

## Voz sobre IP (VoIP): um estudo experimental

**Claudio Luis Sitolino** ([sitolino@apec.unoeste.br](mailto:sitolino@apec.unoeste.br))

Universidade do Oeste Paulista - UNOESTE  
Faculdade de Informática de Presidente Prudente – FIPP

**Juergen Rochol** ([juergen@inf.ufrgs.br](mailto:juergen@inf.ufrgs.br))

Universidade Federal do Rio Grande do Sul - UFRGS  
Instituto de Informática - Departamento de Informática Aplicada  
Av. Bento Gonçalves, 9500 – Campus do Vale – 91501-970 - Porto Alegre RS

### Resumo

Dentro dos serviços de multimídia, o serviço de voz é sempre o primeiro a ser considerado quando se trata de integração de serviços em uma rede. Neste trabalho apresenta-se um estudo experimental sobre o desempenho de um serviço de voz sobre IP (VoIP) quando aplicado a diversas topologias de rede e com diferentes perfis de cargas. Desenvolveu-se uma ferramenta de análise que permite realizar medidas de desempenho do serviço em relação a parâmetros de qualidade de serviço (QoS), como atraso, variação do atraso (*jitter*) e perda de pacotes. Apresentam-se e discutem-se os resultados experimentais obtidos a partir desta ferramenta - tanto para tráfego de voz em um domínio de rede local, como para tráfego em um domínio de rede remoto -, aplicando-se diversas situações de carga nestes domínios. Conclui-se com algumas diretrizes que devem orientar projetos de redes que prevêm integração de serviços de voz e dados.

**Palavras-Chave:** Voz sobre IP, VoIP, avaliação de desempenho, QoS.

### Abstract

*Within multimedia services, voice service is the first to be considered when (we deal) it is about integration of multimedia services in a network. In this work we present an experimental study of the performance of voice over IP (VoIP), when applied to different network topologies with different load profiles (conditions). We develop a tool for performance analysis, based on measurements of performance parameters (or quality of service (QoS) parameters) like: delay, jitter delay and packet loss. We present and discuss the experimental results obtained with the tool, as for voice traffic in a local network domain, as for a remote domain, considering different traffic loads in the domains. We conclude with some guide lines that will be considered in network designs, with integration of voice and data traffic.*

**Keywords:** Voice over IP, VoIP, performance evaluation, QoS.

## 1 Introdução

Voz sobre IP, ou apenas VoIP, é uma tecnologia que permite a digitalização e codificação da voz e o empacotamento destes dados em pacotes IP para transmissão em uma rede que utilize o protocolo TCP/IP. O transporte de dados VoIP, utilizando-se como suporte a rede Internet comercial, tem sido um forte atrativo para os usuários. A experiência, no entanto, demonstrou que os pacotes IP contendo os dados de voz, ao passarem por diversos domínios e roteadores, não tem mais condições de oferecer uma qualidade de voz aceitável no destino. Um dos motivos é que os parâmetros de QoS (*quality of service*) exigidos para este

serviço, relativos ao atraso e à variação deste atraso, não podem ser assegurados pela Internet comercial. O volume de dados gerado por uma aplicação VoIP é outro desafio para a rede TCP/IP, fazendo com que a sua aplicação, muitas vezes, se restrinja a redes corporativas privadas, nas quais é relativamente simples e pouco oneroso a disponibilização de amplos recursos em termos de banda passante.

O paradigma do *melhor esforço* (*best effort*) por parte da rede Internet em relação aos pacotes de dados não é suficiente para uma aplicação de voz, já que esta aplicação é sensível ao atraso. Em uma rede TCP/IP não é possível garantir um atraso constante ou, no mínimo, com uma variação deste atraso dentro de limites definidos, o que torna uma aplicação de voz em tempo real (como por exemplo uma ligação telefônica), inviável ou, como costuma ser, um serviço de baixíssima qualidade - com a voz entrecortada e pouco inteligível.

Neste artigo, apresentam-se os resultados de um estudo experimental do serviço VoIP em relação a diferentes topologias de redes, com diversos perfís de carga, funcionando segundo o paradigma de *melhor esforço* (*best effort*). Para simular um ambiente real, utilizou-se um arranjo experimental envolvendo duas redes Ethernet com vários microcomputadores, interligados por um canal de 33.600 bit/s. Para emular o serviço de voz, foi utilizado o software, Microsoft NetMeeting 2.1, e para a obtenção dos resultados experimentais, desenvolveu-se uma ferramenta de análise específica. A ferramenta consiste de dois aplicativos, sendo que um foi desenvolvido utilizando-se a linguagem C em ambiente Linux, e é responsável pela captura dos pacotes da rede e gerar um *arquivo-log* de texto que contém os pacotes capturados. O segundo módulo do aplicativo, desenvolvido em linguagem Delphi para Windows 95, é responsável pelo processamento do *arquivo-log* de texto que contém os pacotes e assim gerar dados estatísticos e gráficos que permitam a avaliação do desempenho do serviço de voz.

O restante do artigo está organizado como segue. na seção 2, são revisadas algumas características do protocolo TCP/IP em relação a serviços de tempo real e as exigências de QoS mínimas necessárias para um serviço de voz com qualidade mínima. São discutidos também os principais problemas que surgem quando se tenta a implementação deste serviço em uma rede TCP/IP. Na seção 3, apresentam-se e discutem-se tanto a metodologia experimental adotada, como o ambiente que foi definido para avaliar o desempenho do serviço de voz. Na seção 4, são apresentados e analisados os resultados experimentais obtidos. Finalmente, na seção 5, são apresentadas algumas conclusões além de algumas sugestões para trabalhos futuros nesta área.

## 2 Serviço de Voz sobre o Protocolo TCP/IP

TCP/IP é uma tecnologia de rede simples popularizada entre computadores com o sistema operacional Unix e a Internet. Atualmente o TCP/IP está presente na maioria dos sistemas operacionais de rede e é utilizado por muitas empresas em uma variedade de aplicações. TCP/IP é, especificamente, um protocolo de comunicação de dados projetado para aplicações não sensíveis ao atraso, tais como *e-mail*, *web*, *ftp*, etc.

Devido a imprevisibilidade do atraso dos pacotes, os protocolos da camada de transporte do TCP/IP (tanto o TCP como o UDP), não são adequados para aplicações de voz em tempo real [KAT 99]. O TCP (*Transport Control Protocol*) não suporta transmissão de voz em tempo real porque utiliza um mecanismo de recuperação dos dados perdidos por retransmissão; assim, no caso da perda de um pacote, a liberação dos dados para a aplicação deve esperar por todas as retransmissões, o que acarretaria atrasos intoleráveis. Já o protocolo

UDP (*User Datagram Protocol*) evita este problema, pois fornece um serviço de datagrama, no entanto, tem como grande desvantagem o fato de não ser confiável [LIU 99].

Voz sobre IP é um conceito relativamente simples. Basta transformar a voz em um fluxo de bits que pode ser constante (CBR) ou variável (VBR), dependendo do *codec* (codificador/decodificador). O fluxo de bits assim obtido é encapsulado em datagramas do UDP, que por sua vez são encapsulados em pacotes IP como é mostrado na figura 1 [JAI 98]. Os pacotes IP assim obtidos são transportados pela rede, como qualquer pacote de dados IP, sem qualquer distinção.

