

**FUNDAÇÃO DE ENSINO “EURÍPIDES SOARES DA ROCHA”**  
CENTRO UNIVERSITÁRIO “EURÍPIDES DE MARÍLIA” – UNIVEM  
PROGRAMA DE MESTRADO EM COMPUTAÇÃO

**CAIRO GOMIDE**

**MULTIPLEXAÇÃO DE SINAIS DE VOZ  
EM REDES IP**

MARÍLIA  
2004

**CAIRO GOMIDE**

**MULTIPLEXAÇÃO DE SINAIS DE VOZ  
EM REDES IP**

Dissertação apresentada ao Programa de Mestrado do Centro Universitário Eurípides de Marília, mantido pela Fundação de Ensino “Eurípides Soares da Rocha”, para obtenção do Título de Mestre em Computação.

Orientador:  
Prof. Dr. Jorge Luiz e Silva

**MARÍLIA  
2004**

**CAIRO GOMIDE**

**MULTIPLEXAÇÃO DE SINAIS DE VOZ EM REDES IP**

Banca examinadora da dissertação apresentada ao Programa de Mestrado do Centro Universitário Eurípides de Marília, mantido pela Fundação de Ensino “Eurípides Soares da Rocha”, para obtenção do Título de Mestre em Computação.

Resultado: **APROVADO.**

Orientador: **Prof. Dr. Joge Luiz e Silva**

1º Examinador: **Prof. Dr. Ildeberto Aparecido Rodello**

2º Examinador: **Prof. Dr. Hélio Crestana Guardia**

Marília, 03 de setembro de 2004.

GOMIDE, Cairo. MULTIPLEXAÇÃO DE SINAIS DE VOZ EM REDES IP. 110 f. Dissertação (Mestrado em Computação) – Centro Universitário Eurípides de Marília, Fundação de Ensino Eurípides Soares da Rocha, Marília, 2004.

## RESUMO

A convergência na área de telecomunicações introduz aspectos para diferentes tipos de meios físicos de transmissão de sinais. Os tráfegos de voz sobre IP e aplicações multimídia estão mudando e trazendo novos conceitos aos processos de transmissão de dados e voz em meio digital. Novas técnicas de transmissão estão sendo avaliadas a todo instante, pois a diminuição de custos com a aplicação dessas novas técnicas estão tornando cada vez mais atrativa a sua adesão, porém tudo isso pode repercutir em um aumento considerável de tráfego, o que pode também comprometer a qualidade do serviço. Este trabalho apresenta de modo organizado, os conteúdos dos principais artigos e padronizações encontrados na literatura sobre o tema voz sobre IP - VoIP (voice over IP) a fim de avaliar os principais tipos de problemas com o empacotamento dos sinais de voz. Propostas e análises sobre a transmissão de sinais de voz, QoS (Qualidade de Serviço) sobre os sinais em redes, a repercussão sobre a variação dos meios físicos e a forma em que os sinais são inseridos a esses meios, são variantes para a avaliação deste trabalho. O trabalho foca principalmente o impacto do atraso dos pacotes de voz em um método de transmissão de sinais multiplexados. Para essa análise é proposto o desenvolvimento de um modelo matemático, com a aplicação de variados métodos de distribuição do tráfego com e sem atribuição de QoS. Para simulação de implementação do ambiente analisado foi desenvolvido, através do software *Network Simulator*, gráficos comparativos entre os resultados do modelo teórico e o modelo simulado que fazem parte de uma conclusão do estudo proposto.

**Palavras-Chave:** VoIP, Redes, QoS, Protocolos.

GOMIDE, Cairo. MULTIPLEXAÇÃO DE SINAIS DE VOZ EM REDES IP. 110 f. Dissertação (Mestrado em Computação) – Centro Universitário Eurípides de Marília, Fundação de Ensino Eurípides Soares da Rocha, Marília, 2004.

### ABSTRACT

The convergence in the telecommunication area introduces aspects for different types of physical medium of sign transmissions. The voice traffics about IP and multimedia applications are changing and leading to new thoughts of the voice process and data transmission in digital medium. New transmission techniques are being valued all the time, because the cost decrease, through those new technique application, is turning its more and more attractive; however, it can reverberate in a notable traffic increase, which can also endanger the service quality. This work shows, in a organized way, the contents of the main articles and standards found in the literature about the IP voice thesis – VoIP (voice over IP) to value the main types of voice sign binding problems. Proposals and analysis about the voice sign transmission, QoS (Service Quality) about signs in net, the repercussion about the physical medium diversification and the way as the signs are inserted to these means, are variant for this work valuation. The work focuses on specially the impact of the delay of the voice packages in a multiplexed sign transmission system. It is purposed for these analysis, the development of a mathematical model, with the use of diverse traffic distribution methods, with and without QoS attribution. To simulate the analysed site implementation, graphs comparing the results of both the theoretic and simulated models which are part of a conclusion for the proposed study were developed, through the Network Simulator Software.

**Keywords:** VoIP, Networks, QoS, Protocols, Gateway

## LISTA DE ILUSTRAÇÕES

Figura 1: Comutação por circuitos e comutação por pacotes (PENNO, 1999).....	18
Figura 2: Arquitetura PC para PC. ....	21
Figura 3: Arquitetura PC para telefone.....	21
Figura 4: Arquitetura telefone para telefone.....	23
Figura 5: Apresentação completa de uma rede VoIP. ....	30
Figura 6: Rede SS7 (CISCO, 2003) .....	32
Figura 7: Pilha de protocolos SS7 vs. Modelo de Referência OSI.....	35
Figura 8: Processos de comunicação de elementos VoIP baseados em H.323. ....	37
Figura 9: Cabeçalho do Protocolo RTP .....	38
Figura 10: Controle de ações do <i>Media Gateway Control Protocol</i> – MGCP (ARANGE, 2003).....	41
Figura 11: Execução para criação do chamado SIP. ....	43
Figura 12: RSVP: orientado ao multicast e ao receptor. (Kurose, 2003). ....	48
Figura 13: Comparação do tempo de chegada esperado com o tempo de chegada real (ROWE, 1999).....	55
Figura 14: Um exemplo simples de rede Diffserv (ROSS, 2003).....	59
Figura 15: Atraso de fila médio e a intensidade de tráfego.....	68
Figura 16: Breve modelo esquemático do processo de entrada e saída de pacotes.....	68
Figura 17: 1º Caso – Exp. Neg. Voz/Dados – Sem Prioridade .....	74
Figura 18: 1º Caso – Exp. Neg. Voz – Com Prioridade .....	75
Figura 19: 1º Caso – Exp. Neg. Dados – Com Prioridade.....	75
Figura 20: 2º Caso – Comprimento Fixo para Voz e dados – Sem Prioridade. ....	76
Figura 21: 2º Caso – Comprimento Fixo para Voz – Com Prioridade.....	77
Figura 22: 2º Caso – Comprimento Fixo para Dados – Com Prioridade. ....	77

Figura 23: 3º Caso – Comprimento Fixo p/ Voz e Exponencial Negativa p/ Dados – Sem Prioridade. ....	78
Figura 24: 3º Caso – Comprimento Fixo p/ Voz e Exponencial. Negativa p/ Dados – Com Prioridade. ....	79
Figura 25: 3º Caso – Exponencial Negativa p/ Dados – Com Prioridade. ....	79
Figura 26: Visão Geral sobre aplicação e uso do NS (NS,2003). ....	83
Figura 27: Modelos de componentes básicos (NAM, 2004). ....	85
Figure 28: C++ e OTcl: A Dualidade (NS,2003) ....	86
Figure 29: Visão da Arquitetura do NS (NS, 2003) ....	87
Figura 30: NAM Editor com uma implementação de um cenário de simulação. ....	88
Figura 31: Animação de cenário de rede. ....	89
Figura 32: Descrição detalhada dos componentes da ferramenta de animação NAM. ....	90
Figura 33: Modelo de topologia aplicada à simulação. ....	91
Figura 34: Simulação NS / <i>TRACEGRAPH</i> – Exp. Neg. p/ Voz e Dados s/ Prioridade. ....	93
Figura 35: Simulação NS / <i>TRACEGRAPH</i> – Exp. Neg. p/ Voz e Compr. Fixo p/ Dados s/ Prioridade. ....	94
Figura 36: Simulação NS / <i>TRACEGRAPH</i> – Exp. Neg. p/ Dados e Compr. Fixo p/ Voz s/ Prioridade. ....	95
Figura 37: Simulação NS / <i>TRACEGRAPH</i> – Exp. Neg. p/ Voz com Prioridade. ....	96
Figura 38: Simulação NS / <i>TRACEGRAPH</i> – Exp. Neg. p/ Dados com Prioridade. ....	97
Figura 39: Simulação NS / <i>TRACEGRAPH</i> – Comprimento Fixo para Voz com Prioridade. ....	97
Figura 40: Simulação NS / <i>TRACEGRAPH</i> – Comprimento Fixo p/ Dados com Prioridade. ....	98

## **LISTA DE TABELAS**

Tabela 1: Principais VOCODERs e suas características. ....	27
Tabela 2: Vazão típica de Aplicações em Rede (MARTINS, 2001).....	52
Tabela 3: Variação de dados para geração de gráfico do modelo teórico – 1º Caso.....	74
Tabela 4: Variação de dados para geração de gráfico do modelo teórico – 2º Caso.....	76
Tabela 5: Variação de dados para geração de gráfico do modelo teórico – 3º Caso.....	78



## LISTA DE ABREVIATURAS E SIGLAS

CODEC:	Coder-Decoder
CoS:	Class of Service
DoS:	Denial of Service
DSL:	Digital Subscriber Line
FPC:	Flexible PIC Concentrator
HDLC:	High-level Data Link Control
HTTP:	Hypertext Transfer Protocol
IETF:	Internet Engineering Task Force
IP:	Internet Protocol
ITU:	International Telecommunications Union
LAN:	Local Area Network
MB:	Megabytes
Mb:	Megabit
MGCP:	Media Gateway Control Protocol
MTP:	Message Transfer Part
CS-ACELP:	Conjugate Structure Algebraic Code Excited Linear Prediction
PPP:	Point-to-Point Protocol
PSTN:	Public Switched Telephone Network
PBX:	Private Branch Exchange
PCM:	Pulse Code Modulation
LD-CELP:	Low-Delay Code Excited Linear Prediction
PVC:	Permanent Virtual Circuit
QoS :	Quality of Service

RED:	Random Early Detection
RFC:	Request for Comments
RSVP:	Resource Reservation Protocol
RTCP:	Real-time Transport Control Protocol
RTP:	Real-time Transport Protocol
SDP:	Session Description Protocol
SigTran:	Signaling Transport
SIP:	Session Initiation Protocol
SLA:	Service-Level Agreement
SMTP:	Simple Mail Transfer Protocol
SS7:	Signaling System Seven
SSP:	Signaling Switch Point
STP:	Signal Transfer Points
DNS:	Domain name service
TCP:	Transmission Control Protocol
UDP:	User Datagram Protocol
URL:	universal resource locator
VOCODER:	Voice coder-decoder
VoIP:	Voice over IP

## SUMÁRIO

INTRODUÇÃO.....	13
CAPÍTULO 1 – Estudos sobre VoIP.....	17
1.1 Voz sobre IP - Voip.....	17
1.2. Principais arquiteturas VoIP .....	20
1.3. CODECs.....	24
1.4. Componentes de VoIP.....	27
1.4.1. Gateways.....	28
1.4.2. Controladores de Gateways .....	29
1.5 Protocolos de Voz e suas aplicações.....	30
1.5.1 O protocolo de sinalização SS7 (Signaling System Seven).....	31
1.5.1.1. A pilha de protocolos do SS7 .....	33
1.5.2. Protocolo H323 .....	35
1.5.3. Protocolo de Transporte em tempo Real (Real-time Transport Protocol - RTP) ....	38
1.5.4. Protocolo de Controle de Transmissão em Tempo Real (Real-time Transport Control Protocol – RTCP) .....	39
1.5.5. Protocolo de Controle de Gateway (Media Gateway Control Protocol – MGCP)..	41
1.5.6. O Protocolo SIP (Session Initiation Protocol – Protocolo de Inicialização de Sessão) .....	42
1.5.6.1. Processo de execução do SIP .....	42
1.5.7. Transporte de Sinalização .....	44
1.5.8 Megaco/H.248.....	45
1.5.9. Protocolo de Reserva de Recurso (RSVP - Resource Reservation Protocol).....	46
1.6. QoS – Quality of Service (Qualidade de Serviço) .....	49
1.6.1. Princípios – Qualidade de Serviço (QoS).....	51
1.6.2. Vazão .....	51
1.6.3. Latência (Atraso) .....	52
1.6.4. Jitter .....	55
1.6.5. Perdas de pacotes .....	56
1.7. Alternativas técnicas para qualidade de serviço em redes IP.....	56
1.7.1 Um modelo Diffserv .....	58
CAPÍTULO 2 – DESENVOLVIMENTO DE UM MODELO TEÓRICO .....	61
2.1 Multiplexação de sinais de voz em redes IP .....	61
2.1 Avaliação de técnicas para redução de transmissão de pacotes VoIP.....	62
2.1.1. Redução do overhead em cabeçalhos de pacotes.....	63
2.1.2. Redução da carga de processamento dos dispositivos da rede .....	63
2.1.3 Reduzindo o consumo de recursos UDP.....	64
2.2 Princípios de multiplexação do fluxo de voz em redes IP .....	64
2.2.1 Esquema 1 – Encaminhamento por tempo.....	65
2.3. Atraso de fila .....	66
2.4. Desenvolvimento de um modelo teórico de filas.....	68
CAPÍTULO 3 – SIMULAÇÃO DO MODELO TEÓRICO .....	82

3.1 O NS (Network Simulator) .....	82
3.2 Topologia da simulação .....	90
3.3 Aplicação e comportamento do usuário de voz .....	91
3.4 Métricas de avaliação .....	92
3.5 Resultados Obtidos.....	92
CONCLUSÃO.....	99
REFERÊNCIAS .....	101
Anexo A – Script OTCL SEM PRIORIDADE .....	104
ANEXO B – SCRIPT OTCL COM PRIORIDADE .....	106

## INTRODUÇÃO

Desde 1830, a cada novo meio de comunicação adotado, uma rede distinta é estabelecida para torná-lo disponível a seus usuários. Sucessivamente tivemos o telefone, o telex, a comunicação de dados e a TV a cabo, cada um acompanhado por sua própria rede de serviços. Hoje é comum o usuário final possuir conexões separadas para as redes de telefonia, de dados e de TV a cabo.

Quando se fala da convergência na área de telecomunicações, se refere à redução para uma única conexão de rede, fornecendo todos os serviços, com conseqüente economia de escala (STANTON, 2000).

A convergência é um tema bastante discutido, desde dos anos 80, quando foi reconhecida pela primeira vez a importância crescente da comunicação entre computadores. Com a digitalização da rede de telefonia, a voz passou a ser transmitida como dados entre as centrais telefônicas, mantendo-se, porém a rede de terminais analógicos para os usuários finais. Nesta época já foi defendida a extensão do canal de voz digital até o usuário final, substituindo o antigo telefone analógico por um aparelho digital. Foi proposta a criação da Rede Digital de Serviço Integrados (RDSI, em inglês seria ISDN – Integrated Service Digital Network), que levaria ao usuário uma única conexão (digital), podendo ser usado indistintamente para voz (telefonia) e comunicação de dados em até 128 kilobits/s. Quando foi proposto, este serviço seria revolucionário, pois os modems usados na época eram tipicamente de 2400 bits/s. Porém, no Brasil ele demorou muito para chegar, somente passou a ser oferecido ao público no final dos anos 90, quando já existiam modems de 56 kbps, e alternativas ainda mais rápidas.

A segunda tentativa de promover a convergência veio ainda nos anos 80, associada à introdução de fibras óticas e serviços de banda larga. A fibra ótica possui uma capacidade de transmissão digital tão grande, que abriu a possibilidade de enviar por cabo novas aplicações antes impraticáveis, tais como televisão. Enxergava-se a perspectiva de realizar um novo nível de integração entre as redes de comunicação, e foi lançada a proposta de RDSI de Faixa Larga (RDSI-FL), baseada em Asynchronous Transmission Mode (ATM). Nesta proposta, seria criada uma rede mundial ATM, à qual todos os computadores estariam ligados, e através de uma única conexão. Seria realizado acesso a serviços de telefonia e televisão, além de dados. Para dar suporte à transmissão de voz e vídeo, o ATM previu suporte diferenciado para diferentes tipos de serviços, de qualidades e prioridades distintas.

Acredita-se que a convergência nos tempos atuais, entre as redes de serviços será realizada através do TCP/IP, estendido para tratar prioritariamente aplicações como voz e vídeo. A extensão mais promissora se chama Serviços Diferenciados, que permite uma classificação simples de aplicações, de acordo com seu grau de urgência ou importância, ou com o preço cobrado para sua transmissão. Nesta visão, a Internet deverá deixar de ser uma rede onde todos usuários são tratados de forma igual. Será este o preço a se pagar para tornar a rede única para todos os serviços diferentes. Devem aparecer novos serviços, talvez principalmente de entretenimento, pelos quais vai se pagar um preço maior do que pelo serviço tradicional de hoje (SITOLINO, 1999).

A tendência do futuro já está se evidenciando a nível mundial através das fusões entre empresas operadoras de serviços de telecomunicações (telefonia e TV a cabo), e entre fabricantes de equipamentos. Entre fabricantes de telefonia, há uma corrida generalizada para a tecnologia TCP/IP de comunicação de dados, geralmente através da aquisição de empresas especializadas. A consequência mais evidente é o lançamento nos últimos meses de

equipamentos de telefonia IP ou VoIP (Voz sobre IP), termo que caracteriza a comunicação por voz, com qualidade e recursos da telefonia, usando redes de dados para sua transmissão. Alguns casos incluem a aquisição da Bay Networks pela Nortel Networks, da Fore Systems pela Marconi, e da Xylan pela Alcatel (CALVO, 1999). Desconfia-se que não há um futuro muito promissor para um fabricante de equipamentos de telefonia que não acompanhe esta evolução.

Enfim, parece que a tecnologia vai permitir realizar a sonhada convergência para uma rede única. Porém isto só deve trazer benefícios para o usuário final se ele tiver a liberdade de escolher quem será seu provedor destes serviços integrados. Para esta pergunta não se pode postular ainda uma resposta (STANTON, 2000).

O objetivo do nosso estudo é fazer uma análise sobre o impacto do atraso dos pacotes de voz em um método de transmissão de sinais multiplexados. Para esta análise é proposto o desenvolvimento de um modelo matemático, com a aplicação de variados métodos de distribuição do tráfego, com e sem a atribuição de QoS.

Começaremos nosso estudo apresentando as características que compõem a estrutura de uma transmissão de voz sobre IP (VoIP), apresentados no Capítulo 1, descrevendo suas principais arquiteturas dos seus protocolos. Um tópico bastante controverso sobre redes de computadores: a QoS (*Quality of Service – Qualidade de Serviço*), examinando os recentes avanços que visam ao desenvolvimento da arquitetura de rede para fornecer suporte para as exigências de serviços de aplicações de VoIP.

Ainda no Capítulo 1, descreveremos o protocolo RTP (protocolo de tempo real), padrão emergente para a camada de aplicação, onde sua finalidade está voltada à controlar e enquadrar a transmissão de dados multimídia. Alternativas técnicas para qualidades de serviço em redes IP, arquiteturas tais como *Intserv* e *Diffserv*, padrões Internet sensíveis à QoS.

Também no mesmo capítulo examinaremos o protocolo de sinalização RSVP, de grande importância para as arquiteturas *Intserv* e *Diffserv*.

No Capítulo 2, com base na proposta descrita em Yamada (2001), onde são discutidas formas de encaminhamento de pacotes RTP, um estudo e a elaboração de um modelo teórico para verificação do atraso de pacotes nas filas, comparando seus resultados com variados métodos de distribuição de chegada de pacotes.

Finalmente no Capítulo 3, uma simulação de implementação do modelo teórico, apresentado no Capítulo 2, será definida para fins comparativos entre os resultados, bem como para verificação dos cálculos apurados.



# **CAPÍTULO 1 – ESTUDOS SOBRE VOIP**

## **1.1 Voz sobre IP - VoIP**

Normalmente abreviada como VoIP, do inglês Voice over IP, esta tecnologia permite a digitalização e codificação da voz e o encaminhamento de pacotes de dados IP (Internet Protocol – Protocolo Internet) para a transmissão em uma rede que utilize TCP/IP (Transmission Control Protocol/ Internet Protocol – Protocolo de Controle de Transmissão / Protocolo Internet).

Uma das grandes vantagens da utilização da VoIP é a comutação de pacotes, ao contrário da tradicional comutação por circuitos empregada pela telefonia convencional.

Em um circuito telefônico tradicional, é estabelecida uma conexão física que permanece ativa durante toda a chamada, mesmo que não haja conversação. Isso torna impossível que outros usuários usem esta conexão durante os momentos vagos. Já na telefonia com comutação por pacotes, não existe um circuito dedicado entre dois pontos, os dados de voz são convertidos em pacotes à medida que vão sendo criados e são transmitidos pelo caminho disponível na rede naquele instante. Os pacotes podem seguir caminhos diferentes pela rede, mas todos chegarão ao destino pelo caminho livre naquele instante. A Figura 1 ilustra os modelos de comutação de circuitos e comutação por pacotes, descrevendo as diferenças básicas entre uma rede tradicional de circuitos e uma rede de comutação por pacotes (PENNO, 1999).

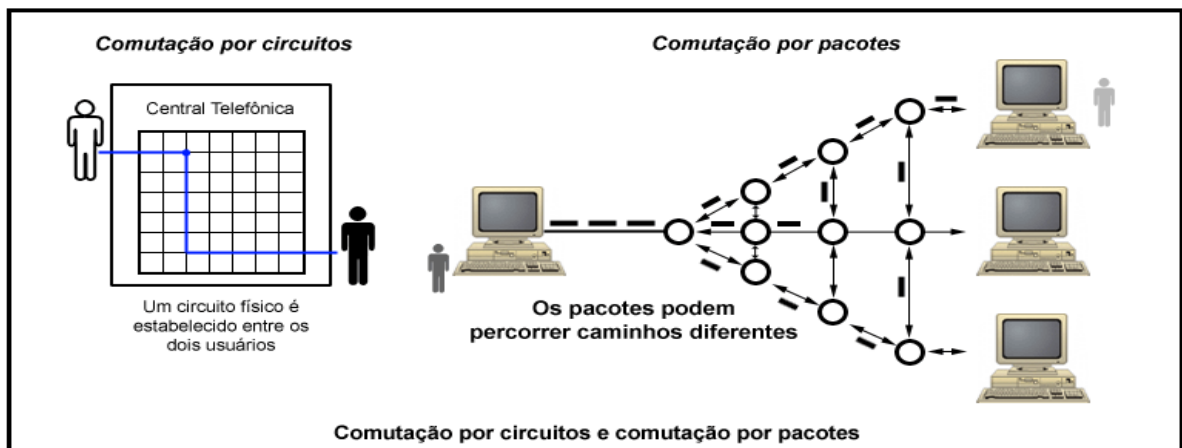


Figura 1: Comutação por circuitos e comutação por pacotes (PENNO, 1999)

Com o crescimento da Internet e das redes corporativas de comunicação de dados, a utilização da VoIP tem se mostrado uma grande vantagem. Ter uma rede de dados separada da rede de voz tornou-se um fator oneroso e de difícil administração. Como explicado anteriormente, dados e voz ocupam o mesmo meio de transmissão do pacote durante somente a transmissão do pacote, não havendo a necessidade de alocar definitivamente uma determinada banda no link de dados para usar a comunicação de voz. Nas empresas, o uso da VoIP em aplicações corporativas tornou-se um fator para redução de custos. Como exemplo disso, uma empresa usando um link de 128 Kbit/s para interligar suas filiais, compartilhando a banda disponível com voz e dados, não será necessário o gasto com ligações interurbanas para comunicação com as filiais, pois o custo do link já está pago, ocupando-se ou não a banda.

Como toda tecnologia, para a VoIP também existem algumas desvantagens que devem ser observadas. Uma delas é o compartilhamento de banda, que pode levar a uma redução da banda reservada aos dados, repercutindo em uma menor velocidade para os *download's*. Um exemplo disso, em um link de 64 Kbit/s, enquanto houver ligações de VoIP, uma parte dessa largura de banda estará sendo utilizada pela voz (10 Kbit/s, dependendo do *codec* utilizado, conforme discutiremos posteriormente), o que pode deixar a rede mais lenta neste período.

Pode-se dizer que a arquitetura de VoIP ainda está em desenvolvimento e não existe uma padronização de protocolos e equipamentos como na telefonia fixa, o que pode levar a algumas dificuldades, mas que tendem a serem resolvidas em breve.

Uma última desvantagem observada são considerações sobre a garantia da qualidade de voz em rede de dados, ou seja, a ligação pode ser bem pior do ponto de vista da qualidade de voz. Isto acontece por que as redes de dados não fazem diferença entre um pacote que contém dados de voz e um pacote de contenha texto de um *e-mail*, por exemplo. A questão de qualidade é um fator que necessita de um estudo mais abrangente, o que será feito no próximo tópico.

Discutiremos alguns pontos críticos para transferências de pacotes de voz pela Internet. Sabemos que o protocolo IP hoje fornece um **serviço de melhor esforço** para todos os datagramas que transporta, ou seja, a Internet faz seu melhor esforço para transportar cada datagrama do remetente ao receptor o mais rápido possível. Mas, as questões de atrasos fim a fim para um pacote individual não são garantidas pelo serviço de melhor esforço.

Na verdade, voz em tempo real impõe rígidas limitações ao **atraso de pacote** e à **variação de atraso**<sup>1</sup> da mesma corrente de pacotes. Voz em tempo real funciona bem em transmissões onde a oferta de largura de banda é abundante e, portanto, o atraso e a variação do atraso são mínimos. Questões que avaliam o atraso e a variação de atraso serão discutidas com mais detalhes nos próximos tópicos.

Existe ainda um grande esforço por parte dos principais fornecedores de equipamentos, incluindo Cisco, 3Com, dentre outros, em promover o uso do protocolo ITU-T H.323, padrão de transmissão de multimídia (voz e vídeo) sobre IP pela Internet. Além do IP,

---

<sup>1</sup> Variação de Atraso é a variabilidade dos atrasos de pacote dentro de uma mesma seqüência de pacotes.

VoIP usa o protocolo real-time ( RTP ) para assegurar-se de que os pacotes comecem a ser entregues de maneira oportuna. Usando redes públicas, atualmente é difícil garantir a qualidade do serviço (QoS). O melhor serviço é possível com as redes privadas controladas por uma empresa ou por um fornecedor de serviço de telefonia, pois em redes públicas como a Internet, as aplicações de serviços sensíveis ao atraso não sofrem qualquer tipo de prioridade nas filas dos roteadores, pois são adotadas posturas igualitárias para o tratamento dos pacotes. Enquanto isso deve contentar com o serviço de melhor esforço.

Nas próximas seções apresentaremos as principais arquiteturas aplicadas ao VoIP, bem como seus *codecs*, métodos de codificações estabelecidos pelas padronizações ITU.

## 1.2. Principais arquiteturas VoIP

Existem basicamente três formas de utilizar as redes IP para transmitir voz: PC para PC, PC para Telefone e Telefone para Telefone.

A forma de PC para PC, como mostrado na Figura 2, descreve um cenário simples de rede com os componentes que definem a arquitetura proposta para a Internet, com a finalidade de fornecer suporte de QoS de aplicações de transmissão de voz. Suponha que dois fluxos de pacotes de aplicações se originem nos *Host* (hospedeiro) 1 e *Host* 2 em uma LAN (*Local Area Network* – Rede Local) e sejam destinados aos *Host* 3 e *Host* 4 em uma outra LAN. Os roteadores das duas LANs são ligados por um enlace de 1.5 Mbps (Mega Bits por Segundo). Vamos supor que as taxas das LANs sejam significativamente mais altas do que 1.5 Mbps e vamos verificar a fila de saída do Router 1; é aqui que vão ocorrer atrasos e perdas de pacotes se a taxa de envio do *Host* 1 e *Host* 2 exceder 1.5 Mbps. A análise deste e outros cenários nos dará avaliações sobre os princípios ao fornecimento de QoS às aplicações de Voz sobre IP. Assuntos discutidos nas próximas seções.

