, o conceito de convergência das redes aponta para o surgimento de uma única rede que servirá de suporte para o transporte de diversos tipos de mídias ou, mais especificamente, como abordada neste trabalho, de uma rede única capaz de prover serviços integrados de comunicação de dados e voz.

Considerando uma empresa na qual, por exemplo, um determinado funcionário de Goiânia desejar falar com um outro funcionário de Recife, é possível fazê-lo sem utilizar a rede de telefonia pública, utilizando a rede corporativa da empresa composta de linhas dedicadas (ou conhecidas como prestação de Serviço por Linha Dedicada para sinais Digitais (SLDD)), devidamente configuradas em VoIP e QoS. Para que isto seja possível, basta que os PABX das unidades de origem e de destino da chamada estejam interligados à rede corporativa de dados existente na empresa, que pode-se obter o tom de discagem, automaticamente ou por um código especial [1, 2]. A chamada será roteada pela rede de dados da empresa até o seu destino, que neste exemplo é Recife.

O investimento inicial na rede corporativa poderá ser alto, mas em pouco tempo ele será pago pela economia em ligações telefônicas urbanas, interurbanas ou mesmo internacionais.

Voz sobre IP (VoIP) [5, 15] é uma das alternativas de convergência das redes de dados e voz. Também existem outras, tais como: voz sobre Frame Relay (VoFR) [1, 15], voz sobre ATM (VoATM) [1, 15] e voz sobre MPLS (VoMPLS) [16]. Todas elas possuem vantagens e desvantagens. Não foi definida ainda uma tecnologia de convergência de redes que pudesse ser eleita como perfeita para a integração plena dos serviços de dados e voz. O projetista deverá pesquisar muito sobre qual caminho tomar, analisando todo o contexto da empresa e das redes de telefonia e de dados existentes em seu ambiente.

A Tabela 2.1 resume as vantagens e as desvantagens dos serviços de comutação de circuitos e VoIP [5, 6].

Tabela 2. 1– Comparação entre os serviços de voz por comutação de circuitos e VoIP

Serviço	Vantagens	Desvantagens
comutação de circuitos (telefonia tradicional)	<ul style="list-style-type: none"> – sem atrasos. – alta disponibilidade – alta qualidade de voz 	<ul style="list-style-type: none"> – alto custo em equipamentos – tecnologia proprietária de difícil atualização – largura de faixa utilizada de 64 kbps – mudança de codificador de voz, como padrão o G.711 (PCM) – alta vulnerabilidade a “grampos” telefônicos
VoIP	<ul style="list-style-type: none"> – baixo custo em equipamentos – possibilidade de adequação de qual o codificador deve ser utilizado: se G.723.1, G.729, ou G.728, etc. – integração de aplicativos de voz e de dados – requisitos de largura de faixa mais baixa – disponibilidade difundida do IP – pode ter alta qualidade de voz dependendo do codificador de voz utilizado 	<ul style="list-style-type: none"> – sujeito a atrasos, ecos, perda de pacotes e <i>jitter</i>

Neste capítulo descrevem-se os principais conceitos básicos do transporte de voz sobre redes de pacotes IP, seus elementos e dispositivos que compõem a rede de VoIP, alguns exemplos de cenários de VoIP, e uma breve introdução à sinalização de voz sobre IP. Finalmente, são realizadas considerações finais deste capítulo.

2.2 – Integração de Voz e de Dados Sobre IP

Pode-se formar uma idéia do funcionamento e dos protocolos associados ao serviço de voz sobre IP em relação ao RM-OSI, como mostrado na Figura 2.1 [1].

Camada	Serviços VoIP
aplicação	aplicações de comunicação /...
apresentação	G.729/G.723.1/G.711/G.728/...
sessão	H.323/SIP/MGCP
transporte	RTP/RTCP/UDP
rede	IP
enlace	PPP/Frame Relay/ATM/Ethernet
física	rádio/fibra óptica/LPCD/par trançado

Figura 2. 1 – Camadas RM-OSI e protocolos VoIP.

2.3 – Exemplos de Cenários de VoIP

A Figura 2.2 mostra um cenário onde o Site 1 é constituído de um PABX que pode ser de uma filial de uma empresa e o Site 2 outra filial. Os PABX das filiais estão interligados pela nuvem PSTN e as redes locais de cada site estão interligadas por uma conexão WAN. Nessa figura a função do gateway é de interligação do sistema telefônico à rede de dados.

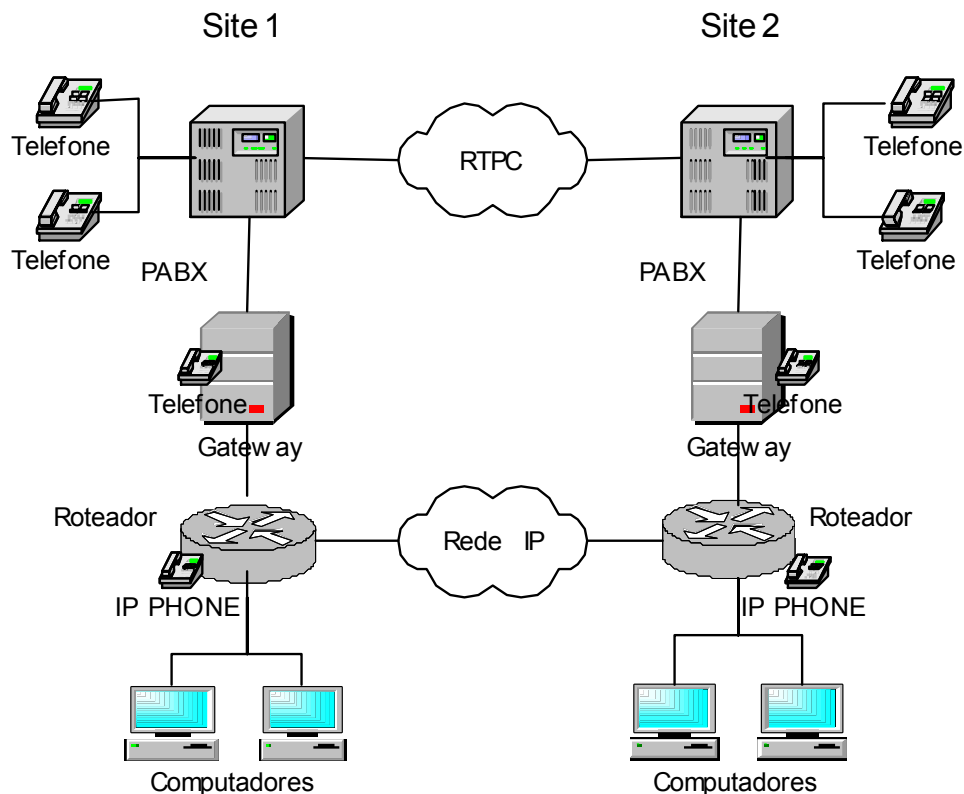


Figura 2. 2 – Convergência das redes de voz e de dados.

Atualmente, o cenário mostrado na Figura 2.2 é o mais encontrado no Brasil. As grandes empresas brasileiras estão optando por esse tipo de solução com o intuito de aproveitar os equipamentos de PABX e dos roteadores já instalados, aproveitando a largura de faixa das conexões WAN existentes.

A Figura 2.3 mostra um cenário de VoIP, onde o PABX pode ser totalmente substituído pelo *Switch Router* (ou *VoIP-Gateway*), como mostrado no site 1. Este equipamento deve realizar todas as funções do PABX, além do roteamento e da comutação dos pacotes de voz e de dados.

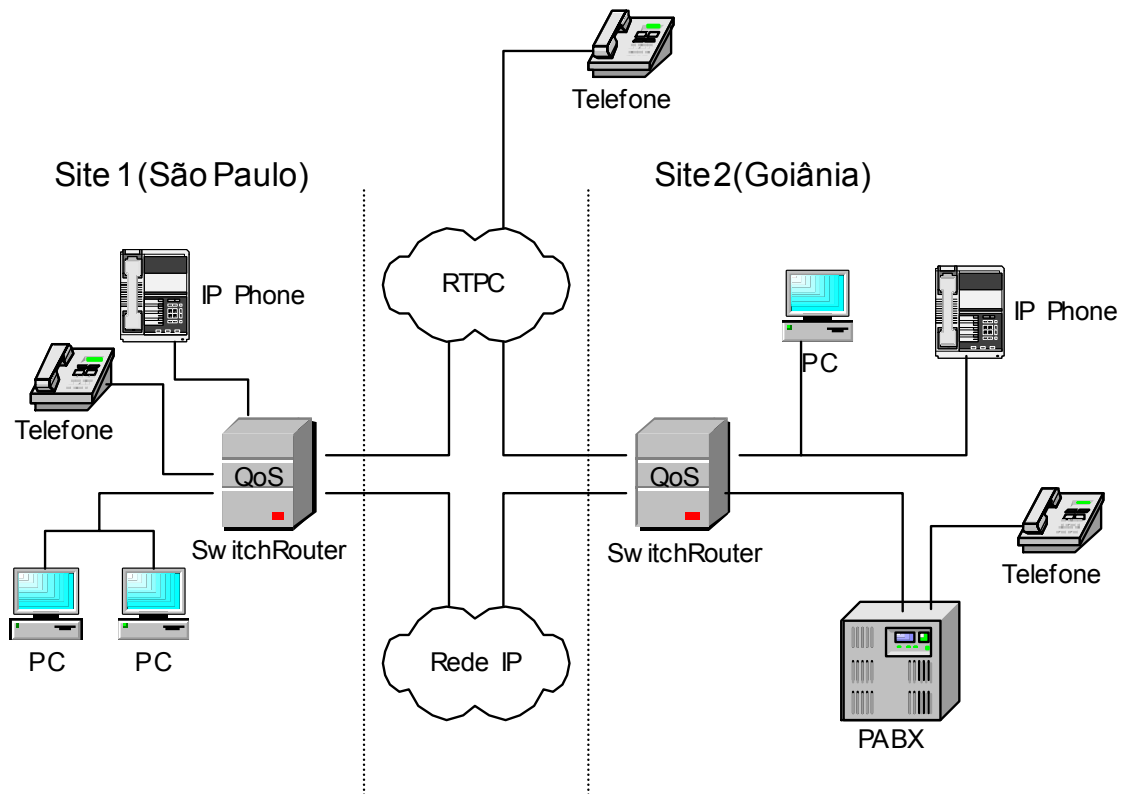


Figura 2. 3 – Telefonia IP.

Outros cenários com interconexões podem ser encontrados em [3] e [6], bem como outros serviços e funções podem ser realizados, tais como *Call Center*.

Analisando as Figuras 2.2 e 2.3, pode-se ter 3 tipos de arquiteturas básicas de implementação de cenários VoIP:

Arquitetura PC-a-PC: nesta arquitetura dois computadores providos de recursos multimídia, conectados a uma LAN ou pela Rede de Telefonia Pública Comutada (RTPC), a um provedor de serviços Internet, comunicam-se para a troca de sinais de voz. O tratamento do sinal de voz, tais como a amostragem, a compressão e o empacotamento, é realizado nos próprios computadores, sendo a chamada de voz estabelecida com base no endereço IP do receptor ou por um serviço de resolução de nomes. A arquitetura PC-a-PC possui uma variante onde o PC pode ser substituído por um equipamento com capacidade de codificação de voz, implementação do protocolo IP e sistema de sinalização de chamada (H.323 ou SIP). Enfim, os próprios computadores são responsáveis pela sinalização e controle das chamadas.

Arquitetura com *gateway*: nesta arquitetura, mostrada na Figura 2.3, um telefone padrão pode ser utilizado para gerar e receber chamadas telefônicas por uma rede (LAN ou WAN). O *gateway* deve estar preparado com os protocolos de sinalização de chamadas (H.323 ou SIP), de comunicação entre *gateways* (MGCP ou MEGACO), bem como dos sistemas de codificação e decodificação de voz (G.723.1, G.729 ou outro) e técnicas de empacotamento e desempacotamento da voz.

Arquiteturas híbridas: utilizam as arquiteturas PC-a-PC e com *gateway*.

2.4 – Padrões de VoIP

Os protocolos padrões da comunicação de VoIP podem ser divididos em [1, 6]:

a) Protocolos de sinalização para controle de chamada

- **H.323-Packet Based Multimedia Communications Systems** – utilizado por telefones IP, computadores, adaptadores IP, controladores de sinalização e *gateways* de estabelecimento, controle e término de chamadas. É o protocolo mais antigo e complexo e atualmente tem sido menos utilizado pelos sistemas de Telefonia IP.

- **Session Initiation Protocol (SIP)** – tem a mesma finalidade do H.323, porém é mais moderno e menos complexo. É mais utilizado nos sistemas de Telefonia IP.

b) Protocolos de sinalização para controle de Gateways

- **Media Gateway Control Protocol (MGCP)** – utilizado pelos controladores de gateways e gateways para estabelecimento, controle e término das chamadas.
- **Media Gateway Control Protocol (MEGACO)** – tem a mesma finalidade do MGCP. Ele foi desenvolvido para ser uma alternativa para esse protocolo, adequando-se também aos controladores distribuídos de gateways, aos controladores de multiponto (conferência) e às unidades interativas de resposta audível.

c) Protocolos para transporte de mídia

- **Real-Time Transport Protocol (RTP)** – protocolo permite o transporte de voz em tempo real.
- **Real-Time Transport Control Protocol (RTCP)** – protocolo responsável pelo controle do transporte de voz realizado pelo RTP.

A Figura 2.4 resume a relação entre os padrões e os protocolos de VoIP.

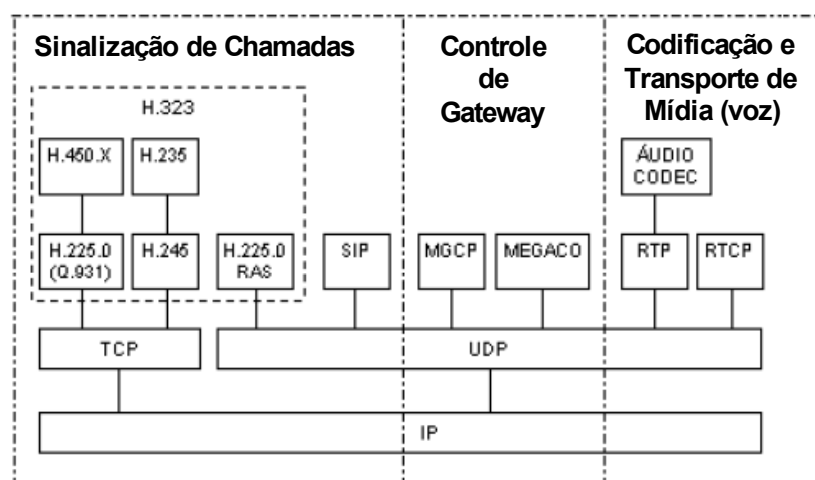


Figura 2. 4 – Relação entre os padrões e protocolos de VoIP.

2.4.1 – O empacotamento em VoIP

A Figura 2.5 mostra o empacotamento em camadas do VoIP em relação à arquitetura TCP/IP.

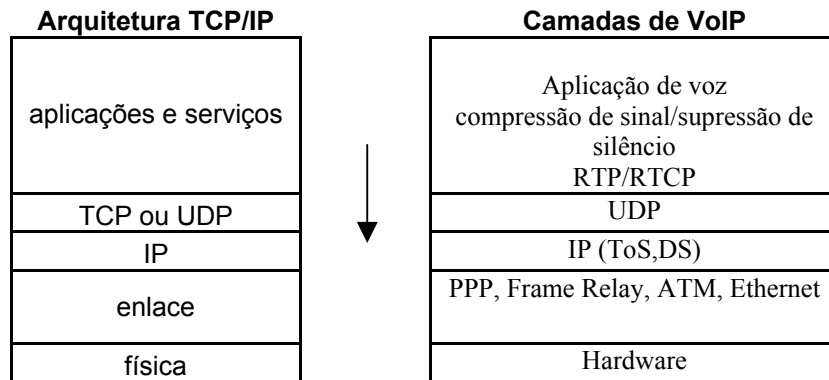


Figura 2. 5 – Arquitetura TCP/IP e as camadas de VoIP.

Pode-se observar na Figura 2.5 que o pacote (cabeçalho + *payload*) criado por uma camada passa a ser o campo de *payload* da próxima camada. O pacote a ser transmitido fica no formato da Figura 2.6 [1].



Figura 2. 6 – Pacote IP para transmissão.

2.4.2 – Sinais DTMF

O funcionamento básico de voz sobre IP resume-se na conversão dos quadros de voz em uma série de pacotes e na transmissão desses pacotes por uma rede IP, reconstruindo no destino o sinal de voz original.

Todos os dispositivos de interligação VoIP devem estar preparados e/ou possibilitar diversas configurações. Assim, desses diversos cenários podem ocorrer diversos tipos de estabelecimento de chamadas telefônicas em um ambiente de voz sobre redes de pacotes, tais como: internos ou externos à empresa e ou interurbanos ou não.

O primeiro tom de discagem é um sinal pausado com interrupções contínuas, onde o sistema espera que o próximo dígito seja discado. Se o sistema não estiver configurado para o roteamento automático das chamadas via rede corporativa, faz-se necessário que o usuário digite um código para obter o tom de discagem do *gateway* de voz.

Na transmissão dos algarismos, sinalização DTMF, deve-se considerar o seguinte [3]:

- Os tons, sinais e dígitos DTMF podem ser transmitidos antes da inicialização ou após a chamada ser estabelecida, e;
- Os codificadores de voz são preparados e aperfeiçoados para a compressão e o tratamento de voz humana. Atualmente, eles são tão sofisticados atualmente que alcançaram eficiência de largura de faixa sem perder a qualidade. Entretanto, não são preparados para a codificação de tons, sinais e dígitos DTMF.

O codificador G.711 PCM [17] suporta muito bem os tons, sinais e dígitos DTMF. E ele é utilizado na telefonia tradicional. Entretanto, já os codificadores G.723.1 [18] e G.729 [19] não reproduzem fielmente tais sinais. As técnicas utilizadas para o transporte de tons, de sinais e de dígitos DTMF, podem ser verificadas em [3, 20].

O esboço de Internet com o título “*RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals*” [21] foi preparado para direcionar o tema de transporte de tons dentro do fluxo RTP. Para suportar o envio de pacotes RTP com identificadores para os tons individuais, o esboço especifica uma grande quantidade de tons e eventos (tais como: dígitos DTMF, tom de ocupado, tom de chamada, etc.) e provê um número para cada um deles.

Os formatos da carga útil RTP para chamadas eventuais de tons, sinais e dígitos DTMF podem ser verificados em [21].

Collins [3] faz uma observação que um *gateway* que detecta um tom deve enviar um pacote de evento assim que o tom é reconhecido e a cada 50 ms enquanto durar o tom. O

gateway deve enviar pelo menos três pacotes de eventos, ainda que o tom seja extremamente breve, só para ter certeza que a outra extremidade foi informada da ocorrência do evento, até no caso de perda de pacotes. Quando vários pacotes são enviados para o mesmo tom, o *timestamp* RTP de cada pacote deve ser o mesmo, indicando o momento que o tom foi detectado, com o campo de duração acrescentado em cada pacote [3].

2.5 – Dispositivos de VoIP

Relembrando o que foi visto até agora nesta dissertação pode-se afirmar com toda segurança que o dispositivo mais importante na rede VoIP é o *gateway*, que implementa basicamente toda a tecnologia de VoIP. Devido a sua onipresença na rede VoIP o *gateway* poderá ter múltiplas funções de interoperabilidade. Ele também pode ser responsável pelas seguintes funções:

- **Verificação:** quando o usuário informa os dígitos do terminal que deseja realizar uma chamada, esses dígitos são recebidos pelo *gateway* que os converte no endereço IP de destino. Esta conversão é baseada na verificação de uma tabela que contém o plano de numeração utilizado e o endereço do IP correspondente;
- **Interface analógica:** as interfaces analógicas mais utilizadas são: E&M, FXO e FXS [2];
- **Interface digital:** as interfaces digitais mais utilizadas são: E&M digital, E1 e ISDN [2];
- **Interface para a rede *Ethernet*:** utilizadas para a conexão do *gateway* à rede IP;
- **Codificador de Voz e DSP integrados:** utilizados para a compressão e a codificação de voz, quando integrados ao *gateway*, e;
- **Suporte à qualidade de voz:** suporte aos algoritmos de compressão, supressão de eco, regeneração de ruído de conforto, cancelamento de eco e comutação de voz.

O cenário da Figura 2.3 deste capítulo ainda mostra o *IP Phone*, outro dispositivo de VoIP, que é ligado diretamente ao *backbone* da rede local, geralmente utilizando uma conexão RJ45. O *IP Phone* é um equipamento que possui uma interface de rede que faz a conversão do número de lista do telefone para o endereço IP, e vice e versa. Ele não é um equipamento barato, tal como o aparelho telefônico convencional e tem diversos outros serviços e funções que aquele não tem, tais como: amostragem, quantização, codificação e empacotamento da voz, transmissão e as características de interface de rede local.

Vários exemplos de cenários de VoIP podem ser encontrados nas referências [1, 3, 5, 6].

2.6 – Considerações Finais deste Capítulo

Este capítulo descreveu os principais conceitos do transporte de voz sobre redes de pacotes IP, bem como ilustrou alguns cenários possíveis de VoIP. Tais conceitos formam o contexto de funcionamento das redes de VoIP, podendo, a partir deste ponto, iniciar a análise dos padrões de VoIP, nos próximos capítulos deste trabalho.

CAPÍTULO III

PROTOCOLOS DE SINALIZAÇÃO DE CHAMADA E DE CONTROLE DE *GATEWAYS*

3.1 – Introdução

No Capítulo 2 desta dissertação foram citadas as duas arquiteturas mais utilizadas de sinalização de chamadas em VoIP: a recomendação H.323 e o protocolo SIP. O H.323 é uma das recomendações do ITU-T que especifica uma arquitetura e a metodologia global que incorpora várias outras recomendações para áudio, vídeo e comunicações de dados sobre redes baseadas em pacotes. Atualmente, a arquitetura H.323 é o padrão desenvolvido mais extensivamente para redes de VoIP, permitindo que os produtos de multimídia e as aplicações de múltiplos fornecedores possam interoperar-se e comunicar-se, sem que os usuários se preocupem com a compatibilidade [1].

A recomendação H.323, brevemente comentada nesta dissertação, refere-se à versão 2 [22], com ênfase na sinalização das chamadas telefônicas das versões 2, 3 e 4 [23].

O protocolo SIP [24] foi desenvolvido como parte da arquitetura global de controle e dados de multimídia do IETF. Ele foi projetado para trabalhar junto com outros protocolos do IETF, como o Protocolo de Descrição de Sessão (*Session Description Protocol* (SDP)) e o Protocolo de Fluxo de Tempo Real (*Real-Time Streaming Protocol* (RTSP)).

Este capítulo descreve os conceitos básicos das arquiteturas H.323, SIP e Softswitch e seus elementos, bem como os dois protocolos de sinalização entre *gateways* e controladores

de mídia mais utilizados, o MGCP e o MEGACO/H.248, e suas principais características. Finalmente, são realizadas as considerações finais sobre este capítulo.

3.2 – Arquitetura H.323

O H.323 é um “guarda-chuva” de especificações, comumente encontrado na literatura, uma vez que descreve um funcionamento conjunto de várias outras recomendações adicionais. A Tabela 3.1 descreve os principais componentes do H.323 para VoIP [1].

Tabela 3. 1 – Principais componentes do H.323 para VoIP.

Componente	Descrição
G.711, G.722, G.723, G.728 e G.729	Especificações para codificadores de áudio
H.225 (Call Signaling Protocols)	Protocolos de sinalização de chamadas H.225
H.245 (Control Protocol for Multimedia Communication)	Protocolo de controle para comunicações multimídia H.245
H.246 - Interworking of H-Series Media Terminals and Global Switched Telephone Networks – GSTN's	Especificações para interconexão entre terminais série H e a rede global de telefonia comutada
H.235 - Security and Encrytion for H-Series	Segurança e criptografia para terminais Série H
H.450.2-7 - Services	Serviços de transferência de chamadas, manter uma chamada em espera, etc.
RTP, RTCP e IP Security (IPsec)	Especificações para Internet

A Tabela 3.2 mostra a pilha de protocolos do H.323 [1, 3].

Tabela 3. 2 – Pilha de protocolos do H.323.

Aplicações de Áudio	Aplicações de Vídeo	Gerenciamento de controle de Terminal			
G.711 G.729 G.723.1	H.261 H.263	RTCP	H.225.0 Registro, Admissão e Status (RAS)	H.225.0 Sinalização de chamada (<i>Call Signaling</i>)	H.245/248 Controle de Sinalização (<i>Control Signaling</i>)
RTP					
UDP (inerentemente incerto)				TCP (confiável)	
Camada de rede (IP)					
Camada de enlace (FR, ATM e <i>Ethernet</i>)					
Camada física					

3.2.1 – Elementos da arquitetura H.323

Além das várias recomendações, a arquitetura H.323 envolve a definição de 4 elementos: *Terminais*, *Gateways*, *Gatekeepers* e *Multipoint Controller Units* (MCU). A extremidade (*EndPoint*) do H.323 pode ser um *PC Phone*, um *IP Phone*, um *gateway* ou um PABX. O principal objetivo do H.323 é ativar a troca de fluxos de mídia entre as extremidades [1, 6].

