

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

**Um Modelo de Videoconferência para Computador Pessoal
Orientado ao Perfil de Aplicação**

por

SAMUEL CARRION

Dissertação submetida à avaliação, como requisito parcial para a obtenção do grau de
Mestre em Ciência da Computação

Prof. Dra. Maria Janilce Bosquioli Almeida
Orientadora

Porto Alegre, agosto de 2002

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Carrion, Samuel

Um Modelo de Videoconferência para Computador Pessoal Orientado ao Perfil de Aplicação por Samuel Carrion. – Porto Alegre:PPGC da UFRGS, 2002.

108 f.:il.

Dissertação (mestrado) – Universidade Federal do Rio Grande do Sul. Curso de Pós-Graduação em Ciência da Computação, Porto Alegre, BR-RS, 2002. Orientadora:Almeida, Maria J.B.

1. Videoconferência. 2. Redes de Computadores. 3. IP. 4. Multimídia. I. Almeida, Maria Janilce Bosquioli. II. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitora : Profa. Wrana Panizzi

Pro-Reitor de Ensino : Prof. José Carlos Ferraz Hennemann

Pró-Reitor Adjunto de Pós-Graduação : Prof. Jaime Evaldo Fensterseifer

Diretor do Instituto de Informática : Prof. Philippe Olivier Alexandre Navaux

Coordenador do PPGC : Prof. Carlos Alberto Heuser

Bibliotecária-Chefe do Instituto de Informática : Beatriz Haro

Agradecimentos

À professora Janilce, que acreditou no meu trabalho quando aceitou me orientar, e acreditou novamente quando mudei de residência;

À minha família e minha esposa, que sempre me apoiaram em tudo;

A todos amigos, colegas e parentes que participaram de alguma forma na elaboração deste trabalho, seja me trazendo uma reportagem, um artigo, comentando o trabalho, me ajudando com o laboratório, me ajudando a administrar o tempo ou mesmo torcendo. Obrigado a todos.

Sumário

Lista de Abreviaturas.....	6
Lista de Figuras	8
Lista de Tabelas	9
Resumo	10
Abstract	11
1 Introdução	12
2 Comunicação e Mídias.....	14
2.1 Comunicação Humana.....	14
2.2 Relevância das Mídias.....	15
3 Panorama e Especificações de Videoconferência	19
3.1 Histórico	19
3.2 Especificações de Tipos de Serviços Multimídia.....	20
3.3 Popularidade Versus Dificuldades	22
4 Padrões para Implementação e Protocolos de Suporte.....	25
4.1 Suporte a Vídeo.....	25
4.2 Suporte a Áudio.....	28
4.3 Suporte a Mídias Discretas.....	33
4.4 Recomendações do ITU-T para Videoconferência	34
4.5 Outros Protocolos e Recomendações	42
5 Aspectos de Qualidade.....	44
5.1 Qualidade de Serviço	44
5.2 QoS na Internet.....	45
5.3 QoS no Contexto de Videoconferência.....	47
5.4 Sincronização	52
5.5 Dimensionamento da Qualidade	53
6 Alternativas para Busca de Qualidade	56
6.1 Infraestrutura – ISDN (H.320) versus IP (H.323)	56
6.2 Soluções Orientadas à Rede	59
6.3 Soluções Orientadas à Aplicação	60
6.4 Caracterização do Problema.....	62
7 O Modelo Proposto	65
7.1 Escopo e Contribuição	65
7.2 Perfis.....	65
7.3 Definição do Modelo.....	69

8	Implementação – Validação do Modelo.....	73
8.1	Tecnologia e Ambiente	73
8.2	Implementação	74
9	Testes e Resultados	85
9.1	Cenário e Metodologia	85
9.2	Execução dos Testes.....	86
9.3	Resultados e Observações	88
10	Conclusões	90
	Anexo – Definições Usadas na Implementação.....	93
	Bibliografia.....	103

Lista de Abreviaturas

A/D	Analógico-Para-Digital
AAL	ATM Adaption Layer
ACELP	Algebraic-Code-Excited Linear-Prediction
ACM	Audio Compression Manager
ACR	Absolute Category Rating
ADPCM	Adaptive Differential Pulse Code Modulation
API	Application Program Interface
ASCII	American Standard Code for Information Interchange
ATM	Asynchronous Transfer Mode
B-ISDN	Broadband ISDN
B-TA	Broadband Terminal Adpter
CBR	Contraint Based Routing
CIF	Common Interchange Format
CLI	Compression Labs, Inc.
CODEC	Codificador/Decodificador ou Compressor/Descompressor
CS-ACELP	Conjugate Structure Algebraic-Code-Excited Linear-Prediction
CSDN	Circuit Switched Data Network
DCR	Degradation Category Rating
DCT	Discrete Cosine Transform
DMOS	Degradation MOS
DPCM	Differential Pulse Code Modelation
EBCDIC	Extended Binary Coded Decimal Interchange Code
FTP	File Transfer Protocol
GSTN	General Switched Telephone Network
HTTP	HyperText Transfer Protocol
IEEE	Institute Of Electrical And Eletronics Engineers
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPX	Internetwork Packet Exchange
ISA	Industry Standard Architecture
ISDN	Integrated Services Digital Network
ISP	Internet Service Provider
ITU-T	International Telecommunications Union – Telecommunication Standardization Sector
JPEG	Joint Photographics Experts Group
Kbps	Kilobytes por segundo
kHz	Kilohertz
LAN	Local Area Network
LAN-E	LAN Emulation
LD-CELP	Low-Delay Excited Linear Prediction
LDU	Logical Data Unit
LPCD	Linha Privativa de Comunicação de Dados
Mbps	Megabytes por Segundo

MC	Multipoint Controller
MCU	Multipoint Control Unit
MIT	Massachusetts Institute of Technology
MOS	Mean Opinion Score
MP	Multipoint Processor
MPEG	Motion Picture Experts Group
MPLS	Multi-Protocol Label Switching
MP-MLQ	Multipulse-Maximum Likelihood Quantization
NEC	Nippon Electric Corporation
N-ISDN	Narrowband ISDN
NTSC	National Television Standards Committee
OSI	Open System Interconnection
PAL	Phase Alternation Line
PBN	Packet Based Network
PC	Personal Computer
PCM	Pulse Code Modulation
POTS	Plain Old Telephone Service
PPP	Point-to-Point Protocol
PSDN	Public Switched Data Network
QCIF	Quarter CIF
QoS	Quality Of Service
RAS	Registration, Admission and Status
RFCs	Request For Comments
RLE	Run Length Encoding
RSVP	Resource Reservation Protocol
RTCP	Real-Time Transport Control Protocol
RTP	Real-Time Transport Protocol
SAP	Session Announcement Protocol
SB-ADPCM	Sub-Band Adaptive Pulse Code Modulation
SCCP	Simple Conference Control Protocol
SDP	Session Description Protocol
SECAM	Système Electronique pour Couleur Avec Mémoire
SIP	Session Initiation Protocol
SLA	Service Level Agreement
TCP	Transmission Control Protocol
TOS	Type Of Service
UDP	User Datagram Protocol
USB	Universal Serial Bus
VFW	Video for Windows
WAN	Wide Area Network

Lista de Figuras

FIGURA 4.1 – (a) Onda de Áudio. (b) Amostragem. (c) Quantização.....	29
FIGURA 4.2 – Segmentação de uma Onda de Áudio codificada com PCM.....	31
FIGURA 4.3 – Terminal Conforme a Recomendação H.323.....	38
FIGURA 4.4 – (a) Níveis do Modelo OSI. (b) Estrutura da Recomendação H.323.	42
FIGURA 5.1 – Relação dos modelos de QoS com o modelo OSI.....	46
FIGURA 6.1 – Distribuição das formas de acesso à Internet no Brasil	60
FIGURA 7.1 – Interação entre os componentes do Modelo.....	71
FIGURA 8.1 – Definição de Atributos por Mídia – Áudio.....	76
FIGURA 8.2 – Definição de Atributos por Mídia – Vídeo	77
FIGURA 8.3 – Definição de Atributos por Mídia – Discreta.....	77
FIGURA 8.4 – Definição de Limites para Adaptação.....	77
FIGURA 8.5 – Tabela de Visualização do Perfil	78
FIGURA 8.6 – Menu Arquivo e Janela.....	79
FIGURA 8.7 – Tela de Chamada	79
FIGURA 8.8 – Telas de Vídeo (Local e Remoto) e Áudio	81
FIGURA 8.9 – Tela de Chat.....	81
FIGURA 8.10 – Tela de Mídias Discretas (simulador).....	82
FIGURA 8.11 – Controle da Sessão – Informações de Perfil e Nível.....	82
FIGURA 8.12 – Controle da Sessão – Auditoria e Descartes	84

Lista de Tabelas

TABELA 4.1 – Formatos de Imagem das Recomendações H.261 e H.263	28
TABELA 4.2 – Codecs de Áudio do ITU-T.....	30
TABELA 4.3 – Modos de operação da recomendação G.722.....	32
TABELA 4.4 – Padrões de Videoconferência do ITU-T	35
TABELA 5.1 – Tempo de compressão típico para Codecs de Áudio	50
TABELA 5.2 – Influência dos problemas de QoS da Infraestrutura nas Mídias de uma Sessão de Videoconferência	51
TABELA 5.3 – Parâmetros de QoS aceitáveis	53
TABELA 5.4 – Valores para MOS e DMOS	54
TABELA 5.5 – Índices MOS típicos dos codecs do ITU-T	54
TABELA 5.6 – Variação do índice MOS do codec G.729 em diversas condições de rede	55
TABELA 6.1 – Comparativo entre Videoconferência em Grupo (H.320) e em Desktop / Internet (H.323)	59
TABELA 7.1 – Níveis de serviço definidos por aplicação.....	66
TABELA 7.2 – Modos de Adaptação – Funcionamento Assimétrico (Exemplo)	72
TABELA 9.1 – Perfil de Videotelefone	87
TABELA 9.2 – Perfil de Vigilância / Monitoração Remota	87

Resumo

A informática vem adquirindo papéis cada vez mais importantes na vida cotidiana. Um dos papéis mais significativos, hoje, é o suporte a comunicações; atualmente, é muito difícil pensar em comunicações – mesmo interpessoais – sem fazer associação às áreas de informática e redes. Dentre as aplicações que utilizam informática e redes como suporte, a tecnologia de videoconferência tem recebido papel de destaque.

Os avanços na tecnologia de redes e conectividade, aliados à padronização e à crescente oferta de produtos de videoconferência, têm aumentado a aplicabilidade e a popularidade destes produtos, sobretudo utilizados sobre arquitetura de redes TCP/IP. Trata-se de uma tecnologia atraente em termos de resultado, por agregar, além do áudio – recurso comum há muito tempo como suporte à comunicação – os recursos de vídeo e aplicações integradas (como quadro-branco compartilhado, Chat, troca de arquivos e outros). Contudo, essas aplicações são bastante exigentes, tanto em termos de banda quanto de qualidade de serviço (QoS) da rede. O primeiro item se justifica pelo volume de dados gerados pelas aplicações de videoconferência; o segundo, pela significativa influência que os problemas de qualidade da infraestrutura de rede (como elevada latência, jitter e descartes) podem exercer sobre tais aplicações.

A busca para as soluções destes problemas não é tarefa simples, pois muitas vezes envolve investimentos que desencorajam a adoção da tecnologia de videoconferência – principalmente para uso pessoal ou por empresas pequenas. Este trabalho propõe uma solução aos problemas mencionados, visando proporcionar uma melhor aceitação e maior disseminação da tecnologia de videoconferência, valendo-se de recursos com pouca demanda de investimento. A estratégia abordada é a adaptação de tráfego, com um enfoque diferenciado: o de levar em conta, para cada aplicação, o comportamento que o processo de adaptação apresentasse. A partir dessa orientação, é proposto um modelo de adaptação de tráfego orientado ao perfil da aplicação, voltado ao interesse do usuário, e que disponibilize uma forma ao mesmo tempo simples e eficiente para que o usuário realize a adequação do mecanismo de adaptação do sistema às suas necessidades e expectativas.

A partir desta proposta, foi implementado um protótipo de aplicação, com o objetivo de verificar a funcionalidade do modelo em termos práticos. As observações dos resultados dos testes, bem como as conclusões geradas, serviram como validação da proposta.

Palavras-Chave : videoconferência, redes de computadores, IP, multimídia.

Title : “A VIDEOCONFERENCING MODEL FOR PERSONAL COMPUTER BASED ON APPLICATION’S PROFILE”

Abstract

Information technology is becoming more and more important in our everyday life. One of its most significant roles today is to support communications. Nowadays it is very difficult to think about communications – even between people – without immediately thinking about computer networks. Among the applications that use computer networks as support, videoconferencing has played a major role.

The advances made in network and connectivity technology, combined with the standardization and growing supply of videoconferencing products, have increased the applicability and the popularity of these products, which are used mainly over TCP/IP network architecture. The technology is attractive in terms of results because it aggregates not only audio features, which have long been a common resource to support communications, but also video features and integrated applications (such as shared white board, Chat, file exchange and others). However, these applications are very demanding both in terms of network bandwidth and network Quality of Service (QoS). The first item is justified by the volume of data generated by the videoconferencing applications; the second by the significant influence that problems of quality of the network infrastructure (such as high latency, jitter and discards) can have on such applications.

The search for solutions to these problems is not an easy task as it often involves investments that discourage the adoption of the videoconferencing technology – mainly for personal use or by smaller companies. This study proposes a solution for such problems in order to allow better acceptance and greater dissemination of the videoconferencing technology, using resources that do not require so much investment. The strategy is traffic conformation, with a differentiated approach: to take into account for each application the conduct of the conformation process. As from this orientation, we propose a model of traffic conformance oriented to the profile of the application, to the interest of the user, involving a form that is simple and efficient for the user to adapt the conformance process of the system to its needs and expectations.

As from this proposal, we have implemented an application prototype in order to determine the operability of the model in practice. The results of the tests as well as the conclusions that were generated support us in this validation.

Keywords : videoconferencing, computer networks, IP, multimedia.

1 Introdução

A área de redes de computadores vem sendo objeto de muito estudo e investimento. O mundo tem assistido uma evolução tecnológica acelerada, aliada à disseminação dos conceitos de produtos para redes de computadores, impulsionada por fatores como :

- O uso crescente de redes (particularmente redes locais) em organizações de diversas áreas e portes (empresas, instituições de ensino, órgãos públicos) .
- A popularização e massificação da Internet.

Nesse sentido, as aplicações e os objetivos das redes também tem sofrido alterações, e dentre as aplicações atuais, uma área que merece atenção é o suporte à comunicação. Esta também é uma área abrangente, onde se podem citar além da comunicação disseminada pela Internet (via e-mail ou WWW) aplicações de tráfego de outras mídias sobre canais de dados convencionais (exemplos de tráfego de áudio e imagem: voz sobre IP, vídeo sob demanda e videoconferência).

As aplicações de videoconferência têm se apresentado como uma alternativa de comunicação muito atraente e em evidente crescimento. Conforme apresentado em [TRE 97][ZAN 99], videoconferência pode ser definida como um meio de comunicação de larga aplicabilidade. Entre as várias áreas a que se destinam tais aplicações, é possível citar ensino a distância, suporte a trabalhos em equipes dispersas geograficamente e reuniões, particularmente as extraordinárias, gerenciamento de projetos, recrutamento de pessoas chave, telemedicina e provisão de cuidados de saúde à distância, consultoria, anúncio de novos produtos, aplicações jurídicas e legais e muitas outras. Em relação à quantidade de participantes, pode-se tratar sistemas 1:N (palestra) ou N:N (CSCW); e, em relação à natureza da aplicação, *soft real-time* (caso geral) ou *hard real-time* (por exemplo, telemedicina). De um modo geral, o mercado divide os tipos de aplicações de videoconferência em dois grandes grupos: Profissional, ou sistemas baseados em estúdio (com o uso de câmeras específicas, placas de captura de vídeo e *hardware* para compressão de dados) e Pessoal, representado por sistemas baseados em computador pessoal, utilizando recursos de *hardware* mais acessíveis como câmeras do tipo WebCam/QuickCam (cujo custo pode ficar abaixo de 200 Reais) e interfaces de áudio convencionais.

Independente da natureza da aplicação, o tráfego multimídia possui requisitos diferenciados de qualquer outra aplicação, podendo ser considerado o tipo de tráfego mais exigente em termos de rede. A infraestrutura a ser provida deve atender tais requisitos, particularmente no que diz respeito aos fatores relacionados ao tempo de transmissão (como latência e jitter) e à largura de banda (que, além da vazão propriamente dita, pode influenciar os fatores anteriores) [ZAN 99][WIL 2000]. O uso profissional normalmente justifica preparação e investimento em infraestrutura – seja esta própria ou acesso terceirizado. No entanto, o segmento de aplicações pessoais – caracterizado principalmente por comunicação via Internet – normalmente não justifica ou mesmo não viabiliza tal investimento. Pode se tornar técnica ou economicamente inviável depender de um acesso em banda larga, ou, mesmo através deste tipo de acesso,

o usuário pode não contar com garantias de qualidade de serviço (QoS), pois as redes IP – base para os acessos Internet – por definição não disponibilizam tais garantias.

Apesar destas dificuldades, a comunicação pessoal é um segmento que merece muita atenção. Como parâmetro, os usuários em potencial podem ser os usuários de Internet, um número que evolui muito a cada dia. Assim, torna-se interessante em termos de mercado e de tecnologia propriamente dita o desenvolvimento de abordagens com o objetivo de suportar tráfego de videoconferência sem a necessidade de investimento em infraestrutura – ou minimizando este investimento. Dentre tais abordagens, escolheu-se trabalhar com adaptação do tráfego de videoconferência, por suas características de implementação voltadas à aplicação. Desta forma, não há demanda ou custo com recursos adicionais, e o sistema desenvolvido a partir desta abordagem é portátil para diversas realidades em termos de infraestrutura de rede.

A abordagem de adaptação, contudo, pode restringir o funcionamento do sistema à forma que o fabricante implementou, sem levar em conta os interesses particulares de um dado usuário em uma dada situação (a importância das mídias para aquela aplicação específica). Se um sistema for considerado tão efetivo quanto ele atende às expectativas do usuário, considera-se também relevante a forma como a adaptação é realizada; assim, este trabalho propõe uma forma de interação com o usuário para definir o que é importante para ele, ou seja, qual mídia deve ser priorizada e qual deve ser relegada em caso de necessidade de adaptação, em determinada circunstância de uso do sistema de videoconferência.

O trabalho está organizado como segue. O capítulo 2 apresenta os conceitos de mídias e as particularidades de cada uma. No capítulo 3, é feito um apanhado da tecnologia – histórico, padronização do serviço e as restrições mais comuns ao seu uso. O capítulo 4 apresenta os padrões de implementação dos serviços, desde a definição dos codecs até os protocolos de suporte utilizados na implementação de sistemas de videoconferência. O capítulo 5 trata das definições de qualidade – tanto da infraestrutura quanto do sistema, este último essencial para o entendimento de como prover o melhor serviço do ponto de vista do usuário. O capítulo 6 apresenta algumas das abordagens possíveis para solucionar os problemas de restrição da tecnologia de videoconferência, disponibilizando serviços de melhor qualidade ao usuário. No capítulo 7 é apresentado o modelo de adaptação baseado em perfis, voltado ao atendimento particular da expectativa do usuário. A implementação de um protótipo que visa validar o mesmo é apresentada no capítulo 8, e os testes do protótipo são apresentados no capítulo 9.

2 Comunicação e Mídias

O termo “mídia” pode ser definido como a forma de distribuição e apresentação de informação. No que diz respeito à interação humana com o ambiente computacional, os diferentes tipos de mídia podem ser classificados de diversas formas, de acordo com [STE 95] :

- Meio de Representação – ou como a informação é codificada (por exemplo, um texto pode ser codificado em ASCII ou EBCDIC);
- Meio de Apresentação – ferramentas e dispositivos para entrada e saída de informação, como mouse, câmera, monitor e outros;
- Meio de Armazenamento – meio físico de armazenamento da informação, como microfilme, CD-ROM, papel;
- Meio de Transmissão – meio que habilita a transmissão de dados, como fibra óptica, cabo coaxial ou o ar livre (para sistemas *wireless*);
- Meio de Percepção – como o ser humano percebe a informação : mídias visuais (texto, imagem) ou auditivas (som, ruído). No escopo deste trabalho, e no que diz respeito à comunicação, o conceito de mídia será tratado com o enfoque de meio de percepção.

Outra forma de classificação de mídias diz respeito às dimensões de representação. Por exemplo, a tela do computador possui duas dimensões espaciais; um holograma exige uma dimensão espacial adicional. O tempo também representa uma dimensão no conceito de sistemas multimídia, e de acordo com ele as mídias podem ser classificadas como [STE 95]:

- Discretas – Algumas mídias, como texto e gráficos, são independentes do tempo. A informação, nesses casos, consiste de uma seqüência de elementos individuais, ou de um elemento contínuo sem um componente tempo. O processamento dessas mídias deve ocorrer o mais rápido possível, mas não é crítico pois a validade – e correção – dos dados não depende de condições temporais.
- Contínuas – Mídias como o áudio e vídeo variam no tempo. A informação não é expressa apenas pelo seu valor, mas também pelo tempo de ocorrência – este tipo de mídia é composto da representação sucessiva de amostras individuais. O processamento dessas mídias é crítico em relação ao tempo, pois cada instância tem uma validade determinada.

2.1 Comunicação Humana

A comunicação humana não se restringe ao uso das palavras. Conforme Patrick, existem três canais para mediar a comunicação humana [PAT 2001]:

- Lingüístico : a comunicação expressa pelo significado puro das palavras;

- Paralingüístico : o resultado da consideração de aspectos como a velocidade do diálogo e o tom de voz, que modificam o significado das palavras;
- Não Lingüístico : a comunicação inclui gestos, postura e posição do corpo, expressão facial e olhar. Mesmo comportamentos como o espaço mantido entre duas pessoas, ou o atraso constante para uma reunião regular, transmitem um determinado significado.

Esses canais abrangem a totalidade da expressão e percepção humana. É interessante observar que, conforme o autor, estudos na área de comunicação dirigidos por psicólogos sociais demonstram que entre 60 e 90 % do significado da comunicação humana é expresso através de canais não verbais. Este dado demonstra claramente que o significado da palavra falada pode ser alterado drasticamente mediante uma alteração de tom, um gesto ou um olhar. Assim, é possível concluir que, mais que o conteúdo e a fluência das palavras, a informação não verbal é bastante significativa na determinação do contexto real da conversa, e de grande auxílio no entendimento.

Transpondo essas considerações para a comunicação mediada por computador, é fácil perceber a importância de se contar com o suporte de diversas mídias, provendo acesso às diversas formas de comunicação mencionadas. Desta forma, é possível disponibilizar recursos para aprimorar a forma de comunicação, tornando-a mais efetiva e até mais simples.

2.2 Relevância das Mídias

No contexto de sistema multimídia, cada componente possui uma orientação específica. Nos parágrafos seguintes, serão apresentados alguns aspectos mais específicos para cada mídia, no que diz respeito à sua natureza, percepção e expectativas.

Vídeo

O recurso de imagem em movimento é produzido mediante aproveitamento da limitação de velocidade do olho humano para perceber alterações de imagens. Estudos demonstram que uma imagem só é percebida pelo cérebro após permanecer projetada na retina por no mínimo 1/40 de segundo. Assim, se uma seqüência de imagens for exibida à taxa de 40 exposições por segundo, o receptor tem a impressão de movimento contínuo. Esse efeito é possível de ser obtido a partir de taxas entre 10 e 20 exposições, conforme o indivíduo; a taxas menores, é possível perceber uma “quebra” na continuidade do movimento.

Em relação às grandezas que definem o formato de imagem de vídeo enquanto recurso de multimídia, pode-se identificar o tamanho ou resolução, a quantidade de cores e a taxa de quadros exibidos por segundo. Maiores detalhes sobre cada um desses parâmetros, bem como sobre os processos relacionados a vídeo (captura, digitalização e compressão) serão apresentados no capítulo 4.

O valor do vídeo para a tecnologia de videoconferência é evidente como recurso para intermediar interação interpessoal, criando a sensação de presença. É a maneira ideal para que os participantes de uma dada reunião possam, de forma natural, obterem noção espacial dos participantes e objetos do ponto remoto, bem como assegurar-se do quanto os outros participantes estão envolvidos no assunto, a razão de pausas no diálogo, as suas atitudes e outros aspectos (humor, sarcasmo) inerentes ao diálogo. Este suporte à interação torna a comunicação mediada por vídeo mais eficiente e efetiva.

Som e Áudio

O som é o resultado da vibração de matéria. Durante esta vibração, são criadas variações de pressão no ar ao redor dessa matéria; o padrão dessa variação é chamado de onda sonora. Uma onda sonora repete a mesma forma em intervalos regulares, denominados períodos. A quantidade de períodos de uma onda sonora em um segundo é chamada de frequência, e medida em hertz (Hz), ou ciclos por segundo. O ouvido humano é capaz de perceber sons entre 20 Hz e 20 kHz; a voz humana compreende ondas sonoras entre 40 e 4 kHz.

Para utilização de som e áudio como recurso de multimídia, é necessária a conversão do formato analógico (a forma originalmente recebida pelo equipamento de captura) para a forma digital. Este processo é denominado digitalização. Os principais parâmetros do áudio nesse formato são a frequência (ou quantidade de amostras por segundo), a quantidade de amostras e o número de canais. Os processos de digitalização e compressão de áudio serão apresentados mais detalhadamente no capítulo 4.

Apesar do valor adicionado pelo vídeo ser de relevante importância, o áudio desempenha um papel fundamental no suporte à interação humana. Alguns estudos apontam que o vídeo tem sua influência maximizada no início da sessão, viabilizando a socialização dos participantes. Passada esta fase inicial, não ocorre a perda de importância do vídeo, e sim um aumento da atenção em relação à fluência e qualidade do áudio. Qualquer tecnologia que degrada a qualidade do canal de áudio (por distorção, eco ou atraso) prejudica o desempenho percebido pelo usuário (ou a qualidade de serviço do ponto de vista deste) por não atender as expectativas baseadas em suas experiências [TAN 92]. Na maioria das aplicações multimídia interativas em tempo real, se houver disponibilidade apenas de áudio, a informação essencial ainda pode ser enviada, ao passo que com a perda do áudio a interação fica muito prejudicada [CHA 98].

Aplicações Integradas e Mídias Discretas

As chamadas aplicações integradas constituem um conjunto de recursos denominado conferência audiográfica (ou conferência de documentos, ou *data conferencing*). O tipo de objeto tratado neste tipo de conferência envolve principalmente o compartilhamento e transferência de gráficos, dados e documentos, como ferramentas de apoio e incremento de produtividade, com restrições de tempo muito menores do que as mídias de áudio e vídeo. Do ponto de vista de aplicações de videoconferência, a padronização do uso deste tipo de aplicação é realizada pela série T.120 do ITU-T (apresentada no capítulo 4).

Expectativas por Mídia

Do ponto de vista do usuário, a satisfação proporcionada pelo sistema é diretamente proporcional aos recursos disponibilizados por este que efetivamente interessam ao usuário. Assim, considerando os atributos das mídias de um sistema de videoconferência como sendo recursos a prover para o usuário, a princípio quanto maior fosse a resolução (medida dos parâmetros das mídias contínuas) e interatividade (medida dos tempos de atraso fim a fim), melhor seria o resultado.

Em relação ao vídeo, é comum que o primeiro contato com a tecnologia de videoconferência traga como parâmetro esperado de qualidade a resolução de televisão, principalmente em termos de fluência de imagem (quadros por segundo). No que diz respeito ao áudio, qualquer sistema que inclua este recurso como forma de interação terá como parâmetro de qualidade esperado, por parte do usuário, as experiências de suas interações por telefone.

No entanto, a maximização das definições das mídias pode não ser interessante, necessária ou mesmo útil. Com base no que já foi apresentado em relação a mídias, a limitação da resolução das mesmas pode se dar em três aspectos:

No que diz respeito aos sentidos humanos, a fluência das mídias deve obedecer aos limites já mencionados. Por exemplo, para a maioria das pessoas, não é possível distinguir a diferença de duas imagens exibidas uma após a outra em tempos menores de 1/40 de segundo, o que pode caracterizar taxas maiores que essas como desperdício de recursos.

No que diz respeito aos recursos disponíveis, a resolução a ser utilizada depende da capacidade do ambiente envolvido, incluindo *hardware* e rede. Os fatores a serem considerados aqui são:

- Os dispositivos de entrada e saída das mídias. Por exemplo, de nada vale a aplicação disponibilizar a manipulação de áudio em cinco canais (*surround*) se os dispositivos de captura manipulam no máximo dois canais (estéreo);
- A capacidade de processamento do *hardware* envolvido, seja ele dedicado para o tratamento da compressão / descompressão das mídias ou um computador ou estação de trabalho;
- A capacidade da rede em prover vazão para o volume de tráfego gerado pela aplicação.

Finalmente, no que tange à aplicação, deve ser levada em conta a necessidade de recursos. Como exemplo, a utilização de som em cinco canais, com frequência de 20 kHz e imagem de alta resolução pode ser o empecilho para uma aplicação de comunicação interpessoal, na qual é possível obter a mesma efetividade utilizando parâmetros muito menores.

É importante que as aplicações levem essas limitações em conta, principalmente em ambientes com restrições de recursos – entenda-se aqui recursos como processamento, capacidade dos dispositivos multimídia, banda passante e outros

relacionados. Em tais casos, a definição da resolução de cada mídia deve ser precedida de uma análise da relação custo-benefício entre tais resoluções e os recursos necessários para utilizá-las. Outra análise que deve ser feita é a relação entre os recursos e a aplicação: aplicações avançadas, que demandam alta resolução de áudio e vídeo, demandam equipamentos e infraestrutura específicos e investimento diferenciado, minimizando ou diminuindo as restrições mencionadas. Todos esses fatores devem ser considerados no momento de escolha da solução para uma dada demanda, sob pena de frustrar as expectativas do usuário e provocar a rejeição da tecnologia.

3 Panorama e Especificações de Videoconferência

A idéia por trás da tecnologia de videoconferência não chega a ser recente. Neste capítulo, será apresentado um breve histórico da tecnologia de videoconferência, abordando desde a origem até o panorama atual do mercado. Também serão apresentadas as especificações de serviços multimídia padronizadas pelo ITU-T (dentro os quais videoconferência é um caso bastante representativo), e alguns fatores que representam restrições à popularização de tal tecnologia.

