

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**TECNOLOGIAS PARA COMUNICAÇÃO COLABORATIVA
EM DIREÇÃO À MELHORIA DA EXPERIÊNCIA E
INTEGRAÇÃO DE PLATAFORMAS**

ALAOR ARRUDA FLORÊNCIO

ORIENTADOR: LUIS FERNANDO RAMOS MOLINARO

DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA

**PUBLICAÇÃO: 400 /DISSERTAÇÃO/2009
BRASÍLIA, DF: OUTUBRO / 2009.**

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

**TECNOLOGIAS PARA COMUNICAÇÃO COLABORATIVA
EM DIREÇÃO À MELHORIA DA EXPERIÊNCIA E
INTEGRAÇÃO DE PLATAFORMAS**

ALAOR ARRUDA FLORÊNCIO

**DISSERTAÇÃO DE MESTRADO SUBMETIDA AO DEPARTAMENTO DE
ENGENHARIA ELÉTRICA DA FACULDADE DE TECNOLOGIA DA
UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS
NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE.**

APROVADO POR:

LUIS FERNANDO RAMOS MOLINARO, DOUTOR, ENE/UNB (ORIENTADOR)

**HUMBERTO ABDALLA JÚNIOR, DOUTOR, ENE/UNB
(EXAMINADOR INTERNO)**

**FLAVIO ELIAS DE DEUS, DOUTOR, ENE/UNB
(EXAMINADOR EXTERNO)**

BRASÍLIA, DF, 1 DE OUTUBRO DE 2009.

FICHA CATALOGRÁFICA

FLORÊNCIO, ALAOR ARRUDA-
TECNOLOGIAS PARA COMUNICAÇÃO COLABORATIVA EM DIREÇÃO À
MELHORIA DA EXPERIÊNCIA E INTEGRAÇÃO DE PLATAFORMAS DE
COMUNICAÇÃO, Distrito Federal, 2009
vii, 124 p., 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2009).
Dissertação de Mestrado – Universidade de Brasília, Faculdade de Tecnologia.
Departamento de Engenharia Elétrica.

1. Comunicações Unificadas

2. Videoconferência

3. H.323

4. SIP

I. ENE/FT/UnB

II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

FLORÊNCIO, Alaor A. (2009). Tecnologias para Comunicação Colaborativa em Direção à Melhoria da Experiência e Integração de Plataformas de Comunicação. Dissertação de Mestrado, Publicação 400/2009, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 124 p.

CESSÃO DE DIREITOS

AUTOR: Alaor Arruda Florêncio

TÍTULO DA TESE: TECNOLOGIAS PARA COMUNICAÇÃO COLABORATIVA EM DIREÇÃO À MELHORIA DA EXPERIÊNCIA E INTEGRAÇÃO DE PLATAFORMAS

GRAU / ANO: Mestre / 2009

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta dissertação de mestrado pode ser reproduzida sem a autorização por escrito do autor.

Alaor Arruda Florêncio

Antonio das Chagas. 828, apto.1007

CEP: 04.714-001, Chácara Santo Antônio, São Paulo - SP

Tel. +55-11-3443-1514 / alaorarruda@gmail.com

AGRADECIMENTOS

Agradeço em primeiro lugar á minha grande família, pais e irmãos, pelo apoio, incentivo e paciência desde muito antes desta etapa.

Á minha querida Deisy por ter acompanhado de perto todo o trajeto que levou à conclusão desta importante etapa e por ter me apoiado e suportado durante esse tempo.

Ao professor e orientador Luiz Fernando Molinaro pelo apoio e oportunidade, além de ter provido um ambiente que me permitiu crescer profissionalmente e pessoalmente durante este trajeto.

Á equipe do Núcleo de Multimídia e Internet, de onde saíram grandes amigos.

RESUMO

Autor: Alaor Arruda Florêncio

Orientador: Luis Fernando Ramos Molinaro

Programa de Pós-graduação em Engenharia Elétrica

Brasília, setembro de 2009

As tecnologias de áudio e vídeo, visando a comunicação entre pessoas, já estão disponíveis há vários anos. Estas foram implementadas como redes independentes entregues por fabricantes diversos e por empresas que tinham visões diferentes sobre o futuro das redes de comunicações. É perceptível nos dias de hoje o aumento do uso de diversas mídias para permitir a comunicação entre pessoas. O que se observa entre os fabricantes, integradores e dentro de organizações é um maior aproveitamento das redes IP (*Internet Protocol*) para fazer com que estas tecnologias trabalhem de forma integrada formando um conjunto conhecido atualmente como Comunicações Colaborativas Unificadas.

Esta dissertação está baseada em um estudo sobre as tecnologias desenvolvidas para possibilitar essas comunicações, mostrando as direções para onde estas evoluíram até se tornarem mais aceitas e confiáveis. A abordagem visa detalhar desde os avanços que afetam diretamente a experiência do usuário final até aqueles que vão influenciar na forma como são projetadas as redes que suportam a transmissão dos dados.

Por fim, são apresentadas soluções que possibilitam a implementação de redes integradas que venham a otimizar o uso das soluções disponíveis atualmente. Uma vez que as redes de comunicações unificadas estão em ascensão e desenvolvimento, as implementações devem estar preparadas para lidar com novas funcionalidades e demandas. Estas demandas vão além de aspectos puramente tecnológicos, sendo provenientes principalmente da necessidade de se aprimorar a experiência do usuário final ao interagir com esses sistemas.

ABSTRACT

Author: Alaor Arruda Florêncio

Supervisor: Luis Fernando Ramos Molinaro

Programa de Pós-graduação em Engenharia Elétrica

Brasília, september of 2009

Audio and video technologies intended to provide means of communication between people have been available for several years. These were implemented as independent networks delivered by a number of manufacturers and by integrator companies that held different visions about the future of communications networks. It's evident nowadays an increased usage of different kinds of medias that allow people to communicate. A tendency one can see these days among manufacturers, integrators and enterprises is a better utilization of IP (Internet Protocol) networks aiming at having these technologies to work in an integrated manner, becoming what is known today as Collaborative Unified Communications.

This dissertation is based on studies regarding technologies developed to make these communications available, highlighting the directions to where those evolved to become more acceptable and reliable. The focus is to detail developments from those that directly affects the end user experience until those that will influence the way these networks are projected to support data transmissions.

Finally some solutions are presented aiming at the implementations of integrated networks to enhance the use of the currently available solution. Since the Unified Communications networks are in development and ascending, the implementations must be prepared to deal with new functionalities and demands. These demands are beyond purely technological aspects, being proceeding mainly from the need to improve the end user experience when interacting with these systems.

SUMÁRIO

1 - INTRODUÇÃO.....	1
1.1 - APLICAÇÕES DE COMUNICAÇÕES COLABORATIVAS	3
<i>1.1.1 - Áudio sobre IP e áudio conferências</i>	<i>3</i>
<i>1.1.2 - Videoconferência</i>	<i>5</i>
<i>1.1.3 - Mensagens e presença</i>	<i>9</i>
1.2 - TECNOLOGIAS E AMBIENTES.....	9
1.3 - MOTIVAÇÃO	11
1.4 - OBJETIVOS	11
1.5 - ORGANIZAÇÃO DO TRABALHO	12
2 - PADRÕES PARA COLABORAÇÃO MULTIMÍDIA.....	13
2.1 - CODECS DE ÁUDIO.....	15
2.2 - CODECS DE VÍDEO	18
2.3 - PADRÃO DE COMUNICAÇÃO H.320.....	23
2.4 - PADRÃO DE COMUNICAÇÃO H.323.....	26
<i>2.4.1 - Componentes da Arquitetura H.323</i>	<i>26</i>
2.5 - SESSION INITIATION PROTOCOL - SIP	33
3 - REDES.....	37
3.1 - REDES COMUTADAS POR PACOTES E REDES COMUTADAS POR CIRCUITOS	37
3.2 - INTEGRATED SERVICES DIGITAL NETWORK - ISDN.....	40
3.3 - TRÁFEGO DE DADOS NAS REDES IP	44
<i>3.3.1 - Transport Control Protocol/ Internet Protocol - TCP/IP.....</i>	<i>45</i>
<i>3.3.2 - User Datagram Protocol (UDP).....</i>	<i>47</i>
<i>3.3.3 - Tráfego Contínuo.....</i>	<i>49</i>
<i>3.3.4 - QoS – Quality of Service.....</i>	<i>53</i>
3.4 - REDES MPLS.....	64
<i>3.4.1 - Vantagens do MPLS:.....</i>	<i>66</i>
4 - TENDÊNCIAS	70
4.1 - PLATAFORMAS INTEGRADAS.....	70
<i>4.1.1 - Implementações</i>	<i>74</i>

4.2 - TELEPRESENÇA	79
4.3 - AVALIAÇÃO	82
4.3.1 - <i>Uso e implementação</i>	82
4.3.2 - <i>ISDN para IP</i>	86
4.3.3 - <i>SIP e H.323</i>	87
5 - CONCLUSÃO	90
REFERÊNCIAS BIBLIOGRÁFICAS	92
APÊNDICES	96
A. REQUERIMENTOS DE SALA	97
A.1 - Iluminação	97
A.2 - Áudio	99
A.3 - Acústica	102

LISTA DE TABELAS

Tabela 2-1 - Padrões de comunicação ITU-T.....	14
Tabela 2-2 – Principais codecs de áudio	16
Tabela 2-3 – Exemplos de relações entre banda de áudio e taxas de bits	18
Tabela 2-4 – Formatos de vídeo suportados pela H.263 e H.261 (ITU-T H.263, 1996 modificado).....	21
Tabela 2-5 – Formatos SD e HD introduzidos com o H.264	22
Tabela 3-1 – Configuração de um acesso BRI	41
Tabela 3-2 – Composição de um acesso PRI	42
Tabela 3-3 – Parâmetros para comunicação de áudio e vídeo.....	53
Tabela 3-4 – Equivalência entre os valores IP Precedence e DSCP (O’Neil, 2002).....	58
Tabela 3-5 – Classes de tráfego usadas pelo CoS	62
Tabela 4-1- Benefícios da videoconferência (Kerravala, 2004).....	83
O lux é a unidade internacional de medida utilizada para indicar a quantidade de luz emitida por um corpo luminoso, por metro quadrado (Serra, 2007), a Tabela A-1 contém alguns exemplos de medidas.	98
Tabela A-1 – Medidas de iluminação em Lux	98
Tabela A-2 – Relação entre tamanho da tela e distância dos participantes.....	105
Tabela A-3 – Sinais e conectores de vídeo.....	106
Tabela A-4 - Padrões primários de TV digital (HSM, 2009)	108

LISTA DE FIGURAS

Figura 1.1 – Prioridades nas organizações (Intellicom Analytics, 2007).....	2
Figura 1.2 – Exemplo de implementação de solução para telefonia	4
Figura 1.3 – Exemplos de dispositivos de áudio	5
Figura 1.4 – Esquemático de um terminal de videoconferência (Davis, 1999).....	6
Figura 1.5 – Equipamento para sala; sala de aula; webcam e headset; equipamento pessoal (Polycom, 2009) (FVG, 2009).....	7
Figura 1.6 – Tipos de videoconferências multiponto	8
Figura 2.1- Padrão H.241 e outros padrões guarda-chuva	13
Figura 2.2 – Espectro de fala para as palavras “sailing” e “failing” com cortes em 3.3 kHz e 22 kHz (Rodman, 2006).....	17
Figura 2.3 – Comparação entre resoluções de imagem.....	21
Figura 2.4 – Requisitos típicos de banda x à resolução de imagem (Polycom, 2006)	23
Figura 2.5 – Estrutura do frame H.221 (C21VIDEO, 2008, modificado).....	24
Figura 2.6 –Estrutura dos frames nos modos 6B e H0 (C21VIDEO, 2008)	25
Figura 2.7 – Pilha de protocolos e recomendações para terminais H.323 (ITU-T H.323, 1996).....	27
Figura 2.8 – Exemplos de terminais de videoconferência.....	28
Figura 2.9 – Zona H.323.....	28
Figura 2.10 – Exemplos de gatekeepers	29
Figura 2.11 – Protocolos e recomendações de um gatekeeper (ITU-T H.323, 1996).....	30
Figura 2.12 – Gateway para interconexão IP e ISDN (Guimarães, 2002)	31
Figura 2.13 – Conexões entre terminais e MCUs para conferências multiponto	33
Figura 3.1- Diagrama representando nós e estações (Stallings - 1997).....	38
Figura 3.2 – Arquitetura padronizada para o acesso BRI (Wilcox, 2000)	42
Figura 3.3 – IMUX (Polycom)	43
Figura 3.4 – Transição ISDN para IP (Weinstein, 2006)	44
Figura 3.5 – Modelo de camadas OSI (Stallings, 1997).....	46
Figura 3.6 – A divisão em 4 camadas das redes TCP/IP (Guimarães, 2002).....	46
Figura 3.7 – Utilização típica de banda para transmissão de dados (modificado - Polycom, 2006).....	50
Figura 3.8 – Utilização típica de banda para transmissão de vídeo+áudio (modificado - Polycom, 2006).....	51

Figura 3.9 – Ilustração de latência e jitter (modificado – Karapetkov, 1999).....	52
Figura 3.10 – Cabeçalho do protocolo IP (IETF RFC 791, 1981)	56
Figura 3.11 – Exemplo de localização da configuração dos parâmetros de IP Precedence	57
Figura 3.12 – Mensagens do RSVP (RNP, 2000)	59
Figura 3.13 – Campo TAG, 32 bits adicionados ao frame Ethernet (GTA/UFRJ, 2002) ...	62
Figura 3.14 – Configuração dos parâmetros de VLAN e CoS	63
Figura 3.15 – Sintaxe de um rótulo MPLS (IETF RFC 3032, 2001)	65
Figura 4.1 – Aplicações para comunicação rumo à comunicação colaborativa (Polycom, 2009).....	72
Figura 4.2 –Protocolos em direção à arquitetura unificada (Karapetkov, 2008).....	73
Figura 4.3 Integração entre plataformas usando MCUs e gateways	75
Figura 4.4 – Economia de custos com Gateways SIP/H.323 - ISDN.....	77
Figura 4.5 – Implementação utilizando um servidor de conferências.....	78
Figura 4.6 –Exemplos de sala de telepresença (Digitaltechnews , 2009) (Tandberg, 2009) (Polycom, 2009)	80
Figura 4.7 – Exemplo de sala de telepresença imersiva (Polycom, 2009)	81
Figura 4.8 – Crescimento da receita em equipamentos de vídeo (Davis, 2009)	85
Figura 4.9 – Receita de equipamentos de videoconferência de grupo (Davis, 2009)	86

LISTA DE SÍMBOLOS, NOMENCLATURA E ABREVIACÕES.

ATM	- Asynchronous Transfer Mode
BAS	- Bit-rate Allocation Signal
BGP	- Border Gateway Protocol
BRI	- Basic Rate Interface
CBR	- Constant Bit Rate
CIF	- Common Intermediate Format
CODEC	- Codificador/Decodificador
COS	- Class of Service
CSMA/CD	- Carrier Sense Multiple Access with Collision Detection
CSU	- Channel Service Unit
DPCM	- Differential Pulse Code Modulation
Ethernet	- Protocolo de rede local sinônimo do padrão IEEE 802.3
FAS	- Frame Alignment Signal
IEEE	- Institute of Electrical and Electronics Engineers
IETF	- Internet Engineering Task Force
IMUX	- Inverse Multiplexer
IP	- Internet Protocol
ISDN	- Integrated Services Digital Network
ITU	- International Telecommunication Union
LAN	- Local Area Network
LDP	- Label Distribution Protocol
LLC	- Logical Enlace Control
MAC	- Media Access Control
MC	- Multipoint Controller
MCU	- Multipoint Control Unit

MP	- Multipoint Processors
MPLS	- Multiprotocol Label Switching
NT1	- Network Terminal 1
NT2	- Network Terminal 2
NTSC	- National Television Standards Committee
OSI	- Open Systems Interconnection
PAL	- Phase Alternation Line
PBX	- Private Branch Exchange
PC	- Personal Computer
PCM	- Pulse Code Modulation
PRI	- Primary Rate Interface
PSTN	- Public Switched Telephone Network
QCIF	- Quarter CIF
QoS	- Quality of Service
RAS	- Registration Admission and Status
RSVP	- Resource reservation protocol
RTCP	- Real-time Transport Control Protocol
RTP	- Real-time Transport Protocol
SDP	- Session Description Protocol
SECAM	- Sequential Couleur Avec Memoire
SIMPLE	- Session Initiation Protocol for Instant Messaging and Presence Leveraging Extensions
SIP	- Session Initiation Protocol
SS7	- Signalling System 7
T.120	- Recomendação do ITU – T para Comunicação e Compartilhamento de Dados
TA	- Terminal Adapter

TCP	- Transport Control Protocol
TE1	- Terminal Equipment 1
TE2	- Terminal Equipment 2
TEI	- Terminal Equipment Identifier
UC	- Unified Communications
UDP	- User Datagram Protocol
VLAN	- Virtual LAN
VoIP	- Voice over IP
WAN	- Wide Area Network

1 - INTRODUÇÃO

Historiadores poderiam se referir à época entre 1997 e 2007 como a década em que as infraestruturas de telecomunicações se prepararam para novas aplicações se tornando redes confiáveis e acessíveis para os diversos tipos e tamanhos de organizações e aplicações (Polycom, 2009). Começando com um seletivo grupo de empresas globais em busca de redução de custos de operações através da união entre telefonia e infraestruturas de dados, observam-se cada vez mais companhias de portes distintos adotando as redes IP (*Internet Protocol*) para suportar aplicações flexíveis e com boa relação custo/benefício.

Um estudo da *IntelliCom Analytics* (Intellicom Analytics, 2007) questionando as prioridades de adoção de tecnologias corporativas foi aplicado em empresas de portes variados. O resultado, representado na Figura 1.1 mostra a evolução das medidas tratadas como prioridades durante o período de 1997 a 2007. As empresas estão gradualmente deixando de focar em infra-estruturas de comunicação em si para procurar aplicações unificadas de colaboração que possam aumentar a produtividade e eficiência de suas operações. Essas aplicações seriam capazes de aproveitar a tecnologia atualmente disponível nas telecomunicações, as quais avançaram de forma a ampliar as suas possibilidades de uso. Dentre essas aplicações, destaca-se o uso de áudio, vídeo e colaboração sobre a rede de dados uma vez que a velocidade na troca de informações dentro das empresas está se tornando um grande diferencial competitivo.



Figura 1.1 – Prioridades nas organizações (Intellicom Analytics, 2007)

Nos dias de hoje, muitas organizações são globais, interligadas, compreendendo empregados, fornecedores, parceiros e clientes de diversas localidades. Da forma como trabalham atualmente, times de projetos, por exemplo, são criados envolvendo pessoas de diversas partes da organização e, possivelmente, em localidades distintas. Esses times precisam de meios de comunicação eficientes capazes de potencializar a comunicação e o relacionamento entre pessoas. Tipicamente é necessário que a interação se dê através de meios de comunicações diversos. Para permitir essa interação e colaboração dinâmica e distribuída, empresas implementaram uma variedade de plataformas de comunicação.

Comunicações Unificadas (de *Unified Communications* - UC) e Comunicações Colaborativas Unificadas são termos que vêm sendo utilizados pelos mercados de telefonia, redes de comunicações e arquiteturas de TI (Tecnologia da Informação) para representar a convergência das redes de voz, vídeo e dados dentro das corporações. Esses conceitos representam uma plataforma de comunicações responsável por unificar uma série de soluções que foram implementadas de forma tecnologicamente independente, como por exemplo, a telefonia IP, videoconferência (salas, desktop, web), serviços de mensagens instantâneas e serviços de presença.

O *Yankee Group* (Kerravala, 2004) define UC como “uma experiência para o usuário envolvendo áudio, vídeo, web e colaboração de dados que ultrapassa distância, tempo e plataformas de comunicação, permitindo que pessoas se comuniquem em qualquer lugar, a qualquer hora através de qualquer meio de comunicação utilizando uma variedade de

dispositivos”. Implícitos nesta definição estão três pontos relevantes para aplicações de colaboração:

Disponibilidade imediata: As aplicações devem estar disponíveis sob demanda e suportar interação em tempo real para que seus usuários possam se encontrar quando precisarem e de onde estiverem. Para participantes que não puderem interagir em tempo real, a solução de colaboração deve disponibilizar todos os elementos envolvidos – áudio, vídeo e conteúdo – para posterior acesso e compartilhamento.

Simplicidade: Este é um ponto importante, pois está relacionado à experiência do usuário do sistema. Uma vez que muitas pessoas são relutantes em mudar os seus hábitos e forma de trabalhar, a solução deve ser amigável para que toda a empresa perceba seus benefícios e o retorno sobre o investimento de fato ocorra.

Interoperabilidade: As soluções de colaboração não envolvem somente um produto ou aplicação, mas sim uma série de soluções integradas e aplicações que incluem telefonia, conferências em voz e vídeo e compartilhamento de conteúdo. Podem ser incluídas também aplicações como conferência web, presença e mensagens instantâneas. O uso de protocolos abertos e homologados por organizações independentes facilita a integração destas soluções e também aumenta a gama de fabricantes e soluções que podem ser escolhidas. Do ponto de vista da usabilidade, a interoperabilidade permite que as soluções interajam entre si independentemente do ambiente onde são implementadas.

1.1 - APLICAÇÕES DE COMUNICAÇÕES COLABORATIVAS

As soluções mais visíveis em um ambiente de comunicação colaborativa são as aplicações de áudio (telefonia analógica e/ou IP) e vídeo. A infraestrutura que suporta essas soluções pode ser implementada por somente um provedor e fabricante ou pode ser mais complexa, podendo incluir a integração de diversas plataformas, sistemas diferentes de um mesmo fabricante e sistemas de diversos fabricantes. Em várias organizações as soluções de áudio e vídeo ainda são implementadas como soluções independentes.

1.1.1 - Áudio sobre IP e áudio conferências

As soluções de áudio ou telefonia para as organizações não se baseiam mais somente em equipamentos analógicos. Atualmente tornaram-se mais complexas e flexíveis permitindo o

uso das redes IP e também a integração entre telefones e centrais (PABX - *Private Automatic Branch Exchange*) do mundo analógico com a telefonia IP. A Figura 1.2 mostra o exemplo de uma arquitetura de solução de telefonia capaz de suportar estas duas plataformas.

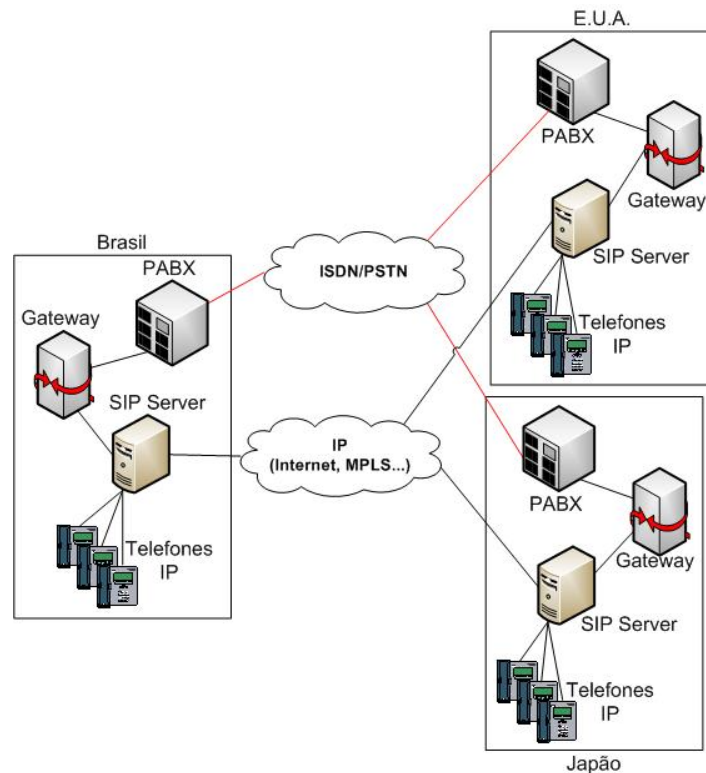


Figura 1.2 – Exemplo de implementação de solução para telefonia

São mostrados na Figura 1.2 telefones IP que se registram em servidores responsáveis por gerenciar estes terminais/ramais. Essas implementações são feitas utilizando o protocolo de sinalização SIP (*Session Initiation Protocol*), H.323 (ITU-T H.323, 1996) ou protocolos proprietários. Na Figura 1.2 um servidor SIP é mostrado como elemento central da rede de telefonia em IP. Através da rede local, os telefones IP podem se comunicar sem custo por chamada. Quando uma chamada precisa chegar à rede de telefonia tradicional, o custo da chamada está condicionado à operadora. Sendo assim, caso uma pessoa no Brasil deseje conversar com outra no Japão, uma opção viável seria trafegar essa chamada até o Japão pela rede IP e de lá iniciar uma chamada local utilizando a rede analógica local.

A telefonia IP oferece vantagens sobre a telefonia analógica como possibilidade de redução de custos, melhoria na qualidade de voz (relacionada agora à qualidade dos codificadores utilizados e taxa de transmissão) e maior quantidade de funcionalidades. Com o uso de gateways conectados às linhas telefônicas convencionais ou PABX é possível fazer

chamadas para telefones convencionais e também ter um melhor aproveitamento das linhas analógicas contratadas.

A Figura 1.3 mostra exemplos de equipamentos terminais utilizados em soluções de telefonia.



Figura 1.3 – Exemplos de dispositivos de áudio

1.1.2 - Videoconferência

A videoconferência (abreviatura do termo áudio-videoconferência), embora tratada como uma única tecnologia, compreende uma coleção de tecnologias que permitem a comunicação audiovisual bidirecional. Isto quer dizer que, pessoas em localidades distintas – seja em salas, cidades ou mesmo países diferentes – podem receber e transmitir áudio e vídeo simultaneamente (Wilcox, 2000).

Esta tecnologia já se tornou um meio bastante utilizado em grandes corporações como forma de treinamento à distância e como forma de diminuir custos e tempo relativos ao deslocamento de pessoas para reuniões importantes. Nas universidades é utilizada como ferramenta para educação à distância, seja em cursos de extensão, pós-graduação, mestrado ou doutorado. Na medicina, é usada como forma de transmitir cirurgias e palestras ao vivo para médicos e participantes de simpósios e congressos.

Um terminal de videoconferência é o elemento fundamental destas aplicações. O codec em um terminal é o coração da videoconferência, pois é o elemento responsável por codificar/decodificar e comprimir/descomprimir os sinais de áudio e vídeo para serem transportados pela rede. O processo implica em transformar os sinais elétricos de áudio e vídeo em zeros e uns (digitalização ou codificação) e, depois seleccionar os bits necessários

para enviar significativamente a mensagem de áudio e vídeo (compressão). A Figura 1.4 mostra o esquemático de um terminal de videoconferência.

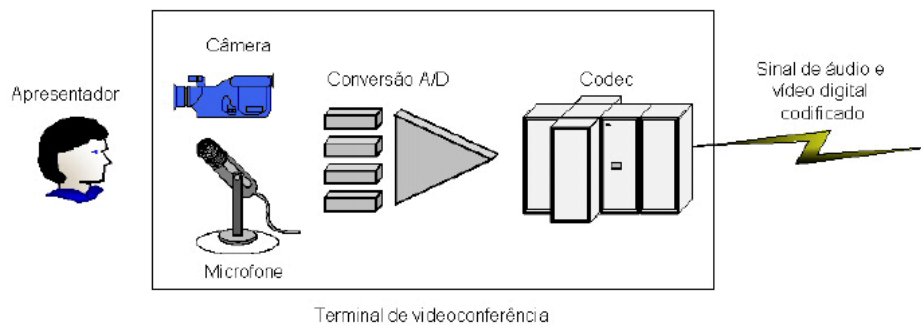


Figura 1.4 – Esquemático de um terminal de videoconferência (Davis, 1999)

Basicamente a videoconferência é utilizada em três tipos de situações genéricas: sala-sala, sala-PC e PC-PC. As situações definem o tipo preferencial de equipamentos a ser utilizado. A Figura 1.5 mostra um equipamento para sala, uma sala de videoconferência, webcam e *headset* para conferências no PC e também um equipamento de videoconferência pessoal.

- **Sala-sala:** como o termo sugere, é baseada em comunicações entre salas distintas. Estas salas são preparadas para fazer videoconferência, com equipamentos de grande porte e alta capacidade de codificação de áudio e vídeo, além de outros equipamentos auxiliares como, por exemplo, câmera de documentos, fontes de vídeo, quadro eletrônico, etc. (Brey, 2000)
- **Sala-PC:** é uma interação entre sala de videoconferência e videoconferência pessoal ou desktop. Neste caso, há um software de videoconferência instalado em um computador pessoal (PC), que irá utilizar recursos próprios para codificar áudio e vídeo. A captação de imagens é via câmera pessoal instalada neste mesmo computador (Brey, 2000);
- **PC-PC:** videoconferência realizada entre computadores pessoais, deixando de ser entre grupos de pessoas e passando a ser entre indivíduos (Brey, 2000).



Figura 1.5 – Equipamento para sala; sala de aula; webcam e *headset*; equipamento pessoal
(Polycom, 2009) (FVG, 2009)

Uma videoconferência pode se dar de duas maneiras: ponto-a-ponto (dois pontos distintos) ou multiponto (três ou mais pontos). No modo multiponto o conceito de sala virtual é aplicado. Todos os equipamentos estão conectados a uma sala virtual configurada de forma a se adequar ao objetivo ou tipo de chamada. Esta sala pode se concentrar em um equipamento ou ainda em uma cascata de equipamentos concentradores. A Figura 1.6 mostra um diagrama com exemplos de salas virtuais multiponto.

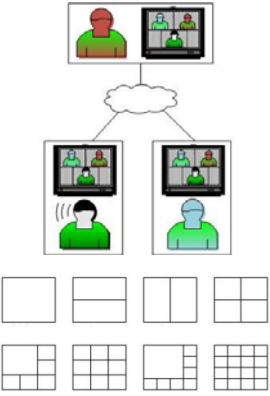
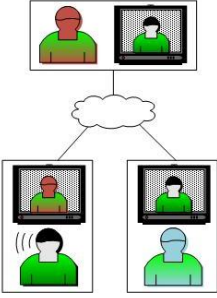
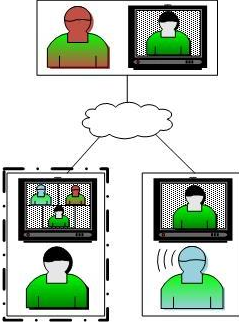
<p>Presença Contínua (CP)</p>		<p>Cada participante é visto em um quadrante dentro da tela.</p> <p>Se houver mais participantes do que divisões na tela, estes são vistos quando começam a falar.</p> <p>Pode ser configurado também para que um participante não se veja na tela ou para que todos vejam a mesma imagem.</p>
<p>Chaveado por voz (VSW)</p>		<p>Também conhecido como “discussão”. Ele permite que uma pessoa de um dos locais participantes, ao falar, seja vista pelos outros participantes da videoconferência multiponto.</p> <p>Assim, quando outra pessoa falar a imagem, automaticamente aparecerá na tela (Davis, 1999)</p>
<p>Travado no interlocutor</p>		<p>Também conhecido como “apresentação”. Através de um comando, ou pré configuração, é possível travar a imagem de qualquer local participante na videoconferência, com isso todos os demais verão apenas o local que foi travado e este, por sua vez, irá ver o local que estiver falando ou, se todos estiverem em silêncio, ficará comutando entre os participantes (Davis, 1999)</p>

Figura 1.6 – Tipos de videoconferências multiponto

1.1.3 - Mensagens e presença

Os serviços de mensagens instantâneas tornaram-se muito populares entre usuários espalhados pela Internet, sendo cada vez mais adotadas como ferramenta de trabalho dentro de organizações. Ferramentas como MSN Messenger¹, Skype² e Google Talk³ estão disponíveis gratuitamente para usuários domésticos. Organizações de todos os portes enxergam nessas soluções uma forma rápida de comunicação, logo fabricantes diversos oferecem soluções que podem ser gerenciadas pela própria organização. Além de mensagens, outras funcionalidades são tipicamente agregadas como compartilhamento de dados, transferência de arquivos e comunicação audiovisual.