A Figura 3.1 mostra um cenário de componentes da arquitetura H.323 [3].

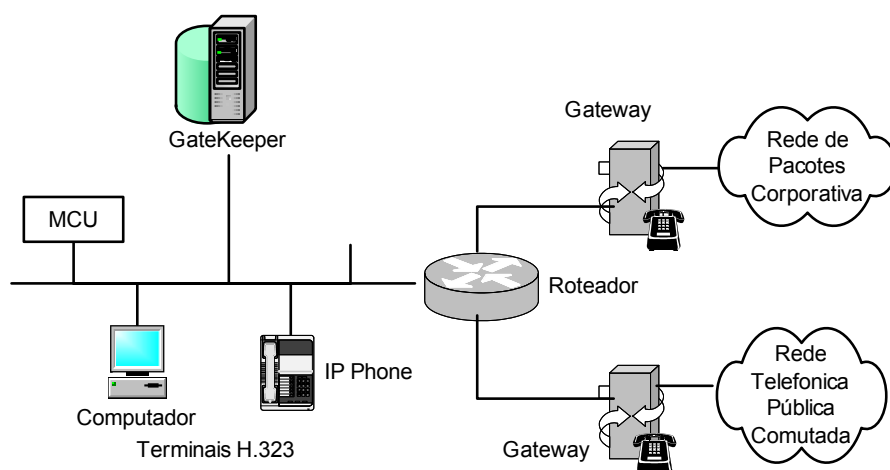


Figura 3. 1 – Elementos da arquitetura H.323.

O **Terminal** é uma extremidade que oferece comunicação com outras extremidades em tempo-real; um dispositivo de comunicação de usuário-final que suporta pelo menos um codificador de áudio (G.711, G.729 e/ou G.723.1) e/ou de vídeo (H.261 ou H.263). Os terminais suportam o protocolo H.245 para permitir o uso dos canais, a recomendação Q.931 [25] para a sinalização e a configuração de chamada, o *Registration Admission Status* (RAS) para interagir com o *Gatekeeper* e o RTP como protocolo de transporte em tempo-real.

No H.323, o *gateway* provê interoperabilidade entre as redes nativas H.323 e as redes não H.323, como a Rede de Telefonia Pública Comutada (RTPC) ou SIP. De um lado o *gateway* tem as características do H.323 de interconexão com os terminais do H.323,

operando com o protocolo H.245 para controle de sinalização quando da utilização de *Composite Gateway*, e o protocolo H.225 [26] para a sinalização de chamadas para o estabelecimento e o encerramento de chamadas e outros protocolos da recomendação H.323.

De outro lado, o *gateway* interliga com a RTPC, implementando os protocolos específicos da telefonia convencional, tais como: Redes Digitais de Serviços Integrados de Faixa Larga (RDSI-FL), Redes Digitais de Serviços Integrados de Faixa Estreita (RDSI-FE), Sistema de Sinalização número 7 (*Signaling System #7* (SS#7)), sinalização associada ao canal (*Channel-Associated Signaling* (CAS)), ou outra arquitetura como o SIP [3].

Os *gateways* também podem servir como canais de comunicação entre os terminais H.323 que não estão na mesma rede, onde a comunicação entre eles precisa passar por uma rede externa. Um *gateway* é um elemento lógico da recomendação H.323 e pode ser implementado juntamente com outros componentes H.323 em um mesmo *hardware*, como o roteador, podendo coexistir com outras funções de codificador de voz, processador digital de sinais, *gatekeeper* e MCU.

Os dois principais componentes que formam um *gateway* são:

- ***Media Gateway Controller (MGC)***: é o elemento responsável pelo gerenciamento de todo o processo de estabelecimento, modificação, monitoramento e encerramento de chamadas pelo *Media Gateway* (MG).
- ***Media Gateway (MG)***: responsável pela conversão de mídia de um tipo de rede para um outro formato exigido pelo outro lado da rede. Por exemplo, pode fazer a conversão da informação de comutação de circuitos para a mídia de comutação de pacotes em uma rede IP.

Os MG e MGC podem ser implementados de duas formas [1, 6]: *Composite Gateway*, quando o MG e o MGC coexistem em um mesmo hardware; e *Decomposed Gateway* definido

na versão 4 do H.323 quando, na implementação dos elementos MG e MGC, eles estão separados fisicamente, mas com o MGC gerenciando múltiplos MG. A comunicação entre o MGC e os MG é feita pelo protocolo de sinalização e controle MGCP ou H.248/MEGACO.

O *Gatekeeper* é um elemento opcional dentro da arquitetura H.323. Comparando com a telefonia de comutação de circuitos, suas funções são similares às do PABX, mas no H.323 ele provê funções adicionais, tais como: encaminhamento de chamadas, manutenção de chamadas em espera e conferência de chamadas. Ele atua como ponto central para todas as chamadas dentro de uma zona H.323. A zona é definida como um conjunto de terminais, *gateways*, MGC e MG controlados por um único *gatekeeper*, que pode atravessar redes ou sub-redes múltiplas, e as entidades dentro dela não precisam ser contíguas. Um *gatekeeper* provê as seguintes funções:

- Tradução de endereço (*Address Translation*): faz a tradução do nome (*alias*) para o endereço de transporte (endereço IP), atualizando uma tabela via registro de mensagens;
- Controle de admissão (*Admission control*): controla e autoriza acesso da rede local usando requisições de autorização, mensagens de confirmação e rejeição pelo protocolo RAS entre as extremidades da rede H.323 e o *gatekeeper*;
- Gerenciamento da faixa (*Bandwidth management*): controla o uso da faixa pelo protocolo RAS, como definir um número de conexões simultâneas em uma zona H.323; e;
- Gerenciamento de zona (*Zone Management*): provê o gerenciamento destas funções para os elementos participantes de uma determinada zona.

A extremidade **MCU** é também opcional no H.323, que administra as conferências de multiponto entre três ou mais terminais e/ou *gateways*. Uma unidade de controle de multipontos consiste em um controlador multiponto (*Multipoint Controller* (MC)) e pelo menos um processador multiponto (*Multipoint Processor* (MP)). Ele controla a negociação do

protocolo H.245 entre os terminais para determinar os meios que podem ser compartilhados entre os vários participantes. O MC também pode alterar a capacidade do meio quando outras extremidades entram ou deixam a conferência. Ele pode fazer parte de uma MCU ou de um *gateway*, roteador, *gatekeeper* ou terminal H.323. Para todo MC existe um MP, que opera sob controle do MC. O MP processa os vários fluxos de mídias de entrada, criando múltiplos fluxos de mídias de saída.

3.2.2 – A sinalização da arquitetura H.323

A recomendação H.225.0 [26] é composta de duas partes. A primeira é uma variante da recomendação Q.931 [25], que especifica a sinalização da camada 3 do ISDN, sendo responsável pelo estabelecimento e pelo encerramento das conexões entre *EndPoints* do H.323. A recomendação H.225 também é conhecida como sinalização de chamadas ou sinalização Q.931 [1, 3].

A segunda parte da recomendação H.225 é conhecida como sinalização de *Registration*, *Admission* e *Status*, ou ainda pelo acrônimo RAS. Esta sinalização possibilita a comunicação entre as extremidades e os *gatekeepers*, habilitando o gerenciamento das extremidades dentro da zona do *gatekeeper* [1, 3].

O H.245 é um protocolo de controle utilizado entre duas ou mais extremidades e seu propósito principal é administrar os fluxos de mídias entre os participantes de uma sessão H.323. Para administrar o controle, o H.245 inclui funções para assegurar que as mídias a serem enviadas de uma extremidade são limitadas ao conjunto de mídias que a outra extremidade pode receber e processar. A recomendação H.245 opera pelo estabelecimento de um ou mais canais lógicos entre as extremidades, que transportam fluxos de mídias entre os participantes e têm várias propriedades, tais como tipo de mídia, taxa de transferência, etc. [3].

As mensagens da arquitetura H.323 são transportadas sobre vários e diferentes tipos de canais. As mensagens do RAS são enviadas em seus próprios canais; as mensagens de sinalização de chamadas são enviadas pelos canais de sinalização de chamadas (*Call-Signaling*); e as mensagens de controle H.245 são encaminhadas em seus canais de controle. Um canal é a referência de um endereço de conexão, ou seja, um endereço IP e a porta associada ao serviço [3].

Todos os três protocolos de sinalização RAS, Q.931 [25] e H.245 [27], podem ser utilizados para estabelecer, manter e encerrar uma chamada, que se constitui de 3 fases distintas [1].

- (1) **Registration**: registro e controle de admissão para as extremidades com o *gatekeeper* para acesso à rede;
- (2) **Call Setup**: localização e re-direcionamento pela rede para chamadas estabelecidas entre duas ou mais extremidades;
- (3) **Media Negotiation**: negociação das capacidades da mídia e controle entre os componentes das extremidades.

A Figura 3.2 mostra as três fases da conexão H.323 [1, 3].

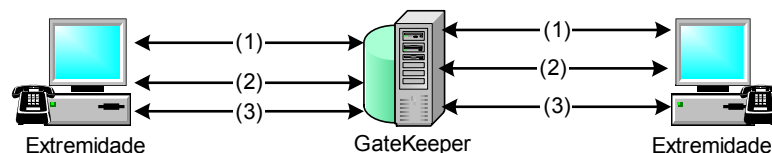


Figura 3. 2 – As três fases da conexão H.323.

3.2.3 – Sinalização RAS

O RAS é um protocolo baseado em UDP, ele é utilizado para registro, admissão de controle, requisição de troca de taxas de transmissão e controle de *status* entre as *EndPoints* e o *gatekeeper*. Como o *gatekeeper* é um elemento opcional na arquitetura H.323, o RAS também é opcional. Se o *gatekeeper* não for utilizado as suas funções precisam ser executadas pela própria extremidade. Entretanto, se o *gatekeeper* for utilizado o RAS é obrigatório [1].

A sinalização RAS é definida na recomendação H.225.0 [26], ela suporta as dez (10) funções: *Gatekeeper Discovery*, *Registration*, *Unregistration*, *Admission*, *Bandwidth Change*, *Endpoint Location*, *Disengage*, *Status*, *Resource Availability* e *Non-Standard*, que podem ser verificadas em [1, 6], bem como são descritas diversas mensagens em [3].

3.2.4 – Sinalização de chamada H.225 (*Call Signaling*)

A sinalização de chamada é utilizada entre as extremidades para ativar o estabelecimento e o encerramento das chamadas. As mensagens utilizadas são as mensagens da recomendação Q.931 [25], modificadas pela recomendação H.225.0 [26] da ITU-T. O Q.931 é um protocolo de sinalização da camada 3 para uma interface do usuário da rede ISDN, e as várias mensagens são definidas nesta recomendação. Essa recomendação aproveita várias mensagens definidas na Q.931 e as reutiliza, com algumas modificações necessárias para a utilização na arquitetura global H.323. Ela também usa mensagem da recomendação Q.932.

A recomendação H.225.0 especifica as poucas modificações que são aplicáveis às mensagens da Q.931 quando utilizadas em uma rede H.323, e várias regras relativas ao uso dos elementos de informações definidos na recomendação Q.931. Portanto, as mudanças especificadas por essa recomendação geralmente envolvem especificar que certos elementos

de informações da Q.931 são obrigatórios, proibidos ou opcionais quando utilizados em uma rede H.323.

As mensagens e exemplos do protocolo H.225 podem ser encontrados nas referências [3, 26, 28].

3.2.5 – Sinalização de controle H.245

O H.245 é o protocolo utilizado entre os participantes de uma sessão para estabelecer e controlar fluxos de mídias. Para uma chamada de voz direta de duas partes, esse protocolo cuida de assegurar que os participantes concordem nos formatos das mídias a serem enviadas e recebidas, bem como dos requisitos de largura de faixa. Para chamadas de multimídias mais complexas, esse protocolo cuida dos fluxos da multiplexação de múltiplas mídias para funções tais como a sincronização dos lábios entre o áudio e o vídeo [27].

Deve-se notar que o protocolo H.245 não é responsável pelo transporte de mídias, ele apenas controla e administra os fluxos de mídias. Por exemplo, não existe um pacote do H.245 que contenha uma amostra de voz codificada, este é um trabalho do RTP.

O protocolo H.245 não é de uso exclusivo do VoIP. Ele é um protocolo mais genérico e é utilizado para o controle de fluxos de mídia. Ele foi projetado para ser usado em um grande número de aplicações. As mensagens do H.245 [27] pertencem aos seguintes grupos:

- **Requests:** são mensagens que exigem do recipiente apresentar alguma ação e uma resposta imediata;
- **Responses:** são mensagens enviadas em resposta às mensagens *Requests*.
- **Commands:** estas mensagens exigem do recipiente apresentar alguma ação, mas nenhuma resposta explícita é necessária, e;
- **Indications:** mensagens apenas informativas.

O padrão H.245 controla fluxos de mídias utilizando os canais lógicos, que são considerados como um caminho unidirecional de mídias entre duas extremidades e possui um número definido pela entidade emissora.

Em uma conversação de duas partes, dois canais lógicos são necessários. A separação da conversação em dois caminhos habilita um terminal a enviar a voz em um formato de mídia e recebê-la em outro formato. Apesar dos canais lógicos serem unidirecionais, na análise da recomendação do H.245 é comumente encontrada a menção de canal bidirecional, que na verdade consiste em dois canais lógicos, embora eles estejam associados entre si.

No estabelecimento de um fluxo de mídia de uma extremidade até a outra, a extremidade que deseja transmitir informações estabelece um canal lógico, indicando o número do canal e as informações lógicas sobre as mídias a serem enviadas, tal como o tipo de carga útil do RTP.

As mensagens H.245 são transportadas no canal de controle H.245. Cada extremidade ou *gatekeeper* estabelece um canal de controle H.245 para cada chamada que ele está participando. O canal de controle H.245 é transportado em um canal lógico especial (canal de número 0). Este canal lógico é especial, isto é, ele não é aberto e fechado como os outros canais lógicos. Ao contrário, este canal é considerado permanentemente aberto desde que a extremidade esteja envolvida em uma chamada [27].

A recomendação H.245 é constituída de vários procedimentos que podem ser pesquisados na ITU-T [27] e [1, 3, 6, 29] descrevem vários exemplos de cenários de sinalização de chamadas com o protocolo H.323.

3.3 – Protocolo SIP

Apesar do H.323 [22] ter sido mundialmente reconhecido como o padrão para a implementação de VoIP, muitos consideram o Protocolo de Iniciação de Sessão (*Session Initiation Protocol* (SIP)) uma alternativa poderosa para o H.323. O SIP é considerado mais

flexível, mais simples, mais fácil de implementar, melhor projetado para oferecer suporte aos dispositivos inteligentes do usuário e para a implementação de características avançadas, do que o H.323. Entretanto, esta comparação refere-se à versão 1 do H.323, uma vez que a versão 2 dele é tão eficiente quanto o SIP, quando da utilização da configuração de chamada *fastStart* [6]. Deve-se considerar ainda que o H.323 já está na versão 5 [30] e muitas implementações foram adicionadas nas versões de 3 a 5, podendo aquecer em muito as discussões de escolha.

O SIP é um protocolo de sinalização que controla o estabelecimento, a configuração, a modificação e o encerramento das sessões de multimídia, como as chamadas telefônicas, e também oferece condições para o transporte de imagens. Ele oferece muito dos recursos da plataforma H.323, mas confia especificamente no protocolo IP.

O SIP pode utilizar tanto do TCP como do UDP como protocolo de transporte, mas devido às características de retransmissão do protocolo TCP, o protocolo UDP é o mais utilizado. Para o transporte de mídia o protocolo utilizado é o RTP, tanto no UDP como no TCP, em conjunto com o protocolo de controle RTCP, RFC3605 [1, 31].

As mensagens de sinalização do SIP são geralmente transportadas pelas mesmas instalações físicas utilizadas para a permutação das mídias, mas os canais são considerados distintos logicamente. Isto deve-se ao fato de que as mensagens de sinalização podem passar por um ou mais servidores *proxy* ou de redirecionamento, enquanto o fluxo de mídia assume um caminho mais direto, muito parecido com a tecnologia do protocolo de sinalização SS7 [8].

3.3.1 – Entidades da rede SIP

A arquitetura do protocolo SIP define duas classes básicas de entidades de rede: **clientes** e **servidores**. Um cliente é um programa aplicativo que envia uma solicitação SIP e

um servidor é uma entidade que responde a esses pedidos. O SIP é um protocolo de plataforma **cliente-servidor** e sua arquitetura é similar à do protocolo *HTTP*.

As chamadas de VoIP que utilizam o protocolo SIP originam-se em um cliente e terminam em um servidor. Um cliente pode ser encontrado dentro de um dispositivo do usuário ou de um servidor. Quando isso ocorre, o SIP habilita o uso de *proxies*, que agem de ambas as formas: **cliente e servidor**. Existem quatro tipos diferentes de servidores: **Proxy** (representante), **Redirecionador**, **Agente de Usuário** e **Registrador** [3]. Apesar das funções dos servidores e clientes serem diferentes, elas podem ser combinadas em um único equipamento físico.

3.3.2 – O estabelecimento de uma chamada SIP

A Figura 3.3 mostra um exemplo da operação de um servidor *proxy*. No exemplo, se a mensagem do *emissor* para o usuário (*heron*) faz um convite (*invite*) para participar de uma chamada, o efeito resultante é que a chamada será encaminhada para ele em sua residência. Neste caso, o servidor *proxy* deverá estar ciente de que o usuário (*Heron*) deverá estar em sua residência em vez de no trabalho, pelo endereço IP disponibilizado por algum serviço de acesso remoto (*Remote Access Service* (RAS)) na rede IP corporativa [3].



Figura 3. 3 – Operação do servidor *proxy* SIP.

O *proxy* pode funcionar como um *proxy* de transcodificação quando uma das partes não puder utilizar o mesmo codificador de voz, talvez por uma questão de largura de faixa ou incapacidade do equipamento [6].

A Figura 3.4 mostra um cenário da operação de um servidor redirecionador.

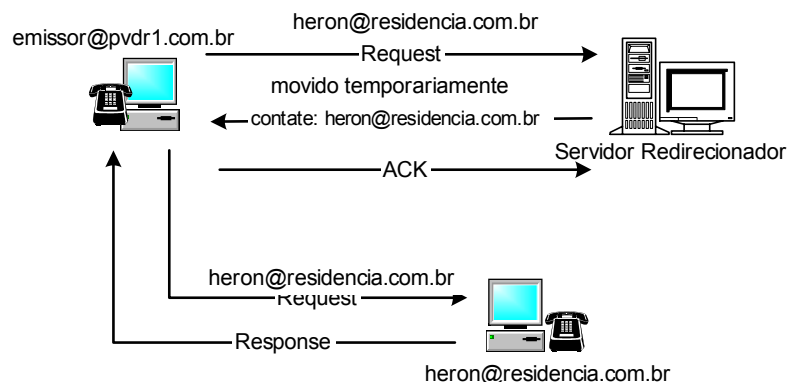


Figura 3. 4 – Operação de redirecionamento SIP.

A operação de redirecionamento pode ter outro significado em prover os serviços de siga-me e encaminhamento de chamadas que um servidor *proxy* pode oferecer. Neste caso do servidor redirecionador, o próprio cliente originador encaminha a chamada. Esse servidor simplesmente provê as informações necessárias para habilitar o cliente originador a executar esta tarefa, depois disso o servidor não estará mais envolvido no processo.

O protocolo SIP possui independência de funcionamento e operação, entretanto não faz reserva de recursos de rede, mas pode utilizar outros protocolos para realizar as tarefas de QoS, tais como o protocolo RSVP.

A simplicidade do SIP está ligada mais especificamente na forma em que ele é escrito e na quantidade necessária de mensagens para o estabelecimento e o encerramento de uma sessão de chamada. As mensagens SIP são codificadas usando a sintaxe de mensagem HTTP/1.1 (RFC 2068) e o conjunto de caracteres ISO 10646 com a codificação UTF-8 (RFC 2279), o que permite sua fácil implementação com linguagens tais como: *Java*, *Perl*, entre outras [1].

Os serviços do SIP para o estabelecimento e o encerramento de sessões multimídia são [1, 3]:

- **Localização de usuário** – determina a localização do usuário e se o mesmo pode ser utilizado para comunicação;

- **Capacidades do usuário** – determinam a capacidade de mídia dos usuários envolvidos na comunicação e os parâmetros de mídia a serem utilizados;
- **Disponibilidade do usuário** – utilizado para verificar se o usuário está disponível para uma nova comunicação, bem como se possui recurso disponível para efetivá-la;
- **Configuração de chamada** – definição dos parâmetros que serão utilizados para o estabelecimento da chamada, e;
- **Controle da chamada** – processo de gerenciamento da chamada, incluindo os processos de transferência e o encerramento das ligações.

Os procedimentos de estabelecimento de uma chamada utilizando o SIP são simples. A Figura 3.5 mostra um exemplo de estabelecimento de uma chamada.

- (a) O estabelecimento de uma chamada começa com uma mensagem de convite (INVITE) de um cliente originador SIP a outro cliente receptor SIP;
- (b) O cliente originador da chamada pode ser informado que o receptor está ocupado, que a chamada está na fila, ou que o receptor está sendo alertado (o telefone está tocando);
- (c) O receptor da chamada responde à chamada com uma mensagem OK ao originador da chamada, informando que está pronto;
- (d) O originador da chamada responde com uma mensagem de reconhecimento (*Acknowledgment* (ACK)), confirmando o convite;
- (e) Neste momento na chamada é estabelecida a conversação se inicia (troca de mídias);
- (f) Ao final da conversação um dos clientes SIP desfaz a chamada encaminhando uma mensagem de saída da sessão (BYE);
- (g) O outro cliente SIP confirma o encerramento da sessão respondendo com uma mensagem OK.

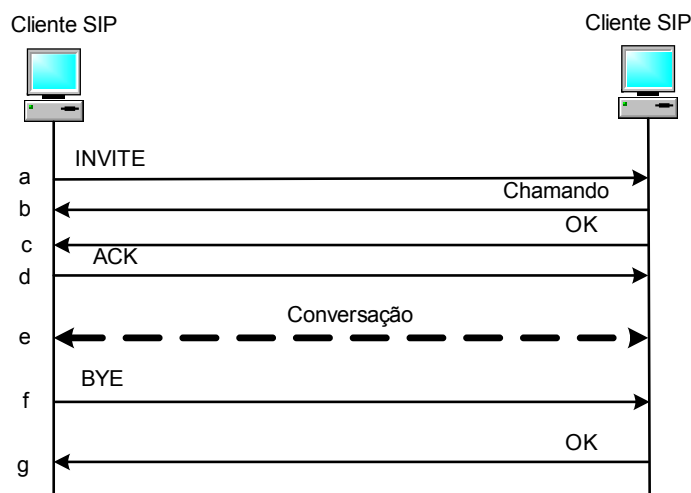


Figura 3. 5 – Exemplo de estabelecimento de uma chamada SIP.

3.3.3 – Resumo da sintaxe das mensagens do SIP

A vantagem da sintaxe do protocolo SIP ter aparência semelhante ao *HiperText Transfer Protocol* (HTTP), é que os programas projetados para serem utilizados com o protocolo HTTP podem ser relativamente adaptados para serem utilizados no SIP. Uma desvantagem óbvia em relação à codificação binária é que as mensagens de texto consomem mais largura de faixa. Entretanto, a vazão de um sistema de sinalização, seja pelo tempo de utilização ou pela quantidade de informação, é insignificante.

As linhas das mensagens são terminadas com CRLF (*Carriage Return* (retorno de carro), *Line Feed* (alimentação de linha)). Mas os receptores devem ter a capacidade de interpretar CR e LF, separadamente [6]. As mensagens SIP são solicitações (*requests*) de um cliente até um servidor ou, então, respostas (*responses*), também conhecidas como mensagens de *status* (estado), de um servidor até um cliente. Cada mensagem, se do tipo *request* ou *response*, compartilha o mesmo formato: uma linha de *start* (início) seguida por nenhum ou mais cabeçalhos e, opcionalmente, seguida por um corpo da mensagem (*message body*), conforme mostra a Figura 3.6 [1, 3, 6].

Exemplo de mensagem	Descrição
message = start line	linha de início
*message-header	cabeçalho da mensagem
CRLF	linha em branco (salto de linha)
[message-body]	corpo da mensagem

Figura 3. 6 – Formato da mensagem SIP.

A linha de *start* pode ser de duas formas [3]:

start line = request-line ou start line = status-line

A Tabela 3.3 mostra os tipos de mensagens do protocolo SIP [1].

Tabela 3. 3 – Mensagem dos tipos *request* e *response*.

Mensagem	start-line	request-line	<i>request</i>
		response-line	<i>response</i>
	message-header	general-header	<i>request e response</i>
		entity-header	<i>request e response</i>
		request-header	<i>request</i>
		response-header	<i>response</i>
	CRLF	linha em branco	<i>response</i>
	message-body	conteúdo da mensagem	<i>request e response</i>
			<i>request e response</i>

A linha de pedido (*request-line*) especifica o tipo de pedido que está sendo emitido, enquanto a linha de resposta (*response-line*) indica o sucesso ou a falha de um determinado pedido, isto é, a resposta do pedido. No caso de falha, a linha indica o tipo ou a razão da falha. As mensagens de cabeçalho (*message-header*) provêm informações adicionais relativas ao pedido ou a resposta, incluindo o originador e o destino pretendido da mensagem. O corpo da mensagem (*message-body*) descreve o tipo de sessão a ser estabelecida, inclusive uma descrição da mídia a ser permutada, indicando qual o codificador de voz deve ser utilizado (G.729, G.728, G.723.1, etc.).