3.1 Histórico

As primeiras demonstrações de videoconferência foram realizadas pela Bell Labs, nos EUA, na década de 20. Cientistas apresentaram uma aplicação bastante primária de videoconferência entre Washington e Nova Iorque, baseada na tecnologia de televisão. Contudo, o interesse nessa tecnologia diminuiu durante a época das duas guerras mundiais, sendo retomado por volta de 1950. Em 1964, a Bell Labs apresentou um produto chamado Picturephone, na World's Fair in Flushing Meadow, em Nova Iorque. Este produto era voltado ao uso pessoal, mas demandava, na época, o equivalente a 100 linhas telefônicas para o tráfego de vídeo, que era separado do tráfego de áudio. A imagem era pouco nítida, mas o que chamava a atenção era o uso de sinal de vídeo compactado: a partir de 45 Mbps originais, um codec de aproximadamente 250 mil dólares produzia o fluxo de vídeo a 6 Mbps [WIL 2000].

A partir da década de 1970, a atenção das empresas se voltou para uma abordagem diferente. Dado o elevado custo dos equipamentos, tornou-se mais viável, do ponto de vista econômico, oferecer sistemas que pudessem ser usados por várias pessoas (sistemas de grupo). Esta linha de produtos teve como pioneira a Nippon Electric Corporation (NEC), mas teve em seu impulso fundamental a atuação da British Telecom, que atuou junto a outras companhias telefônicas européias para conduzir testes de videoconferência de um país para outro. O esforço para fazer estes sistemas funcionarem traduziu-se na concepção inicial dos padrões de interoperabilidade dos sistemas de videoconferência. No final da década de 70 e início dos anos 80, a estratégia de aplicação de videoconferência passou a se traduzir pelo paradigma da substituição, ainda hoje determinante em boa parte das decisões de adoção de videoconferência. A filosofia neste caso é o pagamento pelo sistema de videoconferência através da economia em custos de deslocamento. Algumas das primeiras empresas a utilizarem estúdios de videoconferência foram Boeing, Citybank, Chrysler, Federal Express, Ford, Hewlett-Packard e Texas Instruments [WIL 2000].

Os sistemas utilizados por esses usuários pioneiros tinham em comum o fabricante – a empresa Compression Labs, Inc. (CLI), fundada pelo Dr. Wen-Hsiung Chen em 1976. É de sua autoria o primeiro codec para baixas (para a época) taxas de transmissão, que chegou ao mercado em 1982 e tornou viável o uso de circuitos T-1 para transmissão de vídeo, e possibilitou migrar a arquitetura de comunicação baseada em redes de satélite de alto atraso para redes de comunicação terrestres, diminuindo o custo da aplicação. Em 1984, um grupo de pesquisadores do MIT fundou a PictureTel, que em pouco tempo conseguiu a proeza de disponibilizar serviços de videoconferência à taxa de 224 Kbps [WIL 2000].

Em 1990 o ITU-T aprovou o primeiro padrão de interoperabilidade para sistemas de videoconferência (recomendação H.320), documentando a implementação de tais sistemas em redes ISDN. A partir de 1992, diversos fabricantes começaram a apresentar no mercado sistemas para *desktop*, mas a um custo proibitivo – os primeiros variavam entre 10 e 15 mil dólares. Além do custo, esses produtos tiveram problemas de aceitação por causa da disponibilidade de serviços ISDN. Em 1996, o ITU-T apresentou três recomendações: a recomendação T.120 (que trata-se, na verdade, de um conjunto guarda-chuva de protocolos) para conferência de documentos e compartilhamento de aplicações; e os dois padrões para videoconferência usando o sistema de telefonia PSTN (H.324) e em redes IP (H.323). Este último teve como principal objetivo atender à plataforma de rede de maior capilaridade e base instalada: rede baseada em pacotes, sem garantia de serviço (tipicamente IP) [IMT 99][VSI 99].

3.2 Especificações de Tipos de Serviços Multimídia

O ITU-T, em sua série de recomendações F.700, apresenta o conceito de aplicações de comunicação multimídia, das quais a videoconferência é um caso muito representativo. Especificamente, a recomendação F.702 define serviços de conferência multimídia da seguinte forma [ITU 96]:

“Os serviços de conferência multimídia provêm transmissão em tempo real de voz em conjunto com vídeo e/ou vários tipos de informação multimídia entre grupos de usuários em duas ou mais localidades. (...) Quando estão presentes imagens em movimento, sua qualidade deve ser no mínimo suficiente para a representação adequada de movimentos fluentes de um pequeno grupo de participantes. A mídia de áudio é obrigatória, e uma ou mais das mídias vídeo, texto, gráficos e imagens estáticas deve estar presente”.

Ainda conforme a mesma recomendação, são apresentados os seguintes conceitos:

- Serviço de Videoconferência – é um serviço de conversação audiovisual que provê transferência bidirecional e em tempo real de voz e vídeo em movimento entre grupos de usuários em duas ou mais localidades separadas. Embora as informações sobre o áudio e o vídeo sejam parte essencial do serviço, o intercâmbio de outros tipos de informação, como imagens estáticas em alta resolução, texto ou gráficos também deve ser suportado.
- Serviço de Videofone – é um serviço de conversação audiovisual que provê transferência bidirecional simétrica e em tempo real de voz e vídeo em movimento entre grupos de usuários em duas ou mais localidades separadas. O requisito mínimo deste tipo de sistema é que, sob condições normais, a informação de imagem transmitida seja suficiente para a representação adequada de movimentos fluentes de uma pessoa, considerando a cabeça e os ombros.

Especificamente no que diz respeito à videoconferência, a recomendação F.730 reforça o conceito de aplicação multimídia e a obrigatoriedade do áudio definidos na recomendação F.702, e apresenta conceitos importantes para a classificação dos sistemas. Em relação à qualidade, estes podem ser classificados como:

- Serviço de videoconferência básica – Emprega taxas de transferência mais baixas, até um limite superior de 2,048 Mbps. Devido à restrição a velocidades menores, a largura de banda consumida (principalmente pelo vídeo) deve ser reduzida por meio de compressão. Como resultado, a resolução temporal e espacial das imagens em movimento percebidas nesse tipo de serviço são menores que em uma televisão normal.
- Serviço de Videoconferência de Alta Qualidade – Esses serviços provêm comunicação fim-a-fim de imagem colorida em movimento com alta resolução espacial e temporal (similar ou superior à qualidade de vídeo de televisão convencional), e de áudio equivalente a transmissão estéreo. Opcionalmente, devem prover meios para transferência de gráficos de alta resolução, texto e dados. Esses serviços podem valer-se tanto de redes de televisão analógica como de redes digitais de banda larga (B-ISDN).

No que diz respeito à rede, os sistemas de videoconferência podem trabalhar através das seguintes configurações:

- Ponto-a-Ponto : dois pontos de videoconferência estão conectados diretamente, e o gerenciamento da conferência é feito mediante negociação entre os pontos.
- Multiponto Multicanal : os canais de vídeo de três ou mais pontos de videoconferência são conectados dois a dois; cada terminal permanentemente recebe as imagens de todos outros pontos. Uma Unidade de Controle Multiponto (MCU) é usada para inserir os canais de áudio e gerenciar a conferência.
- Multiponto com Canal Compartilhado : novamente é utilizada uma MCU, que recebe sinais de todos os pontos e os combina para elaborar um sinal único a ser enviado de volta a cada terminal. Isso é feito por multiplexação em um canal de alta velocidade.
- Multiponto com Vídeo Comutado : esta configuração requer no mínimo uma MCU, com as seguintes funções :
 - Receber sinais de cada terminal;
 - Selecionar a imagem a enviar para cada terminal, com base em regras predeterminadas;
 - Adicionar, para cada terminal, os sinais de áudio dos demais. Esse processo deve ser limitado a alguns poucos terminais, selecionados pelo usuário, para evitar que se torne impossível distinguir os sons simultâneos.

Para mensurar a qualidade do serviço de videoconferência prestado, devem ser levados em conta, de acordo com a recomendação F.730, os seguintes fatores:

- Qualidade de Vídeo – trata-se da medida da capacidade de um sistema de transmissão de vídeo de reproduzir corretamente a imagem em movimento. Os objetivos de qualidade de vídeo são geralmente expressos em termos de resolução espacial e temporal, e podem variar conforme a aplicação. Em sistemas básicos,

trata-se de disponibilizar no mínimo os formatos CIF e QCIF; em sistemas de alta qualidade, a percepção do vídeo deve ter no mínimo a qualidade de televisão comum.

- Qualidade de Áudio – em qualquer situação, a qualidade de áudio a ser disponibilizada deve ser no mínimo na qualidade do serviço de telefonia normal, a 3,1 kHz (conforme a recomendação G.711). Contudo, são indicados como ideais os valores de 7 kHz (G.722) para serviços básicos e 15 kHz (transmissão estéreo) para serviços de alta qualidade.
- Sincronização – não deve existir atraso diferencial perceptível entre os sinais de áudio e vídeo.
- Atraso geral – é definido como o atraso de transmissão (introduzido pelos codecs, MCUs e recursos de transmissão) somado ao atraso característico (atraso introduzido pelo terminal quando apenas os lábios e olhos do usuário estão se movendo). Esse atraso deve ser levado em conta, pois pode prejudicar a funcionalidade do sistema por não ultrapassar o tempo de validade da informação.
- Distúrbios causados pela troca de fonte de vídeo – quando a fonte de vídeo é trocada, o codec necessita de algum tempo para construir a nova imagem. Este tempo é considerado como um fator de qualidade, e deve ser o menor possível, por causar uma perturbação visível e desagradável.

3.3 Popularidade Versus Dificuldades

O que vem tornando a videoconferência cada vez mais difundida é a possibilidade de acesso da mesma por parte de pequenas e médias empresas e, principalmente, pelo usuário final. O impulso à disseminação de técnicas de videoconferência se deve principalmente a fatores como:

- Possibilidade de conexão – Conectividade nos dias de hoje é um fato comum, tanto para empresas (pelo aumento significativo do emprego de LANs, presentes em empresas de qualquer porte) como para usuários finais (graças ao explosivo crescimento da Internet nos últimos anos) [BOR 2000].
- Custo do equipamento – Atualmente, é possível participar de uma sessão de videoconferência a partir de um computador pessoal (*desktop*), utilizando recursos de *hardware* bastante acessíveis : câmeras do tipo WebCam ou QuickCam, cujo custo pode ficar abaixo de 100 dólares, e dispositivos de som que normalmente já se encontram instalados no computador [FAY 99].
- Custo do *software* – Na realidade atual, os sistemas que tratavam as sessões de videoconferência baseados em estúdios e enlaces N-ISDN ou B-ISDN ganharam a companhia de programas de custo acessível, e muitas vezes gratuitos. Exemplos são o NetMeeting, da Microsoft, e a versão 0.92b do CU-SeeMe, da White Pine Software. Mesmo considerando que muitas vezes não se tratam de produtos com a intenção de competir no mesmo nicho de mercado (existem soluções mais

adequadas quando se trata de alta performance e confiabilidade), esses sistemas é que têm difundido mais amplamente a tecnologia de videoconferência [TRE 97].

- Implementação de codecs via *software* – Novamente, se forem comparados os sistemas existentes há 10 anos atrás, onde se utilizavam codecs baseados em *hardware* específico, com os sistemas atuais – que têm apresentado codecs via *software*, possíveis graças ao aumento do potencial de processamento dos PCs – percebe-se uma nova redução de custo, colaborando para facilitar o acesso aos sistemas de videoconferência [ZAN 99].
- Avanços na própria tecnologia de videoconferência – Acompanhando a evolução das técnicas de comunicação de dados, como os já citados avanços das LANs e da Internet, o ITU-T publicou recomendações que atendiam ao nicho de mercado em crescimento, a partir da recomendação original H.320 [BOR 2000].

Dentro desse contexto de possibilidades, a maior perspectiva de popularização da tecnologia de videoconferência é na plataforma de computadores pessoais (videoconferência em *desktop*) [DAV 2000a]. Paradoxalmente, a popularização em tais estruturas traz algumas dificuldades, como por exemplo:

- Uso de codecs em *software* pode ser mais barato, mas normalmente possui desempenho inferior aos codecs implementados em *hardware*. Da mesma forma, equipamentos de baixo custo (como câmeras de interface paralela ou USB) possuem desempenho aceitável, mas inferior a outros equipamentos que disponibilizam recursos mais específicos. Uma comprovação desse fato foi apresentada por Zanin [ZAN 2000], ao comparar a taxa de quadros por segundo obtida por uma câmera conectada ao computador via interface paralela à taxa obtida por outro modelo de câmera, conectada por intermédio de placa específica para captura e codificação. É interessante observar que a amplitude dessa diferença está relacionada a diversos fatores, como por exemplo a capacidade de processamento da estação, quantidade de memória RAM, sistema operacional e a implementação do codec.
- No caso de acesso pelo usuário doméstico ou pessoal, a limitação de banda para acesso à Internet, ainda bastante significativa, apesar da crescente oferta de acesso por banda larga [BER 2000].
- A falta de suporte a QoS, tanto para o acesso à Internet quanto no *backbone* (antes e depois do ISP).

Como toda aplicação multimídia, as aplicações de videoconferência possuem requisitos diferenciados de Qualidade de Serviço. No caso específico de videoconferência, no entanto, existe a preocupação adicional com o tempo demandado pelos dados para chegar ao seu destino (o que caracteriza videoconferência como uma aplicação *time-sensitive*, ou sensível ao tempo). Essa preocupação se traduz basicamente nos efeitos sobre o resultado da sessão que podem ser percebidos por influência de fatores como a latência (atraso fim-a-fim), jitter (variação da latência), skew (latência diferenciada para cada mídia, causando perda de sincronismo), erros e descartes da rede. Os dois últimos itens apresentados acima – relacionados à infraestrutura de comunicação – influem diretamente nesses problemas.

Com base no exposto, concluímos que a busca por soluções que minimizem o impacto dos problemas derivados de infraestrutura é fator imperativo para o provimento de serviços e produtos de videoconferência com qualidade e, conseqüentemente, o sucesso e aceitação da tecnologia como um todo.

4 Padrões para Implementação e Protocolos de Suporte

As recomendações da série F do ITU-T especificam serviços multimídia de forma genérica. O caso particular de videoconferência é abordado pelo ITU-T pelas recomendações da série H que serão apresentadas neste capítulo, bem como as recomendações por elas referenciadas. Estas recomendações são referenciadas pelo ITU-T como guarda-chuva, por serem compostas basicamente por três grupos de protocolos, cada um referenciando um aspecto diferente da aplicação de Videoconferência. O primeiro grupo contém os padrões do ITU-T para compressão de vídeo (H.261, H.263 e H.263+), os quais definem os algoritmos de compressão e resolução de vídeo usados em videoconferência. O segundo grupo enfoca as necessidades de áudio para aplicações de Videoconferência, e inclui os codecs G.711, G.722, G.723, G.728 e G.729. O terceiro grupo trata das características de controle da aplicação (série T.120) [VSI 99].

Além destes, a tecnologia de videoconferência prevê a utilização de diversos protocolos de suporte, basicamente definidos pelo IETF na forma de RFCs, como RTP/RTCP e protocolos de sinalização (SIP/SAP/SDP).

Este capítulo apresenta essas recomendações e protocolos, considerados fundamentais no entendimento dos sistemas de videoconferência.

4.1 Suporte a Vídeo

O processo de digitalização de vídeo é composto pela captura de um sinal analógico e sua conversão em tráfego de vídeo digital, usando dispositivos de *hardware* específicos, disponíveis atualmente para a maioria das plataformas. Para este processo, os seguintes componentes devem ser considerados :

- Resolução : tamanho horizontal e vertical da sessão de vídeo. Uma tela cheia (*full screen*) é tipicamente definida como 640 pixels horizontais por 480 pixels verticais – equivalente à taxa de 4:3, aspecto padrão de televisão.
- Quantidade de Cores : na realidade, este componente refere-se à quantidade de bits usados para compor a cor. Essa quantidade varia de 8 bits (256 cores) até 24 bits (usando 8 bits para vermelho, 8 bits para verde e 8 bits para azul, resultando 16.7 milhões de cores).
- Taxa de Quadros por Segundo (qps) : o número de quadros exibidos por segundo. Para apresentar vídeo em qualidade NTSC, por exemplo, são exibidos 30 quadros por segundo. Para os padrões PAL e SECAM, são 25 quadros por segundo [CIS 99]. A velocidade de apresentação de cinema fica em 24 qps, e aplicações de videoconferência apresentam resultados satisfatórios a partir de uma taxa de 6 qps.

Com base nesses parâmetros, é possível determinar a banda necessária para um determinado tráfego de vídeo. Supondo um vídeo em modo tela cheia, 16.7 milhões de cores e uma taxa de 30 quadros por segundo. O tráfego gerado por esse vídeo, por segundo, é dado pela expressão :

$$640 * 480 * 3 * 30 = 27,648 \text{ Mbps}$$

É fácil perceber que isso tornaria impraticável a difusão de tecnologias de transmissão de vídeo. Para prover banda passante para esse tipo de aplicação, seria necessário no mínimo um canal Fast Ethernet (100 Mbps). Em vista disso, foram apresentadas alternativas para viabilizar tal tráfego em redes mais popularmente difundidas:

Manipulação de vídeo

Trata-se da alteração dos parâmetros resolução, quantidade de cores e taxa de quadros da imagem, com o objetivo de diminuir consumo de banda. Por exemplo, algumas aplicações multimídia realizam captura de vídeo a 320 x 240, com cores de 8 bits e uma taxa de 15 quadros por segundo. Com esses requisitos, a banda consumida é dada pela expressão :

$$320 * 240 * 1 * 15 = 9,216 \text{ Mbps}$$

Compressão de Vídeo

Trata-se da substituição transparente das informações originais em pixels por descrições matemáticas mais compactas. O componente responsável pela compressão e descompressão chama-se **codec** (Codificador / Decodificador ou Compressor / Descompressor), que pode ser implementado tanto em *software* quanto em *hardware*. Embora boa parte das placas de captura de vídeo incluam o codec, os processos de captura e compressão são mantidos separados. A taxa de compressão é medida levando em conta a relação entre os tamanhos dos arquivos – original e comprimido. Assim, um arquivo de 60 kbytes, que ao ser comprimido gerou um arquivo de 20 kbytes, sofreu uma compressão de 3:1. Existem dois tipos de técnicas de compressão :

- **Sem perda** : Este processo cria arquivos comprimidos que, quando descomprimidos, geram um arquivo exatamente igual ao original. Normalmente é usado em arquivos executáveis, ou onde os dados não podem ser alterados. O mecanismo consiste na identificação de padrões no conteúdo, de forma a poder descrevê-lo de forma mais eficiente. A taxa de compressão média obtida é de 2:1, o que é insuficiente para viabilizar tráfego de vídeo. Exemplos de produtos que usam essa técnica são compactadores, como o PKZIP, e compressores para aumentar a capacidade de discos de forma transparente, como o STACKER e o DoubleSpace.
- **Com perda** : Esse tipo de mecanismo cria arquivos compactados que parecem similares ao original. Quando na descompressão do arquivo gerado, alguns pixels ou cores são alterados, aproveitando algumas características de percepção do olho humano (que é mais sensível a alterações de luminosidade que a alterações de cor) [TRE 97]. Além disso, quando usada em vídeo, o efeito da compressão com perda é minimizado, pois cada imagem é exposta por uma fração de segundo (1/15 ou 1/30 de segundo, conforme a taxa de qps). Esse método não pode ser usado em arquivos executáveis, onde o resultado deve ser idêntico ao original; em compensação, as taxas de compressão podem variar de 2:1 até 300:1.

Considerando que o vídeo é uma seqüência ordenada de imagens, pode-se analisar a compressão de vídeo em dois âmbitos distintos : Interquadro (*interframe*) e Intraquadro (*intraframe*).

- Compressão interquadro (ou compressão temporal) : Neste método a compensação do movimento é produzida pela compensação das diferenças entre os quadros [TRE 97]. É utilizado um quadro chave (*key frame*) e diversos quadros de variação (delta frames), que contêm apenas as diferenças causadas pelo “movimento”. Durante a descompressão, os quadros de variação buscam a informação em quadros anteriores: a maioria das implementações completa os quadros de variação com base no quadro anterior (de qualquer tipo), enquanto implementações como MPEG referenciam o último quadro chave. Todas as técnicas intraquadro baseiam sua efetividade na redundância entre os quadros – quanto mais redundantes entre si, maior a eficiência deste tipo de método [CIS 99].
- Compressão intraquadro (ou compressão espacial) : São aplicados recursos de compressão em cada quadro componente do vídeo (seja ele um quadro chave ou de variação) já processado pelo método interquadro. Embora o método intraquadro normalmente receba mais atenção, a performance geral do codec é mais beneficiada pela compressão interquadro (temporal). Os métodos de compressão intraquadro existentes são os mesmos aplicados a imagens estáticas, como o RLE (desenvolvido originalmente para fax) e o JPEG.

Padrões do ITU-T : H.261 e H.263

A recomendação H.261, aprovada em 1993, descreve métodos de codificação e decodificação de imagens em movimento em taxas de bits múltiplas de 64 kbit/s. Esse padrão também é referenciado como $p \times 64$ Kbits/s, onde p varia de 1 a 30 (a taxa de bits varia de 64 kbit/s a 2 Mbit/s). Esse codec foi projetado inicialmente para linhas ISDN, que operam nesta faixa. A origem do vídeo é tomada como sendo figuras não-entrelaçadas, a uma freqüência de 30000/1001 (aproximadamente 29,97) vezes por segundo. As imagens são formadas por um atributo de luminosidade e dois atributos de diferença de cor (respectivamente Y , C_B e C_R).

O algoritmo de codificação usado é um híbrido de predição interquadro e intraquadro. A codificação interquadro remove redundância temporal – a compensação do movimento é produzida pela computação das diferenças entre os quadros. A codificação intraquadro, por sua vez, utiliza a codificação de transformação para remover redundância espacial. Inicialmente, as redundâncias temporais entre imagens sucessivas são removidas. As imagens restantes são transformadas usando a técnica DCT. O algoritmo de codificação H.261 é similar ao MPEG, porém estes são incompatíveis entre si. Além disso, o H.261 necessita substancialmente menos poder de processamento para codificação em tempo real do que o MPEG [IMT 99].

São especificados dois formatos de imagem para este padrão:

– CIF (*Common Intermediate Format*) – neste formato, a estrutura de amostragem de luminosidade é de 352 colunas por 288 linhas. A amostragem de cada um dos componentes de diferença de cor é feita em 176 colunas por 144 linhas. A área de

imagem coberta por essas quantidades de colunas e linhas possui uma proporção 4:3, que corresponde à área ativa das entradas de vídeo padrão.

– QCIF (*Quarter-CIF*) – esse formato possui a metade das colunas e linhas do formato CIF. Sua implementação é obrigatória nos codecs que cumpram a especificação H.261, ao contrário do CIF (opcional).

A recomendação H.263, publicada em 1996, especifica a representação codificada que pode ser usada para compressão de imagem em movimento, para aplicações audiovisuais em baixas taxas de bit. Embora os primeiros projetos especifiquem taxas inferiores a 64 kbit/s, essa limitação foi removida, e a recomendação H.263 tende a substituir a H.261 em diversas aplicações [IMT 99].

O algoritmo de codificação é baseado (e muito similar) no algoritmo da recomendação H.261, com a inclusão de algumas opções de codificação para incremento de performance e mecanismos de recuperação de erro. O ITU-T apresenta uma comparação em relação à performance da recomendação H.263, considerada bem superior à da H.261 [ITU 97].

Além dos dois formatos apresentados no padrão H.261, outros três formatos foram apresentados. Suas dimensões e obrigatoriedade de implementação estão relacionadas na TABELA 4.1, juntamente com os formatos da recomendação anterior.

TABELA 4.1 – Formatos de Imagem das Recomendações H.261 e H.263

Formato	Colunas (dx)	Linhas (dy)	Suporte em H.261	Suporte em H.263
Sub-QCIF	128	96		Obrigatório
QCIF	176	144	Obrigatório	Obrigatório
CIF	352	288	Opcional	Opcional
4CIF	704	576		Opcional
16CIF	1408	1152		Opcional

4.2 Suporte a Áudio

Por definição, as aplicações destinadas a videoconferência incluem suporte a áudio. Na realidade, trata-se de um elemento que merece atenção especial. Ao invés do vídeo, onde a perda de um quadro na maioria dos casos não chega a ser prejudicial à aplicação, a perda de trechos de áudio em uma sessão de videoconferência pode comprometer seriamente a interatividade. Conforme Tanenbaum, o ouvido é sensível a variações de som que duram milésimos de segundo, tempos inferiores aos necessários para o olho perceber variações de luz [TAN 97]. Desta forma, um jitter de apenas alguns milésimos de segundo em uma transmissão multimídia afeta mais a qualidade do som percebido do que a qualidade da imagem percebida.

Assim como vídeo digital, o áudio sempre é obtido a partir de fontes analógicas, necessitando a conversão analógico-para-digital (A/D). A conversão de um sinal analógico em um sinal digital envolve a captura de uma série de amostras da fonte analógica. A agregação das amostras forma o equivalente digital de uma onda sonora analógica. Quanto maior a taxa de amostragem, maior a qualidade do som digital, pois

foram utilizados mais pontos de referência para replicar o sinal analógico. O processo de conversão A/D pode ser dividido em três fases: amostragem, quantização e compressão.

Amostragem

Trata-se do processo de medir valores instantâneos de um sinal analógico em intervalos regulares [CHA 98]. O intervalo entre as amostras é determinado por um pulso de clock, e a frequência deste clock é chamada Taxa de Amostragem.

Em 1928, Henry Nyquist provou que a representação digital de um sinal analógico poderia ser funcionalmente idêntica à onda que o originou, contanto que a taxa de amostragem fosse no mínimo o dobro da frequência mais alta presente naquela onda (Teorema de Nyquist). Por exemplo, a voz humana, com uma frequência máxima de 4 kHz, requer 8.000 amostras por segundo para se obter uma representação digital fiel.

Quantização

Trata-se do processo de conversão das amostras contínuas para valores discretos. O uso de oito bits por amostra é considerado satisfatório para áudio em diálogo comum, mas amostras de áudio de alta fidelidade são normalmente digitalizadas a 14 ou 16 bits por amostra, provendo uma granularidade maior.

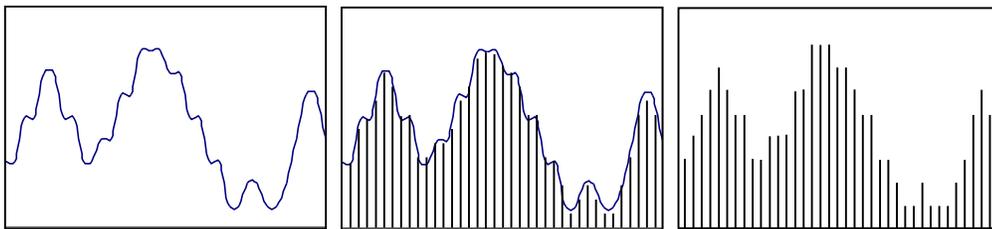


FIGURA 4.1 – (a) Onda de Áudio. (b) Amostragem. (c) Quantização

A Figura 4.1 ilustra como uma onda de áudio (a) é amostrada (b) e quantizada com 3 bits (c) em valores discretos. Convém observar que amostras com diferenças na representação analógica podem ser atribuídas ao mesmo valor discreto após o processo de quantização, devido ao número finito de possibilidades (com 3 bits, podem-se representar 8 quantizações diferentes). Esse erro, introduzido pela quantidade finita de bits, é chamado de Erro de Quantização, e pode ser detectado pelo ouvido humano se for muito grande.

Um exemplo das duas fases iniciais de conversão A/D de som são os CDs de áudio. Estes são codificados em 44.100 amostras por segundo, capturando frequências de até 22.050 Hz (conforme o teorema de Nyquist). São usadas amostras de 16 bits cada, possibilitando 65.536 valores distintos – embora a capacidade de percepção do ouvido humano, considerando intervalos do mais baixo som audível, chegue a 1 milhão (o que comprova que CDs de áudio apresentam ruídos de quantização). Caso se

considerasse a transmissão de áudio de CD, com 44.100 amostras por segundo de 16 bits, a largura de banda necessária seria

$$44.100 * 16 = 705,6 \text{ kbps para som mono, e}$$

$$44.100 * 16 * 2 = 1,411 \text{ Mbps para som estéreo.}$$

A transmissão de som com qualidade de CD em estéreo, conforme apresentado anteriormente, é capaz de demandar um canal T1 completo - o que comprova a necessidade de compressão de tais dados, mesmo que em videoconferência a qualidade de som utilizada seja inferior, em termos de amostras por segundo e bits por amostra.

A TABELA 4.2 apresenta um resumo das características das Recomendações do ITU-T relacionadas a codificação de áudio e referenciadas nas recomendações da série H. São apresentados o tipo de algoritmo usado por cada codec, a taxa de bits a qual se destina, a frequência de trabalho, o atraso típico entre dois pontos (exclusivamente do codec, sem considerar o canal) e aplicações típicas. Maiores detalhes sobre os codecs de áudio serão fornecidos a seguir, em suas respectivas descrições.

TABELA 4.2 – Codecs de Áudio do ITU-T. Fonte : [CRO 97][CHA 98]

Recomendação (Ano de Aprovação)	Algoritmo	Taxa de Bits (kbit/s)	Largura de Banda (Hz)	Atraso entre Pontos (ms)	Aplicação
G.711 (1972)	PCM	56, 64	300 – 3,4k	< 1	Telefonia GSTN, Videoconferência H.320 / H.323
G.722 (1988)	Sub-ADPCM	48, 56, 64	50 – 7k	< 2	Telefonia e Videoconferência ISDN
G.723.1 (1995)	ACELP – 5,3 MP-MLQ – 6,3	5.3, 6.3	300 – 3,4k	67-97	Videotelefonia GSTN, Telefonia H.323, VoIP (básico)
G.728 (1992)	LD-CELP	16	300 – 3,4k	< 2	GSTN, Videoconferência H.320
G.729 (1995)	CS-ACELP	8	300 – 3,4k	25-35	Telefonia GSTN, Modem GSTN, Videotelefonia H.324 GSTN
G.729 – A (1996)	CS-ACELP	8	300 – 3,4k	25-35	Modem GSTN, Videotelefonia H.324 GSTN

Modulação por Código de Pulso

G.711

A recomendação G.711, editada inicialmente em 1972, especifica a modulação de sinais de frequência de voz por código de pulso, por isso também é referenciada como PCM (*Pulse Code Modulation*). A codificação é realizada a uma taxa nominal de 8000 amostras por segundo (8 kHz), utilizando 7 ou 8 bits por amostra. A taxa de transmissão dessa codificação resulta em $7 \times 8000 = 56 \text{ kbit/s}$ ou $8 \times 8000 = 64 \text{ kbit/s}$. Novamente de acordo com o teorema de Nyquist, a especificação G.711 pode codificar frequências de áudio entre 0 e um máximo de 4 kHz.

