



**Universidade Federal de Santa Catarina  
Deptº de Informática e de Estatística  
Curso de Mestrado Fora de Sede -  
UNOESC**



# *Sistemas Multimídia Distribuídos*

**Prof. Roberto Willrich**

Florianópolis, Março de 2000

# Índice

<b>Capítulo 1</b>	<i>Introdução à Multimídia</i>	8
<b>1.1</b>	O que é multimídia?	8
1.1.1	Etimologia .....	8
1.1.2	Uma definição inicial.....	8
<b>1.2</b>	Uso de Multimídia como adjetivo .....	8
<b>1.3</b>	Classificação dos tipos de mídia .....	9
1.3.1	Mídias capturadas <i>versos</i> sintetizadas.....	9
1.3.2	Mídias discretas <i>versos</i> contínuas .....	9
<b>1.4</b>	Motivações e Aplicações de multimídia.....	9
1.4.1	Entretenimento: Aumento da Interatividade .....	10
1.4.2	Informação sob demanda (IOD) .....	11
1.4.3	Educação .....	11
1.4.4	Telemedicina .....	11
1.4.5	Videofonia e Videoconferência.....	11
1.4.6	Trabalho cooperativo suportado por computador (CSCW) .....	12
<b>1.5</b>	Classes de Sistemas Multimídia.....	12
1.5.1	Sistemas Multimídia Standalone.....	12
1.5.2	Sistemas Multimídia Distribuídos .....	12
<b>1.6</b>	Desafios da Multimídia.....	13
<b>Capítulo 2</b>	<i>Dados Multimídia</i> .....	14
<b>2.1</b>	Representação Digital de Informações Multimídia .....	14
2.1.1	Sinais Analógicos para representar informações .....	14
2.1.2	Porque Digitalizar? .....	15
2.1.3	Digitalização, Amostragem e Quantificação .....	16
2.1.4	Problemas da Representação digital.....	18
<b>2.2</b>	Áudio.....	19
2.2.1	Representação digital de áudio .....	19
2.2.2	Características da fala .....	20
2.2.3	Áudio não Voz.....	20
2.2.4	Representação simbólica da música: o padrão MIDI .....	20
<b>2.3</b>	Vídeos e Imagens Analógicos .....	21
2.3.1	Captura e reprodução de imagens e vídeos analógicos .....	21
2.3.2	Vídeos e Imagens Coloridos.....	21
<b>2.4</b>	Representação digital de imagens e vídeos .....	25
2.4.1	Um Sistema Genérico de Vídeo Digital .....	25
2.4.2	Imagens Digitais .....	26
2.4.3	Imagens e Gráficos Animados .....	26
2.4.4	Especificação da Cor.....	28
2.4.5	Sistema RGB .....	28
2.4.6	Sistema CIE 1931 XYZ.....	30
<b>2.5</b>	Principais características e Requisitos das Informações multimídia.....	30
2.5.1	Requisitos de armazenamento e largura de banda.....	30
2.5.2	Relações temporais e espaciais entre mídias .....	31
2.5.3	Requisitos de atrasos e variações de atrasos (Jitter) .....	31
2.5.4	Tolerância a erros e perdas em dados multimídia .....	32
<b>2.6</b>	Qualidade de Serviço.....	32
<b>Capítulo 3</b>	<i>Compressão de Dados Multimídia</i> .....	33
<b>3.1</b>	A Necessidade da Compressão.....	33

<b>3.2 Princípios de Compressão.....</b>	<b>34</b>
3.2.1 Redundância de Dados .....	34
3.2.2 Propriedades Percepção Humana.....	34
<b>3.3 Classificação das Técnicas de Compressão.....</b>	<b>35</b>
3.3.1 Técnicas de compressão com e sem perda .....	35
3.3.2 Técnicas por Entropia, Codificação na Origem e Híbrida.....	35
3.3.3 Codificação com taxa de bits constantes e variáveis .....	35
<b>3.4 Medição do Desempenho de Compressão.....</b>	<b>36</b>
<b>3.5 Técnicas de Compressão sem Perdas .....</b>	<b>36</b>
3.5.1 Codificação Run-Length.....	36
3.5.2 Codificação de Huffman .....	37
3.5.3 Codificação de Lempel-Ziv-Welch (LZW) .....	38
<b>3.6 Técnicas de Compressão de Áudio Digital .....</b>	<b>40</b>
3.6.1 Quantificação não linear.....	40
3.6.2 Codificação Predictiva.....	40
3.6.3 Sumário das recomendações ITU-TS para codificação de voz .....	41
3.6.4 MPEG Áudio .....	41
<b>3.7 Técnicas de Compressão de Vídeo e Imagem Digital .....</b>	<b>43</b>
3.7.1 Técnica de Truncagem.....	43
3.7.2 Codificação de Sub-Amostragem Espacial e Temporal.....	43
3.7.3 Codificação Predictiva.....	44
3.7.4 Preenchimento Condicional.....	44
3.7.5 Estimativa e Compensação de Movimento .....	44
3.7.6 Técnica Transform Coding .....	44
<b>3.8 Padrões de Compressão Multimídia .....</b>	<b>45</b>
3.8.1 Padrão de Compressão de Imagens JPEG .....	45
3.8.2 CCITT H.261 .....	47
3.8.3 H.263 .....	48
3.8.4 ISO/IEC MPEG (Motion Picture Expert Group).....	48
<b>3.9 Implementando Algoritmos de Compressão.....</b>	<b>52</b>
<b>Capítulo 4 Requisitos de Hardware e Sistema Operacional para Sistemas Multimídia.....</b>	<b>54</b>
<b>4.1 Introdução as Bases dos Sistemas Computacionais .....</b>	<b>54</b>
4.1.1 Unidade Central.....	54
4.1.2 Barramentos e Entrada/Saída .....	54
<b>4.2 Requisitos de hardware de terminais multimídia.....</b>	<b>54</b>
4.2.1 Dispositivos de Armazenamento Multimídia .....	55
4.2.2 Sistemas de Captura e Apresentação Multimídia.....	56
4.2.3 Multimídia em Rede.....	56
4.2.4 Servidores Multimídia .....	56
4.2.5 Apresentação de Áudio e Vídeo.....	57
<b>4.3 Requisitos de Sistema Operacional .....</b>	<b>59</b>
<b>Capítulo 5 Aplicações Multimídia.....</b>	<b>60</b>
<b>5.1 Aplicações Pessoa-a-Pessoa e Pessoa-a-Sistema.....</b>	<b>60</b>
5.1.1 Aplicações Multimídia Pessoa a Pessoa.....	60
5.1.2 Aplicações Multimídia Pessoa-a-Sistema .....	61
<b>5.2 Taxonomia Recomendada pela ITU.....</b>	<b>61</b>
<b>5.3 Telefonia Assistida por Computador .....</b>	<b>62</b>
5.3.1 Telefonia a Circuito Assistida por Computador.....	62
5.3.2 Conversação por Pacotes de Voz.....	63
<b>5.4 Videofonia.....</b>	<b>64</b>
5.4.1 Videofonia a Circuito Integrada por Computador .....	64

5.4.2	Videofonia sob redes a pacote.....	65
<b>5.5</b>	<b>Espaço Compartilhado para Trabalho Cooperativo Integrado por Computador (CSCW) ....</b>	<b>65</b>
5.5.1	Ferramentas de Quadro Branco Compartilhado .....	66
5.5.2	Ferramentas de aplicação compartilhada.....	67
<b>5.6</b>	<b>Distribuição de Áudio e Vídeo .....</b>	<b>68</b>
5.6.1	Campos de Aplicação .....	68
5.6.2	Características da distribuição de vídeo de apresentações .....	69
5.6.3	Distribuição de Áudio e Vídeo Institucional.....	69
5.6.4	Distribuição de Áudio e Vídeo em WANs .....	70
<b>5.7</b>	<b>Áudio-Videoconferência.....</b>	<b>71</b>
5.7.1	Principais Características e Requisitos .....	72
5.7.2	Videoconferência baseada em circuitos e pacotes .....	73
5.7.3	Videoconferência a Pacotes com comunicação ponto-a-multiponto.....	74
5.7.4	Requisitos de Rede .....	75
<b>5.8</b>	<b>Aplicações Baseadas em Servidores Multimídia.....</b>	<b>76</b>
5.8.1	Acesso Interativo a Servidores Multimídia .....	76
5.8.2	Comunicação Assíncrona e Síncrona.....	77
5.8.3	Problemas na Transmissão Síncrona (Tempo-Real) .....	78
5.8.4	Vídeo sob-demanda (VOD) .....	78
<b>5.9</b>	<b>Realidade Virtual .....</b>	<b>79</b>
<b>Capítulo 6 Documentos Multimídia e Hipermídia.....</b>		<b>80</b>
<b>6.1</b>	<b>Definição de Documento Multimídia e Hipermídia.....</b>	<b>80</b>
6.1.1	Documentos Hipertexto .....	80
6.1.2	Documentos Multimídia .....	81
6.1.3	Documentos Hipermídia .....	81
<b>6.2</b>	<b>Autoria de Documentos Multimídia.....</b>	<b>82</b>
6.2.1	Evolução histórica .....	82
<b>6.3</b>	<b>Criação de documentos multimídia.....</b>	<b>83</b>
6.3.1	Análise e Projeto Preliminares .....	84
6.3.2	Criação e aquisição de materiais .....	84
6.3.3	Composição de Documentos Multimídia .....	85
6.3.4	Avaliação e refinamento.....	85
<b>6.4</b>	<b>Requisitos para um Modelo Multimídia .....</b>	<b>86</b>
6.4.1	Estrutura de Conteúdo .....	87
6.4.2	Estrutura Conceptual.....	87
6.4.3	Estrutura de Apresentação.....	91
6.4.4	Interações.....	92
<b>6.5</b>	<b>Abordagens para Autoria de Documentos Multimídia.....</b>	<b>92</b>
6.5.1	Linguagens Scripting .....	92
6.5.2	Abordagem Baseada em Informação .....	93
6.5.3	Linha temporal (Timeline).....	93
6.5.4	Modelo de composição via pontos de referência .....	94
6.5.5	Composição Hierárquica .....	95
6.5.6	Modelos baseados em ícones.....	95
6.5.7	Modelos Baseados em Cartões ou Páginas .....	96
6.5.8	Redes de Petri.....	96
<b>6.6</b>	<b>Apresentação de Documentos Multimídia.....</b>	<b>97</b>
<b>6.7</b>	<b>Padronização de Formatos de Documentos Multimídia.....</b>	<b>98</b>
6.7.1	Standard Generalized Markup Language (SGML) .....	98
6.7.2	Office Document Architecture (ODA) .....	99
6.7.3	Hypermedia/Time-based Structuring Language (HyTime) .....	99
6.7.4	Multimedia and Hypermedia Expert Group (MHEG) .....	100
<b>6.8</b>	<b>Exemplos de Modelos Multimídia e Sistemas de Autoria .....</b>	<b>100</b>