3.3.4 – Forma de endereçamento adotada no SIP

É usada uma SIP *Uniform Resource Locators* (URL), similar ao endereço de correio *Simple Mail Transfer Protocol* (SMTP), formada por uma parte referente ao nome do usuário ou número telefônico (*user*) e outra ao nome do domínio ou ao endereço numérico da rede (*host.domain*), ou seja, *user@host.domain*.

Alguns parâmetros genéricos podem ser utilizados em uma SIP URL para identificar o protocolo de transporte, para identificar o número da porta específica para o protocolo de transporte que está sendo utilizado, dentre outros [1].

São exemplos de SIP URL: sip:heron.oliveira@brturbo.com, sip:usuário@10.224.70.22, sip:oliveira@gateway.com;transport=udp.

Apesar da URL SIP ser do formato *user@host*, o mais utilizado comumente é o número do telefone ao invés do usuário (*user*), por exemplo: **55622161599@caixa.gov.br** :*user=phone*, onde **55** é o código internacional do Brasil, **62** é o código DDD da região, **caixa.gov.br** é o domínio do *host* e *user=phone* indica que o usuário é um terminal telefônico [3].

A RFC 2543 define seis diferentes métodos ou pedidos [24]:

- **INVITE:** inicia uma sessão. Em uma chamada simples de duas partes, ele é utilizado para iniciar uma chamada, incluindo informações relativas às partes chamadora e chamada. Também oferece capacidade de iniciar chamadas de múltiplas partes (conferência);
- **ACK:** após ter recebido uma resposta final de um INVITE, a parte chamadora envia um reconhecimento (ACK), ou a confirmação de uma solicitação;
- **OPTIONS:** pergunta ao servidor sobre suas capacidades, por exemplo, o tipo de mídia a ser utilizado;

- **BYE:** encerra uma sessão;
- **CANCEL:** encerra um pedido pendente. Por exemplo, pode ser usado para encerrar uma sessão em que um INVITE foi enviado, mas de uma resposta final que ainda não foi recebida; e
- **REGISTER:** Um cliente agente de usuário usa o método de REGISTER (REGISTRO) para abrir uma sessão (*login*) e registrar seu endereço com um servidor SIP, assim deixando o registrador conhecer o endereço no qual o usuário está localizado.

Na arquitetura SIP, a resposta de uma *start-line* (linha de partida ou início) é uma *status-line* (linha de estado ou situação). Toda linha de estado possui um código de estado (*status*), que possui três dígitos indicando o resultado do pedido. Para cada código de resultado existe um texto explicando a razão ou o motivo do resultado [3].

A sintaxe da *status-line* é:

status-line = SIP version SP status code SP reason-phrase CRLF

A RFC 2543 [24] define também os códigos de estado com valores entre 100 e 699, sendo que os primeiros dígitos do código razão indicam a classe da resposta. As seis classes de código-razão são: 1XX – informativa, 2XX – sucesso, 3XX – redirecionamento, 4XX – erro (falha) de solicitação, 5XX – falha de servidor e 6XX – falha global.

3.3.5 – Interoperando com outros sistemas de sinalização

As redes de comutação de circuitos e de voz sobre pacotes coexistirão por muito tempo. Logo, existe a necessidade óbvia para que as redes baseadas no SIP possam interagir com as redes de comutação de circuitos da RTPC. Mesmo considerando as várias opiniões de que o SIP seja “o futuro da Telefonia IP”, pode-se considerar que uma base embutida já existe

no sistema H.323, e que mais sistemas estão sendo desenvolvidos sobre a recomendação H.323. Então, também existe a necessidade de que as redes baseadas no SIP devem interagir com as redes baseadas na arquitetura H.323, apesar de que a própria RFC 2543, especificamente, não indica que tal interconexão deva ser realizada.

3.3.5.1 – Interoperando com a rede de telefonia pública comutada

Para interagir com a rede de telefonia pública comutada (RTPC) os *gateways* serão exigidos para fornecerem a conversão de mídia entre as redes de comutação de circuitos e as redes de pacotes e vice-versa. Da mesma forma considerada pela arquitetura H.323, não só deve existir a interação das mídias, mas também a interação da sinalização. Como o SIP é um protocolo de sinalização, se as chamadas devem ser estabelecidas entre as redes baseadas no SIP e na RTPC, então a rede SIP deve ser capaz de comunicar com a RTPC de acordo com o protocolo de sinalização utilizado nela. Na maioria dos casos, esse protocolo é o SS7. Para o estabelecimento, a manutenção e o encerramento de chamadas de voz, o protocolo SS7 aplicável é o ISUP [32].

O uso e o controle de *gateways* em uma rede baseada no SIP podem ser encontrados com mais detalhes nas referências [6, 3, 29], e a interação do SIP com a RTPC com a utilização de *gateways* de sinalização é discutida na seção 3.4 deste capítulo.

3.3.5.2 – Interoperação do protocolo SIP com a arquitetura H.323

Vários esboços de Internet têm sido preparados para direcionar o assunto, mas um em particular sob o título “SIP-H.323 *Interworking Requirements*” [33] está sendo preparado como proposta para atender a interoperação entre o SIP e o H.323, com o objetivo de fornecer o serviço de VoIP fim-a-fim entre duas arquiteturas diferentes [33].

A Figura 3.7 mostra a interoperação entre as arquiteturas H.323 e SIP.

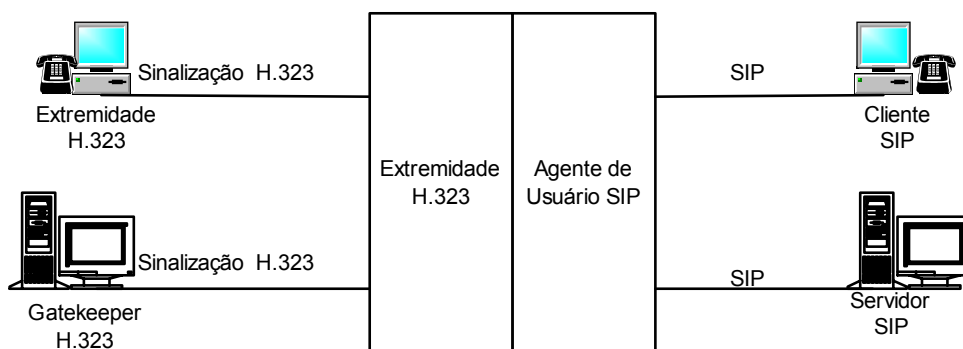


Figura 3. 7 – Interoperação entre as arquiteturas H.323 e SIP.

A abordagem proposta de interoperação (*Interworking*) entre uma rede SIP e uma rede H.323 envolve o uso de um *gateway*, como mostrado na Figura 3.7. Dependendo da função a ser apresentada, o *gateway* pode aparecer para a rede SIP como um cliente de agente de usuário ou um servidor de agente de usuário. Para a rede H.323, o *gateway* deve aparecer como uma extremidade da arquitetura H.323. No lado do SIP, o *gateway* pode também incluir funções de servidores *registrar* e/ou de *proxy*, e no lado do H.323 ele pode conter funções de *gatekeeper*.

Mais detalhes sobre os procedimentos das mensagens do SIP podem ser encontrados em [24] e vários exemplos de cenários de mensagens do protocolo SIP podem ser observados nas referências [1, 3, 6, 29], assim como outros cenários e requisitos de interoperação do protocolo SIP com a arquiteturas H.323 podem ser encontrados no esboço de Internet da referência [33].

3.4 – Protocolos de Sinalização e Controle de *Gateway*

Espera-se que as formas de interconexão dos diversos tipos de redes intermediárias sejam indiferentes e transparentes ao usuário. O desafio está em desenvolver soluções que possam realizar as tarefas de *internetworking* de diversos tipos de redes sem a interrupção indesejada, queda de qualidade e baixo custo. Sempre deve ser lembrado que a função principal dos *gateways*, no tema desta dissertação, é fazer com que a rede de VoIP apareça

para a rede tradicional telefônica como um sistema nativo de comutação de circuitos, e vice-versa.

Um conceito já utilizado no Sistema de Sinalização 7 (*Signaling System 7 (SS7)*) em que a sinalização de uma chamada segue caminhos diferentes da mídia, é também utilizado nas chamadas VoIP, fazendo com que os *gateways* passem a ter duas funções principais: conversão da sinalização e da mídia. A Figura 3.8 mostra a mesma situação em um cenário em que uma rede de VoIP provê um serviço interurbano [32].

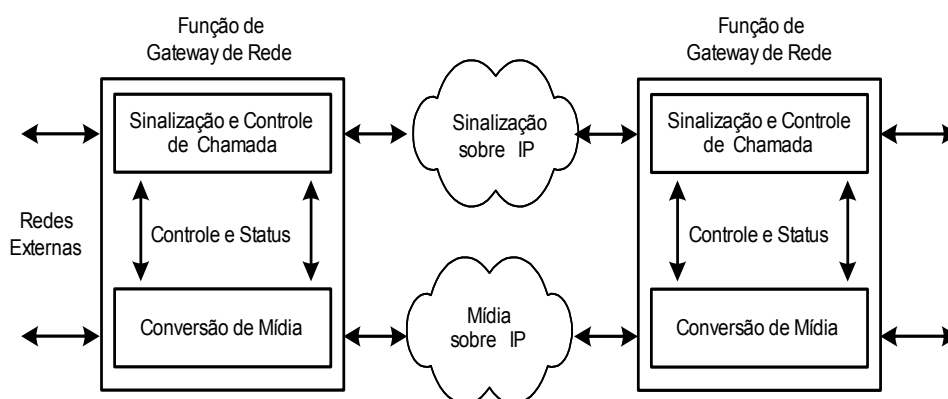


Figura 3. 8 – Tráfego de sinalização e mídia em um serviço interurbano.

Além da separação lógica dos caminhos diferentes da sinalização e da mídia, a separação física também pode ocorrer e tem seus benefícios tal como maior rapidez no tráfego, porque as características de controle não necessitam ser implementadas em todos os nós da rede, mas apenas nos nós centralizados de manipulação das chamadas. Quando são implementados equipamentos com funções técnicas específicas, elas são executadas com maior rapidez.

Como a separação das funções de tráfego de sinalização e de mídia em um só equipamento H.323 e/ou SIP pode acarretar em maior velocidade de tráfego, surgiu a necessidade de protocolos padronizados que pudessem controlar os *gateways* de mídia, conhecidos como Protocolo de Controle de *Gateways* de Mídia (*Media Gateways Controller*

Protocol (MGCP)), desenvolvido pelo IETF. Outro protocolo de controle de *gateways* de mídia é o MEGACO/H.248, conhecido como MEGACO (*MEdia Gateway COntroller* (MEGACO)) no IETF e H.248 [7] no ITU-T, que está sendo desenvolvido em consórcio formado pelos órgãos IETF e ITU-T.

A utilização de qualquer um dos protocolos, MEGACO/H.248 ou MGCP, em conjunto com o SIP e/ou H.323 é chamada de Arquitetura *Softswitch*.

3.4.1 – A arquitetura *softswitch*

A arquitetura *softswitch* envolve a separação do caminho da mídia e as funções de conversão de mídia das funções de controle e sinalização. As entidades que manipulam o controle de chamada são conhecidas como agentes de chamada ou controladores de *gateways* de mídia, enquanto as entidades que executam a conversão da mídia são conhecidas como *gateways* de mídia.

Collins [3] explica que o termo *softswitch* é usado porque muitas funções de comutação (*switching*) tradicionalmente manipuladas por grandes sistemas monolíticos no mundo da comutação de circuitos são emuladas por sistemas de *software*, e que esse termo tem sido usado para referir-se a um agente de chamada ou *Media Gateway Controller* (MGC), ao invés da arquitetura global.

Na interoperação entre o SIP e a Rede Telefônica Pública Comutada (RTPC), o agente de usuário do SIP controla a interoperação funcional de sinalização entre o SIP e a RTPC, que pode ser o *ISDN User Part* (ISUP), enquanto os *gateways* cuidam das funções de conversão de mídia [32].

A Figura 3.9 mostra a interoperação entre as arquiteturas *softswitch* e a RTPC.

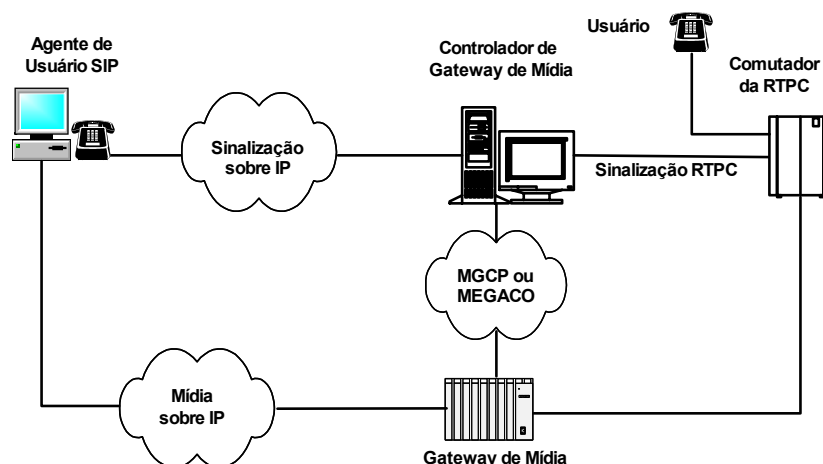


Figura 3. 9 – Interoperação entre a arquitetura *Softswitch* e a RTPC.

3.4.2 – Protocolos para controle de *gateways* de mídia

Além dos conceitos sobre MG e MGC, alguns requisitos genéricos para um protocolo de controle de *gateway* são definidos pela RFC 2805 [34], que descreve os serviços a serem suportados por um sistema distribuído de *gateway*, tal como o sistema mostrado na Figura 3.9. Em síntese, os serviços mais importantes são:

- Recursos de mídia, tais como: resposta audível interativa (*Interactive Voice Response* (IVR)), pontes de conferência (*conference bridges*), etc.;
- Recepção e geração de tons DTMF;
- Controle de cancelamento e supressão de eco;
- Controle de codificadores de voz, tais como G.711, G.723.1, G.729, GSM, etc.;
- Geração de tons, tais como tom de discagem (*dial tone*), tom de chamada (*ring tone*), tom de ocupado (*busy tone*), etc.;
- Monitoramento e coleta de estatísticas;
- Auditoria e teste de extremidade (pontos finais) (*loop back*, etc.);
- Reserva, liberação e bloqueio de extremidades, e;
- Criptografia, dentre outros.

3.4.3 – O MGCP

O MGCP foi desenvolvido inicialmente para atender de forma genérica o controle distribuído do *Gateway* de Mídia (*Media Gateway* (MG)). Atualmente, o MGCP realiza especificamente a interface entre um MGC e os MG [1, 6].

O MGCP foi inicialmente especificado na RFC2705 [35] que ficou obsoleta pela RFC3435 [36] e atualizada pela RFC3660 [37]. Ele é baseado em texto, suporta um modelo de chamada centralizado, um protocolo do tipo cliente/servidor e utiliza o SDP para carregar informações e parâmetros relevantes para o *gateway*, tais como: endereço IP, número da porta UDP, tipo de mídia (áudio ou dados), etc.

A Figura 3.10 mostra a forma de funcionamento do protocolo MGCP, como regras de comunicação entre os controladores dos *gateways* de mídia (MGC) e os *gateways* de mídia (MG) [1, 3, 6].

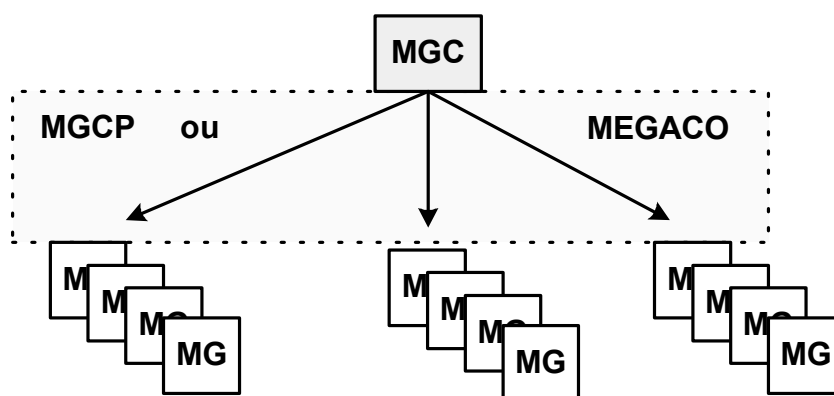


Figura 3. 10 – Síntese de um protocolo controlador de *gateway* de mídia.

Collins [3] explica que o MGCP foi criado provisoriamente pela RFC2705 até a definitiva padronização do MEGACO/H.248, em um esforço comum entre o IETF e ITU-T. Pode-se dizer que o MEGACO é a evolução do MGCP. Logo, devido à grande semelhança entre os dois protocolos, não será tratado nesta dissertação as características do MGCP.

3.4.4 – O MEGACO/H.248

O termo MEGACO (*Media Gateway Control Protocol*) é bastante utilizado pelo IETF e H.248 [7] pelo ITU-T. Nesta dissertação foi adotado o termo MEGACO por ser mais simples e mais curto do que MEGACO/H.248.

O protocolo MEGACO foi inicialmente projetado na RFC2885 [38] Versão 0.8, posteriormente na Versão 1 na RFC3015 [39] e atualmente está na Versão 2 como *Internet Draft* (esboço) [40]. Como esperado, ele apresenta várias semelhanças arquitetônicas com o MGCP e também define o MG como o ponto de conversão do formato de mídia exigido em uma rede para o formato exigido em outra rede.

Entretanto, é inovadora a forma lógica de apresentação do protocolo MEGACO. Nele um MG também pode iniciar um pedido de transação, a escrita fornece as duas formas de codificação: de texto (*Augmented Backus-Naur Form* (ABNF)) utilizada no IETF e binária (*Abstract Syntax Notation One* (ANS.1)) utilizada pela ITU-T. Ambos os códigos são obrigatórios no MGC, mas no MG pode-se escolher um deles [3].

3.4.4.1 – Entidades lógicas do MEGACO

A forma lógica de apresentação do MEGACO é um pouco diferente do MGCP, que define outras entidades lógicas, tais como *Terminator*, *Context*, *Transaction*, *Message* e *Descriptor* [3, 6].

Terminator ou *Termination* (Terminador ou Terminação) é uma entidade lógica dentro do MG que dissipa fluxos e informações de multimídia [6]. Algumas terminações são físicas e têm uma existência semipermanente e podem ser associadas a recursos físicos externos. São similares aos terminais do MGCP. Existem terminações temporárias (ou transientes) enquanto durar uma chamada ou fluxo de mídia. As terminações transientes são

conhecidas como terminações efêmeras e representam fluxos de mídia, como a sessão RTP [3].

Um **Context** (contexto) é uma entidade lógica dentro de um MG que é o resultado da associação de várias terminações com o propósito de compartilhar mídia entre estas terminações. As terminações podem ser adicionadas, removidas ou reposicionadas de um contexto até outro. Uma terminação só pode existir em um contexto em determinado momento, e as terminações de um determinado *gateway* apenas podem trocar mídia se elas estiverem no mesmo contexto. Os comandos utilizados para manipular as entidades lógicas de um MG são [1]:

- *Add* – utilizado para adicionar uma terminação a um contexto;
- *Move* – utilizado para mover uma terminação de um contexto;
- *Subtract* – remove uma terminação de um contexto;
- *Modify* – utilizado para modificar os valores das propriedades de uma terminação;
- *AudiValue* – utilizado para informar ao MGC o estado corrente das propriedades e os eventos de uma terminação;
- *Notify* – utilizado pelo MG para informar ao MGC a ocorrência de um evento ao MG;
- *AudiCapacities* – utilizado para informar todos os possíveis valores para cada propriedade de uma terminação em um MG, e;
- *ServiceChange* – utilizado pelo MG para informar que uma terminação ou um grupo de terminadores está em serviço ou fora de serviço.

A troca de comandos entre o MGC e o MG é realizada pelas **transactions** (transações), que são compostas de uma série de ações, que por sua vez consistem de um conjunto de comandos para manipular as terminações de um contexto específico, envolvendo o transcurso de comandos e as respostas para esses comandos [3, 6].

O exemplo abaixo combina o conceito de mensagem, transações e comandos, além de mostrar o formato de texto de uma mensagem [3]:

```
MEGACO/1 [111.111.222.222] :34567
Transaction = 12345 {
    context = 1111 {
        Add = A5555,
        Add = A6666
    }
    Context = $ {
        Add = A7777
    }
}
```

Neste exemplo, um MGC emite a mensagem de endereço *111.111.222.222* e porta *34567*. *TransactionID 12345* é a identificação da transação. Dois comandos *Add* são relacionados ao contexto *1111*, resultando nas terminações *A5555* e *A6666* a serem adicionadas ao contexto *1111*. Um comando *Add* é relacionado a qualquer contexto (\$) que o MG pode escolher, por meio do qual a terminação *A7777* é adicionada ao contexto. O processamento deste comando resulta na criação de um novo contexto pelo MG, para o qual o *ContextID* deve ser retornado na resposta.

Vários descritores podem ser utilizados com os comandos e as respostas e integra as informações adicionais para qualificar um comando ou uma determinada resposta. Um *Descriptor* (descritor) pode ser obrigatório, proibido ou opcional.

O formato geral de um descritor é [3]:

Descriptorname=<someID>{Parm=value, parm=value, ...}

A relação dos descritores, assim como suas funções, podem ser encontradas na referência [3] e no esboço de Internet da referência [40], bem como os cenários de interoperação entre o MEGACO e os protocolos SIP e H.323 nas referências [1, 6, 29].

3.5 – Considerações Finais deste Capítulo

Este capítulo descreveu os principais conceitos das arquiteturas H.323, SIP e *Softswitch* e seus elementos, assim como os protocolos MGCP, MEGACO e suas principais características.

CAPÍTULO IV

TÉCNICAS DE CODIFICAÇÃO, QUALIDADE DA VOZ E QoS

4.1 – Introdução

A sinalização é essencial na determinação da negociação dos diversos parâmetros entre dois ou mais operadores e/ou provedores de uma comunicação e entre dois ou mais interlocutores, tais como a rota da informação, os codificadores de voz, o encaminhamento de diversos tons, etc.

Em estudos sobre o comportamento do sinal de voz foi verificado que ele é relativamente previsível. Esta previsibilidade está sendo usada nos projetos de sistemas de codificação de voz para reduzir a largura de faixa a ser utilizada para o seu transporte. Um dos fatores do alto custo nas telecomunicações é a largura de faixa utilizada para o transporte das informações. Quanto maior a previsibilidade do sinal de fala, menor será a largura de faixa a ser utilizada na comunicação [6, 18].

Este capítulo apresenta os codificadores de voz mais utilizados na aplicação de VoIP, as técnicas de medição, os fatores que influenciam na qualidade da voz, os conceitos fundamentais de qualidade de serviço (*Quality of Service* (QoS)) em uma rede VoIP, assim como apresenta os principais mecanismos disponíveis atualmente para a sua implementação. Finalmente, são realizadas considerações finais deste capítulo.

4.2 – Codificação da voz

A Figura 4.1 mostra o diagrama de blocos da codificação em um sistema de comunicação independente do tipo de rede utilizado.

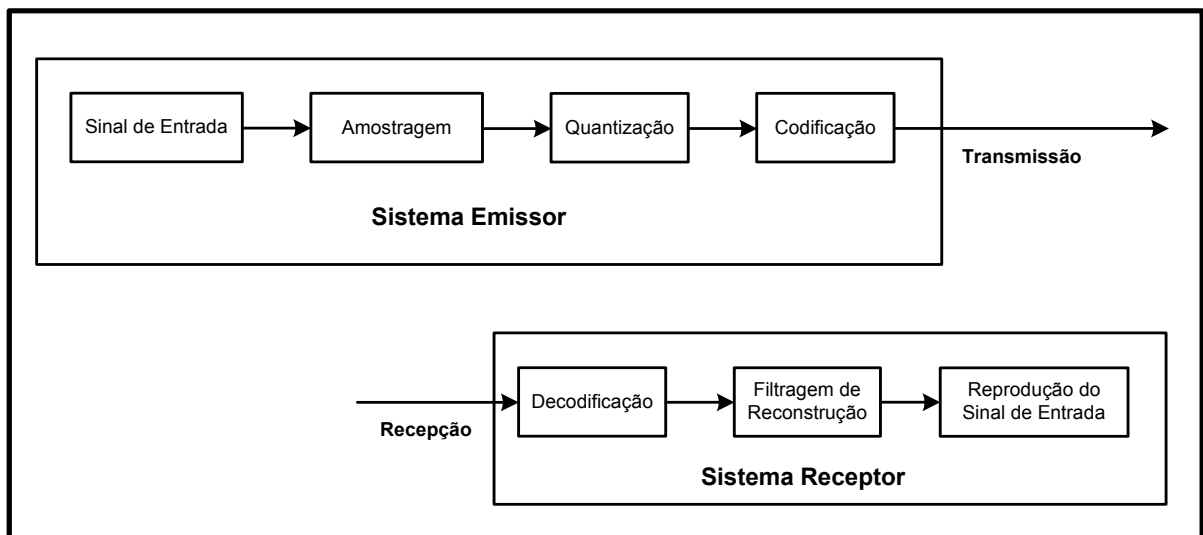


Figura 4. 1 – Diagrama de blocos da codificação de um sistema de comunicação.