As regras de codificação são referenciadas como A-law (utilizado na Europa) e μ -law (utilizado na América). Apesar do formato resultante da codificação se parecer com uma função logarítmica, a conversão dos sinais é definida por tabelas de decisão. São especificadas quatro tabelas diferenciadas, para valores A-law / μ -law, tanto positivos quanto negativos.

O áudio de entrada é dividido em segmentos, cada um deles usando quantidades diferentes de intervalos como valores nas tabelas de decisão. A maior parte dos segmentos contém 16 intervalos, e o tamanho dos intervalos dobra de segmento a segmento. A Figura 4.2 mostra três segmentos com quatro intervalos em cada um. A codificação μ -law usa 8 segmentos de 16 intervalos cada nas direções positiva e negativa, começando com um tamanho de intervalo 2 no segmento 1, e incrementando até um tamanho de intervalo 256 no segmento 8. A codificação A-law usa 7 segmentos. O menor (de tamanho de intervalo 2) contém o dobro de intervalos dos demais (32 intervalos). Os seis segmentos restantes contém 16 intervalos cada, dobrando de tamanho de 4 (segmento 2) até 128 (segmento 7). A codificação A-law, usada internacionalmente, representa sinais menores com maior fidelidade que a μ -law. A recomendação também apresenta tabelas de conversão μ -law \Leftrightarrow A-law.

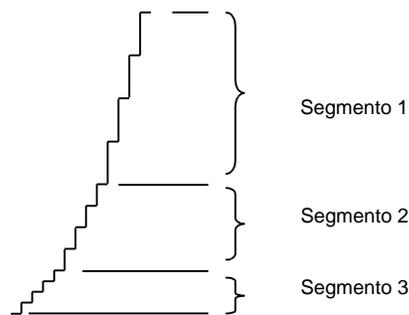


FIGURA 4.2 – Segmentação de uma Onda de Áudio codificada com PCM

Algoritmos baseados em Diferenças

A partir da técnica PCM, foram propostas algumas implementações diferenciadas. Os codificadores e decodificadores DPCM e ADPCM funcionam baseados na expectativa de que as amostras de áudio vizinhas sejam semelhantes entre si. Assim, o codificador DPCM calcula a diferença entre o valor efetivo da amostra e o valor provável (com base na comparação pelas amostras anteriores), e realiza a codificação PCM apenas sobre o resultado da diferença. Desta forma, a taxa de bits gerada é inferior à gerada pela recomendação G.711.

O método ADPCM emprega uma combinação de adaptações de quantização e predição sobre o método DPCM. Adaptação, nesse contexto, significa reação a alterações no nível e no espectro do sinal de áudio de entrada. Existem duas formas de realizar a predição adaptativa :

- Estimativa avançada (*forward estimation*), onde as amostras quantizadas do sinal de entrada são usadas para derivar estimativas;
- Estimativa invertida (*backward estimation*), onde as estimativas são geradas com base na saída do quantificador e no erro de predição ocorrido.

G.722

A recomendação G.722, proposta em 1988, descreve as características de sistemas de codificação de áudio (50 a 7000 Hz), usados em diversas aplicações de conversação de alta qualidade. O sistema de codificação utiliza a técnica SB-ADPCM (*sub-band adaptive differential pulse code modulation*), a uma taxa de 64 kbit/s. Nesta técnica, a banda de frequência é dividida em duas sub-bandas : alta (H) e baixa (L). Os sinais em cada sub-banda são codificados usando ADPCM. O sistema possui três modos de operação para codificação do áudio a 7 kHz, com variações na taxa de bits e na taxa de bits do canal de dados auxiliar (para prover a taxa efetiva de 64 kbit/s usando os bits da sub-banda inferior. Os modos de operação disponíveis são apresentados na TABELA 4.3.

TABELA 4.3 – Modos de operação da recomendação G.722

Modo	Taxa de bits de áudio a 7 kHz codificado	Taxa de bits do canal de dados auxiliar
1	64 kbit/s	0 kbit/s
2	56 kbit/s	8 kbit/s
3	48 kbit/s	16 kbit/s

Conforme a recomendação, independente do modo de operação escolhido para a implementação, é sugerido que o modo 1 sempre seja implementado, mesmo que de forma adicional ao(s) modo(s) escolhido(s).

Codificação por Predição Linear por Ativação de Código

G.723.1

A Recomendação G.723.1, editada em 1996, especifica uma representação codificada para compressão de diálogo ou outros sinais de áudio em serviços multimídia a taxas de bits muito baixas, com o objetivo de servir como suporte de áudio à recomendação H.324. As taxas de bit associadas são 5,3 kbps (mais flexível) e 6,3 kbit/s (melhor qualidade). A codificação do áudio é realizada em quadros de 30 milissegundos. O formato G.723.1 foi selecionado pelo Fórum VoIP (Voice over IP) como o codec base para aplicações de Voz sobre IP a baixas taxas de bit [CHA98].

Esse codificador foi projetado para operar com sinais digitais, obtidos a partir da filtragem da fonte analógica (a filtragem é definida na Recomendação G.712), e então amostradas a 8000 Hz e convertidas para PCM em 16 bits para a entrada do codificador.

A saída do decodificador realiza o processo inverso. O princípio da codificação é a predição análise-por-síntese, seguida de mecanismos para minimizar erros perceptíveis. O codificador trabalha em blocos (quadros) de 240 amostras cada (8000 amostras por segundo x 0,030 segundos).

Os componentes do algoritmo apresentado na Recomendação são definidos em forma de função na própria Recomendação, que inclui também os diagramas de bloco (representação visual da interação dos componentes) para cada componente.

G.728

A recomendação G.728, publicada em 1992, provê codificação de áudio a 3,1 kHz para transmissão a 16 kbps. É normalmente utilizada em sistemas que operam a 56 ou 64 kbps. Com requerimentos computacionais bem mais altos, o padrão G.728 provê a qualidade do padrão G.711 – a um quarto da taxa de transmissão. O algoritmo utilizado para a codificação é o LD-CELP (*low-delay excited linear prediction*).

G.729 e G.729a

As recomendações G.729 e G.729a, aprovadas respectivamente em 1995 e 1996, especificam a codificação de sinais de áudio na faixa de 3,1 kHz, para transmissão a 8 kbps. A recomendação G.729A exige menos poder computacional que a G.729; ambas possuem latência menor que a G.723.1. Existe uma boa expectativa do uso da recomendação G.729A para transmissão de áudio compactado em redes sem fio (*wireless*).

4.3 Suporte a Mídias Discretas

A definição do suporte a mídias discretas é especificada pelo ITU-T através do padrão T.120. Este padrão é composto, na realidade, por uma série de recomendações, que tratam de áreas específicas da teleconferência multimídia – a conferência de documentos e o compartilhamento de aplicações. Essas recomendações especificam como distribuir arquivos e informações gráficas em tempo real durante uma reunião multiponto, de forma eficiente e confiável. O objetivo desses padrões é assegurar interoperabilidade entre terminais sem que nenhum participante necessite conhecimento prévio do(s) outro(s) sistema(s) participante(s) da reunião, mesmo estando os participantes conectados por redes diferentes (a recomendação prevê PSTN, ISDN, CSDN, PSDN, B-ISDN e LANs). Não existe restrição em relação ao volume de dados que trafega, nem em relação às diferenças de taxas impostas por redes distintas, e cada usuário pode participar de uma ou diversas conferências simultâneas.

O escopo é permitir compartilhamento de dados incluindo compartilhamento de imagem (quadro branco), informação gráfica, troca de imagens, compartilhamento de aplicações e especificar protocolos de infraestrutura para aplicações audiográficas e audiovisuais. Além de dar suporte para as recomendações H.320 a H.324, a série T.120 pode ser usada isoladamente como base para desenvolvimento de aplicações de conferência (*data conferencing*) [IMT 99]. A série de protocolos T.120 provê:

- Suporte para estabelecimento de conferência entre um grupo de nós de rede (terminais ou MCUs).
- Mecanismos para identificação dos nós participantes, e sua manutenção em uma lista dinâmica.
- Gerenciamento flexível da comunicação entre qualquer um desses elementos.

As recomendações que compõem o padrão são categorizadas na própria recomendação T.120 como Recomendações de Infraestrutura (T.123, T.122/T.125 e T.124) e Recomendações para Protocolo de Aplicações (T.121, T.126 e T.127). O grupo de Infraestrutura é considerado mandatório; as demais recomendações são de obrigatoriedade condicional.

4.4 Recomendações do ITU-T para Videoconferência

A grande maioria dos sistemas de Videoconferência e Videotelefonia foi desenvolvida de forma a cumprir as recomendações do ITU-T. Essas recomendações são centradas na série H – particularmente as recomendações H.310 e H.320 a H.324. A TABELA 4.4 apresenta um resumo destas recomendações, e as características de cada uma (como codecs de áudio, vídeo, padrões de controle e outras).

TABELA 4.4 – Padrões de Videoconferência do ITU-T. Fonte : [DAT 99]

Recomendação	H.320	H.310/321	H.322	H.323 v1/v2	H.324
Aprovação	1990	1995	1995	1996/1998	1996
Rede	N- ISDN	B-ISDN LAN ATM	Redes de comutação de Pacotes com Garantia de QOS	Redes de comutação de Pacotes sem Garantia de QOS (Ethernet)	PSTN ou POTS, sistema de telefone analógico
Interface de Comunicação	I.400	AAL I.363 AJM I.361 PHY I.400	I.400+ TCP/IP	TCP/IP	Modem V.34
Codecs de Vídeo	H.261 H.263	H.261 H.263	H.261 H.263	H.261 H.263	H.261 H.263
Codecs de Áudio	G.711 G.722 G.728	G.711 G.722 G.728	G.711 G.722 G.728	G.711 G.722 G.728 G.723 G.729	G.723
Dados	T.120	T.120	T.120	T.120	T.120
Multiplexação	H.221	H.221	H.221	H.225.0	H.223
Controle	H.230 H.242	H.242	H.242 H.230	H.245	H.245

H.320

O conjunto de recomendações referenciado a partir da H.320 cobre os requerimentos técnicos para os serviços e equipamentos de videotelefonia em banda estreita (N-ISDN), onde as taxas dos canais não ultrapassem 1920 kbit/s [ITU 97a]. Nesta recomendação são apresentados os conceitos básicos de comunicação de áudio, vídeo e gráficos aplicados a videotelefonia, especificando :

- requisitos para processamento de informação de áudio e vídeo,
- formatos comuns para compatibilidade de entrada/saída de áudio e vídeo e
- protocolos que possibilitam a terminais multimídia utilizar os canais de comunicação e sincronização de sinais de áudio e vídeo.

Assim como outros padrões de teleconferência multimídia, a recomendação H.320 é destinada a aplicação multiponto ou ponto-a-ponto, sobre serviços de circuitos comutados como N-ISDN ou Switched-56 [IMT 99].

A recomendação H.320 prevê inserção de atraso (*delay*) no codec de áudio, com a função de compensar o atraso imposto pelo codec de vídeo, visando manter o sincronismo entre o áudio e a imagem correspondente. Considerando que os atrasos no codificador e no decodificador podem variar de acordo com a implementação, a compensação gerada deve ser transferida juntamente com os dados. De acordo com a recomendação, a inclusão do atraso é opcional – o sistema pode ser implementado com o referido atraso, sem ele ou deixando a cargo do usuário a sua inclusão.

H.310 e H.321

Estas recomendações definem terminais audiovisuais para implementação em redes B-ISDN (ATM). A primeira foi aprovada em novembro de 1996, ou seja, antes da H.320, sob o título de “Sistemas e Terminais para Comunicação Audiovisual em Banda Larga”. A segunda foi aprovada em fevereiro de 1998, pelo mesmo grupo de estudo responsável pela recomendação H.320, e seu título original é “Adaptação de Terminais de Telefonia Visual H.320 a Ambientes B-ISDN”.

H.310

A definição dos terminais para esta recomendação foi realizada visando suportar as seguintes aplicações :

- Serviços de Conversação (como videoconferência e videotelefonia)
- Serviços de Recuperação
- Serviços de Mensagens
- Serviços de Distribuição **com** apresentação individual por receptor (*video-on-demand*)
- Serviços de Distribuição **sem** apresentação individual por receptor (*broadcast TV*)
- Transmissão de Vídeo
- Serviços de Vigilância e Monitoração

H.321

Este tipo de terminal pode ser utilizado para várias aplicações, sem recomendação de serviços em particular. São apresentadas duas possibilidades de implementação : usando um projeto integrado, que inclua a unidade H.320, o AAL e as funções ATM em um terminal ou utilizado um adaptador (B-TA, *Broadband Terminal Adapter*). Nesse caso, os sinais do terminal H.320 (no formato de quadros H.221) são transmitidos entre o terminal e o B-TA, que também recebe a responsabilidade de interação com a rede. Os codecs de áudio, vídeo e o gerenciamento de dados seguem as especificações da recomendação H.320.

Os terminais H.321 também devem ser aptos a comunicação com telefones N-ISDN e PSTN, utilizando para isso áudio no formato G.711.

H.322

A recomendação H.322 cobre os requerimentos técnicos para serviços de telefonia visual para banda estreita, em situações onde os canais de transmissão incluam uma ou mais LANs configurada(s) para prover garantia de QOS equivalente à do N-

ISDN, de forma que nenhum mecanismo de proteção ou recuperação de erro – além dos especificados na recomendação H.320 – seja necessário. Um exemplo de LAN confiável é a ISOEthernet (IEEE 802.9a). Redes locais utilizando ATM estão fora do escopo desta Recomendação, pois dizem respeito à recomendação H.321.

H.323

O padrão H.323 foi originalmente desenvolvido como uma adaptação do H.320. Desde a época de publicação da recomendação H.320 (1990) até os dias de hoje, foi possível assistir a explosão de utilização de redes locais (LANs) e gateways para redes de longa distância (WANs) [ITU 98a]. A recomendação H.323 surgiu como uma extensão lógica e necessária para o padrão H.320, incluindo sua utilização em :

- Redes locais (LAN)
- Redes Metropolitanas (MAN)
- Intranet's corporativas
- Internet
- Conexões do tipo dial-up ou ponto-a-ponto sobre canais GSTN ou ISDN que utilizem protocolos de transporte como PPP ou, de um modo geral, redes PBN (Packet Based Network), que não possuam mecanismos de garantia de QOS [IMT 99].

A especificação H.323 foi aprovada em 1996, e a versão 2 foi aprovada em 1998. O padrão é bastante flexível e abrangente, e inclui a implementação desde dispositivos *standalone* portáteis apenas para voz até estações completas de videoconferência. A DataBeam Corp. apresenta diversos fatores que salientam a importância da recomendação H.323 como padrão para aplicações de videoconferência [DAT 99]:

- A recomendação H.323 direciona os padrões de multimídia para a infraestrutura de rede existente (como redes baseadas em IP). Seu projeto de compensação do efeito da latência altamente variável apresentada em redes LAN possibilita ao usuário final o uso de aplicações multimídia sem alteração da infraestrutura de rede. Além disso, são providos mecanismos para interconexão entre redes diferentes, minimizando os efeitos de transcodificação.
- As redes baseadas em IP estão se tornando mais poderosas em relação à realidade da época da recomendação H.320 – a banda de 10 Mbps do padrão IEEE 802.3 está migrando para os 100 Mbps (Fast Ethernet, 802.3u) e para o Gigabit Ethernet. Convém salientar que embora esses padrões não sejam novidades, o que está sendo considerado aqui é sua disseminação em ambientes instalados.
- O projeto da recomendação H.323 prevê interoperabilidade entre dispositivos, aplicações e fornecedores diferentes, o que retira do usuário a preocupação de compatibilidade entre produtos que atendam à especificação.

- A carga de rede pode ser gerenciada. Além do suporte a multicast (que reduz o consumo de banda), o administrador da rede pode restringir a quantidade de banda passante disponível para as aplicações de conferência, bem como a quantidade simultânea de sessões H.323.

Arquitetura

A recomendação H.323 define quatro componentes fundamentais para sistemas de comunicação: terminais, gateways, gatekeepers e MCUs (Multipoint Control Unit).

Terminais são clientes na LAN que possuam capacidade de prover comunicação em tempo real nos dois sentidos. A Figura 4.3 apresenta o formato de um terminal conforme a recomendação. Com exceção dos componentes de Vídeo e Dados (opcionais), os demais são de implementação obrigatória – ou seja, em termos de mídia, a especificação obriga apenas a implementação de Áudio.

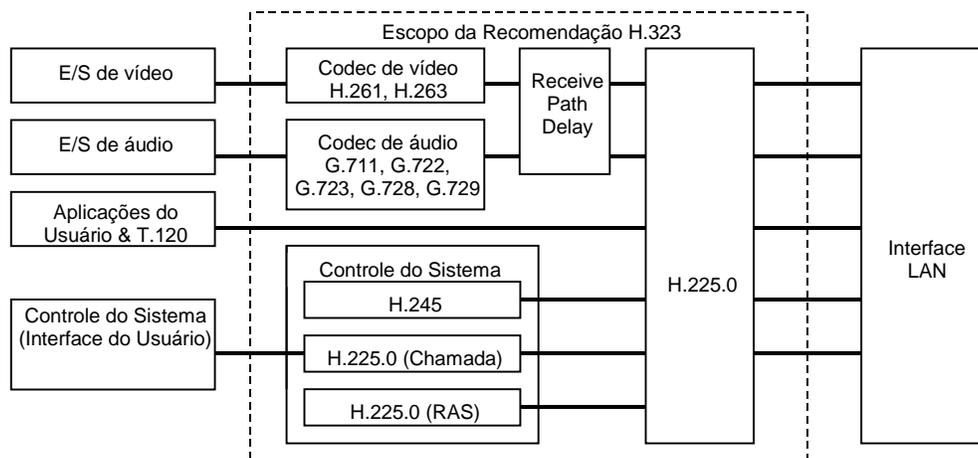


FIGURA 4.3 – Terminal Conforme a Recomendação H.323

- Gateway – Trata-se de um elemento opcional em conferências H.323, destinado a prover conversão entre formatos de transmissão (por exemplo, do formato H.225.0 para H.221 e vice-versa) e entre procedimentos de comunicação (por exemplo, de H.245 para H.242 e vice-versa). Além disso, também são funções do Gateway a tradução entre codecs de áudio e vídeo e a inicialização e encerramento da chamada
- Gatekeeper - Este componente funciona como ponto central para as chamadas dentro da sua zona, a qual é definida como sendo “o conjunto de Terminais, Gateways e MCUs gerenciados por um único Gatekeeper” [DAT 99]. Sua tarefa é prover serviços de controle de chamada para as estações registradas. Em alguns aspectos, o Gatekeeper funciona como um switch virtual.
- MCU - A Unidade de Controle Multiponto suporta conferências entre três ou mais estações. A Recomendação H.323 define uma MCU como um ou mais MC

(Multipoint Controller – mandatório) e zero ou mais MP (Multipoint Processor). O MC gerencia negociações entre todos os terminais (conforme a recomendação H.245), para determinar as capacidades de processamento de áudio e vídeo. O MC também é responsável por controlar os recursos de conferência determinando qual – se algum – dos fluxos de áudio e vídeo será transmitido em multicast. A manipulação – comutação e processamento – dos bits de áudio, vídeo e dados são de responsabilidade do MP.

Conferência Multiponto

Este tipo de capacidade é manipulado de diversas formas. Para efeitos de definição, a recomendação H.323 utiliza os conceitos de conferência Centralizada e Descentralizada.

Conferências centralizadas requerem a existência de uma MCU para implementar as funcionalidades de multiponto. Todos os terminais enviam tráfego de áudio, vídeo, dados e controle para a MCU, como se fosse ponto-a-ponto. O MC gerencia a conferência, usando funções de controle (definidas na recomendação H.245), que também definem as capacidades de cada terminal. O MP reúne sinais de áudio, vídeo e distribuição de dados, enviando os tráfegos resultantes de volta aos terminais. O MP também pode prover conversão entre codecs e taxas de bit diferentes. Uma implementação típica de MCU consiste de um MC e MPs para áudio, vídeo e dados.

Conferências descentralizadas podem usar tecnologia de multicast. Os terminais participantes realizam multicast dos tráfegos de áudio e vídeo para os outros participantes, sem enviar dados para a MCU. Contudo, o controle de multiponto continua sendo processado pela MCU, e as informações de controle de canal (H.245) continuam sendo enviadas em modo ponto-a-ponto para um MC. Neste caso, os terminais receptores são os responsáveis por processar os diversos tráfegos de áudio e vídeo que recebem – ou seja, a técnica de multicast faz uso mais eficiente da banda, as aumenta a carga computacional dos terminais, além de necessitar suporte nos roteadores e switches envolvidos.

Também existe a possibilidade de conferências multiponto mistas, onde alguns terminais estão em uma conferência centralizada e outros em conferência descentralizada. Nesses casos a MCU funciona como uma ponte entre os dois tipos.

Codecs para Vídeo e Áudio

O codec de vídeo é opcional. Todos os terminais H.323 que provenham suporte a comunicação de vídeo devem suportar codificação e decodificação conforme o formato QCIF definido em H.261. Opcionalmente, os terminais podem suportar outros modos das recomendações H.261 e H.263. Caso algum formato H.263 seja suportado, o formato QCIF do H.263 é obrigatório. Outros codecs e formatos de imagem, bem como a taxa de bits para vídeo e algoritmos opcionais que possam ser aceitos pelo decodificador são definidos durante a negociação de capacidades, usando o mecanismo H.245. Após a negociação, o codificador está livre para transmitir qualquer coisa dentro do conjunto de capacidade do decodificador.

O codec de áudio é obrigatório. Todos os terminais H.323 devem ser capazes de codificar e decodificar diálogo de acordo com a recomendação G.711, enviando e recebendo A-law e μ -law. Os terminais ainda podem utilizar as codificações G.722, G.728, G.729, Mpeg 1 / áudio e G.723.1, mediante negociação via H.245. Assim como o fluxo de vídeo, o fluxo de áudio é definido na recomendação H.225.0.

No caso de conferências H.323 executadas em linhas ou segmentos com baixa taxa de bit (< 56 kbit/s), o áudio no formato G.711 não pode ser transmitido. Nesses casos, os terminais devem obrigatoriamente possuir codificadores e decodificadores G.723.1 (terminais multimídia) ou G.729 (terminais exclusivos para voz).

Características Implementadas na Versão 2

Essa versão, destinada a resolver problemas e deficiências da versão 1, introduziu funcionalidades de protocolos já existentes (Q.931, H.245 e H.225), bem como protocolos novos. As principais alterações propostas são :

- Segurança – Foram definidos tópicos para tratamento de segurança : Autenticação – mecanismo para assegurar que um terminal realmente é quem diz ser; Integridade – método para validar que os dados em um pacote não foram alterados durante a transmissão;
- Privacidade – Provida por mecanismos opcionais de criptografia, definidos através da recomendação H.235.
- Chamada Rápida – Na versão anterior, o estabelecimento de uma chamada não disponibilizava o tráfego de dados imediatamente, o que resultava em um considerável atraso entre o momento da resposta da chamada e o início efetivo do diálogo. Nesta versão esse problema foi eliminado.
- Mecanismos para controle de QOS – A versão 2 implementa métodos de interface direta com a camada AAL 5 do ATM – para aproveitamento das características de garantia de QOS deste canal - e o uso do protocolo RSVP - para reserva de recursos.

H.324

Esta Recomendação define os requisitos para terminais multimídia operando a taxas de bit muito baixas, utilizando modems V.34 operando em GSTN. O objetivo é prover vídeo, áudio, dados ou qualquer combinação, em tempo real, entre **dois** terminais multimídia por uma conexão de rede GSTN para voz. Chang reconhece a importância dessa recomendação, qualificando-a como de enorme potencial por ser orientada ao serviço de telecomunicação mais largamente disponível no mundo – linhas de telefone analógicas PSTN [CHA 97]. Entretanto, existe uma credibilidade maior do mercado na recomendação H.323 em relação à H.324, em vista da sua portabilidade : a primeira pode ser implementada em PSTN via modem utilizando protocolos IP. No entanto, algumas estimativas consideram a plataforma à qual a recomendação H.324 se destina como uma importante fundação, onde pode se basear o amplo crescimento e a popularização da tecnologia do “videotelefone” no mercado doméstico.

A principal aplicação desta recomendação é em equipamentos de videofone, em uso isolado (*standalone*) ou em PCs menos potentes. Esses produtos teriam a tendência de utilização de codecs via *hardware*, em placas que incorporassem também as funções de modem – o que os colocaria em uma posição intermediária de custo, sem ser tão baratos quanto os produtos H.320/H.323, nem tendo o custo de aplicações de alta definição. A qualidade ideal também é diferenciada: espera-se do áudio qualidade de telefone, enquanto que o vídeo fica entre 4 e 12 quadros por segundo (dependendo do tamanho da janela).

Baseado na natureza das aplicações deste padrão, duas características são ressaltadas como fundamentais para os produtos H.324 :

- Interoperabilidade : considerando o tipo de aplicação citado (usuário final), é importante que os produtos sejam regularmente padronizados.
- *Voice-Call-First* : a aplicação a que se destinam esses equipamentos exige que a chamada de videotelefone seja tão simples quanto uma ligação telefônica normal. A proposta *voice-call-first* é similar ao mecanismo de transmissão de fax : após a chamada telefônica estar completa, o usuário pode optar por ativar o modo de videotelefone. Para tanto, deve ser suficiente pressionar um ícone, ou tocar um botão.

Videoconferência (H.323) e o Modelo OSI

Um exemplo de relação entre as arquiteturas para sistemas de videoconferência propostas pelo ITU-T e a arquitetura OSI pode ser tomado a partir da recomendação H.323. A FIGURA 4.4 apresenta a relação entre as duas arquiteturas.

APLICAÇÃO	H.323	
APRESENTAÇÃO	G.7XX/H.26X	
SESSÃO	H.225 (RAS)	H.225 (call setup) H.245 (call control) T.120/122/123/125
TRANSPORTE	RTP/RTCP	TCP
	UDP	
REDE	IP	
ENLACE	Enlace (PBN)	
FÍSICO	Físico	

FIGURA 4.4 – (a) Níveis do Modelo OSI. (b) Estrutura da Recomendação H.323

4.5 Outros Protocolos e Recomendações

Além dos codecs e especificações para comunicação de dados, outros padrões e protocolos são de importante conhecimento para a compreensão do funcionamento dos sistemas de videoconferência propostos nas recomendações do ITU-T, apresentadas na seção 4.4. A seguir são apresentados os mais relevantes.

- RTP/RTCP

O protocolo RTP, definido pelo IETF através da RFC 1889, foi projetado para oferecer suporte ao tráfego de aplicações multimídia em tempo real. Trata-se de um protocolo não orientado a conexão, que oferecem funções como identificação do tipo de dado, numeração e informação de tempo (*timestamp*). Embora seja direcionado a transporte fim a fim, o RTP não oferece funcionalidades comuns a protocolos de transporte – não reserva recursos, não garante QoS e não implementa reordenação ou retransmissão de pacotes, relegando estas funções à aplicação. Estas aplicações normalmente dispensam os serviços do TCP, visto que é preferível descartar dados do que introduzir atrasos com retransmissão de pacotes que, se retransmitidos, chegam ao destino sem validade.

O RTP é normalmente utilizado sobre o protocolo UDP, e se constitui de duas partes : o RTP, responsável pelo transporte das mídias contínuas, e o RTCP, que controla informações de retorno dos receptores, referentes a qualidade de recepção, sincronização entre mídias, descartes, jitter e outras. Maiores detalhes sobre os protocolos RTP/RTCP podem ser encontrados em [ZAN 99] e [WIL 2000], além das RFC que os definem (1889 e 1890, respectivamente). A aplicação do RTP enquanto provimento de suporte para busca de QoS na aplicação de videoconferência será abordado no Capítulo 6.

- SIP

O padrão H.323 foi inicialmente desenvolvido para o uso em redes locais (LAN), o que dificulta sua adaptação em situações de endereçamento para redes WAN e/ou Internet. Mesmo a versão 2 da recomendação H.323, que suporta zonas, possui problemas de escalabilidade para redes IP muito grandes. Visando resolver essas limitações, o grupo de trabalho MMUSIC (Multiparty MULTimedia SessIon Control) do IETF desenvolveu protocolos específicos para uso na Internet, referenciados de forma geral como SIP (*Session Initiation Protocol*).

Este padrão (na verdade, conjunto de padrões) provê mecanismos para estabelecimento e encerramento de chamadas, além de outras funções de controle. O SIP foi desenvolvido para uso em TCP ou UDP, e sua implementação é muito similar ao HTTP (enquanto o H.323 envia e recebe mensagens de controle binárias, o SIP implementa requisições e respostas textuais). Esta integração com o HTTP torna sua adaptação ao uso na Web bastante transparente [LUN 99]. Um ponto favorável ao uso do SIP para gerenciamento de sessões de comunicação multimídia é justamente a simplicidade: enquanto uma chamada H.323 pode demandar até vinte pacotes, a mesma chamada realizada via SIP pode ser realizada até com quatro cabeçalhos distintos, implementados como campos de cabeçalho HTTP.

Os outros protocolos do conjunto definido pelo SIP são: SDP e SAP, que tem a função de divulgação de descritores de sessão; RTSP, destinado ao controle de entrega de dados de tempo real sob demanda; e SCCP (*Simple Conference Control Protocol*), para gerenciamento de sessões. Detalhes sobre o funcionamento desta família de protocolos podem ser localizados em [LUN 99], [WIL 2000] e na RFC 2543.

5 Aspectos de Qualidade

O aspecto “qualidade” é fundamental em qualquer serviço, sobretudo nos aspectos referentes à sua aceitação. Neste capítulo, serão apresentados conceitos de qualidade de serviço, nos enfoques pertinentes aos sistemas de videoconferência: o conceito de Qualidade de Serviço, suas considerações em relação à infraestrutura à qual o trabalho é dedicado (Internet) e em relação a videoconferência, e os aspectos de sincronização e dimensionamento, enquanto qualidade do resultado da implementação do sistema.