6.8.1	Modelo de Referência Hipertexto Dexter .....	100
6.8.2	Amsterdam Hypermedia Model (AHM) .....	102
6.8.3	Modelo de Composição Hierárquico de [Hudson, 93] .....	103
6.8.4	Sistema Firefly .....	104
6.8.5	Asymetrix Toolbook .....	104
6.8.6	Sistemas Baseados em Ícones IconAuthor e Authoware .....	105
6.8.7	Macromedia Director .....	106
6.8.8	HTSPN (Hierarchical Time Stream Petri Nets) .....	107

## **Capítulo 7 Requisitos de Comunicação para Sistemas Multimídia Distribuídos..... 110**

<b>7.1</b>	<b>Parâmetros de Desempenho de Redes .....</b>	<b>110</b>
7.1.1	Velocidade da Rede, Largura de Banda e Taxa de bits.....	110
7.1.2	Taxa de bits agregada e individual .....	110
7.1.3	Velocidade de acesso e taxa de bits.....	111
7.1.4	Vazão (Throughput).....	111
7.1.5	Taxa de Erro .....	111
7.1.6	Atraso Fim-a-Fim .....	112
7.1.7	Atraso Ida-e-Volta .....	112
7.1.8	Variação de atraso (Jitter) .....	112
<b>7.2</b>	<b>Caracterização do Tráfego Multimídia.....</b>	<b>113</b>
7.2.1	Tipos de transferência: Assíncrona e Síncrona.....	113
7.2.2	Variação de vazão com o tempo .....	113
7.2.3	Dependência temporal .....	114
7.2.4	Continuidade Temporal .....	114
<b>7.3</b>	<b>Requisitos para Transmissão de Áudio e Vídeo.....</b>	<b>114</b>
7.3.1	Sistemas multimídia versus sistemas tempo-real críticos.....	115
7.3.2	Requisitos de vazão .....	115
7.3.3	Requisitos de confiabilidade (controle de erro) .....	116
7.3.4	Requisitos de atraso e variação de atraso .....	116
7.3.5	Compartilhamento eficiente de recursos de rede .....	118
7.3.6	Garantias de Desempenho .....	120
7.3.7	Adaptabilidade da Rede .....	120
7.3.8	Capacidade de Multicasting .....	121

## **Capítulo 8 Garantias de QoS Fim-a-Fim para Comunicação de Áudio e Vídeo..... 122**

<b>8.1</b>	<b>Componentes dos sistemas multimídia e suas influências no desempenho .....</b>	<b>122</b>
8.1.1	Principais componentes das aplicações pessoa-a-pessoa.....	122
8.1.2	Principais componentes das aplicações pessoa-a-sistema .....	123
<b>8.2</b>	<b>Gerenciamento de Qualidade de Serviço .....</b>	<b>123</b>
8.2.1	Definição de QoS .....	124
8.2.2	Estrutura Geral de QoS .....	124
8.2.3	Especificação da QoS .....	125
8.2.4	Negociação e Renegociação de QoS .....	126
8.2.5	Outros usos da declaração de QoS .....	127
8.2.6	Diferentes Níveis de Garantia .....	127
8.2.7	Fornecendo Garantias de QoS .....	128
8.2.8	Um exemplo de manipulação de QoS .....	128
<b>8.3</b>	<b>Reserva de Recursos em Avanço .....</b>	<b>129</b>
<b>8.4</b>	<b>Aplicações Adaptativas .....</b>	<b>129</b>

## **Capítulo 9 Suportes de Rede para Multimídia..... 130**

<b>9.1</b>	<b>Requisitos de Rede para Comunicação Multimídia.....</b>	<b>130</b>
<b>9.2</b>	<b>Alguns Conceitos Importantes .....</b>	<b>130</b>
9.2.1	Transmissão de Dados Assíncrona e Síncrona .....	130
9.2.2	Multiplexação assíncrona e síncrona .....	131
9.2.3	Comunicações Isócronas .....	132

<b>9.3</b>	<b>ISDN.....</b>	<b>132</b>
<b>9.4</b>	<b>Ethernet.....</b>	<b>135</b>
<b>9.5</b>	<b>Isochronous Ethernet .....</b>	<b>136</b>
<b>9.6</b>	<b>Fast Ethernet (100Base-T) .....</b>	<b>137</b>
<b>9.7</b>	<b>Gigabit Ethernet.....</b>	<b>137</b>
<b>9.8</b>	<b>Token Ring .....</b>	<b>138</b>
<b>9.9</b>	<b>FDDI .....</b>	<b>139</b>
<b>9.10</b>	<b>FDDI-II.....</b>	<b>140</b>
<b>9.11</b>	<b>ATM (Asynchronous Transfer Mode).....</b>	<b>141</b>
9.11.1	Como o tamanho da célula foi determinado .....	143
9.11.2	Modelo de Referência B-ISDN .....	144
9.11.3	Camada Física .....	145
9.11.4	Elementos básicos ATM .....	147
9.11.5	Formato das Células ATM.....	148
9.11.6	Procedimento de estabelecimento de chamada .....	149
9.11.7	Roteamento de células .....	151
9.11.8	Camadas de Adaptação ATM.....	152
9.11.9	Gerenciamento de tráfego .....	153
9.11.10	ATM satisfaz os requisitos de comunicação multimídia .....	154
<b>9.12</b>	<b>Outras tecnologias .....</b>	<b>156</b>

## **Capítulo 10 Protocolos de Transporte para Comunicação Multimídia..... 157**

<b>10.1</b>	<b>Requisitos de Protocolos de Transporte Multimídia.....</b>	<b>157</b>
10.1.1	Alta Vazão .....	157
10.1.2	Capacidades Multicast.....	157
10.1.3	Especificação e Garantias de QoS.....	157
<b>10.2</b>	<b>Protocolos de transporte tradicionais .....</b>	<b>158</b>
10.2.1	Cópia de Dados.....	158
10.2.2	Multiplexação em camadas.....	158
10.2.3	Controle de Fluxo .....	158
10.2.4	Controle de Erro .....	159
10.2.5	Falta de Suporte de QoS .....	159
<b>10.3</b>	<b>Xpress Transfer Protocol (XTP) .....</b>	<b>160</b>
<b>10.4</b>	<b>Protocolos de Reserva de Recurso.....</b>	<b>161</b>
10.4.1	Protocolo ST-II .....	161
10.4.2	Protocolo RSVP .....	163
10.4.3	Comparação de ST-II e RSVP .....	165
10.4.4	Filtragem de mídia .....	166
<b>10.5</b>	<b>Protocolo RTP .....</b>	<b>167</b>
10.5.1	Parte de dados do RTP.....	168
10.5.2	Parte de controle do RTP .....	169
10.5.3	RTP e a comunicação multimídia .....	170
<b>10.6</b>	<b>RTSP (Real-Time Streaming Protocol).....</b>	<b>170</b>

## **Capítulo 11 Servidores Multimídia..... 172**

<b>11.1</b>	<b>Separação do controle e transferência de dados .....</b>	<b>172</b>
<b>11.2</b>	<b>Requisitos de Servidores Multimídia.....</b>	<b>173</b>
<b>11.3</b>	<b>Dispositivos de armazenamento .....</b>	<b>175</b>
11.3.1	Requisitos de capacidade de armazenamento e Largura de banda de transferência .....	175
11.3.2	Tipos de dispositivos de armazenamento .....	175
11.3.3	Vetores de Disco e RAID.....	176

11.3.4	Hierarquia de armazenamento.....	176
11.3.5	Vetor de Servidores .....	177
<b>11.4</b>	<b>Posicionamento de dados no disco.....</b>	<b>177</b>
11.4.1	Posicionamento de dados de maneira contínua.....	178
11.4.2	Posicionamento espalhado de dados .....	179
11.4.3	Posicionamento de dados log-structured.....	179
11.4.4	Posicionamento de dados em vetores de discos .....	179
<b>11.5</b>	<b>Escalonamento de Disco e Controle de Admissão.....</b>	<b>180</b>
11.5.1	Algoritmos de escalonamento de disco tradicionais.....	181
11.5.2	Algoritmo EDF ( <i>Earliest Deadline First</i> ).....	181
11.5.3	Algoritmo Scan-EDF .....	181
11.5.4	Algoritmo Round-Robin .....	182
11.5.5	Escalonamento GSS ( <i>Group Sweeping Scheduling</i> ).....	182
<b>11.6</b>	<b>Interações com o usuário.....</b>	<b>183</b>
11.6.1	Pausa e Retomada.....	183
11.6.2	Avanço e retrocesso rápidos.....	184
11.6.3	QoS e interações com o usuário .....	185
<b>11.7</b>	<b>Configuração de Servidor e Conexão a Rede .....</b>	<b>185</b>
<b>Capítulo 12</b>	<b>Sincronização Multimídia.....</b>	<b>187</b>
<b>12.1</b>	<b>Requisitos de Sincronização .....</b>	<b>187</b>
12.1.1	Classes de aplicação e a requisitos de sincronização .....	187
12.1.2	Medidas do nível de sincronismo .....	188
12.1.3	Causas da perda de sincronização multimídia .....	190
<b>12.2</b>	<b>Mecanismos para Sincronização Multimídia .....</b>	<b>191</b>
12.2.1	Medidas de contenção de variações de atraso na rede .....	191
12.2.2	Medidas de contenção de desvios de processamento.....	193
12.2.3	Medidas de contenção de diferenças de taxa de relógio .....	193
12.2.4	Medidas de contenção dos problemas de perdas de pacote.....	194
12.2.5	Medidas para coordenar múltiplas fontes para transmissão síncrona.....	194
12.2.6	Supporte da estação de trabalho e servidor para mídias contínuas .....	195
12.2.7	Medidas para fornecer sincronização de interações com o usuário .....	195
12.2.8	Técnicas de apresentação para obter efeitos de apresentação ótimos.....	195
<b>Capítulo 13</b>	<b>Internet e a Multimídia.....</b>	<b>196</b>
<b>13.1</b>	<b>Introdução à Rede Internet .....</b>	<b>196</b>
13.1.1	Arquitetura Internet.....	196
13.1.2	Protocolo IP .....	197
13.1.3	IP Multicast.....	198
13.1.4	IPv6.....	200
<b>13.2</b>	<b>TCP/IP e a Multimídia.....</b>	<b>201</b>
<b>13.3</b>	<b>Qualidade de Serviço na Internet .....</b>	<b>203</b>
13.3.1	Serviços Integrados/RSVP .....	204
13.3.2	Serviços Diferenciados.....	205
<b>13.4</b>	<b>Rede MBone.....</b>	<b>207</b>
13.4.1	Arquitetura MBone .....	208
13.4.2	Situação atual .....	209
13.4.3	Ferramentas de Anúncio de Sessão.....	210
13.4.4	Ferramentas de Áudio .....	212
13.4.5	Ferramentas de Vídeo .....	214
13.4.6	MBone VCR: um servidor de vídeo para a MBone .....	217
13.4.7	Ferramentas de Quadro Branco Compartilhados .....	218
13.4.8	Editor de texto compartilhado .....	219
13.4.9	Ferramenta MultiTalk .....	220
<b>13.5</b>	<b>WWW .....</b>	<b>220</b>