4.2.1 – O codificador G.711 ou PCM

A Recomendação G.711 padroniza o codificador mais simples e mais utilizado em todo o mundo. Ele é amplamente utilizado nas redes de comutação de circuitos. Sua amostra é representada em 8 bits, amostrada em uma frequência de amostragem de 8 kHz, com períodos de amostragem de 0,125 milissegundos. Sendo, portanto, necessária uma capacidade de transmissão de 64 Kbps para o transporte do sinal PCM [6].

O PCM tem evoluído com as técnicas diferencial e adaptativa (ADPCM, também conhecido como ITU-T G.726 e G.727), apresentando boa qualidade de voz para as taxas entre 22 e 48 Kbps. Os codificadores PCM e ADPCM utilizam a quantização da forma de onda e têm uma classificação MOS acima de 4, aproximadamente 4,3. A maior desvantagem do G.726 é a largura de faixa consumida [6].

4.2.2 – O codificador G.723.1 (ACELP e MP-MLQ)

O codificador G.723.1, conhecido como codificador preditor linear de codebook-excitado algébrico (*Algebraic Codebook-Excited Linear Predict* (ACELP)) quando opera na taxa de transferência de 5,3 Kbps, ou também conhecido como quantização de probabilidade máxima de pulso-múltiplo (*Multi-Pulse Maximum Likelihood Quantization* (MP-MLQ)) quando opera na taxa de transferência de 6,3 Kbps.

O funcionamento do G.723.1 pode ser sintetizado da seguinte forma [6, 18]:

- o sinal de entrada de voz é amostrado a 8 kHz como da codificação PCM e convertido para PCM uniforme, resultando em um sinal de 16 bits para a entrada do codificador;
- o codificador separa o sinal em blocos ou quadros de 240 amostras por vez, que corresponde a 30 ms de voz;
- um atraso de 7,5 ms, chamado de *look-ahead*, é acrescido ao atraso de 30 ms de voz, totalizando 37,5 ms de atraso mais o atraso de processamento do codificador;
- o quadro é então repassado por um filtro passa-alta e é dividido em quadros de 60 amostras cada um, determinando os coeficientes específicos do filtro;
- dependendo da taxa de transferência utilizada, se 5,3 Kbps ou 6,3 Kbps, os quadros são montados em 20 octetos ou 24 octetos, respectivamente, ACELP ou MP-MLQ.
- tais quadros são compostos com as informações dos coeficientes de predição linear, parâmetros de ganho e índice de excitação *codebook*;
- durante a conversação os quadros de silêncio são inseridos pela técnica de inserção de silêncio (*Silence Insertion Description* (SID)), com capacidade de detecção de presença de voz e geração de ruído de conforto;
- o quadro do SID é composto de 4 octetos, ocupando uma faixa de 1 kbps, mas transmitidos na taxa de 64 kbps do G.711;

- em cada quadro, dois bits menos significativos do primeiro octeto indicam o tamanho do quadro e a versão do codificador em uso, e;
- ao todo são transmitidos 3 tipos diferentes de quadros no G.723.1: um para a taxa de 5,3 Kbps (ACELP), um para a taxa de 6,3 kbps (MP-MLQ) e outro para o SID. A Tabela 4.1 mostra os tipos de quadros existentes no G.723.1.

Tabela 4. 1 – Tipos de quadros existentes no G.723.1.

Bits	Significado	Octetos por quadro
00	alta taxa de voz (6,3 kbps)	24
01	baixa taxa de voz (5,3 kbps)	20
10	quadro SID	4
11		não utilizado

O anexo da Recomendação G.723.1 [41] descreve o esquema de compressão de silêncio com um único usuário, que permite a redução da banda passante. Nesse anexo é descrito tanto o detector de voz ativa (*Voice Activity Detector* (VAD)) como o gerador de ruído de conforto.

4.2.3 – O codificador G.729 ou CS-ACELP

O codificador G.729 [19] é também conhecido como codificador de predição linear de código-excitado algébrico de estrutura conjugada (*Conjugate Structure-Algebraic Code-Excited Linear Prediction* (CS-ACELP)). Esta recomendação descreve um algoritmo de codificação de voz em quadros de 10 ms resultando em uma faixa de 8 Kbits/seg. Um atraso total do algoritmo de 15 milissegundos, utilizando o CS-ACELP.

A voz é reconstruída pelo filtro de síntese de curto prazo, como mostrado na Figura 4.2. Esse filtro é baseado em um filtro de predição linear de décima-ordem (*Linear Prediction* (LP)). O filtro de longo prazo é implementado usando o quadro de *codebook* adaptável. Um pós-filtro é adicionado após a voz ser reconstruída [19].

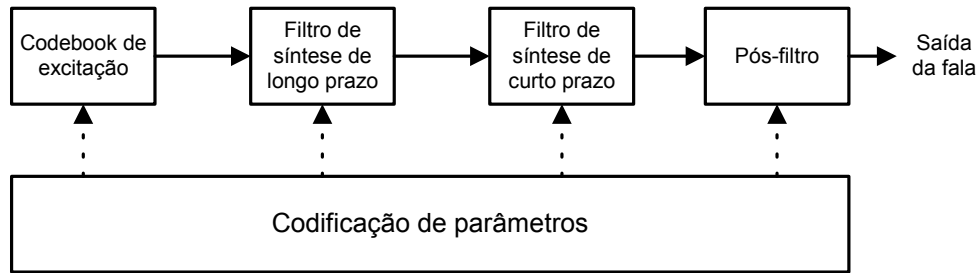


Figura 4. 2 – Diagrama conceitual de blocos do modelo de síntese CELP.

Devido à complexidade do codificador G.729, foram criadas duas implementações chamadas de anexo A [42] e Anexo B [43]. O anexo A, homologado em Genova em 1996, é uma versão de interoperação com a versão completa, isto é, reduziu a complexidade algorítmica do codificador e ao mesmo tempo pode operar com a implementação completa e vice-versa. Entretanto, a qualidade de voz da implementação deste anexo não é tão boa quanto da versão completa original em certas circunstâncias. A classificação do anexo A do G.729 da MOS é 3,7 [42].

As alterações mais importantes implementadas no anexo A do G729 são rotinas simplificadas de pesquisa no *codebook* e uma simplificação no *postfilter* no decodificador. O anexo B inclui novos algoritmos, tais como: detecção da atividade da voz (*Voice Active Detection* (VAD)), transmissão descontínua (*Discontinuous Transmission* (DTX)) e gerador de ruído de conforto (*Comfort Noise Generator* (CNG)).

A Tabela 4.2 mostra os principais codificadores, a duração dos quadros e o tamanho do *payload*.

Tabela 4. 2 – Duração do quadro e tamanho do *payload* por Codificador de Voz/Algoritmo.

Algoritmo	G.711	G.729	G.723.1	G.723.1
taxa (kpbs)	64	8	6,3	5,3
duração do quadro (ms)	20	20	30	30
tamanho do <i>payload</i> (bytes)	160	20	24	20

O cálculo do *payload* é realizado pela Equação (4.1), lembrando que nenhum dado pode ser processado até que o quadro esteja completamente cheio [1].

$$Payload = (taxa * duração) / 8 \quad (4.1)$$

Finalmente, a Figura 4.3 mostra uma relação entre a qualidade da voz dos codificadores mais conhecidos e a largura de faixa utilizada. Os codificadores G.723.1 (6,3 kbps) e G.729 (8 kbps) fornecem a melhor relação entre a qualidade e a largura de faixa [44].

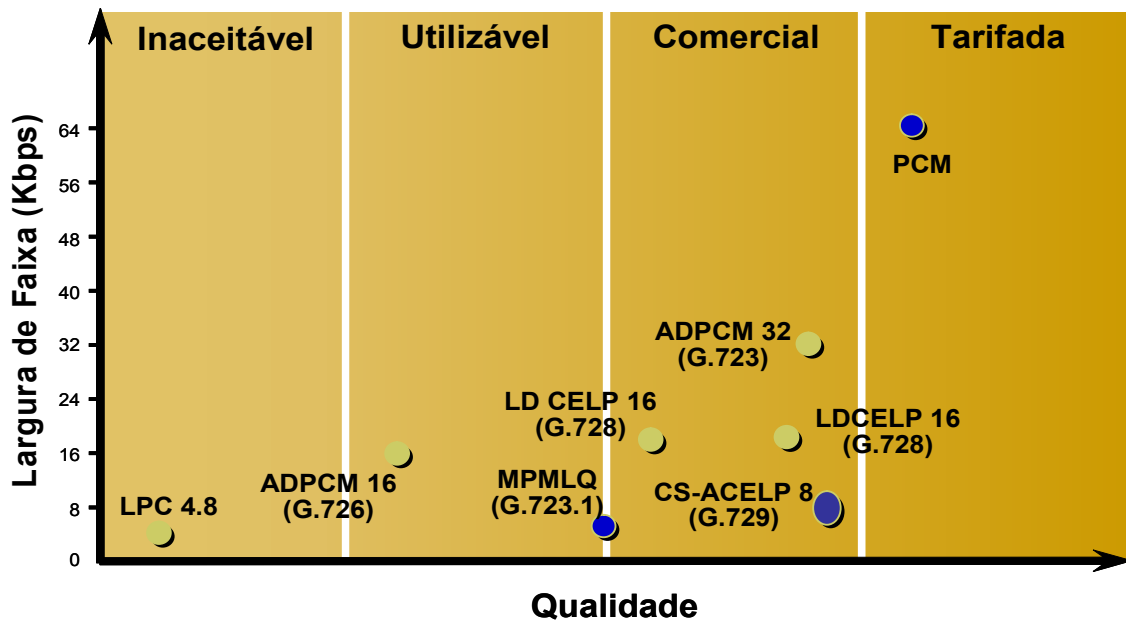


Figura 4. 3 – Relação da largura de faixa e da qualidade de voz dos codificadores.

4.2.4 – Métodos de qualificação do sinal de voz

Medidas padrões têm sido desenvolvidas pelo ITU, mas a qualidade de reprodução de voz na rede telefônica é essencialmente subjetiva. De forma conclusiva, quanto mais próximo o sinal de entrada estiver em comparação ao sinal de saída do sistema, melhor será a qualidade do serviço observado. Além disso, o uso desses métodos é a oportunidade de comparação da resposta dos codificadores de voz.

4.2.4.1 – *Mean opinion score (MOS)*

O ITU-T criou as recomendações P.800 [45] e P.830 [46], também conhecida como pontuação média de opinião (*Mean Opinion Score (MOS)*) ou classificação por categoria absoluta (*Absolute Category Rating (ACR)*) [6].

As notas definidas pelo ITU-T na ACR são como mostrado na Tabela 4.3.

Tabela 4. 3 – Pontuação média de opinião.

Média de Opinião	Nota
excelente	5
boa	4
regular	3
insatisfatória	2
ruim	1

As recomendações são muito extensas e consistem de diversos anexos que podem ser verificados na P.800 [45] e P.830 [46].

4.2.4.2 – *Perceptual speech quality measure (PSQM)*

A recomendação P.861 [47] descreve um método objetivo para estimativa da qualidade de codificadores de voz, operando na faixa telefônica (300 – 3400 Hz). O cálculo da qualidade é baseado em uma medida denominada *Perceptual Speech Quality Measure (PSQM)*, que foi escolhida dentre várias outras estudadas pelo ITU-T por apresentar o melhor desempenho na correlação com a qualidade subjetiva dos codificadores de voz. Outros detalhes podem ser observados em [46].

4.2.4.3 – *Perceptual speech quality measurement plus (PSQM+)*

O *Perceptual Speech Quality Measurement Plus (PSQM+)* é uma evolução do método PSQM que tenta corrigir o reflexo inadequado das interferências da rede de transmissão de dados, como perdas de pacotes e taxa de erro de bit, na pontuação resultante do método

original. O PSQM+ possui o mesmo esquema do PSQM, mantém o mesmo tipo de pontuação e não foi adotado como padrão por nenhum organismo internacional de normatização, mas encontra-se implementado na maioria dos equipamentos disponíveis no mercado, para medição da qualidade da voz na faixa de 0,3 a 3,4 KHz.

O PSQM+ produz um aumento da pontuação que reflete em uma melhor correlação com a pontuação obtida nos testes subjetivos [48].

4.2.4.4 – *Perceptual analysis measurement system (PAMS)*

O *Perceptual Analysis Measurement System (PAMS)* é também um método de medida objetiva da qualidade da voz em sistemas telefônicos, que tenta reproduzir a percepção humana usando técnicas de processamento de sinal e um modelo sensitivo diferente dos empregados no PSQM+: comparando o sinal original da voz à versão degradada de saída do sistema de comunicação [48].

O PAMS teve grande aceitação na Europa e também não foi adotado como padrão por nenhum organismo internacional de normatização, mas encontra-se implementado em muitos equipamentos disponíveis no mercado.

4.2.4.5 – *Perceptual evaluation of speech quality (PESQ)*

O *Perceptual Evaluation of Speech Quality (PESQ)*, recomendação P.862 da ITU-T [49], é também um método de avaliação objetiva da qualidade fim-a-fim de voz operando na faixa telefônica (300 – 3400 Hz). Ele é utilizado em sistemas com fala codificada, atraso variável, filtragem, perda de pacotes ou células, corte no tempo e erros de canal.

O PESQ adiciona novos métodos de equalização e distorção no tempo e combina a robustez das técnicas de alinhamento no tempo do PAMS com a precisão de testes sob condições reais da rede [49].

4.2.5 – Fatores que influenciam a qualidade da voz na rede VoIP

Na transmissão de voz sobre redes de pacotes existem quatro fatores principais que impactam a qualidade do serviço: capacidade do enlace, atraso (fim-a-fim) de pacote, *jitter* de atraso e perda de pacotes.

4.2.5.1 – Largura de faixa

Quanto maior a largura de faixa melhor é a qualidade da voz, isso é uma tendência, mas não é toda a verdade, pois existem outros parâmetros correlacionados, tal como a técnica de codificação de voz. A vazão mínima necessária para a transmissão do sinal de voz é função da técnica de codificação utilizada. É dito, então, que se a largura da faixa não é mínima suficiente para o transporte da voz codificada, a qualidade da voz não será boa.

O codificador G.711 utiliza uma amostragem de 0,125 milissegundos, que requer uma taxa de 64 Kbps ($f = (1/0,000125) * 8 = 8.000 * 8 = 64.000$ bits por segundo). Considerando que nenhuma RFC trata da duração do tempo de transmissão do pacote e que a RFC 1889, que foi substituída pela RFC 3550 [20], fornece um exemplo de duração de pacote de 20 ms, também devido a uma vasta literatura do tema, esta dissertação utiliza esse valor como um parâmetro nos cálculos. A cada transmissão de dados da voz o cabeçalho IP/UDP/RTP deve ser transmitido no pacote. O comprimento total desse cabeçalho é de 40 bytes. Assim, em 1 segundo pode-se transmitir 50 pacotes de 20 ms. Como esse cabeçalho possui 320 bits ($8 \text{ bits} * 40 \text{ bytes}$), ele utiliza uma taxa de 16.000 bits ($320 * 50$) por segundo. Portanto, somente ele utiliza uma taxa de 16 Kbps, sendo que os quadros de voz ocupam 64 Kbps, então o pacote ocupa um total de 80 Kbps ($64 \text{ Kbps} + 16 \text{ Kbps}$) [50].

Porém, se for utilizada a compressão de cabeçalho (*Compressed Real-time Transport Protocol* (CRTP)) da RFC 2508 [51] de 4 bytes de comprimento, o cabeçalho ocupará uma faixa de 1.600 bits por segundo ($4 * 8 * 50$), um total de 65,6 Kbps de faixa por pacote.

Considerando-se, ainda, que o comprimento do cabeçalho do protocolo ponto-a-ponto (*Point-to-Point Protocol* (PPP)) é de 6 bytes, o comprimento total do cabeçalho é de 10 bytes, utilizando uma largura de faixa de 4 Kbps ($10 * 8 * 50$). Assim, a largura de faixa total do pacote de voz para o G.711 com um quadro de voz é de 68 Kbps, com um comprimento de 170 bytes, onde cada quadro da voz possui um comprimento de 160 bytes ($64.000/(50*8)$) [50].

O mesmo cálculo pode ser realizado para o codificador G.729. Para os valores do tamanho do pacote e a taxa de transmissão considerados na simulação desta dissertação, pode-se construir a Tabela 4.4.

Tabela 4. 4 – Cálculo da vazão necessária.

Descrição										
algoritmo	G.711	G.711	G.711	G.711	G.711	G.729	G.729	G.729	G.729	G.729
amostras por pacote	1	2	3	4	5	1	2	3	4	5
taxa transmissão	64	64	64	64	64	8	8	8	8	8
duração (ms)	20	40	60	80	100	20	40	60	80	100
payload (voz)	160	320	480	640	800	20	40	60	80	100
cabeçalho CRTP	4	4	4	4	4	4	4	4	4	4
PPP	6	6	6	6	6	6	6	6	6	6
tamanho do pacote (bytes)	170	330	490	650	810	30	50	70	90	110
vazão necessária em Kbps	68	66	65,33	65	64,8	12	10	9,33	9	8,8

O intervalo de transmissão entre os pacotes deve ter a mesma duração do quadro, isto é, se o quadro tem uma duração de 20 ms, o intervalo deve ser de 20 ms [50], uma vez que o tempo do codificador na geração dos quadros de voz não deve ser desconsiderado.

A Tabela 4.5 mostra a duração da geração dos quadros de voz para o codificador G.729 para 1, 3, 6, 9 e 12 quadros, considerando que para cada janela (quadro) de 10 ms são analisadas 80 amostras de 8 bits [1, 42].

Tabela 4. 5 – Vazão necessária para o G.729 até 12 quadros por pacote.

Descrição					
algoritmo	G.729	G.729	G.729	G.729	G.729
amostras por pacote	1	3	6	9	12
taxa transmissão	8	8	8	8	8
duração (ms)	20	60	60	80	100
payload (voz)	20	60	120	180	240
cabeçalho CRTP	4	4	4	4	4
PPP	6	6	6	6	6
tamanho do pacote (bytes)	30	70	130	190	250
vazão necessária em Kbps	12	9,33	9,33	9	8,8

A Tabela 4.6 mostra a duração da geração dos quadros de voz para o codificador G.723 para 1, 3, 6, 9 e 12 quadros [1, 18].

Tabela 4. 6 – Vazão necessária para o G.723 até 12 quadros por pacote.

Descrição					
algoritmo	G.723.1	G.723.1	G.723.1	G.723.1	G.723.1
amostras por pacote	1	3	6	9	12
taxa transmissão	6,3	6,3	6,3	6,3	6,3
duração (ms)	30	90	180	270	360
payload (voz)	24	72	144	216	288
cabeçalho CRTP	4	4	4	4	4
PPP	6	6	6	6	6
tamanho do pacote (bytes)	34	82	154	226	298
vazão necessária em Kbps	34	7,17	6,74	6,59	6,52

Comparando-se os valores de duração dos quadros para os codificadores G.729 e G.723.1 nas Tabela 4.5 e 4.6, pode-se verificar um crescimento mais rápido na duração dos pacotes para o codificador G.723.1, mesmo considerando o tempo de geração do quadro de 30 ms e não de 67,5 ms (30 ms da amostra + 7,5 ms de *look-ahead* + 30 ms do processador). Por esse motivo o G.729 é o codificador escolhido para realização dos cenários de simulação nesta dissertação, porque apresenta menor comprimento dos dados de voz por pacote. Porém, com o objetivo de enfatizar a análise do comprimento dos dados de voz no pacote IP na rede VoIP, o codificador G.711 foi utilizado no primeiro cenário da simulação, com o comprimento dos dados de voz por pacote, conforme mostrado na Tabela 4.4. Pode-se

verificar nas Tabelas 4.4, 4.5 e 4.6 que o comprimento dos dados de voz em bytes do G.711 é bem superior aos demais codificadores com menor taxa de transmissão, ou seja, o objetivo é analisar os dois extremos dos codificadores: o de maior *payload* (G.711) e o de menor *payload* (G.729).

4.2.5.2 – Atraso

O atraso de pacote é entendido como a diferença de tempo, geralmente em milissegundos, entre o instante em que o terminal de origem envia o primeiro bit do pacote e o instante que o terminal de destino recebe esse bit. Seu comportamento é aleatório em função da carga, do meio e dos equipamentos da rede. No contexto de VoIP, este termo é referenciado como atraso fim-a-fim observado pela aplicação. Assim, o atraso pode ser definido como o tempo gasto entre a geração do sinal de voz pelo orador na origem e sua percepção pelo ouvinte no destino.

A Figura 4.4 mostra a escala de atrasos em milissegundos e os efeitos que esses produzem na qualidade de voz, conforme definido pelo ITU-T [1, 52].

Excelente	Bom	Regular	Pobre	Inaceitável
0	150	250 (ms)	350	450

Figura 4. 4 – Escala de atrasos e os efeitos que estes produzem na qualidade de voz.

Assim, os atrasos de transmissão, bem como os atrasos inseridos ao longo da rede, devem ser mantidos tão pequenos que não sejam perceptíveis aos usuários [1].

Em redes de pacotes, os atrasos podem ser gerados pelos seguintes processos [1]:

- Na formação do pacote: o tempo necessário para o preenchimento do pacote de voz. A média é de 20 a 30 ms para formação do pacote. Estes atrasos podem ser relativos a:

- conversão de sinais em interfaces com a rede telefônica;
 - processamento do sinal envolvendo funções realizadas pelo Codificador de Voz e DSP, tais como codificação e decodificação do sinal;
 - formação do quadro de voz, e;
 - tempo de empacotamento gasto pela CPU na formação dos pacotes a serem enviados pela rede e armazenamento (*buffer*) desse pacote até que o mesmo seja enviado pela rede;
- Na transmissão: o tempo gasto no transporte pela rede, da origem até o destino. Esses atrasos podem ser relativos a:
- atraso de acesso ao meio;
 - atraso de roteamento na rede, e;
 - atraso em *firewalls* e *proxy* com o objetivo de adicionar segurança ao sistema.

Além da avaliação dos atrasos descritos acima, eles podem ser analisados do ponto de vista fixo e variável. O atraso fixo pode ser entendido como o atraso fim-a-fim, que está sempre presente no processo de transmissão do pacote de voz, representando a soma dos vários atrasos provenientes dos diversos elementos da rede. O atraso variável é incremental, ou seja, eventual, causado por situações como congestionamento ou por equipamentos de voz das pontas [1].

A recomendação G.114 [52] da ITU-T define que a variação dos valores do atraso fixo no transporte pela rede está entre 71,5 a 130 ms. Então, o atraso variável pode ocorrer até o limite de 170 ms, para não haver perda na qualidade da voz, considerando o valor de 300 ms, o atraso total no transporte pela rede considerado na Figura 4.4.

4.2.5.3 – Variação de atraso (*jitter*)

Os atrasos variáveis no transporte pela rede, podem ser devido ao *jitter*. O *jitter* é a diferença de tempo na chegada dos pacotes. Para reduzir a incidência do *jitter*, são aplicadas ao processo, em alguns casos nos próprios algoritmos de compressão, regras para suavizar a variação do atraso. Este processo baseia-se em segurar os pacotes em uma parte da memória chamada de *jitter buffer* e, então, liberá-los para reprodução após a inclusão de um atraso previamente definido [1].

A RFC 3550 [20] fornece um exemplo do cálculo do *jitter* de forma mais precisa. Esta equação é utilizada nesta dissertação e reproduzida na Equação (4.2):

$$J_i = J_{i-1} + (|D_{i-1,i}| - J_{i-1}) / 16 \quad (4.2)$$

onde,

J_i - valor do *jitter* até o i-ésimo pacote

J_{i-1} - valor do *jitter* até o pacote anterior ao i-ésimo pacote

$$D_{i-1,i} = (R_i - R_{i-1}) - (T_i - T_{i-1}) = (R_i - T_i) - (R_{i-1} - T_{i-1})$$

R_i - momento (horário) do recebimento do i-ésimo pacote

R_{i-1} - momento (horário) do recebimento do pacote anterior ao i-ésimo pacote

T_i - horário da transmissão do i-ésimo pacote

T_{i-1} - horário da transmissão do pacote anterior ao i-ésimo pacote

4.2.5.4 – Fragmentação

A fragmentação é uma técnica amplamente utilizada na arquitetura de redes de pacotes. Utilizada no protocolo IP, o datagrama IP é segmentado em pacotes menores de modo que este possa trafegar por diferentes redes sem que seu tamanho exceda o limite definido da rede.