5.1 Qualidade de Serviço

Um conceito importante no que diz respeito a aplicações sensíveis ao tempo é o de Qualidade de Serviço (QoS, abreviatura de *Quality of Service*). Apesar da ausência de uma definição formal única, de um modo geral pode-se definir QoS como a habilidade de um elemento da rede (roteador, servidor ou aplicação) prover algum nível de garantia de que os tráfego e os requerimentos de serviço possam ser satisfeitos. Tal habilidade requer a cooperação de todos os níveis da arquitetura de protocolos da rede (apesar de algumas definições atribuírem esta tarefa a níveis inferiores), bem como de todos os elementos fim a fim, o que implica que a garantia será tão efetiva quanto o ponto menos capaz da “cadeia” consiga prover [STA 99]. No caso da efetivação de uma conexão com necessidade de um determinado nível de QoS, a rede deve realizar uma verificação para determinar se este nível é possível de ser mantido. Se não for, o usuário deve ser questionado se o nível possível de ser oferecido pela rede é suficiente para sua aplicação.

Algumas definições mais precisas apresentadas para QoS são apresentadas a seguir.

- O padrão X.902 do ITU-T (*Information Technology – Open Distributed Processing – Reference Model*) apresenta QoS como “um conjunto de requerimentos de qualidade no comportamento coletivo de um ou mais objetos” [HUT 97]. Alguns parâmetros descrevem a velocidade e confiabilidade da transmissão de dados. Os parâmetros normalmente considerados para dimensionar QoS são [TAN 97] :
 - Latência para estabelecimento da conexão
 - Probabilidade de falha no estabelecimento da conexão
 - Vazão
 - Latência de trânsito
 - Taxa de erros residuais
 - Proteção
 - Prioridade

- Resiliência
- O IETF, através da RFC 1946, apresenta que “à medida que a demanda por serviços em tempo real via rede cresce, aumenta também a necessidade das redes compartilhadas proverem serviços de entrega determinísticos. Tais serviços de entrega determinísticos demandam que ambos – a aplicação e a infraestrutura de rede – tenham capacidade de requisitar, configurar e garantir a entrega dos dados. Coletivamente esses serviços são tratados como Reserva de Banda e Qualidade de Serviço (QoS)” [JAC 96].
- Por fim, o IEEE provê uma definição mais genérica para aplicações que demandam comunicação em tempo real: “O conjunto de características qualitativas e quantitativas de sistemas multimídia distribuídos, necessários para obter a funcionalidade requerida de uma aplicação” [VOG 95].

Um ponto importante ao se analisar QoS é apresentado pela Stardust no QoS Forum em: “QoS não cria banda” [STA 99]. Essa afirmação refere-se ao fato de que políticas e mecanismos de garantia de Qualidade de Serviço apenas gerenciam recursos (incluindo banda) de acordo com a demanda das aplicações e a configuração da rede. Logo, para se obter QoS com um nível de serviço garantido, é necessária a alocação de recursos para cada fluxo de dados. Os recursos (particularmente, banda) reservados para uma determinada aplicação deixam de estar disponíveis para aplicações que estejam operando em “best-effort”. Considerando que banda é um recurso finito, o projeto de uso de QoS deve assegurar que restem recursos para essas aplicações após a reserva, ou seja, as aplicações de alta prioridade (contempladas pelo mecanismo de QoS) não devem impedir que as aplicações normais, de baixa prioridade, continuem trabalhando, mesmo que a um desempenho inferior.

A implementação de QoS em qualquer rede implica em três serviços de suporte [STA 99]: **Gerenciamento de Políticas**, no sentido de mapear quais aplicações devem se valer de serviços na rede com QoS; **Autenticação**, para a certificação de quem irá efetivamente usar o recurso com QoS e **Contabilidade (*accounting*) e cobrança (*billing*)**, para realizar auditoria da oferta e do uso dos serviços em QoS e, eventualmente, tarifá-los de forma diferenciada.

5.2 QoS na Internet

Apesar da orientação inicial de atender a atividades de pesquisa e educação – mantida por mais de 20 anos – a Internet hoje é um meio de provimento de muitos serviços. A arquitetura na qual a Internet se baseia tem como fundamento o protocolo IP, cujo projeto original tinha por objetivo ser uma solução de comunicação genérica, baseada em datagramas e não orientada a conexão. Uma premissa levada em conta foi a capacidade de prover apenas serviços de nível *best-effort* (o melhor possível no momento) na camada de rede, relegando às camadas superiores as funcionalidades para obtenção de melhores níveis de serviço. O nível *best-effort* é adequado para aplicações tradicionais de dados, que podem tolerar atraso, perda de pacotes e variações de desempenho. Contudo, não é viável utilizar com este nível de serviço aplicações de tempo real, como telefonia IP, vídeo sob demanda e videoconferência [OPE 2000].

Com a transformação da Internet em infraestrutura comercial, a demanda por qualidade cresceu rapidamente [LI 2000]. Esse crescimento, por sua vez, foi acompanhado pelo debate sobre a validade e real necessidade de mecanismos de garantia de QoS – algumas opiniões chegaram a considerar que o crescimento da oferta e a diminuição de preço de banda (através de tecnologias como WDM) tornariam esse recurso tão abundante a ponto de prover QoS por si só. Contudo, esses fatores tendem a servir como incentivos para o desenvolvimento e uso de aplicações que consumam cada vez mais banda [STA 99][BER 2000][XIA 99].

Para prover os níveis de QoS atualmente – e, provavelmente, cada vez mais – necessários, o IETF apresentou uma série de modelos de serviço e mecanismos de garantia de QoS para redes IP – aplicáveis à Internet – conforme apresentado a seguir. É interessante observar que se tratam de serviços não exclusivos, que podem trabalhar de forma complementar [STA 99]. A Figura 5.1 apresenta a relação entre os modelos de QoS propostos pelo IETF e as camadas do modelo de referência OSI.

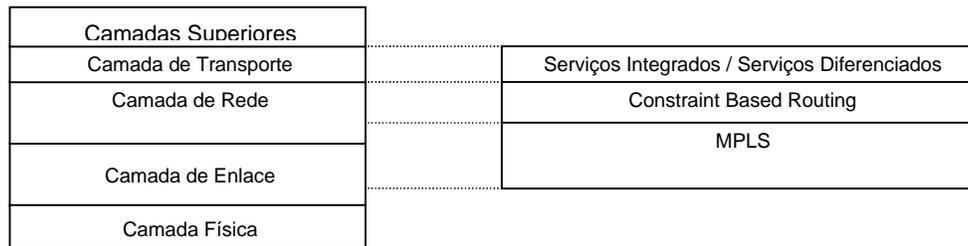


FIGURA 5.1 – Relação dos modelos de QoS com o modelo OSI. Fonte : [XIA 99]

Embora não seja objetivo final desse trabalho, a seguir é apresentada uma breve visão dos principais trabalhos em andamento visando implementação de QoS em redes IP. Serão apresentados a seguir alternativas para tal implementação, através de modelos de alocação de recursos (banda e prioridade) e modelos de otimização ligados ao roteamento.

Serviços Integrados – IntServ, ou Reserva de recursos : os recursos da rede são alocados de acordo com as necessidades de QoS de cada aplicação, e sujeitos às políticas de gerenciamento de banda. Este modelo é definido principalmente na RFC 1633, e especifica duas classes de serviço, além dos serviços do tipo *best effort* : os serviços Garantidos, para aplicações que requerem um limiar de atraso fixo (RFC 2212), e os serviços Preditivos, para aplicações que requerem um limiar de atraso probabilístico (RFC 2211). O protocolo RSVP (RFC 2205) se enquadra nesta categoria [LEI 99].

Serviços Diferenciados – DiffServ, ou Priorização : o tráfego da rede é classificado através da marcação do campo TOS (*Type of Service*) dos pacotes IP e, de acordo com o valor deste campo, são atribuídas a cada pacote determinadas prioridades de encaminhamento e/ou descarte. A arquitetura de Serviços Diferenciados é, essencialmente, um esquema de prioridade relativa [LEI 99]. A definição é apresentada na RFC 2475.

MPLS – Multi-Protocol Label Switching : trata-se de um esquema de encaminhamento de pacotes que atua entre as camadas 2 (enlace) e 3 (rede) do modelo de referência OSI. O *Label Switching* é uma funcionalidade que habilita roteadores a aplicarem rótulos aos pacotes, permitindo que aos demais roteadores da rede chavearem estes pacotes de acordo com os rótulos, minimizando as atividades de leitura/escrita na tabela de roteamento. De um modo geral, o MPLS é um mecanismo eficiente quando aplicado em redes IP, pois além de desempenhar classificação e encaminhamento rápido do tráfego, habilita mecanismos eficientes de tunelamento [LEI 99].

CBR – Constraint Based Routing : é uma ferramenta utilizada para automatizar o processo de Engenharia de tráfego. Trata-se de um incremento à operação dos protocolos que habilitam QoS, onde a forma de operação visa, através de mecanismos de política, computar rotas e distribuir carga de forma a poder atender aos requisitos de QoS. Para a computação das rotas, pode-se usar como métrica medidas como latência, jitter, custo, número de *hops* (saltos) e largura de banda.

Do ponto de vista do usuário que se conecta à Internet buscando acesso com Qualidade de Serviço, o acesso deve ser contratado mediante um acordo de nível de serviço (SLA). Este acordo, firmado por contrato entre o provedor de serviços (no caso, ISP, ou provedor de acesso à Internet) e o cliente (usuário) tem por objetivo especificar o nível e os parâmetros de Qualidade de Serviço que serão fornecidos ao usuário quando da conexão deste [LEI 99]. Para a disponibilização de serviços nestes moldes, é necessário a aplicação das técnicas acima (uma ou mais, conforme a estratégia do ISP) tanto na rede de acesso, ou enlace entre o usuário e o ISP (que deve ser capaz de suportar a banda e os demais parâmetros especificados no SLA) como na rede do provedor (no seu *backbone*). Se for o caso de uma aplicação de videoconferência com o ponto remoto, obviamente os dois pontos devem dispor de conexões similares. Este tipo de serviço, contudo, ainda não vem sendo oferecido no mercado.

5.3 QoS no Contexto de Videoconferência

No caso específico de sistemas multimídia, é possível considerar QoS sem relação com a rede – ou seja, mensurar Qualidade de Serviço para sistemas locais [STE 95]. Para mensurar a Qualidade de Serviço prestado por um dado sistema (no caso, comunicação multimídia), pode-se assumir QoS como a expressão de quão bons estão os serviços fornecidos pelo sistema. Para chegar a tal avaliação, é necessário que se disponha de padrões mensuráveis. Assim, a qualidade de serviço de sistemas multimídia é objetivamente descrita a partir de um conjunto de parâmetros bem definidos, que buscam caracterizar satisfatoriamente a noção subjetiva de qualidade de serviço fornecida pelo sistema em questão (obtida, por exemplo, por métodos de avaliação de opinião como o MOS, apresentado no item 5.5). A qualidade final do sistema pode ser estudada de três pontos de vista distintos [GUE 97][REY 2001]: o da infraestrutura, o da aplicação e o do usuário. Esta separação é conceitual, pois na prática tratam-se de fatores com forte relação de dependência entre si.

QoS do ponto de vista da Infraestrutura

Os parâmetros de QoS enquanto requisitos da infraestrutura provida pela rede (ou seja, responsabilidade dos níveis OSI 1 a 3) são bem definidos, e compreendem basicamente quatro fatores : latência fim-a-fim, jitter, taxa de pacotes perdidos e taxa de bits errados.

Latência fim-a-fim

É definida como o tempo transcorrido entre a captura de uma mídia em um ponto da conexão até a sua apresentação no outro ponto. O impacto causado pela latência do sistema como um todo é sentido na aplicação como prejuízo na interatividade (como algumas ligações por satélite, onde o atraso da comunicação por vezes confunde o usuário). Os valores aceitáveis, de acordo com a recomendação G.114, são (para um trecho, ou *one-way*):

- 0 a 150 milissegundos – aceitável para a maior parte das aplicações.
- 150 a 400 milissegundos – aceitável, com ressalvas ao administrador da rede, quanto a prováveis impactos na qualidade da transmissão das aplicações.
- mais de 400 milissegundos – geralmente inaceitável para planejamento de rede.

Em casos especiais, esses limites podem ser bastante diferentes. Aplicações de alta interatividade podem demandar latência inferior a 50 milissegundos [ITU 96a], enquanto outros casos (como ligação telefônica internacional por satélite, ou o uso de satélites como contingência para circuitos terrestres) podem tolerar atrasos entre 400 e 900 milissegundos [ITU 96a][PIC 2001].

A latência do sistema como um todo pode ter origem na rede ou no processamento das estações [ITU 96a]. No caso das estações, ainda é possível separar a latência que diz respeito a esta, de uma forma genérica, e a latência inerente aos codecs.

- Latência gerada na rede – A rede em si contribui para o aumento da latência das seguintes formas :
 - Atraso de propagação – Tempo necessário para a informação percorrer a distância do circuito. Este tempo é limitado pela velocidade da luz, não sendo portanto afetado pela tecnologia em uso.
 - Atraso de transmissão – Tempo necessário para um pacote percorrer um determinado meio. É determinado pela velocidade do meio em questão e pelo tamanho do pacote.
 - Atraso *Store-and-forward* – Tempo necessário para um dispositivo da rede – switch, roteador, bridge – enviar um pacote que tenha recebido.
 - Atraso de processamento – Tempo necessário pelo dispositivo da rede para tarefas como localizar a rota, substituir cabeçalhos e realizar outras eventuais manipulações na informação (por exemplo, incrementar o contador de *hop*).

- Latência gerada na estação – As tarefas de processamento nas estações (*endpoints*) contribuem para a latência total dos sistemas de videoconferência, pela realização das tarefas de interface com o usuário, gerenciamento da conexão de rede, da sessão de videoconferência, do acesso aos drivers de dispositivo, processos concorrentes na estação entre outros. O tempo consumido e o impacto dessas atividades dependem do equipamento que estiver sendo utilizado (dedicado para videoconferência ou não), das características de *hardware* (processador e memória), do tipo de periférico envolvido (por exemplo, câmeras conectadas via interface USB apresentam melhor performance do que as conectadas via porta paralela), do sistema operacional e de suas características e configurações, do ambiente de desenvolvimento da aplicação, do modo de acesso aos dispositivos (por exemplo, no Windows : DirectSound x WAVE, DirectDraw x Video for Windows) e da forma de desenvolvimento da aplicação propriamente dita.
- Latência gerada no codec – Os codecs de áudio demandam tempo para execução de suas tarefas. Este tempo deve ser levado em consideração no planejamento da aplicação, pois o uso de codecs com maior taxa de compressão – e consequentemente menor uso da rede – pode demandar maior tempo de processamento. O tempo total empregado nessas tarefas é composto de três atrasos distintos [COX 96] :
 - Atraso Algorítmico – A maioria dos codecs processa um quadro de áudio por vez; muitas vezes, é necessário analisar o quadro seguinte para decodificar corretamente o quadro atual – o que, em termos de implementação, implica em colocar o quadro atual em um buffer e receber o próximo (processo denominado *look ahead*). O tempo total é composto pela recepção, atribuição ao buffer, consulta ao próximo quadro e consulta ao quadro atual. Este atraso não pode ser reduzido pela variação da implementação, pois trata do tamanho do quadro (conforme apresentado na TABELA 5.1).
 - Atraso de processamento – Tempo necessário para o codificador analisar o trecho de áudio e o decodificador reconstruí-lo. Este tempo depende da velocidade do *hardware*. A soma do atraso de processamento e do atraso algorítmico é chamada *one-way codec delay*.
 - Atraso de comunicação – Trata-se do tempo necessário para um quadro ser inteiramente transmitido do codificador para o decodificador. A soma dos três tempos é chamada *one-way system delay*.

TABELA 5.1 – Tempo de compressão típico para Codecs de Áudio. Fonte: [CIS 2001][COX 96]

Método de compressão	Tempo de Compressão (ms)
G.711 PCM	0,75
G.723.1 ACELP	30
G.723.1 MP-MLQ	30
G.726 ADPCM	1
G.728 LD-CELP	3 a 5
G.729 CS-ACELP	10
G.729a CS-ACELP	10

Os codecs de vídeo também consomem um tempo considerável para as suas atividades, mas este tempo varia muito conforme a imagem em si, e no patamar de centenas de milissegundos. Por isso, é aceitável que os fabricantes insiram atrasos no fluxo de áudio, para manter sincronismo com o vídeo [ITU 96a][ITU 97].

Jitter fim-a-fim

Pode ser definido como “a distorção aleatória de sinal, ou um efeito subjetivo causado por erros baseado em tempo” [WIL 2000], ou como a variação no atraso da entrega de dois pacotes relacionados. O impacto da existência de jitter é sentido no nível de sincronização, tanto intra-mídia quanto inter-mídia. Os níveis aceitáveis para jitter se situam entre 5 e 10 milissegundos.

A técnica mais comum para minimizar o impacto do jitter é a utilização de um buffer para armazenar os dados recebidos da rede antes de disponibilizá-los à aplicação, referenciado como buffer de dejitter [PIC 2001]. Se o tamanho do buffer utilizado for superior às variações, o efeito do jitter pode passar despercebido. A PictureTel, por exemplo, adota implementações deste tipo de buffer a partir de um valor padrão de 150 milissegundos, configurável para valores inferiores em condições favoráveis de rede.

A adoção de buffer pode ser feita tanto no codec da estação que está recebendo os dados de áudio e vídeo como na rede em si. Contudo, essa técnica não possibilita garantia absoluta de que o tráfego será recebido livre de jitter. Se forem consideradas redes como a Internet ou o ambiente Ethernet, com disputa da rede por diversas aplicações, a previsibilidade do atraso para estimativa do buffer diminui, e se torna impossível garantir ausência de variabilidade. Além disso, o buffer de dejitter contribui para o aumento da latência fim-a-fim, adicionando à transmissão o tempo definido para o buffer.

Taxa de pacotes perdidos

Trata-se do percentual de pacotes descartados durante a transmissão. O impacto sentido na aplicação é a distorção, no caso de mídias sensíveis ao tempo (áudio e vídeo). Os níveis de tolerância a perda são apresentados com bastante variação pela bibliografia disponível, variando entre 0,01% e 3%, de acordo com os mecanismos de predição e recuperação implementados [VOG 95][PIC 2001].

Taxa de bits errados

Exprime o percentual de bits errados devido ao processo de transmissão. A exemplo do anterior, afeta a aplicação distorcendo os fluxos de dados. Os níveis aceitáveis se situam entre 0,01 e 0,1 % [VOG 95].

Os problemas decorrentes dos fatores enumerados acima são tratados e aceitos de forma distinta para cada mídia. Um resumo do impacto destes problemas em cada tipo de mídia presente em uma sessão de videoconferência é apresentado na Tabela 5.2.

TABELA 5.2 – Influência dos problemas de QoS da Infraestrutura nas Mídias de uma Sessão de Videoconferência. Fonte : [LUN 99][SOA 95]

Mídia	Tolerância a Erros e Descartes	Impacto do Jitter	Impacto da Latência
Áudio	Tolerável	Compromete a inteligibilidade	Compromete a interatividade
Vídeo	Tolerável	Compromete a inteligibilidade	Compromete a interatividade e o sincronismo (deve ser compensado)
Discreta	Não Tolerável (compromete a integridade do dado)	Indiferente	Indiferente

QoS do ponto de vista do Usuário

Do ponto de vista do usuário da aplicação, a Qualidade de Serviço de um sistema de videoconferência pode ser descrita por [GUE 97]:

Resolução

Caracteriza a precisão do processo de digitalização de um segmento de fluxo contínuo, e é descrito como uma função de bits por segmento e razão de amostragem.

Distorção

Mede o nível de perda de informação de um fluxo contínuo, e é função de erros de transmissão e perdas geradas, por exemplo, pelos algoritmos de compressão.

Nível de sincronização

Mede quão estável está a apresentação dos segmentos de um fluxo de mídia contínuo (sincronização intra-mídia) e entre mídias relacionadas (inter-mídia). Os aspectos de sincronização serão melhor apresentados no item 5.4.

Interatividade

É uma medida de quanto eficientemente um fluxo pode ser utilizado para trabalhos cooperativos em tempo real. Usualmente, a aplicação oferece um conjunto finito de opções para estes parâmetros, e interage com o usuário para determiná-las. É comum que essas opções sejam subjetivas, como “áudio de CD” ou “qualidade de vídeo pobre”.

QoS do ponto de vista da Aplicação

Trata-se do atendimento às características definidas para as mídias envolvidas na sessão de videoconferência. Entre essas características, pode-se destacar, conforme Reynard [REY 2001]:

- Altura da imagem;
- Largura da imagem;
- Quantidade de cores da imagem;
- Taxa de quadros por segundo;
- Frequência de amostragem do áudio;
- Número de canais de áudio.

5.4 Sincronização

Quando se trata de sistemas multimídia, a sincronização é um ponto chave, pois é necessário garantir ordenação temporal neste tipo de sistema [STE 95].

Um determinado fluxo de dados é composto de unidades lógicas de dados, ou LDU (*logical data unit*) [STE 96]. Assim, em um fluxo de áudio, uma LDU é uma amostra (ou bloco de amostras) de áudio; em um fluxo de vídeo, uma LDU pode tipicamente corresponder a um quadro, e várias LDUs consecutivas representam uma seqüência de quadros. A partir deste conceito, pode-se considerar a sincronização de duas formas :

- Sincronização Intrafluxo (*intrastream*) : a preocupação com o sincronismo aqui é de que as LDUs de um dado fluxo sejam apresentadas no receptor com a mesma relação temporal com que foram capturadas no emissor;
- Sincronização Interfluxo (*interstream*) : trata-se da ordenação temporal a ser aplicada entre dois fluxos de dados distintos e relacionados. Um caso típico é a apresentação de áudio e vídeo com “sincronização labial” (no caso particular de conversa entre pessoas), mas pode envolver diversos tipos de mídia como gráficos, imagens, animações, texto, áudio e vídeo. A diferença de tempo entre LDUs de áudio e vídeo correspondentes é referenciada como **skew**.

Para garantir entrega de dados com sincronismo, uma abordagem freqüentemente utilizada é a multiplexação das diversas mídias (como áudio, vídeo e outros), combinando-os fisicamente em uma unidade no emissor e realizando o processo inverso (demultiplexação) no receptor. Contudo, essa estratégia nem sempre é necessária ou viável [STE 96]. No caso específico de videoconferência, outra alternativa (apresentada pelo ITU-T e já referenciada) é a inserção de atraso entre LDUs da mídia de menor latência (áudio), para acompanhar o tempo (total – algorítmico, de processamento e de comunicação) da mídia de maior latência (vídeo). Esta estratégia é limitada no sentido de que, caso a latência inserida prejudique a interatividade, é preferível que se mantenha esta e o prejuízo seja sentido na sincronização – salvo requisito específico da aplicação [TAN 92].

Especificamente para as mídias de fluxo contínuo, é freqüentemente aceito que o áudio possa ser apresentado até 120 milissegundos antes do vídeo, e este até 240 milissegundos antes do áudio. Tais falhas de sincronismo podem até ser percebidas, mas nem sempre prejudicam a interatividade do sistema. Steinmetz apresenta uma série de experimentos com o objetivo de quantificar os requerimentos de qualidade de serviço necessários para cada sincronização de dados multimídia [STE 96]. O resultado de tais experimentos é apresentado na TABELA 5.3.

TABELA 5.3 – Parâmetros de QoS aceitáveis. Fonte: [STE 96]

Mídia		Modo/Aplicação	Limite
Vídeo	Animação	Correlata	+/- 120 ms
	Áudio	Sincronização labial	+/- 80 ms
	Imagem	Sobreposta	+/- 240 ms
		Não sobreposta	+/- 500 ms
	Texto	Sobreposto	+/- 240 ms
		Não sobreposto	+/- 500 ms
Áudio	Animação	Evento correlato (dança)	+/- 80 ms
	Áudio	Estreitamente acoplado (som estéreo)	+/- 11 μ s
		Fracamente acoplado (diálogo)	+/- 120 ms
		Fracamente acoplado (música de fundo)	+/- 500 ms
	Imagem	Estreitamente acoplado (música com notas)	+/- 5 ms
		Fracamente acoplado (slides)	+/- 500 ms
	Texto	Anotações textuais	+/- 240 ms
	Apontador	Áudio relativo ao objeto apontado	- 500 ms , +750 ms

5.5 Dimensionamento da Qualidade

A avaliação da qualidade de um serviço multimídia é necessariamente subjetiva, e depende da percepção do usuário. Diversos métodos para dimensionamento da qualidade percebida pelo usuário foram desenvolvidos. Um destes foi proposto pelo ITU-T, através da recomendação F.901.

Além destes, outros trabalhos apresentam tais métodos; a maioria tem em comum o emprego do teste ACR (*absolute category rating*), definido na recomendação

do ITU-T P.800 [ITU 96b] cujo resultado é conhecido como índice MOS (*Mean Opinion Score*, ou contagem média de opinião). Trata-se de um método de avaliação desenvolvido pelo próprio ITU-T para determinar a qualidade dos seus padrões de áudio (G.721 a G.729). O sistema consiste em submeter uma determinada quantidade de pessoas, em cabines próprias para tal, a ouvirem diversas amostras de áudio. Para cada amostra, cada indivíduo deve atribuir uma nota de 1 (pior) a 5 (melhor), conforme seu critério; a seguir, é feita uma média aritmética de todos os valores obtidos, cujo resultado corresponderá ao MOS.

Um teste deve incluir uma seleção balanceada de material – ou seja, um número igual de homens e mulheres, além de uma faixa de qualidade de amostras que possibilite aos participantes do teste usar todos os valores. Outro teste comumente usado, definido na mesma recomendação, é o DCR (*degradation category rating*), onde o participante é submetido a duas amostras – uma do áudio original e outra do áudio processado pelo codec – e identifica o nível de degradação sentido da primeira para a segunda amostra. O resultado desta medição chama-se DMOS (*degradation MOS*) [COX 96]. A TABELA 5.4 apresenta a escala de valores do MOS e DMOS.

TABELA 5.4 – Valores para MOS e DMOS. Fonte: [CIS 98]

Índice	Qualidade (MOS)	Nível de distorção (DMOS)
5	Excelente	Imperceptível
4	Bom	Perceptível mas confortável
3	Intermediário	Pouco desconfortável
2	Pobre	Desconfortável
1	Insatisfatório	Muito desconfortável

A TABELA 5.5 apresenta os níveis de avaliação pelo índice MOS dos codecs do ITU-T em condições normais.

TABELA 5.5 – Índices MOS típicos dos codecs do ITU-T. Fonte : [CIS 98]

Codec	MOS
G.711 – PCM	4,1
G.723.1 – ACELP	3,8
G.723.1 – MPMLQ	3,9
G.726 – ADPCM	3,85
G.728 – LD CELP	3,61
G.729 – CSA-CELP	3,92
G.729a – CSA-CELP	3,9

O índice de MOS apresentado por cada CODEC depende dos parâmetros de QoS da rede. A TABELA 5.6 apresenta a variação, também pelo índice MOS, do desempenho do codec CSA-CELP (G.729), em condições anormais de ambiente e de rede.

TABELA 5.6 – Variação do índice MOS do codec G.729 em diversas condições de rede. Fonte : [CIS 98]

Condições	MOS
Baixo nível de entrada	3,54
Dupla codificação	3,46
Tripla codificação	2,68
Taxa de erro de 5 % (bit error)	3,24
Taxa de erro de 5 % (frame error)	3,02

Mesmo se tratando de um método que utiliza a percepção individual, o MOS é uma medida bastante aceita. Ainda pode ser usada como complemento para medições objetivas, como a relação sinal/ruído.

6 Alternativas para Busca de Qualidade

Atualmente pode-se considerar o mercado bastante familiarizado com a tecnologia de videoconferência. Prova disto é o volume de negócios: o Wainhouse Research, empresa de consultoria especializada em videoconferência e comunicação multimídia, estima que no ano 2000 tenham sido comercializados mais de 150 mil unidades de equipamentos para videoconferência – sem levar em consideração câmeras do tipo WebCam [WAI 2001a]. Wilcox também aponta boas expectativas para este mercado, estimando um crescimento de 40% ao ano para chegar a movimentar 640 milhões de dólares em 2003. Uma demonstração do mercado concreta e favorável à tecnologia de videoconferência ficou evidente após os atentados terroristas de 11 de setembro de 2001 [GRI 2001][MAS 2001]. A partir desta data e da opção por evitar viagens de avião, a demanda por tais sistemas aumentou significativamente [MAS 2001], reforçando a tendência previamente existente de usar a videoconferência como substituto para o deslocamento, com seus custos e riscos, e como forma de compartilhamento de recursos (como pessoal especializado ou executivos) [WIL 2000]. Ao mesmo tempo, ainda existe resistência por parte de pessoas e/ou organizações quanto a essa tecnologia, na maior parte das vezes influenciadas por experiências mal sucedidas ou mesmo pela cultura de uso [MAS 2001]. Tais experiências podem ter origem no uso de sistemas de videoconferência em situações de ambiente que não o ideal (no que diz respeito a infraestrutura de telecomunicações ou de ambiente, por exemplo) ou no emprego de sistemas inadequados para a aplicação desejada. Fayers e Davis ilustram da mesma forma este caso: a adaptação de sistemas *desktop* para salas de reunião com 15 a 20 participantes – situação que, certamente, vai gerar má impressão a todos os que estiverem envolvidos [FAY 99][DAV 2000a].

Quando se pensa, portanto, em massificação de tecnologia de videoconferência – a ponto de conexões à Internet serem capazes de suportar sistemas funcionais – é importante analisar como operam os sistemas considerados eficientes e funcionais, bem como analisar as formas de implementar eficiência e funcionalidade em larga escala. A RFC 1946 menciona que a garantia de QoS para tais sistemas – e, conseqüentemente, a eficiência e funcionalidade esperadas – deva ser implementada tanto no nível da rede como da aplicação. As possibilidades de busca de condições para prover sistemas de videoconferência eficientes por meio de soluções orientadas a características da rede ou a componentes adicionais serão apresentadas brevemente, considerando o escopo deste trabalho ser direcionado a uma solução no nível da aplicação. Observe-se que o enfoque do trabalho é direcionado a redes com restrições clássicas de latência e banda, como WAN e Internet. O resultado destas análises, bem como as propostas de soluções, normalmente podem ser aplicados a redes locais (LANs), observando-se as particularidades destas.