O recurso de mensagens instantâneas permite que os usuários se comuniquem de seus computadores usando mensagens de texto. A operação básica desse recurso é a troca instantânea (chegam ao destino assim que o emissor decide enviá-las) de mensagens de texto entre dois usuários. Os usuários podem convidar outras pessoas para uma sessão de mensagens instantâneas. Quando mais um contato é adicionado a uma sessão de mensagens instantâneas, ela se transforma em conferência e o comportamento da janela de conversa muda para oferecer suporte aos recursos de conferência.

O recurso de presença fornece aos usuários informações sobre o estado de outros usuários em suas listas de contatos. O estado de presença de um usuário fornece informações que ajudam outros a decidir se devem contatar este usuário e se devem usar mensagens instantâneas, telefone, vídeo ou email. A presença incentiva a comunicação instantânea quando possível, também fornecendo informações sobre um usuário, como se está em uma reunião ou ausente do escritório, indicando que a comunicação instantânea não é possível.

1.2 - TECNOLOGIAS E AMBIENTES

Uma solução de comunicação não depende somente do equipamento utilizado para a codificação, mas também da interação entre várias tecnologias. Além do codec e dos processadores são necessárias câmeras, microfones, televisões, caixas de som, softwares de

¹ www.msn.com

² www.skype.com

³ www.google.com

gerenciamento, gravadores, infra-estrutura adequada de rede e, certamente, pessoas. Todos esses elementos experimentaram grandes avanços e mudanças nos últimos anos, inclusive as pessoas, uma vez que já se acostumaram com a rapidez proporcionada pela tecnologia e com novas formas de se comunicar e trabalhar.

A qualidade e usabilidade da comunicação em áudio e vídeo melhoraram muito de poucos anos para os dias atuais. Os telefones IP estão utilizando codificadores capazes de prover cada vez mais fidelidade de voz e o mesmo acontece com os codecs de vídeo, possibilitando chamadas em alta definição (Polycom, 2005). Em paralelo, as redes de telecomunicações passaram por avanços que permitem às empresas contratarem enlaces de dados cada vez mais rápidos, confiáveis e com mais opções de configuração.

Ao implementar uma solução envolvendo comunicações em áudio e vídeo, é necessário portanto observar dois aspectos gerais, a tecnologia e a experiência dos usuários. Para uma implementação efetiva em uma organização é necessário, desde o início do projeto, um engajamento das pessoas envolvidas com a infraestrutura de dados e telefonia. É importante também que os usuários dos sistemas também participem do levantamento de requisitos e funcionalidades. A experiência dos usuários é o fator que vai avaliar se houve sucesso na implementação de uma tecnologia.

A tecnologia de voz já está bem estabelecida no mundo todo e é utilizada tanto por corporações quanto usuários domésticos. O mesmo está acontecendo com a videoconferência, não tanto no mercado doméstico, mas bastante recorrente no meio corporativo. Os grandes fabricantes estão ganhando atenção dos diretores de grandes empresas, pois atualmente podem confiar na tecnologia para tomada de decisões.

No intuito de melhorar a experiência para usuários, um novo conceito emerge no mercado, a telepresença imersiva. As soluções tradicionais de videoconferência evoluíram na última década em direção ao uso das redes IP de forma confiável para reduzir custos e permitir maior escalabilidade. Evoluiu também de imagens de baixa resolução para alta definição, atingindo 720p e 1080p (Polycom, 2005) atualmente. Também os codecs de áudio saíram da banda estreita de 3 kHz, usada para codificar voz, em direção à qualidade de CD (22 kHz) estéreo reduzindo a sensação de cansaço nas conferências de voz. A solução de telepresença imersiva se utiliza de toda essa tecnologia disponível, combinada à outros elementos, para criar ambientes onde se pareça o máximo possível que todos os participantes estejam em uma mesma sala.

Nesta dissertação serão tratados tanto os aspectos tecnológicos quanto de ambiente para utilização das soluções de colaboração apresentadas acima.

1.3 - MOTIVAÇÃO

De poucos anos até os dias de hoje, as tecnologias relacionadas à videoconferência avançaram rapidamente. Desde sua introdução no comércio nos anos 70 quando essas soluções eram muito caras e funcionavam sobre redes com recursos muito limitados, a tecnologia de videoconferência se desenvolveu de forma que se tornasse uma solução essencial para proporcionar um meio de comunicação eficiente para muitas organizações.

Soluções de voz e dados evoluíram também para proporcionar mais qualidade e possibilidades de uso, sendo rapidamente adotadas por usuários domésticos e comerciais. Percebe-se essa evolução pela rápida adoção, por usuários domésticos, dos serviços de VoIP (voz sobre o protocolo IP) através de soluções como Skype¹ e MSN Messenger² e também pela grande variedade de importantes fornecedores que se desenvolveram nesse nicho de mercado oferecendo soluções para organizações de diversas escalas de tamanho.

A integração destas soluções proporciona uma nova variedade de aplicações que está sendo bastante visada pelas grandes e médias organizações. Fabricantes de diversos portes estão direcionando suas soluções para que se adéquem às exigências do mercado e também às capacidades proporcionadas pelos avanços na tecnologia.

Avaliando este contexto, esta dissertação é resultante de um estudo sobre as tendências do mercado de colaboração e também das tecnologias envolvidas, sendo estas, padrões de comunicação, infraestruturas de rede e novos padrões para melhorias de qualidade de áudio e vídeo. Tipicamente as informações aqui contidas são encontradas separadamente, fazendo deste trabalho um guia para que se entenda áudio e videoconferência e seus temas relacionados. O capítulo de tendências também apresenta contribuições mostrando como podem ser implementadas redes híbridas para UC e também avaliações sobre mudanças em alguns aspectos dessa tecnologia.

1.4 - OBJETIVOS

¹ www.skype.com

² www.msn.com

O presente trabalho tem como objetivo apresentar um estudo sobre as tecnologias relacionadas à comunicação colaborativa e mostrar a tendência evolutiva das aplicações destas tecnologias. Também objetiva ser uma referência para futuras pesquisas sobre o tema ou para implementações destas soluções.

1.5 - ORGANIZAÇÃO DO TRABALHO

O presente trabalho está organizado em cinco capítulos, incluindo esta introdução, onde o tema a ser destacado nesta dissertação é apresentado.

O capítulo 2 apresenta os padrões usados nas tecnologias para comunicações unificadas. O foco do trabalho está na comunicação multimídia envolvendo áudio e voz. São discutidos aspectos relacionados ao funcionamento e evoluções destas tecnologias e dos padrões utilizados.

O capítulo 3 trata do outro aspecto fundamental para o uso de comunicações unificadas, a infraestrutura de redes. Este aspecto evoluiu bastante de forma a se permitir que as aplicações fossem implementadas de forma confiável e com a qualidade adequada.

Após serem apresentadas as tecnologias e como esta evoluiu, no capítulo 4 são mostrados como as tecnologias são usadas atualmente e quais são suas tendências de uso em ambientes colaborativos. São ressaltados os aspectos e exemplos de infraestrutura e também ambientes para melhoria da experiência para os usuários.

A dissertação é concluída com o capítulo 5 onde é apresentada uma conclusão sobre este trabalho.

2 - PADRÕES PARA COLABORAÇÃO MULTIMÍDIA

Em termos simplificados, as padronizações são necessárias para garantir que equipamentos de fabricantes distintos possam se comunicar. Neste capítulo serão apresentados os principais padrões usados para possibilitar a comunicação colaborativa.

A ITU-T (*International Telecommunications Union - Telecommunication Standardization Sector*), uma agência intergovernamental que congrega mais de 700 organizações públicas e privadas de 192 países, é o órgão responsável por coordenar padronizações relacionadas à telecomunicações¹ (ITU-T, 2009). Dentre as famílias de padrões da ITU-T destacam-se, no contexto desta dissertação, as famílias H (*Audiovisual and multimedia systems*) e G (*Transmission systems and media, digital systems and networks*) as quais contêm os protocolos e padrões mais usados para as tecnologias de colaboração multimídia.

A recomendação H.241 (*Extended video procedures and control signals for H.300-series terminals*) (ITU-T H.241, 2005) é responsável por unificar a sinalização e funcionalidades utilizadas pelos protocolos de comunicação da série H.3XX, além de especificar protocolos para vídeo como o H.264 (ITU-T H.264, 2003) e o H.239 (ITU-T H.239, 2003), estes serão mostrados em maiores detalhes posteriormente nesta dissertação. O H.241, junto com as recomendações da série H.3XX são consideradas recomendações guarda-chuva, pois elas referenciam uma série de outros padrões que devem, ou podem, trabalhar em conjunto. Exemplos de padrões englobados pelo H.241 estão representados na Figura 2.1.

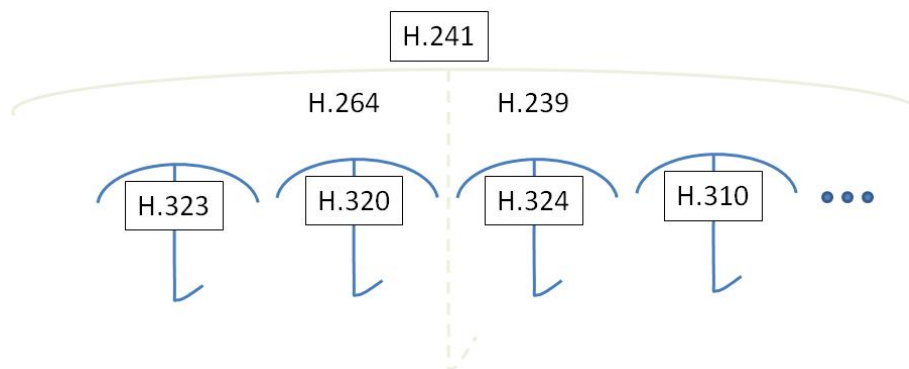


Figura 2.1- Padrão H.241 e outros padrões guarda-chuva

¹ www.itu.int/itu-t

Pode-se classificar os padrões ITU-T relacionados às comunicações multimídia (aplicação mais visível em UC) em três conjuntos principais:

- Padrões de Comunicação
- Padrões de áudio
- Padrões de vídeo

A Tabela 2-1 ressalta os principais padrões de comunicação (também referenciados como recomendações) para áudio e videoconferência e ainda os padrões que são recomendados por estes. Cada recomendação apresentada (H.320, H.323 e H.324) representa um conjunto de outros padrões que devem interagir de modos determinados pelas suas normas para possibilitar a comunicação audiovisual entre os dispositivos em cada tipo de plataforma correspondente (ISDN, IP e PSTN/POTS)¹.

Tabela 2-1 - Padrões de comunicação ITU-T

ITU-T			
	H.320	H.323	H.324
Data de Aprovação	1990	1996	1996
Redes de Acesso	ISDN	IP	PSTN POTS
Vídeo	H.261 (H.263, H.264)	H.261 H.263 (H.264)	H.261 H.263 (H.264)
Áudio	G.711 (G.722, G.728, G.723.1, G.729,G.722.1)	G.711 (G.722, G.728, G.723.1, G.729,G.722.1)	G.723.1
Sinalização/control	H.221 / H.242	H.225 / H.245	H.223 / H.245
Dados	T.120* H.239	T.120* H.239	T.120*

- O T.120 foi especificado inicialmente, mas atualmente está em desuso

Os protocolos de sinalização e controle são os responsáveis pelo estabelecimento e fechamento de conexões, negociação entre os elementos e controle das chamadas. O mecanismo dos protocolos de sinalização está diretamente relacionado com a plataforma usada para a comunicação (H.221 (ITU-T H.221, 1988) para ISDN, H.225 (ITU-T H.225, 2006) para IP e H.223 (ITU-T H.223, 1996) para PSTN). Os protocolos de controle atuam depois da etapa inicial de sinalização e negocia as capacidades e atributos de uma chamada,

¹ ISDN - *Integrated Services Digital Network*, PSTN – *Public Switched Telephone Networks*, POTS - *Plain old telephone service*

além de fazer o controle de uma chamada em andamento (H.242 (ITU-T H.242, 1990) para ISDN, H. 245 (ITU-T H.245, 1996) para IP e PSTN).

Um dos passos da negociação é a definição de quais são os padrões de codificação (codecs) de áudio e vídeo a serem utilizados. Cada elemento da comunicação precisa ser compatível com os codecs obrigatórios e podem trabalhar também com codecs dados como opcionais pela norma, sendo essa negociação feita pelos seus protocolos de controle. A decisão pela capacidade de suportar quais codecs é de responsabilidade do fabricante dos equipamentos estando muitas vezes relacionada à aplicação e capacidade de processamento de cada equipamento.

2.1 - CODECS DE ÁUDIO

Os codecs de áudio para uso em telecomunicações sofrem limitações mais severas que codecs utilizados em aplicações gerais de multimídia. Esse fato decorre da necessidade de se criar algoritmos padronizados, interoperáveis, capazes de prover áudio de boa qualidade e com pouca latência. Além disso, deve exigir pouco processamento computacional e memória para facilitar a incorporação em dispositivos de comunicação que variam desde equipamentos extremamente portáteis, equipamentos de baixo custo até equipamentos de alta qualidade como os sistemas para salas de telepresença.

Os primeiros codecs de áudio para telecomunicações foram desenvolvidos para utilizar a faixa do espectro de áudio correspondente à voz humana. Isso decorre do fato de as primeiras aplicações para estes codecs terem o propósito de trafegar somente voz, não levando em consideração a faixa de espectro necessária para se transmitir sons mais complexos e ricos como uma música, por exemplo. No entanto, com o avanço da demanda das telecomunicações por soluções multimídia, a ITU-T identificou a necessidade de padrões capazes de suportar mais qualidade e possibilidades.

Também com a migração da telefonia corporativa para o VoIP, existe um movimento em busca de uma melhor qualidade de áudio/voz. Uma vez que se trafega áudio de forma digital, a qualidade de voz fica à cargo da implementação de codificadores mais eficientes, não ficando restrita somente à capacidade de banda do canal de comunicação. A qualidade dos codificadores de áudio permitiu que o limite da banda restrita de 3 kHz, proveniente da telefonia analógica fosse ultrapassado para atingir aplicações de alta definição. Estão

disponíveis atualmente codecs capazes de prover uma qualidade de até 22 kHz. A Tabela 2-2 mostra um resumo dos codecs de áudio mais utilizados e suportados pelas recomendações H.320 (G.711, G.722, G.728) e H.323 (G.711, G.723.1, G.728, G.729, G.722) vistas adiante.

Tabela 2-2 – Principais codecs de áudio

Recomendação/ padrão	Algoritmo	Taxa de bits (kbps)	Largura de banda (Hz)	Delay (ms)	Aplicação
G.711	PCM	48, 56, 64	300-3,4K	<1	Telefonia convencional
G.722	Dual Band- ADPCM	48,56,64	50 – 7K	<2	Banda larga de voz. VoIP e videoconferência tradicional
G.722.1	MLT	24, 32	50 – 7K	<2	Muito usado em videoconferência e VoIP devido à larga faixa de frequência e baixo uso de banda
G.723.1	ACELP (5.3) MPC-MLQ (6.3)	5.3, 6.3	300-3,4K	67-97	Tipicamente usado em aplicações VoIP e videoconferências com o H.324
G.728	LD-CELP	16	300-3,4K	<2	Usado em ambientes com restrição de banda
G.729	CS-ACELP	8	300-3,4K	<1	Tipicamente usado em aplicações VoIP devido ao baixo consume de banda

Para poder comparar os diferentes codecs disponíveis alguns pontos importantes devem ser levados em consideração, dentre eles a largura de banda, a taxa de transmissão, o algoritmo usado e a latência.

A largura de banda utilizada pelo codec está fortemente lidada à fidelidade do áudio que se obtém. Sendo esta, a capacidade que o codec tem de codificar sons que vão desde frequências muito graves como a nota mais grave de um violão celo até sons bem agudos como um chimbau. Apesar de maior parte dos sons emitidos pela voz humana estarem em frequências de até 3Khz informações importantes também se encontram acima dos 14 kHz. O ouvido humano, por exemplo, consegue perceber frequências entre 20 Hz e 20Khz.

A captação de uma ampla largura de banda é o que permite a identificação de fonemas semelhantes, tornando a experiência de áudio mais próxima da realidade. Diferenças entre consoantes como “s” e “f”, “p” e “b” são mais perceptíveis nas frequências de 7Khz a 14 kHz (Rodman, 2006). A Figura 2.2 é interessante para representar este aspecto, pois mostra a diferença entre duas palavras da língua inglesa “*sailing*” e “*failing*” e demonstra a importância das frequências mais altas para identificá-las. Na língua portuguesa este mesmo efeito pode ser percebido em palavras como faca e vaca ou sapo e fato.

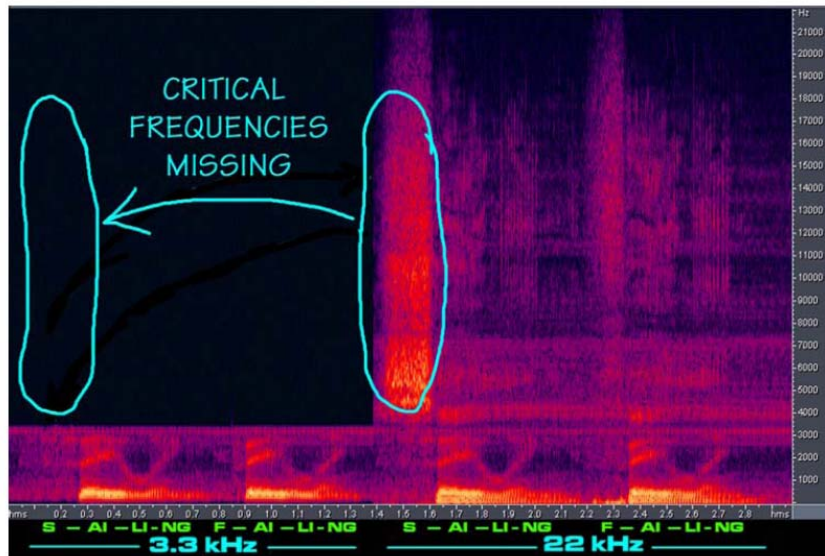


Figura 2.2 – Espectro de fala para as palavras “sailing” e “failing” com cortes em 3.3 kHz e 22 kHz (Rodman, 2006)

A taxa de transmissão é elemento fundamental na comunicação quando a banda (termo agora referenciando a taxa de transmissão em kbps) de um canal de comunicação é limitada. A Tabela 2-3 sumariza uma comparação entre alguns codecs, sua banda em frequência e taxa de transmissão. As taxas de transmissão não necessariamente aumentam com o aumento da faixa de frequência de áudio captada pelo codec. Essa taxa está mais relacionada ao codec escolhido e como este é projetado.

Isso ocorre porque grande parte das informações de áudio está contida em baixas frequências, logo há menos informação a ser codificada em frequências mais altas. Ao mesmo tempo o ouvido humano é menos sensível a imprecisões em frequências mais altas. Levando-se em consideração estes dois fatores, por exemplo, um algoritmo pode ser mais preciso do que outro, o que implica em uma maior ou menor taxa de transmissão.

Tabela 2-3 – Exemplos de relações entre banda de áudio e taxas de bits

BW (kHz)	Taxa típica (kbps)
3.3	8 (G.729), 56 (G.711)
7	10 (G.722.2), 24 (G.722.1), 64 (G.722)
14	32 (G.722.1C)
20	32 (G.719), 64 (AAC-LD)
22	32 (Siren22*)

*Proprietário Polycom Inc.

Outro ponto a ressaltar é que existe uma variedade de taxas de transmissão entre os codecs. Por exemplo, para o G.722 são necessários 64 Kbps para codificar 7 kHz. Essa mesma banda de 7 kHz exige 10 Kbps no G.722.2. Este fato depende das suposições de cada codec ao transformar as informações de sons em bits. O algoritmo ACELP (*Algebraic code excited linear prediction*) do G.722.2 é projetado para voz, então sua abordagem na codificação é diferente do G.722 que pode ser usado tanto para voz, múltiplas vozes ou música.

A latência representa o atraso entre o momento em que uma pessoa fala até que a outra pessoa ouça. Quando essa latência é muito alta, a conversação fica mais difícil e “atravessada”. A latência considera o tempo de processamento do codificador, o tempo de transmissão dos dados e o tempo de decodificação. Logo, codecs mais avançados exigem maior processamento para que o atraso na codificação seja o menor possível, contribuindo assim para um menor valor da latência total.

2.2 - CODECS DE VÍDEO

O processo de digitalização de um sinal de vídeo compreende a conversão de um sinal analógico de vídeo, gerado por um padrão de televisão (PAL¹, NTSC², SECAN³) em um

¹ PAL (*Phase Alternating Line*): Padrão de televisão adotado da Europa com taxa de amostragem de 50Hz e 625 linhas. O PAL possui vários padrões derivados, tal como PAL-M utilizado no Brasil, cuja frequência de varredura é de 60Hz

² NTSC (*National Television Standard Committee*): Padrão de televisão adotado nos E.U.A., frequência de varredura vertical

³ SECAM (*Sequential Couleur Avec Memoire*): Padrão de Televisão adotado na França e em vários países do leste europeu

fluxo digital de vídeo. A digitalização de um sinal puro de vídeo gera um fluxo digital com uma elevada taxa de bits por segundo, sendo necessária para a sua transmissão uma largura de banda extremamente alta. Para uma digitalização de um fluxo de vídeo NTSC, por exemplo, teria-se o cenário seguinte:

$$\begin{aligned} & \textit{Fluxo Digital Gerado (bps)} \\ & = \textit{Resolução da tela (pixels)} \times \textit{Gradação de cores} \\ & \times \textit{Quadros por Segundo} \end{aligned}$$

Que em termos práticos seria:

Resolução NTSC: 720 pixels x 480 pixels

Número de quadros por segundo: 30 qps

Número de gradação de cores: 16 milhões (24 Bits/Pixel)

Fluxo = 720 x 480 x 24 bits x 30 qps = 248 Mbps

Um fluxo digital dessa ordem (248 Mbps) é inviável para transmissão, logo tornam-se necessários os algoritmos de compressão de imagens. A compressão do vídeo envolve basicamente duas técnicas, sendo estas a compressão espacial (*intraframe*) e a compressão temporal (*interframe*).

A compressão espacial ou *intraframe* (intraquadros) tem como referência a imagem do próprio quadro para a realização da operação de compressão. Na compressão temporal ou *interframe* (interquadros) são estabelecidos frames de referência, em intervalos determinados de tempo, que são utilizados como parâmetros de comparação para envio de dados referentes aos frames subsequentes. Essa técnica de compressão permite que somente a variação entre frames em sequência seja enviada para o receptor.

A compressão *interframe* é a mais eficiente para transmissão de vídeo. Esta permite taxas de compressão altas uma vez apenas somente a diferença entre os frames é armazenada e não toda a informação de cada um deles. Esta técnica leva em consideração a presença de informações redundantes transmitidas a cada quadro e atua justamente no sentido de se reduzir o fluxo de dados.

Os principais padrões ITU-T abertos para codificação de vídeo usados em colaboração ou videoconferência são o H.261, o H.263 e o H.264.

O padrão H.261 (*Video Codec for Audiovisual Services at P x 64 kbps*) (ITU-T H.261, 1988) tem como objetivo permitir a interoperabilidade dos sistemas de videoconferência com os três padrões de vídeo existentes – NTSC (Americano), PAL (Europeu) e SECAM (Francês). O padrão foi desenvolvido pelo ITU-T em 1988 e foi idealizado para proporcionar aplicações de videoconferência e videotelefonia sobre linhas digitais ISDN, por essa razão o padrão baseia-se em múltiplos de 64 kbps (os canais utilizados no ISDN são sempre múltiplos de 64 Kbps).

O H.261 definiu inicialmente duas estruturas de imagens compatíveis com os três padrões de transmissão televisiva. Os formatos definidos foram o CIF (*Common Intermediate Format*) e o QCIF (*Quarter CIF*). O primeiro possui 352 pixels por linha e 288 linhas por quadro. O segundo possui resolução de 176 pixels por linha e 144 linhas por quadro. A escolha de um ou outro formato depende da capacidade do canal de transmissão e das capacidades suportadas pelos equipamentos utilizados.

A recomendação H.263 (ITU-T H.263, 1996) é uma extensão da H.261 e utiliza como método de codificação o PDCM/DCT (*Differential Pulse Code Modulation/Discrete Cosine Transform*). Foi projetado para comunicação em baixas taxas de bits. Em versões preliminares (*drafts*) se falava em taxas menores que 64 Kbps, entretanto esta limitação foi ultrapassada e atualmente o H.263 substituiu o H.261 em diversas aplicações.

Os sistemas que utilizam H.263, em geral, têm melhor desempenho com relação ao uso de banda, que outros utilizando o H.261, chegando a uma proporção de 2.5 para 1, significando que quando ajustado para produzir a mesma qualidade de imagem a taxa de bits H.261 é 2.5 vezes maior que a taxa de bits gerada pela H.263.

A recomendação H.263 suporta cinco formatos. Somando-se ao QCIF e CIF suportados na H.261, tem-se, ainda, o SQCIF, o 4CIF e o 16CIF (ITU-T H.263, 1996). O SQCIF é, aproximadamente, metade da resolução do QCIF. Os formatos 4CIF e 16CIF são 4 e 16 vezes a resolução do CIF, respectivamente. O suporte ao 4CIF e 16CIF significa que a H.263 pode competir com outros codecs de vídeo com maior qualidade e em altas taxas de transmissão como o MPEG2¹. A Tabela 2-4 mostra os formatos de quadros suportados pelas

¹ Moving Picture Experts Group 2 - Padrão de codificação para vídeo digital e áudio associado desenvolvido pelo grupo MPEG, este fundado pela ISO – *International Organization for Standardization*.

recomendações H.263 e H.261 e a Figura 2.3 mostra uma comparação proporcional entre os diferentes formatos.

Tabela 2-4 – Formatos de vídeo suportados pela H.263 e H.261 (ITU-T H.263, 1996 modificado)

Formato	Resolução PAL	H.261	H.263
SQCIF	128 x 96		Sim
QCIF	176 x 144	Sim	Sim
CIF	352 x 288	Opcional	Sim
4CIF	704 x 576		Opcional
16CIF	1408 x 1152		Opcional



Figura 2.3 – Comparação entre resoluções de imagem

O padrão mais recente utilizado para videoconferência é o H.264. Este é um padrão de compressão de vídeo relativamente novo (2003), desenvolvido em esforço conjunto pelo MPEG (*Motion Picture Experts Group*), que é um grupo de estudos pertencente à ISO (*International Standards Organization*)¹ e o VCEG (*Video Coding Experts Group*)², grupo de estudos pertencente à ITU.

¹ www.iso.org

² www.itu.int/net/itu-t/info/

O H.264 permite taxas de compressão bem maiores do que já se conseguiu com os padrões anteriores, permitindo a compressão de vídeo com ou sem entrelaçamento de forma bastante eficiente. Mesmo quando utiliza altas taxas de compressão oferece ainda uma qualidade visual melhor do que os outros dois padrões. Além disso, esse padrão viabiliza que a codificação seja feita de forma mais flexível e que se organizem os dados codificados de forma que a taxa de erros e as perdas de dados sejam potencialmente minimizadas.

Devido à qualidade obtida com o H.264 para baixas taxas de transmissão, o ITU-T recomenda o H.264 para videoconferências em alta definição (Polycom, 2005). A Tabela 2-5 mostra as resoluções introduzidas pelo H.264 para *Standard Definition* (SD) e *High Definition* (HD). Apesar de já existir no codec H.263 a resolução 16 CIF, considerada alta definição, a capacidade de processamento e banda necessária na época em que foi introduzido impossibilitava a sua utilização em larga escala.

Tabela 2-5 – Formatos SD e HD introduzidos com o H.264

Formato	Frames por segundo	Resolução (16:9)	SD/HD
1080p	24,30	1920 x 1080	HD
720p	24,30,60	1280 x 720	HD
480p	24,30,60	860 x 480	SD

Uma maior quantidade de pixels resulta em melhor fidelidade de vídeo e melhor definição de imagens. Maiores taxas de frames por segundo resultam em uma melhor percepção de movimento tornando a imagem mais real. As letras “p” ou “i” colocadas após a denominação das resoluções têm significados específicos. O “i” representa formato entrelaçado (*interlaced*), ou seja, as imagens na telas são “desenhadas” em dois momentos: primeiro as linhas ímpares são mostradas, depois as linhas pares são incluídas para compor a imagem completa. Cada frame então é composto por duas varreduras completas na tela de projeção.

O modo “p” representa formato progressivo (*progressive*) onde a cada varredura da tela todas as linhas são mostradas para compor um frame completo. Este formato proporciona um incremento na qualidade da imagem com relação ao modo entrelaçado. Em imagens com muito movimento, as linhas do modo entrelaçado podem ficar muito aparentes uma vez que o

frame é composto em dois tempos. No modo progressivo obtêm-se uma imagem com linhas menos aparentes e contornos mais bem definidos.

A videoconferência em alta definição demanda que todos os equipamentos envolvidos sejam compatíveis com HD caso se queria transmitir e receber nestas resoluções. Apesar de exigir maior processamento, grande parte dos terminais e bridges de vídeo já são fabricados para trabalhar também com o H.264. Além dos requisitos de hardware, é necessário por volta de 1 Mbps de taxa de transmissão no mínimo para se trabalhar em HD (720p), sendo recomendado trabalhar com pelo menos 2 Mbps, permitindo também incluir codec de áudio com mais qualidade. A Figura 2.4 mostra os requisitos típicos de banda de acordo com a resolução desejada.

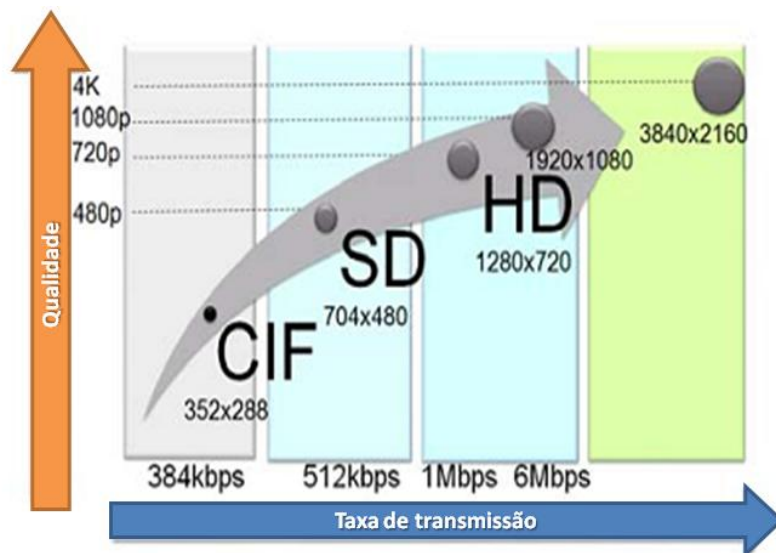


Figura 2.4 – Requisitos típicos de banda x à resolução de imagem (Polycom, 2006)

2.3 - PADRÃO DE COMUNICAÇÃO H.320

H.320 (ITU-T H.320, 1990) é a referência dada à recomendação publicada pelo ITU – T em 1990 para videoconferência e teleconferência em taxas de transmissão que vão de 56 Kbps até 1930 Kbps, sendo definida como “*Narrow-band visual telephone systems and terminal equipment*”. O padrão foi desenvolvido para operar usando redes digitais comutadas, particularmente, usando serviços comutados por circuitos (ISDN, SLDD, *switched 56* e *switched 64*).