13.5.1	Problemas da Web para multimídia .....	221
13.5.2	Suportes de áudio e vídeo na Web.....	222
13.5.3	Compartilhamento multicast de páginas Web.....	224
<b>Capítulo 14</b>	<b>Padrão de Videoconferência H.323.....</b>	<b>226</b>
<b>14.1</b>	<b>Introdução.....</b>	<b>226</b>
<b>14.2</b>	<b>Arquitetura H.323.....</b>	<b>228</b>
14.2.1	Terminais .....	228
14.2.2	Gateways.....	228
14.2.3	Gatekeepers .....	229
14.2.4	Multipoint Control Units .....	230
<b>14.3</b>	<b>Protocolos Definidos pelo H.323.....</b>	<b>230</b>
14.3.1	CODEC de Áudio .....	231
14.3.2	CODEC de Vídeo.....	231
14.3.3	H.225 Layer.....	232
14.3.4	H.225 Registro, Admissão e Status.....	232
14.3.5	H.225 Sinalização de Chamada.....	232
14.3.6	H.245 Sinalização de Controle .....	233
14.3.7	RTP ( <i>Real-Time Transport Protocol</i> ).....	233
14.3.8	RTCP ( <i>Real-Time Transport Control Protocol</i> ).....	233
14.3.9	Dados.....	233
<b>14.4</b>	<b>Características dos Terminais H.323 .....</b>	<b>234</b>
<b>14.5</b>	<b>Unidades de Controle Multiponto (MCUs).....</b>	<b>234</b>
<b>14.6</b>	<b>Procedimentos de Conexão.....</b>	<b>236</b>
<b>14.7</b>	<b>H.323 em Redes IP .....</b>	<b>239</b>
<b>14.8</b>	<b>Ferramentas H.323.....</b>	<b>240</b>
14.8.1	Servidores H.323 .....	240
14.8.2	Cientes H.323 – Características Descritas .....	241
<b>Capítulo 15</b>	<b>Bibliografia.....</b>	<b>242</b>

# Capítulo 1 Introdução à Multimídia

O desenvolvimento tecnológico dos sistemas de informações e de comunicações a alta velocidade permitiram o aparecimento de novas aplicações no domínio de sistemas distribuídos. Uma das principais tendências neste contexto é a integração de diferentes tipos de mídias, como o texto, a voz, as imagens e os vídeos em um vasto domínio de aplicações informáticas distribuídas (por exemplo, tele-ensino, publicações eletrônicas, trabalho cooperativo). Estes sistemas são chamados de sistemas multimídia e hipermídia.

## 1.1 O que é multimídia?

Multimídia é um dos termos mais usados nesta década. Este termo está no cruzamento de cinco grandes indústrias [Fluckiger, 95]: informática, telecomunicações, publicidade, consumidores de dispositivos de áudio e vídeo, indústria de televisão e cinema. Este grande interesse destas indústrias poderosas contribuíram para a grande evolução da multimídia.

### 1.1.1 Etimologia

A palavra multimídia é composta de duas partes: o prefixo multi e o radical mídia:

- **Multi:** originário da palavra latina *multus* que significa “numerosos”. O uso deste prefixo não é recente e muitas palavras de origem latina empregam este radical, como *multiformis* (que tem várias formas) ou *multicolor* (várias cores).
- **Mídia:** plural da palavra latina *medium* que significa meio, centro. Ele é derivado do adjetivo *medius*, que está no centro. No contexto de multimídia, este radical refere-se ao tipo de informação ou tipo de portador de informação, como dados alfanuméricos, imagens, áudio, vídeo, etc.

### 1.1.2 Uma definição inicial

[Fluckiger, 95] define multimídia como segue abaixo:

MULTIMÍDIA É O CAMPO INTERESSADO NA INTEGRAÇÃO CONTROLADA POR COMPUTADOR DE TEXTOS, GRÁFICOS, IMAGENS, VÍDEOS, ANIMAÇÕES, SONS, E QUALQUER OUTRO MEIO ONDE TODO TIPO DE INFORMAÇÃO PODE SER REPRESENTADO, ARMAZENADO, TRANSMITIDO, E PROCESSADO DIGITALMENTE.

## 1.2 Uso de Multimídia como adjetivo

A palavra multimídia pode ser usada como um nome quando nós nos referimos ao campo, mas é mais freqüentemente empregada como um adjetivo. Abaixo é listado vários termos compostos:

- **Aplicação multimídia:** uma função ou conjunto de funções que suportam a provisão de um serviço multimídia específico para o usuário final. P.e. aplicação de áudio-videoconferência.
- **Tecnologia multimídia:** conjunto de campos tecnológicos específicos habilitando aplicações multimídia. P.e. compressão de vídeo.
- **Plataforma multimídia:** um tipo específico e configuração de computador equipado com hardware multimídia, e capaz de suportar softwares de aplicações multimídia. P.e. PC com placas de áudio e vídeo.
- **Placa multimídia:** hardware conectado com o computador e executando uma função multimídia. P.e. placa de captura de vídeo para digitalizar vídeo analógico.

- **Dispositivo de armazenamento multimídia:** um dispositivo capaz de armazenar vários tipos de informações, provendo o acesso a estes. P.e. CD-ROM, tape de áudio digital.
- **Rede multimídia:** rede de comunicação de dados que permite o transporte de múltiplos tipos de informações. P.e. Rede Digital de Serviços Integrados (RDSI ou em inglês ISDN).

## 1.3 Classificação dos tipos de mídia

Nós vimos que sistemas multimídia são aqueles que integram vários tipos de mídia, como texto, imagens, sons e o vídeo. Esta definição certamente não é suficientemente precisa para a definição de multimídia. Nestes termos qualquer sistema integrando dois ou mais tipos de informações poderiam ser classificados como sendo multimídia. Por exemplo, como qualquer arquivo *Word* integra basicamente três tipos de mídia, texto, gráficos e imagens, este editor de texto poderia ser classificado erroneamente como sendo uma aplicação multimídia.

Para definir melhor o termo multimídia, deve-se inicialmente conhecer a classificação de tipos de mídias. Os vários tipos de mídia podem ser classificados de acordo com vários esquemas. A seguir serão apresentadas duas classificações: mídias capturadas versus sintetizadas e mídias discretas versus contínuas.

### 1.3.1 Mídias capturadas versus sintetizadas

Este esquema refere-se a separação tradicional entre tipos de informações capturadas do mundo real (imagens, vídeos e sons) e informações sintetizadas pelo computador (texto, gráficos e animações). A partir desta classificação, pode-se definir aplicações multimídia como sendo aquelas compostas de ao menos um tipo de mídia sintetizado e um tipo de mídia capturada. Na prática, esta regra não é seguida estritamente, em particular pois a classificação em si não casa com a realidade – sons e imagens podem ser sintetizados, textos podem ser capturados do mundo real por sistemas óticos de reconhecimento de caracteres.

### 1.3.2 Mídias discretas versus contínuas

Este esquema distingue entre mídias discretas ou estáticas, que são mídias com dimensões unicamente espaciais (como textos, imagens, gráficos) e mídias contínuas (também chamadas de mídias dinâmicas ou isócronas), que são mídias com dimensões temporais (como sons, vídeos e animações). Estas duas classes são também comumente chamadas de mídias dependentes do tempo e mídias independente do tempo. Uma informação dependente do tempo é apresentada como um fluxo de mídia (fluxo de vídeo, fluxo de áudio).

No caso de mídias contínuas ou mídias dependentes do tempo, seus significados e correções dependem da taxa em que são apresentadas. Por exemplo, um vídeo consiste de um número de quadros ordenados, cada um destes quadros tem uma duração de apresentação fixa.

Mídia discreta ou independente do tempo é qualquer espécie de mídia tradicionalmente utilizada em documentos impressos, como texto e imagens. A semântica da informação não depende do domínio do tempo.

A definição de sistemas multimídia mais aceita é a seguinte:

SISTEMA MULTIMÍDIA É UM SISTEMA CAPAZ DE MANIPULAR AO MENOS UM TIPO DE MÍDIA DISCRETA E UM TIPO DE MÍDIA CONTÍNUA, AS DUAS NUMA FORMA DIGITAL.

## 1.4 Motivações e Aplicações de multimídia

A motivação do uso de dados multimídia em sistemas computacionais é o aumento da transferência de informações pelo uso simultâneo de um ou mais sensores do usuário. Isto

pois humanos aprendem mais, e mais rapidamente, quando vários dos seus sentidos (visão, audição, etc.), são utilizados.