6.1 Infraestrutura – ISDN (H.320) versus IP (H.323)

No que tange ao assunto infraestrutura, os sistemas de videoconferência serão separados por porte – sistemas de grupo e pessoais (*desktop*), dadas as diferenças de demanda de recursos e, conseqüentemente, de infraestrutura impostas por cada tipo. Embora a orientação do trabalho seja para sistemas pessoais, os sistemas de grupo serão

mencionados por constituírem um mercado forte, que merece ser considerado, e um padrão informal de referência para os sistemas pessoais.

Sistemas de Grupo

Atualmente, a grande maioria dos sistemas videoconferência – principalmente sistemas de grupo – está instalada sobre plataforma ISDN. O Wainhouse Research publicou, em julho de 2001, estatísticas apontando que essa é a plataforma usada em aproximadamente 80% dos equipamentos de aplicação comercial nos mercados dos Estados Unidos e Europa [WAI 2001]. Independente da presença de circuitos IP (entre, por exemplo, duas filiais de uma organização, ou entre a organização e outra – como cliente ou fornecedor) o mercado adotou como prática o uso de videoconferência via ISDN, baseado em três motivos de resultado imediato:

- O uso de conexões ISDN sob demanda dispensa a necessidade de contratação ou ampliação de circuitos permanentes – os enlaces ISDN são utilizados quando de interesse, com a largura de banda específica e de acordo com o destino desejado no momento.
- O fato de se utilizar uma rede ISDN à parte da rede corporativa, alocada especificamente para esse fim, pressupõe segregação e conseqüentemente exclusividade do tráfego de videoconferência, eliminando os atrasos correspondentes a tempo de enfileiramento e roteamento, por se tratar de conexão por comutação de circuito.
- Até aproximadamente 1996, a oferta de equipamentos de videoconferência era exclusivamente no padrão H.320. Desta época até hoje, apesar do crescimento da oferta de sistemas de grupo baseados no padrão H.323, os sistemas H.320 continuaram dominantes no mercado de videoconferência em grupo, tanto pela facilidade de aquisição como por se apresentar mais maduro e consolidado para o mercado [WIL 2000].

Não obstante às vantagens mencionadas acima, a mesma pesquisa indica que 25 % dos usuários de videoconferência estão testando ou pretendem migrar para o padrão H.323 (dentro deste percentual de usuários, a média de uso das soluções baseadas em IP chega a 50%). A oferta de soluções baseadas no padrão H.323 ou com capacidade para suportar ambos padrões (*dual mode*) vem crescendo, a ponto de, no ano 2000, a maioria dos equipamentos comercializados ser *dual mode* [WAI 2001]. Essa tendência é baseada em vantagens significativas de redes IP sobre conexões ISDN, que se sobrepõem aos motivos favoráveis mencionados, e são resumidos a seguir.

- Custo – é uma tendência atual do mercado diminuir o custo de manutenção de diversas redes paralelas, simplificando a infraestrutura mediante a convergência de todos os tipos de aplicação, com várias demandas de serviços de rede, na mesma infraestrutura (preferencialmente, IP) [OPE 2000]. A Cisco Systems apresenta uma ilustração prática desta tendência através do exemplo da empresa Large Oil Co., que voltou-se para videoconferência em IP após diagnosticar que o custo da hora de uma sessão de videoconferência em ISDN a 384 kbps oscilava entre 300 e 500 dólares [CIS 2001a].

- Capilaridade – a tecnologia ISDN dificilmente chega até as estações de trabalho dentro de uma organização (salvo conexão externa ou LAN-E, ambos com aceitação questionável), o que restringe o uso a ambientes com infraestrutura especificamente preparada. Isto pode ser viável para estações de grupo (ainda que seu isolamento da rede corporativa, ou a necessidade de duas interfaces – ISDN e LAN seja questionável), mas é crítico no contexto de videoconferência em *desktop*.
- Disponibilidade – acessos ISDN ainda não são tão corriqueiros quanto acessos via LPCDs convencionais, principalmente no Brasil. Além do fator disponibilidade, existem problemas com a continuidade da conexão (*uptime*), particularmente em ligações internacionais [CIS 2001a].

Sistemas Pessoais (*Desktop*)

Os dados apresentados acima referem-se quase que totalmente a sistemas de grupo, demonstrando que mesmo essas aplicações de caráter mais profissional (historicamente baseadas em ISDN) apresentam uma tendência para o padrão H.323. Se a análise for realizada em soluções de videoconferência para aplicações pessoais, baseadas em *desktop*, é possível identificar pontos mais favoráveis ainda às soluções baseadas em IP. Tal análise pode ser dividida em 2 casos:

- Organizações – existe uma tendência do mercado nesses casos ao uso de mecanismos de garantia de QoS e/ou convergência de serviços, utilizando a estrutura IP, antes direcionada apenas ao tráfego de dados, como meio para transmissão de voz e vídeo [LEI 99]. Além disso, é provável nesses casos o emprego de soluções de videoconferência espalhadas pela empresa e valendo-se da arquitetura de rede local (LAN, predominantemente IP no mercado), ou mesmo dos segmentos de WAN entre diversas instalações, dando suporte a filosofias como a de colaboração visual.
- Usuário doméstico – neste caso, o padrão H.323 é o mais indicado pelos fatores :
 - Custo da infraestrutura – a decisão pelo uso de videoconferência em um *desktop* já existente envolve o uso deste, da conexão à Internet, de um *software* muitas vezes gratuito (Cu-SeeMe 0.92, Microsoft NetMeeting) e restringe o investimento a uma câmera de baixo custo.
 - Disponibilidade de conexão – no que diz respeito a Internet, a maior parte dos acessos no Brasil ainda é realizada por linha discada, tornando o protocolo IP – e, conseqüentemente, o padrão H.323 – a grande base para utilização de sistemas por parte do usuário doméstico.

Um resumo das diferenças de características dos sistemas de grupo e *desktop* é apresentado na TABELA 6.1.

TABELA 6.1 – Comparativo entre Videoconferência em Grupo (H.320) e em Desktop / Internet (H.323). Fonte : [DIX 2000]

Tradicional	Internet
Padrão H.320	Padrão H.323
Uso em sala especial	Uso em <i>desktop</i>
Rara	Comum
Uso de linhas telefônicas ISDN	Infraestrutura de Internet
Alto custo de instalação	Baixo custo de instalação
Alto custo de uso	Sem custo de uso
Mercado já definido	Mercado em crescimento
Controle centralizado	Controle Descentralizado
Operação por profissional	Faça-você-mesmo

6.2 Soluções Orientadas à Rede

Uma alternativa sempre presente quando se fala na disponibilização de serviços de telecomunicações é a adequação da infraestrutura da rede. Conforme a Cisco Systems, o projeto de uma rede para suportar serviços de videoconferência deve prever basicamente [CIS 2001a]:

- Mecanismos de garantia de Qualidade de Serviço, para prover classificação e priorização do tráfego fim a fim;
- Planejamento do provisionamento de banda (*capacity planning*) – o enlace (fim a fim) entre os *endpoints* envolvidos deve ser capaz de fornecer banda suficiente para suportar o tráfego gerado.

Ao considerar redes locais (LAN), o provisionamento de banda tende a ser um problema de fácil solução, pois a banda mínima normalmente utilizada é a Ethernet (10 Mbps), e pode-se considerar atualmente como padrão de mercado Fast Ethernet (100 Mbps).

Em se considerando redes WAN privadas, o aspecto QoS é passível de atendimento pelo exposto no capítulo 4. No que diz respeito a dimensionamento de banda, trata-se de uma questão de avaliação de quem está projetando o uso do sistema (em se tratando de WAN privada, considera-se tipicamente alguma organização – empresa, universidade, hospital – interessada no sistema): é possível contratar circuitos da concessionária de telecomunicações conforme a necessidade e a disposição em investimento, desde uma LPCD até um enlace B-ISDN.

Por fim, ao considerar o atendimento a esses requisitos na Internet, é necessário levar em conta :

- A garantia de QoS no *backbone* (responsabilidade do provedor), bem como o tratamento adequado aos pacotes da aplicação conforme já mencionado. Para tal tratamento não seja usufruído indiscriminadamente por qualquer usuário, é necessário que o provedor disponha de mecanismos de autenticação (a menos que sua estrutura de *backbone* possua capacidade de prover QoS a todo o tráfego que

receba, inclusive pacotes de outra natureza de tráfego intencionalmente marcados como sendo de videoconferência). Também é desejável o estabelecimento de um acordo de nível de serviço entre o usuário e o provedor (SLA), para garantir que o serviço contratado está recebendo os recursos especificados no acordo (no caso do usuário) e disponibilizar a contabilização para cobrança diferenciada (no caso do provedor).

- O acesso (ou última milha) – Apesar da oferta crescente de acessos a Internet por banda larga (serviços N-ISDN, cable modem ou xDSL), este tipo de tecnologia ainda está longe de contar uma base instalada que seja capaz de proporcionar o uso de Internet com grande disponibilidade de banda em larga escala. Esta distância se deve a dois fatores: disponibilidade (nem todas localidades tem acesso a este tipo de tecnologia) e custo (nem todos os usuários se dispõem a pagar o custo de assinatura, cumulativo com o custo de assinatura do provedor de internet). Conforme pesquisa realizada pelo Ibope em novembro de 2001, a grande maioria dos usuários ainda acessa a internet por modems de 56 kbps ou menos, conforme apresentado na Figura 6.1 [FRE 2001].

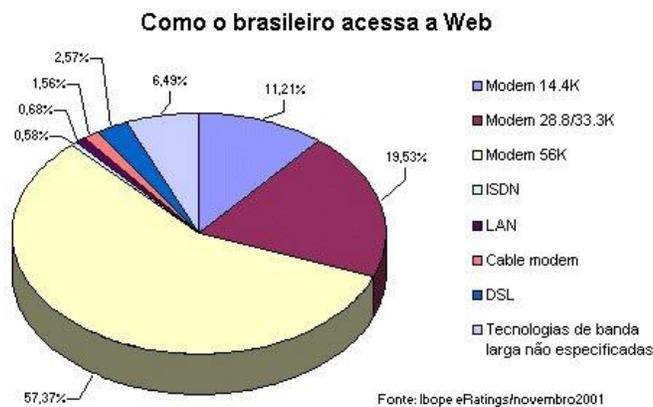


FIGURA 6.1 – Distribuição das formas de acesso à Internet no Brasil.

Fonte : [FRE 2001]

6.3 Soluções Orientadas à Aplicação

Uma forma de contornar os problemas inerentes a redes de recursos limitados, como um acesso à Internet, é relegar à aplicação a administração das perdas impostas pelos níveis inferiores. Neste grupo de soluções, também são identificadas duas estratégias principais: trabalhar com compensação (inferências ou predições) sobre o tráfego ou adaptá-lo, solicitando que o emissor envie apenas o que a rede tem condições de entregar.

Compensação

Este tipo de solução se vale da codificação das mídias para realizar inferências ou compensar perdas. A implementação deste tipo de tratamento se dá no próprio codec de áudio ou vídeo. As recomendações H.261 e H.263 especificam um recurso opcional,

denominado motion compensation. Uma das possibilidades desse mecanismo é a reconstrução no codec que está recebendo um fluxo de vídeo: a ausência de um quadro, ou mesmo a sua recepção incompleta, pode ser compensada pela inferência que o próprio codec faz sobre como deveria ser seu conteúdo, baseado em quadros anteriores (predição). Outra aplicação é a análise do fluxo que será enviado : caso o codec perceba que um dado quadro é idêntico ao anterior, seu envio pode ser suprimido, evitando assim consumo de banda. Maiores detalhes sobre estas técnicas podem ser encontrados em [WIL 2000] e [ITU 97].

Adaptação do tráfego

Os sistemas de comunicação multimídia (voz e videoconferência) possuem requisitos específicos no que diz respeito ao protocolo de transporte, em vista da natureza sensível ao tempo deste tipo de tráfego. Schulzrinne menciona que o TCP até pode ser apropriado para entrega de vídeo unidirecional, em redes com vazão suficiente e utilizando mecanismos de buffer grandes o suficiente para prover entrega próxima de tempo real. Contudo, para aplicações que demandem áudio e vídeo em tempo real, protocolos confiáveis como o TCP são inadequados pelas seguintes razões [SCH 2000]:

- Para que o emissor perceba que um pacote foi descartado e o retransmita, é transcorrido no mínimo um tempo de *round-trip* (ida e volta). Para que o sistema obtivesse algum proveito deste mecanismo, seria necessário que o receptor aguardasse pela retransmissão, aumentando o atraso fim-a-fim e adicionando lapsos na comunicação. Outra alternativa seria que a implementação do TCP descartasse o pacote e não esperasse pela retransmissão, incorrendo na necessidade de modificação do protocolo de transporte para uma dada aplicação.
- O TCP não suporta multicast.
- Os mecanismos de controle de congestionamento do TCP diminuem a janela de congestionamento mediante a detecção de pacotes perdidos. Os fluxos de áudio e vídeo, contudo, tem suas próprias taxas que não podem ser simplesmente diminuídas pelo protocolo de transporte – a mudança de tais taxas deve ser feita no nível da aplicação.

Por esses motivos, a recomendação H.323 define o tráfego de áudio e vídeo sobre UDP [PUR 2001]. Este protocolo provê envio de mensagens orientado a transação e não orientado a conexão, realizando pouco mais do que separar os dados em pacotes IP e enviá-los pela rede com um mínimo de controle (*checksum*) [POS 80].

Para adicionar um mínimo de confiabilidade a esse tipo de transmissão, é utilizado o protocolo RTP encapsulado no UDP. O RTP adiciona um cabeçalho de 8 bytes, que contém informações como número de seqüência, tipo de conteúdo e selo de tempo (*timestamp*) [SCH 2000]. O tamanho do cabeçalho RTP também representa uma vantagem secundária em relação ao TCP no que diz respeito ao *overhead* (8 bytes do primeiro contra 40 bytes do segundo).

No entanto, é importante salientar que o RTP não adiciona, por si só, mecanismos que assegurem qualidade de comunicação ao serviço provido pelo UDP.

Nenhum protocolo fim-a-fim – incluindo o RTP – pode assegurar entrega no tempo adequado. Isso deve ser provido pelas camadas inferiores, mediante controle dos dispositivos da rede. O RTP também não provê mecanismos para recuperação de pacotes perdidos, o que por definição é responsabilidade dos níveis superiores [SCH 2000]. O que o RTP provê é uma forma de avisar as aplicações (através das estatísticas providas pelo RTCP) que algo está errado, mas não é de sua responsabilidade corrigir o problema [HUT 97].

A partir da padronização pelo ITU-T do envio dos sinais de áudio e vídeo sem confirmação, abre-se a possibilidade do eventual descarte de tais sinais. Quando o número de pacotes enviados excede a capacidade de transmissão, estes pacotes são descartados aleatoriamente [MOH 2000].

A idéia da adaptação do volume de tráfego baseia-se no fato de que se for gerado um volume de fluxo de dados que a rede seja capaz de atender, a probabilidade de descarte diminui. Com a diminuição dos descartes, reduz-se a probabilidade de distorção do conteúdo, resultando em uma apresentação mais uniforme para o usuário. Além disso, essa técnica diminui também os tempos de enfileiramento nos equipamentos da rede, com influência direta na latência fim-a-fim e, por conseqüência, na interatividade do sistema.

6.4 Caracterização do Problema

Conforme previamente mencionado, é possível e razoavelmente comum a associação da tecnologia de videoconferência aos seus aspectos negativos, ocasionando até mesmo resistência ao uso desta tecnologia. De fato, devido aos requisitos diferenciados do tráfego multimídia, as implementações profissionais de sistemas de videoconferência pressupõem a garantia de qualidade de serviço por parte da infraestrutura de comunicação. Tais garantias são providas pela arquitetura da rede – ou pelo uso de ISDN, ou pela aplicação de serviços integrados ou diferenciados em redes IP, alternativas que sempre demandam investimento na infraestrutura de comunicação, ou o contrato de infraestrutura capaz de prover o nível de QoS desejado.

Neste capítulo, foram apresentadas algumas alternativas de tecnologias para suportar sistemas de videoconferência com qualidade adequada – proporcionando assim a satisfação do usuário e, conseqüentemente, a aprovação do sistema por parte deste. É claro que se pode utilizar uma combinação das técnicas citadas, obtendo assim o sistema em condições ideais para o melhor uso. Em caso de não se poder contar com as soluções disponíveis, o sistema deve estar preparado para, pelo menos, suportar uma delas.

A proposta de escopo do trabalho – popularização da tecnologia de videoconferência – torna natural a escolha da estratégia de adaptação do tráfego, pelos seguintes motivos:

- Portabilidade e Independência de infraestrutura – a adaptação do tráfego não varia em função dos níveis inferiores da rede.
- Baixo custo – a implementação de mecanismos de adaptação de tráfego não implicam na adição de *hardware* ou *software* adicionais.

- Otimização dos recursos – o comportamento de adaptação do fluxo de tráfego aos recursos disponíveis representa o melhor uso destes, sejam recursos de rede ou outros – como capacidade computacional no receptor, fora do escopo deste trabalho mas passível de atendimento por abordagem similar.

De uma forma geral, a abordagem de adaptação tem por objetivo minimizar os problemas causados por descarte, bem como a possibilidade de ocorrência de jitter quando este está relacionado ao congestionamento do circuito. Como a maioria dos acessos para o segmento pessoal é realizada via linha discada – com restrições claras de banda, agravada por variações da disponibilidade da mesma na estrutura (*backbone*) do provedor de acesso à Internet – o uso de adaptação de tráfego se torna uma alternativa bastante interessante.

Contudo, se forem consideradas as observações já apresentadas em relação à efetividade do sistema e sua influência na satisfação do usuário, é possível incorrer em problemas advindos da própria aplicação da estratégia de adaptação. Como esta implica em que o usuário abra mão de uma parcela do tráfego, esta escolha pode não corresponder exatamente à expectativa do usuário. Assim, o que seria um método para proporcionar a melhor qualidade possível no momento pode restringir a vazão justamente do que mais interessa, pela forma arbitrária com que a adaptação é realizada.

Como exemplo, se o usuário estiver usando o sistema de videoconferência aqui proposto para monitorar a presença em um dado ambiente, o momento inicial do sistema deve visar atender a todas as mídias da melhor forma. Em caso de restrição de recursos, o sistema deve avaliar o que é prioritário caso a caso – neste, a imagem é mais importante para o usuário do que o áudio. Se esta monitoração visa reconhecer a fisionomia de alguém, o tamanho e a resolução da imagem são mais importantes do que a taxa de quadros por segundo. Em uma aplicação clássica de videotelefonia, é preferível manter áudio fluente ao invés de uma imagem grande ou de alta definição. Se esta conversa por videotelefone envolver mais de uma pessoa em cada estação, o tamanho da imagem se torna mais importante do que a taxa de quadros por segundo, e pode ser prioritário em relação à qualidade do áudio – pode ser aceitável receber o som com uma avaliação na escala MOS inferior, mas preservando as características da imagem.

Considerando-se importante para a disseminação de uma dada tecnologia a plena satisfação do usuário a partir do uso da mesma, também é importante fornecer ao usuário uma maneira de definir como ela irá atender a sua necessidade – o que aproxima o que o sistema fornece do que o usuário realmente quer.

O modelo que será apresentado a seguir visa aproximar a estratégia de adaptação de tráfego ao comportamento que um dado usuário considera ótimo para uma dada aplicação, através da interação do mesmo com o sistema para definir a prioridade entre as mídias de acordo com seu interesse específico. Note-se que a forma de interação não é necessariamente voltada a ser manipulada pelo usuário final. Um sistema desenvolvido com esta abordagem deve prover alguns perfis de aplicação previamente definidos, para uma escolha fácil por parte do usuário. Estes modelos são de escopo da implementação, e não da definição do modelo. Além destes, um segundo nível de interação deve disponibilizar uma intervenção mais apurada do usuário, se for de

interesse deste, para realizar a definição conforme sua necessidade e, provavelmente, sua experiência.

7 O Modelo Proposto

No intuito de prover sistemas de videoconferência com a melhor efetividade possível, e vislumbrando o problema de banda passante – ainda bastante significativo, apesar da crescente oferta, tanto por disponibilidade quanto por custo – é apresentado um modelo de adaptação de tráfego direcionado às expectativas do usuário. O modelo é baseado na identificação do perfil da aplicação, ou o conjunto de características e parâmetros que mais atende ao usuário em uma dada utilização.

7.1 Escopo e Contribuição

O trabalho visa apresentar um mecanismo baseado em adaptação para a adequação do tráfego da forma mais agradável ao usuário. Para tanto, é definido o conceito de perfil de aplicação, os componentes necessários para sua utilização e a estrutura do modelo em si. O tipo de sistema produzido a partir do modelo proposto neste trabalho visa atender ao usuário doméstico e a pequenas organizações que desejem acesso a videoconferência via Internet ou Intranet com recursos de banda limitados e com uma finalidade definida.

O escopo do modelo, ou o tipo de sistema a ser construído a partir dele, é direcionado para o mercado de videoconferência pessoal, em redes baseadas em protocolo IP sem garantia de serviço, e com *hardware* – rede e estação – de baixo investimento. A estratégia proposta para atender a este mercado consiste em disponibilizar recursos para a otimização do modelo de adaptação de tráfego proposto por Zanin, de forma a prover o melhor atendimento às necessidades particulares de cada aplicação.

7.2 Perfis

Um ambiente próximo do que o mercado tem informalmente como consenso em infraestrutura para videoconferência profissional (*business quality*) envolve enlaces de banda larga, com mecanismos de garantia de QoS e reserva de banda efetivos (exemplo clássico : ATM) e dimensionamento adequado de latência, jitter e banda [ISO 98][ZAN 99]. Nesse contexto, a discussão sobre a existência de características do sistema mais importantes que outras é praticamente desnecessária, sendo pouco provável que exista interesse em priorizar uma dada característica do sistema em detrimento de outra. Nessa suposta situação, o sistema como um todo é capaz de prover integralmente os recursos necessários ao usuário para o estabelecimento da sessão de acordo com sua expectativa, necessidade ou exigência.

Contudo, esse não é o tipo de situação esperada para impulsionar o uso de videoconferência a patamares de popularização em que se encontra hoje, por exemplo, a tecnologia de telefonia. A situação do mercado, conforme já mencionado, indica que apesar da oferta de meios de acesso de banda larga, a grande maioria dos acessos a conectividade (especificamente, Internet) é realizada por modems de 56 kbps ou menos [FRE 2001]. Mesmo em ambiente de empresas, o recurso banda enquanto infraestrutura

para comunicação por videoconferência ainda possui restrições de disponibilidade e/ou custo [BER 2000].

Para esses ambientes, propõe-se outro tipo de atuação por parte do sistema para atender a demandas específicas : ao invés deste buscar mais recurso na rede para prover mais recurso de interatividade, escolhe-se que recurso (mídia) é relegada para que outra possa atender à demanda. Esse também é o procedimento proposto em caso de congestionamento ou outro motivo que implique em perda de pacotes.

Os sistemas de videoconferência possuem uma vasta gama de aplicações. Alguns casos são apresentados a seguir, ilustrando a diversidade das aplicações e das expectativas dos usuários e, conseqüentemente, dos requisitos :

- Comunicação pessoal (videotelefone)
- Reuniões de negócio
- Apresentações remotas
- Suporte a trabalho cooperativo (colaboração visual)
- Telemedicina

Reynard apresenta em [REY 98] a classificação de algumas aplicações de videoconferência conforme o nível de QoS necessário para atender a um dado interesse do usuário. Em termos de QoS da aplicação, o autor restringiu-se exclusivamente ao parâmetro Quadros por Segundo. Alguns níveis de serviços definidos pelo autor são apresentados na TABELA 7.1.

TABELA 7.1 – Níveis de serviço definidos por aplicação. Fonte : [REY 98]

Aplicação	Objetivo (imagem)	QoS
Vigilância	Ambientes e pessoas	1 quadro em 5 minutos
Monitoração	Ambientes e pessoas	1 quadro por segundo
Comunicação	Pessoas	20 quadros por segundo

Assim como a fluência do vídeo pode ser condição de sucesso para uma certa aplicação, a qualidade e o tamanho da imagem também podem determinar se uma ferramenta é satisfatória ou não. Características como baixa qualidade de vídeo, ou janelas de vídeo muito reduzidas, podem resultar na perda de detalhes importantes em termos de expressões faciais ou outras respostas não verbais que indiquem, por exemplo, o real significado de uma determinada palavra ou entonação de voz. Por outro lado, expressões faciais e demais podem ser perdidas quando, mesmo a uma boa resolução e dimensionamento de imagem, as pessoas se encontram a uma distância tal da câmera que seja suficiente para anular ou impossibilitar o aproveitamento dos sinais associados à imagem dos participantes. Tal distância pode ser imposta, por exemplo, pela disposição da sala ou pela quantidade de participantes (neste caso, quando for necessário o uso de zoom da câmera, para captar a imagem de vários – ou todos –

participantes). Estas dificuldades de uso das características de comunicação não-lingüística começam a aparecer tipicamente a partir de 4 metros [PAT 2001].

O que este trabalho propõe é o uso do conceito de classificação das aplicações de videoconferência direcionada aos requisitos gerados pelo usuário, de forma similar à apresentada por Reynard, a uma arquitetura de adaptação de tráfego. O objetivo é fazer com que as escolhas sobre a adaptação das mídias a serem realizadas pelo sistema de adaptação não sejam feitas de forma arbitrária nem aleatória, pois elas influenciam diretamente o nível de avaliação do usuário, e delas depende o sucesso relativo do sistema (conforme definido por Frater [FRA 2001]). Tais escolhas devem ser feitas de acordo com o critério que mapeia em valores – discretos ou não, conforme a escolha por parte do responsável pela implementação – dos parâmetros configuráveis às características que atendem determinada expectativa do usuário previamente selecionada.

Perfil – Definição

Para a definição do conceito de perfil, é necessário especificar previamente os seguintes conceitos:

Expectativa do usuário – o que um dado usuário espera de um dado sistema de videoconferência para uma dada aplicação, de forma a tornar sua experiência com esta aplicação o mais confortável e eficiente possível. Como exemplo, um usuário de um sistema de videotelefonia baseado no padrão H.323 tem expectativas diferentes de um usuário de um sistema de telemedicina e de um usuário de sistema de monitoração remota, ambos também baseados no padrão H.323. Rettinger menciona a importância desta questão, ao afirmar que “usuários se sentem frustrados se o sistema não suporta o modo de interação que eles esperavam” [RET 95].

Disponibilidade do ambiente – capacidade da rede atender à demanda de tráfego, ou as limitações de recursos impostas pela infraestrutura ao sistema. Restringindo para a técnica proposta – adaptação de tráfego – a disponibilidade a ser considerada é a largura de banda disponível para o tráfego de videoconferência, mas também pode incluir a capacidade de processamento das estações (que pode influir negativamente na recepção) e outros fatores similares.

Resolução do sistema – dimensionamento dos parâmetros variáveis da aplicação, como qualidade de áudio (canais, frequência de amostragem, bits por amostra), qualidade de vídeo (cores, tamanho da janela, quadros por segundo) ou a presença ou não de outros fluxos, como a transferência de arquivos, imagens estáticas, compartilhamento de aplicativos ou quadro-branco. Este último grupo será tratado de forma genérica como “mídias discretas”, conforme conceito do capítulo 2.

A partir destes conceitos, define-se perfil de uma sessão de videoconferência, como sendo o mapeamento da expectativa do usuário para resolução – ou conjunto de resoluções – do sistema, em função da disponibilidade do ambiente.

Para a ilustração da classificação das aplicações de um mesmo sistema de acordo o perfil, foram analisados alguns estudos de caso publicados em [WIL 2000] e pelo

Wainhouse Research, de acordo com três modelos de aplicação escolhidas (videotelefone, vigilância e colaboração visual). Com base nas análises desses estudos, são apresentadas a seguir algumas propostas de exemplos da descrição de perfis, passíveis de definição no modelo proposto.

- Perfil 1 – Videotelefone

É o caso clássico de comunicação interpessoal, onde o áudio é prioritário em relação ao vídeo e a expectativa de qualidade do usuário tem como parâmetro uma chamada de telefone. A princípio, pode ser fornecida uma resolução de áudio igual ou superior ao serviço de telefonia (PCM, com 8 mil amostras de 8 bits por segundo), e o restante da banda fica disponível para vídeo e mídias discretas. Em caso de necessidade de adequar o tráfego a uma oferta menor de banda, pode-se diminuir a qualidade do vídeo. Para isso, pode-se utilizar a seguinte escala de prioridades : inicialmente, quantidade de cores do vídeo, a seguir o tamanho da janela e posteriormente a quantidade de quadros por segundo – esta seqüência compromete a interatividade o mínimo possível. Caso o sistema esteja operando no mínimo possível que mantenha esta característica e seja necessária nova adaptação, o áudio pode ser alterado para uma qualidade inferior à do telefone, mas esta alternativa certamente resultará em percepção desta diferença de qualidade.

- Perfil 2 – Monitoração Remota

Apesar da existência de soluções dedicadas especificamente para este fim, a tecnologia de videoconferência em IP se mostra uma boa alternativa para este tipo de aplicação, principalmente para monitoração de ambientes a grandes distâncias. Nestes casos, é importante que o tamanho da imagem proporcione resolução suficiente para que seja possível identificar o evento monitorado. Pela própria natureza das tarefas de monitoração, também é fundamental que se garanta a atualização dos quadros – mesmo que a baixas taxas (qps). A quantidade de cores, embora com menos prioridade, também é relevante para a monitoração. Por último, o áudio pode ser considerado, neste caso, apenas ferramenta de apoio (pela definição do serviço de videoconferência, este não pode ser desconsiderado).

- Perfil 3 – Colaboração Visual

Um recurso presente na definição da recomendação H.323 é o que se costuma denominar *data conferencing*. Trata-se do suporte a trabalho cooperativo, valendo-se de recursos como compartilhamento de aplicativos, troca de mensagens (*chat*), suporte a transferências de arquivos e espaço para esboços gráficos compartilhados (*shared whiteboard*). Assim, um perfil de aplicação voltado a esse mercado deve dispor dos recursos necessários ao suporte a esse tipo de aplicação. Considerando que a atenção neste caso é voltado às mídias “discretas” em relação ao tempo, estas podem ser tratadas como o recurso de maior prioridade. Em um segundo nível de prioridade, pode se situar o áudio – tanto pela definição das aplicações de videoconferência como pelo seu valor, enquanto meio de suporte a interação. Por último (menor nível de prioridade), pode-se colocar a qualidade do vídeo, recurso que provavelmente terá menos atenção quando os participantes estiverem trabalhando com recursos como, por exemplo, quadro-branco ou compartilhamento de aplicativos.