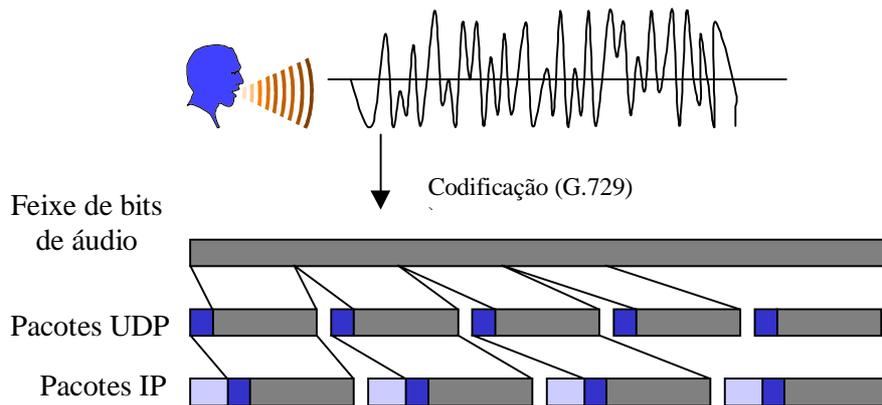


Figura 1 – Encapsulamento do fluxo de dados de voz em pacotes IP

Foi esta relativa simplicidade que estimulou fortemente muitos usuário a implementar serviços de voz e dados dentro de seus ambientes de rede. Aplicações de voz em tempo real, baseadas no protocolo IP, já são uma realidade em muitas empresas, desde que nestas redes seja possível exercer um controle sobre os seus roteadores, sob forma de um tratamento privilegiado dos pacotes de voz IP o que, porém, não é possível quando se trata da Internet comercial [CAL 99].

A escolha do codificador de voz (*codec*) é fundamental para o sucesso da aplicação VoIP [COX 96]. Os *codecs* modernos são inteligentes e executam funções sobre o fluxo de dados de voz como: compressão, supressão de silêncio e cancelamento de eco, o que faz com que o fluxo na saída do *codec* seja normalmente do tipo VBR. O fluxo na saída de um *codec* de voz se apresenta sob forma de conjuntos de pacotes de tamanho fixo, emitidos segundo uma cadência constante, o que caracteriza um serviço do tipo ON/OFF em rajada. Um conjunto de pacotes (rajada) é emitido sequencialmente segundo uma taxa máxima, o que determina o período de ON do serviço. Uma vez emitido um conjunto de pacotes de tamanho fixo, haverá um período de inatividade até que seja emitido o próximo conjunto de pacotes - o que caracteriza o período de OFF do serviço (conferir figura 5). Um exemplo de um codificador com estas características é o codificador G.729 do ITU-T, que apresenta uma qualidade de voz boa, com uma pontuação MOS<sup>1</sup> de 3.92 [MIN 98] e uma taxa de bit média baixa de apenas 8 kbit/s [ITU 96].

O transporte dos pacotes de voz deve ser realizado pela rede segundo exigências específicas deste serviço, também conhecidas como parâmetros de QoS da aplicação [ROC 98]. As redes IP não oferecem nenhum tipo de QoS as suas aplicações, o que é considerado atualmente o principal empecilho para as aplicações multimídia.

<sup>1</sup> MOS (Mean Opinion Score) é uma medida de qualidade da voz amplamente aceita em telefonia que varia entre 1 (ruim) a 5 (excelente).

O padrão dominante para transmitir serviços multimídia em redes de pacotes como IP é a Recomendação H.323 do ITU-T [THO 96] [ITU 96a], que sugere, para o encapsulamento do serviço de voz, a utilização da pilha de protocolos IP/UDP/RTP. No entanto, o H.323 não fornece qualquer garantia de QoS; especifica apenas, como um protocolo de transporte confiável como o TCP, pode ser utilizado para transmissão de informações de controle destes serviços [KOS 98].

O protocolo RTP (*Real Time Protocol*) [SCH 96] é um mecanismo genérico para suportar a integração de voz, vídeo e dados e foi definido pela RFC 1889 [SCH 96]. Sua função principal é agir como uma interface melhorada entre as aplicações de tempo real e os protocolos das camadas já existentes [LIU 99]. De maneira geral, o RTP tenta, mas não garante o fornecimento de pacotes no tempo desejado ou de acordo com uma qualidade de serviço especificada. O cabeçalho do RTP fornece o número de seqüência e a informação de tempo através de uma etiqueta de tempo (*time-stamp*) que é necessária para a remontagem do fluxo de tempo real, a partir dos pacotes IP recebidos. O protocolo, no entanto, não assegura que os pacotes sejam entregues em ordem, nem tão pouco oferece uma garantia de tempo de entrega dos pacotes - o que o torna pouco confiável.

No âmbito do IETF, atualmente estão sendo discutidas cinco abordagens para o oferecimento real de QoS pela Internet [XIA 99]: a arquitetura de serviços integrados (IntServ) [BRA 94], que utiliza o protocolo RSVP [BRA 97] para reserva de recursos; a Arquitetura de Serviços Diferenciados (DiffServ) [BLA 98]; o MPLS (Multiprotocol Label Switching) [ROS 99] [CUN 99]; o Roteamento com QoS [CRA 98] e a Engenharia de Tráfego [KAM 2000]. Destas abordagens, a arquitetura de serviços integrados com RSVP e a arquitetura de serviços diferenciados (DiffServ) são as que oferecem as melhores perspectivas quando se quer implementar serviços multimídia com QoS efetivo em redes corporativas com abrangência limitada.

O protocolo RSVP funciona no topo do protocolo IP, na camada de transporte. É um protocolo de controle e sinalização, comparável com o ICMP (*Internet Control Message Protocol*) ou IGMP (*Internet Gateway Message Protocol*) [MIN 98] [ZAP 98]. O RSVP foi projetado para funcionar com os protocolos de roteamento considerando um único receptor ou um grupo de receptores. Algumas aplicações são direcionadas para apenas um receptor (unicast) enquanto que outras podem enviar dados a mais de um receptor (multicast), sem ter que especificar cada destinatário para a rede [NOR 99]. Com o RSVP, a aplicação é capaz de notificar antecipadamente quais os recursos da rede que serão necessários para atender às exigências de QoS da aplicação (sinalização). Os roteadores envolvidos fazem uma reserva e se comprometem a oferecer estes recursos de forma garantida. Se o roteador não é capaz de disponibilizar de forma garantida estes recursos a aplicação pode não ser implementada. Neste caso a origem é imediatamente notificada, evitando-se assim desperdício de tempo e recursos numa tentativa sem chance de sucesso [NOR 99].

Os componentes do RSVP são o transmissor, o receptor e os roteadores intermediários que intervêm na rota entre a origem e o destino dos pacotes. Seu funcionamento pode ser resumido como segue: o transmissor origina os dados, especifica o destino dos mesmos e define para a rede as exigências de QoS em relação ao transporte destes dados. O receptor, por sua vez, envia notificações para os roteadores envolvidos para que aloquem recursos em termos de banda e *buffers* e assim permitam repassar os pacotes segundo as exigências de QoS definidas. Somente após completados estes passos, em ambos os sentidos (serviço *duplex*), são emitidos os pacotes contendo os dados de voz [LIU 99].