A série de padrões englobados pela recomendação H.320 dita os conceitos básicos para videoconferência, especificando as necessidades de processamento de áudio e vídeo. Fornece

formatos comuns para entradas e saídas de sinais de áudio e vídeo e define protocolos que permitem a um terminal de videoconferência usar a rede de comunicação de dados comutada por circuitos para sincronização dos sinais com outro terminal. Os protocolos e padrões especificados pelo H.320 estão mostrados na Tabela 2-1.

O protocolo de sinalização H.221, especificando pelo H.320 tem como propósito definir a estrutura de quadros para serviços de videoconferência usando um ou vários canais de comunicação B (ITU-T H.221, 1988) Os canais B são definidos pelo padrão de comunicações ISDN, detalhado posteriormente nesta dissertação. Este canal é formado por um enlace de 64 kbps que pode ser combinado com outros canais B para possibilitar comunicações em velocidades até 2.048 kbps.

A Figura 2.5 representa o frame descrito pelo protocolo H.221. Cada frame contém 80 octetos que em um canal B são transmitidos um a cada 125µs (100 frames por segundo). Multiplicando esses valores têm-se a velocidade de 64 kbps de um canal B (100 frames por segundo * 80 octetos * 8 bits). Cada posição de bit (1 a 8) pode ser considerado um canal de 8 kbps. O oitavo bit em cada octeto faz parte do oitavo sub-canal, denominado Canal de Serviço (*Service Channel - SC*). Os primeiro 16 bits do sub-canal SC são utilizados para controle de frame e informações sendo os 8 primeiro bits denominados FAS (*Frame Alignment Signal*) e os outros 8 BAS (*Bit Alignment Signal*).

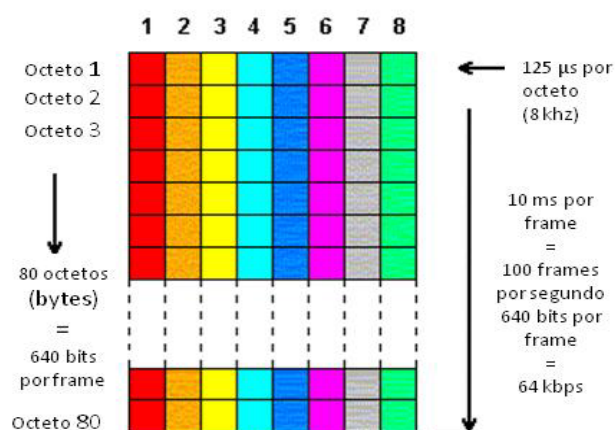


Figura 2.5 – Estrutura do frame H.221 (C21VIDEO, 2008, modificado)

Existem dois modos de funcionamento para conexões que utilizam múltiplos canais B. Por exemplo, em uma conexão tradicional de 384 kbps são utilizados 6 canais B do ISDN

(6*64 kbps). Neste caso os modos possíveis são o 6B ou o modo H0 - 384K *Bonded* (mais utilizado), ambos representados na Figura 2.6.

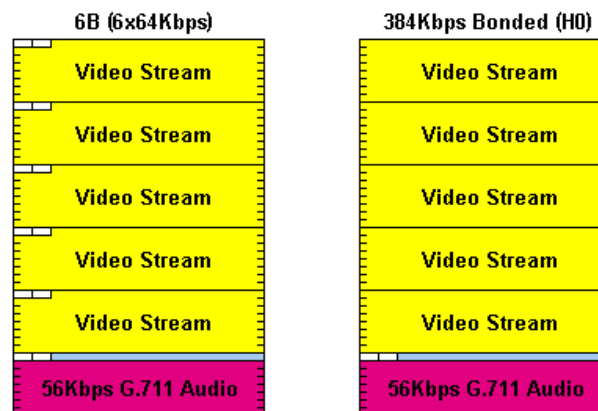


Figura 2.6 –Estrutura dos frames nos modos 6B e H0 (C21VIDEO, 2008)

No modo H0, o canal de 384 kbps é dividido em 6 *Time Slots* – TS, sendo que o primeiro TS contém as informações BAS e FAS, como qualquer outro canal B, e os canais subsequentes não possuem essas informações, carregando somente dados. O oitavo sub-canal de 8 kbps do primeiro TS (TS1) é usado para trafegar BAS e FAS (1,6 kbps) e também dados (6,4 kbps). Os outros sub-canais do TS1 são utilizados para trafegar até 56 kbps de áudio¹.

No modo 6B, todos *Time Slots* carregam as informações BAS e FAS restando os mesmos 56 kbps para áudio e 6,4 kbps para dados no primeiro TS. Somando os outros TSs tem-se 312 kbps para vídeo ($5 \cdot 64000 - 5 \cdot 16 \cdot 100$).

A velocidade de 384 kbps foi usada somente como exemplo. Ambos os modos de funcionamento podem trafegar chamadas em até 2.048 kbps usando 30 canais B.

A recomendação H.242 define todos os processos de sincronização (*handshake*) que estão envolvidos na fase que precede ao estabelecimento efetivo da chamada (ITU-T H.242 -1990). Os principais aspectos envolvidos nesta recomendação são:

- Sequências básicas para procedimentos de in-channel;
- Aspectos da rede como: conectar uma chamada, desconectar uma chamada e transferências de chamadas;
- Procedimentos para ativação e desativação de canais de dados;

¹ As informações de áudio só são trafegadas pelo primeiro *Time Slot*.

- Procedimentos para funcionamento de terminais em redes restritas (canais de 56 kbps).

As chamadas entre terminais H.320 são feitas utilizando números fornecidos pelos provedores do serviço contratado. É possível ter um número para cada canal B e também um número para referenciar um grupo de canais. Como a rede ISDN funciona utilizando uma rede comutada por circuitos baseada na infraestrutura de telefonia existente, cada canal B ocupa uma linha telefônica, logo ao se fazer uma chamada de videoconferência a 128 kbps o usuário deve pagar por duas chamadas de telefone, conseqüentemente uma videoconferência internacional à 384 kbps custaria e equivalente a 6 chamadas internacionais.

2.4 - PADRÃO DE COMUNICAÇÃO H.323

A recomendação H.323 (ITU-T H.323, 1996), assim como a recomendação H.320, compreende um arcabouço de normas e especificações de recomendações ITU-T para “Sistemas Audiovisuais e Multimídia”. A recomendação H.323 tem o objetivo de especificar sistemas de comunicação multimídia em redes baseadas em pacotes e que não provêm uma qualidade de serviço garantida. Além disso, estabelece padrões para codificação e decodificação de fluxos de dados de áudio e vídeo, garantindo que produtos baseados no padrão H.323 de um fabricante interoperem com produtos H.323 de outros fabricantes.

Redes baseadas em pacotes incluem as redes IP como a Internet, redes IPX (*Internet Packet Exchange*), as redes metropolitanas, as redes de longa distância (WAN) e ainda conexões discadas usando PPP (*Point-to-Point Protocol*).

A norma ITU-T H.323, além de referenciar os protocolos associados à comunicação, também descreve todo um sistema e seus componentes envolvidos. A seguir são descritos diferentes elementos definidos pela recomendação H.323.

2.4.1 - Componentes da Arquitetura H.323

2.4.1.1 - Terminais

Os terminais H.323, conhecidos como “codecs”, são os elementos fundamentais em qualquer sistema H.323 uma vez que são os equipamentos com os quais os usuários do sistema têm contato direto. São os dispositivos usados para fazer e receber chamadas

enviando e recebendo áudio e vídeo. Estes terminais podem tanto ser simples, como um telefone IP com funcionalidades básicas, quanto mais complexos, como um sistema de videoconferência de alta definição. Outras opções são softwares que podem ser instalados em PCs para realizarem as funções de um terminal.

Todos terminais H.323 contêm a pilha de protocolos que implementa as funcionalidades definidas pelo sistema devendo suportar obrigatoriamente (Brey, 2000):

- G.711 para comunicação de áudio/voz. Vídeo e dados são opcionais;
- H.245 para negociação das capacidades e do uso dos canais de comunicação;
- H.225 para sinalização e setup de uma chamada;
- RAS (*Registration Admission Status*) para comunicação com entre os terminais e o *gatekeeper*, realizando funções de Registro, Admissão e Status
- RTP (*Real Time Protocol*) e RTCP (*Real Time Control Protocol*) para seqüenciamento de pacotes de áudio e vídeo.

Outros padrões para áudio, vídeo e dados são opcionais, mas normalmente são implementados nas soluções disponíveis no mercado. A Figura 2.7 mostra a arquitetura de um terminal com os protocolos e recomendações suportadas. Na Figura 2.8 são apresentados exemplos de terminais de videoconferência.

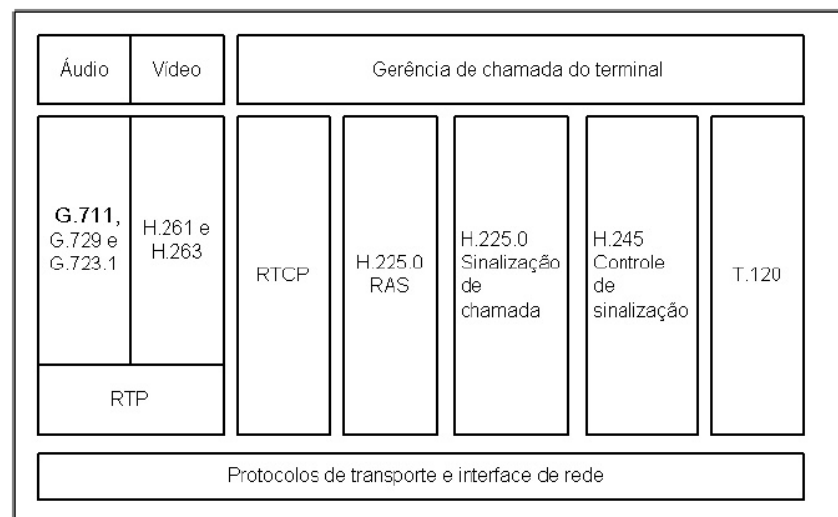


Figura 2.7 – Pilha de protocolos e recomendações para terminais H.323 (ITU-T H.323, 1996)



Figura 2.8 – Exemplos de terminais de videoconferência

Gatekeeper

Em uma rede de videoconferência o *gatekeeper* é de fundamental importância, sendo considerado o cérebro da rede H.323. Enquanto o H.323 permite que terminais conectem e controlem suas próprias chamadas usando endereços IP, um *gatekeeper* é necessário para a construção de uma rede robusta de videoconferência (VIDE, 2004). Uma vez implementado em uma rede, estabelece uma zona. Todos os elementos desta zona devem fazer uso dos seus serviços.

Uma zona H.323 é o conjunto de todos os terminais, gateways e MCUs gerenciados por um único *gatekeeper*, como mostra a Figura 2.9. Uma zona é determinada logicamente e pode ser independente da topologia da rede, podendo ter múltiplos segmentos que são conectados por roteadores ou outros dispositivos (ITU-T H.323, 1996).

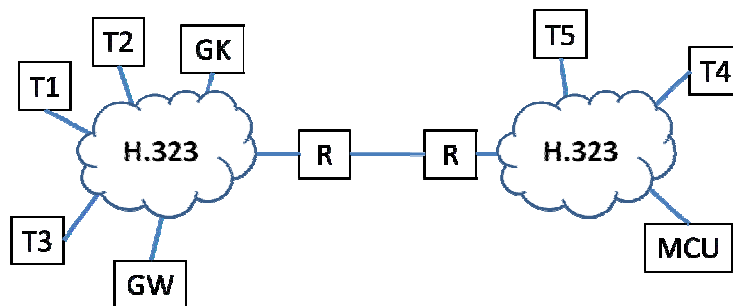


Figura 2.9 – Zona H.323

Um *gatekeeper* é geralmente um software instalado em um computador dentro de uma rede H.323. Atualmente pode-se encontrar *gatekeepers* no mercado de três maneiras: implementações em software para serem instaladas pelo cliente, *appliances* especializados que já possuem o hardware adequado para suportar a aplicação, ou em diversos casos o *gatekeeper* é implementado junto de uma MCU (*Multipoint Controler Unit*), gateway ou software de agendamento. Exemplos de *gatekeepers* são mostrados na Figura 2.10.



Figura 2.10 – Exemplos de *gatekeepers*

As funções principais de um *gatekeeper* são a resolução de endereços, controle de admissão e negociação de largura de banda.

- O processamento de uma chamada de videoconferência, da mesma forma que em uma ligação telefônica, requer um número de endereçamento para ser iniciado. No caso de uma ligação telefônica, utiliza-se um endereçamento internacional definido pela norma ITU-T E.164 (ITU-T E.164, 1988). Para as redes H.323, o número de endereçamento corresponde ao endereço IP dos dispositivos (terminais, gateways, MCUs, etc). A funcionalidade de resolução de nome do *gatekeeper* permite que usuários do sistema utilizem identificações mais amigáveis, associando o endereço IP a apelidos (alias). Estes apelidos podem ser números, letras ou mesmo um endereço de email (H.323 Anexo O).
- A função de controle de admissão é responsável por permitir que somente equipamentos autorizados utilizem os serviços disponíveis na zona. Este processo de admissão se baseia em políticas configuradas no *gatekeeper* que podem variar a critério dos fabricantes.
- O *gatekeeper* pode aceitar ou rejeitar chamadas baseando-se no total de banda disponível ou na quantidade de chamadas simultâneas na rede. Também é possível que durante uma chamada, ocorra a necessidade de se alterar a largura de banda negociada inicialmente. Esse controle pode ser usado para limitar a largura de banda de chamadas individuais, do sistema de videoconferência como um todo e também entre localidades específicas. Dessa forma evita-se que o tráfego de vídeo

sobrecarregue uma rede compartilhada ou que chamadas sejam realizadas com qualidade inferior à esperada.

Além das funções principais do *gatekeeper*, existem funções opcionais como autorização de chamadas de acordo com critérios diversos, controle da sinalização das chamadas, direcionamento de chamadas, modificação de alias e outros. Muitos fabricantes implementam com o *gatekeeper*, funções que são específicas para os seus equipamentos dentre essas funções pode-se citar gerenciamento de versões, agendamento de reuniões, funcionalidades de PABX, monitoramento de chamadas e até mesmo monitoramento da bateria dos controles remotos dos equipamentos. Os protocolos e recomendações de um *gatekeeper* são mostrados na figura Figura 2.11.

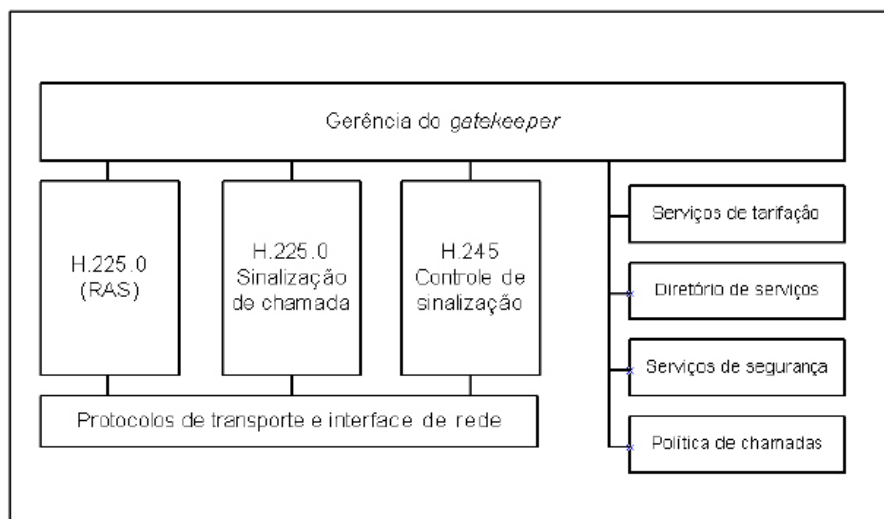


Figura 2.11 – Protocolos e recomendações de um *gatekeeper* (ITU-T H.323, 1996)

2.4.1.2 - Gateways

A função de um gateway H.323 é prover recursos para permitir que uma conferência multimídia seja realizada entre equipamentos aderentes a outros padrões (H.320, H.324, SIP, etc) como mostra a Figura 2.12. Essa habilidade provê não somente a ponte entre tecnologias diferentes e presentes no mercado, como também entre tecnologias que mudaram no decorrer do tempo, dessa forma estende-se o tempo de vida de tecnologias baseadas em protocolos mais antigos. Atualmente este fato vem acontecendo com a rede ISDN, muitas organizações e instituições estão migrando as suas redes para IP e mantendo o legado através do uso de gateways.

Usualmente implementados em MCUs ou em *appliances* especializados, os gateways efetuam a conversão de protocolos entre as redes mediadas por estes. Quando necessário é também feita a adequação das interfaces de acesso, como por exemplo o ethernet usado no H.323 e os padrões estabelecidos pelo ISDN ou PSTN. Outra funcionalidade disponível é a transcodificação de padrões de áudio e vídeo.

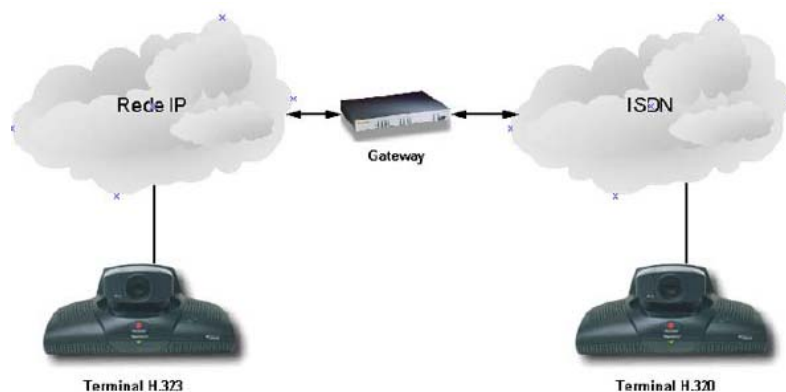


Figura 2.12 – Gateway para interconexão IP e ISDN (Guimarães, 2002)

Uma aplicação muito comum para gateways é o cenário onde se inclui chamadas tradicionais de voz (PSTN) em comunicações utilizando H.323. Essa aplicação pode ter como objetivo desde a comunicação com pessoas que não têm um equipamento específico para videoconferência, até a extensão das funcionalidades de um sistema para incluir usuários quando “móveis” (ex: inclusão de uma chamada celular). Um gateway H.323-H.320 também pode ser utilizado para prover um acesso de backup. Caso o acesso via enlaces IP fique indisponível, codecs IP podem chegar até um gateway e sair da sua rede através de uma conexão ISDN.

Com a popularização do protocolo SIP, o gateway H.323-SIP está se tornando bastante comum em diversos ambientes. O SIP é muito utilizado em aplicações de voz sobre IP (VoIP) e recentemente para aplicações de vídeo. Diversos fabricantes oferecem atualmente terminais e MCUs capazes de comunicar entre os dois protocolos. Os gateways serão comentados novamente nesta dissertação na sessão sobre integração de plataformas.

2.4.1.3 - Multipoint Controler Units - MCUs

A MCU é uma entidade da rede H.323 que tem a finalidade de gerenciar sessões de videoconferência entre três ou mais participantes. Esse recurso é viabilizado através do conceito de salas virtuais, onde vários codecs se conectam para compartilhar áudio, vídeo e

dados. Todos os terminais participantes da conferência devem estabelecer uma conexão com a MCU que gerencia os recursos de processamento e negocia os protocolos e codecs a serem utilizados entre os terminais.

A MCU é composta de duas partes: o *Multipoint Controller* (MC) e um ou mais *Multipoint Processors* (MP). O MC é obrigatório e processa as negociações (H.245) entre os terminais para estabelecimento das chamadas e determinação dos algoritmos de áudio e vídeo a serem utilizados. Em aplicações específicas, determina se algum dos fluxos de dados serão transmitidos em multicast (Packetizer, 2009). No entanto, o MC não realiza nenhum processamento de áudio ou vídeo, essa é a função dos MPs.

O MP é um componente opcional e permite o processamento dos fluxos de dados, sendo que podem existir nenhum ou vários MPs em uma conferência multiponto. O processamento consiste em receber os fluxos de dados dos terminais, mixá-los e devolver para cada terminal um único fluxo de áudio e vídeo. O processamento pode ser feito em *Voice Switched Mode* (tela cheia para todos os participantes e chaveamento de vídeo através do áudio) ou *Continuous Presence* (presença contínua), neste caso os terminais recebem uma tela dividida de acordo com a quantidade de participantes de uma conferência.

A Figura 2.13 mostra os fluxos de dados que ocorrem durante uma conferência multiponto. Já que cada terminal deve se conectar à MCU, enviando e recebendo fluxos de mídia, percebe-se que em uma conferência com N participantes a MCU deve ser capaz de suportar estas N conexões. Também a rede que suporta esta MCU deve ser capaz de suportar a banda requerida para a conferência. Se cada terminal (N terminais) se conecta a 384 kbps, por exemplo, o enlace que chega até a MCU deve ser capaz de suportar $N \times 384$ kbps sendo ainda necessário considerar o *overhead* dos pacotes da rede IP. É possível também, com a maioria das MCUs disponíveis no mercado, que cada terminal se conecte usando uma velocidade diferente. Neste caso, a MCU executa a função de transcodificação para que um terminal conectado a 384 kbps possa ver e ouvir outro terminal conectado a 1 Mbps por exemplo.

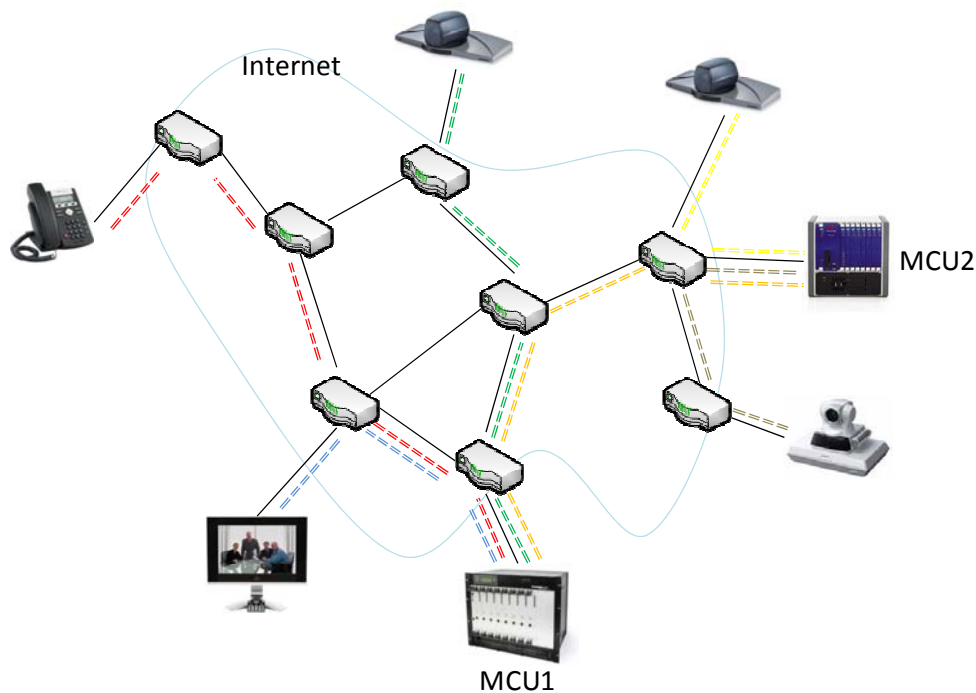


Figura 2.13 – Conexões entre terminais e MCUs para conferências multiponto

Para viabilizar conferência com muitos participantes, usando enlaces restritos ou MCUs com recursos insuficientes, é possível fazer o cascadeamento de MCUs. A Figura 2.13 mostra a MCU2 conectada a dois terminais e também à MCU1. Desta maneira a MCU2 é um participante da conferência realizada na MCU1 e vice versa, fazendo com que todos os terminais possam participar de uma mesma conferência. Este recurso é utilizado também quando alguns terminais (V.35, ISDN, SIP etc) não são compatíveis com a MCU que sedia a conferência, conectando-os à conferência através de uma outra MCU compatível.

Existem diversos modelos de MCUs no mercado e que se diferenciam em diversos aspectos podendo ser estes recursos de vídeo, recursos de voz, *layouts* possíveis para salas virtuais, banda suportada, portas com transcodificação, capacidade de expansão, codecs suportados e protocolos suportados. A MCU, por definição da norma, deve suportar H.323, mas os fabricantes podem disponibilizar MCUs com capacidade para conectar equipamentos usando plataformas como ISDN e SIP.

2.5 - SESSION INITIATION PROTOCOL - SIP

Criado na Universidade de Columbia, em meados da década de 1990, o *Session Initiation Protocol* (SIP) constitui um protocolo alternativo ao H.323 em algumas aplicações. É um

protocolo de aplicação, que utiliza o modelo “requisição-resposta”, similar ao HTTP (*HyperText Transfer Protocol*), para iniciar sessões de comunicação interativa entre usuários. O SIP foi aprovado pela IETF (*Internet Engineering Task Force*) e publicado em junho de 2002 como RFC 3261 (IETF RFC 3261, 2002). A IETF é uma sociedade aberta da qual participam pesquisadores, projetistas, operadores de telecomunicações e provedores de serviços Internet, bem como fabricantes de equipamentos. Todos são voluntários e estão, direta ou indiretamente, relacionados com a arquitetura da Internet, com a especificação e o desenvolvimento de protocolos de comunicação e aplicações, ou com a operação, a segurança e o gerenciamento desta rede.

Este protocolo foi desenvolvido para estabelecer chamadas e conferências em redes baseadas no protocolo IP. O estabelecimento, mudança ou término da sessão é independente do tipo de mídia ou aplicação usada durante a chamada, logo uma chamada qualquer pode trafegar diferentes tipos de dados, incluindo áudio e vídeo. São exemplos de sessões: telefonia via Internet, conferência multimídia, jogos de computador distribuídos, etc. O SIP recebeu também uma rápida adoção com um padrão para comunicações baseadas em presença que é a capacidade de uma aplicação estar alerta sobre a localização e disponibilidade de um usuário (Gonçalves, 2008).

O SIP utiliza uma arquitetura baseada na estrutura cliente-servidor. As requisições geradas por uma entidade cliente são enviadas para uma entidade do tipo servidor. O servidor processa a mensagem e devolve ao cliente uma resposta. Para identificação dos usuários utiliza-se um formato de endereçamento semelhante ao email, o SIP URI (*Uniform Resource Identifier*), com o seguinte formato: sip:nome@domínio. Ex.: sip:aristides@ekiga.net, +55456598953@sipA.com, 7845495@sipB.com.

A arquitetura SIP possui quatro elementos principais: SIP *User Agent*, SIP *Proxy Server*, SIP *Redirect Server* e SIP *Registrar Server* (IETF RFC 3261, 2002),

- **SIP *User Agent* (Agente de usuário)** - pode agir como cliente (UAC - *User Agent Client*) enviando pedido de inicialização de sessão, e como servidor (UAS - *User Agent Server*) respondendo o pedido de sessão. Exemplos de agentes de usuário são os telefones IP e aplicações de conferência (implementadas em servidores denominados *proxy* por exemplo), estes últimos funcionam simultaneamente como clientes e como servidores.

- **SIP Proxy Server (servidor Proxy)** - São entidades mediadores (ao nível de aplicação) que funcionam ao mesmo tempo como servidores e clientes no intuito de fazer requisições no lugar de outros clientes. A sua função primária é fazer o roteamento de mensagens entregando-as para a próxima entidade (IETF RFC 3261, 2002). Os pedidos podem ser atendidos internamente ou encaminhados para outros servidores. É possível que algumas mensagens sejam alteradas antes de encaminhadas. Os *SIP Proxy Servers* podem atuar de dois modos:
 - **Stateful Proxy Server** – Neste modo o servidor armazena as informações de estado das entidades com as quais troca mensagens. Provê maior confiabilidade e controle, no entanto exigem mais processamento e memória do hardware onde são instalados.
 - **Stateless Proxy Server** – Neste modo o servidor somente encaminha mensagens e não armazena informações de estado. O modo *stateless* é usado quando não é necessário que o Proxy controle as chamadas que ocorrem na rede e que passam por ele.
- **SIP Redirect Server (Servidor de Redirecionamento)** – servidor intermediário que atende ao pedido de Agente Usuário fornecendo nome e localização de outro usuário. Esses servidores aceitam um pedido SIP, convertem o endereço recebido como requisição de pesquisa em zero ou mais endereços novos, e devolvem os resultados ao cliente. Ao contrário dos servidores de proxy, o servidor de redirecionamento não inicia o seu próprio pedido SIP e ao contrário dos agentes de utilizador servidores, não aceita chamadas.
- **SIP Registrar Server (Servidor de Registro)** – Registra informações do Agente Usuário para fornecer informações de localidades e status para determinado domínio. Para tal recebe requisições de registro e armazena os dados em um servidor de localização. São os responsáveis por saber onde se encontra um determinado usuário dentro do seu domínio, mesmo que essa informação de localização seja um servidor de redirecionamento.

Já que o SIP é usado para estabelecer sessões independentes do tipo de dado trafegado, o protocolo SDP (*Session Description Protocol*), como seu nome sugere, é utilizado para descrever o tipo de sessão que deve ser estabelecida (IETF RFC 2327, 1998). O protocolo não trabalha somente com o SIP, podendo ser utilizado em sessões chamadas estabelecidas por outros protocolos. Em geral o SDP deve conter as informações suficientes para permitir que

os clientes estabeleçam uma sessão e que os recursos a serem utilizados sejam anunciados. As informações de mídia contêm:

- Tipo de mídia (vídeo, áudio, etc)
- Protocolo de transporte (RTP/UDP/IP, H.320, etc)
- O formato da mídia (vídeo H.261, vídeo MPEG, etc)

Logo, devido à sua versatilidade, o SIP é usado tanto para sessões de voz quanto vídeo. Pode-se, como ocorre com a recomendação H.323, também usar outros protocolos com o SIP para que funções específicas sejam viabilizadas.

Outra aplicação do protocolo SIP é o serviço de presença. Através do SIMPLE (*SIP based Instant Messaging Presence Leveraging and Extensions*) (IETF–SIMPLE, 2009), o SIP oferece funções-chave para presença e mensageria instantânea. O SIMPLE é composto de uma série de especificações que descrevem extensões do protocolo SIP. Seus vários requisitos encontram-se documentados em RFCs individuais ou em documentos não oficiais (drafts). A RFC 3428 (IETF RFC 3248, 2002) acrescenta ao SIP o método MESSAGE, que permite a transmissão de Mensagens Instantâneas semelhantes a um sistema de “*pager*”, ou seja, as mensagens são independentes umas das outras, não existindo um vínculo de sessão entre elas. Na RFC 3856 (IETF RFC 3856, 2004) define-se dois novos métodos; SUBSCRIBE, que solicita o recebimento de notificações de um nó remoto quando certos eventos ocorrerem (tais como alterações na disponibilidade do usuário) e NOTIFY, que envia notificações aos nós remotos quando os eventos ocorrem. As informações enviadas via NOTIFY encontram-se no formato *Presence Information Data Format* (PIDF) (IETF RFC 3863, 2004).

A Presença permite aos usuários fornecer para os dispositivos de outros usuários uma indicação sobre o seu próprio estado, disponibilidade, e como podem ser contatados mesmo antes de uma sessão de comunicação ser iniciada. Quando se integra a aplicações de telefonia, essa funcionalidade oferece o controle sobre como os usuários podem ser localizados. Pode-se fornecer informações sobre dispositivos múltiplos como Telefones IP, celulares, *softphones*, *paggers* (localizadores) e dispositivos sem fio ou dispositivos *Bluetooth*.

3 - REDES

Ressaltados os protocolos e padrões mais usados em comunicações unificadas, este capítulo contém informações sobre as infraestruturas utilizadas.