Outro objetivo das aplicações multimídia é emular a comunicações humanas face-a-face. Isto tem levado à contínua investigação de sistemas de comunicação e computação que se aproximam da velocidade de transmissão, fidelidade e eficiência das comunicações humanas face-a-face.

Áudio, vídeo e imagens são os meios naturais da comunicação entre humanos. Portanto, sistemas multimídia têm aplicações em todas as áreas que estas mídias possam ser utilizadas. Nas seções seguintes são apresentados alguns exemplos de aplicações multimídia.

#### 1.4.1 Entretenimento: Aumento da Interatividade

O uso da multimídia para entretenimento não é recente. Hoje a escala e sofisticação dos jogos e filmes avançarão drasticamente em três direções [Buford, 94]: Vídeo sob demanda, cinema interativo e jogos colaborativos suportados por computador.

##### **Vídeo/Filme sob demanda (VOD)**

Atualmente nós assistimos programas de TV e filmes passivamente: nós não podemos interagir com os programas e não podemos controlar o tempo que queremos assisti-los. Serviços de vídeo/filme sob demanda estão sendo desenvolvidos pelas indústrias de TV, TV a cabo e companhias telefônicas para quebrar esta limitação e fornecer outras funcionalidades.

Em VOD, uma grande coleção de vídeos são armazenados em servidores de vídeo. Usuários ou clientes acessam estes vídeos através de uma rede. As principais vantagens de VOD são:

- Não teríamos que sair de casa para assistir nosso filme favorito. Nossa televisão seria conectado ao servidor de vídeo através de uma rede. Necessitariamnos apenas selecionar o vídeo através de um menu na TV.
- Como os servidores de vídeo são centralizados e fornecem serviços a uma grande população, sua coleção seria grande, atualizada e sempre disponível. Se o servidor não tem o vídeo que queremos, ele pode indicar outros servidores que possuem o vídeo. Adicionalmente, muitos usuários podem ver o mesmo filme sem, não havendo o problema como “Desculpe, o vídeo está fora”.
- Poderíamos assistir nosso filme predileto a qualquer instante que desejarmos.
- Poderíamos interromper, avançar e voltar à traz, ou mesmo pular para uma determinada cena.
- Nós teríamos alta qualidade, pois vídeos são armazenados digitalmente. A qualidade não degradará com o aumento de telespectadores.

As vantagens acima são de sistemas VOD verdadeiros. Outros sistemas, tal como sistemas quase VOD (vistos mais adiante neste curso), não suportam todos os mecanismos de interação dos VOD verdadeiros. Devido ao mercado potencial, muitas companhias e organizações estão pesquisando, experimentando, e prototipando sistemas VOD. O desafio é como suportar um grande número de telespectadores ao mesmo tempo com um custo razoável.

##### **Cinema Interativo**

Em cinemas interativos, a audiência não assistirá mais passivamente o filme pré-registrado, ela poderá ditar os próximos passos tomados pelo herói usando uma tecnologia de baixo custo. Por exemplo, o primeiro filme interativo foi *I'm You Man* de uma companhia de produção nova-iorquina, onde a cada 90 segundos a audiência escolhia uma de três opções em um controle a três botões no braço de cada acento. Os votos eram registrados e a opção mais votada era seguida pelo filme. A imprevisibilidade do filme após várias sessões mostrou-se sendo parte da atração.

### **Jogos Colaborativos Suportados por Computador**

Este tipo de aplicação multimídia não é uma novidade e já existem vários jogos colaborativos, onde vários jogadores podem participar do jogo ao mesmo instante.

#### **1.4.2 Informação sob demanda (IOD)**

É um sistema similar ao VOD, mas outros tipos de informação existem além do vídeo. Para o usuário, o sistema pode ser visto como uma biblioteca vasta e versátil. Nele, usuários emitiriam questões através de uma interface em uma TV ou em um computador e o sistema procuraria, obteria, transferiria e apresentaria informações ao usuário. A habilidade mais importante do sistema é indexar e procurar um vasto conjunto de informações multimídia. Sistemas IOD têm várias aplicações, incluindo:

- Agir como uma enciclopédia de informações gerais.
- Fornecer serviços online de jornais e revistas.
- Fornecer serviços de *home shopping* (poderia-se ver os produtos e serviços na tela e comprá-los).
- Fornecer informações atuais sobre o tempo, a tabela de horários do transporte público, transações bancárias.

O WWW (*World Wide Web*) pode ser visto como um IOD primitivo. O WWW deverá ser aprimorado para suportar acesso, transmissão e apresentação em tempo-real de informações multimídia. O conjunto e a qualidade das informações serão também aumentadas.

#### **1.4.3 Educação**

Sistemas multimídia são aplicados na educação. Pessoas aprendem mais e mais rapidamente quando elas podem ver, ouvir e trabalhar com novos conceitos, que tornam multimídia um meio natural de treinar e educar. Muitos sistemas de educação/treinamento multimídia atuais são sistemas *standalones* com o material armazenado em CD-ROMs que podem ser compartilhados por outros usuários. Isto mudará quando servidores multimídia em uma rede de longa distância (WAN – *Wide Area Network*) estiverem disponíveis:

- os cursos seriam compartilhados por várias pessoas,
- é conveniente que o estudante possa estudar em qualquer lugar e em qualquer hora,
- material de ensino pode ser organizado dinamicamente para seguir cada estudante, assim cada estudante pode aprender no seu estilo e passo,
- interação com técnicos será por e-mail e comunicação de áudio e vídeo.

#### **1.4.4 Telemedicina**

Telemedicina é outra aplicação multimídia importante, especialmente em caso de emergência em localizações remotas. Em telemedicina, todos os registros dos pacientes são armazenados eletronicamente. Instituições e equipamentos são conectados através de uma rede multimídia. A telemedicina oferece as seguintes vantagens:

- consulta com especialistas remotos através do uso de vídeo e áudio de alta qualidade,
- acesso ao registro de pacientes em qualquer lugar e a qualquer tempo,
- acesso mundial da informação tal como disponibilidade e necessidade de um tipo especial de sangue e órgão.

#### **1.4.5 Videofonia e Videoconferência**

Sistemas de Videofonia e videoconferência permitem que usuários dispersos se comuniquem efetivamente através de transmissões de áudio e vídeo. Muitos sistemas atuais utilizam máquinas específicas e redes a circuitos comutados (visto mais adiante). Eles são caros e não são facilmente disponíveis. No futuro, vídeos serão disponíveis nas

nossas janelas da estação de trabalho, comunicando sobre uma rede de pacotes chaveados de propósito geral. Videofones serão tão comuns como os telefones atuais.

#### 1.4.6 Trabalho cooperativo suportado por computador (CSCW)

Sistemas sofisticados de videoconferência suportarão o trabalho cooperativo: pessoas dispersas poderão trabalhar em um mesmo projeto através da transmissão de informações entre eles. Elas poderão acessar facilmente banco de dados multimídia e outros recursos.

### 1.5 Classes de Sistemas Multimídia

Esforços de pesquisa em computação multimídia estão divididos em dois grandes grupos [Furht, 94]: um grupo centra seus esforços em estações de trabalho multimídia *standalone*, sistemas de software associados e ferramentas; o outro grupo combina computação multimídia com sistemas distribuídos oferecendo maiores potencialidades.

#### 1.5.1 Sistemas Multimídia Standalone

Aplicações multimídia em sistemas *standalone* utilizam apenas os recursos presentes no sistema local para prover serviços multimídia. Assim, o sistema local fornece todo o poder de processamento necessário. Tais sistemas devem ser equipados com os dispositivos de captura e apresentação multimídia, tal como microfones e câmeras. Ele não faz uso de capacidades de armazenamento remoto, todas as informações necessárias devem ser armazenadas localmente (comunicação de dados multimídia não é suportada).

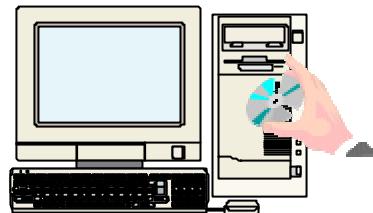


Figura 1. Sistema Multimídia Standalone

Alguns exemplos de aplicações de sistemas multimídia *standalone* são apresentados abaixo:

- Treinamento baseado em computador – técnicos e engenheiros ensinam operações de manutenção e procedimentos operacionais através de documentos multimídia em computadores pessoais ou estações de trabalho *standalone*. O documento é inteiramente armazenado em um dispositivo local, como CD-ROMs ou videodiscos.
- Educação baseada em computador individual - estudantes seguem cursos ou exercícios práticos usando aplicações multimídia em computadores pessoais *standalone*.
- Sistemas de Autoria Multimídia – com a qual autores podem desenvolver documentos multimídia comerciais ou simples apresentações privadas, tal como ferramentas de composição de música, ferramentas de autoria de documentos multimídia e hipermídia.

#### 1.5.2 Sistemas Multimídia Distribuídos

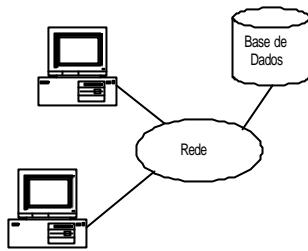
As razões para que sistemas multimídia podem necessitar de uma rede de comunicação são principalmente duas:

- Suportar aplicações em rede. O propósito de muitas aplicações multimídia é fornecer serviços de comunicação a distância. Exemplos são videoconferência, distribuição de pacotes de áudio e vídeo, E-mail multimídia.
- Implementação de modelos cliente-servidor. Aplicações *stand-alone* podem ser de alto custo para armazenamento de informações. Neste caso, a utilização de servidores remotamente acessados pelos sistemas clientes permitem o compartilhamento e armazenamento de informações.

### Classes de Sistemas Multimídia Distribuídos

Sistemas distribuídos podem ainda ser classificados em várias classes. A ITU (*International Telecommunications Union*) identifica quatro classes básicas de aplicações e serviços multimídia:

- **Serviços de conversão:** implica na interação entre um humano e outro humano ou sistema. Esta classe inclui serviços interpessoais tal como videoconferência, videofonia, telesegurança e telecompras.
- **Serviços de mensagem:** cobrem trocas de mensagem multimídia não tempo-real ou assíncronas entre mailboxes.
- **Serviços de recuperação:** cobrem todos os tipos de acesso a servidores de informações multimídia. Tipicamente, o usuário envia um pedido para o servidor e a informação pedida é liberada para o usuário em tempo-real. Vídeo sob demanda é um exemplo deste tipo de serviço.
- **Serviços de distribuição:** cobrem serviços onde a informação é distribuída sob a iniciativa de um servidor. Um exemplo destes serviços é a transmissão de TV.