A abordagem segundo a arquitetura de DiffServ é mais recente que a abordagem do IntServ. O objetivo principal do DiffServ é o oferecimento de uma arquitetura mais escalável e mais gerenciável para obter diferenciação de serviços em redes IP [DOV 99]. O DiffServ

parte do pressuposto de que este objetivo pode ser alcançado, centrando o foco não em fluxos de pacotes individuais, mas em fluxos de tráfego agregados - formados por um conjunto de aplicações com requisitos de serviço semelhantes. A abordagem segundo o DiffServ ainda está em fase inicial do seu desenvolvimento, e muitos aspectos operacionais ainda não foram bem esclarecidos [BER 99].

### 3 Metodologia Experimental

O nosso estudo experimental consistiu na definição de um ambiente de teste para avaliar o tráfego de voz sobre uma rede TCP/IP, além do desenvolvimento de uma ferramenta específica para realizar a análise do tráfego de voz neste ambiente. Na figura 3.1 apresenta-se a topologia do ambiente experimental, que é formado basicamente de dois domínios de redes locais Ethernet (um domínio local e outro remoto), ambos estruturados a partir de um *hub* ligado a porta de um roteador. Estes domínios são interconectados através de dois roteadores, via um enlace de modems com uma taxa de transmissão fixa de 33,6Kbit/s.

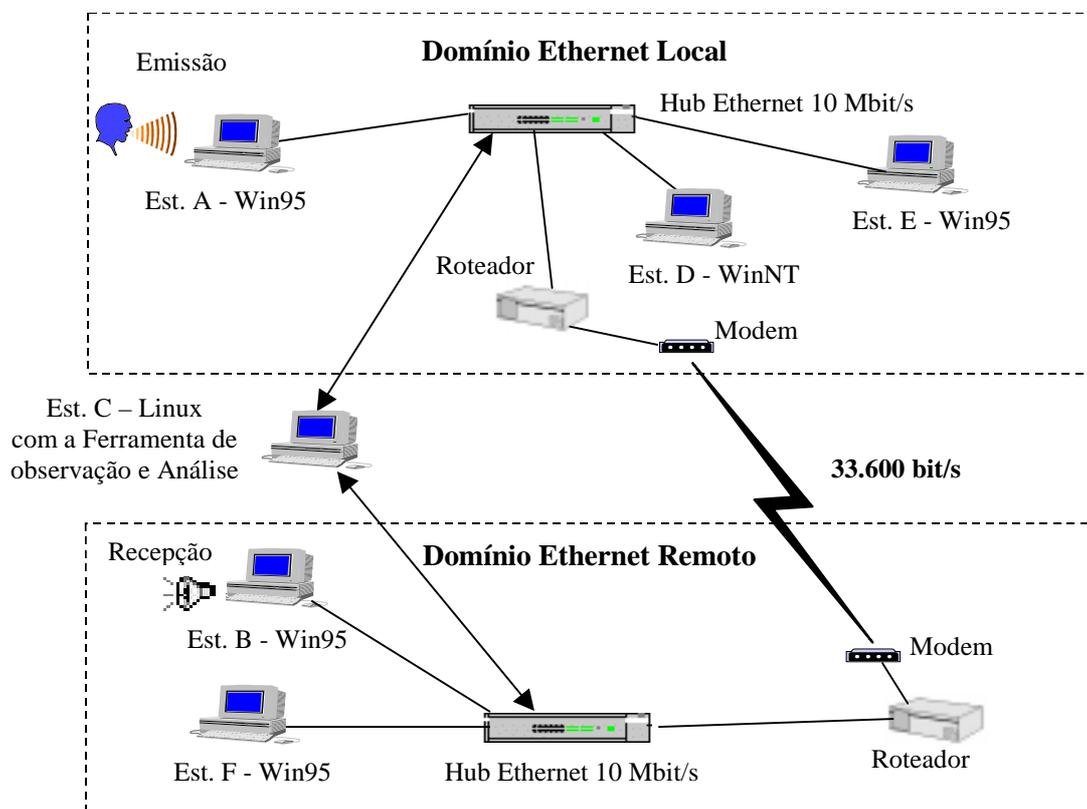


Figura 2 – Arranjo Experimental

Neste ambiente experimental foram definidos dois cenários experimentais: (1.) um cenário formado por um único domínio de rede local Ethernet, contendo tanto a origem e o destino do serviço de voz neste domínio, e (2.) um ambiente de rede formado por dois domínios de Rede Local, um local contendo a origem do serviço de voz e um domínio Ethernet remoto contendo o destino do serviço de voz. A interconexão dos dois domínios se dá através de um enlace de 33,6Kbit/s que contém dois roteadores nas pontas. A seguir,

apresenta-se uma breve descrição dos equipamentos envolvidos, em cada um dos cenários de teste.

### 3.1 Descrição dos equipamentos do Ambiente de Teste

Os equipamentos envolvidos nos domínios Ethernet do arranjo experimental da figura 3 são: *hubs*, microcomputadores, roteadores e modems. Os domínios Ethernet consistem de dois *hub's* Ethernet de 10Mbit/s, seis microcomputadores, dois roteadores Cisco 4000, com diversas portas LAN de 10Mbit/s e uma porta WAN de 2Mbit/s cada, além de um par de modems configurados para transmitir a 33.600 bit/s.

No domínio Ethernet local, foi configurada uma estação para a origem da voz, outra estação para o destino - que pode ser definida tanto no domínio local como remoto. Foi definida uma terceira estação que tem condições de capturar os pacotes que trafegam tanto no domínio local quanto no domínio remoto. Por se tratar de um estudo experimental, e os equipamentos estarem todos dentro do mesmo ambiente, foi-nos possível a colocação de duas placas de rede nesta estação, e desta forma, capturar os pacotes de voz nos dois domínios de colisão. Aos pacotes capturados por esta estação, são acrescentados etiquetas de tempo (*time stamp*) correspondentes aos instantes em que foram capturados. Além disso, foram definidas em cada domínio, mais três estações para gerar tráfego, tanto dentro de um domínio como entre os domínios de rede.

Com este arranjo experimental é possível observar a qualidade da voz de uma conexão, que pode ser tanto remota como local, em condições de carga da rede que pode ser perfeitamente controlada. Foram feitas análises da qualidade do serviço de voz tanto em condições de baixa carga como em condições de carga pesada, ou próximo dos 30% da carga máxima ( $\approx 3,0$  Mbit/s), admitidos em cada domínio de colisão Ethernet.

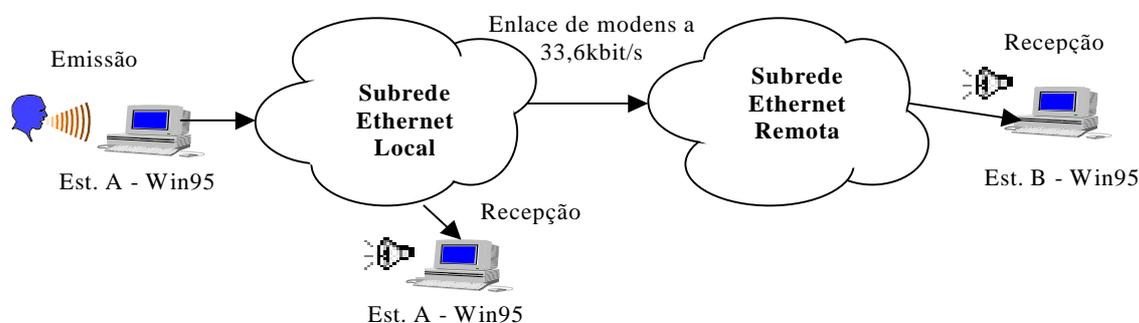


Figura 3 – Configuração local e remota do NetMeeting.