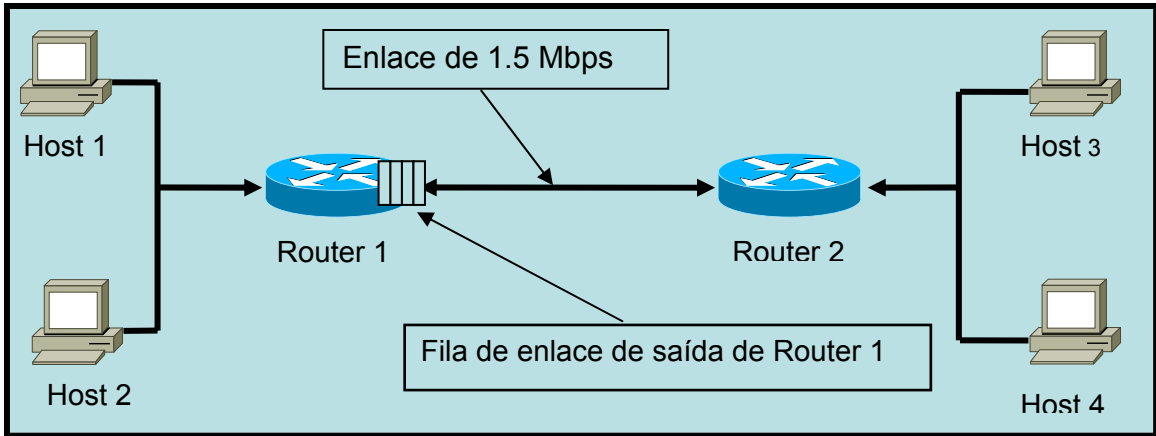


Figura 2: Arquitetura PC para PC.

Na arquitetura de PC para Telefone, um computador configurado conforme descrito anteriormente pode se comunicar com um telefone comum, conectado à rede telefônica tradicional, através de um *gateway*. *Gateways* são dispositivos que farão a conversação entre os pacotes recebidos pela rede IP para um sinal analógico que pode ser entendido e transmitido pela rede de telefonia convencional, conforme apresentado na Figura 3 abaixo.

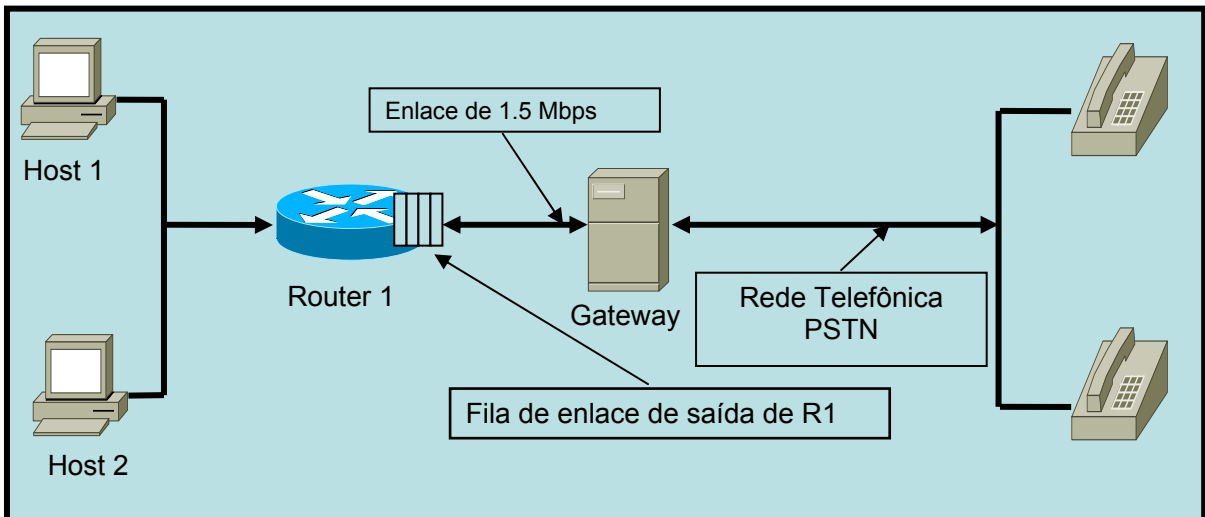


Figura 3: Arquitetura PC para telefone

Existem diferenças significativas entre a Internet e a PSTN (*Public Switched Telephone Network* - Rede Telefônica Pública Comutada), sendo a diferença técnica primária, suas arquiteturas de comutação (SKENE, 2000).

A PSTN é uma rede comutada a circuito que foi otimizada para comunicação em tempo real ou de voz síncrona com Qualidade de Serviço (QoS) garantida. Quando uma chamada telefônica é iniciada, um circuito é estabelecido entre a parte que origina a ligação e a parte chamada. A PSTN garante qualidade desse serviço por dedicar um circuito *full-duplex* de 64k entre as partes de uma conversa telefônica. Independente se as partes estão conversando ou em silêncio, eles estão usando um circuito de 64k dedicado até que a ligação termine. Desde que a largura de banda seja constante, os custos de uma chamada telefônica na PSTN são baseados na distância e no tempo.

Por outro lado, a Internet é uma rede comutada a pacotes que tem sido historicamente usada por aplicações onde uma QoS variável é tolerável, tais como e-mail ou transferência de arquivos. Redes comutadas a pacotes compartilham os recursos entre diversos usuários que desejem transmitir. Assim, por não alocarem um caminho dedicado entre remetente e destinatário, não podem garantir QoS. Além disso, o mecanismo de custo para o uso da Internet não é a distância ou o tempo, mas o uso da largura de banda.

As diferenças na arquitetura de comutação produzem outras significantes diferenças técnicas. A Internet usa roteamento dinâmico, baseado em um endereçamento não geográfico, enquanto a PSTN usa comutações estáticas, baseadas na numeração geográfica do telefone. Além disso, a "inteligência" da Internet é descentralizada, ou distribuída, enquanto na PSTN uma "inteligência" reside em pontos centrais na rede (GOOGMAN, 1999).

Por fim, como um resultado das diferentes arquiteturas, existe alguma diferença cultural. A PSTN contém determinações que se encarregam da tarifação durante a troca do tráfego, enquanto a Internet é, em sua maior parte, isenta de taxas sensíveis ao uso.

Na Figura 4, a arquitetura de comunicação de Telefone para Telefone utilizando uma rede IP, é o que tende a ser mais utilizado no futuro, pois ele não depende que os dois usuários utilizem o computador, além de empregar toda a infra-estrutura das redes já existentes. Neste caso a chamada começa como uma chamada telefônica convencional, em seguida o *gateway* faz a conversão para os pacotes que serão utilizados e transmitidos pela rede IP. Do outro lado, a conversão é realizada pela rede telefônica convencional até o outro usuário.

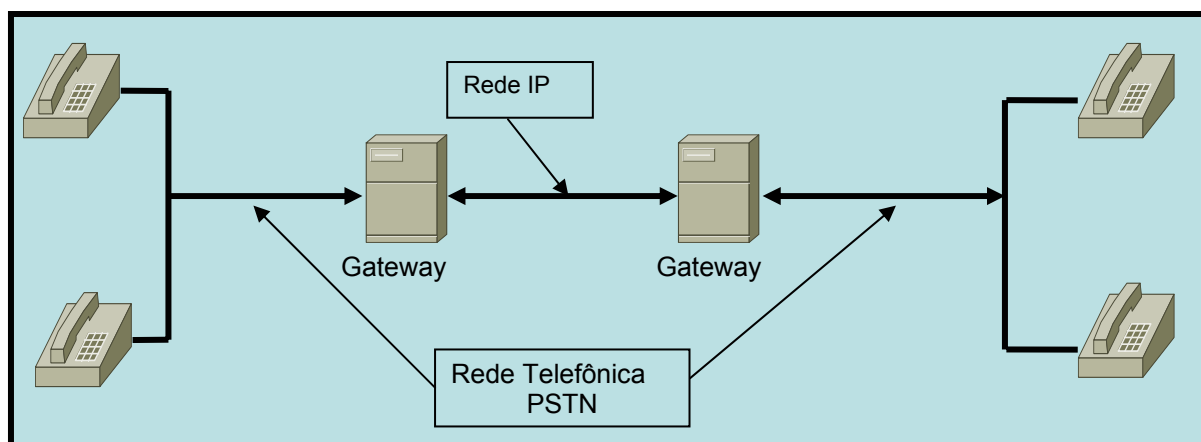


Figura 4: Arquitetura telefone para telefone.

A transmissão de voz codificada em uma rede com o protocolo IP, possui característica peculiar a este ambiente, e importante a se considerar. Como o IP, por si só, não oferece nenhuma garantia de Qualidade de Serviço, categorizado como tráfego de melhor esforço (*best-effort*) como descrito nas seções anteriores, outros protocolos e soluções complementares devem ser agregados na formação da solução final, para permitirem um resultado comparável com o observado na rede de voz convencional (LEFELHOCZ, 1996). A busca por tais protocolos e soluções complementares, nos últimos anos, têm sido o foco de extensivas pesquisas por grupos do IETF, como é a proposta do Grupo de Pesquisa de Serviços Integrados (IntServ) e do Grupo de Pesquisa de Serviços Diferenciados (DiffServ).

### 1.3. CODECS

Os *codecs* são equipamentos responsáveis pela conversão de sinais analógicos de voz em sinais digitais e vice-versa. Os *codecs* são empregados na telefonia convencional, mas a sua utilização na rede de pacotes é também necessária.

A comunicação de voz é analógica, na rede de dados o formato é analógico. O processo de conversão do formato analógico para o formato digital é feito com um codificador-decodificador (CODEC – também conhecido como codificador-decodificador de voz VCODER). Existem várias formas de um sinal de voz analógico ser convertido, e todos seguem variadas padronizações. O processo de implementação de um CODEC é bastante complexo e não discutiremos esse processo em nosso estudo. Basta dizer que a maioria das conversões baseia-se em modulação por código de pulso (PCM – *pulse coded modulation*) ou variações. Cada esquema de codificação tem sua própria história e o mérito, junto com sua largura de banda necessária.

Descrevemos abaixo como um sinal analógico de áudio variável (voz ou música) é convertido em um sinal digital:

1 – O sinal analógico de áudio é primeiramente amostrado a alguma taxa fixa; por exemplo, 8000 amostras por segundo. O valor de cada amostra é um número real arbitrário.

2 – Cada uma das amostras é então “arredondada” para um valor qualquer dentre um número finito de valores. Essa operação é chamada de “quantização”. O número de valores finitos – chamado de valores de quantização – é tipicamente uma potência de 2, por exemplo 256 valores de quantização.



3 – Cada um dos valores de quantização é representado por um número fixo de bits. Por exemplo, se houver 256 valores de quantização, então cada valor – e, portanto, cada amostra – será representado por um byte. Cada uma das amostras é convertida à sua representação de bits. As representações por bits de todas as amostras são, então, concatenadas em conjunto para formatar a representação digitalizada do sinal.

Como exemplo, se um sinal de áudio for amostrado a uma taxa de 8000 amostras por segundo e cada amostra for “quantizada” e representada por 8 bits, então o sinal digital resultante terá uma taxa de 64000 bits por segundo. Esse sinal digital pode então ser reconvertido – isto é, decodificado – em um sinal analógico para ser reproduzido. Contudo, o sinal analógico decodificado é, em geral, diferente do sinal de áudio original. Aumentando a taxa de amostragem e o número de valores da quantização, o sinal decodificado pode se aproximar (e até ficar exatamente igual) ao sinal analógico original. Assim, há uma clara permuta entre a qualidade do sinal decodificado e as exigências de largura de banda do sinal digital.

Além de executar a conversão de sinal analógico para digital, os CODECs comprimem a cadeia de dados, e fornecem a correção de eco (PENNO, 1999). A compressão representada pelo formato de áudio pode implementar recursos para economias da banda. A economia de banda para serviços de voz pode vir em diversos formatos e trabalhar em níveis diferentes. Por exemplo, a compressão analógica pode ser parte do esquema de codificação (algoritmo) e não necessita uma compressão digital adicional das camadas altas de aplicação. Uma outra maneira de conservar a largura de banda é o uso da supressão do silêncio, que é o processo de não emitir pacotes da voz durante o silêncio de uma conversação humana. Usar a supressão da compressão e/ou do silêncio pode resultar em economias significativas de largura de banda. Entretanto, há algumas aplicações que poderiam ser afetadas pela

compressão. Um exemplo disso é o impacto para usuários de modem. Os esquemas da compressão podem interferir com o funcionamento do modem confundindo o codificador utilizado. O resultado poderia ser o modem nunca sincronizaria ou apresentaria uma variação de tráfego (throughput) muito baixo, afetando assim sua performance.

O codec de voz torna possível fornecer alta qualidade de áudio de acordo com a banda requerida. Recentemente, o ITU-T (*International Telecommunication Union – Telecom* - órgão que define normas e regras, bem como propostas de produção de novos serviços em uma convergência de redes), especificou o G.729 e o G.723, padrões codec de voz com a finalidade de fazer melhor uso da banda requerida (H.245, 2000). Por exemplo, em G.729 o tamanho dos blocos codificados é geralmente de 10 bytes e o intervalo do quadro é de 10 ms. Os pacotes IP são constituídos somando o cabeçalho IP/UDP/RTP (visto posteriormente) para o pacote de voz codificado do *codec* e os pacotes são transportados sobre uma rede IP para o lado do receptor.

O melhor codec a ser utilizado depende da relação desejada entre qualidade de voz e largura de banda disponível. Ao utilizar um codec de 64 Kbit/s, tem-se uma boa qualidade de voz e pouco atraso, mas estará utilizando uma grande largura de banda. A avaliação da qualidade de voz é algo subjetivo, ou seja, o que é uma boa qualidade para um usuário pode não ser tão boa para outro, e é o que o indicador de qualidade MOS (*Mean Opinion Score* – Nota média de opiniões) testa entre usuários. Um teste é realizado com um grupo de usuários, onde é levado em conta o nível de percepção que cada um tem da qualidade de voz em uma chamada para determinado codec. Estes resultados são analisados de uma maneira comparativa, a fim de determinar qual codec possibilita uma melhor qualidade de voz em relação a outro.

A solução para os CODECs é uma seqüência de dados transmitidos através de pacotes IP e transportados até seu destino final. Estes destinos devem usar os mesmos padrões de CODECs do remetente. O resultado de usar padrões ou parâmetros diferentes em ambas as extremidades é uma comunicação incompreensível. A Tabela 1 lista alguns dos padrões de CODECs mais importantes descritos pela união de telecomunicações internacional (ITU). De acordo com a RFC 1890 (1996), onde são discutidas as configurações RTP (Real-time Transfer Protocol – Protocolo de Transferência em Tempo Real) para comunicações de áudio, os seguintes VOCODERS de áudio lustrados pela Tabela 1 utilizam frames em sua geração.

Tabela 1: Principais VOCODERS e suas características.

Padrão ITU	Descrição	Bandwidth (Kbps)	Atraso na Compressão (ms)
G.711	PCM	64	< 1.00
G.721	ADPCM	32,16,24,40	< 1.00
G.728	LD-CELP	16	~2.50
G.729	CS-ACELP	8	~15.00
G.723.1	Multirate CELP	6.3, 5.3	~30.00
GSM	RPE/LTP	13	~20.00

#### 1.4. Componentes de VoIP

Os principais componentes de uma rede de VoIP, de alguma forma são similares aos componentes que formam uma rede de circuito compartilhado. As redes de VoIP devem executar as mesmas tarefas que o PSTN, além de executar uma função de gateway para uma rede pública existente. Embora usando diferentes tecnologias, alguns componentes do PSTN também são usados para definir as redes de VoIP. Podemos descrever os principais elementos de uma rede VoIP em três partes principais:

- Gateways

- *Gateways* / Controladores de *Gateways*

- Redes IP

### 1.4.1. Gateways

Os *Gateways* são responsáveis pelas origens da chamada, detecção de chamada, a conversão da voz de analógico para digital, e a criação de pacotes da voz (funções do CODEC). Além disso, os *Gateways* têm características opcionais, tais como a compressão (analógica e/ou digital) da voz, o cancelamento do eco, a supressão do silêncio, e o recolhimento dos dados estatísticos.

Os *Gateways* atribuem a forma ao padrão de voz de maneira que possa ser transportado sobre a rede IP. Os *Gateways* dão origem ao tráfego. Tipicamente, cada conversação (chamada) é uma única sessão IP, transportada por um protocolo de transporte em (RTP) sobre o UDP.

Os *Gateways* existem em diversas formas. Por exemplo: os *Gateways* podem ser equipamentos de telecomunicação dedicada, ou mesmo um PC genérico que executa o software de VoIP. Podemos citar algumas características e serviços:

- *Gateway* de entroncamento conectado entre a rede de telefone e uma rede de VoIP. Tais passagens controlam tipicamente um grande número de circuitos digitais;
- *Gateways* residenciais que fornecem uma interligação da tradicional rede analógica a uma rede de VoIP. Os exemplos de *Gateways* residenciais incluem *cable modem*, dispositivos xDSL, e dispositivos *wireless broadband*;

- Os *Gateways* fornecem uma interligação analógica ou digital tradicional do PBX a uma rede de VoIP. Os exemplos incluem (empresa) *Gateways* VoIP de baixa escala;
- *Gateways* empresariais que fornecem uma interface entre o tradicional PBX digital ou uma interface simples integrada ao PBX a uma rede de VoIP;
- Unidades de telefone IP.

### **1.4.2. Controladores de *Gateways***

Os Controladores de *Gateways* abrigam os serviços de sinalização e de controle que coordenam as funções do *Gateway*. Os controladores de *Gateway* podem ser considerados similares ao *gatekeepers* do protocolo H.323 (como veremos posteriormente). O controlador de *Gateway* tem a responsabilidade para alguma ou toda coordenação de sinalização da chamada, identificação do número de telefone, identificar o host, gerenciar recursos, e sinalizar serviços de acesso ao PSTN (*Gateway* SS7 – descrito posteriormente). Conforme apresentado na Figura 5, podemos verificar todos os componentes de uma rede VoIP (ITU-T, 1998).

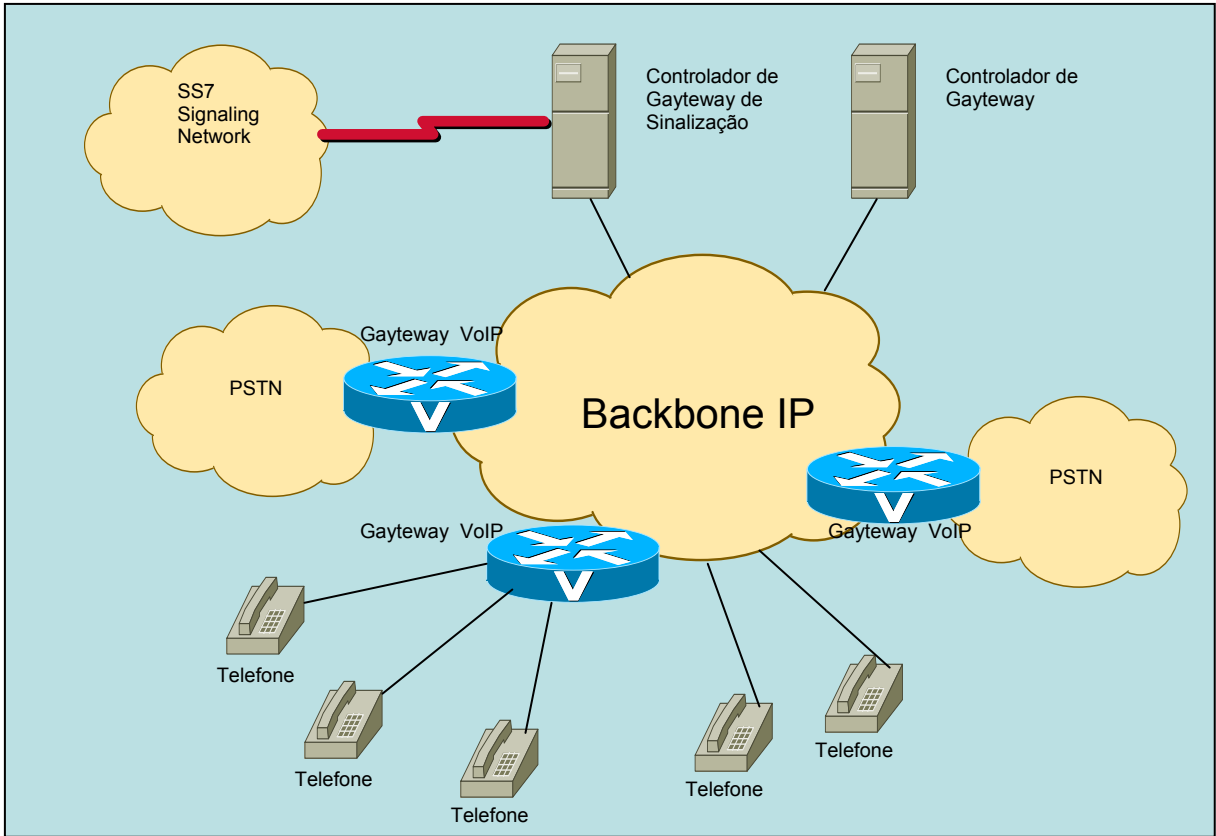


Figura 5: Apresentação completa de uma rede VoIP.

## 1.5 Protocolos de Voz e suas aplicações

Existe uma variedade de produtos VoIP e implementações com características variadas. Tais características são definidas por dois grupos que ditam as padronizações empregadas em redes baseadas em entrega de pacotes multimídia: a ITU e *Internet Engineering Task Force* (IETF – Força Tarefa de Engenharia Internet).

Algumas das implementações são focalizadas nas especificações de ITU mais do que dos padrões do IETF. Também, por causa da sobreposição e do co-desenvolvimento em muitos dos padrões, existem implementações de ambos os grupos. Ainda, alguns fabricantes aplicam características proprietárias que aderem aos padrões ou adicionam novas funcionalidades, totalmente dependente do produto.

Entretanto, nem toda padronização é definida por um grupo ou outro; muitos dos padrões são definidos em ambos os grupos, com o objetivo de resolver os mesmos problemas. O resultado é alguma sobreposição da funcionalidade, bem como as diferenças na formatação e na nomenclatura (GOYAL, 1999). Existem diversos protocolos VoIP, como descrito nas seções seguintes. Nem todos os protocolos são usados em um grupo de produto específico. Os fabricantes codificam seus produtos com o que é o mais aplicável à sua atividade. Cada um destes protocolos tem suas próprias vantagens e desvantagens. Os protocolos apresentados nesta seção não são avaliados profundamente por não ser objeto do nosso estudo.

### **1.5.1 O protocolo de sinalização SS7 (*Signaling System Seven*)**

A rede Sistema de Sinalização 7 (SS7) é uma rede digital dedicada usada para inicialização e controle de chamadas. Cada ponto de sinalização na rede SS7 é unicamente identificado por código numérico de ponto. Códigos de ponto são carregados em mensagens de sinalização trocadas entre pontos de sinalização para identificar a fonte e o destino de cada mensagem (BOLOT, 1993). O ponto de sinalização utiliza uma tabela de roteamento para selecionar o caminho apropriado para cada mensagem.

Existem 3 tipos de pontos de sinalização nas redes SS7 (mostrados na Figura 6):

- Pontos de serviço de comutação (SSPs);
- Pontos de transferência de sinal (STPs);
- Pontos de controle de serviço (SCPs).

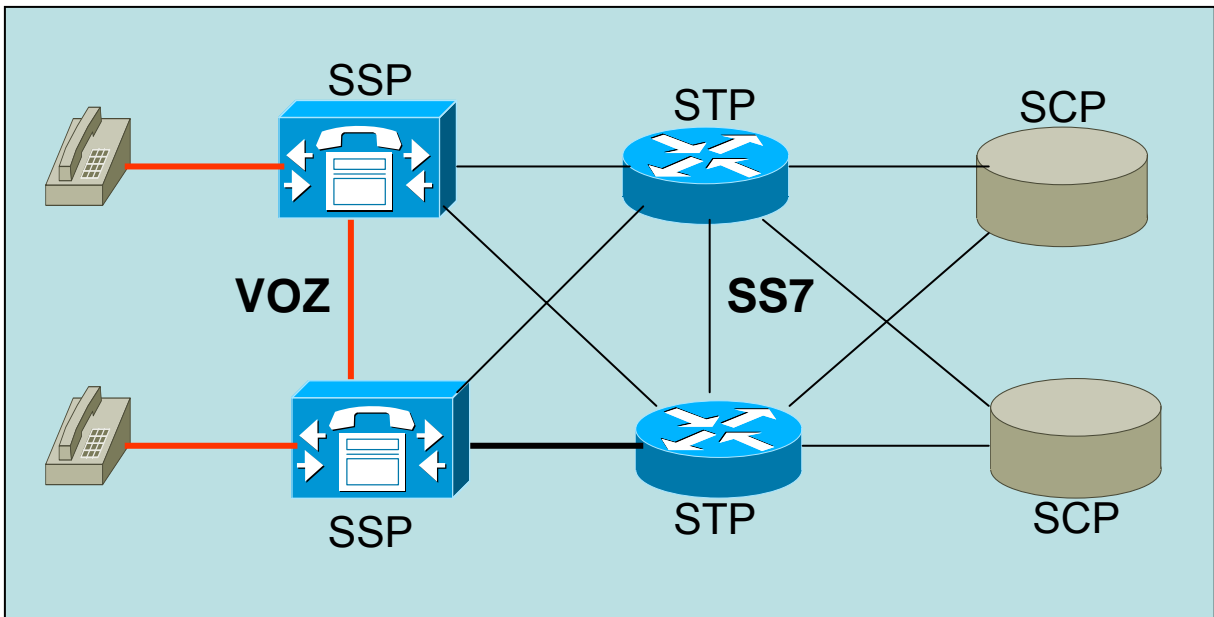


Figura 6: Rede SS7 (CISCO, 2003)

Os SSPs são comutadores que originam, terminam ou enfileiram chamadas. Um SSP envia mensagens de inicialização de chamadas para outros SSPs. Um SSP pode enviar uma mensagem para uma base de dados centralizada (SCP) para determinar como rotear uma chamada (ex. Chamada (0800) *toll-free*). Uma SCP envia uma resposta para o SSP originador contendo os números de roteamento associados ao número discado. Um número de roteamento alternativo pode ser usado pelo SSP caso o número primário esteja ocupado, ou não haja resposta dentro de um período de tempo limitado. As possibilidades para as chamadas vão variar de rede para rede e de serviço para serviço.

O tráfego de rede entre os pontos de sinalização pode ser roteado através de uma STP. Um STP roteia cada mensagem que chega para um enlace de sinalização de saída, baseado na informação de roteamento contida na mensagem SS7 (CISCO, 2003). Devido ao fato de atuar como um concentrador de rede, um STP pode fornecer melhor utilização da rede SS7.



SCPs e STPs são geralmente implantados em pares para assegurar que o serviço na rede permaneça em funcionamento, mesmo que ocorra um evento de falha de comunicação. Enlaces entre pontos de sinalização são fornecidos também em pares. O Tráfego é dividido por todos os enlaces do conjunto. O protocolo SS7 fornece tanto correção de erro quanto retransmissão para permitir a continuidade do serviço na ocorrência de falha nos enlaces dos pontos de sinalização.

O tráfego de voz é carregado nos circuitos de voz. O tráfego SS7 é enviado em enlaces de sinalização bidirecionais a taxa de 56 kb/s ou 64 kb/s. Devido à separação das redes de sinalização e voz, a sinalização pode ocorrer durante uma chamada sem afetar o tráfego de voz. Isto é chamado sinalização fora de banda, que permite com relação a sinalização dentro de banda:

- Inicialização mais rápida;
- Uso mais eficiente dos canais de voz;
- Suporte a serviços que requerem sinalização durante a chamada;
- Melhor controle sobre uso fraudulento da rede.

O SS7 é um padrão internacional definido pelo ITU-T. A definição do ITU-T para o SS7 permite variantes nacionais tais como os padrões ANSI e Bellcore utilizados nos Estados Unidos.

#### **1.5.1.1. A pilha de protocolos do SS7**

As funções de *hardware* e *software* do protocolo SS7 são divididas em abstrações funcionais chamadas níveis. Estes níveis são análogos ao modelo de referência de 7 camadas OSI da ISO, como é mostrado na Figura 7.

A parte de transferência de mensagens (MTP) é dividida em três níveis. O nível mais baixo é equivalente à camada física do modelo OSI. O MTP nível 1 define as características físicas, elétricas e funcionais da sinalização digital do enlace. Entre as interfaces físicas definidas, estão incluídas E-1 (2048 kb/s), DS-1 (1544 kb/s), V.35(64 kb/s), DS-0 (64 kb/s) e DS-0A (56 kb/s).

O MTP nível 2 assegura a transmissão sobre um enlace de sinalização. Quando um erro ocorre no enlace, a mensagem é retransmitida. O MTP nível 3 fornece as funções de roteamento e gerenciamento da rede. As mensagens são roteadas baseadas em um rótulo de roteamento presentes nas mensagens envidas no nível 2.

O controle de sinalização de conexão (SCCP) fornece serviços de rede com e sem conexão e capacidade de tradução de título global (GTT) sobre o MTP nível 3. Um título global é um endereço o qual é traduzido pelo SCCP para um código de ponto destino e número de subsistema (SSN=0,...,255). O SCCP é utilizado como camada de transporte para serviços TCAP incluindo (0800) e serviços de chamada por cartão. Cada serviço tem um único SSN que identifica o usuário do SCCP no ponto de sinalização de destino.