**Figura 2. Sistema multimídia distribuído**

## 1.6 Desafios da Multimídia

Dados multimídia têm características e requisitos diferentes dos dados alfanuméricos para o qual sistemas computacionais tradicionais foram desenvolvidos. Assim, sistemas computacionais convencionais não suportam eficientemente aplicações multimídia. Novas tecnologias devem ser desenvolvidas e utilizadas. Abaixo é dado uma lista das principais características e requisitos das informações multimídia [Lu, 96]:

- Dados multimídia têm dimensão temporal e devem ser transmitidos, processados e apresentados em uma taxa fixa em muitas aplicações, assim processamento e comunicação multimídia devem satisfazer requisitos de tempo-real.
- Aplicações multimídia usam simultaneamente múltiplos tipos de mídia. As relações temporais e espaciais entre estas mídias devem ser mantidas.
- Dados multimídia são dados intensivos, assim eles devem ser compactados, e redes de alta velocidade e computadores potentes são necessários para manipular estes dados.
- Dados multimídia não têm sintaxe e semântica obvia. Banco de dados convencionais não suportam efetivamente dados multimídia. Técnicas de indexação, recuperação, e reconhecimento de informações multimídia são necessários.

## Capítulo 2 Dados Multimídia

No capítulo I foi apresentada a definição de sistemas multimídia como sendo aqueles sistemas capazes de manipularem ao menos um tipo de mídia estáticas (texto, imagens, etc.) e um tipo de mídia contínuas (vídeo, áudio, etc.) na forma digital. A diferença entre um sistema multimídia e um sistema computacional convencional está na diferença de características e requisitos dos dados multimídia<sup>1</sup> dos dados alfanuméricos tradicionais. É necessário então conhecer estas características e requisitos para entender os problemas associados a computação e a comunicação multimídia. Este capítulo discute estas características e requisitos de sistemas de áudio, vídeo e imagens digitais.

### 2.1 Representação Digital de Informações Multimídia

Nas definições de sistemas multimídia apresentadas nesta apostila é ressaltado que todas as mídias deveriam estar na forma digital. Esta seção apresenta como as mídias de apresentação são capturadas do mundo real a partir de sinais analógicos e como estes sinais são transformados numa forma digital.

#### 2.1.1 Sinais Analógicos para representar informações

Esta seção apresenta uma introdução às formas de onda e sinais analógicos e seus uso na descrição de informações detectadas pelos sentidos humanos.

##### **Informações percebidas e variáveis físicas**

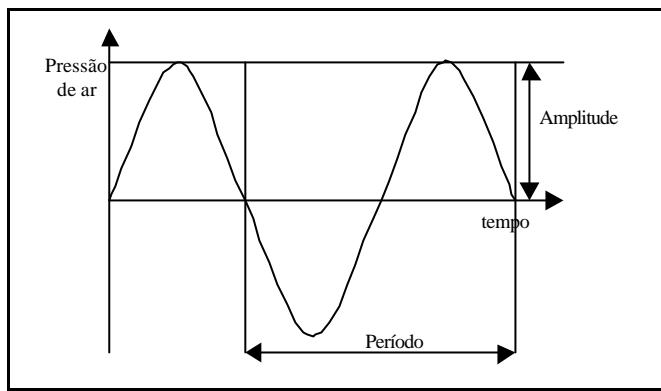
Informações que os sentidos humanos podem detectar podem ser descritas como uma ou várias variáveis físicas cujos valores podem ser funções do tempo e do espaço. Note que por informações é entendido aqui como a representação física em termos de estímulos para um sentido humano, e não o conteúdo semântico, que é o significado deste estímulo.

##### **Descrevendo sons com formas de onda**

Um som, que atravessa o ar, é uma onda de ar comprimido ou expandido cuja pressão altera no tempo e espaço. Na posição de um locutor ou de um detector, os sons podem ser descritos por valores de pressão que variam apenas com no tempo (valores dependentes do tempo). O padrão de oscilação, como mostrado na Figura 3, é chamado de forma de onda (*waveform*). A forma de onda é caracterizado por um **período** e **amplitude**. O período é o tempo necessário para a realização de um ciclo; intervalo de tempo que, num fenômeno periódico, separa a passagem do sistema por dois estados idênticos. A **freqüência** é definida como o inverso do período e representa o número de períodos em um segundo. A freqüência é normalmente medida em Hz (Hertz) ou ciclos por segundo (cps). A amplitude do som define um som leve ou pesado.

Como a onda de som ocorre naturalmente, ela nunca é perfeitamente suave ou uniformemente periódica como a forma de onda da Figura 3.

<sup>1</sup> Neste capítulo, dados multimídia refere-se as representações digitais dos múltiplos tipos de mídias.



**Figura 3.** Forma de onda

### **Descrevendo imagens monocromáticas com variáveis físicas**

As imagens refletem radiações eletromagnéticas (luz) incidentes que estimulam os olhos do observador. A intensidade de luz é uma função da posição espacial do ponto refletido sob a imagem. Portanto, a imagem pode ser descrita pelo valor da intensidade de luz que é uma função de duas coordenadas espaciais. Se a cena observada não foi plana, uma terceira coordenada espacial é necessária.

### **Descrevendo imagens coloridas com formas de onda**

Se a imagem não é monocromática, ela reflete diferentes comprimentos de onda. Assim uma função simples não é suficiente para descrever imagens coloridas, mas um espectro completo de comprimento de onda refletida, cada um com sua própria intensidade, é necessário. Assim as imagens teriam que ser descritas pela conjunção de várias funções bidimensionais. Felizmente, o sistema visual humano tem certas propriedades que simplificam a descrição de imagens coloridas.

A Luz que consiste em uma distribuição espectral de intensidade estimula o sistema visual e cria uma resposta. A resposta nos olhos depende da sensibilidade do sistema visual aos comprimentos de onda. Testes realizados mostram que diferentes distribuições espectrais da luz podem dar a mesma resposta visual. Em outras palavras, é possível criar sensações de cores idênticas com diferentes combinações de comprimentos de onda (isto é diferentes combinações de cores).

A teoria da cor foi desenvolvida por Thomas Young no início de 1802 e afirma que qualquer sensação de cor pode ser reproduzida pela mistura em proporções apropriadas de três luzes coloridas monocromáticas primárias. Cores primárias são independentes no sentido que uma cor primária não pode ser obtida misturando outras duas cores primárias. A *Commission Internationale de l'Eclairage* (CIE) recomendou o uso de uma tripla particular de luz monocromática. Cada fonte de luz é definida pelo seu comprimento de onda ( $\lambda_1 = 70 \text{ nm}$ , vermelho;  $\lambda_2 = 546.1 \text{ nm}$ , verde;  $\lambda_3 = 435.8 \text{ nm}$ , azul). Em vez de ser descrita por uma infinidade de funções bidimensionais, qualquer imagem colorida plana pode ser representada por um conjunto de três funções bidimensionais.

#### **2.1.2 Porque Digitalizar?**

A utilização de informações na forma digital trás as várias vantagens.

#### **Universalidade de representação**

Sistemas computacionais manipulam apenas dados digitais. Quando áudio, imagens, vídeos estão na forma digital, eles podem ser facilmente armazenados e manipulados (processados, transmitidos e apresentados) pelos sistemas computacionais tal qual outros dados. Desta forma, como todas as mídias de apresentação (textos, imagens, som, etc.) são codificadas numa única forma, elas podem ser manipuladas de uma mesma forma e pelo mesmo tipo de equipamento. Além disso, informações multimídia

codificadas são facilmente integradas com outros tipos de dados e são de fácil interação com mídia digitais usando sistemas computacionais.

### **Processamento**

Informações multimídia digitais são processadas, analisadas, modificadas, alteradas, ou complementadas por programas de computador tal qual outros dados. A seguir são apresentados alguns exemplos de processamentos possíveis graças a representação digital de informações multimídia [Fluckiger, 95]:

- reconhecimento de conteúdos semânticos (voz, escrita a mão, formas e padrões);
- estruturas de dados, ligações usando apontadores entre elementos de informações podem ser criados para rápida obtenção de informações;
- editores poderosos com funções *cut-and-paste* para criar monomídia (p.e. som apenas) ou documentos multimídia são possíveis;
- qualidade da informação pode ser aumentada pela remoção de ruídos ou erros como a digitalização de velhos discos de vinil para criar CD's de alta qualidade;
- informações sintetizadas e vídeos podem ser mixadas

### **Qualidade**

Sistemas digitais são mais confiáveis. Sinais digitais são mais tolerantes a ruídos e interferências que os analógicos. Na forma analógica, o valor do sinal é alterado se há ruídos ou interferências. Este erro é acumulativo de um estágio para outro do sistema. Na forma digital, há apenas dois níveis de sinal: alto (1) ou baixo (0). Se o erro causado pela interferência ou ruído é abaixo de um certo limiar, o sinal pode ser reconhecido corretamente. Além disso, em cada estado do processamento digital ou transmissão, o sinal digital são reconstruídos, assim erros não são acumulativos.

### **Segurança**

Se segurança na comunicação é necessária, a representação digital da informação facilita a criptografia.

### **Armazenamento**

A utilização unicamente de mídias digitais permite a existência de um dispositivo único de armazenamento de dados para todas as mídias, sendo que diferenças podem estar ligadas a requisitos de tamanho. Imagens e vídeos necessitam de uma maior capacidade de armazenamento que textos ou gráficos. Som necessita de uma capacidade de armazenamento um pouco menor que imagens.

Dispositivos digitais apropriados podem ser necessários, tal como CD-ROMs (*Compact Disk-Read Only Memories*).

### **Transmissão**

Para a transmissão de informações multimídia qualquer sistema de comunicação de dados podem ser (potencialmente) utilizado para a transmissão de informação multimídia. Uma única rede de comunicação suportando a transmissão digital das informações multimídia é possível (Rede Digital de Serviços Integrados). A este nível existem dificuldades causados pelos requisitos de certas aplicações, em particular aquelas que necessitam o respeito da fidelidade de dependências temporais dos sinais digitais.