No arranjo experimental utilizou-se uma estação para gerar a voz e outra para a recepção do áudio. As estações usam como sistema operacional o Windows 95, e são equipadas com kit multimídia. Utilizamos, como ferramenta de aplicação de voz sobre IP, o Microsoft NetMeeting versão 2.1. Esta ferramenta permite ao usuário configurar o padrão de compactação desejado. Nos testes realizados, foi utilizado o padrão G.729 do ITU-T, ou seja, o codificador gera um feixe de bits contínuo a uma taxa de 8 kbit/s [ITU 96].

Na figura 3, apresentam-se as duas configurações do NetMeeting (local e remota) utilizadas no nosso estudo experimental. Após estabelecer a conexão com o NetMeeting, é possível realizar uma conversação, utilizando-se o microfone do kit multimídia, como se fosse um telefone normal. O NetMeeting, além da compressão de voz, implementa outros recursos para otimizar o uso da banda de rede como, por exemplo, a supressão do silêncio. Sendo o

objetivo do nosso estudo verificar a utilização e os atrasos reais verificados nos pacotes de voz, utilizou-se uma fonte de áudio contínua, sob forma de música, evitando-se assim o supressor de silêncio do NetMeeting.

### 3.2 Descrição da Ferramenta de Captura e Análise

A ferramenta de Captura e Análise dos Pacotes de Voz que foi desenvolvida, divide-se em dois módulos: um módulo (captura de pacotes) funciona em uma estação com o sistema operacional Linux e o segundo módulo (análise estatística) funciona em uma estação com o sistema operacional Windows 95. Para capturar os pacotes da rede, foi desenvolvido um programa utilizando-se a linguagem C em ambiente Linux. O módulo captura todos os pacotes IP que estão trafegando pela rede, acrescentando a cada pacote capturado uma etiqueta com o instante (em microsegundos) em que foi capturado. Para agilizar a captura dos pacotes, estes são armazenados na memória RAM, mas apenas o cabeçalho dos pacotes acrescido da etiqueta com o instante da sua captura. Desta maneira, consegue-se uma drástica redução na ocupação da memória. O tempo de captura é limitado pelo número máximo de pacotes que podem ser capturados e armazenados na RAM. Em nossos testes este limite estava em torno de 30.000 pacotes que, dependendo do tráfego da rede, possibilita capturar de 2 a 8 minutos de áudio. Após capturar os 30.000 pacotes - todos armazenados na memória RAM -, é gerado um arquivo log de texto com todos os pacotes capturados.

Para capturar pacotes nos dois domínios de colisão e ao mesmo tempo não perder os instantes de sua emissão e recepção, adotou-se a seguinte estratégia: foram instaladas duas placas de rede em uma estação com Linux, executando-se o software do módulo de captura de pacotes nesta estação. Desta forma a estação tem condições de capturar simultaneamente os pacotes nos dois domínios de colisão Ethernet e, assim, pode-se determinar o instante em que são capturados os pacotes em cada uma das subredes.

O segundo módulo da ferramenta corresponde à Análise Estatística dos pacotes capturados e foi desenvolvida utilizando-se o Delphi 3.0. Este módulo importa o arquivo log de texto gerado pelo módulo de captura de pacotes. Suas principais funcionalidades são: converter os valores de hexadecimal para decimal, filtrar os pacotes de voz através do endereço IP de origem e de destino das estações responsáveis pelo tráfego de voz, e gerar os gráficos que auxiliam na interpretação dos resultados.

Um dos objetivos principais do módulo de análise é calcular o atraso correspondente a um determinado pacote ao ser transportado da origem até o seu destino. Este cálculo é feito através da diferença entre o instante de tempo em que o pacote é gerado e o instante em que é capturado no destino considerado. O arquivo *log* importado do módulo de captura contém os pacotes duplicados e a identificação de cada dupla de pacotes é feita através do campo de identificação dos pacotes IP. Com esta técnica de duplicação dos pacotes é possível calcular, além do tempo de transporte entre os dois domínios de colisão, também a taxa de perda de pacotes devido a problemas de congestionamento. No caso de um pacote capturado não estar duplicado, significa que o mesmo foi perdido e, desta forma, pode-se identificar se o problema da baixa qualidade do serviço de voz é devido ao atraso dos pacotes, ou se é devido a perda de pacotes.

## 4 Resultados Experimentais Obtidos

Nesta seção apresentam-se os resultados experimentais obtidos através da análise do tráfego de voz gerado pelo Microsoft NetMeeting utilizando-se a ferramenta de Análise de

Pacotes desenvolvida e os arranjos experimentais sugeridos na sessão 3. Foram considerados dois cenários experimentais: (1.) tráfego de voz em um domínio de rede local e, (2.) tráfego de voz entra dois domínios, passando por um enlace com dois roteadores. Em cada cenário experimental foram realizados diversos testes, com diferentes perfis de tráfego de dados de fundo nas subredes, além do tráfego de voz.

1º Cenário Experimental:

Primeiro Teste – Somente tráfego de voz,

Segundo Teste – Tráfego de voz com rajadas de dados aleatórias,

Terceiro Teste – Tráfego de voz com tráfego de dados uniforme e contínuo,

Quarto Teste – Tráfego de voz com tráfego de dados pesada<sup>2</sup>.

2º Cenário Experimental:

Primeiro Teste – Somente tráfego de voz,

Segundo Teste – Tráfego de voz e tráfego de dados intra e entre os domínios.

Inicialmente, para observar com mais detalhes como são emitidos os pacotes de voz pelo NetMeeting, levantou-se o histograma da figura 4, em que foram tabulados os tempos entre a emissão dos pacotes de voz durante o período de um segundo.

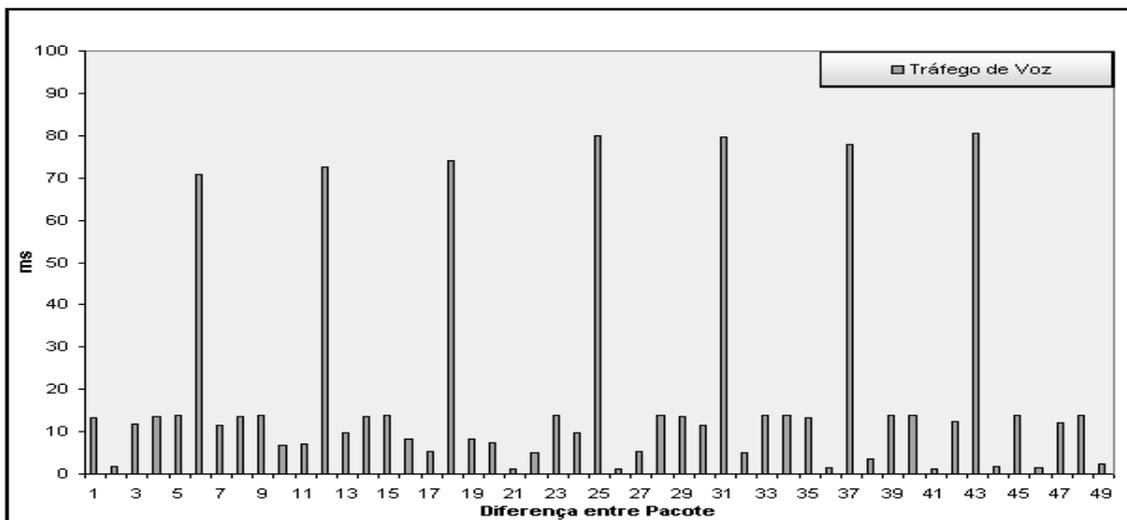


FIGURA 4 – Intervalo entre pacotes de voz de um determinado segundo

Observa-se que o NetMeeting emite em média 50 pacotes de voz por segundo, todos com um tamanho fixo de 75 bytes, o que corresponde a uma banda média de 30 kbit/s para este serviço. Escolheu-se o *codec* padrão G.729 no NetMeeting, que corresponde a emissão um fluxo de voz de 8 kbit/s. Verifica-se que o encapsulamento deste fluxo de voz em pacotes IP, introduz uma redundância  $r$ , dada por:

$$r = \left(1 - \frac{\text{taxa efetiva}}{\text{taxa real}}\right) \cdot 100 = \left(1 - \frac{8}{30}\right) \cdot 100 = 73,33\%$$