Os principais objetivos da técnica de fragmentação são [1]:

- reduzir o tamanho do pacote de voz de modo a evitar que um pacote com várias amostras de voz seja descartado ou chegue com atraso excessivo;
- aumentar a vazão de tráfego da rede;
- reduzir os atrasos;
- combinar o tráfego de pacotes de dados e voz em conexões de baixa velocidade;
- garantir a qualidade da voz, de modo que pequenos pacotes descartados possam ser substituídos sem degradação perceptível pelo usuário, e;
- combinada com a priorização dos pacotes, garante um fluxo constante de informação de voz, reduzindo o *jitter*, já que os pacotes de voz são transmitidos regularmente.

4.2.5.5 – Vazão, precisão e perda de pacotes

A vazão é representada pela quantidade dos dados transmitidos com sucesso e efetivamente recebidos no destino. Já a precisão é a relação do tráfego útil transmitido corretamente e o tráfego total despendido. Em termos ideais, a meta é que todos os dados enviados pela origem sejam recebidos no destino, sem erros, atraso, eco, etc. A medida de precisão é a taxa de erro de bits (*bit error rate* (BER)).

O tráfego de voz em tempo real é extremamente sensível ao atraso e à variação do atraso (*jitter*), mas tolerante à perda de pacotes. As perdas podem ser recuperadas de acordo

com o algoritmo utilizado em cada codificador. Esses algoritmos possuem técnicas de substituição do pacote perdido por outro pacote recebido. Além disso, os codificadores G.729 e G.723.1 utilizam a técnica de *look-ahead*. Esta técnica inclui no quadro corrente da voz amostras do quadro posterior, que podem ser utilizadas para recompor, pelo menos em parte, o pacote perdido. Por exemplo, para o G.729 o valor do *look-ahead* é de 5 ms, que contém 40 amostras do quadro posterior, além das 80 amostras do quadro corrente, ou seja, uma redundância de 50% na transmissão das amostras. O valor do *look-ahead* para o G.723.1 é de 7,5 ms, ou seja 25% de redundância na transmissão das amostras, lembrando que a duração do quadro de voz é de 30 ms [18, 42].

Os codificadores G.711, G.729 e G.723.1 possuem tolerâncias a perda de pacotes de 5%, 2% e 3%, respectivamente [17, 18, 42].

A Tabela 4.7 resume a relação entre os algoritmos dos codificadores e a MOS, largura de faixa, atraso, poder de processamento dos DSP exigidos pelos codificadores em *million instructions per second* (MIPS) e em quais aplicações eles estão mais presentes [1].

Tabela 4. 7 – Algoritmo de compressão e a classificação MOS.

Algoritmo	MOS	Taxa (kbps)	Atraso (ms)	MIPS	Aplicação
G.711 PCM	4,2	64	0,25	-	telefonia
G.726 ADPCM	4,1	32	0,25	10	telefonia
G.729 CS-ACELP	4,1	8	25	30	VoFR, VoATM, VoIP
G.728 LD-CELP	4,1	16	1,25	40	telefonia e VoFR
G.723.1 MP-MLQ	4,05	6,3	67,5	30	multimídia, VoFR e VoIP
G723.1 ACELP	3,55	5,3	67,5	30	multimídia, VoFR e VoIP

4.3 – *Quality of Service* (QoS)

Collins [3] define a QoS como “uma medida coletiva do nível de serviço entregue a um cliente”. Ela é considerada o nível de garantia fornecida a uma aplicação particular que a rede pode satisfazer seus requisitos de serviço. Em uma rede IP, a QoS pode ser medida em termos da largura de faixa, do atraso gerado, do *jitter*, da perda de pacotes e da vazão. Com a

finalidade de oferecer uma alta qualidade de serviço, a rede IP precisa fornecer garantias a uma determinada sessão ou a um conjunto de sessões, assim como os limites dessa garantia.

O IP oferece um serviço do tipo “melhor-esforço”, isto é, não garante a entrega da informação. Por essa razão, o TCP foi desenvolvido para funcionar na camada acima do IP, com a finalidade de assegurar a entrega das informações, livre de erros e na sequência. O TCP oferece esse serviço à custa de algum atraso, o que não é um atrativo para as aplicações em tempo real, tal como a voz. Essas aplicações são sensíveis ao tempo para longos atrasos e perdas de pacotes.

O UDP pode ser utilizado, desde que a perda de pacotes e o nível de congestionamento da rede sejam baixos. O tráfego em uma rede IP pode ser descontrolado, em rajadas e impossível de prever, o que pode levar a uma situação onde uma aplicação consome muito dos recursos da rede, mesmo que em um curto período de tempo, o que força os pacotes de outras aplicações a permanecerem em uma fila de espera ou serem descartados.

Conclui-se que nem o TCP nem o UDP podem resolver esses problemas, então são necessários outros meios para garantir uma melhor qualidade. Muitas soluções existem e outras estão sendo desenvolvidas.

4.3.1 – QoS fim-a-fim

A QoS deve ser fim-a-fim e deve ter o suporte de todas as redes do conjunto. Uma chamada pode originar na rede de um provedor e terminar em uma rede de outro provedor, inclusive percorrer outras redes para fazer isso. Cada uma das redes deve cooperar para assegurar que a qualidade fornecida seja como esperado. Este requisito cria o tema sobre Acordo de Níveis de Serviço (*Service-Level Agreement* (SLA)) entre os operadores ou provedores diferentes. Os SLA são acordos em que os operadores fazem compromissos entre si relativos ao tipo e a qualidade de serviço a serem oferecidos e as penalidades envolvidas se tais compromissos não forem encontrados.

Os SLA dificilmente podem ser implantados na Internet, pelo menos por enquanto, uma vez que a Internet é uma coleção enorme de redes diferentes, que não são administradas por uma só entidade. Na Internet, as partes não são necessariamente administradas da mesma maneira ou de acordo com o mesmo conjunto de critérios. Enquanto isso, os SLA podem ser possíveis entre certas concessionárias ou provedores VoIP, ou mesmo dentro de uma rede corporativa.

Isso não vale afirmar que uma boa política de QoS só pode ser alcançada quando uma chamada ficar dentro de uma única rede. Os SLA podem oferecer o potencial para assegurar a qualidade de múltiplas redes.

4.3.2 – Parâmetros de QoS para aplicações de VoIP

Na análise dos fatores que influenciam a qualidade da voz em uma rede VoIP, na seção 4.2.5 deste capítulo, foram descritos vários parâmetros como a largura de faixa, atraso, *jitter*, vazão, precisão e a perda de pacotes. Além desses, devem ser levados em consideração outros como a confiabilidade e a priorização dos pacotes de voz.

A confiabilidade depende do sistema de roteamento, que pode gerar atrasos, alterar a ordem ou mesmo efetuar descartes de pacotes causando distorções que degradam o desempenho da rede. De certa forma, ela também já foi mencionada na seção 4.2.5 deste capítulo.

O objetivo da priorização ou prioridade é reduzir o atraso dos pacotes pela configuração dos parâmetros de alguns protocolos de QoS. Estes mecanismos podem ser implementados para fornecer um balanceamento entre as transmissões de dados e voz, de maneira a garantir a largura de faixa suficiente para a transmissão dos pacotes de voz sempre que esses estiverem nas filas de transmissão, ou seja, eles terão prioridade de transmissão sobre os pacotes de dados [1]

4.3.3 – Resumo das soluções de QoS

As soluções de QoS devem considerar o problema sob vários ângulos. Uma abordagem é assegurar que os recursos para uma determinada sessão estarão disponíveis e reservados antes mesmo que a sessão seja estabelecida. Conceitualmente, esta solução tem certas semelhanças com a comutação de circuitos, onde a largura de faixa necessária a uma chamada é reservada antes que o tom de chamada (*ringing*) seja executado no telefone receptor.

Outra abordagem é classificar o tráfego em classes ou prioridades diferentes, com valores de prioridade mais alta atribuídos para as aplicações em tempo real e valores de prioridade mais baixa atribuídos para o tráfego de tempo não real. Esta abordagem é mais fácil de implementar, mas exige que nenhuma aplicação possa afetar a outra.

4.3.4 – Mecanismos de implementação de QoS

Alocar mais largura de faixa pode parecer uma abordagem simples e cara de QoS. Simples porque não exige desenvolvimento de sistemas importantes, e cara porque requer equipamentos caros e sofisticados. A disponibilidade desses equipamentos é necessária de forma que os recursos da rede estejam disponíveis em situações de congestionamento do tráfego. Infelizmente, esse investimento ficaria ocioso na maior parte do tempo. Este método seria uma forma ineficiente de resolver o problema de QoS, mas não deve ser desconsiderado completamente, porque uma largura de faixa adicional é necessária, mesmo que seja para suportar um tráfego adicional ou para uma demanda que continua crescente. Assim, são necessários outros mecanismos que administrem a largura de faixa disponível de forma mais eficaz para o suporte dos serviços oferecidos.

A IETF padronizou duas arquiteturas para agregar QoS ao modelo tradicional IP: *Intserv* [54] e *Diffserv* [53]. Ambas possuem o mesmo objetivo de propiciar diferenciação de

serviço. Entretanto, com abordagens distintas, assim como vantagens e desvantagens inerentes a cada uma, uma breve descrição das mesmas é apresentada nas seções 4.4.4.3 e 4.4.4.4 deste capítulo.

4.3.4.1 – *Dejitter buffer*

Os efeitos do *jitter* ou variações de atraso em VoIP podem ser eliminados ou reduzidos pela utilização de *buffers* na recepção denominados de *dejitter buffer*. A sua ação é armazenar temporariamente os pacotes de voz recebidos, introduzindo um atraso adicional antes de enviá-los ao receptor, igualando o atraso total sofrido por todos os pacotes.

A escolha do atraso para o *dejitter buffer* é crítica e para que ele seja capaz de eliminar completamente o efeito do *jitter*, sua janela de atraso deve ser igual à diferença entre o atraso máximo e mínimo na rede.

Se a variação do atraso na rede for muito grande a tentativa de eliminar o *jitter* aumentando a janela de tempo do *dejitter buffer* leva a pacotes sem *jitter*, mas com atrasos totais inaceitavelmente altos e dados inúteis. Neste caso uma solução deve ser adotada e um valor fixo aceitável para a janela de tempo do *dejitter buffer*, mas esta solução pode levar à perda de pacotes devido a atrasos excessivos na rede ou por *overflow* do *dejitter buffer*.

Para tornar a operação do sistema mais eficiente é desejável que o tamanho da janela de tempo do *dejitter buffer* seja adaptativo, isto é, variável de acordo com as circunstâncias de *jitter* da rede.

Pela medição e pela comparação contínua do atraso dos pacotes que chegam com o atraso de referência pode-se conseguir a adaptação do *dejitter buffer*. Esta referência deve ser continuamente atualizada, definindo um ajuste dinâmico do tamanho da janela do *dejitter buffer* ao longo do tempo, considerando que o atraso instantâneo não se afaste muito do atraso de referência adotado.

Em resumo, a escolha do tamanho da janela é uma solução do compromisso entre o atraso do pacote e a taxa de perda de pacotes. Janelas maiores resultam em maiores atrasos e menores perdas de pacotes, enquanto janelas menores resultam em atrasos menores e maiores perdas de pacotes.

4.3.4.2 – Serviço de melhor-esforço

No modelo de serviço de melhor-esforço o tráfego é processado tão rapidamente quanto possível, provê conectividade básica sem nenhuma garantia. As aplicações tolerantes a atrasos e perda de pacotes competem pelos recursos em forma de igualdade com aplicações sensíveis a esses requisitos. Quando existe ocorrência de congestionamento, pacotes são descartados pela ordem de chegada, independente da origem ou da aplicação.

Com o objetivo de atender aos diferentes requisitos das aplicações novos serviços devem ser fornecidos, possibilitando o suporte de diferentes níveis de QoS. Algumas abordagens têm sido propostas pelo IETF visando enfrentar o problema de oferecer garantias de QoS na rede, destacam-se os serviços integrados (*Integrated Services (Intserv)*) [54], os serviços diferenciados (*Differentiated Services (Diffserv)*) [53], a comutação de rótulo multiprotocolo (*Multiprotocol Label Switching (MPLS)*), e o Roteamento Baseado em Restrições (*QoS Routing*).

4.3.4.3 – Arquitetura *Intserv*

A *Intserv* [54] está apoiada em dois pilares: reserva de recursos e controle de admissão, ou seja, antes da transmissão dos pacotes iniciar, a aplicação faz a solicitação dos serviços, o caminho é configurado e os recursos previamente alocados. Especifica dois níveis de serviços, além do melhor esforço: serviço garantido (*Guaranteed*) para aplicações que necessitam de limites fixos de atraso e serviço de carga controlada (*Controlled-load*): seu

desempenho pode ser equiparado ao do melhor esforço sob condições de não-congestionamento [55].

Os elementos da rede no *Intserv* implementam quatro componentes [55]:

- Protocolo de sinalização: utilizado na configuração do caminho e reserva de recursos (*Resource Reservation Protocol (RSVP)*). Implica na manutenção de informações sobre o status de cada fluxo nos nós finais e em todos os roteadores ao longo do caminho;
- Escalonador de pacotes: lida com o envio de pacotes utilizando mecanismos de filas ou temporizadores (*timers*);
- Classificador (classificação *Multi-Field*): com a finalidade de controle de tráfego e contabilização, cada pacote deve ser mapeado para alguma classe, e;
- Rotina de controle de admissão: determina se um novo fluxo pode receber a QoS solicitada, sem gerar impacto às garantias anteriores.

4.3.4.3.1 – Reserva de recursos

A RFC 2205 [56] que descreve o Protocolo de Reserva de Recursos (*Resource Reservation Protocol (RSVP)*), especifica as técnicas de reserva de recursos para as redes IP. O RSVP é parte da série de serviços integrados do IETF e é um protocolo que possibilita a reserva de recursos a uma determinada sessão ou sessões, antes de qualquer tentativa de troca de mídia entre os participantes. Das soluções disponíveis, o RSVP é a solução mais complexa, mas também é a solução mais próxima das redes de comutação de circuitos dentro da rede IP. O RSVP fornece garantia de QoS, alocação significativa dos recursos da rede e retorno significativo às aplicações e aos usuários [3].

Basicamente, o RSVP trabalha como mostrado na Figura 4.5. Um primeiro remetente emite uma mensagem PATH à outra extremidade pelos vários roteadores. Essa mensagem contém uma especificação do tráfego (*traffic specification (TSpec)*) que fornece os detalhes

dos dados que o remetente deseja enviar, em termos de requisitos de largura de faixa e tamanho do pacote. Cada roteador habilitado com RSVP no caminho estabelecido inclui o endereço anterior da origem da mensagem PATH (isto é, o próximo salto anterior na direção ao remetente). O receptor da mensagem PATH responde com um pedido de reserva (*Reservation Request* (RESV)) que inclui uma especificação do fluxo (*flowspec*). Esta mensagem inclui o Tspec e as informações sobre o tipo de serviço da reserva solicitada, tal como o serviço *Controlled-load* ou o serviço *Guaranteed* [3].

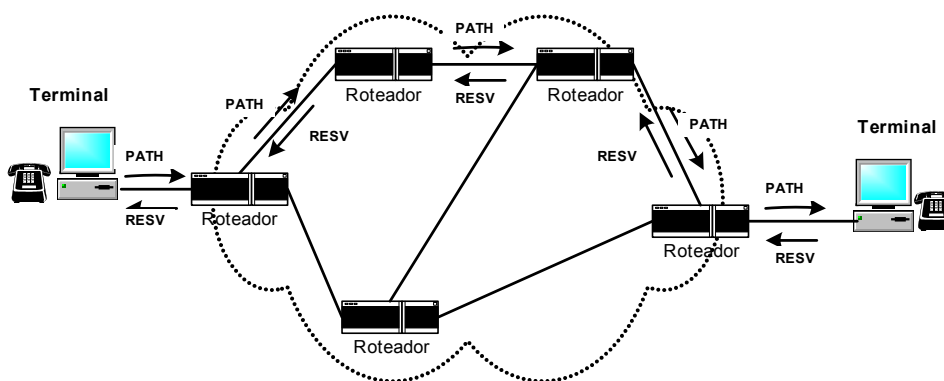


Figura 4. 5 – RSVP utilizado para reserva de recursos.

A mensagem RESV viaja de volta ao remetente ao longo da mesma rota tomada pela mensagem PATH, mas em sentido contrário. Em cada roteador, os recursos solicitados são alocados, assumindo que estão disponíveis e que o receptor tem a autoridade para fazer o pedido. Finalmente, a mensagem RESV alcança o remetente com uma confirmação que os recursos foram reservados.

Um ponto interessante sobre o RSVP é que o receptor, não o remetente dos dados, faz as reservas. Esta tarefa é apresentada com a finalidade de acomodar o transporte *multicast*, onde podem existir apenas um grande número de receptores e um remetente.

O RSVP é um protocolo de controle que não transporta dados do usuário. Estes dados (no caso de VoIP os dados são a voz) são transportados mais tarde usando o protocolo em tempo-real (*Real-Time Protocol* (RTP)). Este transporte acontece somente após os

procedimentos de reserva serem apresentados. As reservas que o RSVP faz são ditas maleáveis, porque elas precisam ser renovadas de forma regular pelo(s) receptor(es) .

4.3.4.4 – Arquitetura *DiffServ*

A RFC 2475 [53], que foi atualizada pela RFC 3260 [57], define os serviços diferenciados (*Differentiated Service (DiffServ)*) designados para priorizar diferentes tipos de tráfego. Basicamente, o protocolo *DiffServ*, faz uso do campo Tipo de Serviço (*Type of Service (TOS)*) da versão 4 do IP e do campo equivalente Classe de Tráfego (*Traffic Class*) da versão 6 do IP. A parte do campo de *TOS/Traffic Class* que o protocolo *DiffServ* usa é conhecida como campo *Differentiated Service (DS)*. Este campo é usado para marcar um determinado fluxo exigindo um tipo de encaminhamento particular. O tipo de encaminhamento a ser aplicado é conhecido como comportamento por salto (*Per-Hop Behavior (PHB)*), no qual o *DiffServ* define dois tipos: encaminhamento rápido (*Expedited Forwarding (EF)*) e encaminhamento assegurado (*Assured Forwarding (AF)*) [57].

A RFC 2597 [58], que também foi atualizada pela RFC 3260 [57], define o AF, que é um serviço em que os pacotes de uma determinada origem são encaminhados com alta probabilidade, desde que o tráfego daquela origem não exceda ao máximo pré-combinado. O AF define quatro classes, cada classe aloca recursos (espaço de *buffer* e largura de faixa) dentro de um roteador. Dentro de cada classe, um pacote de dado pode ter uma das três taxas de descarte (*three-drop*). Se em um determinado roteador existe congestionamento nos recursos alocados a uma determinada classe do AF, então os pacotes com os valores mais altos da taxa de descarte (*drop-rate*) são descartados primeiro, de forma que os pacotes com valores mais baixos da taxa de descarte recebem alguma proteção. Com a finalidade de melhorar a performance, o tráfego entrante não deve conter os pacotes com alta porcentagem de “baixa taxa de descarte”. O propósito é assegurar que os pacotes de mais alta prioridade

consigam atravessar em caso de congestionamento, o que não pode acontecer se todos os pacotes tiverem a prioridade mais alta [58].

A RFC 2598 [59], que foi substituída pela RFC 3246 [60], especifica o EF, que é um serviço em que a um determinado fluxo de tráfego é atribuída uma taxa de saída mínima de um determinado nó, que é maior do que a taxa de chegada no mesmo nó, desde que essa taxa não exceda ao máximo pré-combinado. Este processo assegura que os atrasos de enfileiramento serão removidos. Porque esses atrasos são as causas principais de um atraso fim-a-fim e do *jitter* e esse processo assegura que os atrasos e o *jitter* serão minimizados. O EF pode prover um serviço que é equivalente à alocação virtual de uma linha [60].

4.3.4.5 – MPLS

A comutação de etiquetas ganhou grande interesse na comunidade da Internet e nas redes Intranet, e foi devido à definição de um protocolo chamado comutação de etiquetas multi-protocolo (*Multi-Protocol Label Switching* (MPLS)), descrito pela RFC 3031 [61]. O MPLS é similar ao *DiffServ* na forma que ele marca o tráfego de entrada na rede. A função primária de marcar é não alocar prioridade dentro de um roteador, mas determinar o próximo roteador no caminho da origem até o destino.

O MPLS envolve o anexo de uma pequena etiqueta em um pacote na frente do cabeçalho IP [61]. Este procedimento é igual a inserir uma nova camada entre a camada IP e a camada de enlace subjacente da arquitetura TCP/IP. A etiqueta contém todas as informações que um roteador necessita para encaminhar um pacote. O valor de uma etiqueta pode ser usado para observar o próximo salto do caminho e encaminhar o pacote ao próximo roteador. A diferença entre esse roteamento e o roteamento padrão IP é que a combinação é exata e não a de observar pela combinação mais longa (a combinação com a máscara da sub-rede mais longa). Este processo possibilita as decisões mais rápidas de roteamento dentro dos roteadores.

A etiqueta identifica uma classe de equivalência de encaminhamento (*Forwarding Equivalence Class* (FEC)), que significa que todos os pacotes de uma determinada FEC são tratados igualmente para o propósito de encaminhamento. Todos os pacotes de uma determinada série de dados, como uma chamada de voz, terão a mesma FEC e receberão o mesmo tratamento de encaminhamento. Então, pode-se assegurar que o tratamento do encaminhamento aplicado a um determinado fluxo pode ser configurado de tal forma que todos os pacotes de A até B seguem exatamente o mesmo caminho. Se este fluxo tem um requisito particular de largura de faixa, então esta largura deve ser alocada no início da sessão. Esta característica pode assegurar que um determinado fluxo tem a largura de faixa que necessita e que os pacotes que compõem o fluxo chegarão na mesma sequência transmitida. Conseqüentemente, essa função provê uma QoS mais alta.

O MPLS é tanto um protocolo de engenharia de tráfego como um protocolo de QoS. Ele é parecido com o estabelecimento dos circuitos virtuais do modo de transferência assíncrona (*Asynchronous Transfer Mode* (ATM)) e pode conduzir a semelhantes benefícios de QoS. O MPLS ajuda a fornecer maior QoS ajudando a administrar melhor o tráfego.

4.3.4.6 – Compressão de cabeçalho

A RFC 2508 [51] descreve um método para comprimir os cabeçalhos dos datagramas dos protocolos IP/UDP/RTP para reduzir a sobrecarga do cabeçalho (*overhead*) em enlaces seriais de baixa velocidade. Esse método consiste no não re-envio das informações de cabeçalhos que não mudam após o estabelecimento da chamada, bem como da multiplexação dos dados de voz de dois ou mais canais de chamadas em um pacote IP com um sub-cabeçalho identificando os canais de chamadas, de forma a elevar a eficiência da transmissão de VoIP.

A Figura 4.6 mostra a estrutura do quadro VoIP para Ethernet.

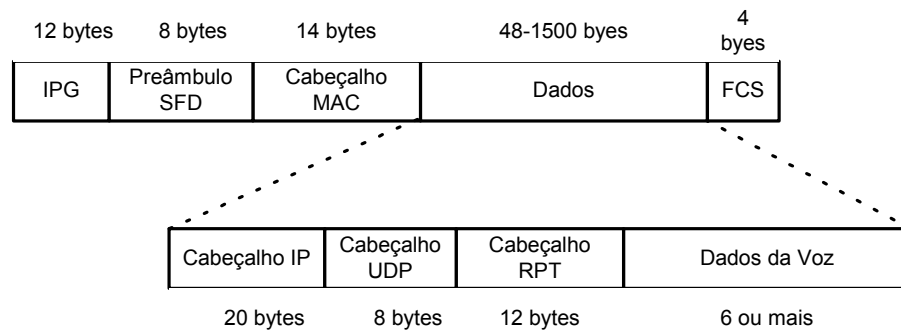


Figura 4. 6 – Estrutura do quadro VoIP na Ethernet.

No cálculo do tamanho do pacote a ser transmitido, um cabeçalho comprimido de 4 bytes é utilizado como resultado do protocolo de transporte em tempo-real comprimido (*Compressed Real-time Transport Protocol (CRTP)*).

4.3.4.7 – Detector de atividade da voz

O detector de atividade de voz (*Voice Activity Detector (VAD)*), é uma técnica eficiente de supressão de silêncio utilizada nos equipamentos de tratamento de voz sobre redes de pacotes. O VAD pode proporcionar uma economia de até 50% da faixa, em que os períodos de silêncio não são transmitidos, permitindo que essa faixa seja alocada para outras aplicações [1].

Na simulação desta dissertação não será utilizada a proporção de redução de faixa com o VAD, porque ele somente afetará os tamanhos dos pacotes em termos proporcionais para todos os pacotes e sua utilização não é o objetivo principal de análise neste trabalho. Porém, na prática o VAD é uma ferramenta poderosa no aproveitamento da largura de faixa do meio.

4.3.5 – Políticas de QoS

Devido à existência de vários mecanismos de QoS, fica a dúvida de qual o nível de QoS deve ser aplicado e a que tipos de tráfego. Este dilema cria o tema das políticas de QoS. Enquanto as arquiteturas tais como o *IntServ*, o *DiffServ* e o MPLS fornecem os mecanismos

para distinguir o tráfego, as políticas de QoS especificam como esses mecanismos devem ser utilizados.