3.1 - REDES COMUTADAS POR PACOTES E REDES COMUTADAS POR CIRCUITOS

Não há comunicação sem rede de acesso, os sistemas têm de estar conectados a uma rede para se comunicarem. A rede de acesso é o elemento básico e fundamental para qualquer comunicação de áudio, vídeo e dados em geral; a infra-estrutura de comunicação é tão importante para esse procedimento quanto os próprios terminais (Stallings -1997).

É importante ressaltar que existem duas grandes classes de redes de acesso: as redes comutadas por circuitos e as redes comutadas por pacotes. A infra-estrutura para a rede de acesso influencia no funcionamento dos equipamentos e protocolos utilizados.

A comunicação em comutação por circuitos implica que existe um caminho de comunicação dedicado entre dois pontos. Esse caminho é uma conexão seqüencial de enlaces entre nós de rede. Em cada enlace físico existe um canal lógico que é dedicado para aquela conexão. Esse tipo de comunicação envolve três fases Figura 3.1:

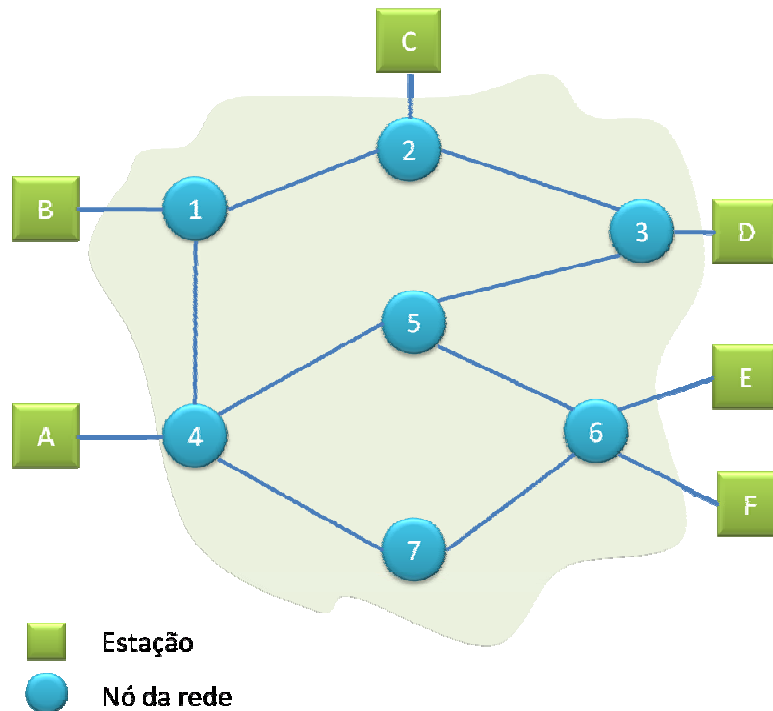


Figura 3.1- Diagrama representando nós e estações (Stallings - 1997)

1. Estabelecimento do circuito – Antes dos sinais serem transmitidos, um circuito fim-a-fim deve ser estabelecido. Por exemplo, a estação A deseja se comunicar com a estação E. O enlace entre A e o nó 4 tipicamente é dedicado, logo essa conexão já está estabelecida. O nó 4 precisa encontrar um rota para chegar ao nó 6. Para isso, aloca um canal livre até o nó 5, estabelecendo então um canal dedicado entre A, 4 e 5. O mesmo procedimento é feito até chegar à estação E.
2. Transferência de dados – A informação agora pode ser transmitida de A para E, sendo esta digital ou analógica, dependendo da natureza da rede. A transmissão digital tornou-se dominante para o tráfego de dados uma vez que é mais confiável e permite melhor utilização de recursos.
3. Encerramento da conexão – Após um período de transferência de dados, a conexão é terminada por um dos lados. Mensagens devem ser propagadas através dos nós 4, 5 e 6 para desocupar os canais alocados.

Este tipo de redes foi criado inicialmente para tráfego de voz, mas existiam também muitas aplicações para transferência de dados utilizando modems. No entanto o tráfego de dados tem um comportamento diferente do tráfego de voz, podendo ser ressaltados alguns problemas (Stallings - 1997):

- Em uma comunicação de dados típica (ex: cliente acessando um banco de dados), em grande parte do tempo a conexão estabelecida não está sendo utilizada, logo, o circuito fica alocado mesmo não havendo tráfego de dados.
- Em uma conexão por circuitos a transmissão deve ser sempre feita em uma velocidade constante. Ou seja, os terminais conectados devem transmitir dados na mesma velocidade. Este fato limita a interconectividade de uma variedade distinta de computadores e terminais.

Por volta dos anos 70, pesquisas foram realizadas para o desenvolvimento de uma nova forma de comunicação digital de longa distância, a comutação por pacotes. Em um meio comutado por pacotes, os dados são divididos em partes (pacotes) e transmitidos pela rede. Os pacotes são roteados independentemente na rede e os recursos são alocados quando necessários. O objetivo principal dessa forma de comunicação é otimizar o uso do enlace disponível. Este tipo de conexão é conhecido pelas características a seguir (Comer - 1998):

- sem conexão – Cada pacote é independente dos outros. Uma seqüência de pacotes pode trafegar por caminhos diferentes e alguns podem ser perdidos.
- melhor esforço – O software de interligação faz uma série de tentativas para entregar os pacotes até que consiga utilizar o meio.
- não confiável – A entrega de pacotes não é garantida. O pacote pode ser perdido, reproduzido, atrasar-se ou ser entregue com problemas

Retornando ao exemplo da Figura 3.1 verifica-se uma diferença no funcionamento do mecanismo de entrega de dados. Assumindo um mecanismo simples de comutação por pacotes, a estação A deseja enviar dados para a estação E. O pacote a ser enviado contém informações de controle indicando que o seu destino é a estação E. O pacote é então enviado ao nó 4. Este nó armazena o pacote, determina a rota e coloca o pacote na fila para transmitir pelo caminho escolhido. Supondo que a melhor rota era pelo nó 5, quando o enlace está disponível, o pacote é transmitido. O mesmo processo acontece passando pelo nó 6 até chegar à estação E. Essa abordagem apresenta algumas características:

- A eficiência da linha é maior, pois um enlace pode ser compartilhado dinamicamente por diversos pacotes de conexões distintas. Os pacotes são enfileirados e transmitidos de acordo com a capacidade do enlace. Na comutação

por circuitos, o enlace é pré-alocado usando TDM (*Time Division Multiplexing*), logo o “espaço no tempo” já está reservado para a transferência do dado, ainda que nada esteja sendo enviado, havendo então um desperdício do canal que continua alocado sem necessidade.

- Quando o tráfego está muito alto em uma rede comutada por circuitos, algumas chamadas são bloqueadas, ou seja, a rede recusa-se a aceitar requisições de conexão até que a carga na rede diminua. Em uma rede comutada por pacotes, os pacotes ainda são aceitos, mas são entregues com atrasos maiores.
- Prioridades podem ser usadas. Logo, se uma rede tem uma quantidade de pacotes enfileirados para transmissão, os pacotes de mais alta prioridade podem ser transmitidos primeiro. Estes, portanto experimentam menos latência do que os pacotes de mais baixa prioridade.

3.2 - INTEGRATED SERVICES DIGITAL NETWORK - ISDN

Como ressaltado no tópico 2.3 - , a recomendação H.320 define as metodologias para transporte de videoconferência sobre as redes ISDN (*Integrated Services Digital Network*). Esta é uma rede comutada por circuitos. De maneira simples, pode-se dizer que este tipo de rede utiliza o sistema telefônico comutado como forma de acesso.

As recomendações ISDN foram publicadas em 1984 pela especificação ITU-T I.411 (ITU-T I.411 - 1993), sendo ao mesmo tempo um conjunto de padrões para transmissão digital e uma infra-estrutura de rede que permite transmissão digital fim-a-fim sobre os fios existentes de telefonia. Através desse serviço digital, um assinante poderia utilizar, além dos serviços de voz da telefonia convencional (analógica), serviços de comunicações de dados, com velocidades superiores as conseguidas com os modems analógicos.

ISDN é baseada originalmente no canal digital T1 que utiliza sinalização de canal comum número 7 (SS7 – Sistema de Sinalização número 7) para controle de comutação dos circuitos telefônicos. Estes circuitos são chaveados de acordo com a demanda para passar tráfego de voz, vídeo ou dados.

Existem dois tipos básicos de acesso para o usuário: o *Basic Rate Interface* (BRI), para usuários com menor demanda de banda e o *Primary Rate Interface* (PRI), para usuários com

maior necessidade de banda. Ambos utilizam multiplexação por divisão de tempo (Wilcox, 2000).

O circuito BRI é formado por dois canais B (Bearer) e um canal D (Delta). Os dois canais B são utilizados para trafegar voz e dados a uma velocidade de 64 kbps. O canal D é usado para sinalização e possui uma largura de banda de 16 kbps. Por essa razão, o circuito BRI é conhecido como 2B+D, totalizando 144 kbps. Essas informações são sumarizadas pela Tabela 3-1.

Tabela 3-1 – Configuração de um acesso BRI

Basic Rate Interface (BRI)	
Velocidade da linha	144 Kbps
Padrão da camada física	ITU-T 1.430
Configuração	2B + D
Canal de informação (B)	64 Kbps
Canal de sinalização (D)	16 Kbps

O circuito PRI compreende o acesso de alta capacidade. Esta forma de acesso oferece uma linha com características T1 ou E1. O padrão T1 é usado na América do Norte e Japão, sendo composto de 23 canais B e um canal D de 64 Kbps (23D+D). O E1 é o padrão usado na Europa e demais países, inclusive o Brasil, e acomoda 30 canais B e um canal D de 64 Kbps (30B+D). Em termos de velocidade, para o T1 tem-se 1.544 Mbps e para o E1, 2.048 Mbps. Esses dados estão sumarizados na Tabela 3-2.

Tabela 3-2 – Composição de um acesso PRI

Primary Rate Interface (PRI)	
Velocidade da linha	1.544 Kbps (América do Norte)
	2.048 Mbps (Europa e Brasil)
Padrão da camada física	ITU-T 1.430
Configuração	23B + D (América do Norte)
	30B + D (Europa e Brasil)
Canal de informação (B)	64 Kbps
Canal de sinalização (D)	64 Kbps

Os canais B de um tronco PRI podem ser usados em conjuntos adequados para se obter a velocidade desejada para uma conexão. Por exemplo, caso se queira fazer uma videoconferência a uma taxa de 512 kbps pode-se usar somente oito canais (8B) ou quatro vezes 128 kbps. Dessa forma a banda restante do tronco pode ser utilizada para a realização de outras conexões de voz ou dados. Outra forma de se obter a mesma velocidade é usando 4 canais BRI (4x128 kbps), neste caso um multiplexador é necessário para que todos esses canais possam ser usados em uma mesma conexão.

A arquitetura padronizada pelo ITU – T do acesso BRI ISDN é mostrada na Figura 3.2

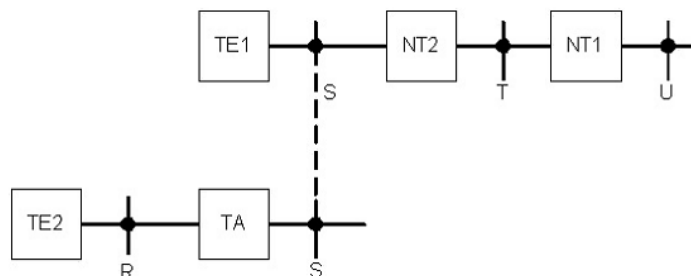


Figura 3.2 – Arquitetura padronizada para o acesso BRI (Wilcox, 2000)

Os elementos da arquitetura são descritos abaixo:

- **NT1 - Network Termination 1:** Este é o equipamento que conecta fisicamente o usuário à rede ISDN. No caso de acesso PRI, este dispositivo é uma CSU/DSU - *Channel Service Unit/Data Service Unit*. Para acessos BRI, ele é chamado NT1, e pode ser integrado ao equipamento final ou não. Repare que na Europa e demais países (exceto EUA), o NT1 é fornecido pelo próprio provedor de serviços ISDN, e portanto não deve estar integrado ao equipamento do usuário.
- **NT2 - Network Termination 2:** Tipicamente, equipamentos que desempenham funções de multiplexação e/ou comutação, como por exemplo PBXs..
- **TA - Terminal Adapter:** Utilizados para conexão de equipamentos não-ISDN à rede.
- **TE1 - Terminal Equipment:** Equipamento final do usuário com suporte a ISDN.
- **TE2 - Terminal Equipment:** Equipamento não-ISDN, conectado à rede via um Terminal Adapter
- **S, T e U** - Interfaces de comunicação entre os equipamentos, onde são definidos os protocolos de comunicação entre eles.

O ISDN se tornou um meio bastante comum para telecomunicações e em especial para videoconferência. Agrupando dois canais B, no acesso BRI, por exemplo, temos uma taxa de 128 Kbps que foi durante muito tempo utilizada para transmissão de áudio e vídeo *full-duplex* em uma qualidade aceitável. É possível também agrupar mais de um acesso BRI para aumentar a velocidade de conexão. Ao se utilizar 4 acessos BRI, pode-se conseguir um acesso de 512 kbps. Para tal é necessário um equipamento conhecido como IMUX (*Inverse Multiplexer*), este é capaz de agrupar essas linhas para que os recursos sejam compartilhados em uma mesma conexão. Um exemplo de IMUX é mostrado na Figura 3.3.



Figura 3.3 – IMUX (Polycom)

Apesar da confiabilidade do enlace ISDN, por ser um canal dedicado, essa plataforma está sendo gradativamente deixada de lado. O gráfico da Figura 3.4 (Weinstein, 2006) mostra que as videoconferências via ISDN estão sendo substituídas pelas realizadas usando o protocolo IP. Este resultado deve-se a uma série de fatores. O custo das chamadas ISDN para organizações pode ser muito alto tanto devido à realização de chamadas quanto à assinatura do enlace. O custo mensal da rede IP não aumenta de acordo com a quantidade de chamadas realizadas. As redes ISDN também são isoladas da rede de dados, não viabilizando um gerenciamento deste enlace, por exemplo, só é possível saber da disponibilidade de um enlace no momento em que se faz a chamada. Para se comunicar com os novos equipamentos em IP é necessário o uso de gateways que é um elemento adicional na rede e com suas limitações de capacidade.

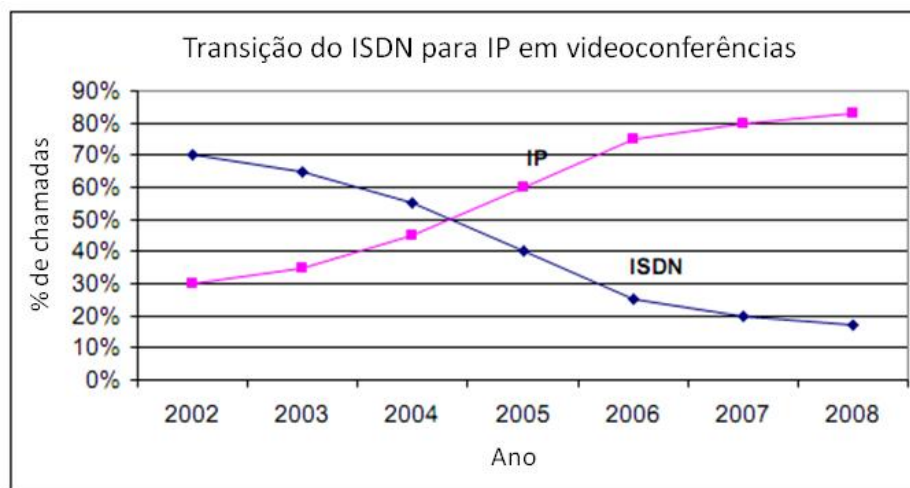


Figura 3.4 – Transição ISDN para IP (Weinstein, 2006)

3.3 - TRÁFEGO DE DADOS NAS REDES IP

As aplicações de voz sobre IP (VoIP) e videoconferência utilizam fluxos (*streams*) de dados com características diferentes das aplicações de dados tradicionais como acessar uma página na internet ou um banco de dados dentro da LAN – *Local Area Networks*. Se uma organização pretende disponibilizar o serviço de áudio e videoconferência sobre redes IP, deve preparar a sua rede para este tipo de tráfego. O fato é que muitas redes empresariais ainda não estão equipadas adequadamente para suportar tráfego para videoconferência, sendo que muitas vezes é necessário avaliar, testar e possivelmente atualizar e reconfigurar a infraestrutura para se obter uma qualidade aceitável.

Em vários casos, quando são feitas pesquisas de pós-implementação de soluções de colaboração em empresas, o relato mais comum obtido é a baixa qualidade de áudio e vídeo. Esses fatores são as principais razões que levam a uma baixa utilização da solução pelos usuários finais. Quando uma análise mais profunda do problema é realizada, as razões para os problemas quase sempre levam à capacidade das redes LAN e WAN em suportar o tráfego de dados (Polycom, 2009).

As grandes corporações, universidades e empresas projetam redes locais que sejam capazes de transmitir todos os tipos de mídias, ou seja, procuram ter uma rede multimídia. A meta é ter uma única solução que suporte toda a informação trafegada numa rede de computadores ligada à Internet, desde acessar uma página web até a realização de aplicações de áudio e vídeo (Black, 1999).

3.3.1 - *Transport Control Protocol/ Internet Protocol - TCP/IP*

As redes locais existentes hoje, na sua maioria, usam a tecnologia ethernet. O que caracteriza esta tecnologia é sua topologia física e os protocolos usados para comunicação entre os equipamentos. Esta sessão da dissertação se concentrará nos protocolos usados, que neste caso são o IP e o TCP.

A popularidade da rede mundial de computadores (Internet) já está bem estabelecida e cresce a cada dia. Juntamente com este crescimento aumenta o uso de redes locais (intranets ou LANs) baseadas na família de protocolos TCP/IP. Esta família de protocolos está dividida em camadas de tal forma que a interoperabilidade entre as camadas é feita por conjunto de serviços, conhecidos como protocolos de comunicações.

O modelo de camadas que pode ser comparado ao TCP/IP é o modelo OSI (*Open Systems Interconnection*) (ITU-T X.200, 1988). O modelo OSI descreve uma arquitetura de 7 camadas e a interação entre elas. Neste tipo de modelo, os protocolos das camadas mais altas utilizam os serviços dos protocolos das camadas inferiores que, por sua vez, não precisam conhecer os protocolos e aplicações das camadas acima (Davis, 2009). A Figura 3.5 mostra o modelo de camadas OSI.

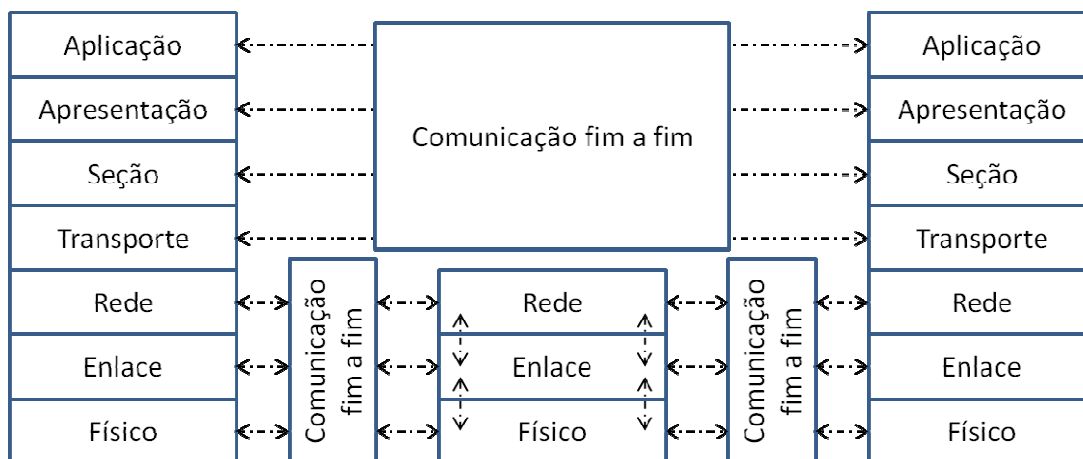


Figura 3.5 – Modelo de camadas OSI (Stallings, 1997)

As redes TCP/IP (Internet, por exemplo) estão organizadas em quatro camadas: física, rede, transporte e aplicação. A Figura 3.6 mostra o desenho da arquitetura de uma rede IP.

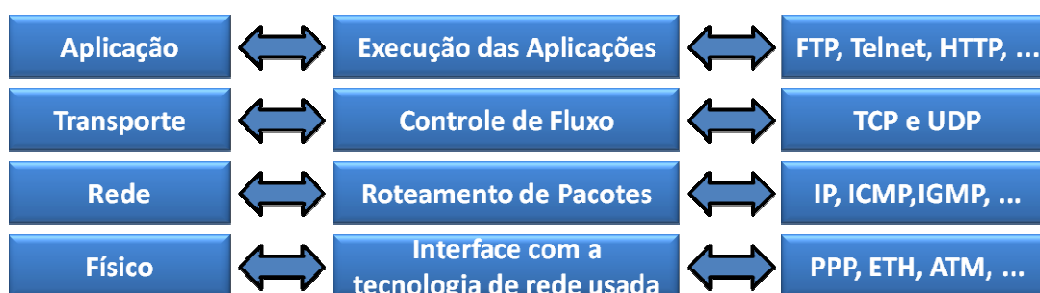


Figura 3.6 – A divisão em 4 camadas das redes TCP/IP (Guimarães, 2002).

O nível físico pode utilizar diversos padrões de redes locais para enviar os sinais (bits) para o cabeamento. Os mais usados são: IEEE 802.2, 802.3 e 802.4 e o HDLC33 (norma X.25)¹. O IEEE 802 é a família de padrões que define as regras de comunicação para redes ethernet. Os padrões 802.2, 802.3 e 802.4 são respectivamente: protocolo de redes locais que define uma implementação de LLC (*Logical enlace control*), protocolo de redes locais que define a implementação da camada física e do MAC (*Media Access Control*) com o uso de CSMA/CD (*Carrier sense multiple access/Collision detection*) para o acesso em várias velocidades e em vários tipos de meio físico (fibra óptica, par trançado, cabo coaxial e etc.) e protocolo de redes locais que define uma implementação da camada física e do MAC, usando o acesso ao meio através de um token (Clayton, 2000).

¹ www.itu.int/tiu-t

O nível de rede é responsável pela conectividade e interoperação das diversas redes existentes na Internet. Neste nível destaca-se o protocolo IP (*Internet Protocol*). O IP implementa um mecanismo de roteamento das mensagens que permite que um programa no nível de aplicação troque informações com outro mesmo estando que redes completamente diferentes (Black, 1999).

O nível de transporte oferece um serviço confiável de transferência de dados fim-a-fim entre aplicações. Os serviços providos por este nível devem oferecer total transparência com respeito aos níveis inferiores e garantir a integridade dos dados trocados na rede utilizando mecanismos de segurança como *checksum*, controle de fluxo, sequenciamento, entre outros. Os principais protocolos definidos para este nível na Internet são: TCP e o UDP (*User datagram protocol*).

O TCP é um serviço (protocolo) confiável e orientado à conexão, ou seja, os dados são entregues ao destinatário através do uso de mensagens de confirmação de recebimento e da possível retransmissão de pacotes quando necessário. O TCP define as regras através das quais dois terminais de videoconferência concordam em estabelecer a conexão e como eles trocam as mensagens. Permite ainda o controle de fluxo de tal forma que o emissor não envie os dados a uma taxa maior que o receptor possa suportar, além de fornecer mecanismos para detecção e correção de erros.

Enquanto o TCP é um protocolo de transporte orientado à conexão, o IP é um protocolo de rede que opera no modo sem conexão. Assim sendo, a combinação TCP/IP pode oferecer um serviço de alta confiabilidade. Para o uso de redes de alta qualidade, onde o problema de confiabilidade não é crítico, pode-se usar o protocolo UDP como protocolo de nível de transporte. Este opera em um modo sem conexão e possui funcionalidades bem mais simplificadas do que o TCP.

O nível de aplicação oferece ao usuário um conjunto de protocolos para implementar os serviços entregues diretamente para um usuário final ou aplicação. Os serviços mais conhecidos e utilizados hoje são entregues por essa camada, como por exemplo, SMTP (*Simple mail transfer protocol*), TELNET, HTTP (*Hypertext transfer protocol*).

3.3.2 - User Datagram Protocol (UDP)

O UDP é um protocolo da camada de aplicação fundamental para as comunicações multimídia. Este protocolo oferece um serviço de transporte de dados sem qualquer tipo de garantia de que um pacote chegue ao destino. Uma prova da simplicidade deste protocolo está em seu cabeçalho que só contém campos que referem-se a porta de origem, porta de destino, comprimento da mensagem e *checksum*. Caso garantias sejam necessárias, é preciso implementar uma série de estruturas de controle, tais como timeouts, retransmissões, acknowledgments, controle de fluxo, etc.

O tráfego de VoIP e videoconferência utilizam uma combinação do TCP e UDP. O TCP é utilizado para controles primários de chamadas e sinalização incluindo estabelecimento de chamadas, controle de fluxo e negociação de codecs. Estes controles de chamadas dependem de uma comunicação confiável para garantir que chamadas possam ser estabelecidas com sucesso e mantidas.

Em contraste, como será visto adiante, os dados que trafegam áudio e vídeo são muito sensíveis à variações no tempo de entrega dos pacotes. Sendo assim, estes dados se beneficiam da baixa latência oferecida pelo UDP quando comparado ao TCP. No entanto, o UDP não possui mecanismos de sequenciamento de pacotes ou para garantir a entrega destes. Essa característica pode comprometer a comunicação pois uma aplicação não “sabe” sobre a qualidade e confiabilidade de chamadas em andamento.

Para tratar essa característica, o UDP trabalha com os protocolos RTP (*Real Time Protocol*) e RTCP (*Real Time Control Protocol*) para implementar essas funcionalidades. O RTP, em conjunto com o UDP, fornece serviços de entrega de áudio e vídeo fim-a-fim. As principais funções fornecidas pelo RTP são identificação do tipo de dados trafegados (garantir decodificação adequada), numeração de sequência de pacotes (permite reordenação e facilita a reconstrução) e monitoramento do tempo de entrega (*timestamping* – permite regular o controle do *jitter*) (IETF RFC, 1889).

O RTCP se baseia na transmissão periódica de pacotes de controle para todos os participantes de uma sessão, utilizando os mesmos mecanismos de distribuição dos pacotes de dados. Os pacotes RTCP contêm informações importantes para a monitoração da entrega dos pacotes de dados, tais como: *jitter*, número de pacotes perdidos, número de pacotes e octetos transmitidos e outros dados úteis para diagnósticos, monitoração e correção de alguns erros na rede.

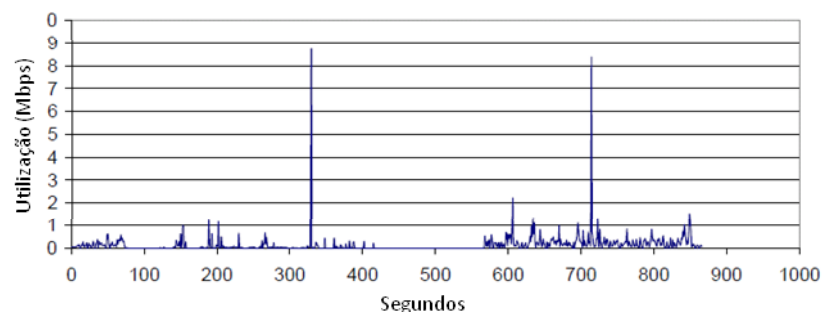
A função primária do RTCP é prover informações sobre a qualidade da distribuição dos dados. Outras funções incluem o transporte de um identificador canônico usado para sincronização de *streams* múltiplos de um mesmo terminal (ex. áudio e vídeo) e garantir que a taxa de transmissão dos dados sejam conhecidas pelos participantes de uma chamada.

3.3.3 - Tráfego Contínuo

O tráfego contínuo suporta aplicações interativas, sendo áudio e vídeo as mais proeminentes. Ambas as aplicações possuem usuários que esperam que o que é falado ou filmado chegue ao outro lado “instantaneamente” para que tudo aconteça como se os participantes estivessem se comunicando diretamente sem a intervenção de outros equipamentos. Alguns dos aspectos mais complicados advêm da necessidade de velocidade nos enlace que trafegam estes dados.

O IP, conforme mostrado anteriormente, está na base da maior parte das redes atualmente e, portanto é responsável por conectar os terminais. No entanto, por projeto, o IP é um protocolo não confiável. O que significa que ele não garante que todos os pacotes enviados cheguem com sucesso à outra ponta. O protocolo TCP por sua vez, apesar de orientado à conexão, somente garante que os dados recebidos sejam os idênticos aos dados enviados.

As aplicações de dados trabalham com rajadas, ou seja, o tráfego não é contínuo. Quando um arquivo ou bloco de informação está pronto para ser transferido pela rede, o emissor dispara o envio do dado assim que possível e depois passa a trabalhar em outras tarefas. O resultado são utilizações intermitentes dos recursos da rede intercalados com períodos de uso



relativamente baixos. A

Figura 3.7 mostra um exemplo típico de uma aplicação de dados mostrando períodos de baixa e alta utilização da rede (Polycom, 2006).

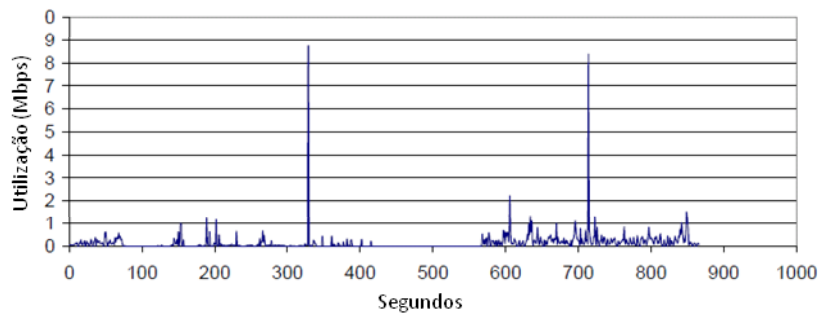


Figura 3.7 – Utilização típica de banda para transmissão de dados (modificado - Polycom, 2006)

As redes IP normalmente experimentam perdas de pacotes sendo este um comportamento normal. A perda de pacotes é até mesmo usada como um mecanismo de controle pelo protocolo TCP. Este mecanismo determina a velocidade que deve ser usada para transmitir dados e funciona aumentando a taxa de transmissão até verificar perdas de pacotes para, por fim, calibrar a velocidade de transmissão. Esse processo ocorre de tempos em tempos na rede durante uma conexão.

O tráfego contínuo possui características bem diferentes das citadas para o tráfego tradicional de dados. Os dados das aplicações de áudio e vídeo são amostragens contínuas de um ambiente real (voz e imagens), sendo transmitidos constantemente para serem reproduzidos na outra ponta. Logo, a utilização dos recursos da rede deve ser constante. Caso um pacote seja atrasado ou fique perdido em trânsito haverá um lapso nas informações disponíveis para o decodificador, conseqüentemente a qualidade do áudio ou vídeo será degradada.

Uma vez que os pacotes para tráfego de áudio/vídeo devem chegar o quanto antes e em seqüências, o controle de erros da camada de transporte (TCP) pode gerar problemas. Quando existe controle de erros de transmissão, o receptor normalmente envia mensagens para o transmissor avisando que não recebeu os pacotes esperados, o emissor então reenvia os pacotes perdidos. Esse processo se torna inviável para comunicação em tempo real uma vez que um pacote retransmitido não faria mais sentido na comunicação caso fosse interpretado.

Essa questão é resolvida com o uso do protocolo da camada de transporte UDP (*User Datagram Protocol*) que não possui um mecanismo de recuperação de pacotes perdidos. Os pacotes são enviados e eles chegam ao destino a tempo, atrasados ou são perdidos.