Esta redundância enorme se deve principalmente ao tamanho do cabeçalho do IP (20 bytes) e os poucos bytes de informação encapsulados em cada pacote de voz. Observa-se também pela figura 4 que os pacotes de voz IP são emitidos em rajadas de 6 pacotes com um espaçamento médio entre os pacotes de  $\Delta t = 8,8$  ms e um espaçamento médio entre as rajadas

<sup>2</sup> Carga pesada no caso corresponde a um fator de carga ( $\rho$ ) da rede próximo de 30% da capacidade máxima ( $C$ ), que é considerado o limite em redes Ethernet, para que não haja problemas de congestionamento (atraso alto) [ROC98].

de 73 ms. O fluxo de pacotes IP corresponde, portanto, a um tráfego em rajadas do tipo ON/OFF com o período médio de ON,  $T_{ON} = 44$  ms, e o período médio de OFF,  $T_{OFF} = 73$  ms, como pode ser observado na figura 5

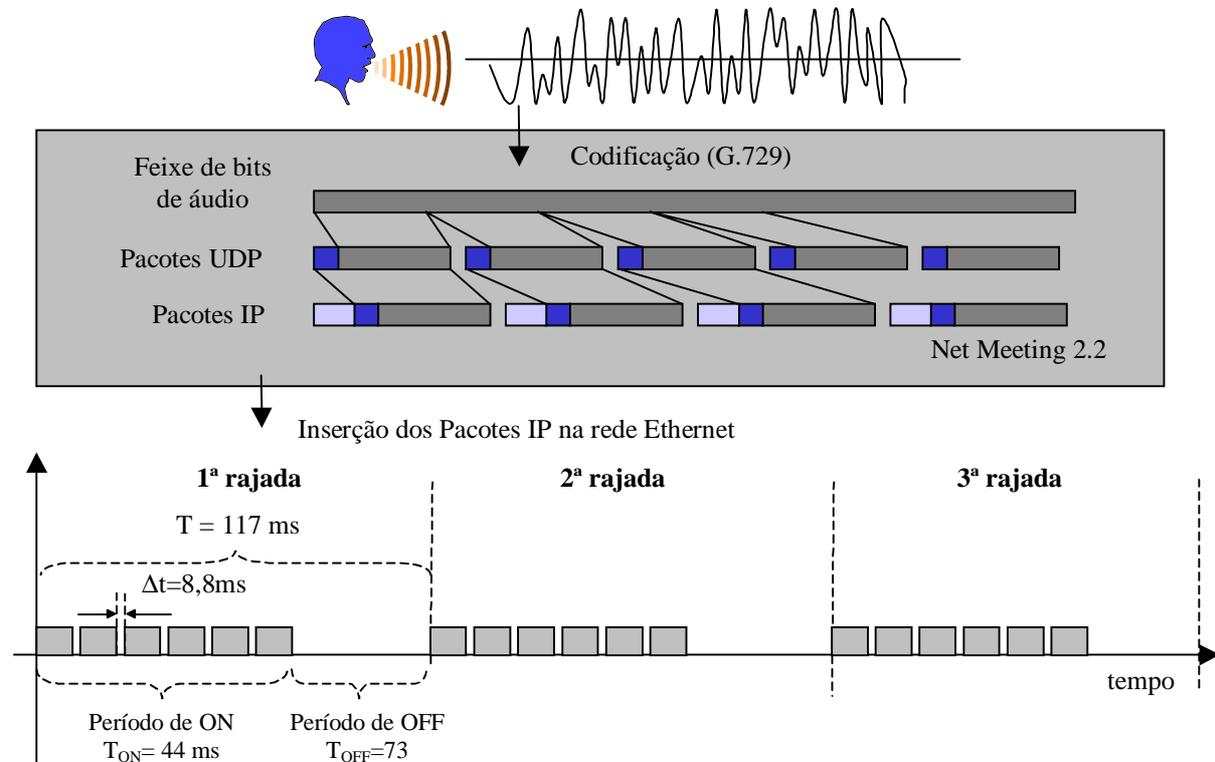


FIGURA 5 – Seqüência de processos para a transmissão de VoIP

#### 4.1 Avaliação dos Resultados do Primeiro Cenário Experimental

Na figura 6, são apresentados, sob forma de histogramas, os espaçamentos verificados entre os pacotes de voz em relação a quatro situações de carga (testes) distintas da rede do cenário experimental local. Observa-se claramente que, tanto o espaçamento entre os pacotes, como o espalhamento temporal dos pacotes de voz (*jitter*), em nenhum dos quatro testes comprometeu a qualidade da voz. Também não se verificaram, em nenhum momento, perdas de pacotes durante os 40 s em que foram realizados os diversos testes.

Para a análise dos resultados obtidos nos diferentes testes, definiram-se os seguintes parâmetros e as seguintes variáveis probabilísticas:

$X_i$  = Intervalo entre pacotes (cenário 1), ou, atraso do pacote de voz (cenário 2)

$R_i$  = Taxa; pacotes de voz em um segundo

$L$  = Tamanho dos pacotes de voz : 75 bytes

$C$  = Capacidade máxima da LAN: 10 Mbit/s (Capacidade máxima efetiva: 0,3C)

$U_i$  = Utilização da banda por segundo

$\sigma_x$  = Desvio Padrão de  $X$

$IC$  = Intervalo de confiança

$T$  = Tempo total considerado

$n$  = Total de pacotes de voz no intervalo de tempo  $T$ .

Utilizaram-se também as seguintes relações estatísticas:

Tempo médio (ou valor esperado) entre pacotes:

$$\bar{X} = E[X] = \frac{\sum_{i=1}^n X_i}{n}$$

Desvio padrão ( $\sigma_x$ ) da variável probabilística [ $X_i$ ]:

$$\sigma_x = \sqrt{\frac{E[X_i - \bar{X}]^2}{n}} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}}$$

Valor médio da taxa de pacotes de voz num período de  $T$  segundos:

$$\bar{R} = \frac{\sum_{i=1}^T R_i}{T}$$

Considerando uma capacidade máxima da rede  $C = 10 \text{ Mbit/s}$ , e um tamanho do pacote  $L$  (50 bytes), então o percentual de utilização da banda ( $U$ ) pelo serviço de voz em um segundo será dado por:

$$U_i = \frac{(R_i * L * 8)}{C} * 100 \quad (\text{rede } 10 \text{ Mbit/s})$$

O valor médio de utilização da banda em relação a um período de  $T$  segundos será:

$$\bar{U} = \frac{\sum_{i=1}^T U_i}{T}$$

Intervalo de Confiança ( $IC$ ):

$$IC = K(\alpha) \left( \frac{\sigma_x}{\sqrt{n}} \right)$$

Considerou-se nas nossas análises um nível de confiança com um percentual de 95%, o que corresponde a um  $K(\alpha) = 1,96$ .