Nível – Definição

Considerando que o mapeamento do interesse em resolução é dependente da disponibilidade, um determinado perfil deve prover a definição de tal comportamento em diversas situações de disponibilidade – ou seja, a definição de como o sistema irá realizar diversas instâncias de mapeamento. Este comportamento inclui a seqüência de resoluções, bem como quais os critérios para passagem para uma resolução superior ou inferior.

O mapeamento das qualidades à resolução da aplicação deve ser realizado de forma progressiva, em um conjunto definido de resoluções possíveis – iniciando da resolução ideal para o usuário na aplicação específica e diminuindo esta conforme a infra-estrutura de rede for impondo limites.

Para que o mapeamento seja realizado desta forma, é proposta a figura de nível – definido como sendo uma determinada resolução do sistema e as condições para passagem para o próximo nível (superior ou inferior). Assim, um perfil é composto de uma seqüência de níveis.

O primeiro nível é o considerado ideal, ou aquele em que a aplicação seja capaz de prover o maior grau de satisfação ao usuário possível para uma determinada largura de banda. O nível seguinte mapeia uma “concessão” do usuário, ou seja, este abre mão da qualidade de uma dada característica do sistema em favor da manutenção da qualidade das demais, bem como os próximos níveis, progressivamente.

Limite de nível para adaptação

A passagem de um nível para outro deve ser realizada mediante um evento específico, relacionado à disponibilidade do ambiente. Define-se limite de nível como sendo o valor que expresse de forma quantitativa qual o volume de restrições de disponibilidade de rede que deve ser levado em consideração para acionar uma mudança de nível (no caso de adaptação, esse valor tende a ser caracteristicamente descartes, podendo também ser expresso pelos valores de latência e/ou jitter). Tal valor deve ser uma grandeza objetiva e expressa preferencialmente por meio de valores discretos, de forma a poder ser comparada com as estatísticas que mensuram os eventos cuja monitoração é responsável pela adaptação do tráfego.

Do ponto de vista de implementação, o Perfil para uma aplicação pode ser representado como uma matriz, onde as linhas correspondem a níveis e as colunas correspondem a resolução esperada para um dado nível, e os critérios de limite para troca do mesmo.

7.3 Definição do Modelo

Direcionando a idéia de Adaptação de Tráfego apresentada no item 6.2 para o escopo do trabalho – Videoconferência para o usuário final (normalmente doméstico ou pequena empresa) em *desktop*, onde o recurso de rede (principalmente acesso) é limitado por questões de investimento ou disponibilidade – a adaptação de tráfego se

apresenta como uma boa alternativa para melhoria da qualidade oferecida, [CHA 94][ZAN 2000], baseado na proposta de suportar apenas serviços do tipo best-effort no nível de rede e deixando para os computadores (*endpoint*) a funcionalidade de obter os níveis de serviço adequados à necessidade [OPE 2000].

Considerando que a avaliação dos sistemas é subjetiva, ela certamente é influenciada pelo sentimento de satisfação das expectativas particulares do usuário para uma dada aplicação – do contrário, a avaliação será influenciada por sentimento de frustração por parte do usuário. Se a aplicação for capaz de atender a esta expectativa, a avaliação tende a fornecer melhores resultados – por proporcionar maior grau de satisfação ao usuário em relação ao que ele deseja [BRO 99].

O modelo ora proposto visa estender o modelo de adaptação de tráfego de Zanin, adicionando a este a capacidade de, mediante a escolha do perfil por parte do usuário, realizar a adaptação do tráfego da maneira mais adequada para o perfil escolhido. Para atingir esta maneira mais adequada e, conseqüentemente, uma maior satisfação do usuário em relação a como o sistema lhe atende, é definida para cada mídia sua prioridade no momento da adaptação do tráfego, bem como os parâmetros que devem ser considerados para que aconteça a adaptação – para baixo (diminuição da resolução por congestionamento) ou para cima (aumento da resolução por disponibilidade). Os componentes propostos para esta funcionalidade são:

- Gerenciador de Perfil de Aplicação – seu objetivo é fornecer ao usuário, de forma fácil e intuitiva, a possibilidade de definir – em alto nível – o modelo de adaptação. Características referentes a prioridade entre mídias, supressão de mídia e opção por descarte devem fazer parte de um perfil, bem como informações sobre taxas mínimas para cada mídia. Ao selecionar um perfil, o usuário automaticamente seleciona o modelo de adaptação que mais lhe é favorável para a aplicação. Além das regras predeterminadas, o usuário também pode opcionalmente valer-se do gerenciador para estabelecer parâmetros que refletem sua necessidade específica. Esses parâmetros seriam mantidos em uma base de dados, lida no momento da carga da aplicação. Também deve ser de responsabilidade do Gerenciador de Perfis o cálculo e a informação sobre a banda necessária para cada nível.
- Gerenciador de Prioridade – seu objetivo é fornecer ao agente de videoconferência, responsável pela adaptação, informações adicionais às fornecidas pelo RTP e pelo RTCP, no sentido de orientar a forma adequada de realizar a adaptação de acordo com o perfil definido pelo usuário. Esta orientação deve se dar, por exemplo, especificando qual a mídia que deve ser adaptada antes, ou qual mídia deve sofrer descarte se a adaptação não for possível ou suficiente. Trata-se da tarefa de fazer com que o Agente de Videoconferência realize a adaptação de acordo com o perfil escolhido. A implementação se dá a partir da recepção das informações referentes à qualidade da sessão enviadas pela estação remota. Tais informações são comparadas com o critério local (o perfil propriamente dito), o que gera decisões sobre a adaptação que devem ser encaminhadas para o agente de Videoconferência.

Esses componentes interagem com os dois componentes definidos por Zanin [ZAN 2000], Agente de Observação da Rede e Agente de Videoconferência. De acordo com o modelo original, estes componentes são responsáveis, respectivamente, pela aquisição dos dados sobre as condições de tráfego da rede e pela configuração da

resolução dos dispositivos de Entrada e Saída. A interação entre os componentes da proposta deste trabalho, bem como a sua interação com os componentes definidos por Zanin, é apresentada na Figura 7.1.

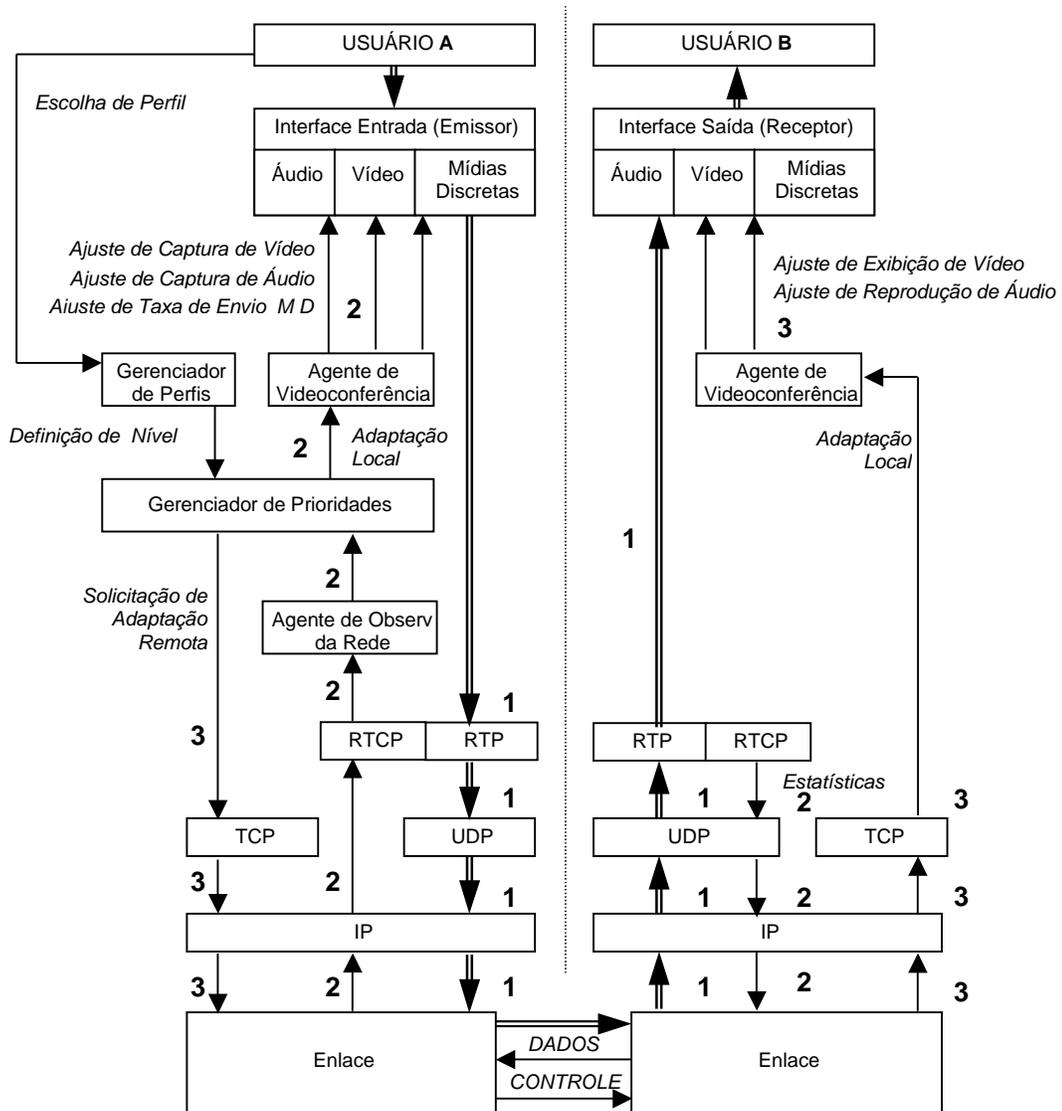


FIGURA 7.1 – Interação entre os componentes do Modelo

O modelo define o seguinte funcionamento para o sistema, para cada estação:

Partindo de uma situação inicial – a carga do sistema, o sistema gera os fluxos das mídias de áudio e vídeo a partir da estação emissora, conforme definido no nível mais alto. Para este tipo de tráfego, é utilizado o UDP, conforme já apresentado no Capítulo 6. Esta fase está representada na Figura 7.1 com o número 1.

A partir do momento em que está recebendo dados da sessão de videoconferência, a estação receptora tem condições de avaliar quanto do tráfego originalmente gerado pelo emissor está sendo recebido e aproveitado. A forma proposta

para a realização desta tarefa é a partir da utilização dos protocolos RTP e RTCP (o primeiro insere informações de tempo – *time stamp* – e seqüência, enquanto que o segundo é responsável pelo retorno da informação à estação geradora). Opcionalmente, a própria aplicação pode realizar auditoria do volume de tráfego que é aproveitado em relação ao volume de tráfego que é enviado. Isto é especialmente útil no caso de a estação não conseguir realizar a exibição do fluxo inteiro recebido (por exemplo, por pouca disponibilidade de memória ou capacidade de processamento). Esta fase está representada na Figura 7.1 com o número 2.

Tal retorno é encaminhado de volta ao emissor do fluxo das mídias da aplicação. Ao receber estes dados, cabe à estação emissora comparar os níveis de perda (na rede, na aplicação ou na própria estação), de latência ou de jitter com os critérios definidos para a aplicação (ou seja, o nível específico do perfil selecionado). Quando o Agente de Observação da Rede identifica que os problemas reportados atingiram o estabelecido pelo Limite de Nível, este informa ao Gerenciador de Prioridades a necessidade de adaptação do tráfego. Este, por sua vez, consulta as informações disponibilizadas pelo Gerenciador de Perfis, para identificar quais os parâmetros (resolução) a serem adotados para o próximo nível.

De posse de tais parâmetros, o Gerenciador de Prioridades realiza duas solicitações de adaptação. A primeira é local, para realizar o ajuste do envio dos dados à disponibilidade do sistema como um todo. Esta alteração implica em configurações de taxa de amostra, de qualidade e ajustes na resolução do(s) dispositivo(s) de captura. A segunda solicitação é remota, e segue para a outra estação como um aviso de que ela deve alterar os seus dispositivos de exibição para a nova qualidade ou resolução. Como este dado é sensível a perda e não pode ser distorcido, é utilizado o TCP para o seu envio. Esta fase está representada na Figura 7.1 com o número 3.

Agindo desta forma, o modelo permite ao sistema trabalhar de forma assíncrona. Suponham-se duas estações em uma sessão de videoconferência implementada conforme este modelo. A estação A está conectada à Internet, por exemplo, via ADSL com taxas de 48 kbps para *upload* e 128 kbps para *download*. A estação B está conectada à Internet, por exemplo, via um circuito de 64 kbps *full duplex*. Nessas condições, é muito provável (desconsiderando o enlace intermediário) que a estação B gere mais solicitações de adaptação que a estação A, encontrando assim seu ponto estável a um nível de recepção inferior ao da estação A. Esta, por sua vez, irá transmitir na sua limitação, e receber na limitação da estação B. Este exemplo se encontra resumido na TABELA 7.2.

TABELA 7.2 – Modos de Adaptação – Funcionamento Assimétrico (Exemplo)

	Modo	Estação A	Estação B
CONEXÃO	Download	128 kbps	64 kbps
	Upload	48 kbps	64 kbps
SESSÃO	Recepção	64 kbps	48 kbps
	Envio	48 kbps	64 kbps

8 Implementação – Validação do Modelo

Para validar o modelo proposto, foi realizada a implementação de um protótipo de sistema de videoconferência baseado no modelo de adaptação de tráfego orientado a perfis. A implementação teve como objetivo observar o comportamento do sistema quando funcionando a partir do referido modelo.

As características desejáveis para tal validação, segundo a análise realizada, seriam as que nos apresentassem o funcionamento de um sistema de videoconferência em *desktop* com um comportamento previsível, mesmo em um ambiente instável. Para isso, foram implementados os seguintes módulos:

- O Gerenciador de Perfis, de forma a poder realizar a definição objetiva de alguns perfis para teste do modelo;
- O Sistema de Videoconferência em si, de forma que fosse possível aplicar os critérios objetivos definidos no perfil da aplicação e observar as alterações no fluxo das diversas mídias que o compõem (vídeo, áudio e mídias discretas).

Este capítulo apresenta uma descrição da implementação realizada, incluindo a tecnologia utilizada, as funcionalidades implementadas e a forma de implementação de tais funcionalidades.

8.1 Tecnologia e Ambiente

O protótipo foi desenvolvido em ambiente Microsoft Windows, utilizando a linguagem C++. A ferramenta de programação utilizada foi o C++Builder da Borland, na versão 5.5.

Para realizar o acesso aos dispositivos de vídeo e áudio e à rede, foram utilizadas as seguintes ferramentas de programação:

- Acesso à entrada e saída de vídeo

A interface com o dispositivo de entrada de vídeo (câmera) e com a saída (monitor) foi realizada mediante acesso direto à API Video for Windows (VFW), sem o uso de componentes intermediários.

- Acesso à entrada e saída de áudio

Para o acesso aos dispositivos de áudio (microfone e caixas de som), foi utilizada a API WAVE, em conjunto com o gerenciador de compressão padrão da mesma (ACM, ou *Audio Compression Manager*). O acesso a estas bibliotecas, por sua vez, foi realizado através de um componente de domínio público, que foi alterado para adaptar suas funcionalidades às necessidades do protótipo.

- Acesso à rede

O acesso à comunicação sobre o protocolo IP (via UDP e TCP) foi realizado através dos componentes padrão do C++Builder (TNMUDP e TSocket, respectivamente). Por questões de implementação, não foi utilizado o conjunto de protocolos RTP/RTCP; contudo, como será apresentado a seguir, as funções de sequenciamento e ordenação foram implementadas na própria aplicação.

8.2 Implementação

Conforme já mencionado, a implementação se deu em duas frentes: no Gerenciador de Perfis, que provê suporte à definição dos perfis e no protótipo propriamente dito.

Implementação do Gerenciador de Perfis

A função do Gerenciador de Perfis, como já mencionado, é prover ao usuário a forma de definir o comportamento da aplicação. O gerenciador deve, portanto, prover ao usuário a interface que apresente as resoluções de mídia disponíveis, bem como os critérios que serão utilizados para realizar a adaptação do tráfego. O resultado desta definição deve ser armazenado de alguma forma, para ser lido pelo sistema de videoconferência. Na implementação, o perfil foi gravado em um arquivo com o formato .INI, similar a um arquivo de configuração do Windows.

Para cada nível dos perfis a serem testados, foram disponibilizadas algumas resoluções para cada atributo de cada mídia. Um dado nível é, portanto, uma combinação destas resoluções. Em termos da implementação, foi criada uma classe – Definição_Perfil – que encapsula a matriz de características por níveis, além dos controles (como o nome do próprio perfil e o nível atual) e os métodos para gravação, recuperação de informações e navegação na estrutura (como subir e descer de nível). Abaixo é apresentada a estrutura de dados das classes Definição_Nivel_Perfil (que define cada nível de um perfil) e Definição_Perfil (que define o perfil propriamente dito, e encapsula uma matriz de níveis).

```

struct Definicao_Nivel_Perfil
{
    bool   _pfVideoHabil;    // vídeo habilitado
    int    _pfQPS;          // 1 a 20
    int    _pfDimVideo;     // 1:160x120, 2:176x144, 3:320x240, 4:352x288
    int    _pfCorVideo;     // bits para cor - 16 ou 24
    bool   _pfAudioHabil;   // áudio habilitado
    int    _pfTagAudio;     // TAG (identificador do codec) 1 = PCM
    int    _pfBitAudio;     // bits por amostra - 8 ou 16
    int    _pfCanAudio;     // canais por amostra - 1(mono) ou 2(stereo)
    int    _pfFreAudio;     // frequencia - 8, 11, 22 ou 44 khz
    bool   _pfDiscretaHabil; // mídia discreta habilitada
    int    _pfMidiaDis;     // 0 = desabilitada, 1 a 9 = kbps
    int    _pfDescLimInf;   // limite de % de descarte para descer nível
    int    _pfDescLimSup;   // limite de tempo em seg com (descarte == 0)
                                // para subir nível
};

class Definicao_Perfil
{
public:

    // Parâmetros gerais e de controle

```

```

AnsiString _pfNomePerfil;
AnsiString _pfDescPerfil;
int _pfNivelAtual;
int _pfMaxNiveis;

// Funções de recuperação de informação

bool fPegaVideoHabil();
int fPegaQPS();
int fPegaDimVideo();
int fPegaAltVideo();
int fPegaLarVideo();
int fPegaCorVideo();
bool fPegaAudioHabil();
int fPegaTagAudio();
int fPegaBitAudio();
int fPegaCanAudio();
int fPegaFreAudio();
bool fPegaDiscretaHabil();
int fPegaMidiaDis();
int fPegaLimInf();
int fPegaLimSup();

// Funções de atribuição de informação

void fSetaVideoHabil(bool);
void fSetaQPS(int);
void fSetaDimVideo(int);
void fSetaCorVideo(int);
void fSetaAudioHabil(bool);
void fSetaTagAudio(int);
void fSetaBitAudio(int);
void fSetaCanAudio(int);
void fSetaFreAudio(int);
void fSetaDiscretaHabil(bool);
void fSetaMidiaDis(int);
void fSetaLimInf(int);
void fSetaLimSup(int);

// Funções de interface para arquivos

void fNovoPerfil(AnsiString, AnsiString);
void fSalvaPerfil(AnsiString, Definicao_Perfil);

void fDesceNivel();
void fSobeNivel();
void fSetaNivel(int);

long int fBandaVideo();
long int fBandaAudio();
long int fBandaTotal();
int fTestaEAdapta(int);

Definicao_Nivel_Perfil Nivel[10];
};

```

Para que o produto final pudesse ser configurado conforme a necessidade do usuário (não necessariamente por este), o Gerenciador de Perfis deve ser capaz de armazenar informação independentemente do programa executável. A forma encontrada para tal armazenamento foi utilizar arquivos de inicialização do Windows (.INI). Abaixo é apresentado um trecho de um arquivo de definição de perfil gerado pelo Gerenciador, e que pode ser utilizado em uma sessão de Videoconferência no Protótipo.

```

[Geral]
NomeAbr=video simples
Descricao=Perfil de teste de vídeo elementar

```

```

[Nivel0]
vid_hab=1
vid_qps=5
vid_dim=2
vid_cor=24
aud_hab=1
aud_tag=1
aud_bit=8
aud_can=1
aud_fre=22050
dis_hab=1
dis_ban=12
lim_inf=5
lim_sup=0

[Nivel1]
vid_hab=1
vid_qps=5
vid_dim=2
vid_cor=16
aud_hab=1
aud_tag=1
aud_bit=8
aud_can=1
aud_fre=11025
dis_hab=1
dis_ban=6
lim_inf=10
lim_sup=0

```

A interface com o usuário permite que sejam definidos os atributos específicos de cada mídia, informando a banda total que será consumida para um determinado nível, conforme apresentado da Figura 8.1 até Figura 8.4. Caso necessário, o usuário pode obter uma visão global do Perfil, conforme a Figura 8.5.

FIGURA 8.1 – Definição de Atributos por Mídia – Áudio

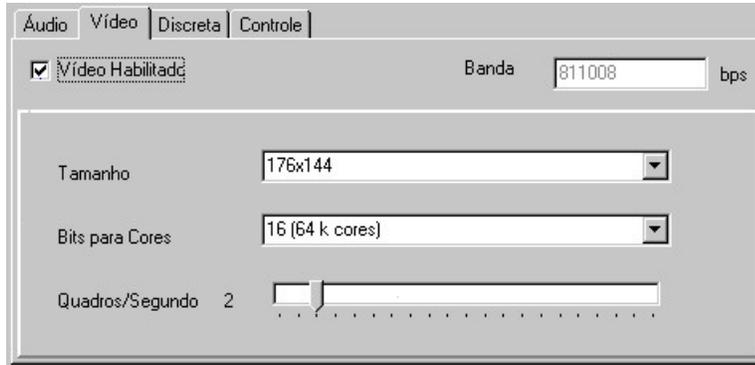


FIGURA 8.2 – Definição de Atributos por Mídia – Vídeo

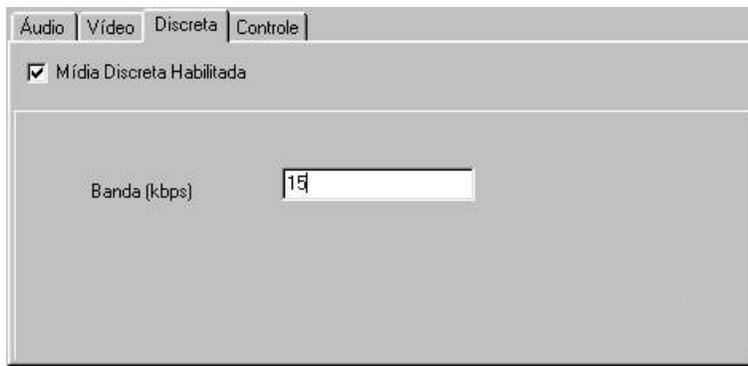


FIGURA 8.3 – Definição de Atributos por Mídia – Discreta

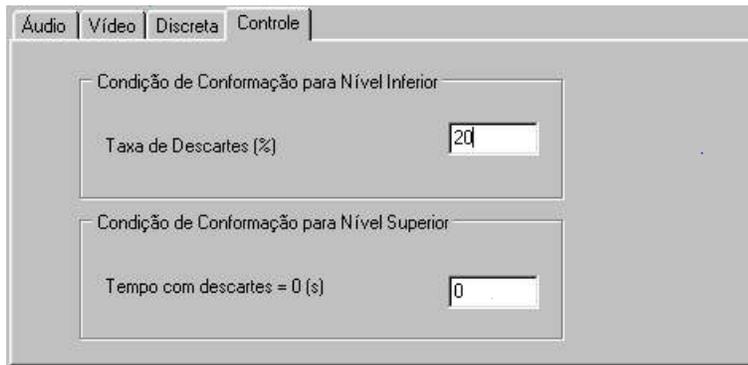


FIGURA 8.4 – Definição de Limites para Adaptação

		Nível 0	Nível 1	Nível 2	Nível 3	Nível 4	Nível 5	Nível 6	Nível 7	Nível 8	Nível 9
Nome do Perfil:		Videotelefone									
Descrição		Comunicação Pessoal com Vídeo									
BANDA NÍVEL		574600	390280	143291	113800	983650	676450	369250	358225	355200	355200
Áudio	CODEC	3	3	3	3	3	3	3	3	3	3
	Amostras/seg										
	Bits/Amostra										
	Qtd Canais										
BANDA ÁUDIO		176400	176400	176400	176400	22050	22050	22050	11025	8000	8000
Vídeo	Tam Janela	240x320	240x320	144x176	120x160						
	Bits Cor/Cores	24 / 16	16 / 64k								
	Quadros/seg	3	3	3	3	3	2	1	1	1	1
BANDA VÍDEO		552960	368640	121651	921600	921600	614400	307200	307200	307200	307200
Discreta	BANDA MD	5	5	5	5	5	5	5	5	5	5
Lim % descarte p/baixar		20	20	20	20	20	20	20	20	20	20
Limite seg para subir		0	0	0	0	0	0	0	0	0	0

FIGURA 8.5 – Tabela de Visualização do Perfil

Implementação do Protótipo

O protótipo foi desenvolvido visando prover o tráfego de áudio, vídeo e mídias discretas. A implementação foi realizada de forma que as configurações referentes à resolução destas mídias fossem feitas de acordo com as definições e critérios presentes no arquivo de definição de Perfil, conforme o modelo definido no Capítulo 7.

- Interface

O protótipo foi implementado a partir de uma interface MDI, com múltiplas janelas. Cada janela representa a implementação do objeto a que diz respeito. É possível ao usuário visualizar simultaneamente as janelas de vídeo local e remoto, a janela de informação sobre o tráfego de áudio, as janelas de mídia discreta (transferência de arquivo e chat) e as informações de estatística da sessão.

A Figura 8.6 apresenta os menus da tela principal do protótipo.



FIGURA 8.6 – Menu Arquivo e Janela

- Negociação de Chamada

O protótipo implementa comunicação multimídia ponto-a-ponto, trabalhando de forma orientada a sessão. Ao executar o protótipo, o usuário o coloca automaticamente em modo de espera (podendo receber chamadas). Para realizar chamadas, o usuário seleciona a opção correspondente, seleciona o perfil que vai definir o comportamento da sua estação e o endereço IP da estação (host) remota, através da tela de chamada apresentada na Figura 8.7. A partir da confirmação destes dados, o sistema realiza a troca de mensagens para estabelecimento da sessão (chamada), por intermédio de portas específicas. As portas utilizadas nas trocas de mensagens para negociação da chamada, dos parâmetros e no envio e recepção das mídias são listadas no Anexo 1.

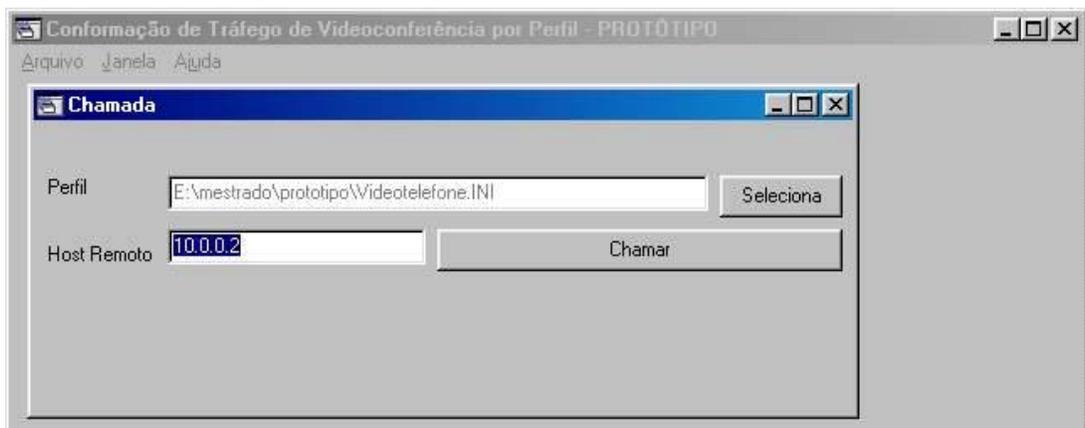


FIGURA 8.7 – Tela de Chamada

- Captura, Tráfego e Apresentação das Mídias

O vídeo é capturado, por intermédio da chamada CALLBACK da Video for Windows, em unidades referidas pela API como Frames (quadros). Dado o fato de o protótipo não implementar compressão de vídeo, o frame possui o tamanho originalmente gerado pelo dispositivo de captura. Suponha-se uma imagem padrão QCIF, com 176 pixels de largura e 144 pixels de altura, e a uma resolução de 24 bits (16 milhões de cores). Este quadro é fornecido pela VFW no seguinte tamanho:

$$176 * 144 * (24 / 8) = 76.032 \text{ bytes}$$

Este frame não pode ser transmitido de uma única vez, por dois motivos. O primeiro é que a implementação dos sockets UDP do C++Builder restringe o tamanho do datagrama a ser enviado em 2048 bytes (2 kbytes). O segundo diz respeito à segurança da transmissão do próprio frame: considere-se o fato de que o protocolo de transporte não ser confiável implica em que não haja garantia de chegada. Supondo que o ambiente de desenvolvimento disponibilizasse o envio de um datagrama deste tamanho, um erro na rede causaria a perda do frame inteiro.

Por essas razões, os quadros de vídeo são divididos em unidades, denominadas segmentos. Os segmentos são encapsulados em uma estrutura que compreende, além do dado propriamente dito (1 byte), o número do quadro ao qual pertencem (2 bytes) e o seu número dentro do quadro (2 bytes). Com este dado é possível reordenar os segmentos na chegada, preenchendo um buffer do tamanho do quadro previamente alocado. Considerando-se a característica não confiável do protocolo responsável por este transporte (o UDP), a recepção não é orientada à quantidade de segmentos que deveria ser recebida, e sim ao quadro ao qual pertence o segmento. Assim, quando o receptor percebe a chegada do primeiro segmento do quadro seguinte, executa a exibição do quadro anterior – mesmo que este esteja incompleto, pois não é possível saber se os segmentos foram perdidos ou desordenados, de forma a chegar depois – e inicia o preenchimento do novo quadro. O formato do pacote de vídeo é apresentado no Anexo 1.