Uma parte importante de QoS é o fato que certos usuários terão um melhor serviço que outros. Esta situação incentiva a outros usuários em utilizar um melhor serviço sem pagar por isso. Logo, existe a necessidade das funções de autenticação para identificar usuários e assegurar que um determinado usuário é quem ele reivindica ser. Além disso, um usuário pode ser certificado a um determinado nível de QoS sob certas circunstâncias, mas não sob outras. Então, certas regras devem ser estabelecidas na especificação de quais circunstâncias e quais ações e métodos devem existir para a certificação. Esta função pode ser considerada um tipo de política de função. Assim, o IETF desenvolveu um protocolo conhecido como protocolo de serviço de política comum aberta (*Common Open-Policy Service Protocol* (COPS)), definido pela RFC 2748 [62]. Este protocolo é um protocolo cliente-servidor, como mostrado na Figura 4.6 deste capítulo. Um ponto de execução de política (*Policy Enforcement Point* (PEP)), como roteador que precisa impor certas regras, pergunta a um Ponto de Decisão de Política (*Policy Decision Point* (PDP)) que faz as decisões da política atual. O PEP pode ser considerado o policial e o PDP pode ser considerado o juiz, fazendo a analogia ao termo *cop* (policial em inglês), o acrônimo COPS fica particularmente ajustado [62].

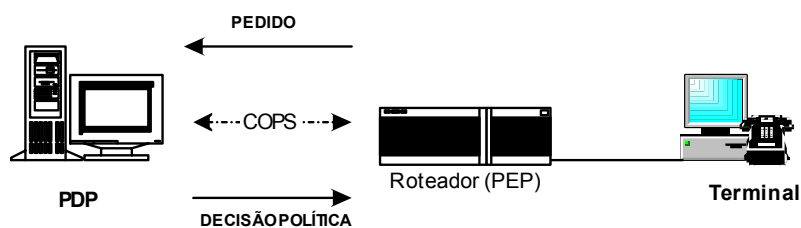


Figura 4. 7 – Pedido e resposta do COPS.

4.3.6 – Combinando as soluções de QoS

As soluções de QoS descritas neste capítulo consideram abordagens diferentes e cada uma possui suas vantagens e desvantagens. Enquanto o RSVP é poderoso, ele possui a

desvantagem de exigir o estado do caminho a ser mantido em cada roteador, algo que pode ser difícil e caro em uma rede grande. O *DiffServ* é mais simples, mas é mais uma técnica de priorização do que um mecanismo de garantia de recurso. O MPLS oferece a promessa como uma solução global, mas exige mudanças significativas para todos os roteadores que querem usar esse método. Então, cada uma das soluções possui suas desvantagens e nenhuma solução única isolada é provável ser a única solução. Combinando soluções diferentes, pode-se possibilitar a utilização em partes diferentes da rede, onde elas podem ser melhor aproveitadas.

4.4 – Considerações Finais deste Capítulo

Este capítulo apresentou os codificadores de voz mais utilizados na aplicação de VoIP, as técnicas de medição, os fatores que influenciam na qualidade da voz, os conceitos fundamentais de qualidade de serviço em uma rede VoIP, assim como apresentou os principais mecanismos disponíveis atualmente para a sua implementação.

CAPÍTULO V

SIMULAÇÕES E RESULTADOS OBTIDOS

5.1 – Introdução

A medição, a simulação e a modelagem analítica são as técnicas de avaliação mais utilizadas. A técnica de simulação é, sem dúvida, a mais utilizada nas áreas de rede e telecomunicação, possibilitando ao pesquisador estudar diferentes cenários devido a sua flexibilidade na alteração dos parâmetros e das situações.

Nesta dissertação foi utilizado o simulador de redes *Network Simulator* (NS). Ele é uma ferramenta bastante conhecida no meio científico e acadêmico por ser um *software* de livre distribuição.

Este capítulo apresenta o simulador NS, descreve as considerações e a topologia da simulação realizada neste trabalho, mostra os resultados obtidos com os testes realizados. Finalmente, são realizadas conclusões sobre os resultados obtidos nesses testes.

5.2 – O Simulador de Redes NS

O NS é um simulador de evento discreto e orientado a objetos, direcionado para pesquisas em redes. Ele fornece suporte significativo para simulação de protocolos TCP , UDP, roteamento *unicast* e *multicast* sobre redes com fio, sem fio, locais, WAN e satélite.

O NS surgiu em 1989 como uma variante do simulador de rede *Realistic and Large* (REAL) e evoluiu substancialmente durante vários anos. Em 1995, o desenvolvimento do NS foi suportado pela Agência de Projetos de Pesquisas Avançadas em Defesa (*Defense Advanced Research Projects Agency* (DARPA)) pelo projeto da *Virtual InterNetwork TestBed* (VINT) no *Lawrence Berkeley National Laboratory* (LBNL) da Universidade da Califórnia, Berkeley (UCB) [63]. Desde a sua criação o NS conta com substanciais contribuições de outros pesquisadores do mundo inteiro.

O núcleo do simulador NS é escrito em C++ e utiliza na modelagem de simulações os comandos *Otcl* na configuração de interfaces e na construção dos *scripts*. Na construção de um *script* pode-se definir uma grande variedade de topologias de rede, como endereço dos nós, taxa de transmissão, atraso dos enlaces, módulos e políticas de roteamento e uma vasta biblioteca de protocolos e vários tipos de aplicações.

A distribuição do NS e de seu código fonte é gratuita. Pode ser obtida pela *web* [63], mantida pela *Information Sciences Institute* (ISI), financiada pela DARPA. O NS é compatível com vários sistemas operacionais, dentre eles o *Windows* e o *Linux*, tem funcionalidade significativa para simular topologias de rede e diferentes modelos de tráfego e possui arquitetura aberta que permite aos usuários adicionar novas funcionalidades.

A versão do NS utilizada na simulação desta dissertação é a *ns-allinone-2.27* [63] e o sistema operacional utilizado foi o *Mandrake Linux* versão 10.0.

5.2.1 – Suporte ao *Diffserv* no NS

O código e o suporte do *Diffserv* foram adicionados ao NS pela *Nortel Networks* [63], em 5 módulos adicionados à hierarquia de classe: um para a funcionalidade do roteador de base *Diffserv* (**dsRED**), um para cada roteador de borda e de núcleo, um para o enfileiramento baseado em RED e um para as políticas [63].

O suporte ao *Diffserv* no NS é realizado pelas funcionalidades capturadas por classes de filas, onde cada módulo define uma classe [63].

O módulo **dsRED** é a base para a implementação do *Diffserv*. Ele define a classe **dsREDQueue**, que é a classe pai para as classes **edgeQueue** e **coreQueue**. Estas classes implementam toda a funcionalidade e declaram todos os parâmetros que são comuns aos roteadores de borda e de núcleo. O módulo **dsRED** está contido nos arquivos **dsred.h** e **dsred.cc** do diretório “~ns-2.27/diffserv/” do simulador [63].

Uma **dsREDQueue** usa a classe **redQueue**, que forma uma estrutura composta de 4 filas físicas e 3 filas virtuais para cada fila física. As filas virtuais definem o nível de preferência de descarte. Os pacotes são enfileirados com um número de preferência de descarte de acordo com sua marca de *code point*. Eles são tratados conforme os parâmetros correspondentes à fila e ao número de preferência. Cada fila da classe de serviço contém os mecanismos de congestionamento providos pela classe **redQueue**. A classe **dsREDQueue** contém uma estrutura de dados conhecida como tabela de comportamento por salto (*Per Hop Behaviour Table* (PHB)). Os dispositivos de borda controlam marcando os pacotes com os pontos de código e os dispositivos de núcleo simplesmente respondem aos pontos de código existentes. Porém, ambos os dispositivos precisam determinar como mapear um ponto de código para uma fila e o nível de preferência particular. A tabela do PHB controla este mapeamento definindo um vetor com três campos: **Code Point** (ponto de código), **Class** (da fila física) e **Precedence** (da fila virtual) [63].

Em síntese, o *Diffserv* do NS provê um classificador de multi-campos (*multi-field* (MF)) baseado no endereço de origem e de destino para medir a seleção e um classificador **DSCP** para selecionar a fila do tráfego de fluxo; medidores tipo **TSW2CM**, **TSW3CM**, **Token Bucket**, **srTCM**, e **trTCM**; temporizadores **RR**, **WRR**, **WIRR** e **PQ**; e os algoritmos de descarte **RIO-C**, **RIO-D**, **WRED** e **Drop-on-Thereshold** [63].

5.2.2 – Geração de tráfego VoIP no NS

As várias formas de geração de tráfego de VoIP no NS, são a própria implementação realizada pelo usuário na definição de uma fonte geradora de tráfego em C++ ou fazendo uso dos agentes geradores de tráfego do NS e a utilização de um arquivo contendo dados de voz gravados e codificados previamente, como o G.723.1, G.729, etc., que utiliza o agente *application/traffic/trace traffic generator*, existente no NS. A mais utilizada no meio acadêmico são os geradores de tráfego do NS, como *Constant Bit Rate* (CBR), *Exponential on/off* (EXP), *Pareto on/off* (PAR), ou Poisson. Estes geradores emulam a aplicação desejada. A utilização desses geradores deve ser feita analisando o comportamento da voz codificada em tempo-real e reproduzindo esse comportamento no simulador.

Os parâmetros mais importantes e disponíveis na configuração dos geradores de tráfego da voz, são:

- **PacketSize_**: tamanho dos pacotes gerados;
- **burst_time_**: tempo médio de **on** para o gerador (intervalos de atividade);
- **idle_time_**: tempo médio de **off** para o gerador (intervalos de pausa);
- **rate_**: taxa de transmissão durante um determinado período de tempo;
- **random_**: flag cujo objetivo é indicar se introduz ou não um ruído aleatório nos tempos marcados como *start* (o padrão é **off**);
- **maxpkts_**: número máximo de pacotes gerados pela fonte (se não for mencionado o padrão é 2^{28});
- **shape_**: forma, parâmetro utilizado pelo objeto de **Pareto on/off**; e;
- **interval_**: intervalo entre os pacotes (campo opcional). O manual do NS informa que nos objetos CBR os campos **rate_** e **interval_** são mutuamente exclusivos,.

5.3 – Métricas

As métricas básicas utilizadas nesta dissertação para a avaliação do impacto do comprimento dos dados de voz no pacote IP em uma rede VoIP com suporte *Diffserv*, são:

- **Vazão (*throughput*):** quantidade de dados transmitidos com sucesso e efetivamente recebidos no destino. Ela é representada pela ocupação da largura de faixa no meio pela fonte;
- **Atraso (*delay*):** atraso fim-a-fim entre a fonte e o destino. Ele é representado pela diferença entre os tempos de saída e de chegada dos pacotes que estão sendo analisados, e;
- **Variação do atraso (*jitter*):** diferença entre os tempos de atraso de dois pacotes consecutivos recebidos no destino.

Visando tornar mais interessante a análise do atraso fim-a-fim por pacote, foram inclusos no *script* do cálculo os atrasos de codificação e decodificação da voz, considerando a origem e o destino com os *gateways* de voz sobre IP (VoIP-GW) na origem e no destino. Esse cálculo tem como objetivo a verificação de possíveis atrasos fim-a-fim superiores a 300 ms, para comparação com os valores definidos pela ITU-T [52], conforme mostrado na Figura 4.4 deste trabalho, porque esses atrasos podem reproduzir uma qualidade de voz de pobre a inaceitável.

O cálculo do atraso fim-a-fim considerando os *gateways* de voz (VoIP-GW) foi realizado somando-se ao atraso do pacote os tempos de codificação na origem e decodificação no destino, mais os tempos de duração dos pacotes. Nesta dissertação, para que esse atraso não seja confundido com o atraso fim-a-fim calculado na simulação até a camada de rede, ele será denominado latência fim-a-fim. Este cálculo foi realizado conforme a Equação 5.1:

$$\text{Latência fim-a-fim} = AF + AV + DP + CO + CD \quad (5.1)$$

onde:

AF – soma dos atrasos fixos dos enlaces da rede simulada (neste caso é 50 ms);

AV – atraso variável gasto entre a fonte geradora do tráfego VoIP e o destino;

DP – tempo de duração dos pacotes;

CO – tempo gasto na codificação da voz no *Gateway* de VoIP na origem, e;

CD – tempo de decodificação da voz no *Gateway* de VoIP no destino.

Os tempos de duração dos pacotes considerados estão informados nas Tabela 5.1 e 5.3, conforme o cenário simulado. Os tempos de duração dos pacotes foram somados ao cálculo da latência fim-a-fim porque a reprodução da voz somente ocorrerá após a chega do último bit do pacote e a decodificação da voz no destino. O tempo de codificação da voz na origem considerado foi de 20 ms e o tempo de decodificação da voz no destino foi de 10 ms, conforme informação do Kostas [64] que os atrasos típicos de decodificação é da ordem da metade dos atrasos de codificação. O resultado desse cálculo fornece o percentual dos pacotes recebidos com atrasos superiores a 300 ms. O gráfico gerado por este cálculo será mostrado em separado do atraso médio fim-a-fim da topologia simulada, para melhor visualização desse cálculo.

5.4 – Modelo da Simulação deste Trabalho

A topologia utilizada na simulação é constituída de um domínio *Diffserv* de quatro *edge routers* (Ein, E1, E2 e Eout) e dois *core routers* (C1 e C2), uma fonte de tráfego VoIP e três fontes de tráfego de fundo nas bordas (Ein, E1 e E2).

A Figura 5.1 mostra a topologia completa da simulação utilizada neste trabalho.

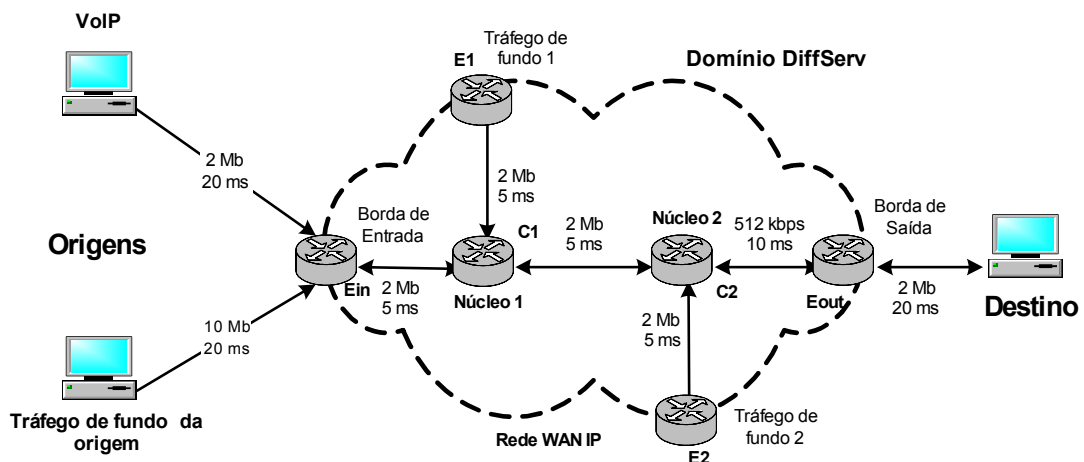


Figura 5. 1 – Topologia da simulação utilizada neste trabalho.

As fontes de tráfego foram configuradas para gerar diferentes tipos de tráfego, como:

- *File Transfer Protocol* (FTP) – para as fontes de tráfego de fundo, com janelas de 4000 bytes;
- *Constant Bit Rate* (CBR) com **on/off** – a fonte de VoIP foi modelada com processos **on/off**, distribuídos com durações de 1,004 segundos nos períodos de atividade (*burst*) e de 1,587 segundos nos períodos de pausa (*idle*) [66]. Os intervalos entre os pacotes, a duração dos pacotes com característica de cada codificador e o número de quadros de voz em cada pacote, foram utilizados conforme mostrado na Tabela 5.2 deste trabalho. Na construção dos pacotes das fontes de geração de tráfego VoIP, os cabeçalhos do IP, UDP, RTP e RTCP foram incluídos no tamanho do pacote gerado, com o cabeçalho comprimido de 4 bytes, conforme RFC 2508 [51];

Parâmetros de simulação considerados são:

- Os valores de atraso fixo dos enlaces foram escolhidos de forma a possibilitar um comportamento da rede simulada mais próximo da prática, dentro das características exigidas pela voz em tempo real;

- as características e as localizações das fontes de tráfego de fundo (*background*) foram escolhidas de forma a gerar uma sobrecarga na rede com uma vazão possível;
- o tempo de *warm-up* (tempo efetivo de espera para a computação das medições) e de *start* das fontes de tráfego foi considerado como, aproximadamente, 10% do tempo total da simulação, ou seja, 30 ms, conforme sugerido por Coutinho [65];
- a técnica de diferenciação de serviços (*Diffserv*) foi escolhida como QoS porque a análise proposta não implica na decisão de roteamento dos pacotes, não sendo necessária a inclusão de outras técnicas como provisionamento para este caso específico ou de engenharia de tráfego, tais como o MPLS e o RSVP;
- O serviço utilizado para tráfego da voz no *Diffserv* foi o AF, em comparação com a telefonia tradicional de comutação de circuitos, que utiliza os enlaces de baixa velocidade para o tráfego da voz e os enlaces de alta velocidade para tráfego da sinalização (SS7). Assim, este trabalho considera que o serviço EF do *Diffserv* seria utilizado para os fluxos de sinalização, devido às suas características serem muito parecidas com as dos circuitos virtuais do ATM, e;
- os enlaces de acesso VoIP, tráfego de fundo e entre os núcleos da nuvem *Diffserv* possuem uma capacidade bastante superior ao tráfego de gargalo, de forma a provocar um congestionamento na rede que está sendo simulada;

Neste trabalho foram realizados dois cenários na simulação, utilizando os mesmos recursos de rede da topologia e as variações nas cargas de trabalho das fontes de VoIP. As fontes para o primeiro cenário foram modeladas com as características de geradores de tráfego como mostrado na Tabela 5.1. A quantidade de quadros por pacote varia de acordo com o tempo de simulação para a fonte de tráfego VoIP. Essa quantidade é de 1 a 5 quadros por pacote para cada codificador, G.711 e Anexo A do G.729.

Tabela 5. 1 – Fontes de tráfego utilizadas no primeiro cenário.

Fonte	Destino	Taxa (kbits)	Atraso (ms)	Suporte	Tamanho do Pacote (bytes)	Duração do pacote (ms)	PHB	Faixa	Tempo (seg)	
									Início	Fim
G711_1	Ein	64	20	DiffServ	170	20	AF	64 K	30	50
G711_2	Ein	64	20	DiffServ	330	40	AF	64 K	55	75
G711_3	Ein	64	20	DiffServ	490	60	AF	64 K	80	100
G711_4	Ein	64	20	DiffServ	650	80	AF	64 K	105	125
G711_5	Ein	64	20	DiffServ	810	100	AF	64 K	130	150
G729_1	Ein	8	20	DiffServ	30	20	AF	8 K	155	175
G729_2	Ein	8	20	DiffServ	50	40	AF	8 K	180	200
G729_3	Ein	8	20	DiffServ	70	60	AF	8 K	205	225
G729_4	Ein	8	20	DiffServ	90	80	AF	8 K	230	250
G729_5	Ein	8	20	DiffServ	110	100	AF	8 K	255	275
SBT	Ein		5	DropTail			BE	2 M	0	285
Ein	C1		5	DiffServ			BE	2 M	0	285
E1	C1		5	DiffServ			BE	2 M	0	285
E2	C2		5	DiffServ			BE	2 M	0	285
C1	C2		5	DiffServ			BE/AF	2 M	0	285
C2	Eout		20	DiffServ			BE/AF	512K	0	285
Eout	Ndest		20	DropTail			BE	2 M	0	285

O mapa dos valores utilizados na topologia da simulação do primeiro cenário é mostrado na Tabela 5.2.

Tabela 5. 2 – Mapa dos valores utilizados na topologia do primeiro cenário.

Nó da Rede	Agente	Aplicação	Fonte	Destino	PHB	Taxa	Faixa do enlace (MB)	Tempo (seg)		Quadros de voz por pacote
								Início	Fim	
VoIP	UDP	CBR	G711_1	Ein	AF	64K	2	30	50	1
VoIP	UDP	CBR	G711_2	Ein	AF	64K	2	55	75	2
VoIP	UDP	CBR	G711_3	Ein	AF	64K	2	80	100	3
VoIP	UDP	CBR	G711_4	Ein	AF	64K	2	105	125	4
VoIP	UDP	CBR	G711_5	Ein	AF	64K	2	130	150	5
VoIP	UDP	CBR	G729_1	Ein	AF	8K	2	155	175	1
VoIP	UDP	CBR	G729_2	Ein	AF	8K	2	180	200	2
VoIP	UDP	CBR	G729_3	Ein	AF	8K	2	205	225	3
VoIP	UDP	CBR	G729_4	Ein	AF	8K	2	230	250	4
VoIP	UDP	CBR	G729_5	Ein	AF	8K	2	255	275	5
SBT	TCP	FTP	SBT	Ein	BE	10 M	10	0	285	
E1	TCP	FTP	E1	C1	BE	2 M	2	0	285	
E2	TCP	FTP	E2	C2	BE	2 M	2	0	285	

As fontes utilizadas na topologia da simulação deste trabalho para o segundo cenário foram modeladas com características de geradores de tráfego como mostrado na Tabela 5.3.

Tabela 5. 3 – Fontes de tráfego utilizadas no segundo cenário.

Fonte	Destino	Taxa (Kbits)	Atraso (ms)	Suporte	Tamanho do Pacote (bytes)	Duração do pacote (ms)	PHB	Faixa	Tempo (seg)	
									Início	Fim
G729_1	Ein	8	20	DiffServ	30	20	AF	8 K	30	50
G729_2	Ein	8	20	DiffServ	50	40	AF	8 K	55	75
G729_3	Ein	8	20	DiffServ	70	60	AF	8 K	80	100
G729_4	Ein	8	20	DiffServ	90	80	AF	8 K	105	125
G729_5	Ein	8	20	DiffServ	110	100	AF	8 K	130	150
G729_6	Ein	8	20	DiffServ	130	120	AF	8 K	155	175
G729_7	Ein	8	20	DiffServ	150	140	AF	8 K	180	200
G729_8	Ein	8	20	DiffServ	170	160	AF	8 K	205	225
G729_9	Ein	8	20	DiffServ	190	180	AF	8 K	230	250
G729_10	Ein	8	20	DiffServ	210	200	AF	8 K	255	275
SBT	Ein		5	DropTail			BE	2 M	0	285
Ein	C1		5	DiffServ			BE	2 M	0	285
E1	C1		5	DiffServ			BE	2 M	0	285
E2	C2		5	DiffServ			BE	2 M	0	285
C1	C2		5	DiffServ			BE/AF	2 M	0	285
C2	Eout		20	DiffServ			BE/AF	512 K	0	285
Eout	Ndest		20	DropTail			BE	2 M	0	285

O mapa dos valores utilizados na topologia da simulação do segundo cenário é mostrado na Tabela 5.4.

Tabela 5. 4 – Mapa dos valores utilizados no segundo cenário.

Nó da Rede	Agente	Aplicação	Fonte	Destino	PHB	Taxa	Faixa do enlace (Mb)	Tempo (seg)		Quadros de voz por pacote
								Início	Fim	
VoIP	UDP	CBR	G729_1	Ein	AF	8K	2	30	50	1
VoIP	UDP	CBR	G729_2	Ein	AF	8K	2	55	75	2
VoIP	UDP	CBR	G729_3	Ein	AF	8K	2	80	100	3
VoIP	UDP	CBR	G729_4	Ein	AF	8K	2	105	125	4
VoIP	UDP	CBR	G729_5	Ein	AF	8K	2	130	150	5
VoIP	UDP	CBR	G729_6	Ein	AF	8K	2	155	175	6
VoIP	UDP	CBR	G729_7	Ein	AF	8K	2	180	200	7
VoIP	UDP	CBR	G729_8	Ein	AF	8K	2	205	225	8
VoIP	UDP	CBR	G729_9	Ein	AF	8K	2	230	250	9
VoIP	UDP	CBR	G729_10	Ein	AF	8K	2	255	275	10
SBT	TCP	FTP	SBT	Ein	BE	10 M	10	0	285	
E1	TCP	FTP	E1	C1	BE	2 M	2	0	285	
E2	TCP	FTP	E2	C2	BE	2 M	2	0	285	

Para o segundo cenário a variação da quantidade de quadros por pacote para a fonte de tráfego VoIP é 1 a 10 quadros por pacote para o codificador Anexo A do G.729.

O objetivo do segundo cenário é a verificação da tendência dos valores da vazão, do atraso e da variação do atraso obtidos no primeiro cenário conforme aumenta o comprimento dos dados de voz no pacote IP, bem como definir o tamanho máximo do comprimento dos dados de voz para o codificador G.729 para a rede simulada, considerando o atraso máximo de 300 ms, conforme recomendação G.114 da ITU-T [52].

5.5 – Condução dos Experimentos e Resultados Obtidos

Com objetivo de verificar o impacto do comprimento dos dados de voz no pacote IP em termos da qualidade de serviço, serão analisadas as consequências devido às variações do comportamento da qualidade de serviço fim-a-fim em um ambiente de serviços diferenciados, verificando os resultados pela variação da ocupação da largura de faixa do meio, do atraso e do *jitter* dos pacotes das fontes geradoras de tráfego de voz.