Na Tabela 1 encontram-se resumidos os parâmetros experimentais obtidos nos quatro testes realizados em relação ao primeiro cenário experimental (tráfego de voz em um domínio de LAN local).

TABELA 1 – Resultados dos testes do Primeiro Cenário Experimental

| Teste                        | $n$  | $T$  | $\bar{U}$ | $\bar{R}$ | $\bar{X}$ | $\sigma_x$ | $IC$                  |
|------------------------------|------|------|-----------|-----------|-----------|------------|-----------------------|
| 1 (só voz)                   | 1975 | 40 s | 0,29%     | 50        | 19,95 ms  | 24,23 ms   | $\pm 1,07 \text{ ms}$ |
| 2 (voz e rajadas de dados)   | 1975 | 40 s | 0,28%     | 49,5      | 19,96 ms  | 33,05 ms   | $\pm 1,46 \text{ ms}$ |
| 3 (voz e dados uniforme)     | 1975 | 40 s | 0,29%     | 50        | 19,94 ms  | 24,42 ms   | $\pm 1,08 \text{ ms}$ |
| 4 (voz e dados carga limite) | 1975 | 40 s | 0,28%     | 49,5      | 19,96 ms  | 32,48 ms   | $\pm 1,43 \text{ ms}$ |
| Médias                       |      |      | 0,285%    | 49,75     | 19,95 ms  | 28,54 ms   | $\pm 1,26 \text{ ms}$ |

O número total de pacotes ( $n$ ) em cada teste foi de 30.000 pacotes (limite máximo de captura da ferramenta) sendo que deste total, 1.975 pacotes foram de voz, o que corresponde a um percentual de 6,58% do total de pacotes que trafegaram pela rede num período de 40 segundos ( $T$ ).

Conclui-se, a partir dos resultados da Tabela 1, que a aplicação de voz sobre IP em uma rede Ethernet de 10 Mbit/s gera um tráfego relativamente baixo - em torno de 30,76

kbit/s, que corresponde aproximadamente a 0,3% da capacidade ( $C$ ) total da rede. Observa-se também que, mesmo com as diferentes variações da carga (testes), o tráfego de voz manteve-se praticamente inalterado, apresentando sempre uma boa qualidade de voz. Isto se deve a dois fatores: (1) a perda de pacotes de voz nos quatro testes foi praticamente nula e (2) a variação do atraso (*jitter*) se manteve baixo.

Na figura 6, apresenta-se um histograma do espalhamento temporal dos pacotes de voz em relação aos quatro testes efetuados, considerando-se intervalos de tempo fixos de 10 ms. Pode-se observar que o valor médio do espaçamento entre os pacotes ( $\bar{X}$ ) manteve-se praticamente o mesmo nos quatro testes. No entanto, o valor do desvio padrão ( $\sigma_x$ ) mostra um aumento claro com o aumento da carga na rede, principalmente com tráfego em rajadas ou carga pesada (testes 2 e 4).

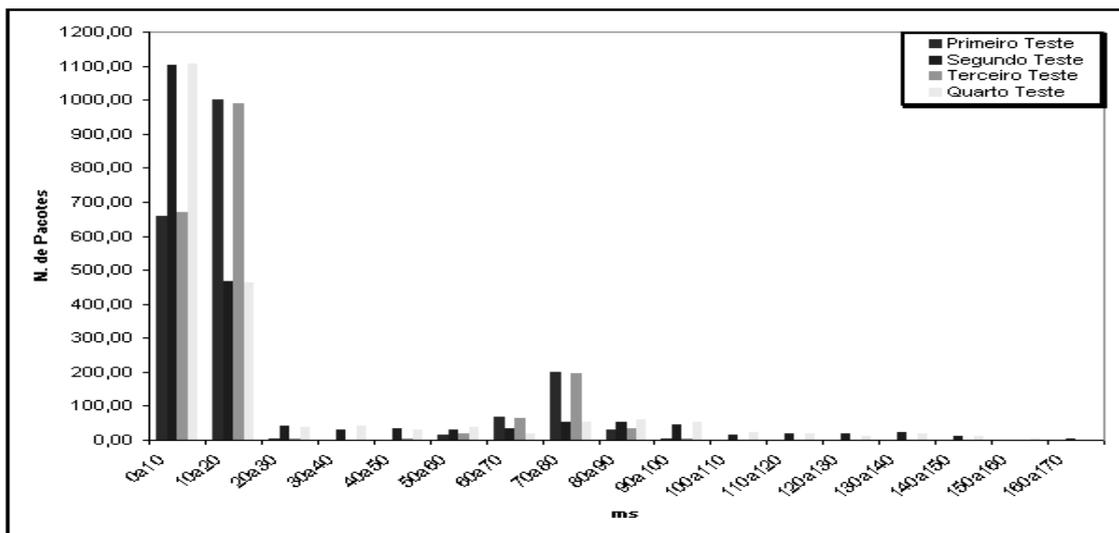


FIGURA 6 – Histograma dos atrasos entre os pacotes de voz dos quatro testes efetuados no Primeiro Arranjo Experimental

Já em relação aos testes 1 e 2, tráfego de voz com tráfego de fundo nulo ou uniforme e baixo, respectivamente, apresentaram valores de *jitter* baixos e equivalentes. Conclui-se que o tráfego de voz em um meio partilhado Ethernet, com carga baixa e uniforme ( $\leq 0,3C$ ), não apresenta perdas significativas de qualidade de voz.

A partir dos testes realizados no primeiro cenário experimental, pode-se concluir que uma única aplicação de voz sobre IP, utilizando-se no nível de enlace uma rede Ethernet de 10 Mbit/s, é perfeitamente viável. Mesmo com as aplicações de dados gerando uma carga pesada na rede, a aplicação de voz sempre manteve uma qualidade aceitável.

#### 4.2 Avaliação dos Resultados do Segundo Cenário Experimental

O segundo cenário experimental é caracterizado pela geração do serviço de voz num domínio local e a sua recepção em um domínio remoto. A interconexão entre dois domínios se dá através de um enlace de 33.600 bit/s, com dois roteadores nas pontas como é mostrado na figura 2. Os pacotes de voz serão roteados, portanto, por duas vezes, antes de chegarem ao seu destino.

Na tabela 2 estão tabulados os principais resultados obtidos a partir dos dois testes efetuados. No primeiro teste considerou-se um tráfego de fundo (*broadcast*) baixo e nenhuma outra aplicação em atividade nos dois domínios além da aplicação de voz. No segundo teste,

além da aplicação de voz, considerou-se um tráfego moderado interno aos domínios, além de um tráfego médio entre os domínios.