A parte de aplicações com capacidades de transação (TCAP) habilita a implantação de serviços de redes inteligentes através da troca de informações entre pontos de sinalização utilizando serviço SCCP sem conexão. Um SSP utiliza uma TCAP para buscar um SCP que determine um número de roteamento associado com o número discado. O SCP utiliza TCAP para retornar a resposta contendo o número de roteamento (ou um erro) de volta para o SSP.

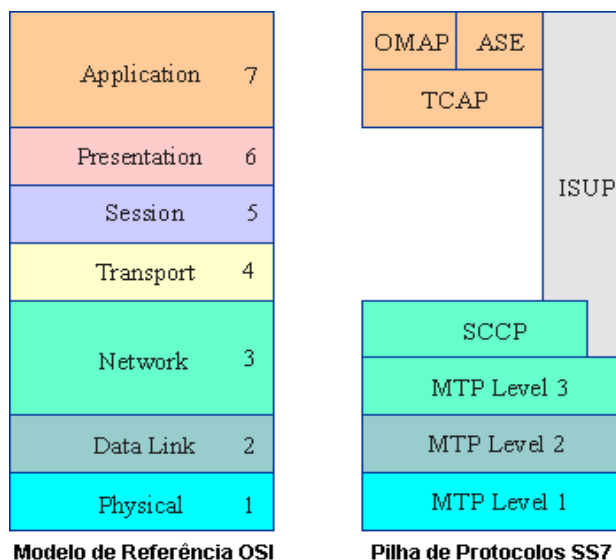


Figura 7: Pilha de protocolos SS7 vs. Modelo de Referência OSI

A última parte é a parte ISDN do usuário (ISUP). Mensagens ISUP são utilizadas para controlar a inicialização e a liberação de circuitos que carregam voz e dados entre duas partes (a chamadora e a respondedora). Chamadas que originam e terminam no mesmo comutador não utilizam sinalização ISUP.

Portanto, o SS7 é um protocolo de sinalização, a princípio, para suportar tráfego de voz (podendo se estender para dados), tem as suas funções divididas em 6 níveis de abstração denominadas níveis.

### 1.5.2. Protocolo H.323

A recomendação H.323 da ITU é um sistema de comunicação multimídia baseado em um conjunto de especificações. Estas especificações definem várias funções de sinalização, bem como os formatos dos meios relacionados a serviços empacotados de áudio e vídeo. Os padrões H.323 foram às primeiras padronizações que classificaram e propuseram a

classificação e solução de entrega dos pacotes multimídia sobre tecnologias LAN (TOGA, 1998).

Entretanto, enquanto a rede IP e a Internet se tornaram cada vez mais presentes, muitas tecnologias e padronizações e protocolos Internet (padrão RFC) foram desenvolvidos e baseados em algumas das idéias H.323.

Hoje há uma cooperação entre o ITU e IETF para resolver os problemas existentes, mas é justo dizer que o processo do RFC de promover os padrões teve um sucesso maior do que as H.323. As redes H.323 consistem em *gateways* (dos meios) e *gatekeepers*. Os *gateways* servem como o ponto final da terminação H.323 e conectam-se com as redes non-H.323, tais como a função de PSTN. Gatekeepers funcionam como uma unidade central para o controle de recepção da chamadas, gerenciamento de largura de banda e sinalização de chamadas.

Um *gatekeeper* e todos os seus *gateways* gerenciados dão forma a uma zona H.323. Embora o *gatekeeper* não seja um elemento requerido em H.323, este pode ajudar às redes H.323 a escalar para um tamanho maior, separando a chamada para controlar as funções da gerência dos *gateways*. As especificações H.323 tendem a ser mais pesadas e com um foco inicial em LAN's. Estes padrões têm baixo retorno em escalonabilidade, especialmente em desenvolvimentos em larga escala. Uma das introduções de escalonabilidade H.323 é sua dependência em sinalização baseada em TCP (orientado a conexão).

Há um desafio em manter um grande número sessões do TCP por causa do *overhead* envolvido. As versões subseqüentes de H.323 têm um foco em resolver alguns destes problemas. Com cada chamada que é iniciada, uma sessão do TCP (protocolo H.225.0) é criada usando um encapsulamento de um subconjunto das mensagens Q.931. Esta conexão

do TCP é mantida para a duração da chamada. Uma segunda sessão é estabelecida usando o protocolo H.245. Este processo de conexão baseada em TCP é para a troca das potencialidades, determinação master-slave, e o estabelecimento e a liberação do meio físico. Este grupo de procedimentos é um adicional aos processos H.225.0. A qualidade de serviço H.323 está na escolha do mecanismo de entrega, que é o protocolo de reserva do recurso (RSVP). Este protocolo não possui boas propriedades de escalonamento devido a seus gerenciamentos de fluxos de tráfego individuais da aplicação (H.245, 2000).

Embora o H.323 não detenha um espaço considerável entre os fornecedores de serviços, ele está muito bem posicionado na utilização de distribuição de aplicações de VoIP empresariais. Na Figura 8, são descritos todos os processos de uma comunicação com elementos VoIP baseados em H.323.

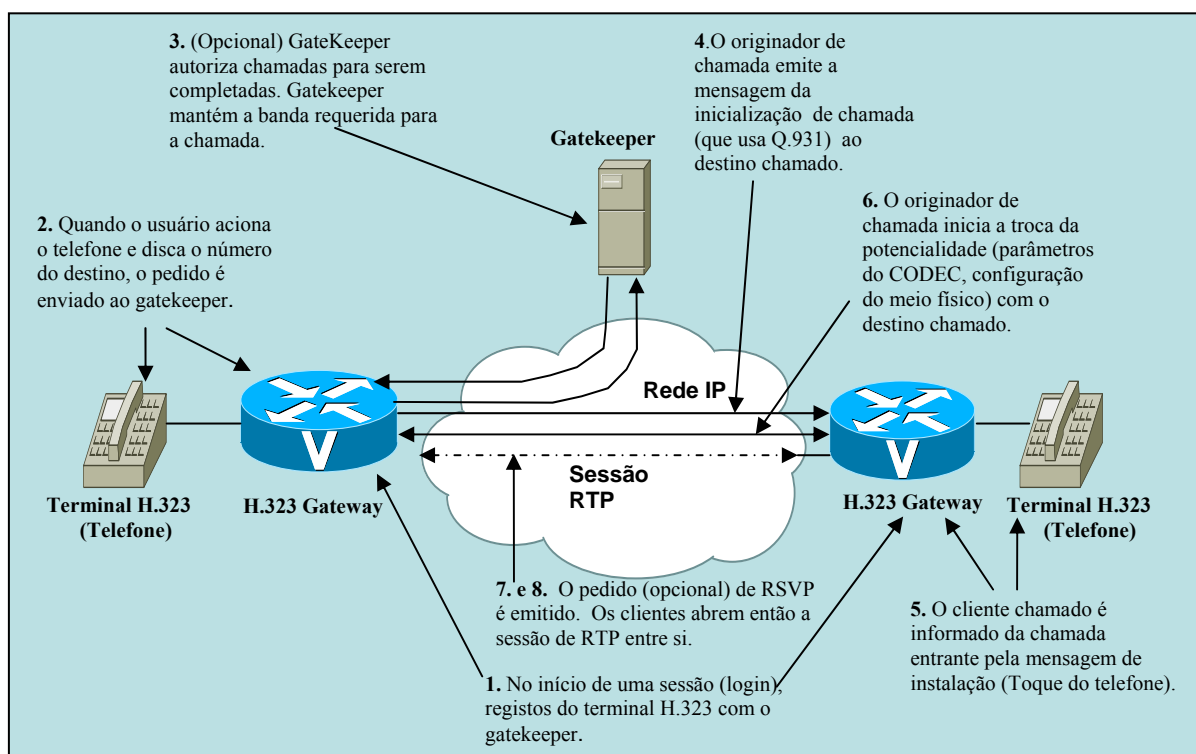


Figura 8: Processos de comunicação de elementos VoIP baseados em H.323.

### 1.5.3. Protocolo de Transporte em tempo Real (*Real-time Transport Protocol* - RTP)

As RFC 1889 (1996) e RFC 1890 (1996) descrevem as características do RTP, que fornece serviços de entrega fim-a-fim para dados, com as características de tempo real, tais como áudio e vídeo interativos. Os serviços incluem a identificação do tipo da carga total, seqüenciamento, *checkpointing*, e monitoração da entrega.

O protocolo de RTP (como demonstrado na Figura 9) fornece características para aplicações em tempo real, com a habilidade de reconstruir o sincronismo, a detecção da perda, a segurança, a entrega e a identificação de codificação de esquemas. Os gateways de mídia que digitalizam a voz, usam o protocolo de RTP para entregar o tráfego da voz (portador). Para cada participante, um par de endereços IP definem a sessão entre os dois *endpoints*, que traduz em uma única sessão de RTP para cada chamada em andamento.

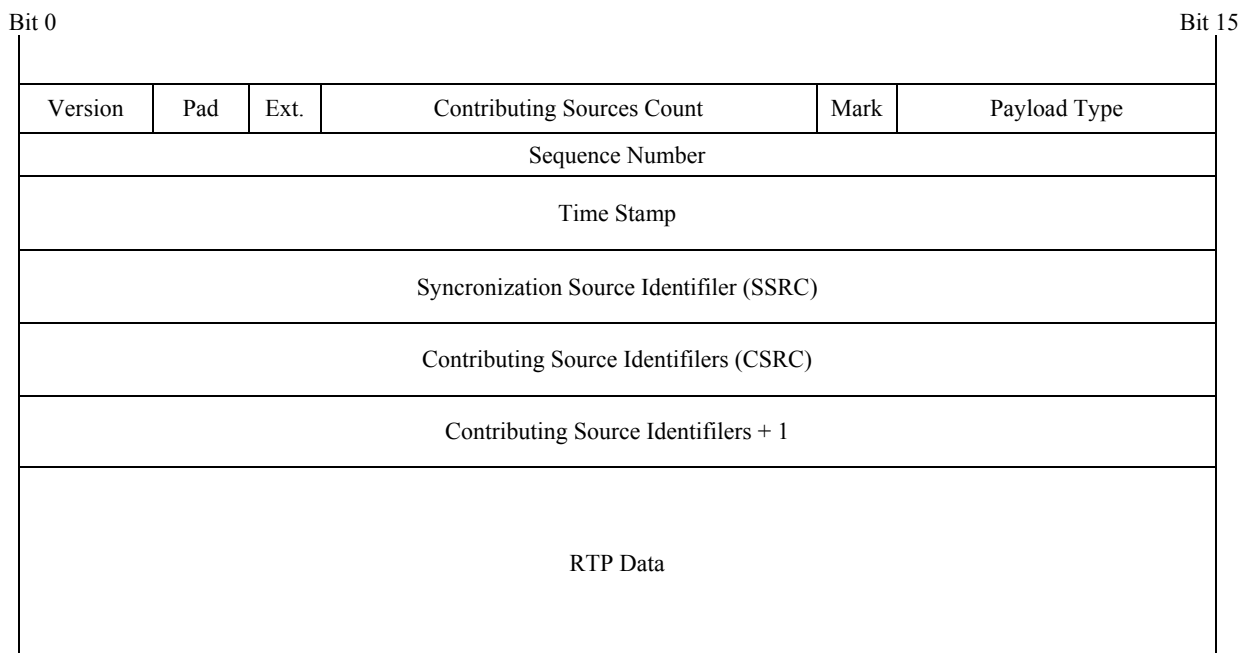


Figura 9: Cabeçalho do Protocolo RTP

RTP é um serviço de aplicação definido sobre o protocolo UDP, livre de conexão, baseado no serviço de entrega de melhor-esforço. Embora RTP seja livre de conexão, tem um sistema que permite a detecção de pacotes faltantes. Como a parte de sua especificação, o campo do tipo do *payload* de RTP inclui o esquema codificação que o gateway usa para digitalizar a voz. Este campo identifica o formato do payload de RTP e determina sua interpretação pelo CODEC no gateway. Um perfil especifica o padrão estático de mapeamento de códigos dos tipos de payload para os formatos do payload. Estes mapeamentos representam a série de ITU-G de esquemas de codificação.

Com os tipos diferentes de esquemas e de taxas de codificação da criação do pacote, os pacotes de RTP podem variar no tamanho e no intervalo. Todos os parâmetros combinados das sessões de RTP ditam quanto de banda é consumido pelo tráfego do portador da voz. O tráfego de RTP que carrega o tráfego da voz é o único contribuinte a carga de VoIP na rede.

#### **1.5.4. Protocolo de Controle de Transmissão em Tempo Real (*Real-time Transport Control Protocol – RTCP*)**

O protocolo de controle de transmissão em tempo real (RTCP) é o protocolo opcional que acompanha o RTP; não é necessário para que RTP trabalhe. A função preliminar de RTCP é fornecer o gabarito na qualidade da distribuição dos dados que está sendo realizada por RTP. Esta função é uma parte do RTP porque é um protocolo de transporte e está relacionado às funções do fluxo e de controle do congestionamento da rede.

Embora os relatórios do gabarito de RTCP não digam onde os problemas estão ocorrendo (somente que são), podem ser usados como uma ferramenta para encontrar problemas. Com a informação gerada de diferentes *gateways* na rede, os relatórios do gabarito de RTCP permitem-no avaliar onde o desempenho da rede pode degradar. RTCP o

permite também monitorar a qualidade de uma sessão da chamada pela perda seguindo do pacote, pela latência (atraso), pelo tremor, e por outros interesses chaves de VoIP.

Esta informação é fornecida em uma base periódica a ambas as extremidades e processada por um *gateway*. Alguns *gateways* não puderam empregar RTCP porque a facilidade para relatar tal informação não é aplicável ao usuário da extremidade. Por exemplo, um único usuário residencial (com um telefone analógico) não pôde ter o acesso ao *gateway* que fornece o serviço. Também, o fornecedor do *gateway* pode usar uma aproximação mais escalonável de estatísticas seguindo da qualidade da chamada. Neste caso, o armazenamento do transporte e a apresentação da informação estatística são dependentes do dispositivo.

Usando RTCP (ou execuções específicas de um fornecedor) na rede, deve ser feita uma verificação de cálculo de largura de banda para o protocolo. É necessário limitar o tráfego do controle de RTCP a uma fração pequena e a largura de banda deve ser informada para cada sessão. Deve ser pequena de modo a não danificar a performance do protocolo do transporte de dados.

Segundo Friedman (1999), uma verificação da largura de banda é necessária de modo que possa incluir o tráfego do controle na especificação da largura de banda. As especificações do RFC recomendam que a fração da largura de banda alocada para RTCP seja fixada em 5% do tráfego de RTP.



### 1.5.5. Protocolo de Controle de Gateway (Media Gateway Control Protocol – MGCP)

O protocolo do controle de *gateway* (RFC 2705, 1999) segue mais a filosofia da arquitetura do *softswitch*<sup>2</sup>. Saindo da linha de *switches* de voz tradicionais, dos componentes dos *gateways*, controlador de *gateway*, e unidades funcionais de *gateway* de sinalização. Isto facilita controlar de cada *gateway* de VoIP como uma entidade separada.

MGCP é um protocolo *master-slave* de controle de *gateway* que coordena as ações de *gateways* (Figura 10). O controlador do *gateway* na nomenclatura de MGCP é consultado às vezes como a um agente da chamada. O agente da chamada controla a inteligência sinalizando *call-related* do controle, quando o *gateway* informar o agente da chamada de eventos do serviço. O agente da chamada instrui o *gateway* para criar e encerrar conexões quando as chamadas são geradas. Na maioria dos casos, o agente da chamada informa aos *gateways* para iniciar uma sessão de RTP entre dois *endpoints* (ARANGE, 2003).

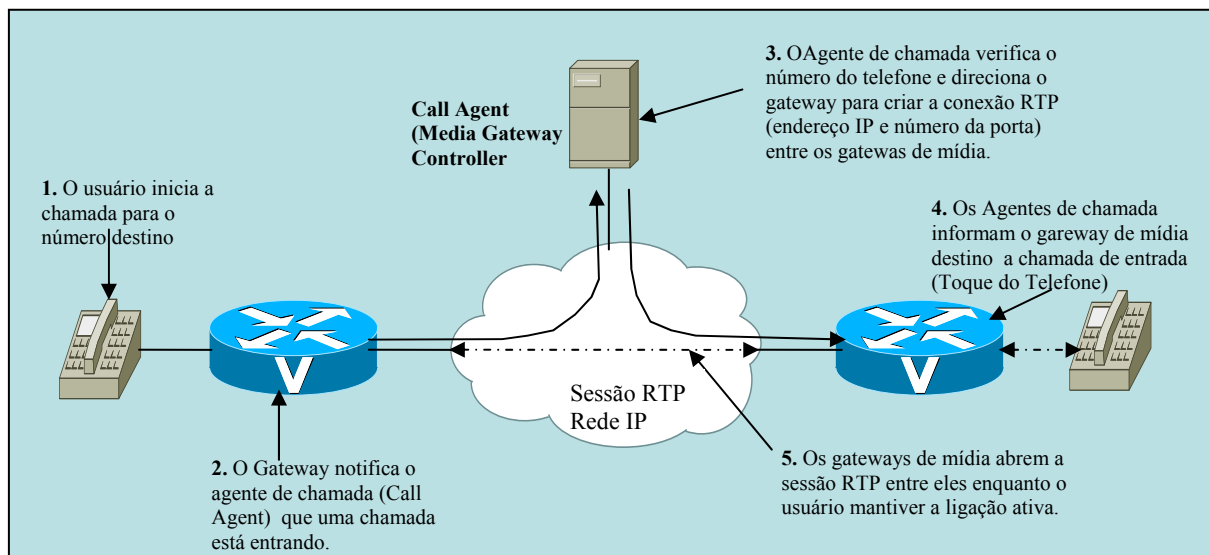


Figura 10: Controle de ações do *Media Gateway Control Protocol* – MGCP (ARANGE, 2003)

<sup>2</sup> Softswitch - interconexão aberta de módulos de software que controlam chamadas de voz e dados em uma rede convergente. O softswitch dita o controle de chamada, estabelece e encerra a sinalização, intermedia protocolos e a criação de serviços em ambientes híbridos. Isto permite que operadoras direcionem qualquer tipo de tráfego por sistemas IP e PSTN e definem ofertas em telefonia IP.

As sinalizações executadas pelo agente e pelos gateways de chamada estão no formulário de mensagens estruturadas dentro dos pacotes do UDP. O agente da chamada e os gateways de mídia têm recursos de retransmissão para estas mensagens.

### **1.5.6. O Protocolo SIP (*Session Initiation Protocol* – Protocolo de Inicialização de Sessão)**

Uma alternativa que compete com o protocolo H.323, é o protocolo SIP, protocolo mais recentemente definido para a telefonia IP. Como os outros protocolos da família Internet (HTTP, FTP, SMTP), ele é baseado em código texto. O protocolo SIP foi publicado pela IETF no documento RFC 2543 (1999). SIP é uma camada de aplicação e um protocolo de controle e sinalização para criar, modificar e terminar sessões entre um ou mais participantes. Essas sessões podem incluir conferências de multimídia, educação a distância, chamados de telefonia IP e distribuição de arquivos com multimídia (CLAUDE, 2002).

O SIP suporta cinco categorias para estabelecimento e terminação dos chamados: pela localização, pela característica e disponibilidade, início de chamados e modificação das características dos chamados.

#### **1.5.6.1. Processo de execução do SIP**

O principal aspecto do protocolo SIP é a sua sinalização, criando, alterando e terminando os chamados (sessões RTP). Os terminais são identificados pelo endereço SIP e não pelos endereços IP. A Figura 11, apresentada a seqüência da execução para criação do chamado SIP:

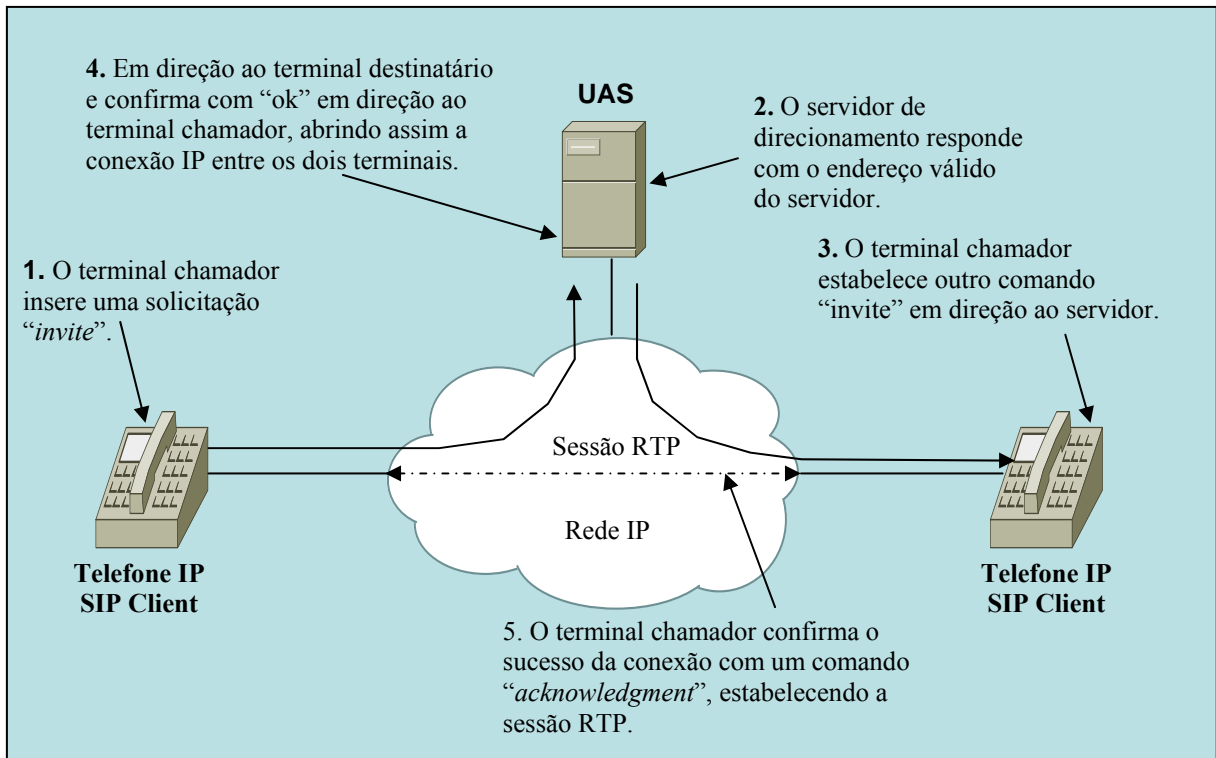


Figura 11: Execução para criação do chamado SIP.

- 1 – O terminal chamador insere uma solicitação “invite” na direção do servidor de direcionamento com base na URL (requer DNS) e informando o endereço do terminal do destino;
- 2 – O servidor de direcionamento responde com o endereço válido do servidor “user agent server – UAS”;
- 3 – O terminal chamador estabelece outro comando “invite” em direção ao servidor “UAS”;
- 4 - Em direção ao terminal destinatário que confirma com “ok” em direção ao terminal chamador, abrindo assim a conexão IP entre os dois terminais;
- 5 – O terminal chamador confirma o sucesso da conexão com um comando “*acknowledgment*”, estabelecendo a sessão RTP.

A transação SIP consiste numa solicitação SIP e uma resposta apropriada. Para facilitar o casamento da solicitação e resposta encontram-se vários campos com valores idênticos, tais como: o identificador de chamado “*call ID*”, o número de seqüência do comando, o campo “*to*”, o campo “*from*” e “*tag*” (BERNI, 2002).

O protocolo SIP utiliza algumas características dos outros protocolos, como por exemplo, os cabeçalhos, erros e regras de codificação do protocolo HTTP. O protocolo de sinalização SIP inicia uma sessão entre usuários.

### **1.5.7. Transporte de Sinalização**

O transporte de sinalização (SigTran) é um projeto dentro do IETF (informativo RFC 2719) que direciona o problema de performance de transporte de sinalização (SS7-to-VoIP). O SigTran foi definido para ser o protocolo de controle entre o *gateway* de sinalização (terminação da sinalização associada com um dado canal/circuito PSTN) e controladores de *gateways* de mídias.

De acordo com a funcionalidade do SigTran, também é possível repassar as mensagens SS7 (conforme descrito no módulo 2.5.1) de sinalização através de uma rede IP, a uma terminação do PSTN em ambas as extremidades. O SigTran trabalha geralmente como um controlador de gateway de sinalização. Estes dispositivos definem uma ponte sobre a rede SS7 e a rede VoIP.

O SigTran garante a interoperabilidade que as redes heterogêneas apresentam. Torna-se crítico quando as chamadas de telefone de VoIP que têm um fluxo *end-to-end* que terminam nas conexões de redes PSTN-VoIP-PSTN.

As mensagens do SigTran necessitam uma maior prioridade para que as redes de VoIP funcionem corretamente. Os controladores de *gateway* de sinalização e de mídia geralmente não mudam suas propriedades; uma vez configurados, não mudam suas posições ou endereços. Desde que as origens e os destinos de mensagens do SigTran se mantenham estáticos, classificar os mecanismos de sinalização de CoS (*Class of Service* – Classe de serviço) é relativamente rápido.

### **1.5.8 Megaco/H.248**

O MGCP foi desenvolvido pelo IETF (órgão que padroniza protocolos da Internet), e o Megaco é resultado de um desenvolvimento conjunto entre o IETF e o ITU (órgão de padronização de protocolos de telecomunicações), resultando num melhoramento do MGCP. A arquitetura de Megaco define os *gateways* de mídia que fornecem a conversão dos meios e as origens das chamadas, quando os controladores do *gateways* fornecerem o controle da chamada.

Megaco atribui as mesmas exigências que MGCP e, em consequência, há algum esforço em unir os protocolos. O Megaco define uma série das transações coordenadas por um controlador de gateway de mídia para o estabelecimento de sessões da chamada. O foco preliminar do Megaco é o promotor de estabilização para o equipamento de telefonia IP. Seguem abaixo a descrição de alguns objetivos do projeto Megaco:

- O telefone IP de Megaco reconhece as necessidades básicas do usuário, de acordo com o objetivo de aplicação;
- Fornece um rápido caminho de expansão para suportar características sofisticadas de telefonia.

- Permite uma larga escala de telefones e de dispositivos similares.
- Implementação simples e rápida.
- Melhor custo benefício para o dispositivo implementado, de acordo com as facilidades oferecidas.
- Tipos de terminações e pacotes implementam características que garantem a confiabilidade.
- A telefone do IP se depara com exigências que se encaixam perfeitamente ao protocolo Megaco/H.248, como proposta na versão original do Megaco.

O objetivo desse tópico foi justamente apresentar as principais características do protocolo Megaco, citando suas potencialidades, a interatividade entre o protocolo MCGP e suas principais funções. Não faz parte desse estudo, entretanto, uma descrição aprofundada sobre as características técnicas do referido protocolo, como descrito em (RFC 3015, 2000).

### **1.5.9. Protocolo de Reserva de Recurso (RSVP - *Resource Reservation Protocol*)**

RSVP (RFC 2205, 1996) não é especificamente um protocolo aplicado a VoIP; inicialmente foi definido como um mecanismo que permite que redes baseadas em alta carga de roteamento, consigam oferecer QoS para aplicações multimídia (WROCLAWSHI, 1997).

RSVP foi criado originalmente para suportar a reserva de recursos (largura de banda ou conectividade) para aplicações específicas, não se tratando de protocolo de roteamento. Cada aplicação sinaliza seus elementos da rede (recursos) emitindo um pedido de RSVP. Este pedido permite que o recurso seja reservado ao longo do seu trajeto (fluxo de tráfego).

Os roteadores (*routers*) identificariam a aplicação de acordo com seus endereços, tipo de protocolo e números de portas. Um *scheduler* de pacote ou é usado para determinar quando os pacotes foram enviados. Em operações *full-duplex*, são necessários dois pedidos que vão em ambas as direções, de forma que as garantias possam satisfazer as reservas de RSVP.

Na essência, o protocolo RSVP permite que as aplicações reservem largura de banda para seus fluxos de dados. Ele é usado por um hospedeiro, em nome do fluxo de dados de uma aplicação, para requisitar uma quantidade específica de banda na rede. O RSVP também é usado por roteadores para repassar requisições de reserva de largura de banda. Para implementar o RSVP, um software RSVP deve estar presente nos receptores, remetentes e roteadores. De acordo com (KUROSE, 2003) podemos citar duas características do RSVP:

1. Fornece **reservas de largura de banda em árvores *multicast*** (o *unicast* é tratado como um caso de *multicast* degenerado).
2. É **orientado ao receptor**, isto é, o receptor do fluxo de dados é quem inicia e mantém a reserva de recursos usada para aquele fluxo.