A Figura 3.8 mostra um gráfico de uso da rede durante uma videoconferência ocorrendo a 384 kbps. Estão representados os fluxos de áudio e vídeo, sendo possível notar o uso mais uniforme da rede.

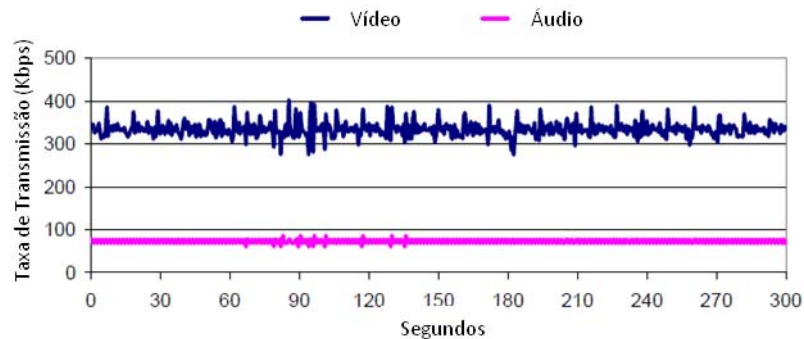


Figura 3.8 – Utilização típica de banda para transmissão de vídeo+áudio (modificado - Polycom, 2006)

Assim sendo, é necessária uma rede com largura de banda disponível, sem perda de pacotes e sem atraso. Estas são características fundamentais que as redes IP devem ter. Existem alguns fatores que influenciam as aplicações em tempo real de videoconferência, que são:

- **Largura de banda ou *throughput*.** Largura de banda é a taxa de transmissão de dados (bits por segundo) de uma rede. François Fluckiger define largura de banda como: “A taxa de bits entre dois terminais é o número de dígitos binários que a rede é capaz de aceitar e entregar por unidade de tempo” (Fluckiger, 1995). Vale ressaltar que as aplicações de videoconferência em IP requerem mais banda do que aplicações em ISDN, por exemplo. Uma aplicação de videoconferência em que ocupe uma banda de 384 Kbps (320 Kbps para vídeo e 64 Kbps para áudio) em ISDN, vai ocupar 25% a mais de banda em redes IP, ou seja, uma banda total de 480 kbps (*fullduplex*). Se a rede for *half-duplex* a banda deve ser multiplicada por 2, resultando em uma banda de 960 kbps. Isto se deve ao *overhead* (bits que fazem parte do pacote IP, mas que são descartados após a retirada dos dados principais) dos pacotes IP utilizados para a sinalização do protocolo (Karapetkov, 1999);
- **Atraso de transmissão ou latência.** Nenhuma rede transmite bits instantaneamente. Há um tempo de transmissão dos dados do emissor até chegar no receptor. Este tempo é a latência da rede. A latência é introduzida pelo processo de codificação e decodificação (também dependendo do equipamento a ser

utilizado) e também do tempo que se leva para os dados “atravessarem” a rede. Pouco se pode fazer localmente para mudar o atraso na rede em larga escala a não ser negociar diretamente com a empresa provedora de serviços de telecomunicações. Em um ambiente ótimo, uma conexão dos EUA até a Europa sobre fibra ótica introduz 90ms de latência. O excesso de latência gera o efeito de conversações “atravessadas” (ou truncadas) em uma videoconferência. Em chamadas com latência menor do que 50ms esse efeito não é tão grave, mas se torna bem aparente com latências maiores do que 150ms (VIDE, 2004).

Pode existir também a latência que entre os fluxos (*streams*) de áudio e vídeo. Se o fluxo de áudio chegar 30 ms na frente do fluxo de vídeo percebe-se uma falta de sincronismo entre vídeo e áudio. Para um tempo maior que 40 ms tem-se um efeito considerável e incômodo, chamado de *lip sync*. Em redes com aplicações de tempo real a latência total não dever exceder 300 ms (Karapetkov, 1999);

- **Variação do atraso ou *Jitter*.** A técnica de se medir a variação do atraso consiste em avaliar a diferença entre o maior e o menor atraso de transmissão sobre um período de tempo. Esta diferença de valor é chamada de *Jitter*. Ele refere-se à variação do atraso gerado pela rede e por seus elementos ativos. O *jitter* também ocorre em redes comutadas por circuitos, mas está tipicamente na ordem de nanossegundos em ligações diretas de fibra ótica e em microssegundos para circuitos de longa distância. Em redes IP, o *jitter* pode ser na ordem de milissegundos ou valores maiores, dependendo da capacidade e do tráfego da rede naquele instante. Para uma videoconferência com qualidade, o *jitter* não deveria ultrapassar 40ms (Polycom, 2006). A Figura 3.9 mostra um exemplo de latência e *jitter* em uma rede IP.

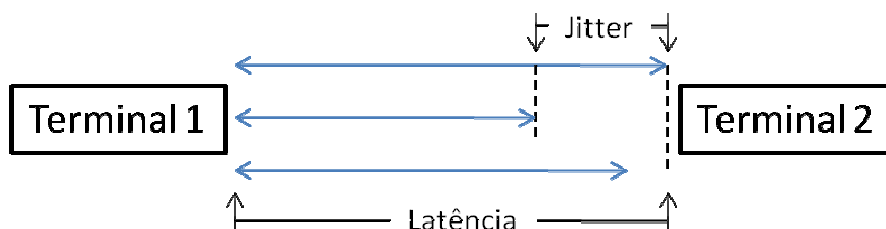


Figura 3.9 – Ilustração de latência e *jitter* (modificado – Karapetkov, 1999)

- **Isocronismo.** Refere-se à capacidade da rede de garantir uma taxa de transmissão de dados e assegurar que o valor do *jitter* seja pequeno e constante (Radvision, 2001);
- **Taxa de erro.** Característica que a rede tem em manter a integridade dos dados transmitidos. Diz respeito à transmissão dos dados sem alteração, perda, duplicação ou fora de ordem ou de sincronismo (Radvision, 2001).

A Tabela 3-3 mostra um resumo dos parâmetros desejáveis para garantir qualidade de voz e vídeo em redes corporativas.

Tabela 3-3 – Parâmetros para comunicação de áudio e vídeo

Parâmetro	Valor
Perda de pacotes	< 0,1 %
Latência	<=150 ms
<i>Jitter</i>	< 40 ms

A perda de pacotes é tipicamente o parâmetro mais difícil de ser controlado. Vários fabricantes indicam que podem entregar boa qualidade de áudio e vídeo com uma taxa de perda de pacotes maior do que 0.1%. Cada fabricante usa um algoritmo para maquiagem o efeito da perda de informação. Quando um pacote é perdido, perde-se informação, logo há degradação de qualidade, então os fabricantes utilizam algoritmos que fazem deduções sobre a informação perdida para recuperar a imagem (Polycom, 2006). Apesar de existirem essas soluções, a melhor opção é não perder pacotes e se utilizar estes algoritmos como mecanismo de backup.

Avaliando estes fatores, de alguma maneira a rede deve garantir que os pacotes associados à voz e vídeo transitem pela rede em tempo hábil, sem se perderem e sem a ajuda de um protocolo de transporte. QoS (*Quality of Service*) é um termo que referencia um conjunto de mecanismos utilizados para tentar garantir a qualidade aceitável de um serviço especificado na rede. Neste caso das aplicações de áudio e vídeo estes mecanismos dão prioridade para os fluxos com a intenção de garantir que estes sejam entregues no tempo adequado e corretamente.

3.3.4 - QoS – Quality of Service

Nem todas as empresas utilizam os mesmos mecanismos para a implementação de QoS em videoconferência. Logo, nem sempre a qualidade é garantida. Existem alguns mecanismos que são mais utilizados, tornando-se necessários devido às redes não confiáveis e limitações do controle de uso da banda feito pelo *Gatekeeper*. Os métodos mais conhecidos são o uso do campo ToS (*Type of Service*) do cabeçalho do pacotes IP, reserva de caminhos na rede para determinado tráfego e a aplicação de QoS na camada 2 conhecida padronizada pelo IEEE 802.1p¹.

A seguir é feita uma descrição sobre as limitações do *Gatekeeper* em prover qualidade de serviço. Nas sessões seguintes são apresentados de forma sucinta, os mecanismos de QoS mais utilizados para videoconferência.

3.3.4.1 - QoS no *Gatekeeper*

O objetivo da videoconferência é prover um serviço de áudio e vídeo de qualidade para os usuários. Um dos desafios para se atingir esse objetivo é conseguir os recursos suficientes da rede. Atualmente o *gatekeeper* é o dispositivo disponível, na solução H.323, para se conseguir qualidade de serviço em uma rede baseada em IP.

A maior parte dos *gatekeepers* executa o controle de banda através de processos de admissão, ou seja, as chamadas são aceitas ou rejeitadas pelo *gatekeeper*. Já que o *gatekeeper* não sabe sobre os pacotes que estão trafegando na rede tampouco sobre os que não fazem parte de sua zona de gerenciamento, a banda disponível pode existir de fato ou não. O *gatekeeper* não consegue avaliar na verdade o quanto existe de disponibilidade de recursos na rede.

O controle do *gatekeeper* se baseia em um *pool* (valor) fixo configurado pelo administrador da rede. A cada ligação que ocorre, a banda a ser utilizada é subtraída deste valor total. Quando a chamada é finalizada, o valor referente à banda utilizada é somado novamente. Se não existir banda suficiente para corresponder ao requerido por uma nova chamada, esta é rejeitada. Os *pools* podem ser configurados para controlar a banda dentro de uma zona, entre sites diferentes dentro da zona ou mesmo entre zonas diferentes.

Esse mecanismo de controle de banda baseado em admissão não é o mais adequado para garantir qualidade de serviço:

¹ www.ieee.org

- O controle de banda é estático – As decisões sobre admissão de chamadas são baseadas em um valor pré-configurado. O *gatekeeper* não tem conhecimento dos recursos que estão sendo utilizados por chamadas que fluem na rede sem terem sido admitidas por ele.
- *Gatekeepers* só reconhecem tráfego H.323 – O *gatekeeper* não é capaz de reconhecer outro tráfego que não seja o H.323. Logo, as decisões feitas podem não coincidir com a disponibilidade real de banda na rede.
- *Gatekeepers* não conseguem fazer policiamento – Um terminal pode fazer um pedido de chamada para o *gatekeeper* requisitando 384kbps e enviar 400kbps ou mais. Se a sinalização de controle não passar pelo *gatekeeper* durante uma chamada, este não pode forçar uma redução da banda.

As limitações para o uso de QoS somente com *gatekeeper* fazem com que sejam necessárias outras ferramentas de QoS mais eficientes. Atualmente essas ferramentas são implementadas na maioria dos fabricantes de equipamentos de VoIP e videoconferência. As ferramentas mais conhecidas são o *Differentiated Services* (DiffServ) (IETF RFC 2475, 1998) e o *Integrated Services* (IntServ) (IETF RFC 1633, 1994). Estas podem ser usadas separadamente, porém são mais utilizadas em conjunto para garantir os objetivos da implementação.

3.3.4.2 - DiffServ e IP Precedence

O DiffServ utiliza um campo no cabeçalho do protocolo IP denominado *Type of Service* (ToS) (IETF RFC 2475, 1998). Terminais configurados para enviar tráfego que necessita de QoS, quando em uma rede que suporta DiffServ, marcam cada pacote enviado com um valor determinado. Pacotes de aplicações diferentes podem ser marcados com valores distintos. Esses valores são utilizados para classificar o tráfego em classes de serviço diferentes. De acordo com a classe escolhida, os pacotes são enviados, enfileirados e priorizados pelos roteadores.

Em resumo, essa classe de ferramentas oferece um modelo onde o tráfego de dados é tratado pelos sistemas intermediários (roteadores e switches) de acordo com prioridades relativas configuradas no campo ToS. Definida pelas RFCs 2474 (IETF RFC 2474, 1998) e 2475 (IETF RFC 2575, 1998), o DiffServ se sobrepõe à especificação original para priorização de pacotes com o uso do campo ToS, o IP Precedence da RFC 791 (IETF RFC

791, 1981). No DiffServ são adicionados mais níveis de prioridade para a marcação dos pacotes em relação ao IP Precedence.

O DiffServ define o campo DiffServ (DS Field), que substitui o campo ToS (usado primeiramente pelo IP Precedence). Sua função é prover um comportamento nó-a-nó (*Per-hop-behavior* – PHB), ou seja, todos os elementos da rota utilizada devem ser configurados com filas separadas de acordo com alguns parâmetros como prioridade, tamanho, pesos e algoritmo de rejeição (IETF RFC 2475, 1981). A configuração para suportar essa classificação é necessária em cada nó (roteador), pois um pacote marcado não receberá tratamento especial se o roteador não estiver configurado para reconhecê-lo.

A Figura 3.10 mostra o cabeçalho IP onde observa-se a localização do campo *Type of Service*.

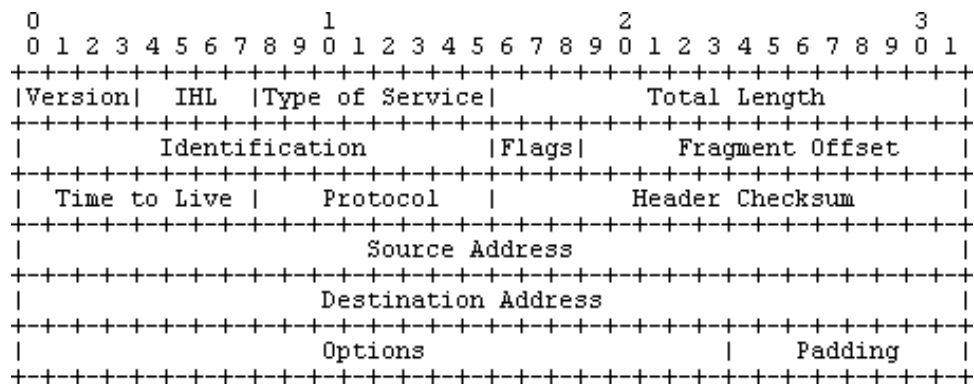


Figura 3.10 – Cabeçalho do protocolo IP (IETF RFC 791, 1981)

O campo ToS, quando usado para IP Precedence utiliza 3 bits para o tipo de serviço (suportado por equipamentos de VC como mostra a Figura 3.11). Um serviço de classe mais alta recebe maior prioridade e o contrário ocorre para marcações mais baixas.

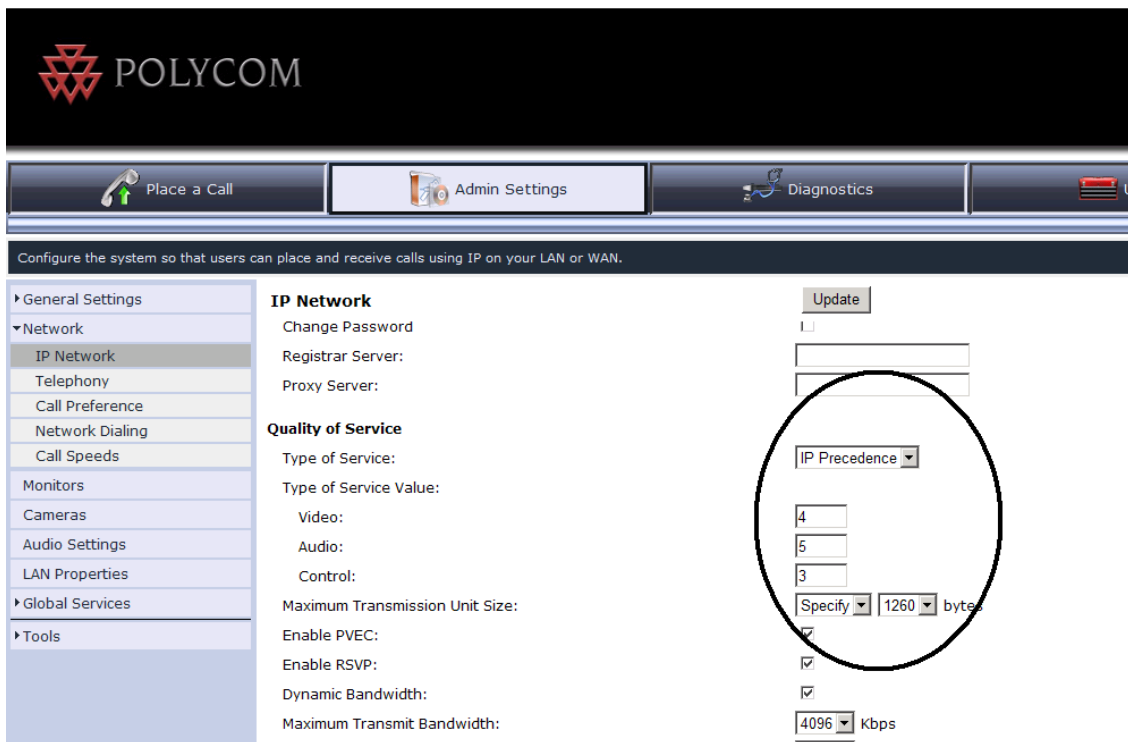


Figura 3.11 – Exemplo de localização da configuração dos parâmetros de IP Precedence

O DiffServ define o campo DSCP (*Differentiated Services Code Point*) que utiliza 6 bits e amplia os campos que antes eram utilizados pelo IP Precedence que utilizavam 3 bits. A ferramenta foi desenvolvida dessa forma para manter uma compatibilidade com o modelo IP Precedence. A Tabela 3-4 mostra um comparativo entre os valores usados para o IP Precedence e o DSCP. Os três bits mais significativos continuam valendo para o IP Precedence e os três bits menos significativos estendem os valores para uso do DSCP. Grande parte dos equipamentos de videoconferência e clientes SIP são compatíveis também com esse modo de QoS.

Tabela 3-4 – Equivalência entre os valores IP Precedence e DSCP (O’Neil, 2002)

Valor IP Precedence	DSCP
0 (000000)	0-7 (000000 - 000111)
1	8-15
2	16-23
3	24-31
4	32-39
5	40-47
6	48-55
7	56-63

3.3.4.3 - IntServ -RSVP

Este modelo IntServ de qualidade de serviço é caracterizado essencialmente pela reserva de recursos (largura de banda, atraso e *jitter*), antes do estabelecimento da comunicação. Este serviço utiliza o protocolo de sinalização RSVP (*Resource ReSerVation Protocol*) que permite que as próprias aplicações solicitem da rede as reservas dos recursos necessários para seus diversos serviços. Na sinalização RSVP existe troca de mensagens de controle entre emissor e receptor de forma que seja possível negociar uma faixa da largura de banda para a transmissão dos dados através do caminho a ser utilizado.

O RSVP é usado para gerenciar recursos ao longo do caminho pelo qual deseja-se utilizar aplicações que necessitem de QoS. Ele não realiza transporte de dados, é apenas um protocolo de sinalização que atua juntamente com o ICMP (*Internet Control Management Protocol*) (IETF RFC 762, 1980) e o IGMP (*Internet Group Management Protocol*) (IETF RFC 2236, 1997). O processo de sinalização se dá antes da transmissão dos dados e é renovado sempre que necessário. Uma série de mensagens (ou objetos do RSVP) deve ser trocada entre as aplicações e os elementos de rede para corretamente requisitar a qualidade de serviço para uma determinada sessão, são elas: PATH e RESV (IETF RCF 2205, 1997).

A Figura 3.12 ajuda a mostrar um exemplo simplificado da forma como o protocolo funciona.

- A mensagem PATH é enviada pelo transmissor, informando para o endereço de destino (em *unicast* ou *multicast*) especificações de tráfego, isto é, largura de banda, latência e *jitter*. Cada roteador ao longo do caminho, com RSVP habilitado estabelece, então, um PATH-state. As mensagens desse tipo armazenam o estado de cada nó (cada roteador), por onde ela transitou.
- Quando a mensagem PATH chega no receptor, este analisa as informações contidas na mensagem e seleciona os parâmetros de reserva desejados. Iniciando-se assim o procedimento de reserva de recursos através da mensagem RESV. Essa mensagem é enviada de volta aos mesmos roteadores estabelecendo um RESV-state. Além das especificações do tráfego contidas no PATH-state, o RESV contém especificações de pedido, que indica qual tipo de serviço, dentro do IntServ, está sendo requerido (carga controlada ou serviço garantido) e especificações de filtro, que caracterizam os pacotes para os quais a reserva está sendo feita.
- Cada roteador ao longo do caminho pode aceitar ou não as requisições da mensagem RESV. Se a mensagem for rejeitada, o roteador envia uma mensagem de erro para o receptor e o processo de sinalização termina.

Quando o último roteador recebe a mensagem RESV, inicia-se então a comunicação propriamente dita. Neste caso o último roteador o mais próximo do transmissor. O processo de tráfego de mensagens é resumido na Figura 3.12.

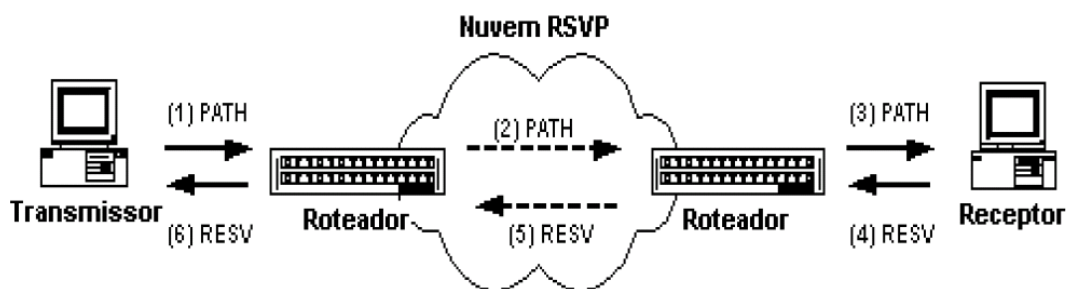


Figura 3.12 – Mensagens do RSVP (RNP, 2000)

O RSVP pode ser implementado tanto para transmissão unicast quanto multicast. Como já foi dito anteriormente o RSVP trabalha em conjunto com outros componentes (controle de

admissão, classificador e escalonador de pacotes) que atuam tanto nos elementos finais, transmissores e receptores como nos roteadores. Por essa razão é tipicamente implementado em conjunto com as soluções de DiffServ.

A utilização do RSVP com o protocolo H.323 para obter QoS está descrita no apêndice da recomendação H.323 desde a versão 2 (atualmente a recomendação está na versão 6). Para suportar o uso do RSVP, vários parâmetros são disponibilizados e várias formas de se estabelecer o “túnel” RSVP são disponíveis. Dentre os parâmetros estão os que definem quem deve iniciar o pedido de reserva e quais são os requerimentos de rede para reserva (Dhesikan, 2002).

A reserva RSVP pode ser iniciada tanto por um terminal (terminal, gateway ou MCU) quanto pelo *gatekeeper*, sendo que as mensagens de reserva devem trafegar pelo mesmo caminho do *stream* de mídia. Vale observar que uma vez que os *gatekeepers* não trafegam mídia e também podem estar em uma rede diferente de onde estão os terminais, é comum que o pedido RSVP parta do terminal para garantir que a reserva usará o mesmo caminho que os pacotes e áudio e vídeo.

Para que um terminal possa iniciar a reserva, este deve pedir permissão para o *gatekeeper* especificando sua intenção na mensagem de ARQ (*Admission Request*). O *gatekeeper* então, de acordo com sua configuração rejeita ou permite a chamada (ITU-T H.323, 1996). Usualmente as reservas são feitas somente para os canais de áudio e vídeo. Os outros canais abertos em uma comunicação, conteúdo e FECC (*Far-end Camera Control Protocol*), podem tipicamente trabalhar com melhor esforço. Como são abertos canais de envio em ambos terminais H.323, é necessário estabelecer reservas RSVP em ambas as direções da comunicação.

As reservas são feitas somente após a negociação de portas e capacidades realizadas entre os terminais. Isso ocorre porque a reserva depende dessas informações para “pedir” os recursos adequados de cada nó.

Para o caso do protocolo SIP, as implementações para uso de RSVP são tratadas tipicamente nos *proxies* junto de um novo elemento denominado RSVP server, sendo que esta integração Proxy/RSVP Server é transparente para o usuário final. A informação de requerimento de banda trafegada pelo SDP é enviada pelo RSVP server que se encarrega de estabelecer a reserva necessária (Jung, Allamiyarov, Seo, Kim, 2005).

3.3.4.4 - *Class of Service* (CoS) IEEE802.1p

O CoS, ou IEEE802.1p¹ é um tipo de serviço que utiliza 3 bits do cabeçalho de um frame ethernet, quando este se utiliza de outro padrão, o IEEE802.1Q, usado para definir Virtual LANs (VLANs). É uma maneira de se gerenciar tráfego de dados em uma rede através do agrupamento de tipos similares de tráfego (ex.: vídeo, áudio, transferências de arquivos) e tratá-los como classes diferentes com níveis de prioridades distintos. O mecanismo é semelhante ao DiffServ ou IP Precedence, mas neste caso o tratamento é feito na camada de enlace (camada 2) do modelo OSI. É importante ressaltar mais uma vez que as tecnologias de QoS que utilizam classes de serviços não garantem o nível de serviço desejado, mas priorizam a tentativa de melhor esforço.

O IEEE 802.1p rotula cada quadro Ethernet com três bits reservados para a prioridade do quadro. Esses bits estão localizados no campo TAG, especificado pela norma IEEE 802.3ac¹, comentada adiante. É importante não confundir esses três bits com os três bits de precedência do cabeçalho IP mostrados nas sessões anteriores: os três que carregam a prioridade do quadro Ethernet estão no cabeçalho MAC (Ethernet) na camada de Enlace. Os quadros marcados (*tagged*) têm sua prioridade explícita. A marcação não deriva do endereço MAC de origem ou do endereço MAC de destino, nem é computada através de informações retiradas do quadro. Mas é explicitamente definida no campo reservado para essa finalidade de acordo com o tipo de dado trafegado.

Os elementos de uma rede que trabalham na camada de enlace do modelo OSI são os switches. Para fazerem uso do CoS, é necessário que os switches da rede em questão tenham algum mecanismo para controlar QoS bem como algoritmos de filtragem em caso de congestionamento. Ou seja, os switches devem implementar filas separadas, com políticas de encaminhamento específicas para quadros com prioridades diferentes, e conseqüentemente com necessidades de QoS também diferentes (Niclas, 1999).

Apesar de todas as padronizações propostas e homologadas, o 802.1p e 802.1Q não poderiam ser efetivados na prática se não fosse feita uma proposta de modificação do quadro Ethernet. O campo necessário para estas implementações não está definido na especificação inicial do Ethernet. Este campo seria o responsável pela identificação do quadro para uma

¹ www.ieee.org

determinada VLAN e pela identificação de qual prioridade este quadro teria diante de outros na fila de encaminhamento.

O IEEE 802.3ac, resolve esta questão propondo a adição de quatro octetos ao quadro Ethernet. Esses quatro octetos são os bits responsáveis tanto pela identificação da VLAN da norma IEEE 802.1Q, quanto pela marcação da prioridade do quadro, referente à norma IEEE 802.1p. Logo, as duas implementações compartilham o mesmo campo. O novo campo, denominado TAG, foi inserido logo após o campo de Endereço de Origem e antes do campo denominado "EtherType" (para Ethernet v.2), ou Tamanho do Quadro (para IEEE 802.3), como mostra a Figura 3.13.

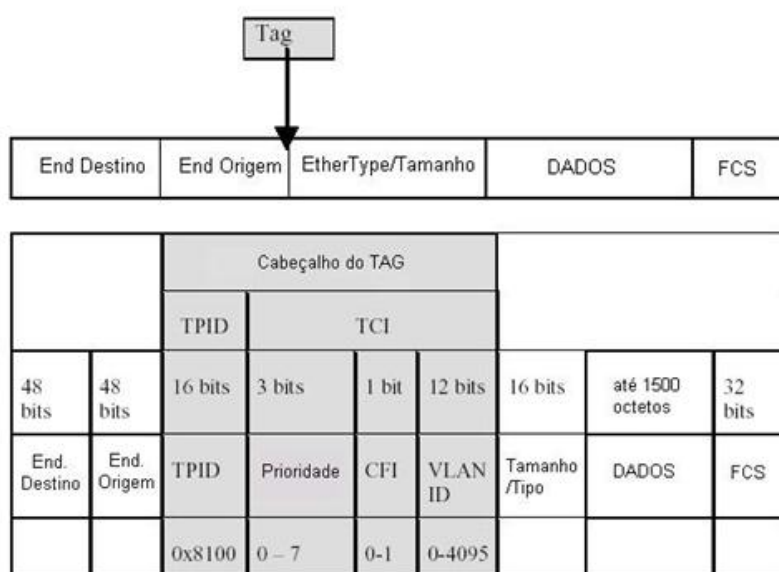


Figura 3.13 – Campo TAG, 32 bits adicionados ao frame Ethernet (GTA/UFRJ, 2002)

As sete classes de tráfego são mostradas na Tabela 3-5.

Tabela 3-5 – Classes de tráfego usadas pelo CoS

Network Control	7	Usado para manter a rede em operação. O controle de tráfego de rede deve ter prioridade máxima
Voice	6	Tráfego que requer menos do que 10 ms de delay e jitter
Video	5	Tráfego que requer menos do que 100 ms de delay
Controlled Load	4	Usado para aplicações de negócios que exigem certo nível de controle de admissão.
Excellent effort	3	Tratado como o melhor uso do “melhor-esforço”
Best Effort	1	Tráfego que se utiliza na maioria das redes. Ex: tráfego de arquivos em uma LAN
Background	0	Tráfego que pode estar na rede e que não impacta relevantemente no uso dos recursos. Ex, tráfego esporádico em rajadas.

A Figura 3.14 mostra uma tela de configuração de um equipamento de Videoconferência. Percebe-se que a opção IEEE 802.1p/Q é disponível e ao selecioná-la, é necessário especificar tanto o valor da VLAN quando a prioridade dos frames.

The screenshot displays the 'LAN Properties' configuration page. A sidebar on the left contains a menu with 'LAN Properties' selected. The main content area includes an 'Update' button and several input fields: 'Site-Local', 'Global Address', 'Default Gateway', 'Host Name' (pre-filled with 'demo-spo-hdx4k'), and 'Domain Name' (pre-filled with 'polycom.com'). A yellow box highlights the 'Enable 802.1p/Q' section, which contains a checked checkbox, a 'VLAN ID' field with the value '1', and three dropdown menus for 'Video Priority', 'Audio Priority', and 'Control Priority', all set to '0'. Below this section, there are checkboxes for 'Enable EAP/802.1X' and 'Enable 802.1p/Q'. A warning message at the bottom of the highlighted area reads 'Any changes made to this page will cause...'.

Figura 3.14 – Configuração dos parâmetros de VLAN e CoS

As Figuras Figura 3.11 e Figura 3.14 representam exemplos de configuração de QoS realizadas nos próprios terminais envolvidos em uma comunicação. No entanto, somente marcar as opções adequadas não garante que pacotes sejam priorizados na rede ou que melhorias na qualidade das comunicações sejam percebidas. Estas configurações são somente o último passo realizado após a preparação da uma rede para suportar QoS.

Cada ferramenta de QoS atua em domínios específicos delimitados pela sua área de atuação e pela sua compatibilidade com a camada suportada pelos ativos de rede. O QoS implementado no gatekeeper, por exemplo está na camada 7 do modelo OSI. Sua área de atuação são os terminais dentro de uma zona H.323. O controle de acesso ao uso de recursos de rede requer que a infraestrutura de rede esteja adequadamente preparada para o tráfego de vídeo.

O método IEEE 802.1p ou CoS refere-se ao uso de 3 bits na camada 2 do modelo OSI para que os switches ou até mesmo roteadores de camada 2 sejam capazes de priorizar o tráfego de dados de acordo com a marcação configurada. A marcação dos bits pode ocorrer no

terminal, de acordo com a configuração determinada pelo gerente de rede, ou pode ocorrer nos switches. Neste segundo caso, de acordo com a análise de alguns campos pré-determinados, o switch pode marcar os pacotes apropriadamente. Uma vez marcados, os pacotes serão priorizados pelos switches da rede.