TABELA 2 – Resultados dos testes do Segundo Arranjo Experimental

| Teste             | n   | T   | $\bar{U}$ |             | $\bar{R}$ | $\bar{X}$ | $\sigma_x$ | IC            |
|-------------------|-----|-----|-----------|-------------|-----------|-----------|------------|---------------|
|                   |     |     | 10 Mbit/s | 33,6 kbit/s |           |           |            |               |
| 1 (só voz)        | 800 | 16  | 0,29%     | 89,5%       | 50        | 74,51 ms  | 12,70 ms   | $\pm 0,56$ ms |
| 2 (voz com dados) | 800 | 148 | 0,15%     | 46,7%       | 26,1      | 128,17ms  | 98,05 ms   | $\pm 4,32$ ms |
| Médias            |     |     | 0,22%     | 68,1%       | 38,1      | 101,34ms  | 55,38 ms   | $\pm 2,44$ ms |

Observa-se que o valor médio do atraso do primeiro teste do cenário dois, da ordem de 74,51 ms, aumentou sensivelmente se comparado com os 19,95 ms observados no primeiro teste do cenário experimental um. Saliente-se, no entanto, que todos os pacotes chegaram dentro do intervalo de tempo máximo suportado pela aplicação de voz sobre IP, que é em torno de 300 ms. No segundo teste do cenário 2, tráfego de voz com diversas aplicações de dados, o valor do atraso médio aumentou de 74,51 ms para 128,17 ms, muitos pacotes chegaram depois do tempo mínimo, o que prejudicou sensivelmente a qualidade da voz, pois além dos pacotes perdidos em função do aumento do tráfego, houve descarte de pacotes.

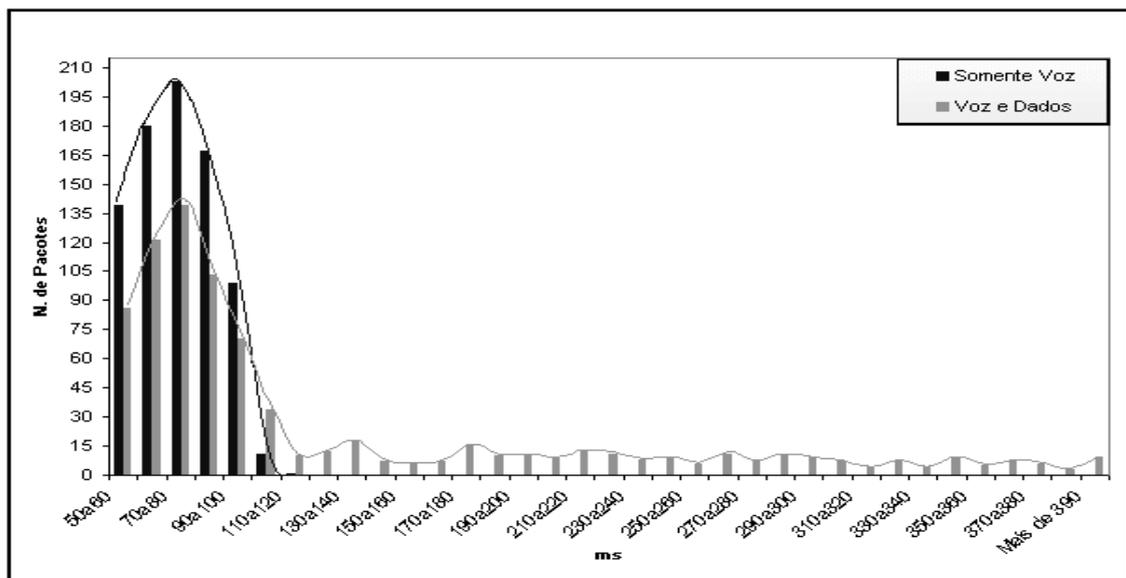


FIGURA 7 – Histograma dos atrasos nos testes realizados segundo o Segundo Cenário Experimental

Na figura 7, apresenta-se um histograma dos atrasos em relação aos dois testes. Observa-se que, em relação ao teste 1 - somente o tráfego de voz -, a curva de distribuição do atraso apresenta-se próxima de uma distribuição normal (curva preta da Figura 7). Já a curva de distribuição do atraso, considerando-se tráfego de voz e dados (teste 2), está mais próxima de uma distribuição exponencial (curva cinza da Figura 7). Esta observação é reforçada pelos valores  $\bar{X}$  e do desvio padrão  $\sigma_x$  da Tabela 2.

Os valores calculados para os testes do segundo cenário experimental utilizam as mesmas expressões e parâmetros já descritas anteriormente. A única exceção é em relação à variável  $X_i$ , que no cenário experimental 1 correspondia ao intervalo de tempo entre os

pacotes de voz, e no cenário experimental 2, corresponde ao atraso que sofre um pacote de voz ao ser transmitido entre as subredes. A variável  $n$  foi fixada em 800 pacotes de voz, devido ao limite do tamanho do arquivo de pacotes capturados pela primeira ferramenta. Com a existência de pacotes de voz e dados na rede, logo esgotou-se a capacidade do arquivo, o que nos obrigou a usar um número menor de pacotes de voz.

No primeiro teste do arranjo experimental 2, foram capturados em torno de 26.000 pacotes de voz, mas utilizou-se um valor de  $n$  menor, para que pudessem ser apresentados no mesmo histograma os dois testes (Figura 7). Com o valor de  $n$  fixo em 800 pacotes de voz, o primeiro teste apresentou uma taxa média de 50 pacotes por segundo em um período ( $T$ ) de 16 segundos, portanto sem perdas. No segundo teste, devido ao tráfego de dados, a taxa de pacotes caiu para 26,1 pacotes por segundo em um período ( $T$ ) de 148 segundos. Observa-se que houve uma grande perda de pacotes, o que acarretou um aumento no tempo necessário para capturar os 800 pacotes de voz.

## 5 Conclusões e Trabalhos Futuros

Através deste estudo preliminar, conclui-se que ainda são necessários muitos avanços na tecnologia de voz em tempo real sobre redes IP, principalmente no que se refere ao fornecimento de QoS para este serviço. Os resultados experimentais apresentados mostraram que a aplicação VoIP baseada em um *codec* tipo G.729, que gera um fluxo de 8 kbits e origina um tráfego do tipo ON/OFF que apresenta rajadas com 6 pacotes de voz com  $T_{ON}= 44 ms$  e  $T_{OFF}=73 ms$  e um período fixo  $T=117 ms$ . A banda ocupada pelo fluxo de voz, após o encapsulamento em pacotes IP, é de aproximadamente 30 kbit/s. Apesar da escolha do tipo de *codec* influir na banda ocupada pelo serviço, é o encapsulamento em pacotes IP o que introduz o maior desperdício de banda, podendo chegar a uma redundância em torno de 70% da banda total ocupada pelo serviço.

Pelos testes realizados no primeiro cenário experimental, pode-se concluir que o tráfego de voz em um domínio de rede local é perfeitamente viável, com qualidade de voz boa, desde que possa ser assegurado que o tráfego neste domínio não ultrapasse os 30% da capacidade máxima ( $C$ ) desta rede. Tendo em vista que a banda necessária para um conexão de voz no nível IP se situa em torno de 25-35 kbit/s, acreditamos que em ambientes de redes locais que apresentam segmentação dos domínios através de um *switch* – e, portanto, uma boa reserva de banda -, um serviço de voz local será perfeitamente possível. Picos de carga esporádicos e de pouca duração afetam a qualidade da voz, mas não a ponto de comprometer o serviço. Muito do que se consegue em termos de compressão pelos modernos codificadores, que conseguem transformar um feixe PCM de 64 kbit/s (canal de voz) para 8 kbit/s, é perdido novamente pelas redundâncias introduzidas através dos cabeçalhos do protocolo TCP/IP no momento do encapsulamento destes feixes.