As duas características apresentadas por Kurose (2003) podem ser verificadas no diagrama da Figura 12. Neste diagrama, é apresentada uma árvore *multicast* com dados partindo do topo para os hospedeiros na base da árvore. Embora os dados se originem no remetente, as mensagens de reserva se originam nos receptores. Quando um receptor transmite uma mensagem de reserva pela corrente ascendente em direção ao remetente, ele mistura a mensagem de reserva com outras mensagens de reserva que estão chegando da corrente inferior.

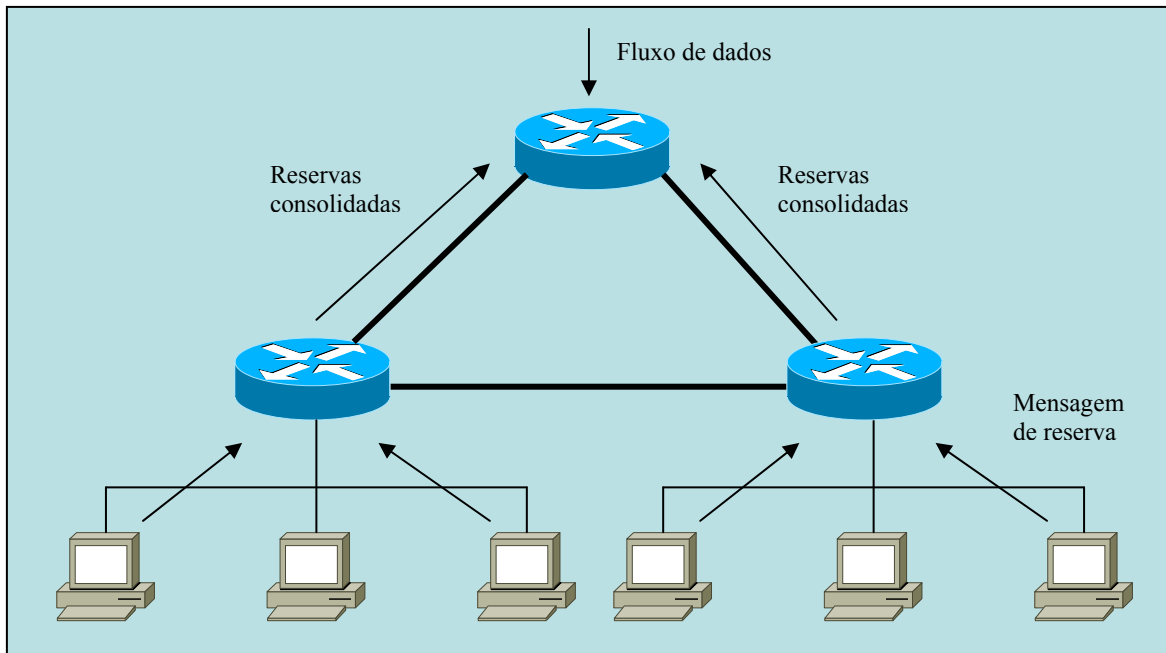


Figura 12: RSVP orientado ao multicast e ao receptor (Kurose, 2003)

O padrão RSVP (RFC 2205, 1996) não especifica como a rede fornece largura de banda reservada aos fluxos de dados. O RSVP é um protocolo que permite que as aplicações reservem a largura de banda necessária. Uma vez efetuadas as reservas, é de responsabilidade dos roteadores da Internet o fornecimento de largura de banda reservada aos fluxos de dados. É importante esclarecer que o RSVP não é um protocolo de roteamento – ele determina em quais enlaces que as reservas devem ser feitas. Ao contrário disso, o RSVP depende de um protocolo de roteamento para determinar as rotas de fluxos.

O RSVP é muito caracterizado como um protocolo de sinalização (conforme descrito na sessão 1.5.7), ou seja, é um protocolo que permite que os hospedeiros estabeleçam e encerrem as reservas para os fluxos de dados.

A seguir é apresentado um breve estudo sobre QoS (*Quality of Service* – Qualidade de Serviço), um importante aspecto de implantação e operação, tanto para redes de pacotes



como um todo, como para redes IP em particular. Considerações importantes sobre o tráfego de VoIP também serão levantadas nas próximas sessões, tais como: Vazão, Latência e Jitter.

## **1.6. QoS – *Quality of Service* (Qualidade de Serviço)**

A qualidade de serviço (QoS) em redes é um importante aspecto de implantação e operação tanto para redes de pacotes como um todo, como para redes IP em particular (Claude, 2002).

As redes TCP/IP, ou simplesmente redes de IP possuem uma grande base instalada, e seu crescimento e aceitabilidade se dão por dois fatores: o crescimento da Internet e a aceitação pelas empresas da base TCP/IP como plataforma de suporte às aplicações em rede. E para os próximos anos esse cenário não tende a mudar, pois teremos cada vez mais computadores utilizando o TCP/IP para comunicação via Internet.

Em um âmbito geral, as redes IP foram desenvolvidas e têm como uma de suas premissas básicas, o poder de ser utilizadas com diversos tipos de meios físicos e tecnologias existentes, de forma a viabilizar a comunicação entre as aplicações fim – a – fim em redes.

Para o cenário atual, a utilização das redes IP demanda que “qualquer aplicação” possa executar com qualidade sobre o IP. Dessa forma, a questão que segue vem a ser a identificação das eventuais limitações do IP e procedimentos necessários para adequá-lo à realidade de utilização das redes.

O desafio para utilização do IP como plataforma principal para aplicações em redes é: o protocolo IP sem qualquer garantia de qualidade de serviço e as dificuldades de mudanças de protocolo, por causa da grande base IP instalada (SILVA, 2000).

A primeira questão é de caráter técnico, e diz respeito ao paradigma previsto para o protocolo que enfatiza a simplicidade de concepção. Um exemplo disso podemos citar o protocolo IP, que não tem nenhuma garantia de vazão constante para uma aplicação em

particular. Além disso, uma aplicação não pode obter do IP nenhuma garantia de entrega dos próprios pacotes que eventualmente são descartados ou perdidos sem que nenhum tipo de correção ou ação seja tomada. Não existe também nenhuma garantia de tempo de entrega para os pacotes (YAMADA, 2001).

A segunda questão é de como se adequar a um novo protocolo (IP), sem efetivamente mudá-lo. O IP (versão 4) deverá mudar para IP (versão 6). O IPv6 ou Ipng (*new generation*) está focado para outras questões (endereçamento, segurança). E não apresenta nenhuma solução para a questão citada.

Pode-se dizer que a proposta de novos algoritmos com mecanismos de tratamento de deficiências são algumas das formas de se sanar as “deficiências” do IP, e permitirem o suporte efetivo de qualquer tipo de aplicação sobre redes IP.

O IP possui uma grande base instalada e a tendência é que ele suporte as aplicações em rede, tais como:

- Telefonia e Fax sobre IP (VoIP – *Voice over IP*);
- Comércio eletrônico (*E\_commerce*);
- Vídeo sobre IP;
- Educação a distância (EAD) (*Distance Learning*);
- Vídeo conferência;
- Aplicações de grupo;
- Aplicações multimídia;
- Aplicações em tempo real;
- Outras

Para obtenção de uma QoS, o entendimento dos seus princípios, parâmetros, mecanismos, algoritmos e protocolos são de grande importância, pois a qualidade de serviço

(QoS) nas redes é um aspecto fundamental para o desempenho fim – a – fim das aplicações (VoIP, multimídia,...).

### **1.6.1. Princípios – Qualidade de Serviço (QoS)**

Pode-se dizer, que Qualidade de Serviço é um conjunto de padronizações que impõem limites (valores mínimos e máximos) para determinados parâmetros (perdas, atrasos, vazões, etc).

A QoS é garantida pela rede, seus componentes e equipamentos utilizados. Do ponto de vista dos programas de aplicação, a QoS é expressa em termos de uma “solicitação de serviços” ou “contrato de serviços”. Esse contrato é denominado de SLA (Service Level Agreement) (McCABE, 1998).

Algumas aplicações, como as de multimídia, são as que normalmente exigem os parâmetros de QoS mais satisfatórios. Entretanto, em aplicações multimídia, envolvendo dados, gráficos e arquivos com animação (vídeo...), em que não há necessidade de sincronização ou aplicação de QoS. Por outro lado, em uma aplicação envolvendo áudio, somente garantir a vazão não é o suficiente. Para este caso, os atrasos e as perdas influenciam tanto na interação com o usuário, como nos recursos das redes, exigindo-se assim uma maior necessidade de suporte da qualidade de serviço da rede.

Na seqüência, algumas considerações sobre os parâmetros de QoS em aplicações multimídia são feitas para identificar suas exigências.

### **1.6.2. Vazão**

Dentre os vários parâmetros necessários para satisfazer as exigências da QoS, a vazão é a mais básica, devendo esta ser atendida pela rede. Abaixo seguem alguns parâmetros de vazão para algumas aplicações típicas:

<b>Aplicação</b>	<b>Vazão (Típica)</b>
Aplicações Transacionais	1 Kbps a 50 Kbps
Quadro Branco (Whiteboard)	10 kbps a 100 kbps
Voz	10 kbps a 120 kbps
Aplicações Web (www)	10 kbps a 500 kbps
Transferência de arquivos	10 kbps a 1 Mbps
Vídeo (Streaming)	100 kbps a 1 Mbps
Aplicação de Vídeo Conferência	500 kbps a 1 Mbps
Vídeo MPEG	1 Mbps a 10 Mbps
Aplicações com Imagens Médicas	10 Mbps a 100 Mbps
Aplicação em Realidade Virtual	80 Mbps a 150 Mbps

Tabela 2: Vazão típica de Aplicações em Rede (MARTINS, 2001)

### **1.6.3. Latência (Atraso)**

A latência (ou atraso) é o tempo que o pacote leva para percorrer todo o percurso de uma rede por toda a sua extremidade. Em termos de telefonia, a latência é o tempo médio que a voz leva para chegar até o ouvinte. Os grandes valores de latência não degradam necessariamente a boa qualidade uma chamada de telefone, mas o resultado pode ser uma falta da sincronização entre os clientes. Geralmente, aceita-se que a latência fim-a-fim deve ser menos de 150ms. Para garantir que o custo da latência permaneça abaixo de 150 ms, uma verificação sobre as causas da latência deve ser feita (KAMIENSKI, 2000).

Ao projetar uma rede de multi-serviço, o total de atraso que um sinal ou as amostragens de pacotes sofrem uma somatória de todas as latências em todos os pontos da rede.

A origem de uma latência é o tempo que o *endpoint* leva para criar os pacotes usados para o serviço de voz. Este atraso no empacotamento é causado pela demora ao completar os pacotes com os dados. Geralmente, quanto maior o tamanho do pacote, maior o tempo para completá-lo com dados. O atraso no empacotamento é administrado pelo padrão do CODEC que está sendo usado. Este problema existe também no lado da recepção porque o *gateway* de mídia deve retirar os pacotes da fila e promover o processo de completar os dados do pacote de retorno. Se os pacotes forem mantidos em um tamanho menor, é geralmente pequena a dependência de hardware e software na implementação de gateways de mídia. Todas as operações de gateway de mídia não devem exceder 30 ms.

Uma outra fonte de latência é o atraso na digitalização dos dados nos links físicos nos equipamentos de interconexão (PERCY,1999). Este atraso é inversamente proporcional à velocidade do link. Ou seja, quanto mais rápidos os links, mais baixa é a latência. Este valor depende da tecnologia de conexão, bem como o seu método de acesso. Por o exemplo, leva-se o tempo de 125 microssegundos para inserir 1 byte em um circuito 64-Kb. O mesmo byte colocado em um circuito Oc-3/stm-1 é transmitido em 0,05 microssegundos. Embora este atraso não seja variável (independente da largura de banda usada), mantendo o número de link pequeno e usando taxas de largura de banda mais altas, a latência é reduzida.

A propagação do atraso é o tempo que um sinal elétrico leva para percorrer o comprimento de um meio físico. A velocidade destes sinais é sempre mais lenta do que a velocidade de luz. Há sempre um atraso na propagação; entretanto, o atraso somente é

considerado quando o sinal (ou o pacote) viaja a uma longa distância. A fórmula para cálculo de propagação do atraso é a seguinte:

$$\text{Propagação do atraso} = \text{Circuito km} / (299,300 \text{ km} \times 6).$$

A fila de atraso se dá por uma grande origem de latência, é a quantidade de tempo que um pacote fica armazenado no buffer em algum elemento da rede, enquanto aguarda ser transmitido. As cargas do tráfego da rede resultam em enfileiramento variável do atraso. A quantidade de *buffes* usados em uma fila é geralmente um parâmetro configurável; quanto menor o número de *buffers*, melhor será o valor de latência. Entretanto, este atraso é baseado também na quantidade de tráfego que o elemento está tentando passar com um determinado link, e conseqüentemente aumenta com a carga da rede. Com isso, torna-se necessário reservar a largura de banda adequada e os recursos para o tráfego de voz. Se a fila usada para o tráfego da voz não estiver suportando a demanda e essa fila estiver crescendo muito, o resultado é um aumento da latência.

O atraso no chaveamento de pacote trata-se de uma outra origem de latência, onde o tempo de transferência de um roteador ou de um *buffer de switch* de um pacote para definir em qual interface o pacote deve ser dirigido. Embora este atraso seja geralmente pequeno, a arquitetura do roteador ou do *switch* deve ser um fator a se considerar. Se um pacote deve ser armazenado como parte de seu processamento, isso irá repercutir em uma maior latência.

A latência é o objeto principal do nosso estudo. Nos próximos capítulos discutiremos em (Yamada, 2001) o atraso de pacotes de voz em uma rede IP.

### 1.6.4. Jitter

O jitter é o valor médio de todas as diferenças de atrasos incrementais, do início da medição até o pacote atual. Os pacotes mais recentes têm um peso maior do que o dos pacotes menos recentes (ROWE, 1999). Ou seja, com uma taxa constante de transmissão de pacotes de 20 ms, cada pacote esperar-se-ia chegar exatamente no destino nos 20ms. Isso nem sempre acontece dessa forma. Por exemplo, a Figura 13 mostra o pacote (P1) e o pacote (P3) chegando dentro do esperado, mas o pacote (P2) chegando 12 ms chegando mais tarde e o pacote quatro (P4) chegando 5 ms mais tarde.

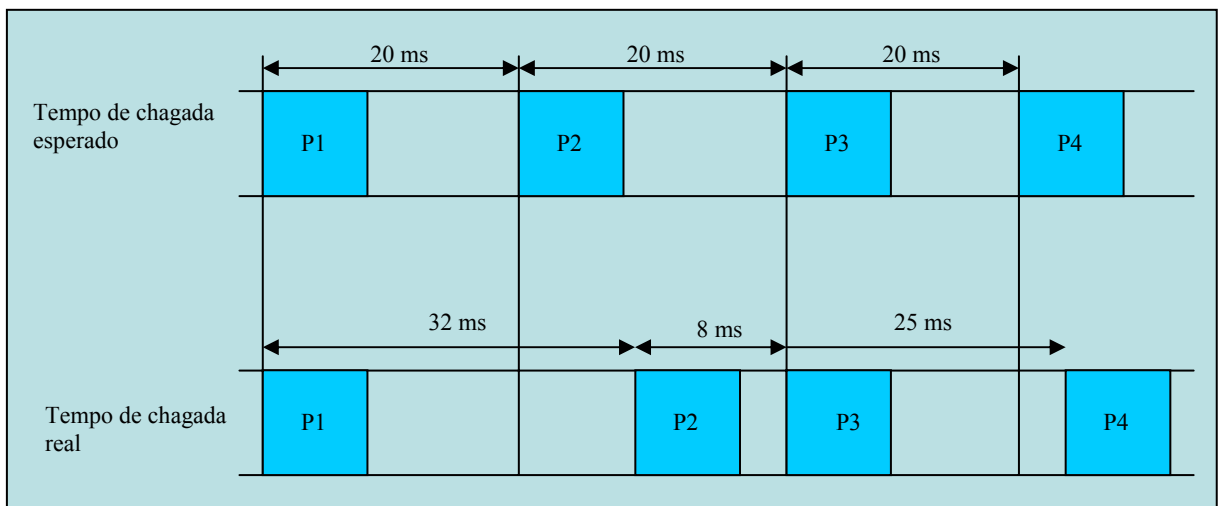


Figura 13: Comparação do tempo de chegada esperado com o tempo de chegada real (ROWE, 1999).

O maior culpado pelo *jitter* é a variação no enfileiramento devido a mudanças dinâmicas nas cargas de tráfego da rede.

Os *gateways* de mídia têm os *play-out buffers* (*buffers* de apresentação) que armazenam uma cadeia de pacotes de modo que a onda de voz reconstruída não seja afetada pelo jitter do pacote. Os *play-out buffers* podem minimizar os efeitos do jitter, mas não podem eliminar o jitter mais acentuado.

Embora alguma quantidade de jitter deva se esperar, o jitter mais acentuado pode causar alterações na qualidade da voz porque o gateway de mídia pode rejeitar a chegada de pacotes avariados. Nesta circunstância, o gateway de mídia podia iniciar seu *play-out buffers* e causar aberturas na reconstrução das ondas de voz.

### **1.6.5. Perdas de pacotes**

As duas principais causas de perdas de pacotes em redes IP são os descartes de pacotes por roteadores e switch routers, provocados por erros, congestionamentos, etc, e por perdas ocorridas na camada 2 (PPP - *Point – to – Point protocol*, *Ethernet*, *Frame Relay*, *ATM*, etc.) durante o transporte dos mesmos (SAWASHIMA, 1996).

Em aplicações como VoIP, a perda de pacotes é um problema sério, pois a perda de pacotes com trechos de voz digitalizada implica numa perda de qualidade não aceitável para a aplicação. A solução para a perda de pacotes é específica para cada aplicação.

Em QoS, a perda de pacotes é um fator preocupante e seus limites são controlados definindo-se taxas de perdas que permitam uma operação adequada da aplicação.

## **1.7. Alternativas técnicas para qualidade de serviço em redes IP**

Existem essencialmente duas alternativas técnicas para abordagem sobre a qualidade de serviços em redes IP: o modelo de serviços integrados (*Intserv Integrated Services Architecture*) com RSVP (*Resource Reservation Protocol*), discutido na sessão 1.5.9, e o DifServ (*Differentiated Services*), modelos de serviços diferenciados.

Além da IntServ e da DifServ, outras alternativas também podem ser citadas, a MPLS (*Multiprotocol Label Swithing*), a SBM (*Subnet Bandwidth Managment*) e o dimensionamento, onde os recursos são dimensionados na fase de projeto de forma a não



haver congestionamento. Todas as alternativas citadas são iniciativas do IETF (*Internet Engineering Task Force*), o qual está empenhado em propor um conjunto de soluções para os mecanismos de controle de QoS, que garanta a interoperabilidade dos sistemas entre diferentes fornecedores. Isso enfatiza a importância das redes IP para o suporte às novas aplicações multimídia, com restrições de tempo real, etc.

Como apresentado na sessão 1.5.9, o RSVP pode ser usado para reserva de recursos por fluxo nos roteadores de rede. A capacidade de requisitar e reservar recursos por fluxo, por sua vez, faz com que seja possível que a estrutura de Intserv forneça qualidade de serviço a fluxos individuais. Com a progressão de trabalhos com o Intserv e RSVP, pesquisas na área (ZHANG, 1998) apontaram algumas dificuldades no modelo Intserv:

- Escalabilidade. A reserva de recurso por fluxo usando o RSVP está diretamente proporcional à capacidade de processamento do roteador em processar as reservas de recursos. Com isso, esse processamento representa uma considerável sobrecarga para redes de grande porte.

- Modelos de serviços flexíveis. A estrutura Intserv fornece um pequeno número de classes de serviços (CoS). Os conjuntos de classes fornecidas não comportam características como priorização de atividades ou mesmo relacionamento de processos de dependências entre elas.

Novas propostas de técnicas para garantia de qualidade de serviços em redes IP estão sendo discutidas pela IETF. Um grupo de trabalho, o Diffserv (Diffserv,1999) (*Differentiated services* – serviços diferenciados) está desenvolvendo uma arquitetura para fornecer diferenciação de serviço escalável e flexível, ou seja, a capacidade de operar diferentes “classes” de tráfego de maneiras diferentes dentro da Internet.

Segundo Ross (2003), a necessidade de escalabilidade surge do fato que centenas de milhares de fluxos de tráfego simultâneos podem estar presentes em um roteador de *backbone* Internet. Ao contrário da Intserv, a Diffserv apresenta uma maior integração entre as classes de serviços através da interoperabilidade entre os processos ou classes de serviços.

A flexibilidade está relacionada à facilidade de tratamento de classes de serviços, onde antigas classes de serviços podem ser todas tornadas obsoletas, enquanto novas classes podem ser implementadas, sem a definição de serviços específicos em novas classes de serviços, fornecendo componentes funcionais (objetos) para a arquitetura de rede. A maioria desses objetos é descrita detalhadamente em (RFC 2475, 1998), mostrando suas possíveis variações.

A seguir apresentaremos um modelo de arquitetura Diffserv, descrevendo aspectos de serviços diferenciados, sem nos estender a seus modelos arquitetônicos.

### **1.7.1 Um modelo Diffserv**

Em Ross (2003) é descrito um modelo Diffserv que apresenta uma estrutura que define os componentes arquitetônicos do modelo de serviço diferenciado. Conforme modelo apresentado na Figura 14, são apresentados alguns aspectos dos serviços diferenciados com seus conjuntos de elementos:

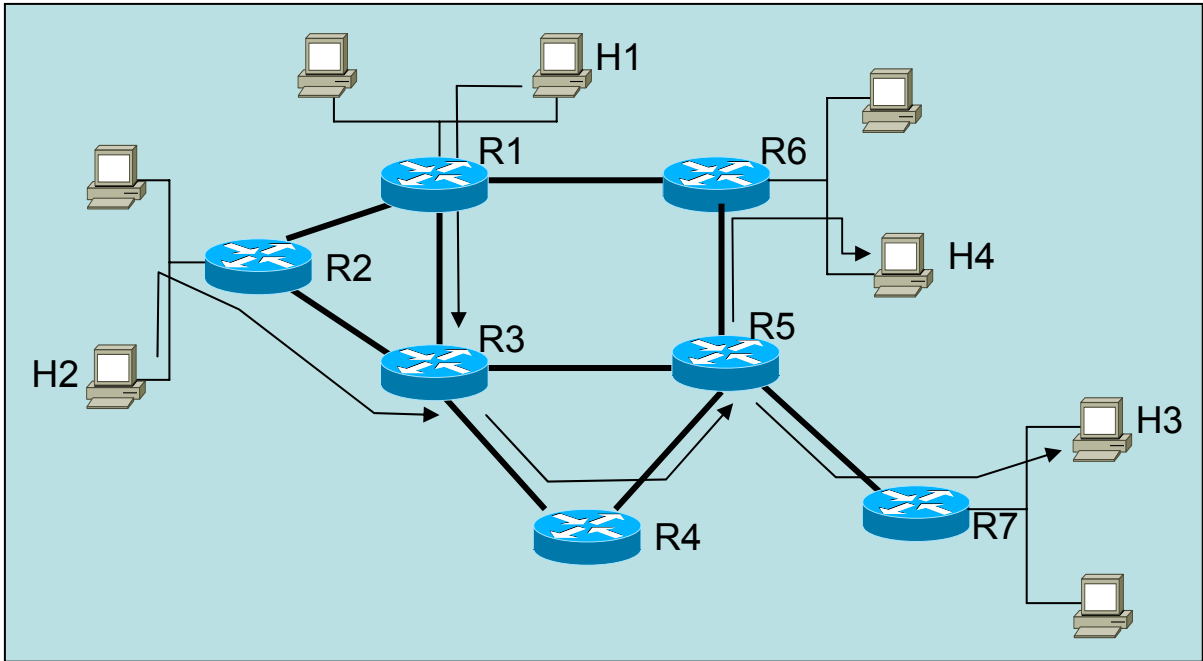


Figura 14: Um exemplo simples de rede Diffserv (ROSS, 2003).

Duas funções estabelecidas como classes de serviço propostas pelo autor são descritas de acordo com o modelo apresentado acima, são elas:

- **Função de borda:** *classificação de pacotes e condicionamento de tráfego*. O tráfego de borda (host ou o primeiro roteador habilitado a Diffserv), os pacotes que chegam são marcados, ou seja, é atribuído um determinado valor ao campo *differentiated service* (DS) do cabeçalho do pacote. Na Figura 14, os pacotes sendo enviados de H1 para H3 e os pacotes enviados de H2 para H4 recebem marcas diferente para cada tipo de tráfego. Tal marca identifica a classe de tráfego à qual ele pertence, podendo receber serviços diferenciados dentro do núcleo da rede, por diferentes classes de tráfego.

- **Função do núcleo:** *envio*. Em um roteador habilitado a diffserv, ao receber um pacote marcado com DS, este é repassado até seu próximo salto de acordo com as características estabelecidas à classe do pacote. Este comportamento é descrito pelo autor como comportamento por salto. Esta característica é diretamente ligada à forma como o buffer e a largura de banda de um roteador são compartilhados entre as classes de tráfego

concorrentes. Dessa forma, conforme o modelo apresentado na figura 14, se os pacotes enviados de H1 para H3 receberam a mesma marca dos pacotes que estão indo de H2 para H4, os roteadores não diferenciarão se os pacotes foram originados de H1 ou H3.

A arquitetura Diffserv ainda se encontra em estágio de desenvolvimento, com uma evolução bastante ascendente e para estudos mais aprofundados deve ser mantido uma atenção especial sobre as atualizações das RFCs 2474 (1998) e RFC 2475 (1998) (BLAKE, 1998).

Por hora concluímos nosso estudo sobre os elementos de VoIP que interagem diretamente ao escopo de uma avaliação sobre atrasos de pacotes de voz em uma rede IP. Apresentaremos nossa proposta de dissertação direcionada a essa avaliação e a dividiremos em duas etapas: no capítulo 3, faremos a apresentação de um estudo segundo Yamada (2001), onde é proposto um estudo sobre o tráfego de pacotes de voz e sobre este estudo levantaremos um modelo teórico de filas para avaliação do atraso dos pacotes. Na seqüência, discutiremos no capítulo 4 um modelo de simulação que comprova a análise do modelo teórico de filas. Para isso usaremos o software *Network Simulator* (NS), também descrito no mesmo capítulo.

## CAPÍTULO 2 – DESENVOLVIMENTO DE UM MODELO TEÓRICO

### 2.1 Multiplexação de sinais de voz em redes IP

Para a utilização da rede IP como plataforma de convergências de redes, vários aspectos dessa rede integrada necessitam ser investigados. O principal problema a solucionar é a garantia de QoS para os serviços que apresentam restrições de tempo. Várias soluções estão sendo propostas, e uma das soluções apresentada pelo IETF, como discutido no capítulo 1, é o DiffServ. Nesta solução, os pacotes são classificados em fluxos, e os fluxos por sua vez são subdivididos em conjuntos menores de classes de serviços. Uma ordem de prioridade é atribuída às classes de serviços. Dessa maneira, os pacotes classificados em uma categoria de classe de serviço são tratados de uma mesma maneira em toda rede. Por ex., os pacotes de voz podem ser classificados em uma classe com alta prioridade de processamento dentro da rede (HANDLEY, 1998) (SCHULZRINNE, 1996).

Um outro aspecto bastante importante para investigar é relativo a multiplexação de pacotes de voz. Para entender o problema envolvido, considere a arquitetura telefone a telefone mostrada na Figura 4. Os *gateways* mostrados na figura devem exercer várias funções. Eles devem tratar de todas as fases da sinalização telefônica e também das fases de conexões *gateway-gateway*. Além disso, os *gateways* necessitam ter codecs, empacotar os sinais de voz digitalizados, acrescentar os cabeçalhos RTP e UDP e transmitir através da rede IP.

A parte de dados (sinais digitalizados de voz) de um pacote de voz tem um tamanho pequeno, para não haver um longo atraso de empacotamento (SUBBIAH,1998). Desse modo, pode haver um overhead muito grande (um cabeçalho longo em relação à parte de dados) na transmissão de pacotes na rede IP. Para reduzir esse overhead, várias técnicas são discutidas em Yamada (2001). Uma técnica é a compressão do cabeçalho através da

eliminação da redundância existente no cabeçalho. A outra técnica é multiplexar vários pacotes de voz de diferentes conexões em um único pacote contendo um único cabeçalho.

Em Yamada (2001), é feito um estudo aprofundado de como a técnica de multiplexação pode ser implementada, e é feita também uma avaliação da qualidade da voz quando ocorrem perdas de pacotes de voz quando são transmitidos na rede IP. Entretanto, o artigo não estuda o impacto do atraso dos pacotes em cada nó da rede, na qualidade da voz. Um atraso longo pode tornar uma comunicação telefônica não natural, dificultando em muito a conversação.