O tráfego é gerado por uma fonte de tráfego CBR e o condicionamento é feito pelo medidor e policiador TSW2CM, as políticas de descarte são realizadas pela detecção antecipada aleatória ponderada (*Weighted Random Early Detection* (WRED)), onde todas as probabilidades de descarte são baseadas em um único comprimento de fila.

O experimento realizado neste trabalho consiste em enviar fluxos de dados de voz, empacotados com compressão do cabeçalho de 4 bytes, variando a quantidade de quadros de voz por pacote nos codificadores G.711 e G.729, conforme o cenário, de um nó de origem a um nó de destino, passando pelos roteadores de borda e de núcleo no domínio *Diffserv*.

Com a finalidade de possibilitar as medidas de desempenho da rede em cada cenário, foram determinadas as políticas de encaminhamento dos pacotes do tráfego agregado AF e BE nas filas de saída dos roteadores da borda de entrada **Ein**, do núcleo **C2** e da borda de saída **Eout**.

Os cálculos dos valores da média das 10 simulações de cada cenário, a variância, o intervalo de confiança e os gráficos, foram obtidos pelas planilhas do programa Microsoft Excell XP. Esses valores foram obtidos como resultados da simulação pelo arquivo *trace* de saída do *script* utilizado (**wtrace.tr**). Os valores da média da largura de faixa, do atraso e do *jitter* foram obtidos pelo *script* TCL, e posteriormente foram transferidos para o sistema Windows XP, para facilitar a visualização desses resultados.

5.5.1 – Primeiro cenário – G.711/G.729 de 1 a 5 quadros por pacote

O fluxo do tráfego agregado **AF**, isto é, da fonte geradora VoIP, foi definido em 64 Kbps para as fontes G.711 e 8 Kbps para as fontes G.729, com intervalos interpacotes determinados pela duração dos pacotes conforme mostrado na Tabela 5.2 deste capítulo.

Os PHB, as métricas e o perfil de descarte dos pacotes foram definidos como mostrado na Tabela 5.5 deste capítulo, com o objetivo de evitar o descarte dos pacotes de voz, ao mesmo tempo minimizando o descarte dos pacotes do tráfego de melhor esforço (BE) para a topologia utilizada neste trabalho. Os outros parâmetros necessários para a configuração do domínio *Diffserv* foram mostrados nas Tabelas 5.1 e 5.2, da seção 5.4 deste capítulo.

Tabela 5. 5 – Parâmetros utilizados no primeiro cenário da simulação.

Classe do Tráfego	Tipo de Tráfego	Codificador	PHB	DSCP	Métricas	Descarte
AF	voz	G711	AF	10	TSW2CM, CIR 64 Kbps	descarte fora do perfil
AF	voz	G729	AF	10	TSW2CM, CIR 8 Kbps	descarte fora do perfil
BE	ftp		BE	0	TSW2CM, CIR 2 Mb	descarte fora do perfil

Como o tráfego de fundo foi gerado com o objetivo de congestionamento no enlace de gargalo **C2Eout** da rede utilizada nos testes, o mesmo não será analisado. O objetivo deste trabalho é analisar o comportamento dos pacotes no tráfego de voz.

A Figura 5.2 mostra a vazão média obtida nos testes realizados no primeiro cenário deste trabalho.

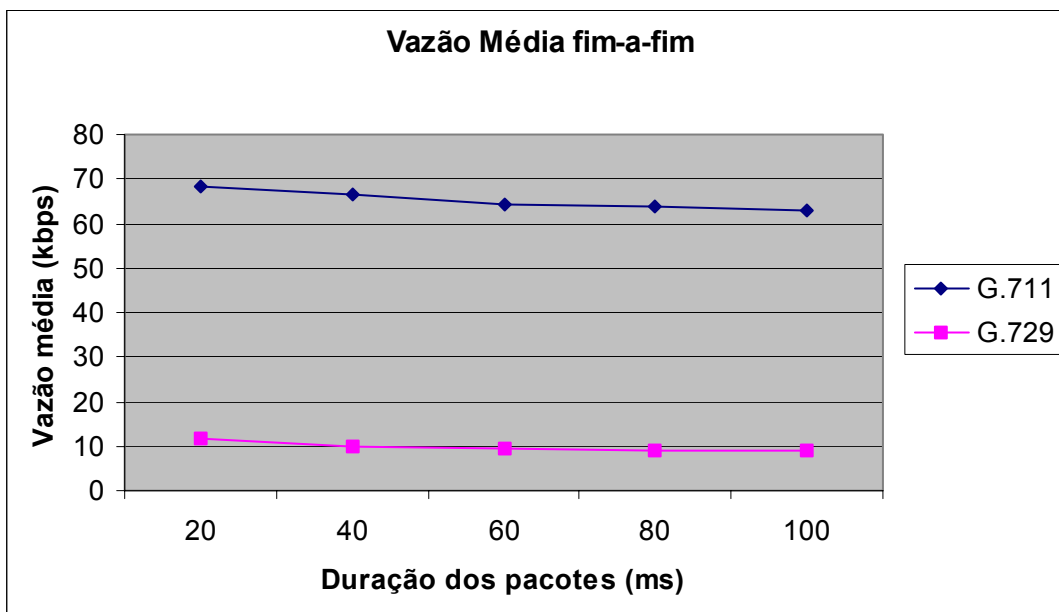


Figura 5. 2 – Vazão média obtida no primeiro cenário.

Pode-se observar na Figura 5.2 que ocorreu uma redução da vazão média fim-a-fim com o aumento do tamanho dos dados de voz. Como não ocorreram perdas de pacotes de voz, isso representa a média da largura de faixa ocupada no meio em relação à quantidade de pacotes de voz transmitidos e efetivamente recebidos no destino.

Pode-se verificar na Figura 5.2 uma pequena redução na vazão fim-a-fim do codificador G.729, variando a quantidade de quadros por pacote de 1 a 5, em relação a uma maior redução na vazão do codificador G.711 com a mesma variação da duração dos pacotes, sendo que o tamanho dos quadros de voz no PCM é bastante superior ao tamanho dos quadros do G.729 A, em bytes.

Os resultados obtidos na Figura 5.2 mostram uma redução da vazão fim-a-fim para os dois codificadores enquanto aumenta a quantidade de quadros por pacotes. Conseqüentemente, melhorando a eficiência de transmissão.

Utilizando os tempos de transmissão na origem e de recebimento no destino dos pacotes, foi calculado o atraso por pacote. A Figura 5.3 mostra o gráfico da média do atraso fim-a-fim dos pacotes de cada codificador.

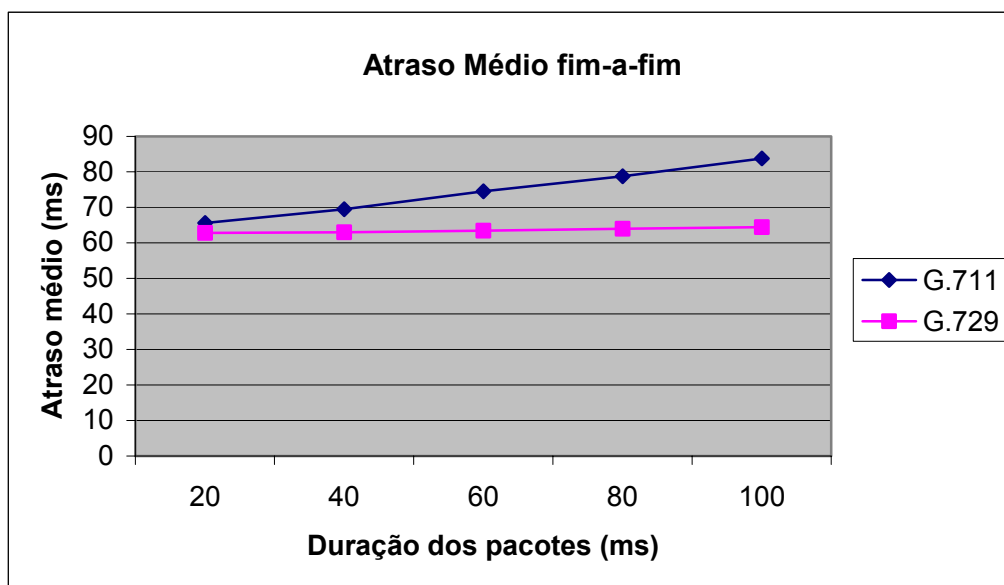


Figura 5. 3 – Média do atraso fim-a-fim dos pacotes de voz do primeiro cenário.

Pode-se observar na Figura 5.3 que a variação da quantidade de quadros por pacote causa um atraso mais acentuado nos pacotes das fontes do tipo G.711, uma vez que os pacotes da fonte de tráfego G.729 são bem menores em quantidade de bytes que os pacotes gerados pela fonte PCM.

Utilizando-se da fórmula da Equação (4.2) da seção 4.2.5.3 do capítulo 4 deste trabalho, foram gerados os valores da variação do atraso e calculadas as médias do *jitter*. A Figura 5.4 mostra os resultados obtidos nesses cálculos.

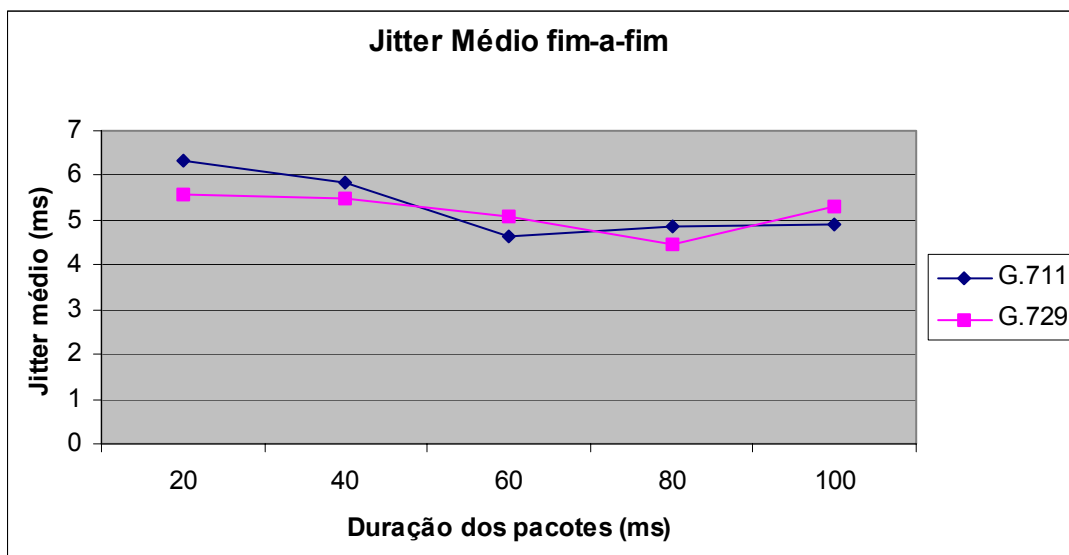


Figura 5. 4 – Média da variação do atraso do primeiro cenário.

O comportamento da variação do *jitter* ilustrado na Figura 5.4 mostra uma pequena redução da média de variação do atraso com o aumento do tamanho dos dados de voz nos pacotes. Apesar de, conforme mostrado nessa figura, os valores de *jitter* para os dois codificadores aumentam e diminuem no mesmo intervalo que a duração dos pacotes aumenta, mas indicando uma tendência de redução na média desses valores.

Para o percentual de pacotes com atraso superior a 300 ms no cálculo considerando os atrasos referentes ao tempo de codificação na origem e decodificação no destino, conforme descrito na seção 5.3 deste capítulo, não é apresentado neste cenário, pois não ocorreram pacotes com atrasos superiores a 300 ms no primeiro cenário simulado.

5.5.2 – Segundo Cenário – G.729 de 1 a 10 quadros por pacote

A taxa de transmissão de fluxo de tráfego agregado **AF** da fonte geradora VoIP, foi definida em 8 Kbps para as fontes G.729, com intervalos interpacotes conforme mostrado na Tabela 5.4, determinados pela duração dos pacotes.

Os PHB, as métricas e o perfil de descarte dos pacotes foram definidos como mostrado na Tabela 5.6. Os outros parâmetros necessários para a configuração do domínio *Diffserv* foram mostrados nas Tabelas 5.3 e 5.4, da seção 5.4 deste capítulo.

Tabela 5. 6 – Parâmetros utilizados no segundo cenário da simulação.

Classe do Tráfego	Tipo de Tráfego	Codificador	PHB	DSCP	Métricas	Descarte
AF	voz	G729	AF	10	TSW2CM, CIR 8Kbps	descarte fora do perfil
BE	ftp		BE	0	TSW2CM, CIR 2 Mb	descarte fora do perfil

Em síntese, a diferença entre os dois cenários deste trabalho é que apenas foi mantida como fonte de tráfego VoIP o codificador G.729, com variação do tamanho dos dados de 1 a 10 quadros por pacote. A topologia e as condições de QoS da rede simulada do segundo cenário permaneceram as mesmas do primeiro cenário.

A Figura 5.5 mostra a vazão média fim-a-fim obtida nos testes realizados no segundo cenário deste trabalho.

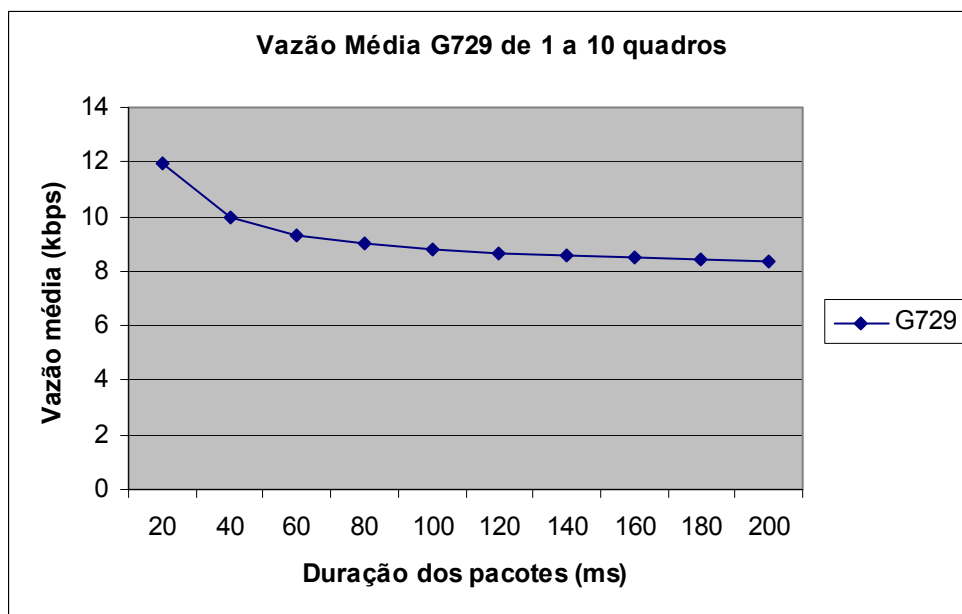


Figura 5. 5 – Vazão média obtida no segundo cenário.

Pode-se observar na Figura 5.5 que ocorreu uma redução na vazão média fim-a-fim com o aumento do tamanho dos dados de voz. Como não ocorreram perdas de pacotes de voz, isso representa a média da largura de faixa ocupada no meio em relação à quantidade de pacotes de voz transmitidos e efetivamente recebidos no destino.

O gráfico da Figura 5.5 confirma a tendência de queda da vazão em relação ao aumento do comprimento do pacote de voz observado na Figura 5.3 do primeiro cenário deste trabalho.

A Figura 5.6 mostra o gráfico da média do atraso fim-a-fim dos pacotes para o segundo cenário deste trabalho.

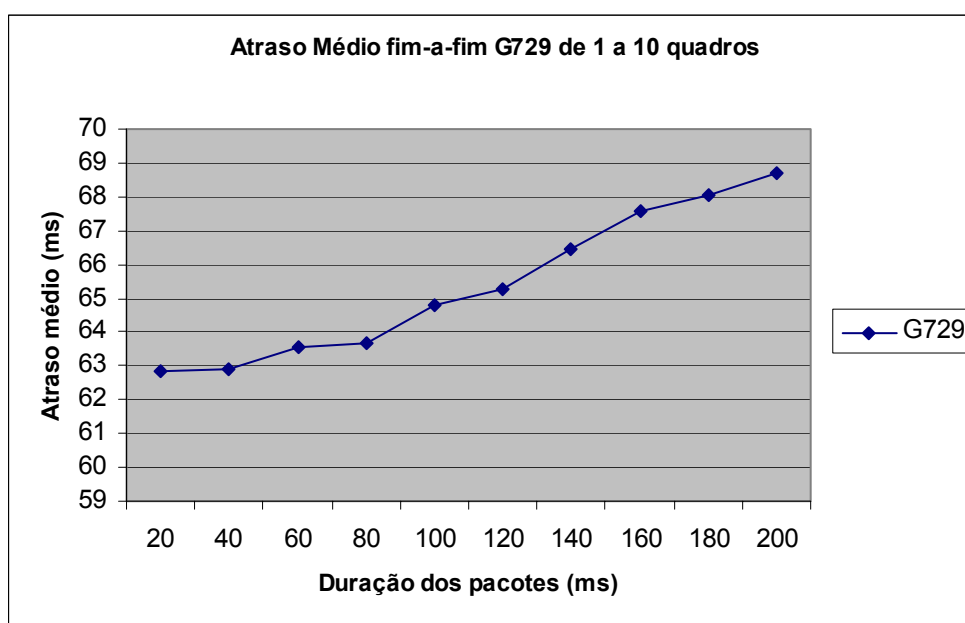


Figura 5. 6 – Média do atraso fim-a-fim dos pacotes de voz do segundo cenário.

Comparando os gráficos da Figura 5.3 da média do atraso fim-a-fim do primeiro cenário com a Figura 5.6 da média do atraso fim-a-fim do segundo cenário, pode-se verificar que a média permaneceu entre 60 e 70 ms para a rede simulada. Confirmando a tendência de aumento desse atraso, com o aumento do comprimento dos dados de voz por pacote IP.

O *jitter* médio do segundo cenário deste trabalho é mostrado na Figura 5.7.

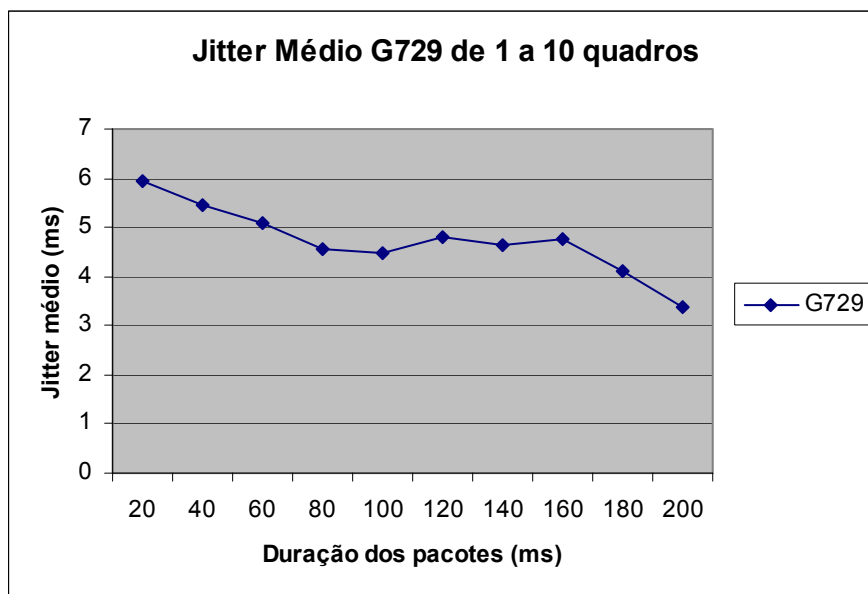


Figura 5. 7 – Média da variação do atraso do segundo cenário.

Comparando com os resultados do primeiro cenário da Figura 5.4, que apresenta apenas uma flutuação dos valores de *jitter* em um intervalo de 2 ms, com os resultados obtidos no segundo cenário mostrados no gráfico da Figura 5.7, pode-se concluir, para a topologia simulada neste trabalho, que o *jitter* reduz com o aumento do comprimento dos dados de voz por pacote IP.

Nos estudos realizados por Zheng e outros [10], foi demonstrado que quando aumenta a carga de tráfego, o tráfego *em rajadas* e o comprimento do estado **on/off**, o comportamento do *jitter* piora significativamente, reduzindo a qualidade de voz.

Pelos testes realizados neste trabalho, mais especificamente no segundo cenário, pode-se concluir que a variação do comprimento dos dados da voz provoca uma pequena redução no *jitter* de atraso referente às condições da rede simulada, melhorando a QoS.

A Figura 5.8 mostra o percentual de pacotes com latência fim-a-fim superior a 300 ms no cálculo em que foram considerados os atrasos referentes ao tempo de codificação na origem e decodificação no destino, duração dos pacotes, conforme Equação 5.1 da seção 5.3

deste capítulo. O gráfico da Figura 5.8 ilustra como a latência fim-a-fim influencia a qualidade de serviço, considerando a topologia simulada como uma situação real.

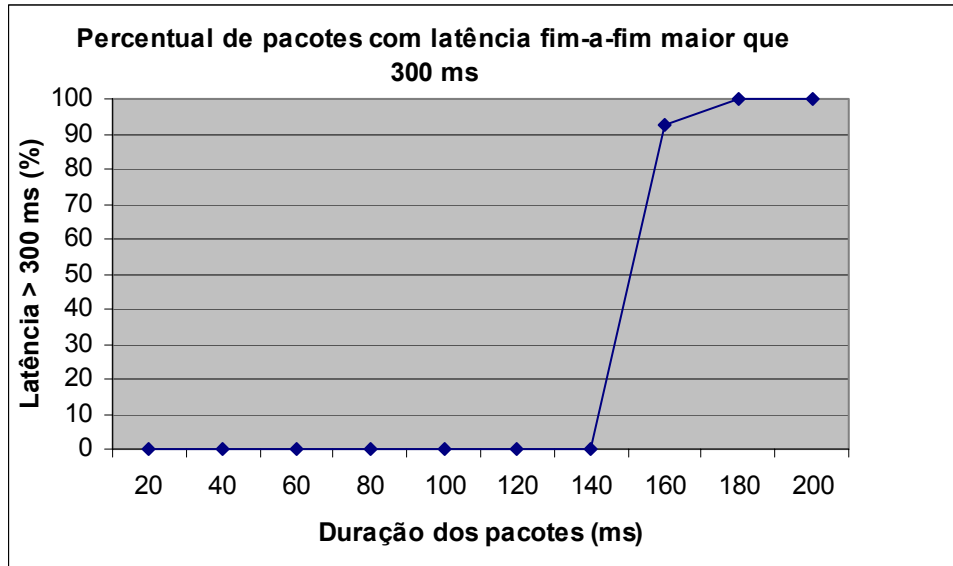


Figura 5. 8 – Percentual de pacotes com latência maior que 300 ms do segundo cenário.

Como esperado, não houve descarte dos pacotes de voz durante a simulação de nenhum dos cenários simulados neste trabalho, devido à implementação de QoS com suporte *Diffserv*, ocorreu apenas descarte dos pacotes de tráfego de fundo, por isso não é mostrado o gráfico comparativo das perdas de pacotes. O tráfego de fundo não foi suficiente para provocar descarte de pacotes de voz.

5.6 – Validação dos Resultados Obtidos

Com o objetivo de validar os resultados obtidos nas simulações foram realizados os cálculos do nível de confiança considerado de 95% em um total de 10 replicações dos *scripts* da simulação, conforme citado por Leon-Garcia [67], onde foram calculados a média, a variância e o intervalo de confiança das variáveis de saída na simulação.

As Tabelas 5.7, 5.8 e 5.9 mostram os intervalos de confiança da largura de faixa média fim-a-fim, do atraso médio fim-a-fim e do *jitter* médio fim-a-fim, respectivamente, para cada fonte geradora de tráfego de voz do primeiro cenário deste trabalho.

Tabela 5. 7 – Intervalo de confiança da vazão média fim-a-fim do primeiro cenário.

Intervalo de Confiança da vazão média fim-a-fim do primeiro cenário					
Codificador	Duração dos Pacotes (ms)	Média Amostral (ms)	Variância	Limite Inferior (ms)	Limite Superior (ms)
G.711	20	68.424	2.42532E-12	68.4239989	68.4240011
G.711	40	66.728	0	66.728	66.728
G.711	60	64.42	8.0844E-13	64.41999937	64.42000063
G.711	80	63.896	8.0844E-13	63.89599937	63.89600063
G.711	100	62.792	0	62.792	62.792
G.729	20	11.796	0	11.796	11.796
G.729	40	10.012	0	10.012	10.012
G.729	60	9.568	3.78956E-14	9.567999863	9.568000137
G.729	80	9.172	3.78956E-14	9.171999863	9.172000137
G.729	100	8.968	0	8.968	8.968

Tabela 5. 8 – Intervalo de confiança do atraso fim-a-fim do primeiro cenário.