Foram, porém, os resultados obtidos no segundo cenário experimental, que revelaram a fragilidade do TCP/IP em relação às exigências de QoS do serviço de voz. Neste cenário, implementou-se um serviço de voz entre dois domínios de redes locais, interconectados através de um enlace de 33,6 kbit/s, passando por dois roteadores. Constatou-se que neste caso só era atingida uma qualidade de voz razoável quando o suporte como um todo; enlace, roteadores e redes locais, era disponibilizado integralmente para este serviço, sem tráfego de dados de qualquer outra aplicação.

Verifica-se que a distribuição dos atrasos dos pacotes de voz, à medida que aumenta o número de saltos (*hops*) em roteadores ao longo do caminho dos pacotes de voz, mais próxima fica de uma função de distribuição exponencial; modelo de tráfego largamente

utilizado na modelagem de fluxos de tráfego genéricos. No teste 2 do cenário experimental 2 (tabela 2), pode-se observar que o *jitter* do fluxo de voz, que pode ser associado ao desvio padrão ( $\sigma_x$ ) associado ao atraso ( $X_i$ ) dos pacotes, aumenta dramaticamente com apenas dois níveis de roteamento dos pacotes, tornando a qualidade da voz inaceitável.

Acreditamos que uma plataforma TCP/IP para suportar de forma adequada serviços multimídia com exigências de QoS específicas, só será obtida através das arquiteturas de rede como DiffServ [BL A98] [DO V99], IntServ [XIA 99], MPLS [ROS 99] ou roteamento com QoS (engenharia de tráfego) [KAM 2000], todas recentemente propostas. Julgamos que a ferramenta de análise desenvolvida, com algumas modificações, também será útil na avaliação do desempenho dos serviços multimídia nestas novas arquiteturas. Em projetos de redes corporativos (Intranets) em que se queira integração de dados com serviços multimídia de tempo real, não basta, portanto, a disponibilização de enormes quantidades de banda para resolver o problema dos atrasos: é imperativo que também sejam incorporados os novos avanços em termos de arquitetura, para que possam ser assegurados os QoS exigidos por estes serviços.

## 6 Referências Bibliográficas

- [BER 99] BERNET, Y. et al., “**A Framework for Differentiated Services**”, Internet Draft, disponível em: <http://www.ietf.org/internet-drafts/draft-ietf-diffserv-framework-02.txt>, Feb. 1999.
- [BLA 98] BLACK, D. et al. “**An Architecture for Differentiated Services**”, RFC 2475, The Internet Society, dezembro de 1998.
- [BRA 94] BRADEN, R. et al. “**Integrated Services in the Internet Architecture: na Overview.**” RFC 1633, The Internet Society, junho de 1994.
- [BRA 97] BRADEN, R. et al., “**Resource ReSerVation Protocol (RSVP)**”, Version1, Functional Specification. RFC 2205, The Internet Society, setembro de 1997.
- [CAL 99] CALVO, S. **A Linguagem dos Negócios**. Network Computing, São Paulo, v.1, n.4, p.34-42, jun de 1999.
- [CRA 98] CRAWLEY, E. et al. **A Framework for QoS-based Routing in the Internet**, RFC 2386, The Internet Society, agosto de 1998.
- [COX 96] COX, R. V.; KROON, P. **Low Bit-Rate Speech Coders for Multimedia Communication**. IEEE Communications Magazine, v. 34, n.12, Dec. 1996, p.34-41
- [CUN 99] CUNHA, D. **Tecnologia MLPS e DiffeSer com ATM**. Uberlândia: Universidade Federal de Uberlândia, 1999. Disponível em: <daniela@lrc.deene.ufu.br>.
- [DOV 99] DOVROLIS, C.; RAMANATHAN, P. **A case for Relative Differentiated Services and the Proporcional Differentiation Model**. IEEE Network, v.13, n.5, Sep/Oct 1999.
- [ITU 96] ITU Rec. G.723.1, “**Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 and 6.3 kbit/s**” Mar. 1996
- [ITU 96a] ITU Rec. H.323, “**Visual Telephone Systems and Equipment for Local Area Networks wich provide a Non-Guaranteed Quality of Service**” Nov. 1996.
- [JAI 98] JAIN, R. **Voice over IP: Issues and Challenges**. Nortel, Canada, August 1998. Em janeiro de 2000 disponível em: <http://www.cis.ohio-state.edu/~jain/talks/voip.htm>.
- [KAM 2000] KAMIENSKI, C. A.; SADOK, J. “**Engenharia de Tráfego em uma Rede de Serviços Integrados.**” In: 18º Simpósio Brasileiro de Redes de Computadores, Anais:...SBRC 2000, Belo Horizonte, MG, 23 a 26 de maio de 2000, p.151-166.

- [KAT 97] KATSOUROS, M. S. “**Voice over IP: Policy and Regulatory Issues**”, 1997.  
Disponível em: <http://www.dpbs.umd.edu/~mark/631paper.html>, em janeiro de 2000.
- [KOS 98] KOSTAS, T.J., BORELLA, M.S., SIDHU, I., SCHUSTER, G.M., GRABIEC, J., MAHLER, J. “**Real-Time Voice Over Packet-Switched Networks**”, IEEE Network, Jan./Feb. 1998.
- [LIU 99] LIU, C. “**Multimedia Over IP: RSVP, RTP, RTCP, RTSP**”. disponível em: [http://www.cis.ohio-state.edu/~jain/cis788-97/ip\\_multimedia/index.htm](http://www.cis.ohio-state.edu/~jain/cis788-97/ip_multimedia/index.htm), em dezembro de 2000.
- [MIN 98] MINOLI, D.; MINOLI, E. **Delivering Voice over IP Networks**, Wiley Computer Publishing, John Wiley & Sons Inc. New York, 1998, 276p.
- [NOR 99] Nortel, “**RSVP: Resource ReserVation Protocol**”, Nortel/Micom Networks.  
Disponível em: <http://www3.nortelnetworks.com/WhitePapers/rsvp/wprsvpte.htm>, em dezembro de 2000.
- [ROC 98] ROCHOL, J. **Redes de Computadores. Apostila Interna do Grupo de Redes do Computadores do CPGCC/UFRGS**. Porto Alegre, 1998.
- [ROS 99] ROSEN, E. et al., “**Multiprotocol Label Switching Architecture**”, Internet Draft, The Internet Society. Disponível em: <http://www.ietf.org/internet-drafts/draft-ietf-mppls-arch-06.txt>, em agosto de 1999.
- [SCH 96] SCHULZRINNE et al., “**RTP: A Transport Protocol for Real-Time Applications**”, RFC 1889, The Internet Society, Jan. 1996.
- [THO 96] THOM, G. A. **H.323: The Multimedia Communication Standard for Local Area Networks**. IEEE Communications Magazine, v.34, n.12, December 1996, p.52-56.
- [XIA 99] XIAO, X.; NI, L. M. **Internet QoS: A Big Picture**, IEEE Network, v.13, n.2, Mar/Apr 1999; p.8-18.
- [ZAP 98] ZAPPALA, D. “**RSVP Protocol Overview**”, Disponível em: <http://www.isi.edu/div7/rsvp/overview.html>, em dezembro de 2000.