O objetivo desse estudo é analisar a influência dos atrasos dos pacotes multiplexados na qualidade da voz. É de conhecimento geral que um atraso em torno de 150 ms, já é suficiente para haver uma dificuldade de conversação em uma ligação telefônica. Assim, dada uma configuração de rede, com as capacidades de enlaces definidas, o objetivo deste estudo é estimar o número de ligações telefônicas que a rede pode acomodar satisfazendo a restrição de atraso em torno de 150 mseg. Essa estimativa é fortemente afetada pela quantidade de pacotes de outros serviços (por ex., pacotes de dados e/ou pacotes de vídeos) e do tipo de escalonador utilizado. Inicialmente, o estudo vai focar o caso em que não exercerá nenhum controle na saída do buffer (esquema FIFO) e a taxa de pacotes de outros serviços será um parâmetro variável. No outro caso, será utilizado um escalonador com prioridade. Os pacotes multiplexados de voz terão maior prioridade no processamento em cada nó da rede. Neste caso também, o parâmetro variável será a taxa de pacotes de outros serviços, tais como FTP (*file transfer protocol*) em canais TCP e UDP.

## **2.1 Avaliação de técnicas para redução de transmissão de pacotes VoIP**

Para o sistema VoIP são utilizadas técnicas de digitalização de sinais e o fluxo é segmentado em blocos de dados e estes transportados em forma de pacotes IP. O *gateway* de VoIP é um dispositivo de transição entre a rede de telefonia pública compartilhada (PSTN) e uma rede IP (conforme descrito no capítulo 1), por usar técnicas de transporte de pacotes de voz.

Em sistemas VoIP, algumas exigências devem ser identificadas e levadas em conta na implementação da arquitetura do sistema. Nos próximos tópicos, veremos como as exigências do esquema de transporte de dados são uma questão crucial na transmissão de pacotes VoIP.

### **2.1.1. Redução do *overhead* em cabeçalhos de pacotes**

Conforme apresentado no capítulo 1, a exigência da questão de *overhead* na definição do tamanho do pacote é um fator a se considerar. O pacote VoIP tem um cabeçalho IP/UDP/RTP seguido por um *payload* de sinal de voz. Normalmente, o referido cabeçalho tem um tamanho de 40 bytes e assim quando transportamos uma pequena carga, a eficiência da banda usada para a banda de voz é muito pequena. Como exemplo disso podemos dizer que para um bloco codificado pelo *codec* G.729, e codificado em um quadro e em um pacote IP, o cabeçalho do pacote alcança o *overhead* de quase 80% (ITU-T, 1996).

Com a redução do *overhead* em cabeçalhos IP, pode-se dizer que dois fatores estarão diretamente relacionados às exigências:

- Aumento da eficiência, reduzindo o *overhead* no cabeçalho do pacote;
- Redução de largura de banda exigida da rede IP.

### **2.1.2. Redução da carga de processamento dos dispositivos da rede**

O aumento do número de pacotes VoIP de tamanhos pequenos pode trazer um aumento considerável na carga de processamento de roteadores e servidores que compõem a rede. O estouro na capacidade de processamento dos roteadores se dá porque a maioria desses roteadores utiliza o sistema de fila, causando a perda de pacotes. Com esses fatores, a redução da carga de pequenos pacotes VoIP nos meios de roteamento torna-se uma exigência imprescindível.

### **2.1.3 Reduzindo o consumo de recursos UDP**

Uma única conexão UDP é estabelecida entre o *gateway* VoIP e um cliente, em uma chamada. Esta conexão requer um par de portas e recursos de gerenciamento no *gateway* VoIP. Com isso, recursos de comunicações UDP são consumidos, comprometendo a performance do *gateway* VoIP. A priorização também é um processo requerido em redes IP, bem como o gerenciamento de banda entre os *gateways* para manter uma melhor garantia de QoS. A redução do número de portas UDP reduz o número do fluxo para o gerenciamento em redes IP.

## **2.2 Princípios de multiplexação do fluxo de voz em redes IP**

Com base nos estudos de Yamada (2001), objeto do nosso estudo, a técnica de multiplexação de fluxo de voz dentro de um único pacote RTP, uma carga útil de pacotes RTP pode ser compartilhada por vários fluxos de voz, e estes fluxos serem multiplexados em um quadro de voz codificado.



No esquema de multiplexação, um pequeno pacote é constituído por um bloco de códigos gerados em um quadro identificado por um pequeno cabeçalho de pacote para demarcar o fluxo de voz, o tamanho da carga e outras características.

No princípio de multiplexação de pacotes baseado em redes IP, muitos algoritmos podem ser usados para determinar o comprimento de pacotes RTP multiplexados. Neste caso, o autor avalia a qualidade da voz em dois esquemas (como veremos a seguir) baseado em um meio de transmissão que dependem de funções de controle de conexões de fluxo de voz multiplexados. Para que justifique o tamanho do pequeno cabeçalho de pacote, que tem informações para reconstruir o cabeçalho original, apenas duas indicações são propostas pelo autor: 8-bit *channel ID* e 8-bit *payload length*, incluídos no cabeçalho do pacote. O tamanho total do cabeçalho é definido em 2 bytes, considerando a limitação da fragmentação do datagrama IP sobre o ethernet.

### **2.2.1 Esquema 1 – Encaminhamento por tempo**

Neste esquema, o autor propõe um tempo constante para determinar o intervalo de tempo de pacotes RTP que estão sendo enviados. Dessa forma são estabelecidos dois processos:

- 1) É definido um tempo fixo para o tempo de conclusão do pacote RTP. O envio é acionado periodicamente do começo ao fim da operação de multiplexação.
- 2) Quando o envio é ativado no tempo especificado inicialmente, deve ser feita uma coleta dos pequenos pacotes gerados e armazenados para construir o próximo pacote RTP a ser enviado.

Após essa análise do estudo apresentado pelo autor Yamada (2001), apresentaremos o estudo sobre a análise do atraso de pacotes nos mecanismos de transporte propostos. Um modelo teórico de filas é abordado.

### **2.3. Atraso de fila**

A definição de uma análise de atraso em filas é aqui descrita em uma avaliação de atrasos pacote a pacote, diferentes de outros atrasos: de propagação, de transmissão e processamento, talvez o mais complexo.

Podemos caracterizar o atraso em uma fila com o seguinte exemplo: se vinte pacotes chegarem a uma fila vazia ao mesmo tempo, provavelmente o primeiro pacote a chegar na fila não sofrerá nenhum atraso de fila, porém, o último pacote deverá sofrer um atraso considerável, na espera da transmissão dos primeiros pacotes. As medições que caracterizam o atraso de fila são definidas com variações estatísticas tais como atraso de fila médio, a variância de atraso de fila e a probabilidade de que o atraso de fila exceda um valor especificado (TANEMBAUM, 1996).

Para dizermos que um atraso de fila é relativamente grande, dependemos de algumas variantes tais como a taxa com que o tráfego chega à fila, da velocidade de transmissão do enlace e de que forma o tráfego chega à fila (periodicamente ou aleatoriamente).

Em nossa amostragem trabalharemos com algumas nomenclaturas, onde aproveitaremos para apresentá-las desde já, exemplificando o modelo de atraso e facilitando o entendimento. Chamamos de “ $\lambda$ ” a taxa média com que os pacotes chegam à fila (medido em pacotes/segundo). “ $C$ ” é a taxa de transmissão em *bits*/segundo com que os bits saem da fila e “ $L$ ” o comprimento do pacote em *bits*. Com isso, temos uma taxa média de chegada dos *bits* à fila dado por  $L\lambda$  *bits*/s.

Em uma fila longa, com um número infinito de *bits*, a intensidade de tráfego é dada pela razão  $(L\lambda)/C$ , que desempenha um importante papel no tamanho do atraso de fila. Em  $(L\lambda)/C > 1$ , a velocidade média com que os bits chegam à fila excederá a velocidade com que eles possam ser transmitidos para fora, aumentando a fila sem um limite, tendendo ao infinito. Com isso é desejável que a intensidade de tráfego não seja maior do que 1.

Em uma outra situação, consideramos que  $(L\lambda)/C \leq 1$ . Por exemplo, se existe uma chegada periódica de pacotes (1 pacote a cada  $L/C$  segundos), os pacotes chegarão a uma fila vazia e não haverá atraso de fila. Em outro caso, os pacotes chegando em rajadas, porém periodicamente, poderá haver um atraso médio significativo, ou seja,  $N$  pacotes chegam ao mesmo tempo a cada  $(L/C)N$  segundos. Dessa forma, o primeiro pacote transmitido não sofrerá atraso na fila; o segundo pacote terá um atraso de  $L/C$  segundos, e, o  $n$ ésimo pacote transmitido terá um atraso de fila  $(n-1)L/C$  segundos.

No processo de chegada aleatório, ou seja, as chegadas não seguem nenhum modelo, os pacotes são separados por chegada de tempo aleatório. Nessa uma forma mais realista, se a intensidade de tráfego for próxima de zero, as chegadas dos pacotes serão poucas e bem espaçadas, diminuindo a possibilidade de um pacote encontrar outro na chegada da fila. Por outro lado, quando a intensidade de tráfego tender a 1, poderá ocorrer uma situação em que a velocidade de chegada excederá a capacidade de transmissão, aumentando o comprimento médio da fila. Neste caso, à medida que a intensidade do tráfego se aproxima de 1 (um), o comprimento médio da fila vai ficando maior, como demonstrado na Figura 15.

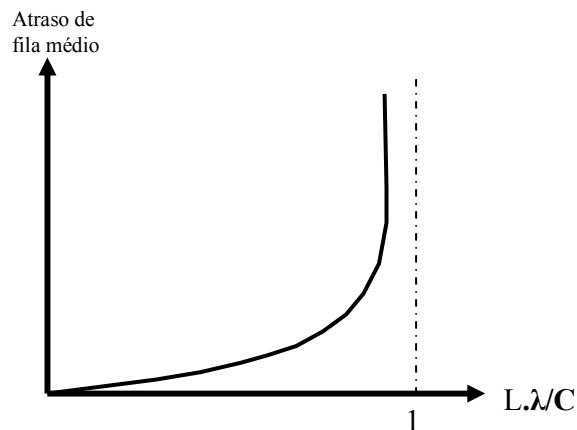


Figura 15: Atraso de fila médio e a intensidade de tráfego

## 2.4. Desenvolvimento de um modelo teórico de filas

A definição do modelo teórico de filas para análise de atraso de filas médio em um tráfego de pacotes de dados e voz é a proposta desse estudo. Para demonstrarmos claramente as características dos tipos de pacotes e os componentes representativos do atraso nodal, apresentamos abaixo um breve modelo esquemático (Figura 16), simbolizando o processo de entrada e saída de pacotes, bem como um *scheduler* ou escalonador do tipo de classes de serviços a serem implementados.

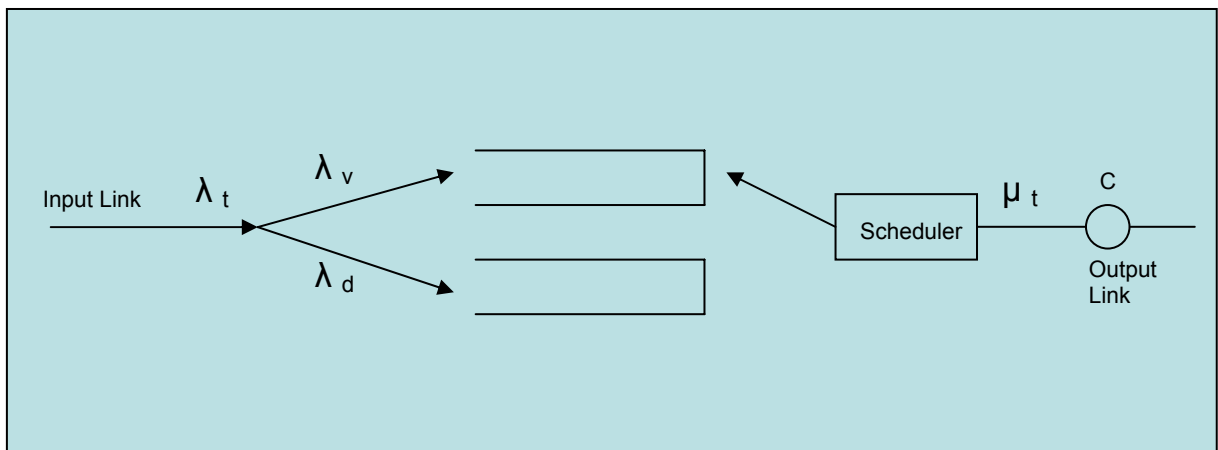


Figura 16: Breve modelo esquemático do processo de entrada e saída de pacotes.

Na seqüência, faremos a proposta dos modelos matemáticos que estabelecem os tipos de classes de serviços que irão atender os requisitos do tráfego voz e dados dentro de um circuito multiplexado. Sabemos que o atraso de 150 msec provoca uma certa dificuldade de conversação. Para tanto, iremos propor dois tipos de classes de serviços as quais avaliarão a melhor situação de tráfego que atenderão as exigências conforme descritas em Yamada (2001).

Com a análise projetada pelo modelo teórico, seremos capazes de estabelecer quais serão as melhores condições dos demais componentes (largura de banda, processamento, meio físico, etc) do circuito que envolve a transmissão. Seguem então, as duas classes de serviço:

- Classe de serviço **Sem Prioridade**:

$$\lambda_v + \lambda_d = \lambda_t ; \quad \mu_v + \mu_d = \mu_t \quad ( 1 )$$

Onde:  $\lambda_t$  – Taxa média total de pacotes que chegam na fila (pacotes / segundo)

$\lambda_d$  – Taxa média total de pacotes de dados que chegam na fila (pacotes por segundo)

$\lambda_v$  – Taxa média total de pacotes de voz que chegam na fila (pacotes / segundo)

$\mu_t$  - Taxa média total de pacotes que saem da fila (pacotes / segundo)

$\mu_v$  - Taxa média total de pacotes voz que saem da fila (pacotes / segundo)

$\mu_d$  - Taxa média total de pacotes de dados que saem da fila (pacotes / segundo)

$$\lambda_v = x\lambda_d ; \quad \mu_v = y\mu_d \quad ( 2 )$$

Onde:  $x$  – porcentagem de pacotes de voz em relação a taxa total de pacotes entrantes  $\lambda_t$

$y$  – porcentagem de pacotes de voz em relação a taxa total de pacotes entrantes.

∴

$$\lambda_d = \frac{\lambda_t}{x+1} \quad ; \quad \mu_d = \frac{\mu_t}{y+1} \quad (3)$$

$$\rho_d = \rho_t \frac{y+1}{x+1} \quad ; \quad \rho_d = \frac{\lambda_d}{\mu_d} \quad ; \quad \rho_t = \frac{\lambda_t}{\mu_t} \quad (4)$$

Com o cálculo de  $\rho$ , são informados intensidade de tráfego com relação aos pacotes de dados e voz.

$$E\{W\} = \frac{E\{T_0\}}{1 - \rho_t \frac{(x+y)(y+1)}{y(x+1)}} \quad (5)$$

Onde :  $E_{\{T_0\}}$  – Tempo gasto no sistema

$E_{\{w\}}$  – Tempo gasto de espera na fila

A fórmula 5 representa o cálculo de variação de atraso de pacotes de dados e voz sem prioridade. A relação é feita com os dois tipos de pacotes, pois no tipo de classe de serviço sem priorização, torna-se desnecessária a diferenciação do tráfego. Considerando-se o *buffer* infinito, a carga de pacotes UDP e o “gargalo” no processamento, possivelmente

provocado por uma demanda muito grande de terminais VoIP, o atraso pode ser considerável e o limite tolerável para o tráfego de voz não respeitado. Veremos em nossa simulação uma demonstração clara sobre a interferência do atraso.

- Classe de serviço **Com Prioridade**:

A classe de serviço com prioridade, atribui uma propriedade de escalonamento de tráfego, onde existe um reconhecimento do tipo do tráfego (voz ou dados) e um algoritmo de priorização é empregado sobre os pacotes de voz, diminuindo assim o impacto sobre o atraso provocado pelas filas nos elementos de roteamento do circuito. Conforme discutido no capítulo 2, na prática são aplicadas técnicas de serviços diferenciados (Diffserv) entre cliente e servidores. Veremos então o modelo matemático (fórmulas 5 e 7) para o cálculo de variação de atraso para os pacotes de dados e voz.

$$E\{W_v\} = \frac{E\{T_0\}}{1 - \rho \frac{x(y+1)}{y(x+1)}} \quad (6)$$

$$E\{W_d\} = \frac{E\{T_0\}}{\left(1 - \rho \frac{x(y+1)}{y(x+1)}\right) \left(1 - \rho \frac{(x+y)(y+1)}{y(x+1)}\right)} \quad (7)$$

Onde:  $E_{\{W_v\}}$  – Tempo médio de espera na fila de pacotes de voz.

$E_{\{W_d\}}$  - Tempo médio de espera na fila de pacotes de dados.

Em ambas as classes de serviços, com e sem prioridade, faremos ainda uma análise em três casos diferentes de formas de disposição de entrada de pacotes na fila. Essas formas interferem diretamente no resultado de uma determinada avaliação quando dizemos que os pacotes entram na fila de forma periódica ou mesmo em “rajadas” (REZENDE, 1999). Tais

processos de chegada de pacotes podem ser considerados não muito realistas, na verdade o processo de chegada deve ser considerado *aleatório*, sem um modelo específico. Apesar disso estaremos trabalhando em nosso modelo teórico com processos de chegada de pacotes em *exponencial negativa e pacotes com comprimento fixo*, tanto para voz, como para dados.

Seguem abaixo os três casos propostos com os modelos representativos de chegada de pacotes na fila. Nas r fórmulas 8, 9 e 10 calculamos o tempo gasto no sistema ( $E_{\{T_0\}}$ ), considerando as variações de intensidade de tráfego, a porcentagem de pacotes da taxa total de entrada de dados e voz e a taxa saída de pacotes da fila.

1º Caso: Exponencial negativa para voz e para dados.

$$E\{T_0\} = \frac{(y+1)^2(y^2+1)}{y^2(x+1)} \frac{\rho}{\mu} \quad (8)$$

2º Caso: Comprimento fixo para voz e para dados.

$$E\{T_0\} = \frac{(y+1)^2(y^2+1)}{2y^2(x+1)} \frac{\rho}{\mu} \quad (9)$$

3º Caso: Comprimento fixo para voz e exponencial negativa para voz e para dados.

$$E\{T_0\} = \frac{(y+1)^2(2y^2+1)}{2y^2(x+1)} \frac{\rho}{\mu} \quad (10)$$

Exibiremos a seguir, modelos gráficos representativos do cálculo de tempo médio de espera na fila, considerando os três casos de processo entrada de pacotes. Para tanto, devemos fornecer algumas variáveis como referência, para que possamos chegar a um resultado teórico que condiz com uma possível situação real. Teremos a oportunidade de efetuar simulações



que irão comprovar o resultado desse modelo teórico. Seguem abaixo as variáveis necessárias para o planilhamento gráfico.

Logo informamos:

$C$  – Taxa de transmissão – 512 kbps

$L - (1/\mu_d)$  – comprimento do pacote (*bits/pacote*) = 1000 *bits/pacote*

$X$  – Porcentagem de pacotes de voz com relação a  $\lambda_t = 0,1 = 10\%$

$Y$  - Porcentagem de pacotes de dados com relação a  $\lambda_t = 0,9 = 90\%$

Para cada caso, apresentamos uma tabela de variação de  $\lambda_t$ , onde informamos a taxa total de entrada de pacotes na fila. Dessa forma, o leitor poderá fazer uma breve análise e chegar à conclusão que o modelo matemático condiz exatamente com o que foi apresentado na Figura 15, quando a intensidade de tráfego tender a 1, poderá ocorrer uma situação em que a velocidade de chegada excederá a capacidade de transmissão, aumentando o comprimento médio da fila. Neste caso, à medida que a intensidade do tráfego se aproxima de 1, o comprimento médio da fila vai ficando maior.

Tabela 3: Variação de dados para geração de gráfico do modelo teórico – 1º Caso.

1º Caso					
$\lambda_t$	$\rho_t$	$E\{T_o\}$	$E\{W\}$	$E\{W_v\}$	$E\{W_d\}$
10	0,019531	0,000141	0,000146	0,000141	0,000147
30	0,058594	0,000422	0,000475	0,000427	0,000481
50	0,097656	0,000703	0,000865	0,000717	0,000882
70	0,136719	0,000985	0,001335	0,001011	0,001371
90	0,175781	0,001266	0,00191	0,00131	0,001977
110	0,214844	0,001547	0,002633	0,001614	0,002746
130	0,253906	0,001828	0,003566	0,001922	0,003749
150	0,292969	0,00211	0,00482	0,002235	0,005107
170	0,332031	0,002391	0,006591	0,002554	0,00704
190	0,371094	0,002672	0,009285	0,002877	0,009997
200	0,390625	0,002813	0,011238	0,003041	0,012148
220	0,429688	0,003094	0,017646	0,003372	0,019232
240	0,46875	0,003376	0,033628	0,003709	0,036952
250	0,488281	0,003516	0,055906	0,00388	0,061687

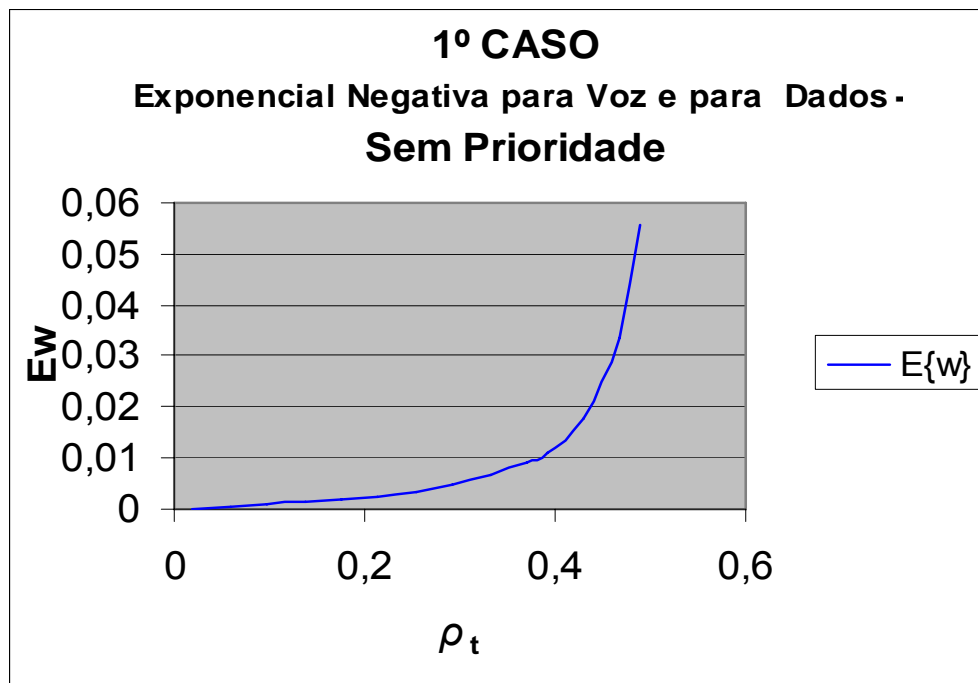


Figura 17: 1º Caso – Exp. Neg. Voz/Dados – Sem Prioridade

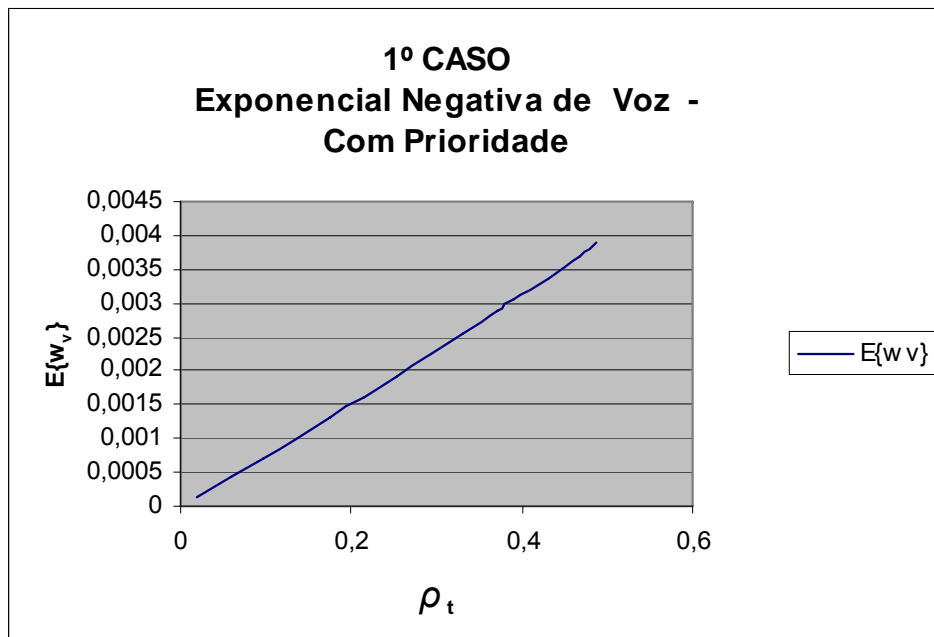


Figura 18: 1º Caso – Exp. Neg. Voz – Com Prioridade

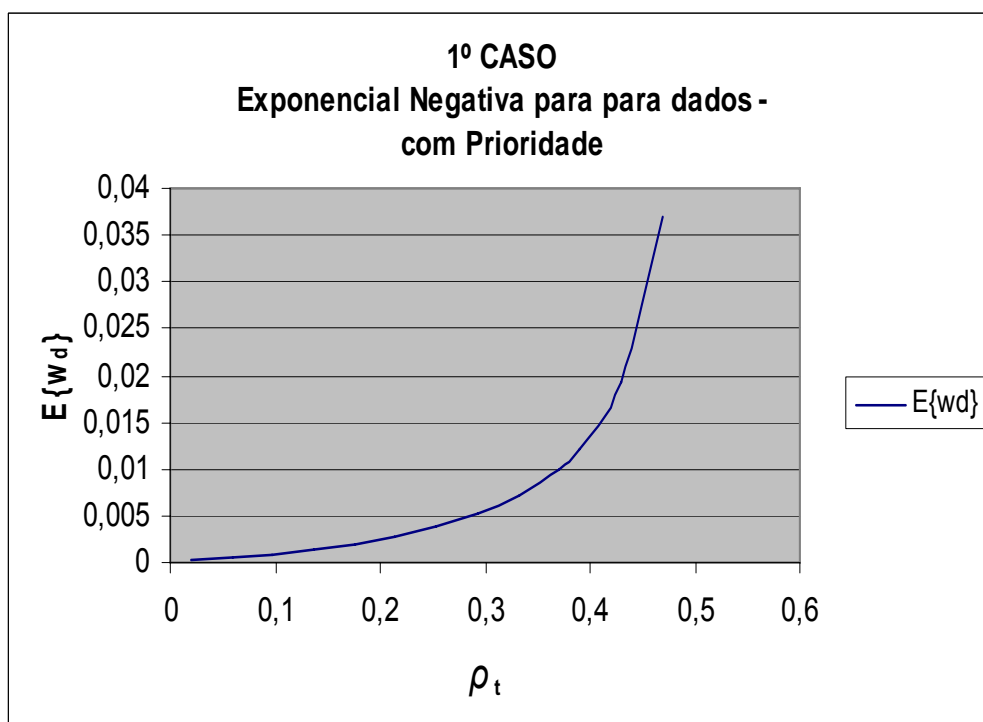


Figura 19: 1º Caso – Exp. Neg. Dados – Com Prioridade.

Tabela 4: Variação de dados para geração de gráfico do modelo teórico – 2º Caso.