Enquanto a especificação 802.1p é efetiva na LAN, ela não provê QoS fim a fim quando o tráfego de dados atravessa diferentes redes ou mesmo a Internet. Essa limitação decorre do fato de esta especificação estar presente na camada 2 do modelo OSI. O roteamento entre redes é avaliado através da camada 3 ou outro mecanismo que depende de informações do protocolo IP (camada 3).

Para extrapolar esses limites, o Diffserv pode ser usado para que as informação de prioridade dos pacotes estejam presentes na camada 3. Dessa forma a priorização de pacotes pode ser interpretada desde o roteador na borda da rede local até o destino, passando por roteadores diversos. Todos os roteadores que fazem parte do caminho por onde os pacotes trafegam devem ser capazes de reconhecer a classificação dos pacotes e incluí-los na fila adequada. Somente dessa forma a marcação realizada no terminal ou mesmo no roteador de borda terá efeito.

Para que os requerimentos de QoS possam ser atendidos tanto na LAN quanto entre LANs, pode-se contar com roteadores na borda da rede capazes de traduzir a classificação do IEEE 802.1p para Diffserv. Dessa forma, a priorização de pacotes vindas da camada 2 é transferida para outros roteadores de camada 3.

Sendo assim, conclui-se que é possível definir “camadas” de QoS para que o tráfego seja adequadamente priorizado dentro de uma mesma rede ou entre redes distintas.

3.4 - REDES MPLS

O *Multiprotocol Label Switching* (MPLS) (IETF RFC 3032, 2001) foi desenvolvido na intenção de resolver alguns problemas inerentes às redes IP. Quando usada para trafegar dados IP, seu funcionamento básico se dá por meio da adição de um rótulo nos pacotes IP na entrada do *backbone*. A partir deste ponto, todo o encaminhamento passa a ser feito com base nesse rótulo e não mais no endereço IP, simplificando assim o processo de roteamento. Os rótulos MPLS são divulgados entre os roteadores para que estes possam formar seu mapeamento de rótulo-a-rótulo dentro de todo o *backbone*. Antes de ser entregue à rede do

cliente o rótulo é retirado preservando o IP de destino para que seja usado após sair do *backbone* MPLS.

Essa rede busca evitar que na transmissão de pacotes ocorram perdas ou atrasos perceptíveis, para isso se utiliza de uma gestão de tráfego mais eficaz que possibilita melhor desempenho aos serviços. Além disso, oferece a possibilidade de uso de QoS (*Quality of Service*), priorizando tráfego de aplicações críticas e distribuindo de forma mais eficiente o tráfego entre os diversos caminhos disponíveis. Criam-se, dessa forma, as condições necessárias para o melhor uso dos recursos da rede permitindo ainda o fluxo de áudio e vídeo. Apesar de ter sido desenvolvido para redes com camada IP (camada 3) e de enlace ATM (*Asynchronous transfer mode*) (camada 2), o mecanismo de encaminhamento dos pacotes no MPLS pode ser utilizado para quaisquer outras combinações de protocolos de rede e de enlace, o que explica o nome de *Multiprotocol Label Switching*.

O elemento mais importante de uma arquitetura MPLS é o rótulo. A comutação por rótulos faz com que não sejam mais tratados pacotes IPV4, IPV6 nem mesmo quadros de camada 2, e sim pacotes rotulados independente do que está abaixo do rótulo.

A figura abaixo representa a estrutura de um rótulo.

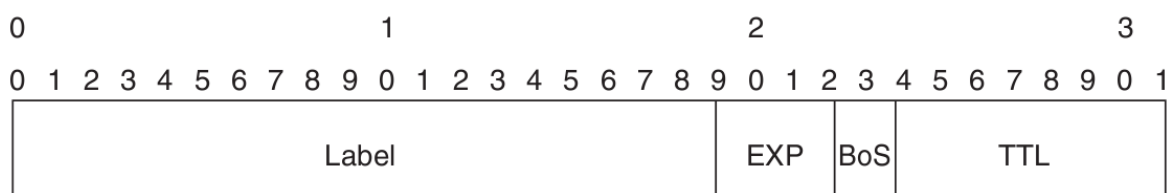


Figura 3.15 – Sintaxe de um rótulo MPLS (IETF RFC 3032, 2001)

Os primeiros 20 bits são de fato o valor do *label* que pode estar entre 0 e 1.048.575. Os bits 20 a 22 são conhecidos como bits experimentais (EXP) por questões históricas visto que em momento de desenvolvimento, não se havia definido a utilização dada a esta informação. Hoje estes bits são utilizados somente para prover qualidade de serviço. O bit 23 é o bit de final de pilha ou “*Bottom of Stack (BoS)*”. Uma pilha é um conjunto de rótulos inseridos no topo de um pacote, cada um com seu propósito, para serem interpretados nos devidos nós de rede. Os bits 24 a 31 são utilizados para definir o tempo de vida do pacote e possui a mesma função do TTL (*Time to Live*) encontrado no cabeçalho IP, ou seja evitar um loop de roteamento.

O MPLS não pode ser considerado um protocolo de camada 2, visto que o encapsulamento de camada 2 ainda está presente mesmo em pacotes rotulados, nem é exatamente um protocolo de camada 3 pois o mesmo também está presente. Logo, a melhor forma de se enxergar o MPLS é como um protocolo de camada 2,5 .

Além dos rótulos, um papel importante para o bom funcionamento de uma rede MPLS é desempenhado pelo protocolo de distribuição de rótulos como o LDP (*Label Distribution Protocol*). Ele é responsável por distribuir os rótulos entre os roteadores definindo assim os LSPs (*Label Switching Paths*). O LDP exerce função análoga aos protocolos de roteamento de camada 3, orientando o caminho a ser percorrido pelos pacotes, porém cada rótulo distribuído pelo LDP possui significado apenas local, ou seja, ao contrário do IP, o rótulo de destino muda a cada nó de rede.

3.4.1 - Vantagens do MPLS:

1. Maior rapidez no encaminhamento:

No modelo convencional de roteamento de pacotes (IP) cada roteador da rede precisa executar um algoritmo de camada 3 para determinar como será feito o encaminhamento dos pacotes. Enquanto o pacote transita pela rede cada roteador precisa tomar uma decisão independente sobre qual será o próximo enlace a ser utilizado. Esse algoritmo basicamente lê o endereço IP de destino mais a máscara e compara linha a linha com sua tabela de roteamento até que o cálculo do campo de rede do endereço coincida com alguma entrada da tabela. Esta análise e classificação do cabeçalho do pacote nível 3, dependendo da complexidade da rede (número de rotas e rotas alternativas) pode demandar muitos recursos do processador. Em redes de datagramas, este processo ocorre para cada pacote em cada nó entre a origem e o destino.

Já em um ambiente MPLS, as rotas otimizadas entre pontos da rede são identificadas previamente. No momento da entrada de um pacote em um *backbone* MPLS, este recebe um rótulo (correspondente ao caminho a ser seguido pelo pacote) e a partir de então será encaminhado com base nesta informação e não mais em seu endereço de destino camada 3. Esta tradução de um endereço IP em um rótulo não leva em consideração apenas o endereço de destino, mas todas as peculiaridades referentes à classe de serviço à qual pertence o tráfego, ou seja, se há alguma definição de QoS, VPN, etc. Os roteadores ao longo do

caminho utilizam a informação do rótulo para determinar qual o próximo nó que deve receber o pacote. Como estes equipamentos manipulam exclusivamente a informação do rótulo, a análise e classificação do cabeçalho do nível 3 ocorre exclusivamente no ponto de ingresso da rede (roteadores de borda) facilitando assim o processo de encaminhamento dentro do *backbone*. Esse modelo traz a vantagem de tornar mais rápido o encaminhamento diminuindo de forma generalizada o tempo de tráfego dos pacotes no *backbone* independente se este possui garantia de qualidade de serviço implementada ou não.

2. Engenharia de Tráfego e Qualidade de Serviço

A engenharia de tráfego é o processo de seleção de caminhos na rede visando à obtenção de padrões de tráfego que utilizem os recursos de forma mais eficiente e balanceada.

O roteamento convencional baseado em algoritmos IGP (*Interior Gateway Protocol*) (IETF RFC 1371, 1992) tende a selecionar caminhos na rede que acabam por não resultar em uma utilização balanceada, pois normalmente o tráfego é sempre enviado por caminhos específicos a não ser que estes se tornem indisponíveis. Dessa forma, alguns recursos podem ser subutilizados enquanto outros podem ser penalizados por cargas excessivas de tráfego, diminuindo a eficiência e aproveitamento de caminhos redundantes de rede.

Um método limitado de engenharia de tráfego pode ser usado manipulando as métricas do IGP associadas com os enlaces da rede, entretanto estas informações são complicadas de administrar em ambientes com várias opções de caminhos entre dois pontos.

O MPLS pode ser utilizado em conjunto com os algoritmos IGP para implementar a engenharia de tráfego com as informações sobre rotas específicas a serem seguidas por um pacote na rede. Este roteamento explícito de pacotes, definido pela engenharia de tráfego, faz com que todo o fluxo de informações que possuam as mesmas características de tráfego parametrizadas pelo administrador da rede transite pelo mesmo caminho. A administração e monitoramento dos fluxos de dados facilitam a avaliação adequada da utilização dos recursos da rede. Logo, em termos práticos, se uma rede possui 2 enlaces de 1 Gbps para o mesmo destino e apenas um é utilizado, devido a limitações de IGP, com a engenharia de tráfego dobra-se a capacidade de tráfego, já que os fluxos podem ser divididos entre os dois caminhos disponíveis (Ghein, 2007).

Associados à engenharia de tráfego podem ser implementados parâmetros de qualidade de serviço, provendo assim a capacidade de escolha inteligente de caminhos pela rede com base

nas necessidades específicas de cada classe de serviço. Esta característica agrega inteligência para administrar níveis de serviço diferentes de acordo com as políticas da rede. Em ambientes MPLS, o tratamento da Qualidade de Serviço para o roteamento é administrado de duas formas:

- a) O rótulo MPLS contém informações sobre a Classe de Serviço (EXP). Na medida em que tráfego flui na rede, esta informação é utilizada para priorizar o tráfego em cada nó.
- b) A rede MPLS pode estabelecer múltiplos caminhos entre equipamentos de entrada e saída. Para cada fluxo de informações é estabelecido um nível de serviço apropriado, e o tráfego é direcionado para o caminho adequado quando entra na rede.

Esta possibilidade de utilização de recursos da rede requer um nível de granularidade de informações superior ao utilizado pela engenharia de tráfego tradicional.

A implementação mais difundida e utilizada hoje da tecnologia MPLS é a VPN (*Virtual Private Network*). Esta aplicação aumenta em muito a escalabilidade das redes corporativas, visto que ao contrário das tecnologias como Frame-Relay e ATM, não existe a necessidade de conexões ponto a ponto entre as várias redes de um mesmo cliente.. Com MPLS, a VPN é configurada no roteador de borda do provedor em cada localidade do cliente e a comunicação entre eles segue uma topologia distribuída dentro da nuvem do provedor.

Considerando que o tráfego de uma determinada rede privada pode passar por diversos roteadores compartilhados, a segregação é feita por meio de tabelas de roteamento isoladas para cada VPN. Dessa forma, as informações de IGP de cada VPN são trocadas apenas entre os roteadores que se ligam diretamente à rede cliente, preservando-se a topologia do cliente e do *backbone*, posto que se estes não compartilham nenhuma informação de roteamento, podem até mesmo utilizar o endereçamento de camada 3 idênticos.

No que tange à videoconferência, este é um aspecto decisivo em momento de escolha de tecnologia de *backbone*, pois a VPN, por ser fechada entre os roteadores e não entre as aplicações de videoconferência, se torna transparente ao cliente, e além disso, dependendo da topologia adotada dispensa a utilização de NAT (*Network Address Translation*). Este aspecto facilita ainda a administração de redes de provedores de serviço de videoconferência, visto que não há preocupação quanto a sobreposição de endereços de camada 3 dos diversos

clientes, sendo esta informação encapsulada e traduzida em rótulos antes de ingressar o *backbone*.

4 - TENDÊNCIAS

Ao se tratar de redes de comunicações unificadas dois aspectos gerais devem ser levados em consideração. Estes devem ser observados e trabalhados em conjunto para que se obtenham os melhores resultados com a solução implementada.

No aspecto das aplicações e interface com o usuário final, deve ser avaliado o grau de complexidade da interação usuário-aplicação. Esta interação deve ser transparente e amigável o suficiente para que mesmo sem uma qualificação tecnológica qualquer pessoa consiga utilizar as funcionalidades disponíveis que desejar. Uma vez tendo o domínio sobre o uso da aplicação, a experiência do usuário deve ser satisfatória o suficiente para que a solução torne-se de fato uma ferramenta de comunicação eficiente.

No quesito infraestrutura, o desafio é projetar e gerenciar uma rede convergente de forma que as aplicações diversas possam coexistir. Para isto é necessário a compreensão do papel exercido por cada plataforma envolvida na comunicação e suas possíveis integrações. Com o uso das redes integradas, os dados de diferentes aplicações trafegam pelos mesmos ativos de rede demandando assim uma gerência centralizada e uma equipe com uma visão mais completa de toda a necessidade do negócio.

A evolução destes aspectos é complementar e contínua, sendo que à medida que a infraestrutura oferece mais capacidade o usuário demanda mais funcionalidades. Neste capítulo estes dois aspectos são tratados ressaltando suas condições no cenário atual e projeções futuras.

4.1 - PLATAFORMAS INTEGRADAS

Diversas organizações já perceberam o valor da migração de suas plataformas para redes IP. De acordo com uma pesquisa realizada em 2004 pelo Yankee Group (Kerravala, 2004) em organizações de diversos portes, 60% dos respondentes já haviam migrado para as redes IP e 30% planejavam fazer essa mudança em no máximo 2 anos. Os gerentes de TI compreendiam que essa migração facilita o gerenciamento, o desempenho e a confiabilidade da infraestrutura de TI.

Nos dias de hoje, muitas organizações são globais, interligadas, compreendendo empregados, fornecedores, parceiros e clientes. Da forma como trabalham atualmente, times de projetos, por exemplo, são criados envolvendo pessoas de diversas partes da organização e, possivelmente, em localidades distintas. As pessoas envolvidas estão sempre interagindo com outras e através de meios de comunicações diversos. Para permitir essa interação e colaboração dinâmica e distribuída, empresas implementaram uma variedade de plataformas de comunicação, como representado na Figura 4.1, no entanto, muitas vezes são tratadas como plataformas independentes de comunicação. Ou seja, as soluções não se comunicam ou o gerenciamento de cada uma independe das outras.

A comunicação colaborativa visa integrar essas aplicações em um meio compartilhado com foco nas capacidades de comunicações em tempo real. Dentro das organizações a capacidade de se utilizar vídeo e áudio conferência através de uma variedade de tipos de terminais, de forma integrada com as soluções de mensageria instantânea e também com as aplicações de negócios, provê maior usabilidade para os empregados. Implementadas sozinhas, as soluções oferecem benefícios isolados de produtividade. Quando entregues de forma unificada, alteram a forma de se trabalhar. Em algumas empresas a comunicação unificada já existe, mas dependem da interação do usuário com uma equipe técnica. Entregar a interação de soluções deste tipo nas mãos dos usuários finais é o ideal e facilita a adoção dessa tecnologia. No entanto, além de serem amigáveis, essa adoção depende das plataformas e infraestruturas de redes serem capazes de prover a capacidade necessária para uma experiência de qualidade.

Os aspectos apresentados no capítulo 3 devem ser estudados e levados em consideração para se estabelecer um projeto adequado capaz de suportar as aplicações de comunicação. Sendo o meio compartilhado com outras aplicações que não são multimídia, a rede de dados deve ser capaz de prover a qualidade adequada para cada tipo de tráfego. Dentre os aspectos que podem ser analisados para avaliar essa capacidade estão a qualidade dos enlaces utilizados, equipamentos de rede e telecomunicações, projeto de priorização de tráfego e segmentação da rede.

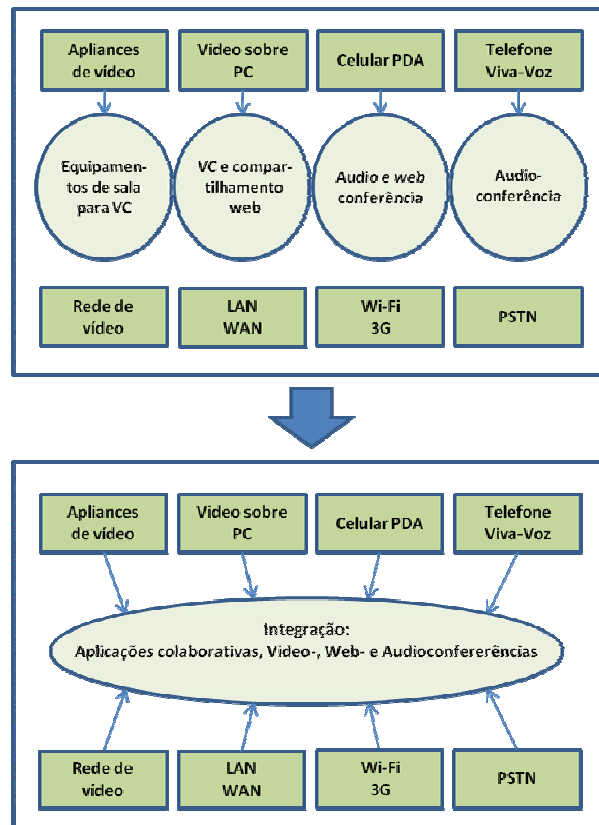


Figura 4.1 – Aplicações para comunicação rumo à comunicação colaborativa (Polycom, 2009)

Duas tendências estão mudando o mercado de comunicação áudio-visual, afetando as formas de uso e as tecnologias utilizadas. A primeira tendência é a mudança na forma de agendar conferências. Conferências de áudio ou vídeo inicialmente eram disponibilizadas para usuários através de eventos agendados e reservas de recursos, sendo que os fatores limitantes poderiam ser portas nas MCUs, banda disponível, canais ISDN disponíveis, etc. Os agendamentos estão perdendo espaço para as conferências ad-hoc (sob demanda), ou seja, aquelas que são criadas assim que se desejar realizar uma chamada.

As conferências de áudio já migraram para as conferências ad-hoc há algum tempo, sendo feitas sem a ajuda de operadores em 96% das ocorrências já em 2004 segundo o Yankee Group (Ghein, 2007). As conferências de vídeo, por sua vez, sempre foram em sua maioria agendadas. Fabricantes nos dias de hoje trabalham para que as conferências em vídeo sejam cada vez mais flexíveis e tenham mais recursos como presença e opções de acesso (de computadores ou celulares, por exemplo). Com opções de presença, basta clicar em uma sala ou participante, verificar sua disponibilidade e iniciar uma chamada.

A segunda tendência é a integração de aplicações, esta implica na integração das plataformas de vídeo com outras soluções. Como ressaltado, tecnologias de vídeo foram implementadas em redes independentes, como aplicações individuais separadas das aplicações principais de uma organização. As soluções de vídeo sempre tiveram, por exemplo, diretórios de usuários que eram diferentes dos usados pelas aplicações de TI, além de estarem em redes distintas e com equipes de gerenciamento independentes. Atualmente fabricantes e operadoras estão unindo as soluções de voz, vídeo e dados em plataformas integradas, como ilustrado na Figura 4.2. Para o usuário final, isso implica em uma maior extensão da funcionalidade destas aplicações. Visando atingir este objetivo, diversas soluções são possíveis.

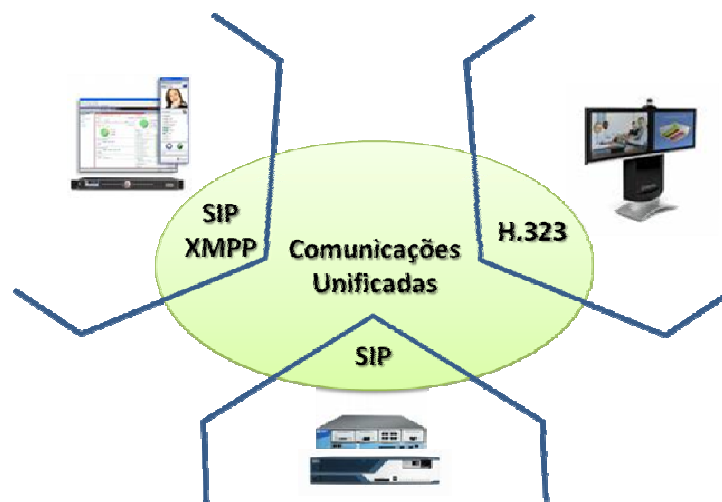


Figura 4.2 –Protocolos em direção à arquitetura unificada (Karapetkov, 2008)

Entre os fabricantes percebe-se também uma migração de padrões proprietários para padrões abertos de mercado e portanto, inter-operáveis¹. Dentre estes padrões abertos destacam-se os citados no capítulo 2. Como ressaltado anteriormente, a grande parte das redes de vídeo tradicionais são baseadas no protocolo H.323. No entanto, para implementação de serviços de voz (e em alguns incluindo vídeo) sobre redes IP, grandes fabricantes, em casos onde não implementam com o SIP, desenvolvem os seus próprios protocolos ou utilizam variações proprietárias do SIP. Por trabalharem com protocolos diferentes (H.323 ou SIP), essas soluções de áudio e vídeo não podiam se comunicar inicialmente.

¹ Apesar de adotarem padrões, em alguns casos pode-se perder algumas funcionalidades específicas quando equipamentos de fabricantes diferentes são conectados. Tipicamente as funcionalidades básicas são compatíveis.

O desafio técnico para que a unificação dessas plataformas possa ocorrer está na escolha de um protocolo capaz de comunicar todos esses elementos. Uma vez que o SIP é um protocolo de comunicação genérico, escalável, leve e desenvolvido para a Internet, este vem sendo utilizado como a principal solução para a unificação destas tecnologias. O IETF concebeu o SIP para ser um protocolo genérico capaz de estabelecer qualquer tipo de sessão, não somente áudio e vídeo. O SIP pode ser usado também para mensagens instantâneas e transferência de dados, por exemplo.

Nesse sentido, centrais telefônicas IP já estão se adaptando para serem inter-operáveis com o protocolo SIP ou mesmo mudando suas plataformas, antes proprietárias, para terem alguma compatibilidade com o SIP (Kerravala, 2004). Dessa forma, estarão baseadas em um protocolo padrão e aberto. Essa migração abre espaço para que as soluções de um fabricante sejam usadas em implementações com outros fabricantes e também com outras aplicações. A funcionalidade de presença é implementada sobre o SIP usando-o através do protocolo *SIP for Instant Messaging and Presence Leveraging Extensions* (SIMPLE).

No “mundo vídeo”, como dito, as aplicações em sua grande maioria funcionam usando o H.323, mas vários fabricantes estão adotando uma implementação *dual-stack* para terminais e MCUs. Dessa forma, os equipamentos conseguiriam trabalhar em ambas plataformas SIP e H.323. Esta aplicação é vista na sessão a seguir, sendo fundamental para as implantações de ambientes UC atualmente.

4.1.1 - Implementações

A Figura 4.3 mostra como pode ser implementada uma rede visando comunicações unificadas capaz de integrar aplicações em SIP, H.323 e PSTN/ISDN (para acesso à terminais ISDN ou ao serviço de telefonia tradicional). Nesta representação verificam-se vários elementos já mostrados nesta dissertação interconectados através de gateways e MCUs. O gateway, por definição, tem o papel de possibilitar a comunicação entre protocolos distintos. A MCU, além da função de ser um concentrador de chamadas para permitir chamadas multiponto, neste exemplo assume também uma funcionalidade semelhante ao gateway, é o equipamento que permite que equipamentos de protocolos diferentes participem de uma mesma chamada. Esta capacidade das MCUs de trabalhar com múltiplos protocolos não é obrigatória de acordo com a norma H.323, sendo um avanço em direção à UC.

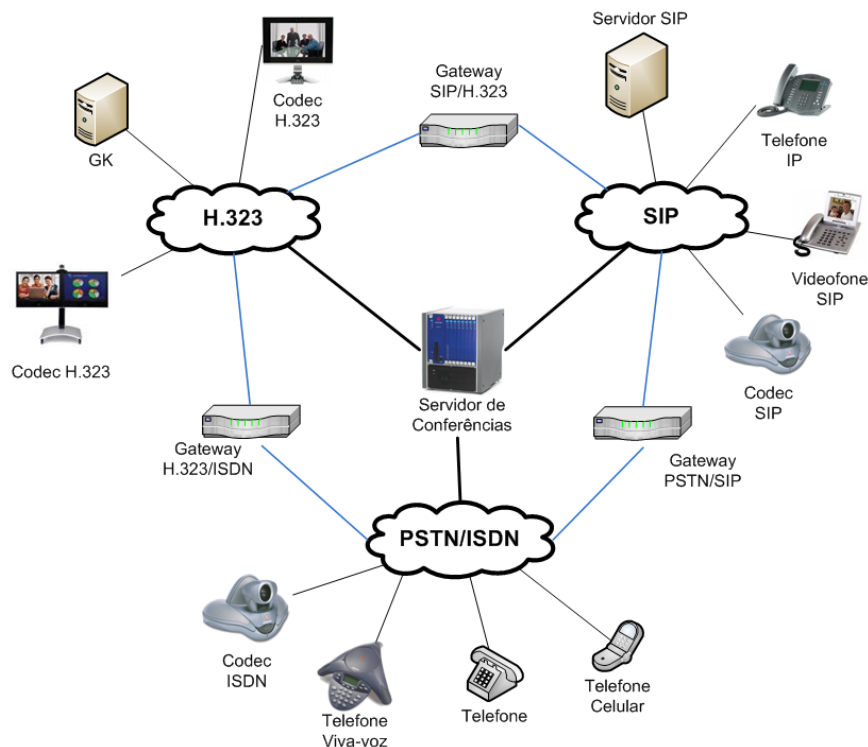


Figura 4.3 Integração entre plataformas usando MCUs e gateways

A Figura 4.3 representa os elementos principais de soluções em SIP, H.323 e ISDN/PSTN separadas de acordo com as plataformas sobre as quais cada um trabalha. As plataformas SIP e H.323 podem ser implementadas em uma mesma infra-estrutura de rede comutada por pacotes TCP/IP, compartilhando os seus recursos (switches, roteadores, etc.). Já a plataforma ISDN/PSTN possui requisitos diferentes de infra-estrutura uma vez que trabalha com a comutação por circuitos (modems, PBX, etc.).

A rede SIP é representada pelo Servidor SIP, onde os terminais telefônicos IP e terminais de vídeo se registram. Uma vez registrados, esses elementos podem usufruir das funcionalidades providas por esse servidor podendo utilizar o formato de endereçamento SIP, ramais e suas mensagens específicas. Tipicamente, em uma solução de telefonia SIP são implementados gateways para a rede de telefonia pública como apresentado na Figura 1.2.

O mesmo ocorre com a rede H.323, onde os terminais se registram ao *gatekeeper* (GK), podendo então se comunicar através de apelidos (*alias*) H.323. Os terminais H.323 podem obedecer à um plano de numeração estabelecido pelo administrador da zona H.323. Tipicamente a recomendação E.164, proveniente da telefonia convencional é usada como referência. O protocolo H.323 nativamente oferece uma série de funcionalidades para

videoconferência como controle de câmera remota e compartilhamento de conteúdo, muitas das quais ainda não estão disponíveis no protocolo SIP (Packetizer, 2009).

A rede ISDN/PSTN tem como elementos centralizadores os PABX que concentram diversas linhas telefônicas para controlar as chamadas realizadas ou fazer um melhor aproveitamento dos troncos de linhas contratados.

Neste ambiente se destacam os dois elementos usados para prover a integração entre as soluções, os gateways e MCUs, estes últimos mais conhecidos atualmente por *bridge* de conferências. Os gateways são conversores de protocolos e interfaces capazes de possibilitar a comunicação entre protocolos e tecnologias diferentes logo, no ambiente da Figura 4.3 podem ser utilizados gateways como ISDN/PSTN – H.323, ISDN/PSTN – SIP e também SIP – H.323. Através da conversão dos protocolos, os gateways permitem chamadas ponto-a-ponto entre terminais que usam estes protocolos.

Tomando como exemplo um gateway de sinalização SIP – H.323 pode-se resumir o seu funcionamento da seguinte maneira. As mensagens SIP chegam de um dispositivo SIP ao gateway, os parâmetros de comunicação são extraídos das mensagens SIP e utilizados para recriar mensagens equivalentes no protocolo H.323. O gateway então envia essas mensagens “traduzidas” para a linguagem H.323 para os terminais de destino. O mesmo acontece no sentido inverso H-323 para SIP. Logo, o gateway é visto como um agente de usuário para a rede SIP e como um terminal H.323 para a rede H.323.

Convenientemente, ambos SIP e H.323 utilizam os mesmos protocolos, RTP e RTCP, para trafegar áudio e vídeo deixando que o gateway possa focar principalmente na mediação da sinalização dos protocolos (Karapetkov, 2008). Este ponto é relevante para a implementação de gateways, pois é necessário um alto poder de processamento para lidar com o processamento de mídia. Enquanto o tráfego gerado pelos pacotes de sinalização pode alcançar alguns kbps, o tráfego de vídeo pode atingir alguns Mbps (é necessário por volta de 1Mbps para comunicações em alta definição com resoluções maiores ou iguais a 720p).

Os gateways ISDN/H.323 têm sido bastante utilizados para estender o uso da rede ISDN para equipamentos que só utilizam H.323. Podem ser usados como ponte para uma rede de backup caso a infra-estrutura IP falhe, como forma de aumentar o tempo de vida de redes legadas que só funcionam em ISDN e também para estender o uso da videoconferência para participantes que somente possam se conectar por um telefone convencional.

As mesmas aplicações citadas anteriormente valem para os gateways SIP – ISDN/PSTN, no entanto a aplicação mais comum é como forma de economizar custos para chamadas de longa distância. Como uma conexão ISDN/PSTN é cobrada por tempo de uso, diferente da rede IP, no intuito de reduzir custos é possível usar a rede IP para chegar até uma localidade distante e então, à partir deste ponto realizar uma chamada local na rede ISDN, como mostra a Figura 4.4.

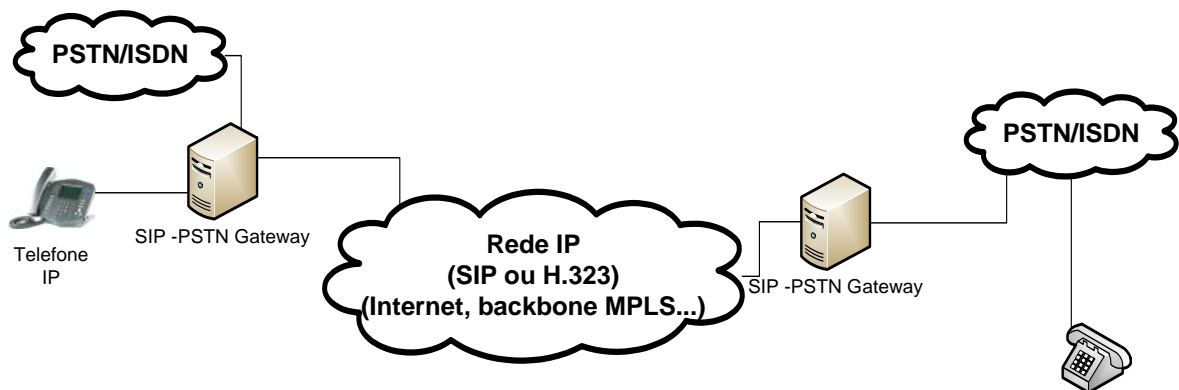


Figura 4.4 – Economia de custos com Gateways SIP/H.323 - ISDN

As MCUs, ou simplesmente servidores de conferências, são equipamentos desenvolvidos para concentrar chamadas provenientes de vários terminais e mixá-las dentro de uma mesma conferência. Ou seja, diversos dispositivos enviam seus fluxos (áudio e vídeo) para uma sala virtual e a MCU envia um único fluxo para cada participante contendo o áudio e/ou vídeo de todos os participantes mixados. A MCU representada na Figura 4.3 apresenta uma funcionalidade a mais, sendo capaz de fazer a tradução dos protocolos usados pelas diversas plataformas, sendo estas SIP, H.323 e ISDN/PSTN.