A vantagem da transmissão digital em relação a transmissão de sinais analógicos é que ela é menos sensíveis a ruídos, a detecção de erros, recobrimentos e a criptografia são facilitadas.

#### **2.1.3 Digitalização, Amostragem e Quantificação**

Esta seção apresenta a forma de digitalização dos vários tipos de mídias de apresentação. Digitalização aqui é o processo envolvido na transformação de sinais analógicos em sinais digitais:

- **Sinal analógico** é uma medida física que varia continuamente com o tempo e/ou espaço. Eles são descritos por uma função dependente apenas do tempo ( $s=f(t)$ ), dependente apenas do espaço ( $s=f(x,y,z)$ ), ou dependente do tempo e do espaço ( $s=f(x,y,z,t)$ ). Sinais analógicos são produzidos por sensores que detectam fenômenos físicos (que simulam os sentidos humanos) e os transformam em uma medida que toma a forma de uma corrente ou tensão elétrica. A precisão é ditada pelas características dos sensores.
- **Sinais digitais** são seqüências de valores dependentes do tempo ou do espaço codificados no formato binário.

Para a conversão de sinais analógico em digital é necessário a realização de três passos: amostragem, quantificação e codificação. A figura 1 ilustra o processo de digitalização de um sinal analógico no domínio do tempo.

### **Amostragem**

Nesta etapa um conjunto discreto de valores analógicos é amostrado em intervalos temporais (p.e., para sons) ou espaciais (p.e., para imagens) de periodicidade constante, como apresentado na figura 1a. A freqüência de relógio é chamado de taxa de amostragem ou freqüência de amostragem. O valor amostrado é mantido constante até o próximo intervalo. Isto é realizado através de circuitos *sampling and hold*. Cada uma das amostras é analógica em amplitude: ele tem qualquer valor em um domínio contínuo. Mas isto é discreto no tempo: dentro de cada intervalo, a amostra tem apenas um valor.

Segundo o teorema de Nyquist: se um sinal analógico contém componentes de freqüência até  $f$  Hz, a taxa de amostragem deve ser ao menos  $2f$  Hz. Na prática, esta freqüência é um pouco maior que  $2f$  Hz. Por exemplo, a taxa de amostragem de CD-audio é de 44,1 kHz, e dos tapes de áudio digital (DAT) é de 48kHz para cobrir uma faixa audível de freqüência de 20 kHz. Outro exemplo, os componentes principais de freqüência da voz humana estão dentro de 3,1 kHz, com isto os sistemas de telefonia analógicos limitam o sinal transmitido a 3.1 kHz; é comum usar uma freqüência de amostragem de 8 kHz para converter este sinal em digital.

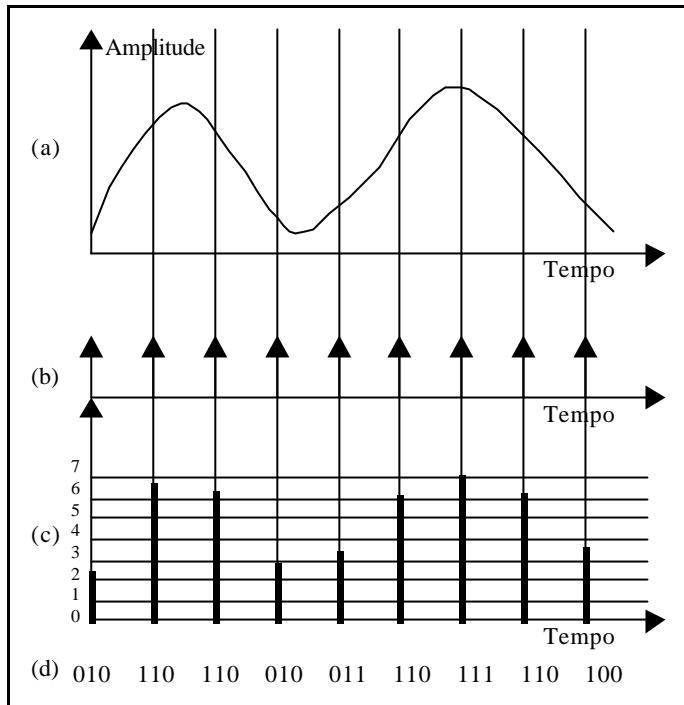
### **Quantificação**

O processo de converter valores de amostras contínuas em valores discretos é chamado de quantificação. Neste processo nós dividimos o domínio do sinal em um número fixo de intervalos. Cada intervalo tem o mesmo tamanho e recebe um número. Na figura 1c estes intervalos são numerados de 0 a 7. A cada amostra dentro de um intervalo é atribuído o valor do intervalo. O tamanho deste intervalo de quantificação é chamado de passo de quantificação. A técnica que utiliza o mesmo passo de quantificação é chamada modulação PCM (*Pulse Coded Modulation*). Algumas vezes, nem todos os valores amostrados são retidos após a quantificação. No caso ilustrado pela figura 1c, todos os valores amostrados foram retidos.

### **Codificação**

A codificação consistem em associar um conjunto de dígitos binários, chamado de *code-word*, a cada valor quantificado. No caso da figura 1d, oito níveis de quantificação são usados. Estes níveis podem ser codificados usando 3 bits, assim cada amostra é representada por 3 bits.

Em algumas aplicações de telefonia, a digitalização da voz humana utiliza 16 bits por amostra, que então leva a  $2^{16}$  ou 65.536 passos de quantificação. Em outras aplicações de compressão de voz, algumas vezes, apenas 8 quantificações por bits são necessários, produzindo apenas 256 passos de quantificação.



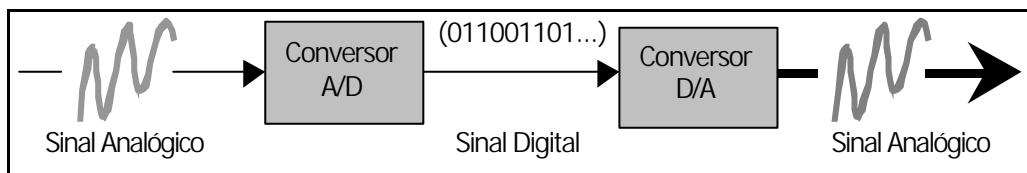
**Figura 4.** Conversão A/D [Lu, 96]: (a) sinal analógico; (b) pulsos de amostragem; (c) valores amostrados e intervalos de quantificação; (d) seqüência digital

### Taxa de bits

Taxa de bits é definida como o produto entre taxa de amostragem e o número de bits usados no processo de quantificação. Por exemplo, supondo uma freqüência de 8k Hz e 8 bits por amostra, a taxa de bits necessária à telefonia é igual a  $8000 \times 8 = 64$  kbps.

### Conversão analógica/digital e digital/analógica

Em sistemas multimídia todas as informações multimídia são representadas internamente no formato digital. Mas humanos reagem a estímulos sensoriais físicos, assim a conversão digital-para-analógico (ou conversão D/A) é necessária na apresentação de certas informações (figura 2).



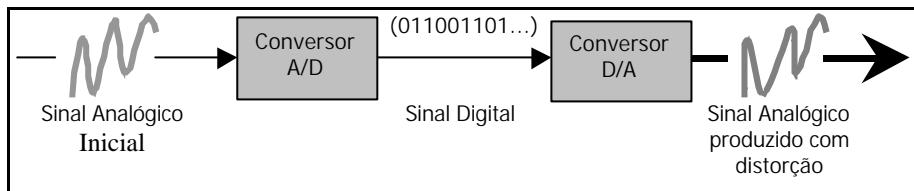
**Figura 5.** Conversão analógico/digital e digital/analógica

### 2.1.4 Problemas da Representação digital

Apesar de aportar várias vantagens, a digitalização de informações multimídia apresenta algumas deficiências que são apresentadas nesta seção.

#### Distorção

O maior problema da utilização de informações multimídia na forma digital é a distorção de codificação: amostragem, quantificação e codificação dos valores introduz distorção ao sinal analógico restituído. O sinal gerado após a conversão D/A não é idêntico ao original (como ilustrado na figura 3), assim a informação apresentada ao usuário não é idêntica aquela capturada do mundo real.



**Figura 6.** Conversão analógico/digital e digital/analógica

Aumentando a taxa de amostragem e o número de bits usado para codificação reduz esta distorção. Segundo o teorema de Nyquist: se um sinal analógico contém componentes de freqüência até  $f$  Hz, a taxa de amostragem deve ser ao menos  $2f$  Hz. Mas há uma clara limitação tecnológica neste ponto: capacidades de armazenamento não é infinita, e os sistemas de transmissão tem largura de banda<sup>2</sup> limitadas.

Na maior parte dos sistemas multimídia, os usuários finais das informações multimídia são os humanos. Como nem todos os componentes de freqüências são percebidos pelos humanos, a solução para reduzir a distorção é escolher um balanço apropriado entre a precisão da digitalização e a distorção percebida pelo usuário.

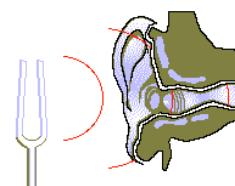
### Necessidade de grandes capacidades de armazenamento

Outro problema gerado pela digitalização de informações multimídia é a necessidade de meios de armazenamento digital com grandes capacidades principalmente para o armazenamento de vídeos, imagens e áudios. Por exemplo, oito minutos de som estereofônico de qualidade CD são suficientes para completar 80 megabytes do disco duro de um PC padrão. Para reduzir este problema, faz-se uso de algoritmos de compressão que reduzem este requisito.

## 2.2 Áudio

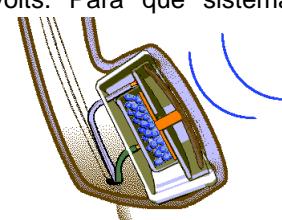
Áudio é causado pelo distúrbio da pressão de ar que alcança o tímpano. Quando a freqüência do distúrbio de ar está na faixa de 20 Hz a 20.000 Hz ele é audível. A maioria dos sistemas multimídia trabalham com esta faixa de freqüência. Outro parâmetro usado para a medição do som é a amplitude (medido em decibéis - dB), variação que causa o som leve ou pesado. Por exemplo, o limiar da dor é de 100 a 120 dB.