O tráfego de vídeo é enviado através da porta UDP 2000. Devido a particularidades do ambiente de desenvolvimento, não foi possível realizar o envio direto a uma porta única, pois o receptor mostrou-se limitado em relação à quantidade de pacotes recebidos (dentro de uma mesma chamada do CALLBACK na captura do vídeo no emissor); assim, no lado do receptor, foi necessário disponibilizar mais de uma porta (faixa de portas UDP de 2001 até 2018), que no momento da recepção encaminham o buffer recebido à função de remontagem. No lado do emissor, é realizado um rodízio para enviar cada segmento a uma destas portas.

A captura e reprodução do áudio foram realizadas mediante a API ACM, que adiciona serviços de interface com os codecs de áudio do sistema operacional à API WAVE. Esta, por sua vez, trabalha de forma similar à API VFW : quando um buffer de áudio é preenchido, a API WAVE emite uma mensagem para o sistema operacional; na ocorrência desta mensagem, a API ACM captura os dados disponibilizados pela WAVE, executa a compressão/codificação selecionada e disponibiliza os mesmos para a aplicação, para que possa ser feita a segmentação e envio pela rede.

Para a implementação do protótipo, foram selecionados alguns dos formatos suportados pela API ACM. Os procedimentos de segmentação e remontagem utilizados para o envio e recepção do fluxo de áudio são idênticos aos descritos para o vídeo. O formato do pacote de áudio também é apresentado no Anexo 1, bem como as portas utilizadas e os formatos suportados na implementação.

A Figura 8.8 apresenta a interface do sistema, com o tráfego real das mídias de vídeo (janela local e remota) e o tráfego de áudio.



FIGURA 8.8 – Telas de Vídeo (Local e Remoto) e Áudio

Foi disponibilizado ao protótipo um módulo de Chat (bate-papo via texto), sem restrições específicas ou relações com o nível ou perfil selecionado em vista do pouco consumo de banda gerado por este módulo. A janela corresponde ao módulo de Chat é apresentada na Figura 8.9.

O tráfego de mídias discretas (como compartilhamento de aplicativo, quadro-branco) foi simulado no protótipo mediante uma interface própria. Através desta, o usuário pode enviar tráfego simulado, na forma de rajadas, informando para isso as características deste tráfego: a quantidade de rajadas, tamanho de cada uma e o tempo entre elas (que pode ser especificado ou aleatório). Como estas mídias são por definição mais sensíveis a erros e descartes e menos sensíveis ao tempo em comparação às demais, é utilizado o protocolo TCP para seu transporte. A interface para a simulação de mídias discretas é apresentada na Figura 8.10.

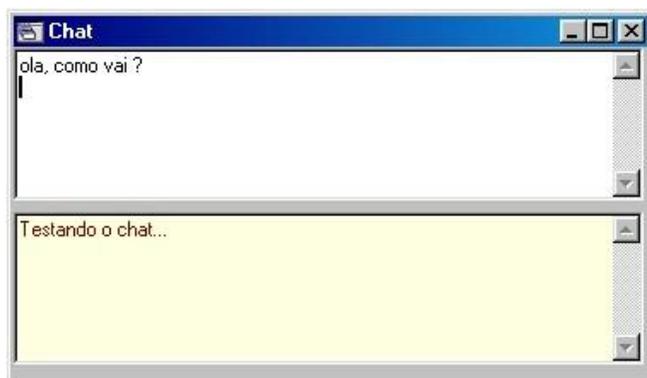


FIGURA 8.9 – Tela de Chat



FIGURA 8.10 – Tela de Mídias Discretas (simulador)

- Auditoria e Adaptação do Tráfego

No estabelecimento da sessão, cada estação configura apenas os seus parâmetros locais, conforme definido no perfil. Esta configuração (bem como todas alterações que acontecerem) é apresentada na tela de controle, conforme a Figura 8.11. Após a conclusão destas configurações (de captura de áudio e vídeo), cada estação envia para outra seus parâmetros locais (que os utilizará como parâmetros de exibição). Desta forma, ambas estações possuem controle do nível que está sendo utilizado inicialmente na outra estação.



FIGURA 8.11 – Controle da Sessão – Informações de Perfil e Nível

A manutenção de uma variável adicional que reflete a configuração do nível remoto atual disponibiliza a informação sobre os tamanhos dos quadros que deviam ser recebidos – tanto para áudio quanto para vídeo. Como mencionado anteriormente, um quadro é dividido em diversos segmentos. A cada segmento recebido, é incrementado um contador, que representa quantos bytes referentes ao quadro atual já foram recebidos. Quando o quadro muda (ou seja, quando a estação recebe um segmento que

pertence ao próximo quadro), o valor do contador é armazenado em outra variável (bytes recebidos para o último quadro) e o contador é zerado.

Em um intervalo de tempo definido (a implementação considerou 2 segundos), o sistema calcula a taxa de descartes, considerando quantos bytes deviam ter sido recebidos pelo quadro e quantos efetivamente o foram. Essa comparação se dá por amostragem, visto que no intervalo de tempo implementado são recebidos diversos quadros. A taxa de descartes é calculada separadamente para cada mídia; após isso, é calculada uma taxa de descartes geral (referente a áudio e vídeo), que representa a média ponderada dos descartes, conforme a expressão abaixo.

$$\text{Descarte Geral} = ((DV * QV) + (DA * QA)) / (QV + QA)$$

Nesta expressão, as variáveis iniciadas por D representam as taxas apuradas de descarte (DV : Descarte de Vídeo, DA : Descarte de Áudio) e as variáveis iniciadas por Q representam os tamanhos de quadros (QV : Quadro de Vídeo, QA : Quadro de Áudio).

Uma vez calculado este percentual, o sistema o envia para a outra estação. Na estação que recebe o valor que representa o percentual de descartes, o sistema realiza a comparação entre o percentual de descarte reportado pela outra estação e o percentual definido no nível atual do seu perfil como limite para realizar a adaptação. Caso essa comparação indique que o descarte reportado foi maior que o limite, a estação diminui um nível no seu perfil e envia a sua nova configuração local (de captura) para a estação remota (que utilizará como configuração de exibição). O limite para o envio do tráfego referente às mídias discretas também é limitado pela definição do nível. Observe-se que este tráfego não é contabilizado na auditoria da taxa de descartes, uma vez que é utilizado um serviço confiável (TCP) para seu transporte.

O fluxo de mensagens que efetivamente implementa este processo é apresentado no Anexo 1, bem como os pacotes e portas envolvidos. A tela que apresenta os percentuais de descartes apurados, bem como os volumes totais recebidos por mídia, é apresentada na Figura 8.12. Esta tela possui ainda um grupo chamado “Adaptação”, onde é possível especificar o tipo de reação do sistema ao receber a indicação de descartes da estação remota. As opções são ignorar a indicação (“Nenhuma”), habilitar uma barra de seleção para simular o descarte remoto (“Simulada”) ou realizar a adaptação com base na estatística recebida (“Real”). As duas primeiras opções destinam-se a depuração e testes, e a terceira destina-se ao funcionamento normal do sistema.



FIGURA 8.12 – Controle da Sessão – Auditoria e Descartes

Comparando-se estas funções com o modelo conceitual apresentado no Capítulo 7, percebe-se a presença e atuação de todos os componentes definidos no modelo: o Agente de Videoconferência (na configuração dos parâmetros de captura e exibição), o Agente de Observação da Rede (na recepção das informações enviadas pela estação remota) e o Gerenciador de Prioridades (que realiza a adaptação de acordo com a definição do perfil). As funções que o modelo propõe serem executadas pelos protocolos RTP/RTCP são executadas no protótipo pela auditoria de descartes, o que proporciona o subsídio necessário para o funcionamento do protótipo de acordo com o modelo.

9 Testes e Resultados

A partir da conclusão da implementação do protótipo, o mesmo foi submetido a alguns testes, em ambiente e situações controlados. Este capítulo apresenta os testes realizados e as conclusões obtidas a partir dos mesmos.

O principal objetivo dos testes realizados foi o de avaliar a funcionalidade do modelo, utilizando o protótipo de sistema desenvolvido em todas as funções propostas no modelo. Para isso, foram definidos perfis, utilizando o Gerenciador de Perfis; a seguir, os perfis definidos foram utilizados para o estabelecimento de sessões, acompanhando (ou provocando) a ocorrência de situações que gerassem o critério escolhido para realizar adaptação do tráfego – no caso da implementação realizada, ocorrência de um determinado percentual de descartes.

9.1 Cenário e Metodologia

Conforme apresentado na seção 8.2, a implementação se valeu de utilização de vídeo sem compressão. O envio desta mídia, neste formato, gera um volume de tráfego considerável: considerando 3 quadros por segundo e a resolução QCIF (176x144) e 64 mil cores (16 bits), a banda necessária seria de aproximadamente 152 kilobytes (1,21 Megabit) por segundo.

Tendo em vista tais volumes de tráfego da sessão (principalmente o tráfego de vídeo), os testes foram realizados em dois equipamentos conectados via Rede Local (Fast Ethernet), a 100 Mbps. A conexão entre as estações foi feita via cabo UTP com pinagem *cross-over* (ou seja, sem o uso de hub ou *switch*). Abaixo são apresentadas as configurações de *hardware* e sistema operacional dos dois equipamentos utilizados.

Estação 1

Processador	: Pentium III 700 MHz
Memória RAM	: 320 Mb
Sistema Operacional	: Microsoft Windows 98
Câmera	: Creative Labs WebCam (USB)
Dispositivo de áudio	: Genius SM32X

Estação 2 :

Processador	: AMD K62 533 MHz
Memória RAM	: 64 Mb
Sistema Operacional	: Microsoft Windows 98
Câmera	: Cubic Videocomm CVideo-Mail Express (Paralela)
Dispositivo de áudio	: CMedia C3DX

Em relação ao método dos testes, foi considerada a existência de duas fases de testes distintas. A primeira delas é a definição dos perfis, e a segunda é a utilização efetiva dos mesmos no protótipo.

Em relação à fase de definição, o método utilizado para testar a implementação foi a validação das funcionalidades do Gerenciador de Perfis. Considere-se que a fase de

definição é efetiva se, pelo uso exclusivo do Gerenciador de Perfis, for possível definir perfis de aplicações conforme apresentado no Capítulo 7.

Para a fase de execução, a validação se deu pelo estabelecimento de sessões de videoconferência utilizando o protótipo e os perfis definidos pelo Gerenciador. A avaliação desejada, neste caso, é a manutenção dos níveis de qualidade de acordo com a estratégia de adaptação – ou seja, adequando a resolução para minimizar a ocorrência de problemas causados por congestionamento (no caso do protótipo implementado, tais problemas referem-se especificamente a descartes).

A avaliação da taxa de descartes no decorrer dos testes poderia ser feita de duas formas. A primeira delas seria a contagem dos pacotes que chegam à estação receptora no nível da rede, utilizando uma ferramenta para capturá-los (como um “sniffer”) e outra ferramenta para identificar os quadros, separá-los e contabilizar quantidade de segmentos para cada quadro de cada mídia. A outra forma seria utilizando o próprio módulo de Controle do protótipo, e as estatísticas referentes a descartes geradas pelo mesmo (cujo mecanismo de cálculo é apresentado no Capítulo 8). O primeiro método pode informar com precisão o volume que não trafegou na rede, e fornece informações precisas para todos os quadros enviados. O segundo é menos preciso e trabalha por amostragem; contudo, as estatísticas calculadas pelo protótipo disponibilizam a apuração de eventuais perdas causadas por overhead de processamento (CPU). Como este também é um aspecto que contribui para a qualidade do sistema, optou-se pelo segundo método – baseado na estatística fornecida pelo módulo de Controle.

9.2 Execução dos Testes

Teste 1 – Definição dos Perfis

Este teste teve como objetivo específico a elaboração do arquivo de configuração de perfis com a ferramenta que se destinava a esta tarefa (o Gerenciador de Perfis). Foram escolhidos dois perfis para serem definidos – Videotelefone e Vigilância, ambos apresentados no Capítulo 7. Observe-se que as definições realizadas não foram validadas com uma massa estatisticamente significativa de usuários. É interessante salientar, contudo, que não se trata do escopo do teste definir perfis comprovadamente adequados ao mercado, e sim suportar a definição dos mesmos através de exemplos hipotéticos.

Perfil 1 – Videotelefone

A definição do perfil de videotelefone foi feita priorizando a qualidade de áudio. O objetivo aqui foi suportar uma forma de comunicação verbal (como o telefone comum), com a adição do recurso de imagem em tempo real, na medida da disponibilidade do ambiente. Assim, os primeiros níveis de adaptação reduzem as cores da imagem, a seguir as dimensões da janela de vídeo, e depois a fluência do vídeo (quadros por segundo). O som é mantido com qualidade de 22 kHz até o sexto nível, utilizando codec PCM e 8 bits por amostra.

TABELA 9.1 – Perfil de Videotelefone

Nível	Vídeo			Áudio	
	Tamanho	Cores	QPS	CODEC	Resolução
0	320 x 240	16 M	6	PCM	22 / 8 / M
1	320 x 240	64 k	6	PCM	22 / 8 / M
2	176 x 144	64 k	5	PCM	22 / 8 / M
3	176 x 144	64 k	4	PCM	22 / 8 / M
4	176 x 144	64 k	3	PCM	22 / 8 / M
5	176 x 144	64 k	2	PCM	22 / 8 / M
6	176 x 144	64 k	2	PCM	11 / 8 / M
7	160 x 120	64 k	2	PCM	11 / 8 / M
8	160 x 120	64 k	1	PCM	11 / 8 / M
9	160 x 120	64 k	1	PCM	8 / 8 / M

Perfil 2 – Monitoração / Vigilância

A definição do perfil de vigilância, por sua vez, priorizou a qualidade da imagem em detrimento da qualidade do áudio. O objetivo deste perfil é proporcionar a capacidade de observação visual – por exemplo, de um dado ambiente – complementada se possível por áudio. Assim, o perfil foi definido utilizando o áudio no formato fornecido pelo codec GSM, com resolução de apenas 2 bits por amostra, e tentou-se manter a qualidade do vídeo (especificamente, a dimensão da janela) em níveis que suportassem boa identificação da imagem.

TABELA 9.2 – Perfil de Vigilância / Monitoração Remota

Nível	Vídeo			Áudio	
	Tamanho	Cores	QPS	CODEC	Resolução
0	320 x 240	16 M	6	GSM	22 / 2 / M
1	320 x 240	16 M	5	GSM	22 / 2 / M
2	320 x 240	16 M	4	GSM	22 / 2 / M
3	320 x 240	64 k	4	GSM	22 / 2 / M
4	320 x 240	64 k	4	GSM	11 / 2 / M
5	320 x 240	64 k	3	GSM	11 / 2 / M
6	176 x 144	64 k	2	GSM	11 / 2 / M
7	176 x 144	64 k	1	GSM	11 / 2 / M
8	176 x 144	64 k	1	GSM	8 / 2 / M
9	160 x 120	64 k	1	GSM	8 / 2 / M

Teste 2 – Execução do Sistema

Inicialmente, foram testadas as execuções do sistema com cada um dos perfis definidos anteriormente.

Observou-se que nos primeiros instantes de execução do sistema, o mesmo iniciava o processo de adaptação, o qual se repetia até chegar ao nível mais baixo. Para avaliar a origem deste comportamento, o perfil de Videotelefone foi alterado, deixando o primeiro nível idêntico ao último, mas o fato se repetiu.

A seguir, foi alterada a opção “Adaptação” do sistema, de “Real” para “Nenhuma”. Nesta situação, foi possível perceber uma alta taxa de descartes no início da sessão, que aproximadamente cinco segundos depois diminuiu até próximo a zero.

Desta forma, foi possível concluir que o sistema passa por um período de instabilidade logo após o estabelecimento da sessão, bem como após a adaptação do tráfego. Através de monitoração do consumo de CPU, pôde-se observar um aumento instantâneo do consumo desta, coincidente com o ajuste das janelas de exibição e captura de vídeo e/ou o ajuste da configuração de captura e reprodução de áudio. Além deste período (durante o qual o processo de adaptação deve ficar suspenso), foi considerada a existência de eventuais picos de descarte acima da média.

A ocorrência desses instantes de descarte acima da média demonstrou, pelas observações dos testes, não ser efetivamente prejudicial à qualidade do sistema; ainda, por serem ocorrências isoladas, conclui-se que não vale a pena realizar o processo de adaptação – pois até este se completar o descarte retornou a um valor tolerável. “pico” desapareceu. Com base nisto, foi implementada no sistema a auditoria da média dos últimos descartes; esta grandeza é capaz de caracterizar a ocorrência de congestionamento com maior precisão, avaliando a “tendência” do comportamento do tráfego e não o instante específico. Realizando novamente os testes, desta vez utilizando os dois recursos mencionados (o tempo de tolerância após o estabelecimento da chamada ou a adaptação e a apuração do descarte médio), o sistema manteve-se estável desde o primeiro nível, para ambos os perfis.

Teste 3 – Tráfego Concorrente

Este teste teve por objetivo diminuir a disponibilidade da rede, de forma que se pudesse diferenciar a influência percebida nas estatísticas referentes a descartes entre o que era gerado pela rede e o que era gerado pela CPU. A inclusão do tráfego concorrente foi feita utilizando-se dois programas – um servidor e um cliente de FTP, que viabilizaram a transferência de arquivos em taxas determinadas.

Os testes foram realizados com os dois perfis utilizados no Teste 2. De acordo com as execuções do sistema realizadas desta forma, foi possível observar o sistema realizando adaptação até níveis inferiores aos do Teste 2 quando da entrada de tráfego concorrente.

Como o consumo de capacidade de processamento dos programas cliente e servidor de FTP, observados isoladamente, foi bastante baixo, nos foi possível concluir que a grande maioria dos descartes reportados pela aplicação realmente tem origem na capacidade da rede.

9.3 Resultados e Observações

Com base nas observações, foi possível chegar a algumas conclusões sobre a aplicabilidade da estratégia de adaptação orientada a perfil.

A primeira observação foi referente à efetividade do modelo de adaptação, como solução válida para adequar o volume de tráfego à disponibilidade da rede.

A segunda observação relevante foi quanto à previsibilidade do sistema. O conjunto de recursos providos por esta abordagem tornou possível disponibilizar uma maneira do usuário manifestar suas preferências quanto ao comportamento da adaptação.

Em relação à implementação, também foi possível perceber a limitação da mídia de vídeo, quanto à taxa de quadros por segundo: a própria captura não apresentou desempenho satisfatório quando foram solicitadas taxas mais altas que 15 quadros por segundo (apenas em modo captura) ou 6 quadros por segundo (quando da captura e envio dos quadros). Este fato já havia sido mencionado por Zanin, e deve-se à necessidade de processamento da imagem na CPU, competindo com os processos em execução.

Ainda no que diz respeito às taxas de vídeo (quadros por segundo) obtidas, sabe-se que a utilização de mecanismos de compressão pode elevá-las em relação ao observado. Por exemplo, a adoção de compressão intraquadros pode diminuir bastante o consumo de rede; o uso de compressão interquadros pode, além de demandar menos banda, diminuir o consumo de CPU necessária para o processamento.

Finalmente, foi possível observar a influência das características da implementação no desempenho do sistema. Em algumas situações, o uso excessivo de CPU (principalmente em estações de menor capacidade de processamento) foi responsável pelo aumento no volume de descartes reportados pela aplicação, visto essa considerar apenas o que efetivamente recebia (através dos sockets) da rede.

10 Conclusões

Este trabalho apresentou um modelo de adaptação de tráfego de videoconferência voltado ao mercado pessoal. O modelo teve como proposta ser uma extensão do modelo de Zanin, desenvolvido na UFRGS, mediante a implementação de mais mídias e a distribuição de prioridade à adaptação de cada uma destas mídias. Esta distribuição foi realizada de forma orientada aos interesses do usuário, interesses estes específicos de cada situação de aplicação do sistema.

O levantamento bibliográfico realizado foi o ponto de partida para o trabalho, no sentido de fornecer a base conceitual necessária. O entendimento dos padrões de implementação e das necessidades de infraestrutura proporcionou o domínio teórico da tecnologia de videoconferência, bem como as questões de rede associadas a ela; o estudo dos problemas e enfoques para soluções, por sua vez, auxiliou na assimilação do papel da adaptação de tráfego e do modelo original no contexto de videoconferência.

De posse de tal conhecimento, foi possível construir o modelo de acordo com a proposta do trabalho – modelo de adaptação orientado ao perfil de aplicação. A construção do modelo se valeu dos componentes apresentados e definidos no modelo original, com a adição de mecanismos para possibilitar ao usuário definir o comportamento da adaptação, no que diz respeito à mídia e característica a serem relegadas. Tais componentes foram concebidos para proporcionar ao sistema um comportamento previsível do ponto de vista do usuário, capaz de atender suas expectativas específicas e, desta forma, aumentar a satisfação deste com o uso do sistema e, conseqüentemente, a aceitação da tecnologia.

Em relação às alternativas existentes para prover videoconferência com qualidade aceitável em ambientes de baixo custo (como o caso para o qual este trabalho foi direcionado: *desktop* via Internet), ficou clara a validade do modelo. A recente introdução de oferta de banda larga para o usuário final não influencia no sentido de tornar esse modelo obsoleto: por mais que esse tipo de conexão à Internet esteja em crescimento, a base de usuários em modems de 56 kbps ainda é muito grande, conforme já mencionado. No mercado corporativo, especificamente nas empresas menores – ou menos dispostas a investir em infraestrutura com suporte a banda larga – também é possível considerar que, apesar da crescente oferta, banda ainda não é um recurso abundante. Da mesma forma, as alternativas que visam proporcionar acesso à Internet com garantia de QoS compatível com os requisitos de aplicações são restritas em termos de oferta no mercado para o usuário comum.

O uso do formato de aplicação específica, proposto através do modelo de perfis de aplicações, pôde ser comprovado na prática mediante o protótipo de implementação. Através de tal implementação e de sua aplicação em testes, foi possível observar que concluir que o modelo é efetivo em relação à idéia a que se propõe. Tal conclusão foi baseada principalmente em três focos de observação distintos.

A primeira observação diz respeito ao uso do Gerenciador de Prioridades. Utilizando o protótipo específico deste componente, foram definidos os perfis utilizados nos testes. Com isto, comprovou-se ser possível e funcional disponibilizar ao usuário a definição do comportamento da adaptação e do que mais lhe interessa em uma sessão.

A segunda foi uma confirmação dos experimentos de Zanin, voltada à técnica da adaptação de tráfego em si. Na implementação realizada, foi possível observar a concorrência do tráfego de vídeo com os tráfegos de áudio e de outras mídias (tipicamente as rajadas geradas por compartilhamento de aplicações, quadro-branco, Chat – que foram tratadas de forma genérica como mídias discretas).

A terceira observação – maior e mais significativa delas, por estar diretamente relacionada à contribuição do trabalho – foi em relação ao uso do protótipo em aplicações diferentes. O protótipo foi submetido a testes com o uso dos perfis criados previamente – conseqüentemente, em situações distintas em termos de aplicação e de expectativa do comportamento da adaptação por parte do usuário. Nesses testes, foi possível validar o uso da orientação a perfil de aplicação como forma de priorizar as características (resoluções de cada mídia) desejáveis pelo usuário para cada aplicação.

Ainda em relação aos testes, convém observar novamente que os exemplos de perfis utilizados nos testes foram extraídos da interpretação de alguns estudos de caso. Como o objetivo do modelo foi construir uma forma de suportar a definição dos perfis, e não propriamente defini-los, tal definição foi considerada adequada.

Como recomendação do uso do modelo em si, é recomendável que o produto final desenvolvido a partir dele deva incluir algumas definições de perfis básicos, providas pelo próprio fornecedor. Essa recomendação se deve ao fato de que é necessário conhecimento e experiência prévios das aplicações de videoconferência, coisa que não se pode esperar de um usuário iniciante. Ao mesmo tempo, é necessário dispor da ferramenta de definição (tal como o foi na implementação), de forma a possibilitar que usuários mais experientes definam seu próprio perfil de aplicação, de acordo com suas experiências e necessidades.

Como perspectivas de trabalhos futuros, existem várias frentes de pesquisa a serem exploradas. Dentre essas, devem ser prioritários os estudos para evolução da implementação, da infraestrutura e o estudo dos perfis de aplicação.

O desenvolvimento da infraestrutura é um tema que vem sendo bastante explorado. As principais contribuições desse enfoque para o nicho de aplicação dizem respeito à pesquisa das formas de prover QoS para usuário final, o que pode ser utilizado em complemento à técnica de adaptação explorada nesse trabalho.

No que tange à evolução da implementação do protótipo, é importante implementar a adequação aos padrões definidos pelo ITU-T, no sentido de consolidar o protótipo como uma ferramenta efetiva. Nesse sentido, os principais tópicos a serem observados são a implementação dos protocolos RTP/RTCP, compressão de vídeo, suporte a multicast e aos mecanismos de sinalização especificados na recomendação H.323, conforme definidos na recomendação H.245.

Finalmente, um tópico muito interessante a ser explorado é o mapeamento dos recursos por tipo de aplicação. Esta tarefa, que teria por objetivo criar os perfis a serem usados no modelo proposto, deve ser realizada mediante testes práticos com uma massa significativa de usuários, a exemplo da mensuração de MOS definida pelo ITU-T, e preferencialmente assistido por profissionais das áreas afins (como comunicação, para avaliar a efetividade da mesma, e psicologia, para dimensionar a aceitação por parte do

usuário). O resultado final desta análise pode ser aplicado em aplicações desenvolvidas a partir desse modelo, e novamente submetidas ao usuário, para avaliar a satisfação deste com a utilização do sistema. Este tende a ser este um importante passo para proporcionar uma melhor aceitação da tecnologia, utilizando o máximo potencial do modelo definido neste trabalho.

Anexo – Definições Usadas na Implementação

Este anexo apresenta a documentação das definições utilizadas na implementação do protótipo, e visa esclarecer como as funcionalidades implementadas funcionam do ponto de vista da aplicação.

Trocas de Mensagens

As funções do sistema que demandam interação entre as duas estações são realizadas por meio de envio e recepção de mensagens, com conteúdos específicos e para portas específicas. A seguir serão apresentadas as principais trocas de mensagens implementadas no protótipo.