Diferente do Gateway que somente faz a tradução de protocolos, a MCU além de traduzir protocolos entre diferentes plataformas, gerencia conferências envolvendo os terminais que trabalham com os protocolos suportados. É comum muitas MCUs deste tipo suportarem também funções de gateway, ou seja, disponibilizam números/ramais que em conjunto com extensões ou outros códigos permitem chamadas ponto-a-ponto entre terminais de protocolos diferentes. Visto de outra forma, para executar essa funcionalidade a MCU cria uma conferência com somente dois terminais para simular uma chamada ponto-a-ponto.

A Figura 4.5 mostra um outro exemplo de implementação mais atual e flexível, sendo alternativo ou complementar ao da Figura 4.3.

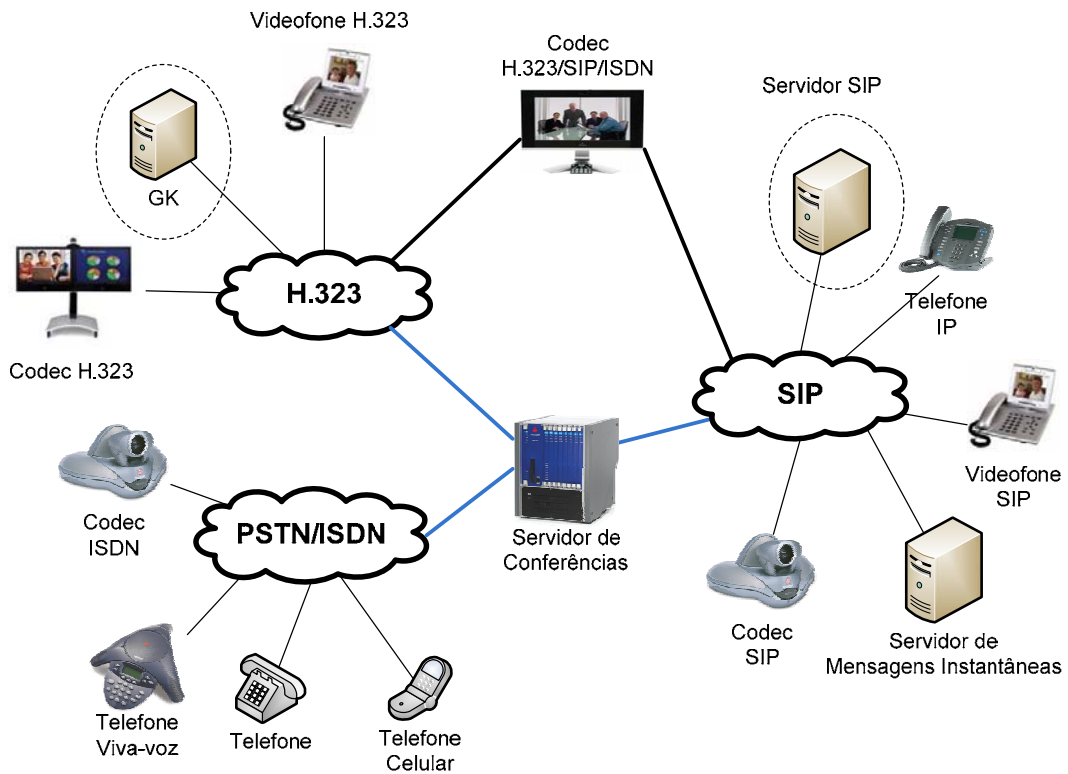


Figura 4.5 – Implementação utilizando um servidor de conferências

Nesta implementação existem terminais, não somente as MCUs, que são capazes de suportar diversos protocolos. Os terminais funcionam em um modo que usualmente é denominado *dual-stack* ou pilha dupla referindo-se ao suporte das pilhas de protocolos H.323 e SIP. É importante ressaltar que um *gatekeeper* não gerencia terminais SIP. Analogamente um servidor de registro SIP não gerencia um terminal H.323. Cada pilha trabalha independentemente, sendo que um terminal pode se encontrar registrado simultaneamente em cada um dos servidores das redes SIP e H.323. Ou seja, os terminais e MCUs *dual stack* é que são capazes de serem gerenciados pelos dois servidores em paralelo.

Esta implementação é uma solução bastante comum nos dias de hoje, pois com a compatibilidade com o SIP, os terminais de videoconferência podem suportar uma série de outros servidores que existem em uma organização e também fazer uso da infra-estrutura de DNS (*Domain Name System*) utilizada mundialmente. Vale recordar que o formato de endereçamento SIP é semelhante ao do email (ex: sipcliente@servidor.com), usufruindo dessa estrutura de resolução de nomes.

Outra vantagem do uso destes terminais é a capacidade de chamar diretamente outros elementos de protocolos distintos, sem consumir recursos de *gateways* ou MCUs. Assim, um

terminal de videoconferência, tipicamente H.323, pode chamar um ramal da telefonia IP usando a sua funcionalidade SIP e também usar a infraestrutura de gateways implementados na rede SIP para chamar telefones convencionais.

Ao que pode parecer, devido à introdução do protocolo SIP em terminais *dual-stack*, existe uma tendência de desenvolvimento destas soluções para que todas passem a utilizar também este protocolo. Apesar dos elementos das redes de vídeo suportarem SIP atualmente, essas redes são raramente desenvolvidas completamente em SIP. A razão para isso é que o SIP ainda não possui as mesmas funcionalidades do H.323 para essas redes.

Uma rede de vídeo implementada em H.323 oferece maior interoperabilidade e funcionalidades entre fabricantes do que um rede implementada em SIP. Dentre essas funcionalidades únicas do H.323 pode-se citar *dual stream video* (baseado no protocolo H.239), controle de velocidade dos fluxos (implementado com o protocolo H.245) e controle de câmera remota (*Far End Camera Control* – baseado nos protocolos H.224 (ITU-T H.224, 1994) e H.281 (ITU-T H.281, 1994)).

4.2 - TELEPRESENÇA

Nas seções anteriores, foi mostrado que a tecnologia para uso em colaboração de áudio e vídeo evoluiu bastantes nos últimos anos. Os codecs de áudio estão cada vez mais avançados possibilitando a audição de mais informações. É possível participar de uma videoconferência com qualidade estéreo em 22 KHz, permitindo identificar os participantes pela sua voz, mesmo entre vozes parecidas, e também pela localização destes em uma sala.

Os codecs de vídeo avançaram em direção à alta definição suportando atualmente resoluções de vídeo que vão até 1080p. É possível observar detalhes da expressão facial, de objetos e de conteúdos enviados de um computador ou outra fonte. Essa alta definição possibilita o uso de monitores e projetores cada vez maiores e que provêm uma experiência melhor e mais próxima da realidade para os participantes de uma videoconferência.

Para aproveitar essa tecnologia em constante avanço, as salas de videoconferência em si estão se tornando cada vez mais preparadas para essas soluções. Na busca por um melhor retorno sobre os investimentos sobre equipamentos de videoconferência, as salas são trabalhadas para que as soluções sejam cada vez mais utilizadas e efetivamente entrem na rotina das organizações.

Dois conceitos usados pelo mercado de videoconferência atualmente são a telepresença e a telepresença imersiva. Ambos estão relacionados à capacidade da solução de videoconferência de propiciar uma experiência onde todos os participantes sintam como se estivessem em uma mesma sala. Ou seja, é possível ver imagens com detalhes das pessoas e do ambiente e ainda ouvir de forma clara cada participante sem demandar esforço. Muitas das soluções de telepresença procuram ainda manter uma proporção de vídeo que apresente o tamanho real das pessoas. Na Figura 4.6 são mostrados exemplos de soluções de salas preparadas especialmente com o intuito de atender ao conceito de telepresença. Este conceito está sendo expandido no mercado e atualmente salas que utilizam equipamentos de alta definição, porém não obedecem aos layouts especializados como os apresentados abaixo também já podem ser tratadas como telepresença.



Figura 4.6 –Exemplos de sala de telepresença (Digitaltechnews , 2009) (Tandberg, 2009)
(Polycom, 2009)

Outro tipo de experiência são as chamadas salas de telepresença imersiva. Essas salas são desenhadas com o mesmo princípio das salas de telepresença, mas a diferença é que os participantes têm o mínimo de contato possível com os equipamentos envolvidos em questão. Alguns artifícios usados para se obter esse efeito são a utilização de uma interface única para fazer as chamadas, inexistências de separações aparentes entre as telas, microfones em localizações não visíveis e câmeras à altura dos olhos dos participantes, muitas vezes

embutidas nas telas onde as imagens são visualizadas. Dessa forma aproxima-se ainda mais de um ambiente real. A Figura 4.7 mostra em exemplo de sala ampla de telepresença imersiva.



Figura 4.7 – Exemplo de sala de telepresença imersiva (Polycom, 2009)

Fabricantes ao vender essas soluções tipicamente requisitam do cliente uma sala vazia. O fabricante então fica responsável por construir e preparar a sala adequadamente para a solução, não sem antes ter revisado o tamanho, localização e infra-estrutura básica da sala (fiação, ar-condicionado etc).

As salas de telepresença são projetadas levando-se em consideração diversas técnicas que visam um maior aproveitamento dos recursos audiovisuais. Dentre os fatores a serem levados em consideração estão: iluminação; recepção e transmissão de áudio; recepção e transmissão de vídeo; acústica; layout e decoração (cores e mobiliário) adequada para videoconferências.

Para estes projetos deve-se avaliar o efeito do ambiente nos dois sentidos da comunicação. Por exemplo, com relação à transmissão de áudio deve-se avaliar os microfones adequados, verificar a fiação para que ruídos indesejados não sejam transmitidos, a disposição destes microfones considerando a qualidade de captação do áudio e aparência da sala, além de avaliar também o material e qualidade da mobília para que também não emitam ruídos altos. No sentido da recepção pode-se avaliar a qualidade dos alto-falantes, disposição espacial na sala, distância dos microfones e outros fatores.

Como ressaltado anteriormente, a intenção destas salas é poder usufruir das informações que agora são transmitidas pelo meio digital. Desta forma utiliza-se da melhor forma a

tecnologia adquirida, trazendo para as organizações um retorno maior com a interação mais eficiente e real entre as pessoas.

No Apêndice A são levantados pontos que devem ser levados em consideração quando se visa um projeto de sala de videoconferência capaz de aproveitar ao máximo os benefícios da solução de videoconferência.

4.3 - AVALIAÇÃO

Levando em consideração as tendências atuais no mercado de comunicação colaborativa e a forma como esta solução vem sendo implementada, alguns pontos são passíveis de avaliação. O primeiro ponto a ser levado em conta é que o mercado, que há algumas décadas enfrentou a transição das comunicações analógicas para as digitais, neste momento enfrenta uma nova mudança de foco como ressaltado no início deste trabalho. A intenção atual é desenvolver soluções integradas capazes de absorver as soluções já implementadas no passado e flexíveis quanto às perspectivas de crescimento futuro.

Além disso, a forma de trabalho dentro das organizações está mudando, valorizando cada vez mais a interação entre as pessoas como forma de geração de conhecimento. Em uma era onde a troca de informações (mensagens, relatórios, vídeos sob demanda, emails, etc) é facilitada, as soluções de UC surgem com uma ferramenta para extrapolar os limites geográficos e atender à necessidade de se desenvolver as relações pessoais.

Partindo desses aspectos, tecnológicos e comportamentais, diversos caminhos podem ser seguidos cabendo às organizações avaliar de forma adequada a situação do mercado atual, a sua própria posição neste mercado e sua direção de crescimento.

4.3.1 - Uso e implementação

Como visto na Figura 1.1 os desafios técnicos para se manter uma infraestrutura básica de telecomunicações não são mais o foco de preocupação das organizações. Não que esse aspecto não seja importante, mas o que se observa é que com o desenvolvimento de tecnologias novas e mais confiáveis, além de equipamentos e técnicas para controle de infraestrutura, o tempo antes gasto com a manutenção desta pode ser utilizado para avaliar melhor suas aplicações.

Uma vez que redes bem projetadas conseguem suportar altas taxas de transmissão, os fabricantes de diversos equipamentos aproveitam o momento para encontrar formas de melhorar a qualidade da experiência provida pelas soluções de áudio, vídeo e colaboração de dados. Logo, a confiabilidade das redes, associada à melhoria da experiência faz com que uma tecnologia como a videoconferência, por exemplo, não seja vista somente como uma alternativa barata em relação à viagens, mas sim como mais uma ferramenta disponível para o dia-a-dia.

Organizações entendem que a comunicação entre pessoas usando diversos meios (mensagens, áudio e vídeo) permite que estas lidem com problemas como a dispersão de times globais, a necessidade de se tomar decisões e implementá-las cada vez mais rápido e também a necessidade de responder rapidamente às pressões competitivas do mercado. A Tabela 4-1 retrata benefícios da videoconferência que vão além da redução de custo com viagens.

Tabela 4-1- Benefícios da videoconferência (Kerravala, 2004)

Aumento do aprendizado	Participantes aprendem 200% mais em reuniões face-a-face do que em reuniões que utilizam somente áudio (Fonte: Universidade de Wisconsin)
Aumento da taxa de Absorção	Participantes absorvem as informações 40% mais rápido em reuniões com vídeo (Fonte: Escola de negócios de Wharton)
Aumento da retenção de conteúdo	Participantes retêm 38% mais informações em reuniões face-a-face do que em reuniões em áudio (Fonte: Universidade de Harvard e Universidade de Columbia)
Melhoria da persuasão	Reuniões face-a-face aumentam o poder de persuasão em 43% quando comparadas as reuniões em áudio somente (Fonte: 3M Co.)
Aumento no impacto da comunicação	55% do impacto na comunicação vem das expressões faciais e linguagem corporal enquanto 38% vem da inflexão vocal (Fonte: UCLA)

Como essas tecnologias estão disponíveis à quase 20 anos, seus usuários já encontraram diversas formas de aplicação nos diversos nichos de mercado. São usadas em variadas áreas como educação (escolas e universidades), saúde, governos, indústrias de diversos ramos,

home offices, eventos, provedores de serviços, comunicação em empresas (pequena, médias e grandes) etc.

Diante de tantas oportunidades e aplicações é comum que as soluções sejam implementadas antes mesmo de se compreender todas suas capacidades. Logo, dentro de organizações, é necessário um planejamento adequado que leve em consideração as intenções das equipes de TI, equipes de telecomunicações e as necessidades dos usuários dos serviços. Dois tipos de organizações podem existir nesse ambiente:

- Organizações que ainda não possuem soluções de comunicações colaborativas. - Essas organizações estão buscando formas de se comunicar e colaborar conteúdo, no entanto ainda não passaram por experiências com essas ferramentas. Essa falta de experiência implica no risco de se investir em uma tecnologia que pode não ser bem aceita e cair em desuso. O conhecimento das soluções e a compreensão das funcionalidades necessárias é incrementado com erros e sucessos.
- Organizações que possuem uma solução ou mais, sendo estas implementadas em épocas diferentes e funcionando em ambientes independentes. - Essas organizações já conhecem os benefícios e eventuais contratempos das soluções de comunicação, no entanto sentem a necessidade de integrar os ambientes para que seja alcançado um nível mais alto de benefícios.

Para ambos os casos algumas recomendações são relevantes:

- Começar com um grupo pequeno de usuários: se organizações ainda não começaram a integrar suas soluções, pode-se começar com um grupo de usuários avançados de um departamento. Essa prática permite que as equipes de tecnologia entendam o impacto das soluções para desenhar melhor os requisitos e soluções finais.
- Desenvolver métricas para avaliação: a avaliação do impacto no negócio depende da capacidade de relacionar o seu uso com ganhos de produtividade.
- Levantar casos de uso: é importante levar em consideração as necessidades atuais dos usuários que terão acesso à tecnologia. Deve-se aliar a viabilidade tecnológica com o atendimento de necessidades dos usuários. Muitas vezes os usuários podem não saber o que é possível ser feito, logo é necessário apresentar e discutir as possibilidades com todos.

- Estabelecer uma estratégia de adoção: ao compreender as necessidades deve-se decidir sobre gerir a rede internamente ou terceirizar o gerenciamento, quais são as tecnologias a ser implementadas primeiro e como será dado o suporte às aplicações após implementação.

Para o usuário final destas soluções, procura-se simplificar ao máximo a forma de utilização, mas como visto é necessário avaliar bem as soluções disponíveis. Por trás da simplicidade apresentada aos usuários estão plataformas avançadas de gerenciamento e integração.

Apesar da complexidade de implementação, o mercado de UC está crescendo ano após ano e existe uma tendência de aumento principalmente no que diz respeito a integrações e soluções de videoconferência. Estudos mostram que o mercado de vídeo corporativo tende a crescer uma vez que os enlaces se tornam mais baratos assim como os equipamentos. Pode-se observar na Figura 4.8 uma previsão de crescimento para o mercado de vídeo corporativo nos próximos anos. Este fato se dá também devido às capacidades nativas de integração de novos equipamentos e protocolos diversos.

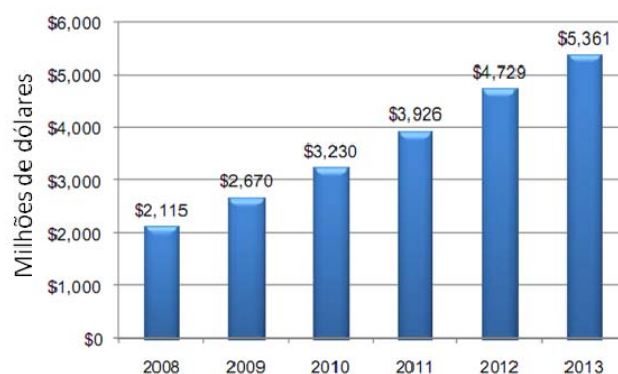


Figura 4.8 – Crescimento da receita em equipamentos de vídeo (Davis, 2009) .

Fato importante a ser notado é que usuários já perceberam as vantagens de se trabalhar com alta-definição e com soluções para melhorar a experiência de vídeo. O gráfico da Figura 4.9 mostra que as soluções de telepresença em 2008 já representavam 15% do mercado de equipamentos de comunicações em vídeo. Esses dados justificam os novos investimentos que devem ser feitos em equipamentos preparados para o futuro, ou seja, que tenham tecnologia o suficiente para se aproveitar de investimentos feitos em enlaces e infra-estrutura de dados.

■ Desktop Executivo ■ Sistemas de Sala
■ Sistemas de telepresença

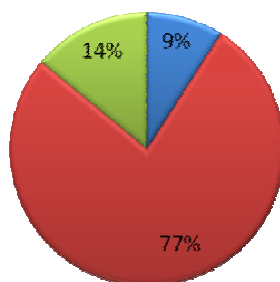


Figura 4.9 – Receita de equipamentos de videoconferência de grupo (Davis, 2009)

4.3.2 - ISDN para IP

Como ressaltado em seções anteriores, as redes ISDN utilizam circuitos que são estabelecidos para cada chamada (áudio e/ou vídeo) realizada. Pela operadora que provê os enlaces ISDN é cobrado um valor fixo para a assinatura dos enlaces e um valor variável que depende do tempo e velocidade das chamadas. Conferências internacionais realizadas a 384 kbps, por exemplo, equivalem a 6 chamadas internacionais de voz. Empresas que utilizam a infraestrutura de dados IP para comunicação de áudio e vídeo tipicamente contratam de sua operadora um enlace adequado para as suas necessidades e pagam um valor fixo que independe do tempo de uso ou velocidade de cada chamada.

Apesar do alto custo, durante muito tempo a rede ISDN era a forma mais utilizada para assegurar a confiabilidade de conexão, uma vez que as redes IP não conseguiam prover a mesma garantia. Este cenário está sendo alterado nesta última década devido, entre outros fatores, à melhoria nos serviços IP fornecidos pelas operadoras e redução do custo da banda larga para empresas e usuários domésticos. As redes ISDN já não são o meio principal utilizado para realização de comunicações de áudio e vídeo. As tecnologias disponíveis para redes IP permitem atualmente que se priorize o tráfego de pacotes de áudio e vídeo, garantindo assim a qualidade de comunicação em um meio compartilhado. Esta priorização, como já visto, depende de contratos feitos com a provedora de serviços.

O uso de redes IP, além do menor custo, oferece outras vantagens como:

- melhor gerenciamento local dos elementos da rede;
- escalabilidade;

- maiores velocidades de conexão (ex: chamadas até 6 MPs);
- monitoramento de chamadas e dispositivos;
- simplicidade de discagem;
- integração com novos serviços;
- segurança e criptografia.

Além de possibilitar comunicação com menor custo e flexibilidade, o uso de redes IP trouxe as aplicações de UC para o computador pessoal, permitindo uma difusão desses serviços em larga escala e usando equipamentos periféricos de baixo custo.

Em ambientes de UC as redes ISDN ainda podem ser aproveitadas para agregar funcionalidades, mas não devem mais ser consideradas o foco da infraestrutura. Estas redes podem ser utilizadas para permitir a comunicação com equipamentos legados sem suporte à redes IP, integração com as redes PSTN (chamadas para linhas telefônicas convencionais e aparelhos celulares) e comunicação com regiões onde a conectividade IP não é confiável ou não é possível estabelecer serviços de QoS (priorização de pacotes).

4.3.3 - SIP e H.323

Como apontado anteriormente, os protocolos H.323 e SIP exercem funções semelhantes dentro dos universos VoIP, videoconferência e Comunicações Colaborativas Unificadas (UC). O H.323 foi inicialmente desenvolvido para suportar aplicações de vídeo, logo implementa funcionalidades específicas para estas aplicações com confiabilidade. O SIP é bastante utilizado em aplicações de voz oferecendo funcionalidades adequadas para estas aplicações. Além da comunicação de voz provida pelo SIP, este protocolo, por sua flexibilidade, permitiu o desenvolvimento de outras aplicações como ‘presença’ e mensagens instantâneas.

Devido à essa capacidade flexível do SIP de iniciar sessões para tipos diversos de aplicações, este é o protocolo mais adotado para novas aplicações de UC. As aplicações de voz usando SIP já estão bastante difundidas dentro de empresas, logo novas aplicações que também suportam SIP podem alcançar rapidamente toda a organização. Existem diversas aplicações para vídeo em SIP, no entanto, quando se trata de videoconferência, o H.323 é superior ao SIP em funcionalidades e interoperabilidade.

Para suprir essa deficiência, o IETF tem agregado novas RFCs ao SIP para que este possa oferecer as mesmas funções do H.323 (Karapetkov, 2008), como por exemplo:

- FECC (*Far end camera control*) – Funcionalidade que permite o controle de câmeras remotas em uma chamada de vídeo. Também usado para controlar funcionalidades (ex.: menus, layouts) de MCUs durante uma conferência multiponto. No H.323 essa capacidade é descrita no Anexo Q, em que os protocolos H.281 (sintaxe dos comandos) e H.224 (sinalização) são usados para criar e usar um canal adicional que provê essa funcionalidade. Para o SIP, o protocolo SDP (IETF RFC 2327, 1998), através da RFC 4573 define um tipo de mídia H.224 que ‘encapsula’ a sinalização H.281. Dessa forma é possível que haja interoperabilidade entre o SIP e o H.323 para FECC, seja através de um gateway ou MCU.
- *Dual vídeo* – Tecnologia usada para envio de conteúdo (ex.: apresentações, vídeos, tela de PC). Este é tratado como um fluxo secundário de vídeo enviado em paralelo ao fluxo principal. O H.323 utiliza o protocolo H.239 para realizar essa funcionalidade associando papéis (*‘Live’* e *‘Presentation’*) para os diferentes fluxos de mídia. Estes são interpretados de maneira diferente pelos dispositivos H.323. O SIP, para a mesma função, utiliza uma série de RFCs (4574 (IETF RFC 4574, 2006), 4796 (IETF RFC 4796, 2006), 4388 (IETF RFC 4388, 2006), 4582 (IETF RFC 4582, 2006) e 4583 (IETF RFC 4583, 2006)) usadas para definir um tipo de mídia, identificá-la e associá-la ao fluxo principal.
- Controle de fluxos – O H.323 utiliza mensagens *FlowControlCommand* do protocolo H.245 para limitar a banda ou parâmetros de determinados fluxos de mídia. Este recurso pode ser usado em casos onde há súbita perda de pacotes, situação onde essas mensagens são trocadas para reduzir a banda utilizada mantendo estabilidade da chamada. O processo de modificar parâmetros de sessões abertas no SDP é descrita pela RFC 3264 (IETF RFC3264,2002). O SIP, através da RFC 3264 envia uma mensagem de reINVITE incluindo os novos parâmetros a serem usados.

Do ponto de vista de padronização ambos os protocolos conseguem prover as mesmas funcionalidades. Mas sob o ponto de vista de suporte às funcionalidades implementados nos produtos a situação é diferente. As funcionalidades citadas funcionam com a sinalização básica do H.323, sendo suportadas pela grande parte dos dispositivos H.323. No caso do SIP a situação é um pouco diferente. O SIP FECC já é suportado por diversas MCUs, mas o *dual vídeo* ainda encontra muitos problemas de interoperabilidade e as primitivas do controle de

fluxo ainda dão muita liberdade para a implementação, dificultado sua adoção. Outro ponto importante onde se encontram problemas é o suporte à segurança, tanto para autenticação, quanto para criptografia de mídia e sinalização. Cada protocolo utiliza métodos diferentes, o que dificulta bastante a compatibilidade.

O fato é que os dois protocolos, SIP e H.323, ainda devem coexistir durante muito tempo sendo adotados de acordo com as necessidades específicas da solução a ser implementada. Como ressaltado anteriormente, o mais adequado é optar por equipamentos capazes de suportar ambos protocolos, avaliar as funcionalidades disponíveis e também os planos dos fabricantes visando o futuro destes equipamentos.

O ITU-T está trabalhando atualmente na proposta de uma solução para o dilema de se decidir entre SIP ou H.323. Percebe-se que a interoperabilidade entre os dois protocolos demanda esforços de ambos os lados e para implementação de novas funcionalidades, ambos os protocolos devem procurar formas de se adequar. Cada protocolo tem suas vantagens e desvantagens quando se avalia seu papel nas comunicações colaborativas. Avaliando o cenário atual, o ITU-T está desenvolvendo um novo protocolo, o H.325.

O novo protocolo visa uma plataforma de aplicação para permitir que aplicativos de comunicação multimídia proveniente de diversos tipos de dispositivos possam estabelecer sessões entre si. Serão viabilizadas, através do protocolo, aplicações de áudio, vídeo, telefonia, compartilhamento de dados, aplicações corporativas, mensagens instantâneas, presença, quadro branco e compartilhamento de aplicações (ITU-T H.325, 2006). Todas essas aplicações, utilizando um mesmo protocolo, poderão implementar diversas novas funcionalidades.

O ITU-T argumenta que o SIP e o H.323 somente reproduzem o ambiente PSTN e possuem limitações como: troca de capacidades deficiente; problemas de interoperabilidade; questões como QoS e segurança consideradas tardiamente gerando soluções não ideais e sub-utilização da tecnologia IP (ex.: não poder fazer downloads sob demanda). O que se percebe é que, como acontece com qualquer tecnologia, o SIP e o H.323 tendem a ser substituídos por protocolos mais inovadores, capazes de atender às demandas dos usuários e se aproveitar da infraestrutura de telecomunicações existente. Apesar de ambos protocolos já terem mais de 10 anos, esse processo ainda pode levar tempo.

5 - CONCLUSÃO

Ao longo deste trabalho foram abordados os temas principais relacionados às comunicações colaborativas multimídia. Foi mostrado que as aplicações de áudio e vídeo, antes implementadas em ambientes separados, estão sendo tratadas cada vez mais como uma solução única, se tornando para as organizações uma nova ferramenta de trabalho. Neste sentido essa dissertação se apresenta como um guia para que se compreendam os diversos aspectos dessas tecnologias e se possa avaliar de forma mais objetiva as soluções disponíveis no mercado.

Foram apresentadas também algumas formas de se implementar soluções de comunicações colaborativas procurando unificar soluções de áudio e vídeo através dos dois protocolos mais implementados atualmente, SIP e H.323. Compreender como se trabalhar com os dois protocolos ajuda a perceber que não é necessário um protocolo somente para que se consiga uma solução completa. Durante os próximos anos uma solução completa será implementada ainda através de *gateways*, *bridges* e protocolos diversos que serão implementados de acordo com as necessidades e aplicações desejadas.

O que se observou é que as tecnologias para comunicações caminham junto nas suas diversas etapas. Enquanto as infraestruturas de redes avançam para se tornarem mais confiáveis e rápidas, novos equipamentos são desenvolvidos e novos serviços são criados. Com o *feedback* das experiências obtidas, o ciclo se repete continuamente.

Conclui-se também que a implementação de soluções de integração exige planejamento e contribuições vindas da experiência obtida com as soluções já implementadas. O planejamento deve envolver as pessoas, a infraestrutura e decisões focadas no futuro e expansão da solução.

O que pôde ser notado também é que fabricantes de soluções de UC estão atendendo uma demanda que está em sintonia com a tendência das organizações em valorizar cada vez mais os seus funcionários e a geração de conhecimento. As comunicações unificadas aparecem neste cenário como facilitadores para permitir uma melhor interação entre pessoas que estão separadas geograficamente.

Com a introdução cada vez maior de comunicações realizadas utilizando redes IP, que por sua vez não introduzem custos adicionais para cada chamada realizada, as soluções de comunicações colaborativas podem ser difundidas para uma grande quantidade de pessoas. Chega-se a um ponto onde as soluções não são somente usadas para reduzir custos de viagens, mas também para permitir que pessoas passem a ser conhecer melhor, obtendo assim mais agilidade e qualidade na comunicação.

Este trabalho teve enfoque nas soluções de áudio e vídeo que demandam na maioria das vezes equipamentos especializados como telefone VoIP, videofones e equipamentos de videoconferência. Um tópico que pode ser abordado no futuro são as aplicações web e como estas estão posicionadas neste ambiente de UC. Essas aplicações tipicamente requerem taxas de transmissão menores e permitem também uma grande variedade de funcionalidades.

Outros trabalhos futuros que podem complementar este estudo estariam associados à redes IP seguras (criptografia, proxies, firewalls), integração com as redes móveis sem fio, frameworks para planejamento de soluções de integração e desenvolvimento de novos serviços suportados pelas novas tecnologias.