Intervalo de Confiança do atraso fim-a-fim do primeiro cenário					
Codificador	Duração dos Pacotes (ms)	Média Amostral (ms)	Variância	Limite Inferior (ms)	Limite Superior (ms)
G.711	20	65.6087619	8.0844E-13	65.60876127	65.60876254
G.711	40	69.45473715	1.61688E-12	69.45473626	69.45473805
G.711	60	74.5338811	4.85064E-12	74.53387955	74.53388265
G.711	80	78.72487398	0	78.72487398	78.72487398
G.711	100	83.79207692	0	83.79207692	83.79207692
G.729	20	62.78960652	0	62.78960652	62.78960652
G.729	40	62.97323904	0	62.97323904	62.97323904
G.729	60	63.44709357	0	63.44709357	63.44709357
G.729	80	63.95536471	0	63.95536471	63.95536471
G.729	100	64.4593399	0	64.4593399	64.4593399

Tabela 5. 9 – Intervalo de confiança do jitter fim-a-fim do primeiro cenário.

Intervalo de Confiança do jitter fim-a-fim do primeiro cenário					
Codificador	Duração dos Pacotes (ms)	Média Amostral (ms)	Variância	Limite Inferior (ms)	Limite Superior (ms)
G.711	20	6.32795578	1.26319E-14	6.327955701	6.32795586
G.711	40	5.854810435	0	5.854810435	5.854810435
G.711	60	4.653999408	0	4.653999408	4.653999408
G.711	80	4.856077275	0	4.856077275	4.856077275
G.711	100	4.884125752	1.57898E-14	4.884125664	4.884125841
G.729	20	5.554139093	0	5.554139093	5.554139093
G.729	40	5.500088899	0	5.500088899	5.500088899
G.729	60	5.090367637	6.31594E-15	5.090367581	5.090367693
G.729	80	4.457126023	0	4.457126023	4.457126023
G.729	100	5.314064054	0	5.314064054	5.314064054

As Tabelas 5.10, 5.11 e 5.12 mostram o intervalo de confiança da largura de faixa média fim-a-fim, do atraso médio fim-a-fim e do *jitter* médio fim-a-fim, respectivamente, para cada fonte geradora de tráfego de voz do segundo cenário deste trabalho.

Tabela 5. 10 – Intervalo de confiança da vazão fim-a-fim do segundo cenário.

Intervalo de Confiança para a vazão média do segundo cenário					
Codificador	Duração dos Pacotes (ms)	Média Amostral (Kbps)	Variância	Limite Inferior (Kbps)	Limite Superior (Kbps)
G729	20	11,928	7,57912E-14	11,92799981	11,92800019
G729	40	9,972	0	9,972	9,972
G729	60	9,316	0	9,316	9,316
G729	80	8,992	0	8,992	8,992
G729	100	8,792	0	8,792	8,792
G729	120	8,676	0	8,676	8,676
G729	140	8,572	0	8,572	8,572
G729	160	8,492	0	8,492	8,492
G729	180	8,428	5,05275E-14	8,427999842	8,428000158
G729	200	8,392	0	8,392	8,392

Tabela 5. 11 – Intervalo de confiança do atraso fim-a-fim do segundo cenário.

Intervalo de Confiança para o atraso fim-a-fim do segundo cenário					
Codificador	Duração dos Pacotes (ms)	Média Amostral (ms)	Variância	Limite Inferior (ms)	Limite Superior (ms)
G729	20	62,85371173	8E-13	62,8537111	62,85371236
G729	40	62,92383232	0	62,92383232	62,92383232
G729	60	63,57213636	0	63,57213636	63,57213636
G729	80	63,65913061	0	63,65913061	63,65913061
G729	100	64,78933684	0	64,78933684	64,78933684
G729	120	65,24920468	0	65,24920468	65,24920468
G729	140	66,46942446	0	66,46942446	66,46942446
G729	160	67,5666664	8E-13	67,56666337	67,56666463
G729	180	68,07178704	8E-13	68,0717864	68,07178767
G729	200	68,68918947	2E-12	68,68918838	68,68919057

Tabela 5. 12 – Intervalo de confiança do jitter fim-a-fim do segundo cenário.

Intervalo de Confiança para o jitter fim-a-fim do segundo cenário					
Codificador	Duração dos Pacotes (ms)	Média Amostral (ms)	Variância	Limite Inferior (ms)	Limite Superior (ms)
G729	20	5,929054208	0	5,929054208	5,929054208
G729	40	5,443941213	1,26319E-14	5,443941134	5,443941292
G729	60	5,086364835	1,26319E-14	5,086364756	5,086364914
G729	80	4,566206621	3,15797E-15	4,566206581	4,566206661
G729	100	4,471506344	0	4,471506344	4,471506344
G729	120	4,819582937	1,26319E-14	4,819582858	4,819583016
G729	140	4,628285521	6,31594E-15	4,628285465	4,628285577
G729	160	4,759512094	0	4,759512094	4,759512094
G729	180	4,129433197	0	4,129433197	4,129433197
G729	200	3,396070234	0	3,396070234	3,396070234

5.7 – Conclusões deste Capítulo

Neste capítulo foi realizada uma breve apresentação do simulador NS, foram descritos os parâmetros e as variáveis da simulação da topologia utilizada nos testes, e os resultados obtidos foram apresentados e analisados.

Pode-se verificar pelos resultados obtidos nos testes realizados neste capítulo que o comprimento dos dados de voz possui grande influência sobre a qualidade de serviço através da métrica de atraso fim-a-fim. Isso pode ser mostrado pelas variações das métricas utilizadas. E, conseqüentemente, essas variações afetam a qualidade da voz.

O gráfico da Figura 5.2 do primeiro cenário e o gráfico da Figura 5.5 do segundo cenário mostram que conforme o codificador de voz utilizado, com o aumento do comprimento dos dados de voz por pacote IP ocorre uma melhoria na eficiência da transmissão com uma sensível redução na vazão média fim-a-fim.

O gráfico da Figura 5.3 do primeiro cenário e o gráfico da Figura 5.6 do segundo cenário mostram que quando aumenta o comprimento dos dados de voz por pacote IP, aumenta o atraso médio fim-a-fim. O gráfico da Figura 5.8 mostra, para as características do modelo simulado, uma redução da qualidade da voz de pobre à inaceitável para os pacotes com comprimento dos dados de voz a partir de 160 ms ou 8 quadros de voz por pacote IP, para o codificador G.729.

A variação do atraso mostrada no gráfico da Figura 5.4 do primeiro cenário, indica uma pequena flutuação da média do *jitter* fim-a-fim, com tendência de redução em relação à variação do tamanho dos dados de voz por pacote IP. Essa tendência é confirmada no segundo cenário, a variação do comprimento dos dados de voz por pacote provoca uma pequena redução na média do *jitter* fim-a-fim. Pode-se concluir que a variação do comprimento dos dados de voz por pacote provoca uma redução na variação do atraso em uma rede com suporte QoS.

Considerando o cálculo do atraso fim-a-fim por pacotes mostrado em percentual na Figura 5.8, nas condições da topologia simulada para o codificador G.729, pode-se verificar uma degradação da qualidade da voz a partir do comprimento dos dados de voz com duração 160 ms do pacote ou 8 quadros por pacote IP.

O cenário em estudo deste trabalho mostra que, com o objetivo de melhorar o aproveitamento da capacidade de transmissão do meio, pode-se transmitir o máximo de quadros de voz por pacotes, desde que não exceda os limites máximos do atraso e do *jitter* definidos pelas recomendações da ITU-T.

Não ocorrem perdas de pacotes da voz em nenhum dos cenários simulados, por esse motivo a perda de pacotes não foi analisada como métrica de QoS na redução da qualidade da voz em relação à variação do comprimento dos dados de voz por pacote IP.

A solução para o problema de redução da qualidade da voz em relação ao atraso fim-a-fim (e possivelmente de perdas de pacotes) e com aproveitamento da eficiência da transmissão da largura de faixa utilizada no meio, pode-se sugerir, pelos resultados obtidos neste trabalho, que a variação da quantidade de quadros por pacote seja feita de forma dinâmica na fonte de geração do tráfego.

CAPÍTULO VI

CONCLUSÕES, CONTRIBUIÇÕES DESTE TRABALHO E SUGESTÕES PARA FUTUROS TRABALHOS

6.1 – Considerações Finais

Foram descritos neste trabalho vários conceitos e informações relacionadas ao transporte de voz sobre redes IP, os padrões mais utilizados de VoIP e seus principais dispositivos. Foram apresentados a interoperação da sinalização de chamadas entre as redes de VoIP e a rede telefônica pública comutada, as técnicas de codificação, as características da qualidade da voz e os requisitos essenciais para implementação da qualidade de serviço em uma rede de VoIP. No contexto de qualidade de serviço foram descritos os principais fatores que influenciam a qualidade da voz e os mecanismos implementados de QoS em uma rede VoIP.

Neste trabalho, foi desenvolvida uma simulação para topologia proposta, com QoS, em dois cenários, utilizando o NS para avaliação das métricas de vazão, do atraso médio, da variação média do atraso fim-a-fim e do percentual de pacotes com atraso fim-a-fim superior a 300 ms por pacote considerando os *gateways* de voz.

Os valores dos resultados obtidos da vazão fim-a-fim para os dois cenários da topologia proposta apresentaram-se próximos dos cálculos mostrados nas Tabelas 4.4 e 4.5 da seção 4.2.5.1 do capítulo 4. Mostrando uma redução da largura de faixa utilizada no meio pela fonte de tráfego de VoIP.

Os valores dos resultados obtidos do atraso médio fim-a-fim, para os dois cenários da topologia simulada neste trabalho, mostraram uma tendência de aumento desse atraso quando aumenta o comprimento dos dados de voz por pacote. O aumento desse atraso indica uma redução da qualidade da voz.

Nas variações da média do atraso mostradas no primeiro cenário, os valores comportaram-se bastante alternados em relação ao aumento do comprimento dos dados de voz durante a simulação. Porém, já para o segundo cenário, a variação do comprimento dos dados de voz por pacote mostra uma pequena redução na média do *jitter* fim-a-fim. Pode-se concluir que a variação do comprimento dos dados de voz por pacote indica uma tendência de redução na variação do atraso em uma rede com suporte QoS.

A técnica de QoS utilizada neste trabalho, o *Diffserv*, mostrou-se eficaz na implementação de uma rede VoIP, para os cenários da topologia simulada, mostrando que em uma rede com gerenciamento de QoS definido, não haverá perdas de pacotes de voz, porque a largura de faixa está garantida pelas políticas de QoS implementadas.

6.2 – Contribuições deste Trabalho

A contribuição deste trabalho foi a análise do impacto do comprimento dos dados de voz em uma rede de VoIP prática com QoS.

6.3 – Sugestões para Futuros Trabalhos

Vários estudos podem ser desenvolvidos a partir deste trabalho e vários acréscimos nos cenários analisados também podem ser realizados para estudo.

Como trabalhos futuros, alguns temas podem ser sugeridos para desenvolvimento a partir desta dissertação, tais como:

- Introdução de amostras reais de voz dentro dos padrões definidos pelas recomendações da série P.800 da ITU-T [68] e análise da qualidade da voz utilizando-se métodos de qualificação da voz não-subjetivos;
- Com a introdução de amostras reais de voz dentro dos padrões definidos pelas recomendações da ITU-T [68] pode-se realizar um estudo do comportamento do *jitter* em relação ao *bursty* de tráfego na rede;
- Análise da geração do *look-ahead* em relação ao comprimento do pacote, sendo que ele é gerado a cada quadro de amostras; e;
- Utilização de outras variáveis de QoS para o ajuste dinâmico do comprimento dos dados de voz no pacote IP e outras políticas e contratos de níveis de serviço.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Soares, L. C., Freire, V. A., “Redes Convergentes”, 1ªEdição, Editora Alta Books, 2002.
- [2] Neto, V. S., Carvalho, F. T. A., “Tecnologia de Centrais Telefônicas”, 2ªEdição, Editora Érica, 1999.
- [3] Collins, D., “Carrier Grade Voice Over IP”, 2ªEdição, Editora McGraw Hill, 2003.
- [4] Teleco, “Prestadores de Serviço de VoIP via Internet no Brasil”, <http://www.teleco.com.br/voip.asp>, Maio, 2005.
- [5] Minoli, D., Minoli, E., “Delivering Voice over IP Networks”, 1ªEdição, Editora John Wiley & Sons Inc,1998.
- [6] Hersent O., Guide, D., Petit, Jean-Pierre, “IP Telephony”, 2ªEdição, Editora Addison Wesley, 2000.
- [7] ITU-T H.248, “Gateway Control Protocol”, Recommendation H.248, disponível em: <http://www.itu.int/ITU-T/publications/recs.html>, acesso em: Junho, 2004
- [8] Performance Technologies, “SS7 Tutorial”, disponível em: <http://www.pt.com/tutorials/ss7/>, acesso em: Dezembro, 2003.
- [9] Oouchi, H., Takenaga, T., Sugawara, H. and Masugi Masao, “Study on Appropriate Voice Data Length of IP Packets for VoIP Network Adjustment”, GLOBECOM 2002 – IEEE, Global Telecommunications Conference, Novembro, 2002
- [10] Zheng, L., Zhang, L., Sugawara, H. and Xu, D., “Characteristics of Network Delay and Delay Jitter and its Effect on Voice over IP (VoIP)”, GLOBECOM 2001 – IEEE Network, Vol.1 pags 122-126, Global Telecommunications Conference, Junho, 2001

- [11] Sun, L.F., Wade G., Lines B.M., Ifeachor E.C., “Impact of Packet Loss Location on Perceived Speech Quality”, disponível em http://www.iptel.org/2001/pg/final_program/15.pdf, Abril, 2001, acesso em: Abril, 2004.
- [12] Melo, E.T.L., Westphall, C., “Qualidade de Serviço em Redes IP com Diffserv: Avaliação através de Medições”, Maio, 2001.
- [13] Fernandes, N.L.L., Moraes, L.F.M., “Relação Entre a Qualidade de Respostas das Recomendações G.723.1 e G.729, e o Comportamento da Rede IP de Suporte”, Dissertação de Mestrado, UFRJ, Março, 2003.
- [14] Neto, J.P.M., Kelner, J., “Um Mecanismo de Provisionamento de Tráfego para Serviços Diferenciados”, Dissertação de Mestrado, UFPE, Fevereiro, 2003.
- [15] Ahmed, Fatima, “A Comparison of Voice Technologies (VoIP, VoFR, and VoATM)”, disponível em: <http://www.developer.com/voice/article.php/3112781>, acesso em: Dezembro, 2003
- [16] Cherukuri, R.; Walsh, T.; “Voice over MPLS”, disponível em: <http://www.mplsforum.org/>, acesso em: Novembro, 2004.
- [17] ITU-T G.711, “Pulse Code Modulation (PCM) of Voice Frequencies”, Recommendation G.711, disponível em: <http://www.itu.int/ITU-T/publications/recs.html>, Novembro, 1982, acesso em: Janeiro 2004.
- [18] ITU-T G.723.1, “Speech Coders : Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 and 6.3 kbit/s”, Recommendation G.723.1, disponível em: <http://www.itu.int/ITU-T/publications/recs.html>, Março, 1996, acesso em: Janeiro, 2004
- [19] ITU-T G.729, “Coding of Speech at 8 kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear-prediction (CS-ACELP)”, Recommendation G.729, disponível

- em: <http://www.itu.int/ITU-T/publications/recs.html>, Março,1996, acesso em: Março, 2004.
- [20] Schulzrinne, H., et al., “RTP: A Transport Protocol for Real-Time Applications”, Internet RFC3550, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Julho, 2004.
- [21] Schulzrinne, H., et al., “RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals”, Internet RFC2833, disponível em: <http://rfc.net/rfc-index.html>, Maio, 2000, acesso em: Janeiro 2005..
- [22] ITU-T H.323, “Coding of Speech at 8 kbit/s using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP)”, Recommendation H.323 Version 2, disponível em <http://www.itu.int/ITU-T/publications/recs.html>, Março,1996, acesso em: Abril, 2004.
- [23] ITU-T H.323, “Coding of Speech at 8 kbit/s using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP)” (This version 4 of H.323 integrates without further Modifications Annex G (17/02/2000), and Includes Annexes J, K, M1 et M2 approved on 17 November 2000, and Annex L that was Approved Independently on 1 March 2001), Recommendation H.323 Version 4, disponível em <http://www.itu.int/ITU-T/publications /recs.html>, Novembro, 2000, acesso em: Março, 2004.
- [24] Handley M., et al., “SIP: Session Initiation Protocol”, Internet RFC2543, Março, 1999.
- [25] ITU-T Q.931, “ISDN User-Network Interface Layer 3 Specification for Basic Call Control”, Recommendation Q.931, disponível em <http://www.itu.int/ITU-T/publications/ recs.html>, Maio, 1998, acesso em: Março, 2004.

- [26] ITU-T H.225.0, “Call Signaling Protocols and Media Stream Packetization for Packet-based Multimedia Communication Systems”, Recommendation H.225.0 Version 4, disponível em <http://www.itu.int/ITU-T/publications/recs.html>, Novembro, 2000, acesso em: Abril, 2004.
- [27] ITU-T H.245, “Control Protocol for Multimedia Communication”, Recommendation H.245 Version 8, disponível em <http://www.itu.int/ITU-T/publications/recs.html>, Julho, 2001, acesso em: Maio, 2004.
- [28] ITU-T H.225.0, “Call Signalling Protocols and Media Stream Packetization for Packet-based Multimedia Communication Ssystems”, Recommendation H.225.0, disponível em <http://www.itu.int/ITU-T/publications/recs.html>, Julho, 2003, acesso em: Junho, 2004
- [29] Douskalis, B., “IP Telephony”, 1ª Edição, Editora Hewlett-Packard, 2000.
- [30] ITU-T H.323, “Coding of Speech at 8 kbit/s using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP) ” (This version 5 of H.323 Integrates Without further Modifications Annexes M3 (07/2001), P (01/2003), Q (07/2001) and R (07/2001) that were published separately, and Annex O that was approved independently on 07/2003), Recommendation H.323 Version 4, disponível em: <http://www.itu.int/ITU-T/publications/recs.html>, Julho, 2003, acesso em: Julho, 2004.
- [31] Huitema C., et al., “Real Time Control Protocol (RTCP) attribute in Session Description Protocol (SDP)”, Internet RFC3605, disponível em: <http://rfc.net/rfc-index.html>, Outubro, 2003, acesso em: Agosto 2004.
- [32] SS8 Networks, “Simply SS7”, disponível em: <http://www.ss7.com/>, Dezembro, 2002, acesso em: Junho, 2004.

- [33] Agrawal, H. et al., “Session Initiation Protocol (SIP)-H.323 Interworking Requirements”, disponível em: <http://ietfreport.isoc.org/all-ids/draft-agrawal-sip-h323-interworking-reqs-07.txt>), acesso em: Outubro, 2004.
- [34] Greene, N. et al., “Media Gateway Control Protocol Architecture and Requirements”, Internet RFC2805, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Abril, 2004.
- [35] Greene, N. et al., “Media Gateway Control Protocol Architecture and Requirements”, Internet RFC2705, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Abril, 2004.
- [36] Andreassen, F. et al., “Media Gateway Control Protocol (MGCP)”, Internet RFC3435, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Janeiro, 2004.
- [37] Foster, B. et al., “Basic Media Gateway Control Protocol (MGCP) Packages”, Internet RFC3660, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Dezembro, 2003.
- [38] Cuervo, F. et al., “Megaco Protocol version 0.8”, Internet RFC2885, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Agosto, 2004.
- [39] Cuervo, F. et al., “Megaco Protocol Version 1.0”, Internet RFC3015, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Novembro, 2004.
- [40] Groves, C., Pantaleo, M. et al., “The Megaco/H.248 Gateway Control Protocol, version 2”, Internet Draft, disponível em: <http://www.ietf.org/internet-drafts/draft-ietf-megaco-h248v2-04.txt>, acesso em: Abril, 2004.
- [41] ITU-T G.723.1 ANNEX A, “Speech coders : Silence Compression Scheme”, Recommendation G.723.1 ANNEX A, disponível em: <http://www.itu.int/ITU-T/publications/recs.html>, Novembro, 1996, acesso em: Abril, 2004.

- [42] ITU-T G.729 ANNEX A, “Reduced Complexity 8 kbit/s CS-ACELP Speech Codec”, Recommendation G.729 ANNEX A, disponível em: <http://www.itu.int/ITU-T/publications/recs.html>, acesso em: Novembro,1996, acesso, Abril, 2004.
- [43] ITU-T G.729 Annex B, “A silence Compression Scheme for G.729 Optimized for Terminals Conforming to Recommendation V.70”, Recommendation G.729 ANNEX B, disponível em: <http://www.itu.int/ITU-T/publications/recs.html>, Outubro,1996, acesso em: Abril, 2004.
- [44] Cavalcante, J. at all, “Voice and Data Integration”, Cisco do Brasil, disponível em: <ftp://ftp.registro.br/pub/gter/gter09/>, Outubro, 1999, acesso em: Setembro, 2004.
- [45] ITU-T P.800.1, “Mean Opinion Score (MOS) Terminology”, Recommendation P.800.1, disponível em: <http://www.itu.int/ITU-T/publications/recs.html>, acesso em: Maio, 2004.
- [46] ITU-T P.830, “Subjective Performance Assessment of Ttelephone-Band and Wideband Digital Codecs”, Recommendation P.830, disponível em: <http://www.itu.int/ITU-T/publications/recs.html>, Fevereiro, 1996, acesso em: Maio, 2004.
- [47] ITU-T P.861, “Objective Quality Measurement of Telephone-Band (300-3400Hz) Speech Codecs”, Recommendation P.861, disponível em: <http://www.itu.int/ITU-T/publications/recs.html>, acesso em: Março, 2004.
- [48] Anderson, J., Agilent Technologies, “Methods for Measuring Perceptual Speech Quality, White Paper”, disponível em: <http://literature.agilent.com/litweb/pdf/5988-2352EN.pdf>, acesso em: Maio, 2004.
- [49] ITU-T P.862, “Perceptual Evaluation of Speech Quality (PESQ), an Objective Method for end-to-end Speech Quality Assessment of Narrowband Telephone Networks and

- Speech Codecs”, Recommendation P.862, disponível em: <http://www.itu.int/ITU-T/publications/recs.html>, acesso em: Novembro, 2004.
- [50] Voice Over IP Calculator, disponível em: <http://www.voip-calculator.com/bandwidth.html>, acesso em: Janeiro, 2005.
- [51] Casner, S., et al., “Compressing IP/UDP/RTP Headers for Low-Speed Serial Links”, Internet RFC2508, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Fevereiro, 2004.
- [52] ITU-T G.114, “One-way Transmission Time”, Recommendation G.114, disponível em: <http://www.itu.int/ITU-T/publications/recs.html>, acesso em: Maio, 2004.
- [53] Blake, S., et al., “An Architecture for Differentiated Services”, Internet RFC2475, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Setembro, 2004.
- [54] Braden, R., et al., “Integrated Services in the Internet Architecture: an Overview”, Internet RFC1633, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Junho, 2004.
- [55] Oyama, C. S. O., Lucena, S. C., News Generation, “Considerações Acerca do Estabelecimento de QoS no RNP2”, disponível em: http://www.rnp.br/newsgen/0205/qos_rnp.html#ng-2-1, acesso em: Maio, 2004.
- [56] Braden, R., et al., “Resource ReSerVation Protocol (RSVP)”, Internet RFC2205, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Maio, 2004.
- [57] Grossman, D., et al., “New Terminology and Clarifications for Diffserv”, Internet RFC3260, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Abril, 2004.
- [58] Heinanen, J., et al., “Assured Forwarding PHB Group”, Internet RFC2597, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Junho, 2004.
- [59] Jacobson, V., et al., “An Expedited Forwarding PHB”, Internet RFC2598, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Junho, 2004.

- [60] Davie B., et al., “An Expedited Forwarding PHB”, Internet RFC3246, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Março, 2004.
- [61] Rosen, E., et al., “Multiprotocol Label Switching Architecture”, Internet RFC3031, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Janeiro, 2004.
- [62] Durham, Ed. D., et al., “The COPS (Common Open Policy Service) Protocol”, Internet RFC2748, disponível em: <http://rfc.net/rfc-index.html>, acesso em: Janeiro, 2004.
- [63] Network Simulator (Version 2.27), disponível em: <http://www.isi.edu/nsnam/ns/>, acesso em: Janeiro, 2004. Contributed code: Diffserv Model, Nortel Networks, <http://www7.nortel.com:8080/ctl/>.
- [64] Kostas, T. J., et al., “Real-Time Voice Over Packet-Switched Networks”, IEEE Network, disponível em: <http://www.eecs.harvard.edu/cs143/protected/papers/VoIP-Kostas98.pdf>, acesso em: Março, 2004.
- [65] Coutinho, M. M., “Network Simulator”, disponível em: [http://www.cci.unama.br/margalho/network simulator/](http://www.cci.unama.br/margalho/network%20simulator/), acesso em: Outubro, 2004.
- [66] Chuah, C., “Workload Model for Packet Audio Traffic”, disponível em: [http://www.ece.ucdavis.edu/~chuah/research/ voip/](http://www.ece.ucdavis.edu/~chuah/research/voip/), acesso em: Janeiro, 2005
- [67] Leon-Garcia, A., “Probability Random-Process for Electrical Engineering”, 2ªEdição, Editora Addison Wesley, Publishing Company Inc, 1994.
- [68] ITU-T P.800, “Methods for Subjective Determination of Transmission Quality”, Recommendation P.800, disponível em: <http://www.itu.int/ITU-T/publications/recs.html>, acesso em: Agosto, 2004.