Troca de mensagens para Estabelecimento da Chamada

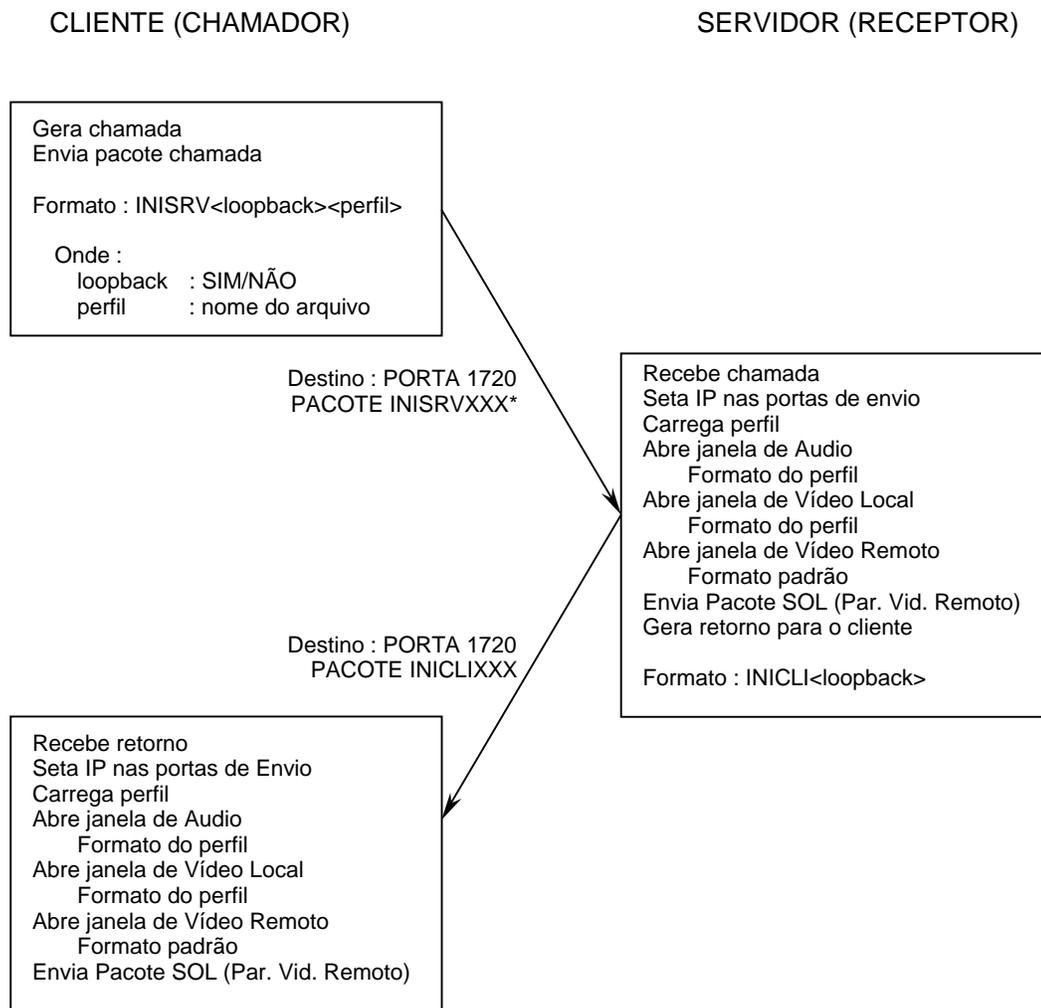


FIGURA 1 – Estabelecimento da Chamada

Troca de Mensagens para Solicitação de Parâmetros de Captura Remotos

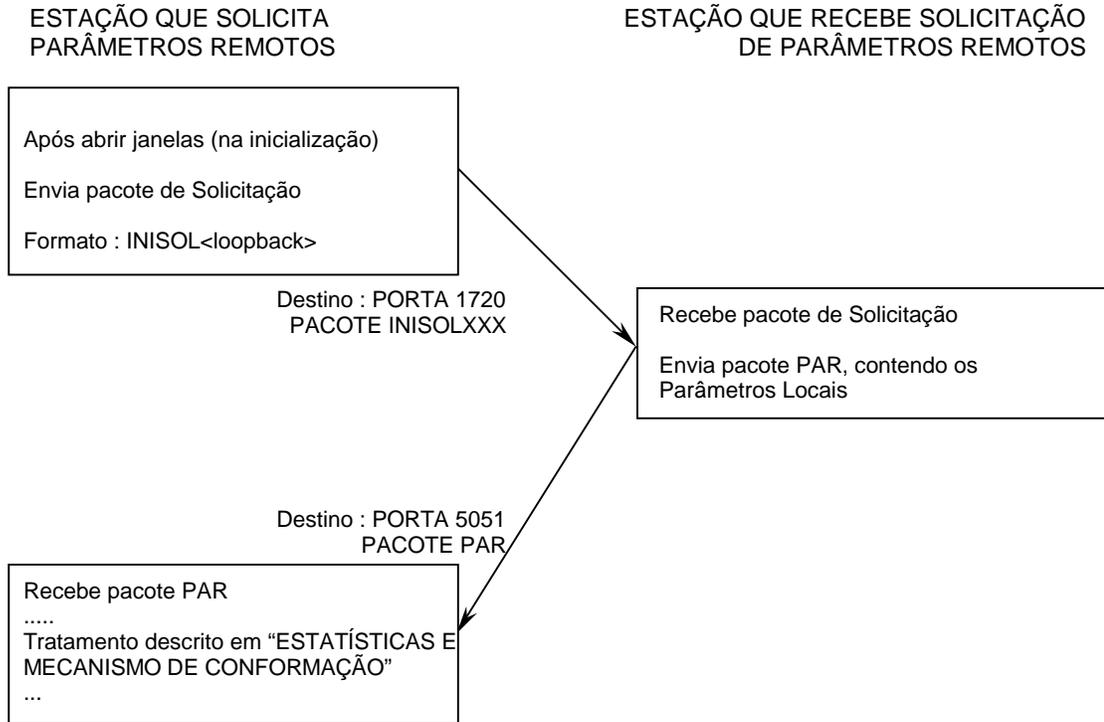


FIGURA 2 – Solicitação dos Parâmetros de Captura Remotos

Troca de Mensagens para Encerramento da Chamada

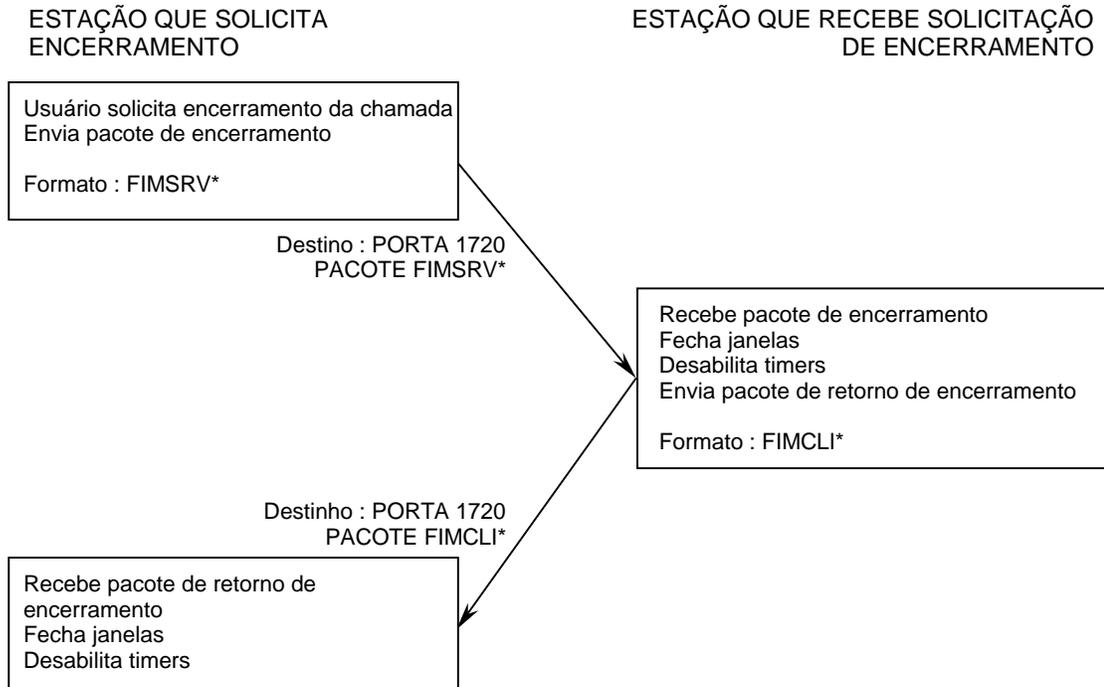


FIGURA 4 – Encerramento da Chamada

Formato e Conteúdo dos Pacotes

A seguir são apresentados os formatos dos pacotes implementados no protótipo.

Pacote de Áudio

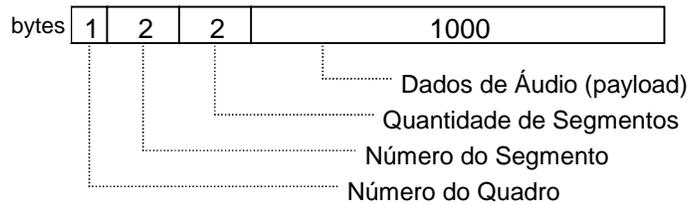


FIGURA 5 – Formato do pacote de Áudio

Pacote de Vídeo

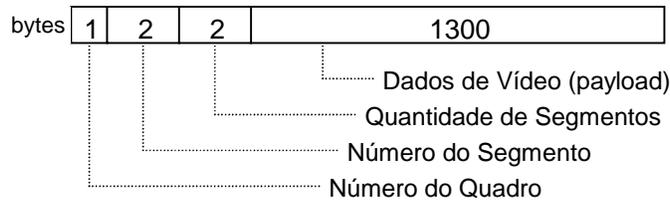


FIGURA 6 – Formato do pacote de Vídeo

Pacote INI Estabelecimento de Chamada – Solicitação

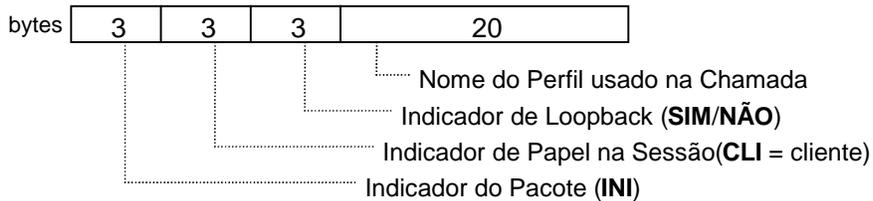


FIGURA 7 – Formato do pacote de Solicitação de Estabelecimento de Chamada

Pacote INI – Estabelecimento de Chamada – Aceite

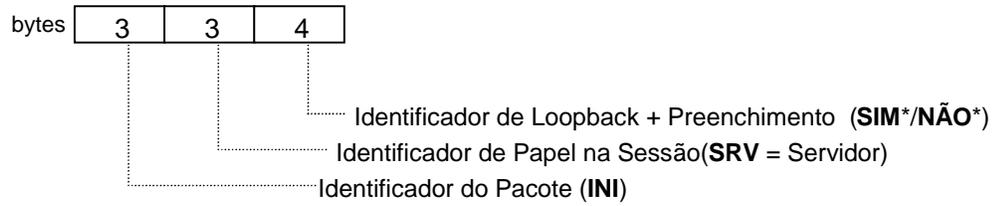


FIGURA 8 – Formato do pacote de Aceite de Estabelecimento de Chamada

Pacote FIM – Encerramento de Chamada – Solicitação

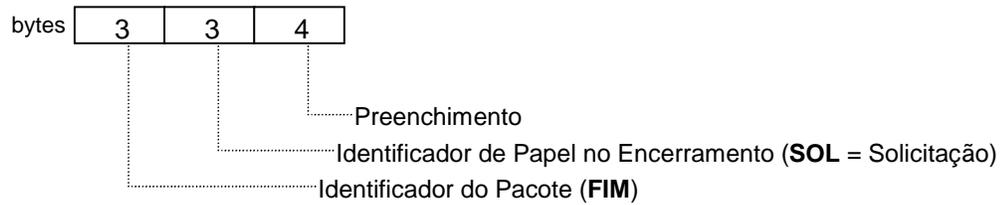


FIGURA 9 – Formato do pacote de Solicitação de Encerramento de Chamada

Pacote FIM – Encerramento de Chamada – Aceite

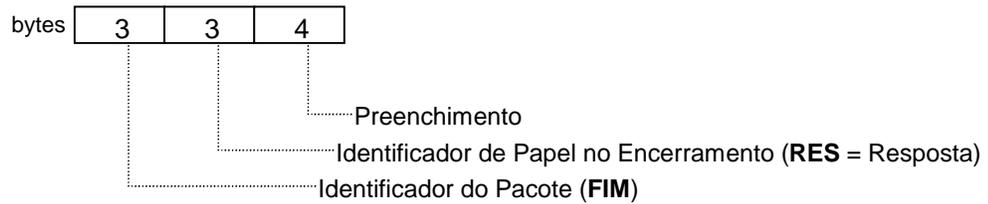


FIGURA 10 – Formato do pacote de Aceite de Encerramento de Chamada

Pacote EST – Envio de Estatísticas

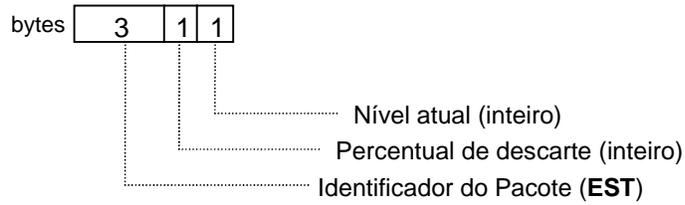


FIGURA 11 – Formato do pacote de Estatísticas

Pacote PAR – Adaptação / Parâmetros Locais

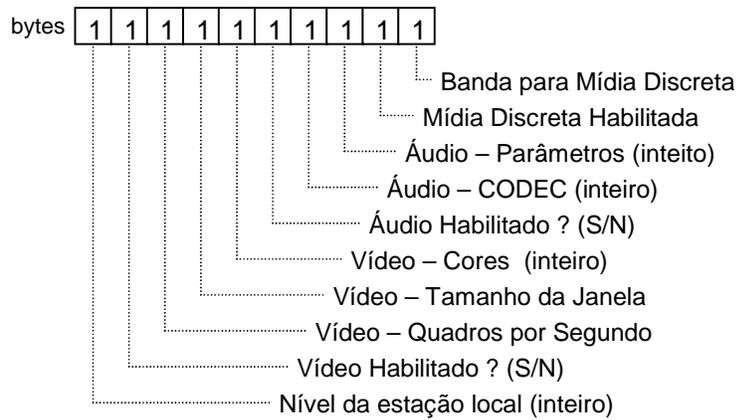


FIGURA 12 – Formato do pacote de Informação de Parâmetros Locais

Pacote RET – Retorno (confirmação de ajuste)

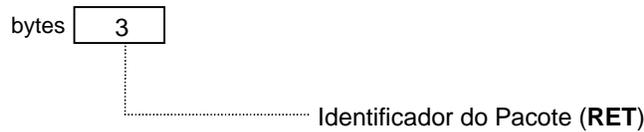


FIGURA 13 – Formato do pacote de Confirmação de Ajuste

Pacote de Abertura/Fechamento do Chat

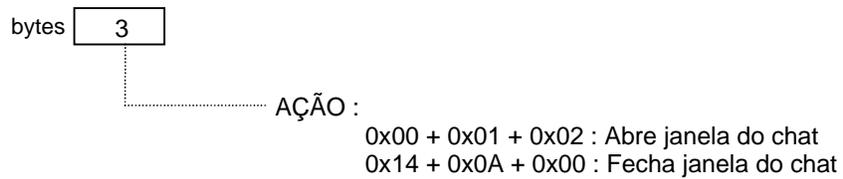


FIGURA 14 – Formato do pacote de Abertura / Fechamento do Chat

Pacote de Dados do Chat

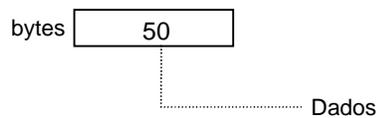


FIGURA 15 – Formato do pacote Dados do Chat

Pacote de Abertura/Fechamento do Simulador de Mídias Discretas

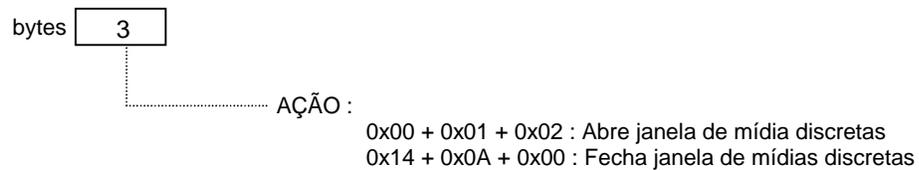


FIGURA 16 – Formato do pacote de Abertura / Fechamento do Simulador de Mídias Discretas

Pacote de Dados de Mídias Dscretas

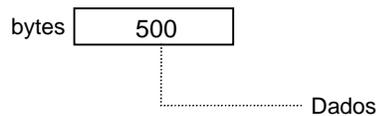


FIGURA 17 – Formato do pacote de Dados de Mídias Discretas

Portas utilizadas

A implementação considerou o uso de diversas portas socket, de forma a dedicar números de porta específicos para determinadas funções do sistema. A Tabela 1 apresenta as portas que foram utilizadas na implementação.

TABELA 1 – Portas utilizadas pelo Protótipo

Protocolo	Porta	Função
TCP	1720	Estabelecimento e Encerramento da Chamada
UDP	2000	Envio de Vídeo
UDP	2001 a 2020	Recepção de Vídeo
UDP	3000	Envio de Áudio
UDP	3001 a 3010	Recepção de Áudio
TCP	5050	Envio de Estatísticas e Retorno de Adaptação
TCP	5051	Envio de Parâmetros Locais
TCP	3050	Chat (Abertura/Fechamento e Conteúdo)
TCP	3100	Mídias discretas (Abertura/Fechamento e Envio)
TCP	3101 a 3110	Recepção de mídias discretas

Formatos suportados na Implementação

Na implementação do protótipo, foram utilizados alguns formatos para cada atributo das mídias contínuas (vídeo e áudio). Os formatos são relacionados nas Tabelas 2 e 3 a seguir, juntamente com a identificação (código) que os mesmos recebem pelo protótipo e pelo Gerenciador de Perfis. Os códigos apresentados também são enviados de uma estação para a outra, através do pacote PAR, para informar seus parâmetros locais.

TABELA 2 – Valores dos atributos de Vídeo suportados pelo Protótipo

	Atributo	Valor	Código
Vídeo	Cores	16 bits (64 k cores)	16
		24 bits (16 M cores)	24
	Tamanho da Janela	160 x 120	1
		176 x 144 (QCIF)	2
		320 x 240	3
		352 x 288 (CIF)	4

TABELA 3 – Valores dos atributos de Áudio suportados pelo Protótipo

	Formato (CODEC)	Código	Resolução	Código
Áudio	PCM	3	8000 Hz / 8 bits / Mono	1
			11025 Hz / 8 bits / Mono	2
			22050 Hz / 8 bits / Mono	3
			44100 Hz / 16 bits / Stereo	4
	ADPCM	1	8000 Hz / 4 bits / Mono	1
			11025 Hz / 4 bits / Mono	2
			22050 Hz / 4 bits / Mono	3
			44100 Hz / 4 bits / Stereo	4
	GSM	2	8000 Hz / 2 bits / Mono	1
			11025 Hz / 2 bits / Mono	2
			22050 Hz / 2 bits / Mono	3
			44100 Hz / 2 bits / Mono	4

Bibliografia

- [BER 2000] BERNET, Yoram. The Complementary Roles of RSVP and Differentiated Services in the Full-Service QoS Network. **IEEE Communications Magazine**, New York, v.38, n.2, p.154-162, Feb. 2000.
- [BRO 99] BROWN, Dave. Videoconferencing and Multimedia Delivery. **Network Computing**, Feb. 1999. Disponível em: <<http://www.networkcomputing.com/netdesign/1003videoconf.html>>. Acesso em: 10 dez. 1999.
- [CHA 98] CHANG, Kai; SEMERIA, Chuck. **Standards for Multimedia Applications on Converging Networks**. 3Com Corporation, May 1998. Disponível em: <<http://www.summitonline.com/techtrends/papers/3com2.html>>. Acesso em: 27 jan. 2000.
- [CIS 2001] CISCO SYSTEMS. **VoIP – Understanding Codecs: Complexity, Support, MOS, and Negotiation**. [S.l.], 2001. Disponível em: <http://www.cisco.com/warp/public/788/voip/codec_complexity.html>. Acesso em : 10 set. 2001.
- [CIS 2001a] CISCO SYSTEMS. **Session VVT-230 – Designing and Deploying IP Videoconferencing**. São Paulo : NETWORKERS, 2001.
- [CIS 98] CISCO SYSTEMS. **Cisco voice over Frame Relay, ATM e IP – Student Guide**. San Jose, 1998. v.1.
- [CIS 99] CISCO SYSTEMS. **Designing Internetworks for Multimedia**. [S.l.], 1999. Disponível em: <<http://www.cisco.com/cpress/cc/td/cpress/ccie/ndcs798/nd2013.htm>>. Acesso em : 26 jan. 2000.
- [COX 96] COX, Richard V. ; KROON, Peter. Low Bit-Rate Speech Coders for Multimedia Communication. **IEEE Communications Magazine**, New York, v.34, n.12, p.34-41. Dec.1996.
- [CRO 97] CROSSMAN, A. **Summary of ITU-T Speech / Audio Codecs used in the ITU-T Videoconferencing Standarts**. PictureTel Corporation, July 1997. Disponível em: <http://standards.pictel.com/reference/summary_itu_codecs.htm>. Acesso em : 10 nov. 1999.
- [DAT 99] DATABEAM CORP. **A primer on the H.323 Series Standard - Version 2.0**, 1999. Disponível em: <<http://www.databeam.com/h323/h323primer.html>>. Acesso em: 21 out 1999.

- [DAV 2000a] DAVIS, Andrew W. **Why Desktop Video is Inevitable**. Wainhouse Research / Teleconference Magazine, Mar. 2000. Disponível em: <<http://www.wainhouse.com/articles/otd0300.html>>. Acesso em: 11 ago. 2001.
- [DIX 2000] DIXON, Robert S., Internet Videoconferencing: Coming to Your Campus Soon!. **Educause Quaterly**, [S.l.], n.4, p.22-27, 2000. Disponível em <<http://www.educause.edu/pub/eq/eqm00/eqm004.html>>. Acesso em: 15 jul. 2001.
- [FAY 99] FAYERS, G. J. **PPNCG Entry-Level Product Recommendations for Videoconferencing in HEP**. Londres : Imperial College of Science, Technology and Medicine, 1999. Disponível em: <<http://www.hep.ph.ic.ac.uk/ppncg/VidConRecommendations2.doc>>. Acesso em: 22 set. 2001.
- [FRA 2001] FRATER, Michael; ARNOLD, John F.; VAHEDIAN, Abedin. Impact of Audio on Subjective Assessment of Video Quality in Videoconferencing Applications. **IEEE Transactions on Circuits and Systems for Video Technology**, New York, v.11, n.9, p. 1059-1062, Sept. 2001.
- [FRE 2001] FREIRE, Simone. **Ibope eRatings mapeia uso da banda larga no Brasil**. IDG Now! 21 nov. 2001. Disponível em: <<http://idgnow.terra.com.br/idgnow/internet/2001/11/0046>>. Acesso em: 24 nov. 2001.
- [GRI 2001] GRIGONIS, Richard. Conferencing Technology in the Spotlight. **Computer Telephony**, [S.l.], 2001. Disponível em: <<http://computertelephony.com/article/CTM20011031S0001>>. Acesso em: 19 dez. 2001.
- [GUE 97] GUEDES, L. A.; CARDOZO, E. Especificação de um Protocolo para Negociação de Qualidade de Serviço em Sistemas Multimídia. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES, SBRC, 15., 1997, São Carlos. **Anais...** São Carlos:UFSCAR, 1997. p.101-116.
- [HUT 97] HUTCHISON, David; EL-MARAKBY, Randa; MATHY, Laurent. A Critique of Modern Internet Protocols : The Issue of Support for Multimedia. In: EUROPEAN CONFERENCE ON MULTIMEDIA APPLICATIONS, SERVICES AND TECHNIQUES, ECMAST, 2., 1997, Milan. **Multimedia Applications, Services and Techniques : Proceedings**. Berlin:Springer-Verlag, 1997. p. 507-522 (Lecture Notes in Computer Science, v.1242).
- [IMT 99] IMTC – INTERNATIONAL MULTIMEDIA TELECONFERENCING CONSORTIUM, INC. **ITU-T Standards Process**. Disponível em: <<http://www.imtc.org/standardsbody.htm>>. Acesso em : 9 out. 1999.

- [ING 98] INGVALDSEN, Tone; KLIVNING, Espen; WILKINS, Miles. A study of Delay Factors in CSCW Applications and Their Importance. In: INTERNATIONAL WORKSHOP ON INTERACTIVE DISTRIBUTED MULTIMEDIA SYSTEMS AND TELECOMMUNICATION SERVICES, IDMS, 5., 1998, Oslo. **Interactive Distributed Multimedia Systems and Telecommunication Services: Proceedings**. Berlin:Springer-Verlag, 1997, p.159-170.
- [ISO 98] ISOKANGAS, Jari. **H.320 based videoconferencing in ISDN, Ethernet and ATM Networks**. TUT - Tampere University of Technology, Tampere, Finlândia, 1998. Disponível em: <<http://matwww.ee.tut.fi/kamu/distancedocs/jarii.htm>>. Acesso em: 19 dez. 1999.
- [ITU 93] ITU-T – INTERNATIONAL TELECOMMUNICATIONS UNION – TELECOMMUNICATIONS STANDARDIZATION SECTOR. **ITU-T Recommendation G.711 – Pulse Code Modulation (PCM) of Voice Frequencies**. Geneva, 1993. 10p.
- [ITU 96] ITU-T – INTERNATIONAL TELECOMMUNICATIONS UNION – TELECOMMUNICATIONS STANDARDIZATION SECTOR. **ITU-T Recommendation F.702 – Multimedia conference services**. [S.l], 1996, 27 p.
- [ITU 96a] ITU-T – INTERNATIONAL TELECOMMUNICATIONS UNION – TELECOMMUNICATIONS STANDARDIZATION SECTOR. **ITU-T Recommendation G.114 – One-way Transmission Time**. [S.l], 1996, 13 p.
- [ITU 96b] ITU-T – INTERNATIONAL TELECOMMUNICATIONS UNION – TELECOMMUNICATIONS STANDARDIZATION SECTOR. **ITU-T Recommendation P.800 - Methods for subjective determination of transmission quality**. Geneva, 1996, 37 p.
- [ITU 97] ITU-T – INTERNATIONAL TELECOMMUNICATIONS UNION – TELECOMMUNICATIONS STANDARDIZATION SECTOR. **ITU-T Recommendation H.263 – Video Coding for Low Bit Rate Communication**. Geneva, 1996, 44p.
- [ITU 97a] ITU-T – INTERNATIONAL TELECOMMUNICATIONS UNION – TELECOMMUNICATIONS STANDARDIZATION SECTOR. **ITU-T Recommendation H.320 - Narrow-band Visual Telephone Systems and Terminal Equipment**. [S.l], 1997, 19p.
- [ITU 98] ITU-T – INTERNATIONAL TELECOMMUNICATIONS UNION – TELECOMMUNICATIONS STANDARDIZATION SECTOR. **ITU-T Recommendation H.323 – Packed-based multimedia communications systems**. [S.l], 1998, 114 p.

- [JAC 96] JACKOWSKI, S. **Native ATM Support for ST2+**: RFC 1946. [S.I.]: IETF, maio 1996. 21 p.
- [LEI 99] LEITE, Carlos A. B. B. **Implementação de QoS na Internet: Uma Visão**. 1999. Relatório (Doutorado em Ciência da Computação). LCMI-DAS, UFSC, Florianópolis. Disponível em: <<http://www.lcmi.ufsc.br/redes/redes98/cbrandao/qosinternet/qosinternet.html>>. Acesso em: 14 jan. 2002.
- [LI 2000] LI, Bo et. al. QoS-Enabled Voice Support in the Next-Generation Internet: Issues, Existing Approaches and Challenges. **IEEE Communications Magazine**, New York, v.38, n.4, p.54-61, Apr. 2000.
- [LUN 99] LUNARDI, Sediane C. **Protocolos de Suporte a Integração de Serviços**. 1999. Trabalho Individual (Mestrado em Ciência da Computação) – Faculdade de Informática, PUC, Porto Alegre.
- [MAS 2001] MASETTI, Ivan. Terrorismo impulsiona videoconferência. **RNT - Revista Nacional de Telecomunicações**, São Paulo, n.267, p. 26-28, nov. 2001.
- [MOH 2000] MOHR, Alexander E.; RISKIN, Eve A.; LADNER, Richard E. Unequal Loss Protection : Graceful Degradation of Image Quality over Packet Erasure Channels Through Forward Error Correction. **IEEE Journal on Selected Areas in Communications**, New York, v.18, n.6, p.819-828, June 2000.
- [OPE 2000] OPERAX NET SERVICES. **The challenge of enabling QoS in IP networks**. [S.I.], 2000. Disponível em: <<http://www.operax.com/docs/whitepaper-QoS-challenge-C.pdf>>. Acesso em: 20 dez. 2001.
- [PAT 2001] PATRICK, Emilee J. **Barriers to Collaboration : User-Centered Research and the Acces Grid**. Motorola Labs : Applications Research, January 2001. Disponível em: <http://internet2.motlabs.com/agnode/papers/epatrick_paper-barierstocollaboration.pdf>. Acesso em: 22 jan. 2002.
- [PIC 2001] PICTURETEL CORPORATION. **Reliability & IP Performance**. [S.I.], 2001. Disponível em: <<http://www.picturetel.com/pdf/whitepapers/PictureTelReeliabilityandIPPerformanceWhitePaper.pdf>>. Acesso em: 14 ago. 2001.
- [POS 80] POSTEL, J. **User Datagram Protocol**: RFC 768. [S.I.]: IETF, ago. 1980. 3p.
- [PUR 2001] PURI, Rohit; LEE, Kang-Won; RAMCHANDRAN, Kannan; BHARGHVAN, Vaduvur. An Integrated Source Transcoding and Congestion Control Paradigm for Video Streaming in the Internet. **IEEE Transactions on Multimedia**, New York, v.3, n.1, p. 18-32, Mar. 2001.

- [RET 95] RETTINGER, Leigh A. **Desktop Videoconferencing**: Technology and use for Remote Seminar Delivery. 1995. Dissertação de Mestrado, North Carolina State University, Raleigh.
- [REY 2001] REYNARD, Gail. **Multimedia Quality of Service**. Disponível em: <<http://www.crg.cs.nott.ac.uk/~gtr/MVR/QoS.pdf>>. Acesso em: 12 fev 2002.
- [REY 98] REYNARD, Gail; BENFORD, S.; GREENHALGH, C. Awareness Driven Video Quality of Service in Collaborative Virtual Environments. In: ACM CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, CHI, 1998, Los Angeles. **Proceedings...** Disponível em: <<http://www.crg.cs.nott.ac.uk/~gtr/CHI98.ps.gz>>. Acesso em: 4 jan. 2000.
- [SCH 2000] SCHULZRINNE, Henning, **Some Frequently Asked Questions About RTP**. Universidade de Columbia, 2000. Disponível em: <<http://www.cs.columbia.edu/~hgs/rtp/faq.html>>. Acesso em: 17 maio 2002.
- [SOA 95] SOARES, Luis Fernando Gomes et al. **Das LANs, MANs e WANs às Redes ATM**. Rio de Janeiro: Campus, 1997. 576 p.
- [STA 99] STARDUST.COM, INC. **The Need for QoS**. QOS Forum, 1999. Disponível em: <http://www.qosforum.com/white-papers/Need_for_QoS-v4.pdf>. Acesso em: 15 mar. 2001.
- [STE 95] STEINMETZ, Ralf; NAHRSTEDT, Klara. **Multimedia** : Computing, Communications & Applications. New Jersey: Prentice Hall, 1995. 854p.
- [STE 96] STEINMETZ, Ralf. Human Perception of Jitter and Media Synchronization. **IEEE Journal on Selected Areas in Communications**, New York, v.14, n.1, p.61-72, Jan. 1996.
- [TAN 92] TANG, John C.; ISAACS, Ellen A. **Why do Users Like Video? Studies of Multimedia-Supported Collaboration**. Santa Clara: Sun Microsystems, 1992. Disponível em: <<http://research.sun.com/techrep/1992/sml-tr-92-5.pdf>> Acesso em: 3 dez 2001.
- [TAN 97] TANENBAUM, Andrew. **Redes de Computadores**. Rio de Janeiro: Campus, 1997. 923p.
- [TRE 97] TRENTIN, Marco A. S. **Serviços de Rede para Apoiar um Centro de Ensino Remoto Interativo**. 1997. Dissertação (Mestrado em Ciência da Computação) – Instituto de Informática, UFRGS, Porto Alegre
- [VOG 95] VOGEL, A.; KERHERVÉ, B. Distributed Multimedia and QoS: Survey. **IEEE Multimedia**, New York, v.2, n.2, p.10-18, Summer 1995.

- [VSI 99] VIDEOCONFERENCING SOLUTIONS INC. **International Telecommunications Union Standards for Videoconferencing**. [S.l.], 1999. Disponível em: <<http://h323standards.com/technical/h320-h323standards.htm>>. Acesso em: 15 ago. 1999.
- [WAI 2001] WAINHOUSE RESEARCH. **Videoconferencing and Visual Collaboration Endpoint – Deployments and Plans**. [S.l.], 2001. Disponível em <<http://www.wainhouse.com>>. Acesso em: 3 ago. 2001.
- [WIL 2000] WILCOX, James R. **Videoconferencing – the whole picture**. 3rd ed. New York : Telecom Books, 2000. 611p.
- [XIA 99] XIAO, Xipeng; NI, Lionel M. Internet QoS : A Big Picture. **IEEE Network Magazine**, New York, v.13, n.2, p.8-18, Mar./Apr. 1999.
- [ZAN 2000] ZANIN, Fábio A. **Um Modelo para Videoconferência em Computador Pessoal sobre Redes IP**. 2000. Dissertação (Mestrado em Ciência da Computação) – Instituto de Informática, UFRGS, Porto Alegre.
- [ZAN 99] ZANIN, Fábio A. **O Ambiente de Videoconferência sobre redes ATM**. 1999. Trabalho Individual (Mestrado em Ciência da Computação) – Instituto de Informática, UFRGS, Porto Alegre.