A onda sonora é uma onda contínua no tempo e amplitude. A onda apresentada na Figura 3 pode ser um exemplo de onda sonora.



### 2.2.1 Representação digital de áudio

A forma de onda de áudio é convertida em um sinal elétrico contínuo (analógico) por um microfone. Este sinal elétrico é medido normalmente em volts. Para que sistemas computacionais processem e comuniquem sinais de áudio, o sinal elétrico deve ser convertido em um sinal digital. Este processo de conversão de um sinal analógico em digital foi apresentado na seção 2.1.3. O mecanismo que converte o sinal de áudio digital em analógico é chamado de Conversor Analógico para Digital (CAD).



Áudio digital necessita ser amostrado continuamente em uma taxa fixa. Cada amostra é representada por um número fixo de bits. A tabela abaixo mostra a taxa de amostragem e o número de bits usados para cada amostra para várias aplicações de áudio. Relembre, quanto maior a taxa de amostragem e maior o número de bits por amostragem, maior é a qualidade do áudio restituído, mas com isso

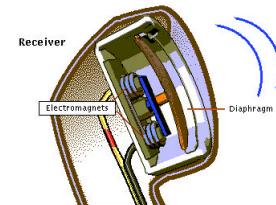
<sup>2</sup> Em transmissão, a faixa de freqüências que o sistema pode transmitir sem excessivas atenuações. Em redes de dados, algumas vezes este termo é usado para referenciar a capacidade total de uma rede expressa em bits por segundo. Em tais casos, o termo apropriado é taxa de bits.

maior é a taxa de bits. Note na tabela que para áudio estéreo, tal como CD-Audio, dois canais são necessários.

Aplicações	Nº de canais	Taxa de amostragem	Bits por amostragem	Taxa de bits
CD-Audio	2	44.1 kHz	16	1,41 Mbps
DAT	2	48 kHz	16	1,53 Mbps
Telefone Digital	1	8 kHz	8	64 Kbps
Rádio digital, long play DAT	2	32 KHz	16	1,02 Mbps

Para a apresentação do áudio digitalizado é necessário realizar a transformação de uma representação artificial do som em uma forma de onda física audível pelo ouvido humano. Para isto são utilizados Conversores Digital para Analógico (CDA).

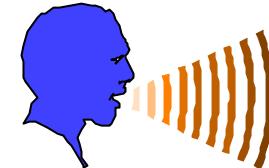
Normalmente os conversores CAD e CDA são implementados em uma única placa. Um exemplo de placa de audio é Creative Sound Blaster AWE64, possibilitando até 16 bits por amostras, produzindo áudio qualidade CD.



## 2.2.2 Características da fala

A fala é a forma dominante de comunicação entre humanos. Ela suporta linguagens faladas, portanto ela tem um conteúdo semântico. Esta característica da fala tem duas consequências imediatas quando consideramos seu uso em sistemas computacionais:

- Parte do conteúdo semântico pode ser reconhecido pelo computador. Os componentes individuais da fala: fonemas, e grupos de fonemas, que são as palavras, podem ser reconhecidas. Isto é chamado de reconhecimento de voz.
- A translação pelo computador de uma descrição codificada de uma mensagem em uma voz é possível. Isto é chamado de síntese de voz. Um tipo particular de síntese é a conversão texto para voz.



### Digitalização da voz

A faixa de freqüência da voz é de 50 Hz a 10 kHz. Apesar disto a largura de banda necessária à transmissão ou registro é menor que 10 kHz. Isto pois humanos não falam todo o tempo: seqüências de fonemas seguem períodos de silêncio.

## 2.2.3 Áudio não Voz

Sons não sejam voz e que os humanos podem ouvir em geral não têm semânticas complexas. Assim há poucas razões para trabalhos de reconhecimento computadorizado destes tipos de sons. Há algumas aplicações militares de reconhecimento de sons não-fala.

## 2.2.4 Representação simbólica da música: o padrão MIDI

Como visto anteriormente, qualquer som pode ser representado como um sinal de som digitalizado, que é uma seqüência de amostras, cada uma codificada por dígitos binários. Esta seqüência pode ser descompactada como nos discos compactos de áudio ou compactados. Uma característica deste modo é que ele não preserva a descrição semântica do som. A menos que seja utilizado técnicas de reconhecimento complexas, o computador não sabe se a seqüência de bits representa, por exemplo, uma fala ou música, e se musicas que notas são usadas e por que instrumentos.

Algumas representações de som preservam a semântica da informação. No caso da codificação da fala, pode-se usar um texto e atributos como voz masculina ou feminina, sotaque e taxa de palavras. A música também pode ser descrita de uma maneira simbólica usando técnicas similares as pautas musicais. O formato mais utilizado para isto é aquele definido no padrão MIDI (*Musical Instrument Digital Interface*). Este padrão define como codificar todos os elementos musicais, tal como seqüências de notas,

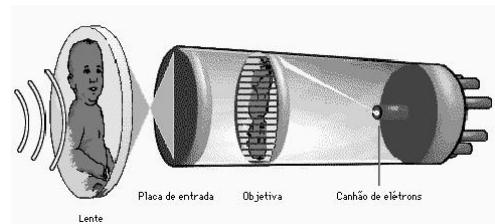
condições temporais, e o “instrumento” que deve executar cada nota (são 127 instrumentos e outros sons como aqueles produzidos por helicóptero, telefone, aplausos, etc.).

Arquivos MIDI são muito mais compactos que amostragens digitalizadas: um arquivo MIDI pode ser 1000 vezes menor que um arquivo CD áudio. Além disso, a representação MIDI é revisável (modificáveis). As desvantagens são a necessidade de um processamento extra de informação, e imprecisão dos instrumentos de som (variam com o dispositivo usado para a apresentação).

## 2.3 Vídeos e Imagens Analógicos

### 2.3.1 Captura e reprodução de imagens e vídeos analógicos

As imagens são capturadas usando câmeras da seguinte maneira: as lentes da câmera focam uma imagem de uma cena em uma superfície foto-sensível de sensores CCD (*Charge-Coupled Device*); o brilho de cada ponto é convertido em uma carga elétrica por uma camada foto-sensível, estas cargas são proporcionais ao brilho nos pontos; a superfície foto-sensível é rastreada por um feixe de elétrons para capturar as cargas elétricas, devendo ser feito rapidamente antes que a cena mude. Desta maneira a imagem ou cena é convertida em um sinal elétrico contínuo.



Nesta seção, por simplificação assume-se a captura e reprodução de vídeos monocromáticos, onde apenas um sinal de luminância é produzido (apenas a luminosidade é capturada, temos a imagem em tons de cinza). Neste caso são usadas **câmeras de Luminância**, que captam a imagem em tons de cinza, e gera um sinal só com a luminância da imagem. A imagem é gerada por um CCD monocromático que capta o tom de cinza que incide em cada célula do circuito. Este tipo de câmera é utilizada em geral para aplicações em visão computacional e nos casos onde a informação sobre a luminosidade da imagem é suficiente.

O dispositivo de apresentação de imagens mais utilizado é o tubo de raios catódicos (CRT). Eles são usados nos aparelhos de TV e monitores de computadores. Há uma camada de fósforo fluorescente no interior do CRT. Esta camada é rastreada por um feixe de elétrons na mesma forma do processo de captura na câmera. Quando o feixe toca o fósforo ele emite luz durante um curto instante. O brilho da luz depende da força do feixe. Quando quadros repetem-se suficientemente rápidos a persistência da visão resulta na reprodução de um vídeo. Em implementação prática, o sinal elétrico enviado da câmera para o dispositivo de apresentação deve conter informações adicionais para assegurar que o rastreamento esteja sincronizado com o rastreamento do sensor na câmera. Esta informação é chamada *sync information*.

### 2.3.2 Vídeos e Imagens Coloridos

Nesta seção é discutido a captura e a reprodução de imagens e vídeos coloridos.

Todos os sistemas de TV a cores são baseados na teoria *Tristimulus* de reprodução da cor. Esta teoria afirma que qualquer cor pode ser reproduzida com a mistura das três cores primárias: vermelho, verde e azul. Para capturar imagens coloridas, uma câmera divide a luz nos seus componentes vermelho, verde e azul. Estas três componentes de cor são focalizados em sensores de vermelho, verde e azul, que convertem estes três componentes em sinais elétricos separados.

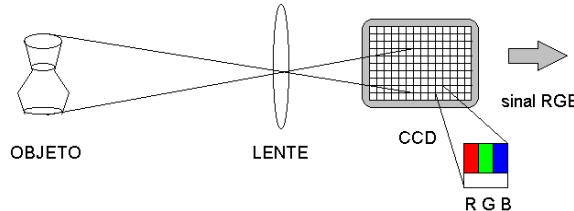
Em um monitor colorido, há 3 tipos de fósforos fluorescentes que emitem luzes vermelha, verde e azul quando tocadas por 3 feixes de elétrons. Estes fósforos são arranjados de tal forma que cada posição do vídeo tem 3 tipos de fósforo. A mistura da luz emitida destes 3 fósforos produz um ponto de cor.