2º Caso					
$\lambda_t$	$\rho_t$	$E\{T_o\}$	$E\{W\}$	$E\{W_v\}$	$E\{W_d\}$
10	0,019531	7,03E-05	7,30623E-05	7,0588E-05	7,33372E-05
30	0,058594	0,000211	0,000237701	0,00021337	0,000240404
50	0,097656	0,000352	0,000432718	0,00035833	0,000440983
70	0,136719	0,000492	0,000667378	0,00050553	0,000685361
90	0,175781	0,000633	0,000955134	0,00065501	0,000988482
110	0,214844	0,000774	0,001316308	0,00080683	0,001372917
130	0,253906	0,000914	0,001783104	0,00096104	0,001874444
150	0,292969	0,001055	0,00240979	0,0011177	0,002553356
170	0,332031	0,001196	0,003295494	0,00127687	0,003519786
190	0,371094	0,001336	0,004642629	0,00143861	0,004998632
200	0,390625	0,001406	0,005618792	0,00152046	0,006074162
220	0,429688	0,001547	0,008823172	0,00168617	0,009616173
240	0,46875	0,001688	0,01681397	0,00185461	0,018476121
250	0,488281	0,001758	0,027952961	0,00193988	0,030843304

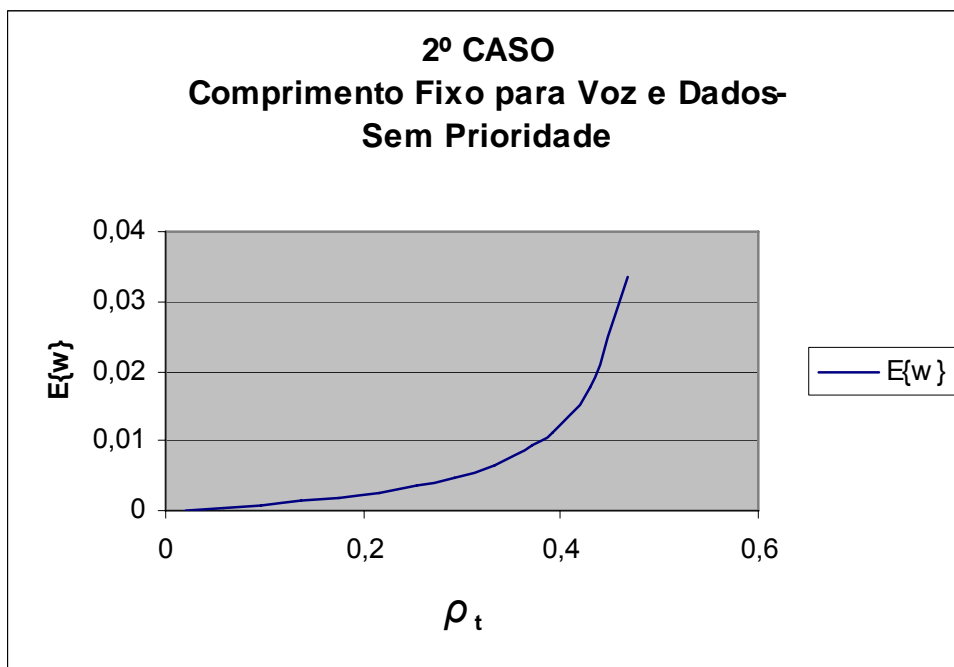


Figura 20: 2º Caso – Comprimento Fixo para Voz e dados – Sem Prioridade.

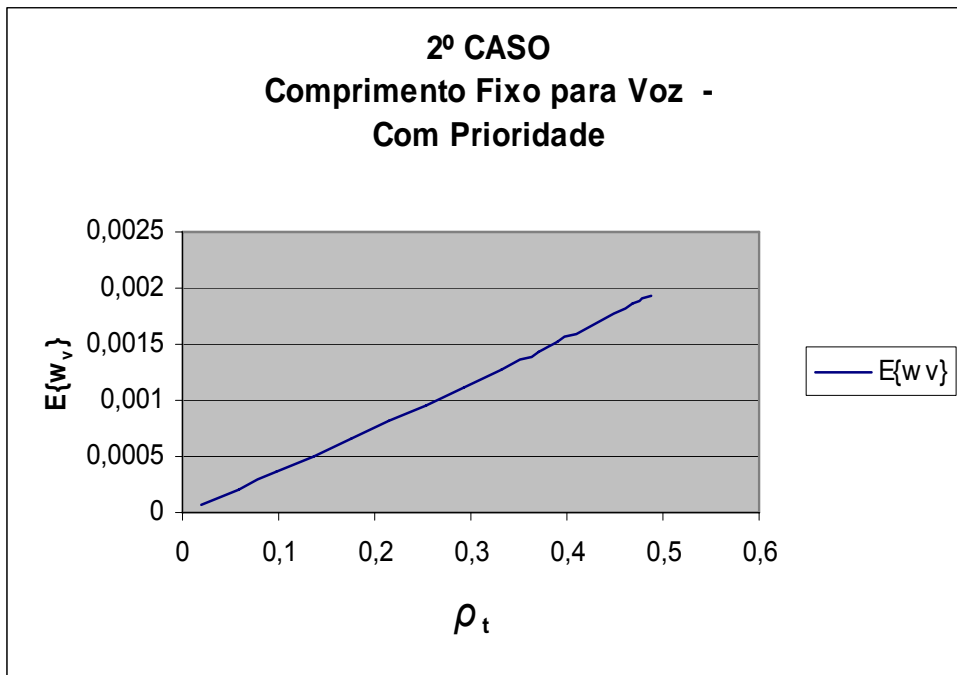


Figura 21: 2º Caso – Comprimento Fixo para Voz – Com Prioridade.

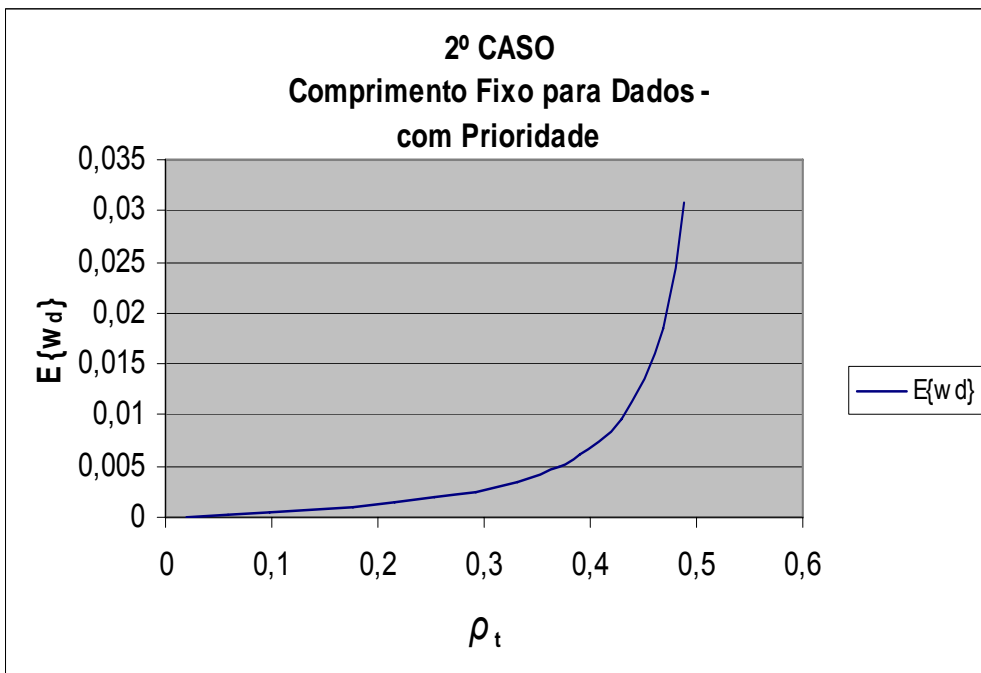


Figura 22: 2º Caso – Comprimento Fixo para Dados – Com Prioridade.

Tabela 5: Variação de dados para geração de gráfico do modelo teórico – 3º Caso.

3º Caso					
$\lambda_t$	$\rho_t$	$E\{T_o\}$	$E\{W\}$	$E\{W_v\}$	$E\{W_d\}$
10	0,019531	0,000133	0,000138096	0,0001334	0,000138615
30	0,058594	0,000399	0,000449281	0,0004033	0,00045439
50	0,097656	0,000665	0,000817885	0,0006773	0,000833507
70	0,136719	0,00093	0,001261418	0,0009555	0,001295408
90	0,175781	0,001196	0,001805309	0,001238	0,001868339
110	0,214844	0,001462	0,002487966	0,001525	0,002594964
130	0,253906	0,001728	0,003370262	0,0018165	0,003542906
150	0,292969	0,001994	0,004554768	0,0021126	0,004826123
170	0,332031	0,00226	0,006228846	0,0024134	0,006652783
190	0,371094	0,002525	0,008775079	0,0027191	0,009447963
200	0,390625	0,002658	0,010620134	0,0028738	0,011480834
220	0,429688	0,002924	0,016676765	0,003187	0,018175623
240	0,46875	0,00319	0,031780251	0,0035054	0,034921899
250	0,488281	0,003323	0,052834169	0,0036666	0,058297234

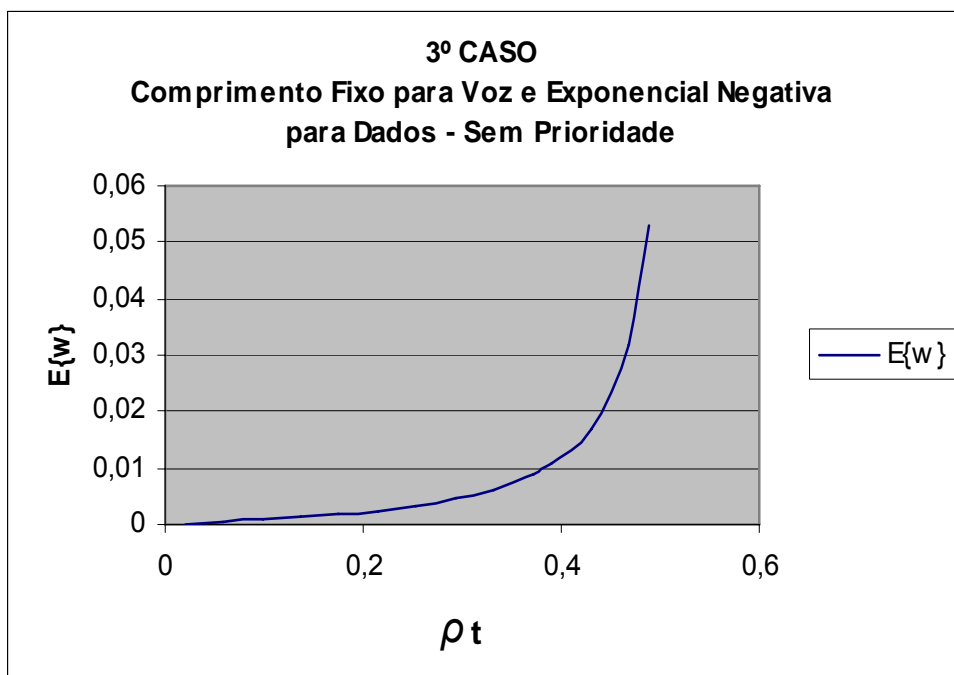


Figura 23: 3º Caso – Comprimento Fixo p/ Voz e Exponencial Negativa p/ Dados – Sem Prioridade.

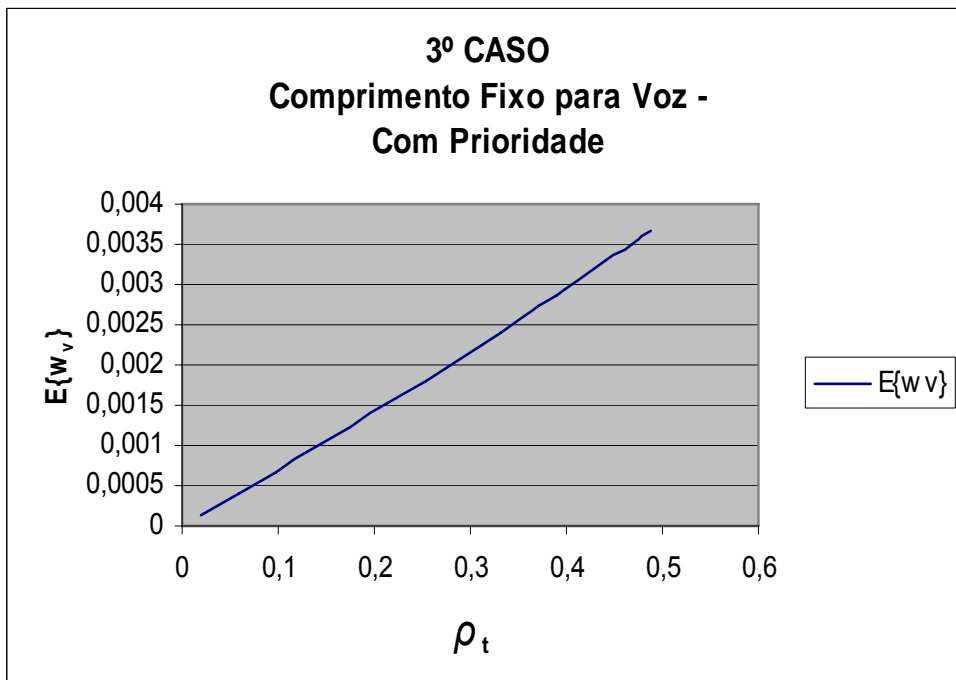


Figura 24: 3º Caso – Comprimento Fixo p/ Voz e Exponencial. Negativa p/ Dados – Com Prioridade.

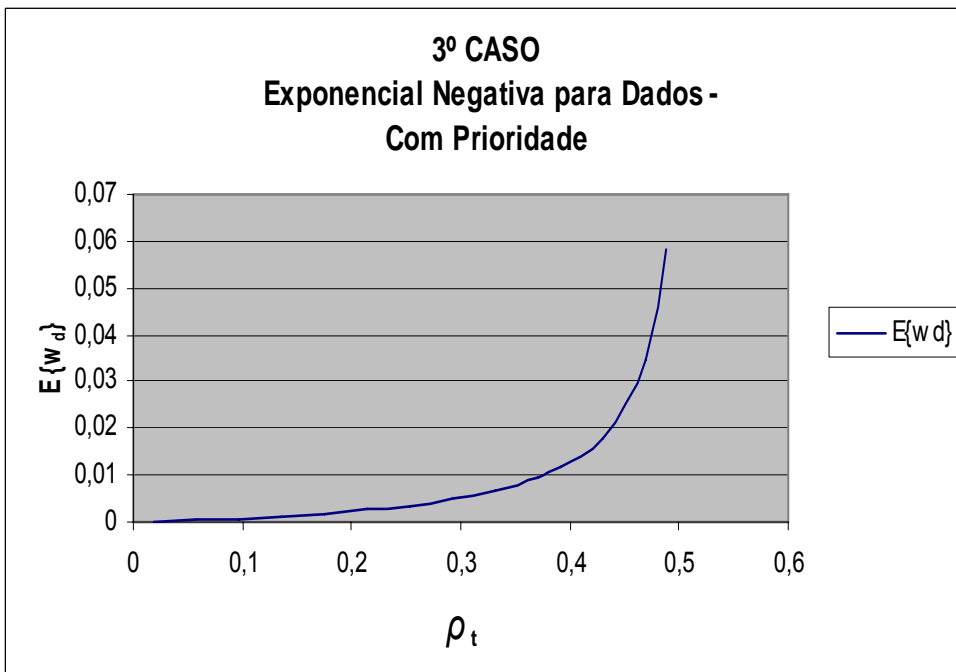


Figura 25: 3º Caso – Exponencial Negativa p/ Dados – Com Prioridade.

Concluimos assim uma análise matemática, conforme demonstrado nos gráficos acima, o modelo comportamental do atraso de pacotes de voz, considerando-se o tempo de permanência do pacote na fila em função da intensidade do tráfego – o mesmo resultado poderia ser obtido colocando-se o tempo de permanência na fila, em função da taxa de entrada de pacotes de dados e voz – onde as amostragens definidas através dos métodos de distribuição do tráfego (Exponencial Negativa / Comprimento Fixo, Com / Sem Prioridade), demonstrados nos três casos propostos, poderíamos discutir inúmeras formas para chegarmos em resultados diversos, porém conseguimos concluir a análise com os modelos distribuição apresentados acima.

Podemos observar no 1º Caso (Exponencial negativa para voz e para dados, Figuras 17, 18 e 19), quando não estabelecemos a prioridade, devemos analisar o tráfego total, ou seja, os 10 % de pacotes de voz sendo transmitidos juntamente com os 90 % de pacotes de dados. Isso nos leva a concluir que à medida que aumentamos a intensidade dos tráfegos, tanto de dados como de voz, o tempo de espera na fila também aumenta. Ainda no 1º Caso, porém agora avaliando-se a distribuição em Exponencial negativa (Tabela 3) apenas para voz, pois com a aplicação da prioridade para voz, devemos avaliar o comportamento do tráfego para cada tipo de pacote transmitido, como descrito nas Figura 18 do 1º Caso, onde o tempo de permanência do pacote na fila também aumenta com o crescimento da intensidade do tráfego, porém em uma forma mais constante, devido à porcentagem de pacotes de voz (10%) em relação aos pacotes de dados (90%).

Nos demais casos (conforme Figuras 21, 22, 23, 24 e 25) os comportamentos dos gráficos são bem parecidos, porém o leitor deve observar naturalmente, o tempo em que o pacote permanece na fila (Tabelas 3, 4 e 5), de acordo com a intensidade do tráfego estabelecida e a aplicação da prioridade para pacotes de dados e voz.



No capítulo 3, voltaremos a discutir o modelo teórico, porém executando a simulação em software, sobre cada um dos modelos propostos. Uma breve apresentação do software aplicado, o *Network Simulator* (NS), bem como os critérios definidos para essa simulação. Finalmente, serão apresentadas apurações dos resultados e uma análise comparativa do modelo teórico com a simulação.

## CAPÍTULO 3 – SIMULAÇÃO DO MODELO TEÓRICO

Para comprovarmos o resultado efetivo do modelo matemático, atribuímos o uso do simulador de redes *Network Simulator – NS* (VINT, 2003) versão 2, que está sendo desenvolvido dentro do projeto VINT, por algumas universidades e centros de pesquisas americanas. O *NS* possui funcionalidades específicas para simular a Internet, o que o torna uma ferramenta poderosa para configurar rapidamente simulações complexas e também, para comparação de resultados de pesquisas.

Veremos nas próximas seções uma descrição mais detalhada das ferramentas de simulação e análise utilizadas em nosso estudo.

### 3.1 O NS (Network Simulator)

O *Network Simulator* (NS) é um software de simulação de redes (NS, 2003). Ele trabalha na simulação de protocolos como o TCP, protocolos de multicast em redes com ou sem fio e simulando até mesmo roteamento de pacotes, entre outros recursos. O *Network Simulator* é um excelente simulador de redes locais (LANs) e WANs.

A primeira versão do NS surgiu em 1989, provinda do software *REAL Network Simulator*. Com o apoio do DARPA, em 1995 este software passou a ser difundido pelo projeto VINT na LBL, Xerox PARC, UCB, e USC/ISI (DARPA/NIST, 2004). Atualmente o desenvolvimento do NS é apoiado pelo DARPA com o SAMAN e pelo NSF com o CONSER, ambos em colaboração com outros pesquisadores incluindo o ACIRI. O NS também sempre inclui contribuições substanciais de outros pesquisadores, incluindo códigos de wireless do UCB Daedalus e dos projetos do CMU Monarch e da Sun Microsystems.

Atualmente na versão 2.1, o NS é um software desenvolvido na Universidade de Berkeley em C++ e OTcl (linguagem de script TCL com extensão orientada a objeto desenvolvido no MIT), sendo orientado a objeto.

O NS é um simulador de redes dirigido por eventos que simula vários tipos de redes IP. Ele pode trabalhar simulando protocolos de rede como o TCP e o UDP, comportamento de tráfego em FTP, Telnet, Web, CBR e VBR, mecanismos de gerenciamento de filas de roteadores como DropTail, RED e CBQ, algoritmos de roteamento como o Dijkstra e outros. O NS também implementa multicasting e alguns dos protocolos da camada MAC para simulação de LANs. O projeto NS é agora parte de projeto VINT que desenvolve ferramentas para mostrar resultados de simulações, análises e conversores que convertem topologias de redes geradas por geradores bem conhecidos para o formato NS.

A Figura 26, descreve uma visão geral sobre a aplicação e uso do NS, com seus scripts OTcl, interpretadores, bibliotecas, formas de apresentação de resultados e análises e ferramentas de animação e gráficos para avaliação de resultados.

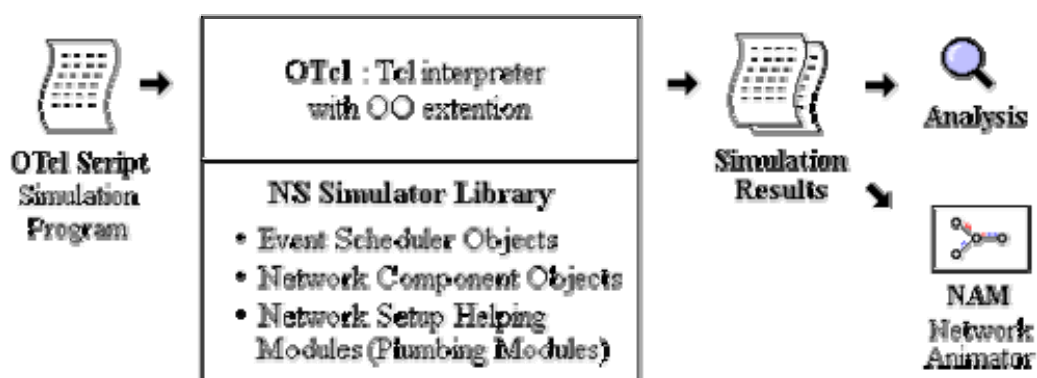


Figura 26: Visão Geral sobre aplicação e uso do NS (NS,2003).

Para uma visão simplificada, o NS é um interpretador de script Tcl (OTcl) orientado a objeto e uma biblioteca, a qual contém objetos de controlador de evento, objetos de componentes de rede e módulos de ajuda de configuração de rede. A programação em NS baseia-se em desenvolvimento de scripts em OTcl.

Em uma programação em NS, o usuário deve escrever um script OTcl que inicia um controlador de evento, configura a topologia da rede utilizando os objetos de rede e funções de ligações às bibliotecas, as quais definem quando as fontes de tráfego serão iniciadas. As ligações são utilizadas para configurar a rede, pela configuração do ponteiro do "vizinho" de um objeto para o endereço de um objeto apropriado (conforme modelo descrito na Figura 26). Quando um usuário quer construir um novo objeto de rede, este pode facilmente fazê-lo construindo um novo objeto ou utilizando um objeto de uma biblioteca, e ligando o caminho de dados pelo objeto.

Um controlador de eventos, lançado através de um ID de pacote, é um outro componente do NS que reside nos objetos de rede do controlador de evento. Esse controlador de eventos garante que um pacote seja único com um tempo de escalonamento e um ponteiro para o objeto que controla o evento. Um controlador de evento coloca o evento na fila de eventos escalonados que faz a chamada aos componentes de rede apropriados no tempo pré-definido. A comunicação dos componentes de rede se dá por passagem de pacotes, contudo isto não consome tempo de simulação real.

Todos os componentes de rede que precisam gastar algum tempo de simulação segurando o pacote (isto é, necessita de um delay) utilizam o controlador de evento para isso. Por exemplo, um componente de chaveador de rede que simula um chaveador com 20 microssegundos de atraso de chaveamento associado a um evento para um pacote ser chaveado no controlador de evento como um evento de 20 microssegundos de atraso. O

controlador de evento depois de 20 microssegundos tira da fila o evento e grava-o no para o componente de chaveador, que então passa o pacote para um componente de ligação de saída apropriado.

Um outro uso do controlador de evento é o timer. Por exemplo, o TCP precisa de um timer para manter o tempo de retransmissão do pacote (transmissão do pacote no mesmo número de pacote TCP, mas diferente ID do pacote NS). Os timers utilizam controladores de evento em um modo similar ao delay. A diferença é que o timer mede o valor do tempo associado com um pacote e a relacionada ação apropriada para o pacote depois de um certo tempo, e não simula um delay. Na Figura 27, enumera alguns modelos de objetos básicos em linguagem OTcl e C++.

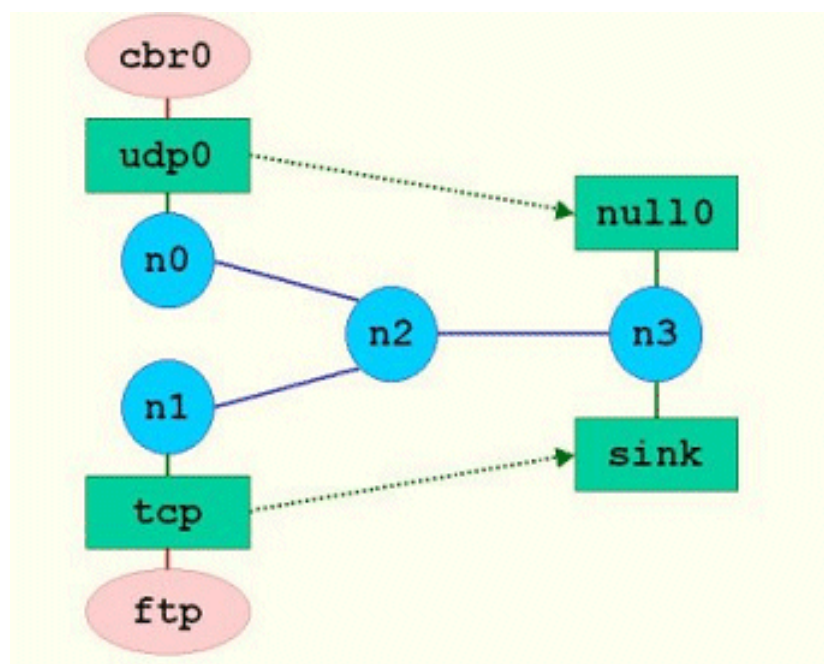


Figura 27: Modelos de componentes básicos (NAM, 2004).

Uma composição de objetos escritos em OTcl e C++ estabelece a estrutura do NS. As características de performance de processamento do NS são definidas por estruturas de implementações de caminhos de dados e caminhos de controle, onde os caminhos de dados

(objetos de componentes de redes básicos) são escritos e compilados em C++, para redução de tempo de processamento. Tais objetos são colocados à disposição para o interpretador de OTcl por uma ligação OTcl que cria uma correspondência do objeto OTcl para cada um dos objetos C++, fazendo com que as funções de controle e as variáveis de configuração especificadas pelo objeto C++ se comportem como funções de membros e variáveis de membros de objetos OTcl correspondentes. Desta forma, os controles de objetos C++ são dados pelo OTcl. Isto também é possível para adicionar funções de membros e variáveis para o C++ ligados aos objetos OTcl. O objeto em C++ que não é necessário ser controlado na simulação ou internamente utilizado por um outro objeto, não precisa ser ligado ao OTcl. A Figura 28 mostra um exemplo de hierarquia de objetos em C++ e OTcl.

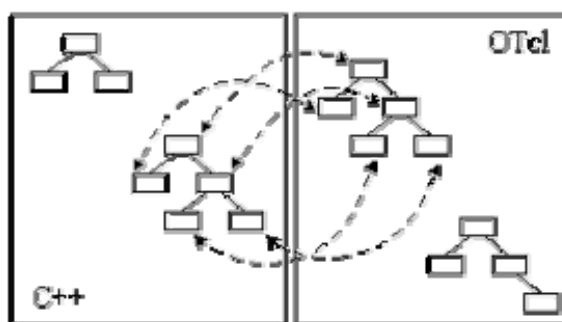


Figure 28: C++ e OTcl: A Dualidade (NS,2003)

A Figura 29 mostra a arquitetura geral do NS. Nesta figura um usuário qualquer (não um desenvolvedor do NS) pode estar no canto mais à esquerda e abaixo, planejando e executando simulações em Tcl utilizando os objetos de simulação das bibliotecas OTcl. Os schedulers de evento e a maioria das componentes de rede são implementados em C++ e disponíveis em OTcl pela ligação que é implementada utilizando tclcl.

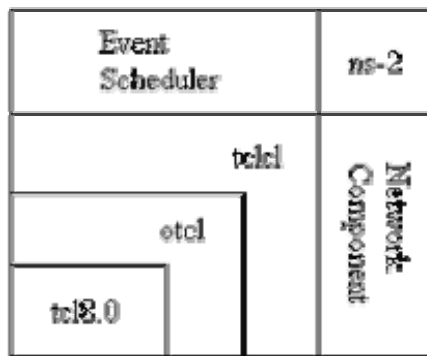


Figure 29: Visão da Arquitetura do NS (NS, 2003)

Como resultado de uma simulação, o NS gera um ou mais arquivos de saída baseados em texto que contém dados detalhados da simulação. Os dados podem ser utilizados para análise de simulações ou como entrada para uma ferramenta de simulação gráfica chamada Network Animator (NAM) que é desenvolvidos com uma parte do projeto VINT. O NAM tem uma interface gráfica de usuário bastante amigável e tem um controlador de velocidade. Com isso pode apresentar graficamente informações presentes como throughput e números de pacotes descartados por cada link, contudo as informações gráficas não podem ser utilizadas para análise de simulações precisas.

Abaixo seguem ilustrações do NAM Editor e de um modelo de aplicação em execução no animador de simulação (NAM, 2004). Na Figura 30, é apresentado um modelo de simulação com quatro *nodes*, um cenário que representa uma situação onde fontes FTP(*File Transfer Protocol*), conectadas a roteadores por enlace variando entre 1Mbps a 2Mbps com um atraso de 10ms.