REFERÊNCIAS BIBLIOGRÁFICAS

- Acustica. (2008). www.acustica.ind.br/Gesso%20Acartonado.htm - acessado em 05/07/2009.
- Avispl. (2009). www.avispl.com/SLIS/AVI/ProjectImages/79_lg.jpg, acessado em 25/08/2009.
- Black, U. (1999). "Voice over IP.". New Jersey: Prentice Hall.
- Brey, S. (2000). "Handbook of Internet Computing.", Borko Furht.
- C21VIDEO. (2008). "H.221 Framing used in ISDN Conferences." c21video.com/h221.html, acessado em 13/05/2009.
- Clayton, J. (2000). "Illustrated Telecom Dictionary.", New York: McGraw – Hill, Second Edition.
- Comer, D. E. (1998). "Interligação em rede com TCP/IP.", vol.1, 3a. edição, Editora Campus.
- Das, K. "Multiprotocol Label Switching.", www.ipv6.com/articles/general/Multi-Protocol-Label-Switching.htm, acessado em 25/07/09.
- Davis, A. (1999). "The Evolution from the Corporate Boardroom to the Departmental Conference Room and Brach Office.", Wainhouse Research.
- Davis, A., Brent, K. (2009). "The compelling Case for Video Telephony in Unified Communications.", Weinhouse Research.
- Dhesikan S. (2001). "Quality of Service for IP Videoconferencing.", White Paper, CISCO Systems.
- Dhesikan, S. (2002). "H.323/RSVP Synchronization for Video over IP.", White Paper, CISCO.
- Digitaltechnews. (2009). http://www.digitaltechnews.com/news/images/cisco_telepresence_3000_system.jpg, acessado no dia 14/08/2009
- ENGEL Engenharia Elétrica. (2002) "Investir na iluminação do seu escritório é obter vantagens imediatas e retorno a curto prazo.", www.grupozug.com.br/ENGEL/escritorios.htm, acessado em 08/05/2009.
- FGV-Faculdade Getúlio Vargas. (2009). <http://www.eaesp.fgvsp.br/default.aspx?pagid=IMNCNQPQ>, acessado em 10/08/2009
- Fluckiger, F. (1995). "Understanding Networked Multimedia – applications and technology.", Herefordshire: Prentice Hall.
- Ghein, L. De. (2007). "MPLS Fundamentals.", Cisco Press, 1ª edição.
- Gonçalves, F. E (2008). "Building Telephony Systems with OpenSer.", Packet publishing.
- GTA/UFRJ Grupo de Teleinformática e automação. "IEEE 802.1p - Qos na camada MAC." www.gta.ufrj.br/grad/02_2/802.1p/tag.htm, acessado em 05/07/2009.
- Guimarães, P. O. I. (2002). "Arquitetura de um Ambiente de Videoconferência Integrando Redes IP e Redes Comutadas.", Dissertação.
- Hertig. (2009). <http://www.hertig.com.ar/productos/7/33/>, acessado em 24/07/2009.

HSW. (2009) eletronicos.hsw.uol.com.br/televisoes-de-projecao1.htm - acessado em 05/07/2009

IETF RCF 2205. (1997). “Resource ReSerVation Protocol (RSVP).”

IETF RFC 1633. (1994). “Integrated Services in the Internet Architecture: an Overview.”

IETF RFC 2236. (1997). “Internet Group Management Protocol.”, version 2.

IETF RFC 2327. (1998) “SDP: Session Description Protocol.”

IETF RFC 2474. (1998). “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers.”

IETF RFC 2475. (1998). “An Architecture for Differentiated Services.”

IETF RFC 3032. (2001). “MPLS Label Stack Encoding”

IETF RFC 3261. (2002). “SIP: Session Initiation Protocol.”

IETF RFC 3264. (2002). “An Offer/Answer Model with the Session Description Protocol (SDP).”

IETF RFC 3428. (2002). “Session Initiation Protocol (SIP) Extension for Instant Messaging.”

IETF RFC 3856. (2004). “A Presence Event Package for the Session Initiation Protocol (SIP).”

IETF RFC 3863. (2004). “Presence Information Data Format.”

IETF RFC 4388. (2006). “Dynamic Host Configuration Protocol (DHCP) Leasequery.”

IETF RFC 4573. (2006). “MIME Type Registration for RTP Payload Format for H.224.”

IETF RFC 4574.(2006). “The Session Description Protocol (SDP) Label Attribute.”

IETF RFC 4582. (2006). “The Binary Floor Control Protocol (BFCP).”

IETF RFC 4583. (2006). “Session Description Protocol (SDP) Format for Binary Floor Control Protocol (BFCP) Streams.”

IETF RFC 4796. (2007). “The Session Description Protocol (SDP) Content Attribute.”

IETF RFC 762. (1980). “Internet Control Message Protocol.”

IETF RFC IP 791. (1981). “Internet Protocol.”

IETF RFC1371. (1992). “Choosing a Common IGP for the IP Internet.”

IETF SIMPLE. (2009). “SIP for Instant Messaging and Presence Leveraging Extensions (simple).”, <http://www.ietf.org/dyn/wg/charter/simple-charter.html>, acessado em 29/06/2009.

ISO. (2009). “About ISO”, www.iso.org/iso/about.htm, acessado em 10/04/2009.

Itest. (2009).
http://www.itest.com.br/config/imagens_conteudo/produtos/imagensGRD/GRD_59_LD-240.jpg, acessado em 07/08/2009.

ITU-T E.164. (1988). “The international public telecommunication numbering plan.”

ITU-T H.221. (1988). “Frame structure for a 64 to 1920 kbit/s channel in audiovisual teleservices.”

ITU-T H.223. (1996). “Multiplexing protocol for low bit rate multimedia communication.”

ITU-T H.224. (1994). “A real time control protocol for simplex applications using the H.221 LSD/HSD/MLP channels.”

- ITU-T H.225. (2006). “Call signalling protocols and media stream packetization for packet-based multimedia communication systems.”
- ITU-T H.239. (2003). “Role management and additional media channels for H.300-series terminals.”
- ITU-T H.241. (2005). “Extended video procedures and control signals for H.300-series terminals.”
- ITU-T H.242. (1990). “System for establishing communication between audiovisual terminals using digital channels up to 2 Mbit/s.”
- ITU-T H.245. (1996). “Control protocol for multimedia communication.”
- ITU-T H.261. (1988). “Video Codec for Audiovisual Services at P x 64 kbps.”
- ITU-T H.263. (1996). “Infrastructure of audiovisual services – coding of moving video.”
- ITU-T H.264. (2003). “Advanced video coding for generic audiovisual services.”
- ITU-T H.281. (1994). “A far end camera control protocol for videoconferences using H.224.”
- ITU-T H.320. (1990). “Narrow-band visual telephone systems and terminal equipment.”
- ITU-T H.323. (1996). “Infrastructure of audiovisual services – Systems and terminal equipments for audiovisual services.”
- ITU-T H.325. (2006). “Third Generation Multimedia System, IMTC Forum 'H.323, SIP: is H.325 next?'.”, www.itu.int/ITU-T/worksem/h325/200605/presentations/s1p2-jones.pdf, acessado em 01/08/09.
- ITU-T I.411. (1993). “ISDN user-network interfaces - Reference configurations.”
- ITU-T X.200. (1988). “Information technology - Open Systems Interconnection - Basic Reference Model: The basic model.”
- ITU-T. (2009). “ITU-T in brief.”, www.itu.int/ITU-T/, acessado em 05/04/2009.
- Jung, Y.; Allamiyarov, T.; Seo, J.; Kim, Y. (2005). “SIP and RSVP-TE for QoS-guaranteed DiffServ Provisioning.”, Dept. of Information & Communiation Engineering, Yeungman University.
- Karapetkov, S. (1999). “Deploying H.323 for Voice, Video and Data Communication in IP Networks.”, Siemens Information and Broadband Networks.
- Karapetkov, S. (2008). “Migrating Visual Communications from H.323 to SIP.”, White Paper.
- Kerravala, Z., Hamilton, G. (2004). “Unified Collaborative Communications for Real-time Enterprise.”, The Yankee Group Report.
- Mancera. (2009). www.mancera.cl/productos-acustica.htm
- Microphone Reviews. (2009). www.microphonereviews.com, acessado em 23/09/09
- Niclas E. (1999). “IEEE 802.1 p,Q – QoS on the MAC level.”, www.tml.tkk.fi/Opinnot/Tik-110.551/1999/papers/08IEEE802.1QosInMAC/qos.html, acessado em 25/06/2009.
- O’Neil, T. M. (2002). “Network based Quality of Service.”, White Paper, Polycom.
- Packetizer. (2009) “A Primer on the H.323 Series Standard”, www.packetizer.com/ipmc/h323/papers/primer/, acessado em 05/06/2009.

- Packetizer. “H.323 vs SIP: A comparison.”, www.packetizer.com/ipmc/h323_vs_sip/, acessado em 04/07/2009.
- Polycom, (2005). “High Definition the evolution of videoconferencing, Technology Brief.”, White Paper.
- Polycom. (2006). “Supporting real time traffic - Preparing IP network for videoconferencing.”, White Paper.
- Polycom. (2008). “Integrators Reference Manual for Polycom HDX systems.”, versão 2.5.
- Polycom. (2009). “The Power of collaboration within Unified Communications.”, White Paper.
- Polycom. (2009). “Art of Teleworking.”, White Paper.
- Polycom. (2009). Tela do software Polycom SoundStructure Studio – capturada em 24/07/2009
- Radvision Ltd. (2001). “L2W – 323 Gateway version 2.2 User Guide.”, Tel Aviv.
- RAU-Revista Arquitetura & Urbanismo. (1999). www.catep.com.br/dicas/gesso%20acartonado.htm – acessado em 05/07/2009.
- Refax. (2002) “Sistema Modular Tegal.”, http://www.refax.com.br/img/produtos/modular_tegal.jpg, acessado em 10/05/2009
- RNP-Rede Nacional de Pesquisa. (2002). “Perspectivas sobre Qualidade de Serviço nos Protocolos da Internet - Estudo de Caso: Aplicações de Vídeo Sob Demanda.”, 2002 www.rnp.br/newsgen/0007/art1.html, acessado em 25/07/09.
- Rodman, J. (2006). “The Effects of bandwidth on Speech Intelligibility.”, POLYCOM.
- Serra, J. (2007). “Curso Fotografia Básica.”, users.linkexpress.com.br/juliano/afotema4.htm, acessado em 12/07/2009.
- Stallings, W. (1997). “Data and Computer Communications.”, New Jersey: Prentice Hall, Fifth Edition.
- Tandberg. (2009). http://www.tandberg.com/images/products/telepresence_t1_med.jpg, acessado no dia 14/08/2009
- Tennessee. (2009). www.videoconferencingtennessee.com/_/rsrc/1238179561453/Home/Video%20Conferencing%20Conference%20Room.jpg, acessado em 25/08/09.
- USC. (2009). www.usc.edu/schools/sppd/private/assets/photos/programs/captureroom.jpg, acessado em 25/08/09.
- VIDE. (2004), “Videoconference Cookbook.”, www.vide.net/cookbook, acessado em 10/06/2009.
- Vitra. (2009). www.vitra.com/_assets/published/12/d2a285c7f39b115.jpg, acessado em 25/08/09.
- Weinstein, I. M. (2006). “The ISDN to IP Migration for Videoconferencing.”, Weinhouse Research.
- Wilcox, J.(2000). “Videoconferencing the whole picture.” New York: Telecom Books.
- Wilson, T. V.; Freudenrich, C. (2009) “How Projection Television Works.”, www.electronics.howstuffworks.com/projection-tv.htm, acessado em 05/07/2009.

APÊNDICES

A. REQUERIMENTOS DE SALA

A.1 - ILUMINAÇÃO

A qualidade da imagem obtida numa sala de videoconferência está diretamente relacionada à iluminação do ambiente. Nos últimos anos, o avanço tecnológico das câmeras de vídeo e dos equipamentos de iluminação proporcionaram facilidades na obtenção de uma boa qualidade de imagem, porém, para garantir um padrão aceitável, algumas recomendações podem ser destacadas:

- evitar sempre que possível, salas com janelas, de modo que as fontes externas de iluminação sejam neutralizadas; não sendo possível, deve-se procurar utilizar cortinas do tipo *black-out*;
- a sala deve ser iluminada por meio de fontes de luz homogêneas, utilizando-se, preferencialmente, lâmpadas fluorescentes do tipo “luz do dia”. Não utilizar outros tipos de fontes de iluminação, como lâmpadas incandescentes misturadas à iluminação fluorescente; esta combinação pode provocar na câmera uma confusão de cores e produzir resultados indesejáveis;
- a intensidade luminosa, medida com um luxímetro (Figura A-1) em uma sala de videoconferência deve estar situada na faixa de 700 a 1.100 lux (VIDE, 2004)0, ou seja, um pouco acima do estabelecido pela norma da ABNT para iluminação de escritórios;



Figura A-1 – Exemplo de luxímetro (Ittest, 2009)

O lux é a unidade internacional de medida utilizada para indicar a quantidade de luz emitida por um corpo luminoso, por metro quadrado (Serra, 2007), a Tabela A-1 contém alguns exemplos de medidas.

Tabela A-2 – Medidas de iluminação em Lux

luz do luar	1 lux
vela acesa a 30cm de distância	10 lux
isqueiro aceso a 30cm de distância	15 lux
vela acesa a 5cm de distância	100 lux
lâmpada incandescente de 100W em um ambiente de 9m quadrados	100 lux
flash comum a 1 metro de distância	250 lux
dia claro, uma hora após o pôr do Sol	1000 lux
dia nublado, uma hora depois do nascer do Sol	2000 lux
refletores de uma sala cirúrgica	10.000 lux
dia nublado às 10 horas da manhã	25.000 lux
luz do Sol em dia claro	100.000 lux

- a sala deve ser uniformemente iluminada, não devendo haver áreas de sombra em parte alguma. Para isto, as luminárias devem possuir aletas difusoras (Figura A-2 e Figura A-3) e refletores de alumínio anodizado de alta pureza;



Figura A-2 – Luminária de sobrepor com refletor em alumínio anodizado brilhante de alta pureza (2 lâmpadas) (Engel, 2009)



Figura A-3 – Luminárias com 4 lâmpadas em forro modulado (Refax, 2009)

- devem-se utilizar reatores do tipo eletrônico de alto desempenho, visando evitar a ocorrência de ruídos indesejáveis;
- as paredes devem ser lisas, de cores uniformes e sem superfícies muito refletoras. As cores recomendadas são: cinza, azul e verde claros, creme ou branco gelo.

A.2 - Áudio

A seleção e a localização de microfones são decisivas, devendo haver bastante cuidado quanto a sua utilização. A Figura A-4 mostra exemplos de microfones que são tipicamente utilizados em videoconferências. Alguns fabricantes optam por utilizar interfaces proprietárias, mas disponibilizam adaptadores ou entradas adicionais em seus equipamentos para suportar microfones que são tipicamente usados em salas convencionais.



Figura A-4 – Microfones unidirecionais e multidirecionais (Microphone Reviews, 2009)

Microfones unidirecionais têm uma sensibilidade muito maior para sons frontais do que para qualquer outra direção, na média de 20db-30db a mais de sensibilidade para ondas sonoras que chegam pela frente, do que por trás do microfone. São muito utilizados em salas grandes ou auditórios que exijam o uso de alto-falantes. Esse tipo de microfones pode ser

disposto de forma que apontem para a direção contrária do alto falante, assim diminuindo o efeito de microfonia.

Os microfones omni-direcionais são usualmente colocados em salas de reuniões sobre a mesa ou pendurados no teto. Estes microfones têm a função de captar áudio igualmente de todas as direções. Tipicamente têm um alcance maior do que os microfones unidirecionais, podendo chegar a alguns metros. A maioria dos fabricantes permite que o ganho destes microfones seja ajustado para que ruídos indesejados não sejam captados.

Muito comum em ambientes amplos onde se utiliza videoconferência é o uso de equipamentos para auxiliar no cancelamento do efeito de eco. Esse efeito ocorre quando o áudio gerado pelo emissor (alto-falante) é captado pelo microfone local, sendo retransmitido para o ponto remoto. A maneira mais fácil de perceber esse efeito é quando um participante fala e logo depois ouve o seu próprio som pelo seu alto-falante.

Em ambientes amplos com várias fontes e saídas de áudio, como os exemplos da Figura A-5, muitas vezes é necessária a ajuda de equipamentos específicos de gerenciamento de áudio. É comum se ver em auditórios, matrizes de som que são utilizadas para distribuir e controlar as diferentes entradas para as diferentes saídas de um ambiente. No entanto, as matrizes tradicionais são analógicas e não trabalham para suprimir problemas como ruídos, atrasos, ajustes de ganho, eco e microfonia.

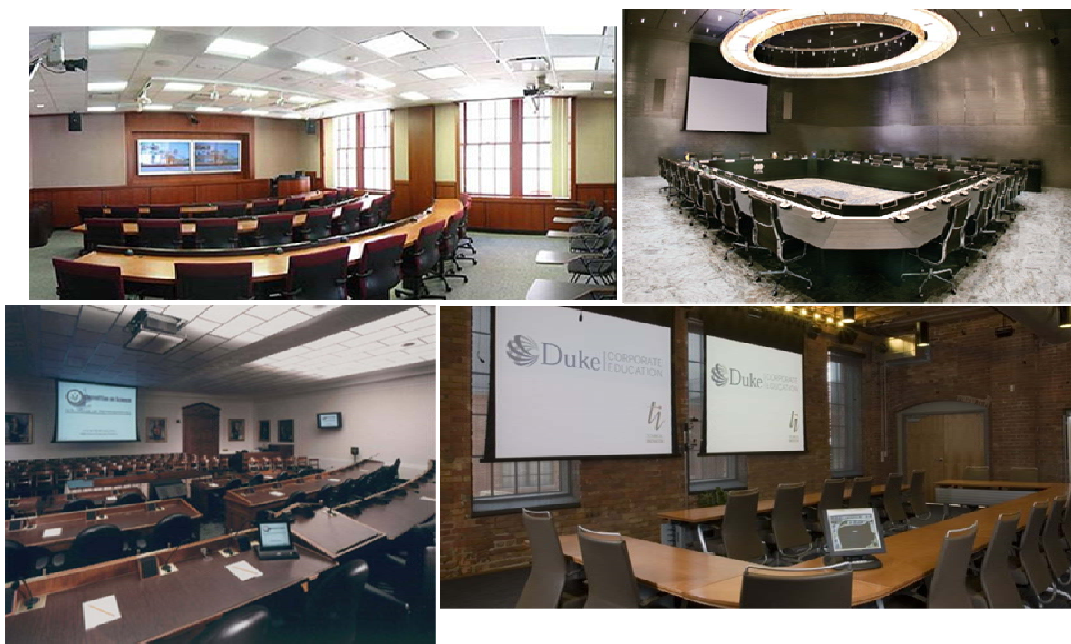


Figura A-5 - Exemplos de salas amplas de videoconferência (Avispl, 2009) (Vitra, 2009)
(USC, 2009) (Tennessee, 2009)

Atualmente se utilizam matrizes digitais capazes de digitalizar os sinais analógicos, processá-los e depois enviá-los para as saídas determinadas. Essa solução é composta tipicamente de um hardware, onde se plugam os elementos de áudio, e um software para controlar esses elementos. É possível, através do software instalado em um PC controlar as entradas de som (ex.: microfones mono e estéreo, linhas de telefone, entradas de DVD ou PC) para enviá-las para as saídas corretas (ex.: amplificadores mono e estéreo e saída para gravação). Todos esses sinais, antes de serem enviados para as saídas, passam antes por um processo de digitalização e processamento visando resolver os problemas citados no parágrafo anterior. A Figura A-6 mostra a interface de um software para gerenciamento de áudio. É possível verificar em verde as diferentes entradas (4 microfones de mesa e uma linha de telefone) e as saídas (equipamento de videoconferência e amplificadores de parede e de teto). É possível definir, portanto, quais fontes de áudio sairão por quais saídas e que tipo de melhorias devem receber.

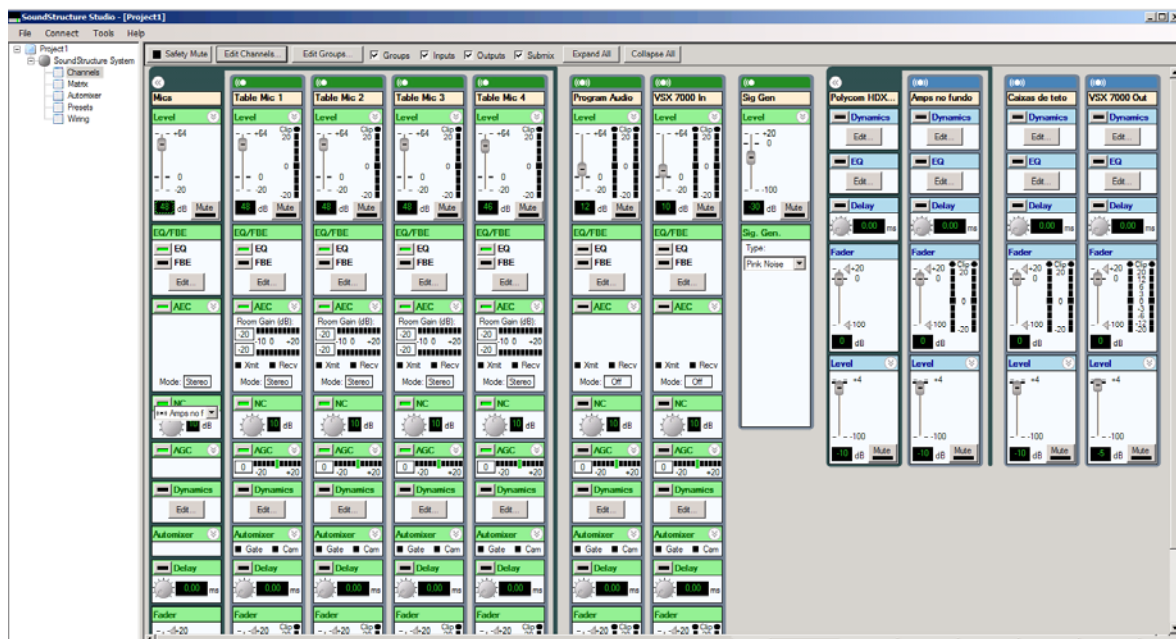


Figura A-6 – Interface de um software para gerenciamento de áudio (Polycom, 2009)

A.3 - Acústica

A qualidade do áudio é de extrema importância para uma videoconferência. Por isso, deve-se observar o tratamento acústico do ambiente objetivando minimizar os ruídos indesejáveis, tanto externos quanto internos.

Recomenda-se que o nível de ruído máximo admissível em uma sala de videoconferência não ultrapasse 50 dB – SPL (decibel – Sound Pressure Level). Esta medida deve ser tomada com o ambiente vazio, todos os equipamentos ligados e com o medidor (decibelímetro - Figura A-7) no centro da sala. Caso não seja possível atingir valores abaixo do citado, é necessário tratar o ambiente acusticamente, considerando-se os seguintes aspectos: absorção e refletividade, níveis de ruído do ambiente, efeito de reverberação (eco), tamanho da sala e sua geometria.



Figura A-7 – Decibelímetro (Hertig, 2009)

Inicialmente, na escolha do local, deve-se evitar grandes salas com paredes paralelas, pois isto acarretará a reflexão do som, prejudicando a videoconferência. Ambientes que possuam superfícies refletoras, como mesas de vidro, teto de gesso, paredes de madeira e piso de cerâmica, também não são recomendáveis.

A redução das reflexões sonoras a níveis aceitáveis no interior de uma sala exige a utilização de técnicas de neutralização com emprego de materiais acusticamente absorventes. Uma opção é a aplicação de carpete no piso e o uso de placas absorventes de espuma acústica (Figura A-8) nas paredes e no teto. Outras opções muito utilizadas são a aplicação de tecido para parede e o jateamento de celulose, este último mais usado em ambientes grandes onde se quer reduzir o tempo de reverberação.

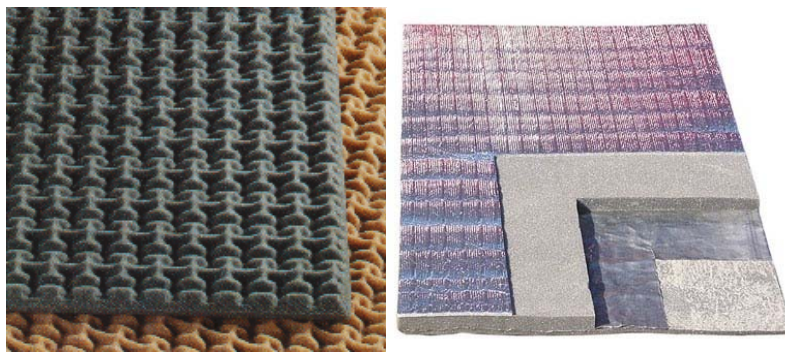


Figura A-8 – Placas para controle de reverberação (Mancera, 2009)

Caso haja excesso de ruído externo prejudicando a qualidade da videoconferência e da teleconferência, é necessária a utilização de paredes apropriadas para o isolamento acústico. Com isso, é possível reduzir a influência de fontes externas no interior da sala e garantir o sigilo da informação, dificultando a escuta.

Uma solução bastante utilizada para isolamento acústico em paredes é o Gesso Acartonado ou *Dry Wall*. É de construção seca e rápida, possui 10% do peso de uma parede de alvenaria e o isolamento acústico e térmico é igual ou superior a uma parede de tijolos da alvenaria. Uma divisória de gesso acartonado com 12cm de espessura, com lã mineral no interior, pode proporcionar o mesmo isolamento acústico que um muro de concreto de 12cm de espessura, +-40dB, mas com 10% do peso. Dependendo da obra, é possível instalar até 50 metros quadrados de uma nova parede em um dia e deixá-la pronta para pintura (Acustica, 2008).

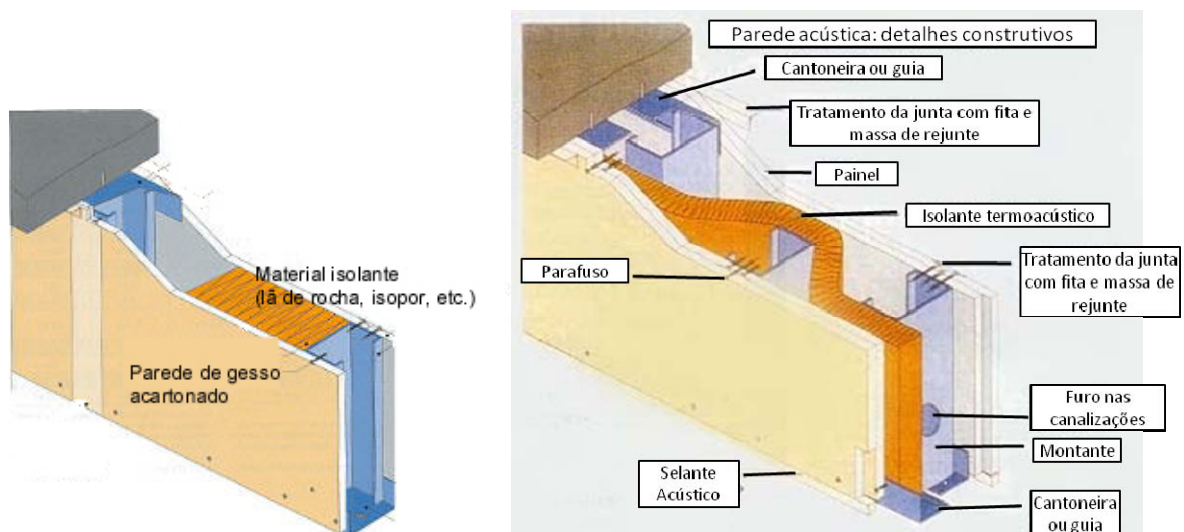


Figura A-9 – Detalhe de uma parede para isolamento acústico (RAU, 1999)

A.1.3 - Vídeo

Os avanços em direção à alta definição possibilitaram uma grande melhoria de qualidade de imagem, mas para que essa alta definição seja aproveitada é necessário que os equipamentos periféricos de exibição trabalhem também com alta definição. Quando se adquire um equipamento de videoconferência capaz de enviar e receber imagens com resoluções de 1280x720 ou 1920x1080 já se percebe uma melhoria de qualidade mesmo quando se utilizam TVs e projetores tradicionais. Isso ocorre porque a capacidade das câmeras e dos codificadores é mais avançada.

Para salas de videoconferência, atualmente opta-se pelo uso de TVs de LCD (*Liquid Crystal Display*) ou plasma. Essas TVs estão disponíveis no mercado atual em tamanhos

maiores do que as TVs CRT e oferecem imagens mais claras, menor uso de energia elétrica e maior praticidade, podendo se adequar à maioria dos layouts e mobiliário das salas de videoconferência. Existem televisores CRT (*Cathode ray tube*) capazes de reproduzir imagens em até 1080p, mas devido ao tamanho destes e a entrada no mercado dos monitores LCD e plasma, essa tecnologia tornou-se praticamente obsoleta no mercado de videoconferência.

É preciso também levar em consideração a distância entre os participantes de uma conferência e o monitor. Para salas grandes de videoconferência, onde os participantes estão 3 a 5 metros distantes do monitor, é recomendado um monitor de pelo menos 50 polegadas. A Tabela A-2 sugere tamanho de telas com a distância dos participantes.

Tabela A-3 – Relação entre tamanho da tela e distância dos participantes

Tamanho da tela (polegadas) (16:9)	Distância (metros)
30	1,91
32	2,13
42	2,67
50	3,17
57	3,63
60	3,81
65	4,11

Sinais digitais em alta definição exigem também o uso de conectores específicos. Os conectores comuns RCA e S-Video utilizados em equipamentos anteriores à alta definição em muitos casos não são mais compatíveis com os novos equipamentos. A Tabela A-4 mostra os conectores utilizados para videoconferência e os sinais que são trafegados.

Tabela A-4 – Sinais e conectores de vídeo

<p>Vídeo Composto, S-Video</p>	<p>Carregam o mesmo sinal de vídeo analógico Y/C. Com o conector S-Video os sinais Y e C são transmitidos separadamente permitindo maior qualidade de vídeo.</p>	
<p>VGA 15 pinos</p>	<p>Transporta três sinais de cores separadamente (RGB) e opcionalmente também dois sinais para sincronismo. Muito usado por PCs e também telas de LCD.</p>	
<p>Vídeo Componente (suporta HD)</p>	<p>Trafega o sinal YPbPr (ou RGB em alguns casos) e exige menor banda quando comparado ao RGB. Provê um sinal de luminância e dois de crominância.</p>	
<p>DVI-D (suporta HD)</p>	<p>Conector DVI projetado para trafegar dado puramente digital.</p>	
<p>DVI-A</p>	<p>Conector DVI projetado para trafegar vídeo analógico. Tipicamente usado para conectar em monitores VGA trafegando RGB ou YPbPr</p>	
<p>HDMI (suporta HD)</p>	<p>Projetado para trafegar vídeo e áudio digital.</p>	

Em várias aplicações para videoconferência é comum o uso de projetores para que se obtenha uma imagem com tamanhos maiores do que as disponibilizadas pelas TVs. Para estes

ambientes duas opções estão disponíveis: a projeção frontal e a traseira. A projeção frontal é a mais barata e fácil de implementar, mas existe um conflito de interesses entre a projeção e a captação de imagens pela câmera local. Tipicamente, as projeções frontais são melhor visualizadas em ambientes com baixa iluminação o que dificulta a captação das imagens na sala degradando a qualidade da videoconferência. A melhor opção é usar a projeção traseira como mostrado na Figura10, pois opera em ambientes com maior luminosidade.



Figura A-10 – Televisores de projeção traseira (HSM, 2009)

Alguns pontos devem ser levados em consideração tanto para projeções frontais quanto traseiras. Uma projeção frontal pode oferecer um maior tamanho de tela, mas é necessária uma distância boa até a projeção. Uma regra prática (válida também para projeção traseira) é que a distância da tela seja pelo menos uma vez e meia o tamanho da tela em diagonal. Em outras palavras, se você quer uma tela de projeção frontal de 100 polegadas (250 cm), precisará sentar-se a cerca de 3,7 metros de distância (Polycom, 2008).

Outros fatores a serem considerados são formatos de imagem suportados (4:3 e 16:9), resolução (para *full* HD deve suportar 1920x1080p), taxa de amostragem (até 60 quadros por segundo), nível de preto, nível de contraste, qualidade da tela, conexões aceitas, manutenção e luminosidade (para salas de videoconferência, 2500 lumens ou mais).

Tabela A-5 - Padrões primários de TV digital (HSM, 2009)

HDTV	1920 x 1080	16:9	24p, 30p, 60i
	1280 x 720	16:9	24p, 30p, 60p
SDTV	704 x 480	16:9	24p, 30p, 60i, 60p
	704 x 480	4:3	24p, 30p, 60i, 60p
	640 x 480	4:3	24p, 30p, 60i, 60p