Para este modelo de simulação considera-se que todo o tráfego é de Melhor Esforço com filas gerenciadas pelo mecanismo *Drop Tail* (FIFO), como uma melhor forma para modelar a Internet atual. Da mesma forma estaremos estabelecendo a simulação do modelo matemático proposto em nosso estudo, criando parâmetros nas mesmas condições de origem de dados e voz.

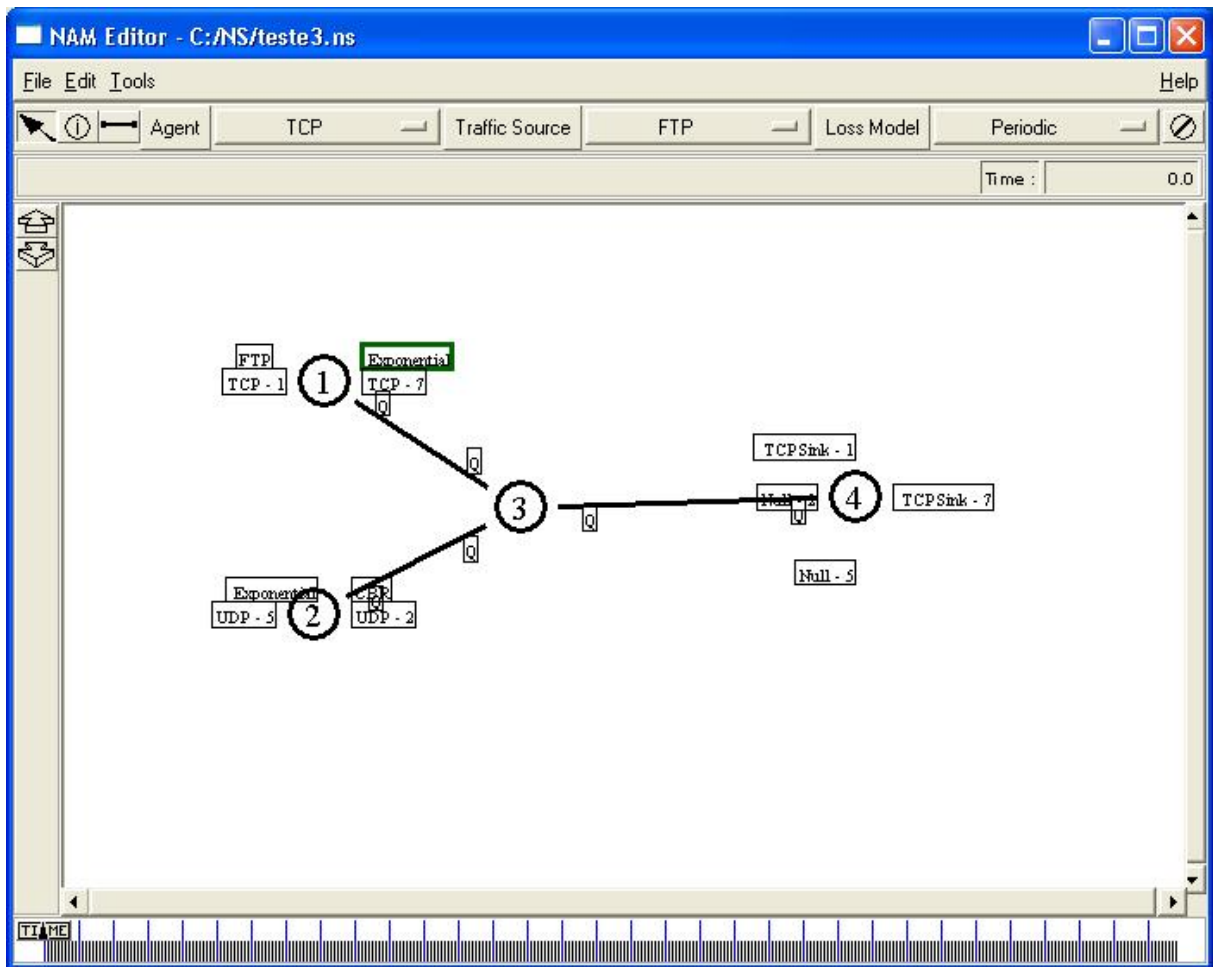


Figura 30: NAM Editor com uma implementação de um cenário de simulação.

Na Figura 31, é apresentado o modelo de animação do cenário previamente criado, demonstrando os processos de geração de pacotes de dados e voz, encaminhados para um nó concentrador onde os pacotes são colocados em fila (FIFO), considerando-se o *buffer* infinito e o histórico todo o tráfego – incluído os processos de envio e recebimento de cada pacote – sendo armazenado em um arquivo *trace*, utilizado para uma análise de resultados, com a aplicação de gráficos tais como *XGRAPH* e *TRACEGRAPH*, conforme descrito posteriormente.



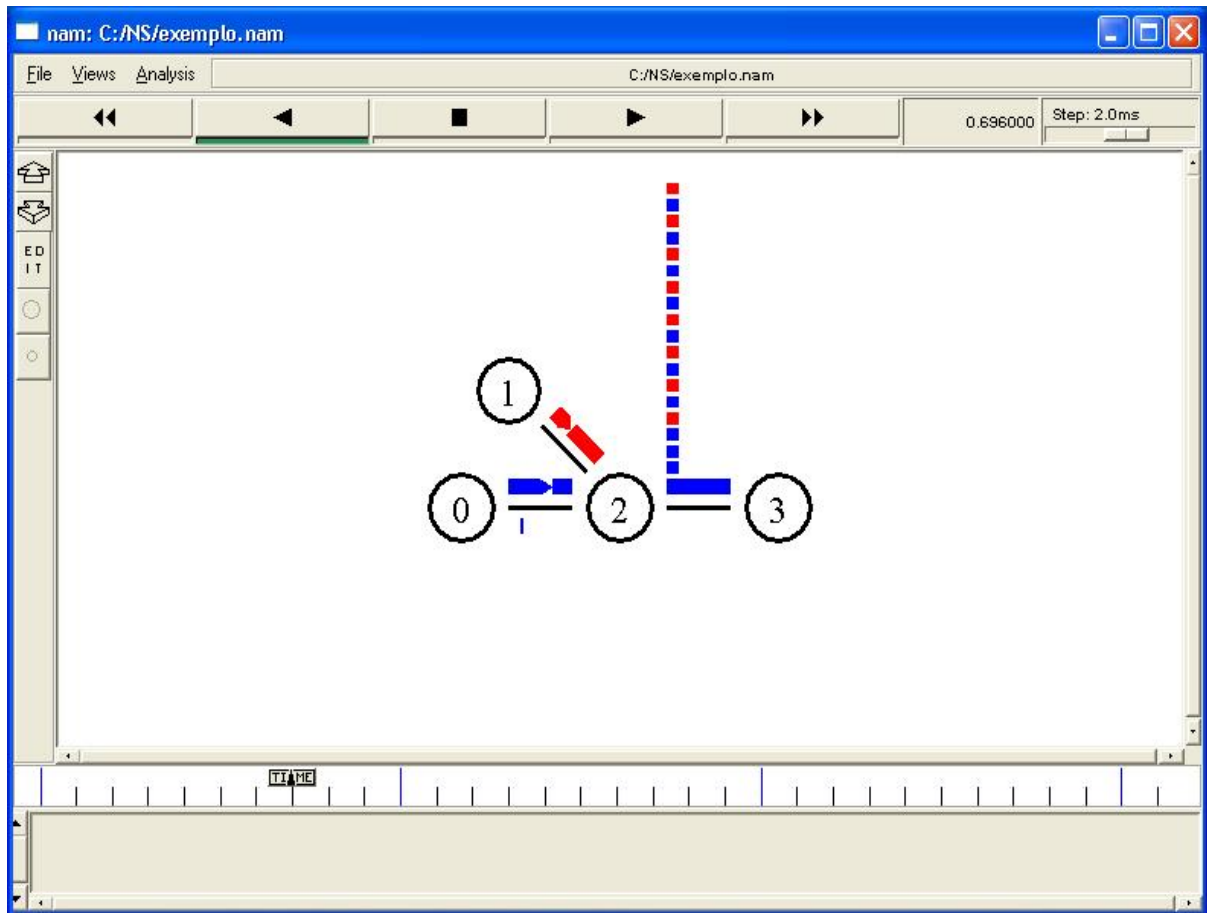


Figura 31: Animação de cenário de rede.

Uma descrição detalhada dos componentes da ferramenta de animação é apresentada na Figura 32, mostrando de forma clara, todos os seus recursos, dando ênfase aos procedimentos de análise comportamental de maneira visual, proporcionando uma avaliação mais objetiva e mais próxima a uma implementação real.

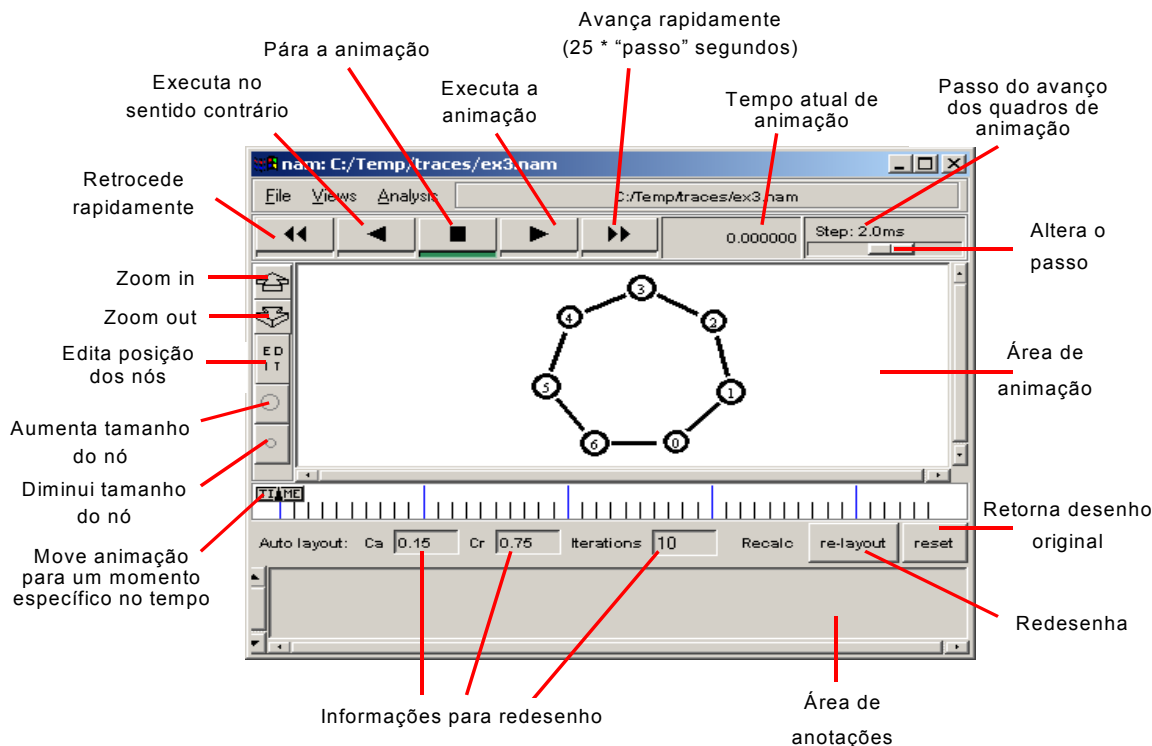


Figura 32: Descrição detalhada dos componentes da ferramenta de animação NAM.

Na próxima seção iniciaremos o processo de simulação do nosso modelo matemático, objeto do nosso estudo. Uma proposta de topologia, a definição de scripts em OTcl e uma apuração dos resultados obtidos, farão parte do processo de conclusão do estudo.

### 3.2 Topologia da simulação

A topologia adotada para a simulação é um modelo fictício. Estabelecendo quatro nós de roteamento, caracterizando pontos de entrada de tráfego (dados e voz), um nó concentrador e um último ponto de recepção dos sinais enviados. O cenário representa uma situação onde fontes FTP (*File Transfer Protocol*) estão conectadas a roteadores por enlaces de 512Kbps com um atraso de 10ms. Os roteadores 3 e 4 estão também interligados por um enlace de 512 Kbps que representa um enlace ponto-a-ponto E1 de longa distância, com um atraso de 0ms representando o “gargalo” na topologia, como ilustra a Figura 33.

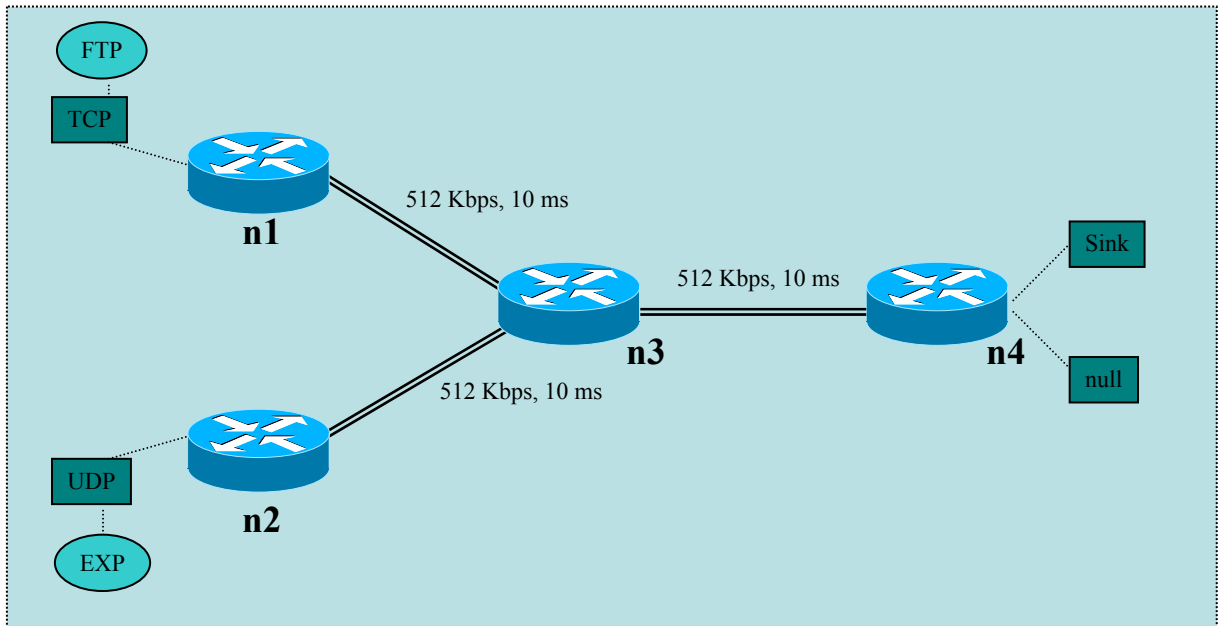


Figura 33: Modelo de topologia aplicada à simulação.

Para esta simulação considera-se que todo o tráfego é de Melhor Esforço com filas gerenciadas pelo mecanismo *Drop Tail* (FIFO), como uma melhor forma para modelar a Internet atual (BRADY, 1999).

Neste cenário é representada uma situação onde uma conexão de voz é estabelecida e fontes-origem enviam pacotes FTP para fontes-destino para gerar o tráfego da rede, representado o fluxo em *background* na rede. O tráfego de voz foi parametrizado em pacotes UDP de acordo com as aplicações de áudio na Internet .

### 3.3 Aplicação e comportamento do usuário de voz

Para simular um usuário de voz serão utilizados pacotes UDP com fontes ON/OFF (pareto) com o tempo ocioso (*idle time*) sendo distribuído exponencialmente, o que caracteriza melhor uma conversação na Internet (KAMEINSKI, 2000). Com uma média ON (fonte ativa) e média OFF (fonte desativada).

### 3.4 Métricas de avaliação

Basicamente foram utilizadas duas métricas para avaliar o atraso do pacote de voz em uma rede IP sob sinais multiplexados:

- Atraso (*delay*) : representado pelo atraso fim a fim entre fonte e destino;
- Quantidade de carga útil: taxa de dados úteis que estão sendo recebidos pelo receptor.

Entende-se por dados úteis ou carga útil, apenas os *frames* de voz de um pacote, excluindo-se os cabeçalhos.

### 3.5 Resultados Obtidos

Para obtermos uma avaliação dos dados gerados pelo *NS*, através dos arquivos *trace*, conforme descrito na seção 3.1, fizemos uso da ferramenta de análise de resultados de tráfego do *NS*, o *TRACEGRAPH* para geração de gráficos do resultado simulado e finalmente compará-los com o modelo teórico.

O *TRACEGRAPH* é uma poderosa ferramenta gráfica (Malek, Jaroslaw / *Wroclaw University of Technology*), especialmente criada para analisar resultados obtidos por simulações do *NS*. Com avaliações e tipos de gráficos pré-definidos, assim como o *XGRAPH*, o *TRACEGRAPH* possui formas variadas (2D, 3D e histogramas) de exibições que auxiliam na avaliação e apuração de resultados de arquivos *trace*.

Inicialmente apresentaremos os gráficos dos dados gerados pelo *NS*, de acordo com a topologia e métricas citadas anteriormente, as quais condizem com a mesma arquitetura e parâmetros, métricas e de distribuição de entrada de dados. Foram utilizados *nodes* e módulos geradores de tráfego que simulam as características dos pacotes os quais queremos avaliar.

No *TRACEGRAPH* trabalhamos com grandezas métricas proporcionais às discutidas no modelo teórico, como poderemos observar nos gráficos apresentados a seguir. É válido lembrar que as características dos resultados obtidos pela ferramenta foram mantidas, para garantir a integridade da avaliação, sem qualquer interferência nos cálculos ou mesmo nos processos.

Na Figura 34, o gráfico apresenta a média de atraso fim a fim em função da intensidade de tráfego durante todo o intervalo de duração da transmissão, em uma distribuição de entrada de pacotes em exponencial negativa para dados e voz. Para análise do tráfego, devem ser considerados pacotes de dados e voz sem priorização (QoS), o que torna a transmissão de uma comunicação de voz algo mais suscetível aos problemas gerados pelo atraso. O resultado foi avaliado com pacotes de 1000 bits, com o empacotamento utilizando a técnica de multiplexação propostos por Yamada (2001). Assim teremos uma carga útil melhorada em 30 % com relação aos outros métodos de empacotamento, porém é importante lembrar que, de nada adianta o aumento da eficiência da carga útil dos pacotes de voz, se aumentar demasiadamente o atraso geral dos pacotes. Devido a este fator, as medições dos atrasos finais dos pacotes transmitidos devem ser consideradas.

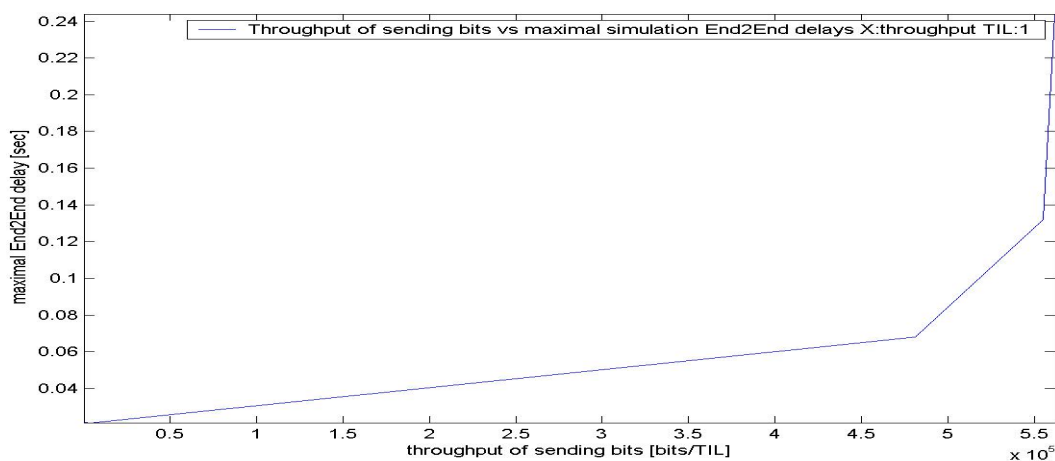


Figura 34: Simulação NS / *TRACEGRAPH* – Exp. Neg. p/ Voz e Dados s/ Prioridade.

Para o 2º caso, também sem priorização, com distribuição de pacotes de dados e voz, ambos com comprimento fixo, a média de atraso na fila tem seu valor reduzido pela metade (Figura 35), porém a intensidade de tráfego pode prejudicar o tempo médio de chegada de pacotes de voz. Sem a aplicação de QoS e a diferenciação dos tipos de serviços de entrega de sinais, o método de empacotamento mais eficiente não apresentará o resultado esperado. Essas comparações também podem ser avaliadas através das Tabelas 3, 4 e 5.

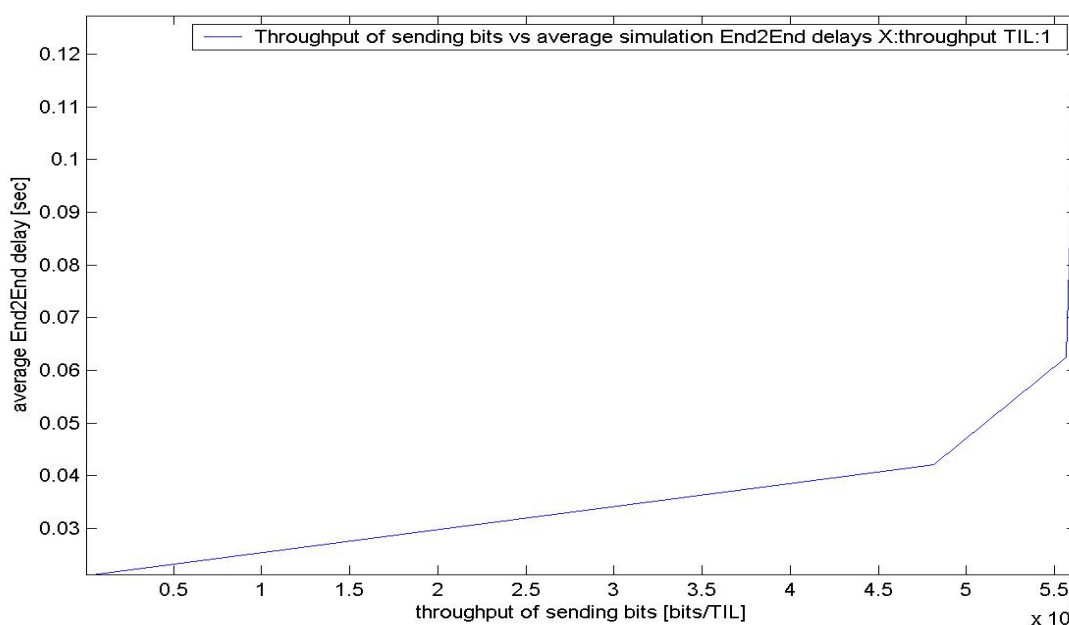


Figura 35: Simulação NS / *TRACEGRAPH* – Exp. Neg. p/ Voz e Compr. Fixo p/ Dados s/ Prioridade.

Ainda sem considerar o fator QoS, no 3º caso foi atribuído comprimento fixo para voz e exponencial negativa para dados. Com esse método de distribuição verificamos que a média de atraso fim a fim dobra seu valor com relação ao método aplicado no 2º caso. No modelo teórico, a variação do atraso médio na fila é dada pelo cálculo do tempo gasto no sistema  $E\{T_o\}$ , vinculado à taxa de pacotes de voz em relação aos pacotes de dados. Na Figura 36 é exibido o resultado da simulação do 3º caso comparado ao modelo matemático.

Nesse caso também cabe ao leitor uma breve avaliação comparativa pelas Tabelas 3, 4 e 5, onde pode ser visto uma pequena diferença entre os valores do 1º e 3º casos.

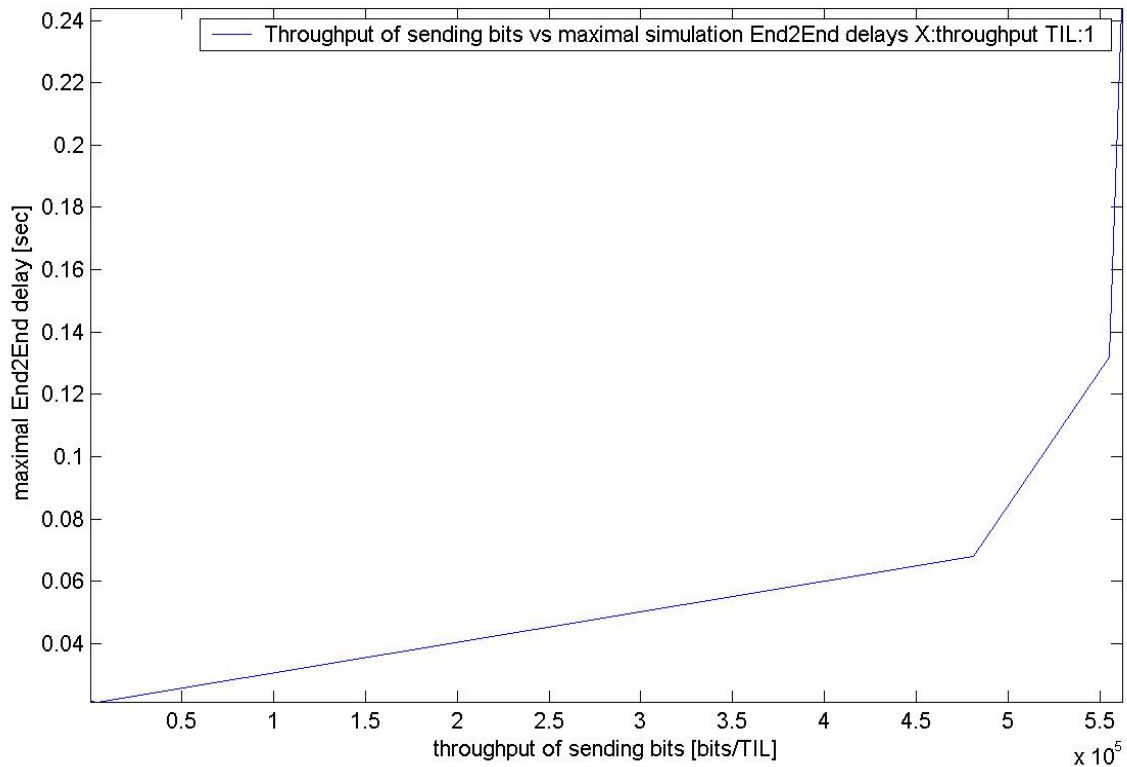


Figura 36: Simulação NS / *TRACEGRAPH* – Exp. Neg. p/ Dados e Compr. Fixo p/ Voz s/ Prioridade.

Nos casos das simulações com prioridade, podemos verificar que a aplicação de QoS e da diferenciação dos tipos de serviços (*diffserv*) influenciam consideravelmente sobre as formas de distribuição de tráfego. No modelo matemático são tratados separadamente os pacotes de dados e os pacotes de voz - da mesma forma foram avaliados na simulação, aplicado sobre o tráfego de voz, representando diferentes casos de distribuição de chegada, bem como a atribuição de um algoritmo de priorização de pacotes (PQ) em uma determinada fila.

No 1º caso com prioridade, com a chegada de pacotes de dados e voz em exponencial negativa, é observado que o tráfego de voz tem um atraso crescente em função da intensidade do tráfego. A priorização garante que o tráfego de voz não sofra a interferência dos pacotes de dados. A implementação de uma *Priority Queue* (PQ) no NS é executada sobre um canal DropTail, para definição de filas (FIFO), com pacotes UDP de fontes ON/OFF. Para efeito de implementação simulada, os *codecs* são representados por essas fontes de geração de pacotes.

Se compararmos o resultado simulado (Figura 37) com o modelo teórico (Figura 18), veremos que o comportamento dos gráficos representando os pacotes de voz e os pacotes de dados (Figura 38) e seus respectivos valores, são proporcionalmente iguais. Para os demais casos, o comportamento do tráfego não deve ocorrer de forma diferente para os pacotes de voz, porém com métodos de distribuição como do 2º caso – comprimento fixo de pacotes de voz (Figura 39) e comprimento fixo de pacotes de dados (Figura 40) – obtemos um melhor resultado, com um menor tempo de espera na fila.

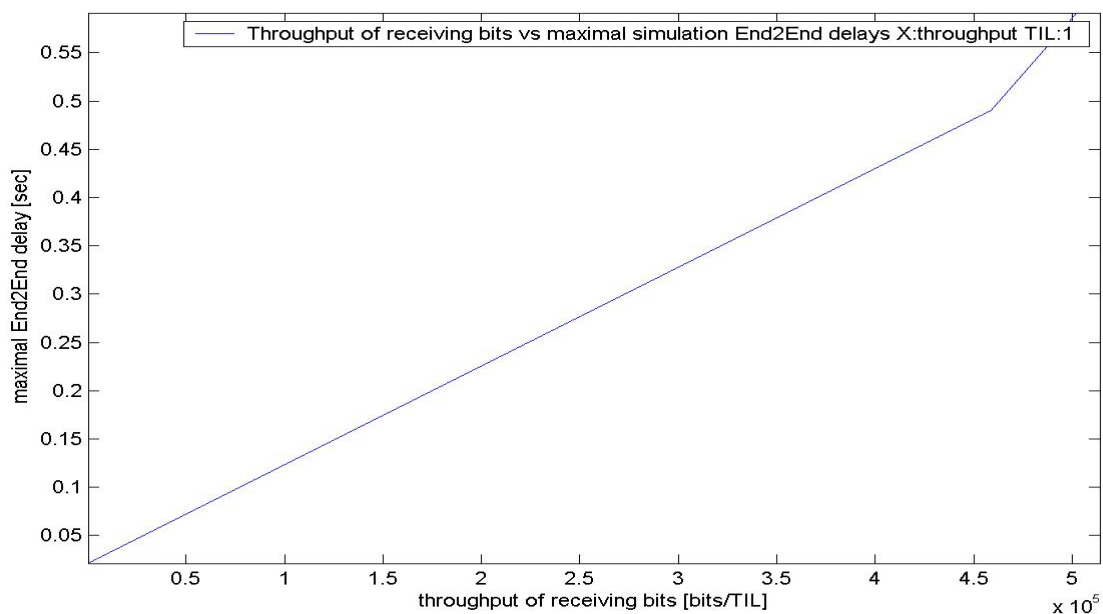


Figura 37: Simulação NS / TRACEGRAPH – Exp. Neg. p/ Voz com Prioridade.



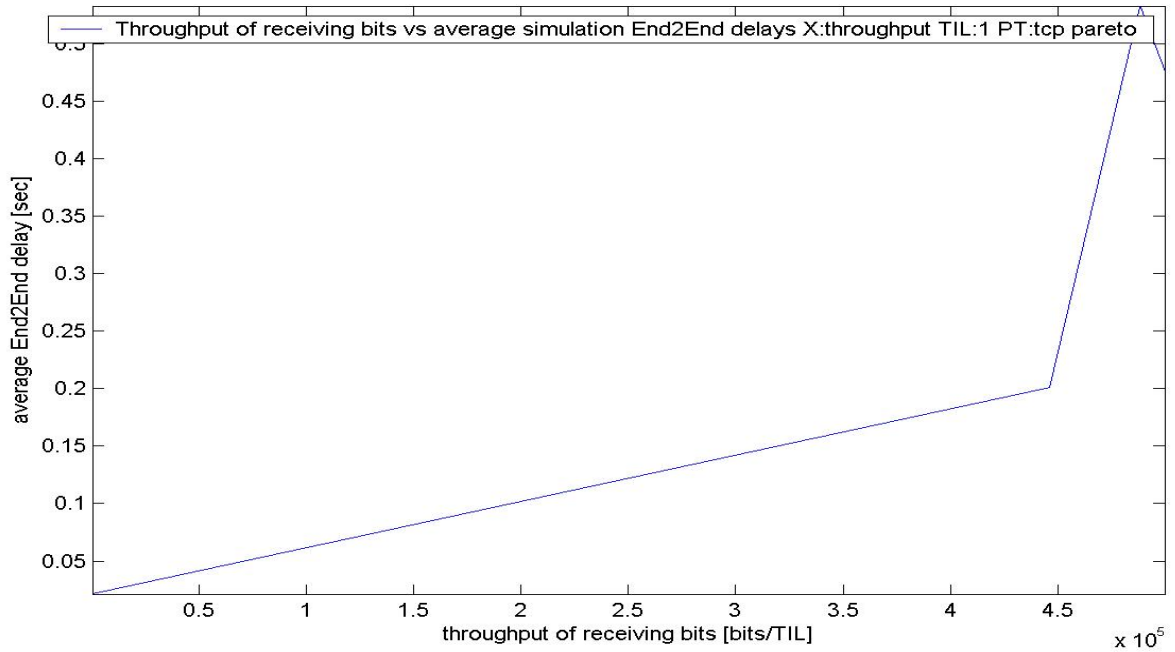


Figura 39: Simulação NS / *TRACEGRAPH* – Comprimento Fixo para Voz com Prioridade.

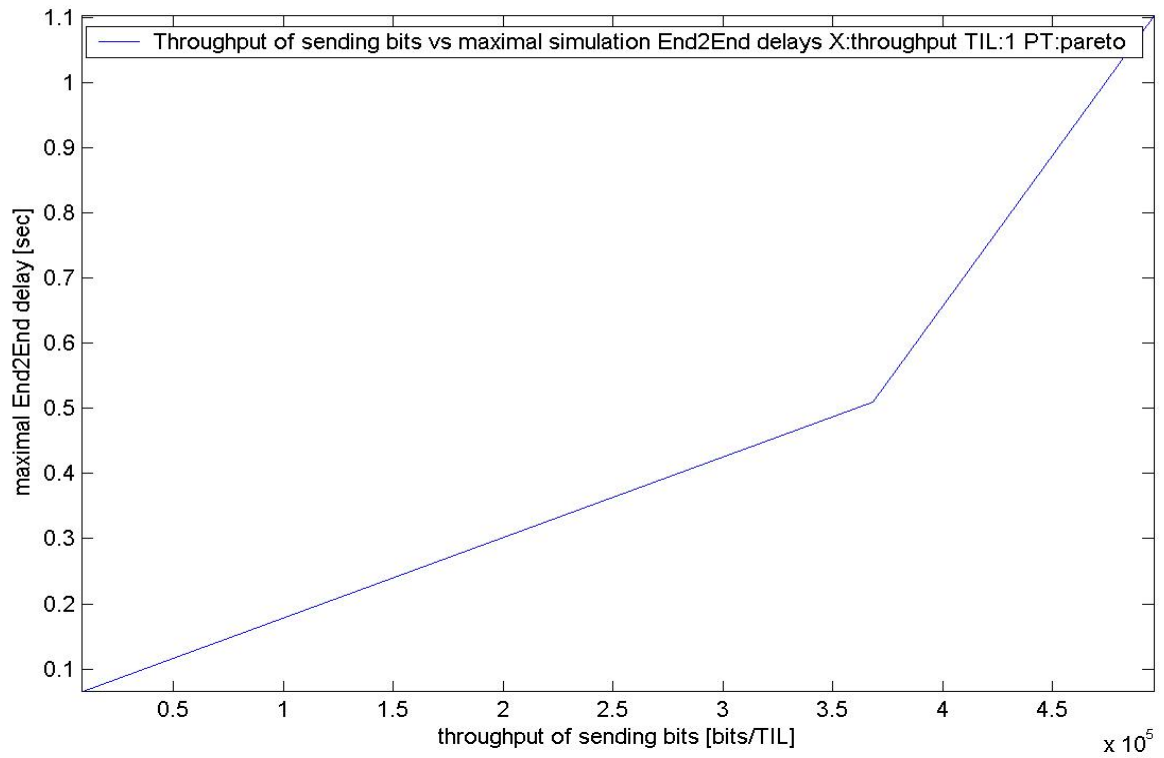


Figura 38: Simulação NS / *TRACEGRAPH* – Exp. Neg. p/ Dados com Prioridade.

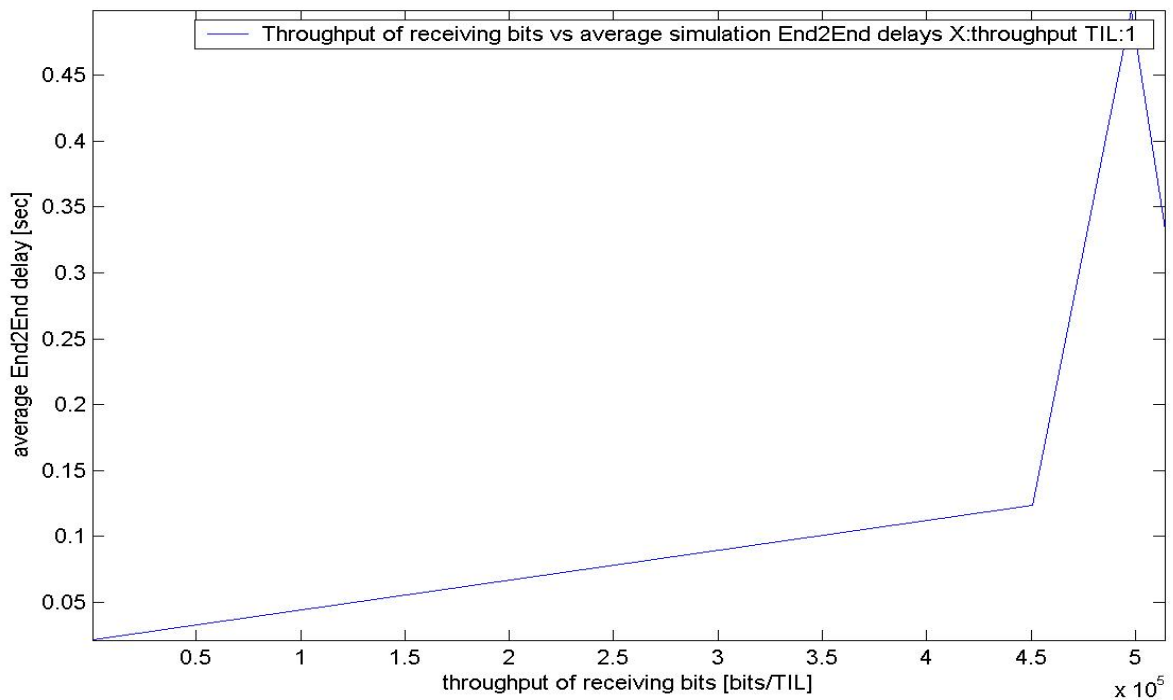


Figura 40: Simulação NS / *TRACEGRAPH* – Comprimento Fixo p/ Dados com Prioridade.

Finalmente, com o resultado da simulação, pode-se verificar que na transmissão de sinais em redes IP utilizando um melhor método de encaminhamento de pacotes, pela técnica de multiplexação de pacotes, proposto por Yamada (2001), os recursos que garantem o desempenho na transmissão dos sinais, a diferenciação dos tipos de serviços, métodos de garantia de recursos, não podem ser desconsiderados, visto que em uma transmissão de sinais, quanto maior o tráfego, maior é a necessidade de controle e gerenciamento dos pacotes, para garantir os limites necessários de um bom resultado se tratando de Voz sobre IP.

## CONCLUSÃO

Este trabalho apresentou um estudo sobre a análise do impacto do atraso na transmissão de pacotes voz sob rede IP, tomando-se como base a avaliação do resultado de um estudo apresentado por Yamada (2001), sobre formas de encaminhamento de pacotes IP/UDP/RTP.

Técnicas apresentadas por aquele autor, estabelecem diversos critérios a serem investigados, tais como: redução de *overhead* no cabeçalho dos pacotes, redução da carga de processamento nos dispositivos de roteamento e redução de recursos UDP. Esquemas de transportes de pacotes RTP são avaliados: Compressão de cabeçalhos de pacotes IP/UDP/RTP. Este último caracterizando a multiplexação de sinais, segundo o autor, apresentou os melhores resultados nas diversas simulações executadas, porém, o mesmo não avaliou a questão do atraso na transmissão dos pacotes.

Estudos sobre VoIP (Voz sobre IP) descrevem vários parâmetros a serem considerados na condição de um resultado do objeto desse trabalho. O atraso é um dos principais aspectos de QoS tratados em nosso estudo, porém, levantamentos sobre as arquiteturas VoIP, *codecs*, componentes de VoIP e seus protocolos (H323, RTP, RTCP, SIP, MEGACO, RSVP), técnicas de implementação de QoS nas redes (*difserv*), foram necessários para uma melhor compreensão sobre as causas do atraso e as implementações possíveis e necessárias para um melhor resultado da transmissão de voz em redes IP.

Um modelo teórico sobre a avaliação do atraso foi elaborado com base na proposta apresentada por Yamada (2001). Topologias, métricas, definições sobre os processos de distribuição de pacotes (Exponencial Negativa e Comprimento Fixo) de pacotes de dados e voz. Com a análise projetada pelo modelo teórico, fomos capazes de estabelecer quais foram

as melhores condições dos demais componentes (largura de banda, processamento, meio físico, etc) do circuito que envolve a transmissão.

Para verificação da veracidade dos cálculos do modelo teórico, uma simulação de implementação elaborada no NS (*Network Simulator*) - software de simulação de redes que trabalha na simulação de protocolos como o TCP / UDP, protocolos de multicast em redes com ou sem fio e roteamento de pacotes, entre outros recursos – foi criado o mesmo ambiente estabelecido para o modelo teórico, com a mesma topologia, capacidade de canais, tamanho de pacotes e gerados scripts para a simulação de tráfego com as mesmas formas de distribuição de fontes de tráfego.

Como resultados obtidos da simulação, comprovamos a veracidade do modelo teórico, e observamos que na transmissão de sinais em redes IP utilizando um melhor método de encaminhamento de pacotes (multiplexação de sinais, Yamada (2001)), recursos que garantem o desempenho na transmissão dos sinais (processamento), e diferenciação dos tipos de serviços, um bom resultado na transmissão de sinais de Voz sobre redes IP serão garantidos.

## REFERÊNCIAS

- ARANGE, M. , **Media Gateway Control Protocol (MGCP)**, Internet draft, draft-huitema-MGCP- June 1999.
- BRADY, P. – **A Model for Generating ON-OFF Patterns in Two-Way Conversations** , Bell System Technology Journal, vol 48, 1999
- BERNI, Daniel; **Voz Sobre Protocolo da Internet**; Edição 347; Saber Eletrônica; Edição 347; Dezembro 2002.
- BLAKE, S.; Black, D.; Carson, M.; Davies, E.; Wang, Z.; Weiss, W.; **An Architecture for Differentiated Services**; RFC-2475; Network Working Group; Dezembro 1998.
- BOLOT, J. and Garcia, A., **Control Mechanisms for Packet Audio in the Internet**, INRIA - 1993.
- CLAUDE, Falbriad; **Protocolos e Aplicações para Rede de Computadores**, Ed. Érica; Janeiro 2002
- CISCO, **Enterprise IP Packet Telephony Solutions Guide**, <<http://www.cisco.com/>>, Nov.2003
- CALVO, S.; **A Linguagem dos Negócios**; Revista Network Computing; junho de 1999
- DARPA/NIST, **Network Simulation Validation Workshop**, <<http://www.dyncorp-is.com/darpa/meetings/nist99may/>>, Maio 2004.
- FRIEDMAN, T., **Multicasting Session Membership Size Estimation**, *Proceedings of IEEE Infocom '99*. Nova York, <[ftp://gaia.cs.umass.edu/pub/Friedman99\\_Infocon99.ps.gz](ftp://gaia.cs.umass.edu/pub/Friedman99_Infocon99.ps.gz)>, Março 1999.
- GOOGMAN, Bill - **Internet Telephony and Modem Delay**, IEEE Network Communicator, May 1999
- GOYAL,P.; Greenberg , A – **Integration of Call Signaling and Resource Management for IP Telephony**, IEEE Network, May/June 1999
- HANDLEY, M., Jacobson, V.; **SDP: session description protocol**, Request for Comments (Proposed Standard) –RFC-2327, Internet Engineering Task Force, Apr. 1998.
- H.245, International Telecommunication Union Telecommunication Standardization Sector ITU-T, **Security and Encryption for H-Series (H-323 and other H.245-based) Multimedia Terminals**, novembro de 2000.
- ITU-T, Recommendation H.323 V2, **Packet-Based Multimedia Communications Systems**, 1998.
- ITU-T, Recommendation G.729, **Coding of speech at 8 kbits/s using conjugate-structure algebraic-code-excited linear-prediction (CD-ACELP)**, Março 1996.

- KAMIENSKI, C. , **Qualidade de Serviço na Internet**, SBRC 2000, Maio 2000.
- KUROSE, J.F, *Unix Network Programming*, < <http://manic.cs.umass.edu> > , 2003.
- LEFELHOCZ, C. et al., **Congestion Control for Best-Effort Service : Why We Need A New Paradigm**, IEEE Network, Janeiro 1996.
- MCCABE, James D.; **Practical Computer Network Analysis and Design**; Morgan Kaufmann Series in Networking, 1998.
- MARTINS, Joberto; **Qualidade de Serviço (QoS) em Redes IP Princípios Básicos, Parâmetros e Mecanismos**; JSMNet – Itelcon – UNIFACS; Disponível em <<http://www.networkdesigners.com.br/artigos/qos/qos.html>>. ; fevereiro 2001.
- NAM, **NAM: Network Animator**, < <http://www.isi.edu/nsnam/nam/>>, Maio 2004.
- NS, **The Network Simulator — ns-2**, < <http://www.isi.edu/nsnam/ns/>> , Dez. 2003
- PENNO Filho,R.– Minicurso **Voz sobre IP e Voz sobre Frame Relay**, SBRC, Maio 1999
- PERCY,Alan, **Understanding Latency in IP Telephony**, Brooktrout Technology, Needham, MA, February 1999
- ROWE, M. – **Measure VoIP Networks for Jitter and Loss**, Test & Measurement World, December 1999
- REZENDE, J. and Ziviani, Artur **Tráfego de Voz em um Ambiente de Diferenciação de Serviços na Internet**, SBRC , Maio 1999.
- ROY, R. R, AT&T, **Framework for H.323 Inter-Gatekeeper Communications**, APC 1385, Rapporteur Meeting, Cannes, France, Jun , 1998.
- RFC 1889, Schulzrinne H., **RTP: A Transport Protocol for Real-Time Applications**, IETF, Jan. 1996.
- RFC 1890, Schulzrinne H., **RTP Profile for Audio and Video Conferences with Minimal Control**, IETF, Jan. 1996.
- RFC 2205, R. Braden, **Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification**, IETF, Set. 1996.
- RFC 2474, R. Braden, **Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers**, IETF, Dez. 1998.
- RFC 2475, S. Blake, **An Architecture for Differentiated Services**, IETF, Dez. 1998.
- RFC 2543, M. Handley, **SIP: Session Initiation Protocol**, IETF, Mar. 1999.

RFC 2705, M. Arango, **Media Gateway Control Protocol (MGCP) Version 1.0**, IETF, Out. 1999.

RFC 3015, F. Cuervo, **Megaco Protocol Version 1.0**, IETF, Nov. 2000.

STANTON, Michael; **A Convergência das Redes de Comunicação**; 2000

SITOLINO, Cláudio Luiz; **Voz sobre IP – um estudo experimental**; 1999

SKENE, Jerry; **IP and next generation networks, ITU-T\* Rapporteur for studies on voice over Internet protocol (VoIP) gateways, and Standards Director at Tellabs, Inc.** 2000.

SCHULZRINNE, H., Casner, S., Frederick, R.; **RTP: a transport protocol for real-time applications**, Request for Comments (Proposed Standard) 1889, Internet Engineering Task Force, Jan. 1996.

SUBBIAH, B.; Sengondan, S.; **User Multiplexing in RTP payload between IP Telephony Gateways**; draft-ietf-avt-mux-rtp-00.txt; Internet Engineering Task Force – IETF; Aug 21, 1998

SAWASHIMA, H. – **Characteristics of UDP packet Loss: Effect of TCP Traffic**, Inet 1997. <[http://www.isoc.org/isoc/whatis/conferences/inet/97/proceedings/F3/F3\\_1.HTML](http://www.isoc.org/isoc/whatis/conferences/inet/97/proceedings/F3/F3_1.HTML)>

SILVA, A., **Qualidade de Serviço em VoIP**, RNP NewsGeneration, volume 4, número 3, Maio 2000.

TOGA, J., Elgebaly, H. **Demystifying Multimedia Conferencing Over the Internet Using the H.323 Set of Standards**, Intel Architecture Labs, Intel Corporation, 1998. Disponível em <<http://www.intel.com.tw/technology/itj/q2199/articles/art4.htm>>.

TANEMBAUM, Andrew; **Computer Networks**; 3<sup>rd</sup> edition, Prentice-Hall, 1996.

VINT, Project (Virtual InterNetwork Testbed). <<http://www.isi.edu/nsnam/vint>>, Dez. 2003

YAMADA, Hideaki; Higuchi, Norio; **Voice Quality Evaluation of IP-Based Voice Stream Multiplexing Schemes**, KDDI R&D Laboratories Inc.; Japan, 2001.

WROCLAWSKI, J.; **The Use of RSVP with IETF Integrated Services**; RFC-2210; Network Working Group; Setembro 1997.

ROSS, Keith W., **Redes de Computadores e a Internet – Uma nova abordagem**, Addison Wesley, 2003.

## ANEXO A – SCRIPT OTCL SEM PRIORIDADE

```
SET NS [NEW SIMULATOR]

$NS COLOR 1 BLUE
$NS COLOR 2 RED
$NS COLOR 3 GREN

$NS USE-NEWTRACE
SET NF2 [OPEN EXEMPLO3.TR W]
$NS TRACE-ALL $NF2

SET NF [OPEN EXEMPLO3.NAM W]
$NS NAMTRACE-ALL $NF

PROC FINISH {} {
    GLOBAL NS NF
    $NS FLUSH-TRACE
    CLOSE $NF
    EXEC NAM EXEMPLO3.NAM &
    EXIT 0
}

SET N0 [$NS NODE]
SET N1 [$NS NODE]
SET N2 [$NS NODE]
SET N3 [$NS NODE]

$NS DUPLEX-LINK $N0 $N2 0.5MB 10MS DROPTAIL
$NS DUPLEX-LINK $N1 $N2 0.5MB 10MS DROPTAIL
$NS DUPLEX-LINK $N2 $N3 0.5MB 10MS DROPTAIL

$NS QUEUE-LIMIT $N2 $N3 100

$NS DUPLEX-LINK-OP $N0 $N2 ORIENT RIGHT-DOWN
$NS DUPLEX-LINK-OP $N1 $N2 ORIENT RIGHT-UP
$NS DUPLEX-LINK-OP $N2 $N3 ORIENT RIGHT

#AGENT/TCP SET NAM_TRACEVAR_ TRUE ;# HABILITA "TRACE" DE VARIÁVEIS
TCP (NAM)

$NS DUPLEX-LINK-OP $N2 $N3 QUEUEPOS 0.5

# SET FILE [OPEN DROP.TR W]
# $NS CREATE-TRACE TCP $FILE $N1 $N2

SET TCP [NEW AGENT/TCP]
$TCP SET CLASS_ 2
$NS ATTACH-AGENT $N0 $TCP
$TCP SET PRIO_ 0

SET SINK [NEW AGENT/TCPSINK]
$NS ATTACH-AGENT $N3 $SINK
$NS CONNECT $TCP $SINK
```



```
$TCP SET FID_ 1

SET FTP [NEW APPLICATION/FTP]
$FTP ATTACH-AGENT $TCP
$FTP SET TYPE_ FTP
$FTP SET PACKET_SIZE_ 1000

SET UDP [NEW AGENT/UDP]
$NS ATTACH-AGENT $N1 $UDP

SET NULL [NEW AGENT/NULL]
$NS ATTACH-AGENT $N3 $NULL
$NS CONNECT $UDP $NULL
$UDP SET FID_ 2

SET CBR [NEW APPLICATION/TRAFFIC/CBR]
$CBR ATTACH-AGENT $UDP
$CBR SET TYPE_ CBR
$CBR SET PACKET_SIZE_ 1000
$CBR SET RATE_ 1MB
$CBR SET RANDOM_ FALSE

#SET EXP [NEW APPLICATION/TRAFFIC/EXPONENTIAL]
#$EXP ATTACH-AGENT $UDP
#$EXP SET TYPE_ EXP
#$EXP SET PACKET_SIZE_ 1000

$NS AT 0.1 "$CBR START"
$NS AT 0.1 "$FTP START"
#$NS AT 0.1 "$EXP START"
$NS AT 3.0 "$FTP STOP"
$NS AT 3.0 "$CBR STOP"
#$NS AT 3.0 "$EXP STOP"

$NS AT 3.5 "$NS DETACH-AGENT $N0 $TCP ; $NS DETACH-AGENT $N3 $SINK"

$NS AT 3.5 "FINISH"

PUTS "CBR PACKET SIZE = [$CBR SET PACKET_SIZE_]"
PUTS "CBR INTERVAL = [$CBR SET INTERVAL_]"

$NS RUN
```

## ANEXO B – SCRIPT OTCL COM PRIORIDADE

```
set ns [new Simulator]

$ns color 1 Blue
$ns color 2 Red
$ns color 3 gren

set f1 [open pareto.tr w]
set f2 [open ftp.tr w]
#set f3 [open exp.tr w]

$ns use-newtrace
set nf [open exemplo.tr w]
$ns trace-all $nf

set nf2 [open exemplo.nam w]
$ns namtrace-all $nf2

proc finish {} {
    global ns nf nf2 f1 f2

    $ns flush-trace
    close $nf
    close $nf2
    exec nam exemplo.nam &

    close $f1
    close $f2

    exit 0
}

proc record {} {
    global sink1 sink2 f1 f2

    set ns [Simulator instance]
    set time 0.2
    set bw1 [$sink1 set bytes_]
    set bw2 [$sink2 set bytes_]
    set now [$ns now]
    puts $f1 "$now [expr $bw1/$time*8/1000000]"
    puts $f2 "$now [expr $bw2/$time*8/1000000]"
    $sink1 set bytes_ 0
    $sink2 set bytes_ 0
    $ns at [expr $now+$time] "record"
}

#-----
#Create nodes and links
#-----

set n0 [$ns node]
set n1 [$ns node]
set n2 [$ns node]
set n3 [$ns node]

$ns duplex-link $n0 $n2 0.5Mb 10ms DropTail
```

```

$ns duplex-link $n1 $n2 0.5Mb 10ms DropTail

$ns simplex-link $n2 $n3 0.5Mb 10ms CBQ
$ns simplex-link $n3 $n2 0.5Mb 10ms DropTail

#-----
#Configure CBQ - PQ
#-----

set cbqlink [$ns link $n2 $n3]
set cbqueue [$cbqlink queue]

$cbqueue algorithm formal

set topclass [new CBQClass]
$topclass setparams none 0 1 auto 8 2 0

set class1 [new CBQClass]
set queue1 [new Queue/DropTail]
$class1 install-queue $queue1
$class1 setparams $topclass true 0.5 auto 1 1 0

set class2 [new CBQClass]
set queue2 [new Queue/DropTail]
$class2 install-queue $queue2
$class2 setparams $topclass true 0.3 auto 1 1 0

set class3 [new CBQClass]
set queue3 [new Queue/DropTail]
$class3 install-queue $queue3
$class3 setparams $topclass true 0.2 auto 1 1 0

$cbqlink insert $topclass
$cbqlink insert $class1
$cbqlink insert $class2
$cbqlink insert $class3

$cbqlink bind $class1 1 ; # fid 1
$cbqlink bind $class2 2 ; # fid 2
$cbqlink bind $class3 3 ; # fid 3

$ns queue-limit $n2 $n3 100

$ns simplex-link-op $n2 $n3 orient right
$ns simplex-link-op $n3 $n2 orient right
$ns simplex-link-op $n0 $n2 orient right-UP
$ns duplex-link-op $n1 $n2 orient right-down

#-----
#Create Agents
#-----

#----- Trafego FTP -----
$ns duplex-link-op $n2 $n3 queuePos 0.5

set tcp [new Agent/TCP]
$tcp set class_ 2

```

```

$ns attach-agent $n0 $tcp

set sink2 [new Agent/LossMonitor]
$ns attach-agent $n0 $sink2
$ns connect $tcp $sink2
$tcp set fid_ 2

set sink [new Agent/TCPSink]
$ns attach-agent $n3 $sink
$ns connect $tcp $sink
$tcp set fid_ 1

set ftp [new Application/FTP]
$ftp attach-agent $tcp
$ftp set type_ FTP
$ftp set packet_size_ 1000

#----- Trafego On-Off -----
set udp [new Agent/UDP]
$ns attach-agent $n1 $udp

set sink1 [new Agent/LossMonitor]
$ns attach-agent $n1 $sink1
$ns connect $udp $sink1
$udp set fid_ 1

set null [new Agent/Null]
$ns attach-agent $n3 $null
$ns connect $udp $null
$udp set fid_ 2

set pareto [new Application/Traffic/Pareto]
$pareto attach-agent $udp
$pareto set type_ pareto
$pareto set packet_size_ 1000
$pareto set rate_ 1mb
$pareto set random_ false

#-----
# Schedule events
#-----

$ns at 0.0 "record"

$ns at 0.1 "$pareto start"
$ns at 0.1 "$ftp start"

$ns at 3.0 "$ftp stop"
$ns at 3.0 "$pareto stop"

$ns at 4.5 "$ns detach-agent $n0 $tcp ; $ns detach-agent $n3 $sink"

$ns at 3.1 "finish"

$ns run

```