A visão acima descreve o sinal RGB. Na realidade, o sinal analógico pode ser gerado da seguinte maneira [França Neto, 98]:

- Sinal RGB (red, green, blue) - O sinal é separado pelas cores básicas, com isso é possível ter uma imagem mais pura. Ele é utilizado em câmeras e gravadores profissionais, imagens geradas por computador, etc.
- Sinal de vídeo composto colorido - os sinais das cores (RGB) são codificados em um único sinal seguindo um determinado padrão (NTSC, PAL-M, SECAM, etc) ;
- Sinal de luminância e crominância ou Y/C (S-video) - o sinal é composto por duas partes, a luminância e a crominância; como isso a imagem tem uma melhor qualidade do que no vídeo composto. Muito usado por vídeos SVHS, laser disc, DVD e outros aparelhos que geram imagens de boa qualidade (acima de 400 linhas);

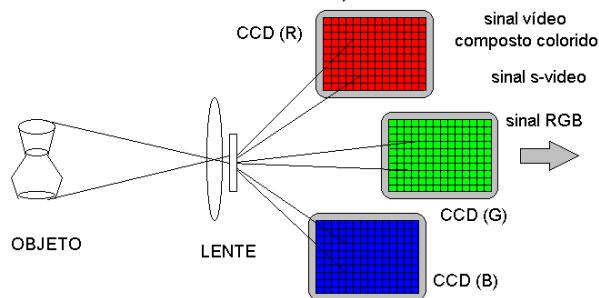
Existem vários tipos de câmeras que geram imagens coloridas, entre elas temos:

- Câmera de crominância (1 passo - 1 CCD) - Capta a imagem em cores, e gera um sinal de vídeo composto colorido, em apenas uma passagem. A imagem, em geral, não é profissional, pois é usado um único CCD com filtros RGB em cada célula, como pode ser visto na Figura 7. Este tipo de câmera é utilizado em aplicações multimídia ou em casos onde não é necessário uma imagem com muita qualidade. Esta é uma câmera do tipo doméstica (VHS, 8mm, VHS-C, etc), desta forma tem um custo baixo.



**Figura 7.** Câmera de crominância (1 passo - 1 CCD)

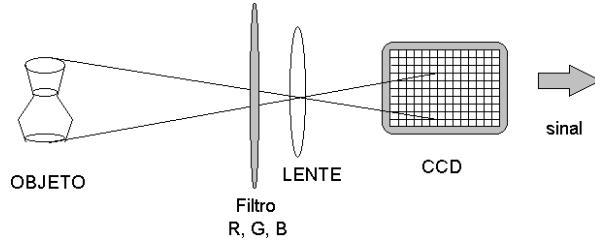
- Câmera de crominância (1 passo - 3 CCD) - Capta a imagem em cores, e pode gerar sinal de vídeo composto colorido, S-video ou sinal RGB. Tem uma qualidade de imagem profissional, pois são usados 3 CCDs com filtros separados R, G e B em cada um, como pode ser visto na Figura 8. Por ter 3 CCDs independentes, cada um pode ter uma resolução maior, garantindo uma melhor resolução da imagem. É utilizada em aplicações profissionais, onde é necessário uma imagem com boa qualidade. Esta é uma câmera do tipo usado em produtoras e emissoras de TV (U-matic, BetaCAM, SVHS, Hi8, etc), desta forma tem um custo elevado.



**Figura 8.** Câmera de crominância (1 passo - 3 CCD)

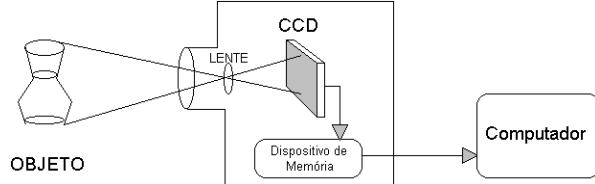
- Câmera de crominância (3 passos - 1 CCD) - Capta a imagem em cores, porém este processo é feito em 3 passos. É utilizado um único CCD para captar a imagem, sendo que para gerar uma imagem colorida é colocado um filtro externo para cada componente R, G e B (Figura 9). A digitalização da imagem então é feita em 3 passos, ou seja, para cada filtro é feito uma digitalização. Assim temos a informação das intensidades de cada componente RGB. Com esta informação é composta uma imagem colorida, pois para cada ponto temos a contribuição R, G e B. Este processo tem uma desvantagem pelo fato de que as imagens devem ser estáticas, pois é preciso trocar os filtros e fazer nova captação para os outros

filtros. Tem uma boa qualidade de imagem, pois este CCD pode ter uma boa resolução, proporcionando uma melhor resolução da imagem. É utilizada em geral para aquisição de imagens de telescópio, onde é necessário uma imagem com alta definição e as imagens são relativamente estáticas. Esta é uma câmera que pode ter um custo baixo, no caso de CCDs de pouca qualidade (baixa resolução), ou de alto custo se o CCD tiver alta resolução.



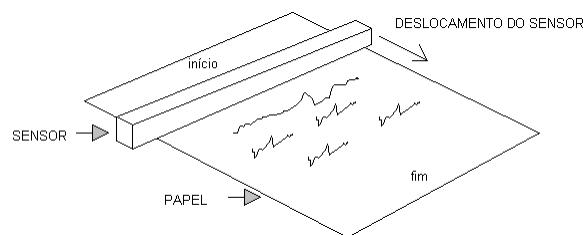
**Figura 9.** Câmera de crominância (3 passos - 1 CCD)

- A câmera fotográfica digital é um dispositivo de funcionamento semelhante a uma câmera fotográfica tradicional, porém a imagem não é armazenada em um filme e sim de forma digital em memória. Esta imagem é digitalizada através de um CCD, e armazenada de forma compactada ou não em um dispositivo de memória. A qualidade da imagem depende da qualidade e resolução do CCD e da compressão utilizada para armazenar a imagem digitalizada. A resolução, atualmente, varia entre 320x240 até algo em torno de 1600x1200 pontos por imagem. A imagem pode ser armazenada em vários tipos de memória como memórias não voláteis, cartões de memória, disquetes magnéticos, etc. As imagens podem ser transferidas para um computador por cabos ou leitores dos dispositivos de memória, e então são processadas, como é visto na Figura 10. O custo deste dispositivo pode ser baixo no caso de câmeras domésticas com poucos recursos e baixa resolução, ou muito alto quando a câmera possui recursos profissionais e alta resolução.



**Figura 10.** Câmera fotográfica digital

O scanner digitaliza a partir de imagens em papel. A imagem é colocada sobre uma superfície transparente, em geral plana ou cilíndrica, que se move numa direção ortogonal à um elemento de digitalização de linha (Figura 11). Este elemento se compõe de uma fonte de luz e de um sensor que mede a luz refletida linha por linha, em sincronismo com o deslocamento da imagem, ou do sensor. A resolução deste dispositivo está situada entre 50dpi a 4000dpi (pontos por polegada). Um dispositivo semelhante, é o film scanner, que obtém uma imagem digital a partir de imagens em transparências, utilizando o laser para maior resolução. Este dispositivo também atinge resoluções superiores a 2000 dpi.



**Figura 11.** Esquema de funcionamento do scanner

Existem scanners de vários modelos, desde simples scanners de mão, até potentes scanners utilizados em grandes gráficas para captar imagens com um grau de detalhe

muito grande. O scanner de mão, é geralmente usado em aplicações domésticas onde não se tem a necessidade de muita qualidade, é composto por um sensor que é arrastado sobre a imagem. Por outro lado os scanners de mesa proporcionam uma melhor qualidade na aquisição das imagens.



**Figura 12.** Exemplos de scanners

Vários fatores influenciam na qualidade dos scanners, entre eles temos:

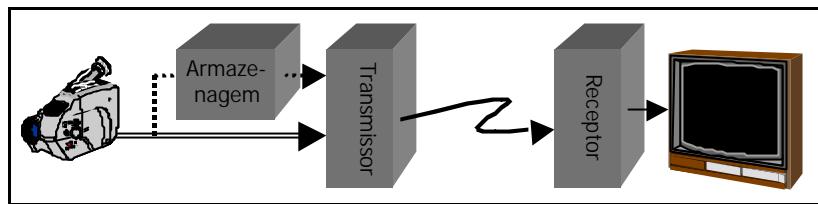
- **Resolução óptica:** É a resolução via hardware que o scanner pode atingir. Outras resoluções acima desta podem ser obtidas com técnicas de interpolação, mas com perda de qualidade, ou seja, a máxima resolução obtida com toda a qualidade é a sua resolução óptica. A escolha da resolução de um scanner depende diretamente da aplicação para qual este vai ser usado, quanto mais detalhes forem necessários na captação das imagens, mais resolução óptica este scanner deve ter. Geralmente esta resolução começa de 300dpi, o que atende a uma boa parte das aplicações, até algo em entre 1200dpi a 2000dpi, que são dispositivos mais sofisticados para aplicações mais específicas.
- **Quantidade de bits para representar cada componente de cor:** este número é diretamente proporcional à quantidade de intensidade de cada componente. Geralmente temos cada componente representada por 8 bits, o que dá 256 níveis de intensidade para cada componente e um total de 24 bits o que corresponde a 16.777.216 cores. Alguns scanners já trabalham com 10 bits, ou seja 30 bits no total, e podem representar mais de um trilhão de cores.
- **Tamanho da área de leitura:** É a área máxima que pode ser usada para digitalizar uma imagem ou documento. A maioria dos modelos trabalham com os formatos A4 ou CARTA.
- **Velocidade de captação da imagem:** É o tempo para realizar a digitalização de uma imagem. Esta velocidade pode variar de acordo com o mecanismo de captação e da forma como os dados são transferidos para o computador. Geralmente é especificado alguns padrões para aferir esta velocidade, como uma imagem de tamanho padrão tipo A4, resoluções e número de cores fixas. Com isso é possível fazer uma análise do desempenho de um determinado scanner em relação a outro, e de acordo com a aplicação desejada, escolher qual o melhor. Outro fator que deve ser levado em conta é a forma como os dados são transferidos para o computador. Geralmente é feito através de interface paralela ou interface padrão SCSI (Small Computer System Interface). No caso da interface paralela, a velocidade é bastante limitada, porém é mais fácil a instalação, pois a maioria dos computadores já vem com uma interface paralela, bastando conectar o cabo. No caso da interface SCSI, a velocidade pode ser bem maior dependendo da qualidade dos dispositivos, este tipo de interface é mais complicada de instalar, pois caso o computador não possua uma porta SCSI, sendo necessário a instalação de uma placa SCSI para fazer a ligação do scanner ao computador. A velocidade do sistema (cpu, memória, disco rígido, sistema operacional, etc) pode ter importância, dependendo da velocidade do scanner, ou seja no caso de scanners mais rápidos o sistema pode tornar o processo de aquisição lento. A velocidade do scanner é geralmente proporcional ao seu custo, com isso deve ser levado em conta na hora de escolher o scanner ideal para a aplicação.
- **Qualidade do sensor:** Esta qualidade diz respeito a capacidade de representar as cores de forma mais fiel possível. Quanto melhor o sensor, mais semelhante a imagem gerada será da imagem real.

## 2.4 Representação digital de imagens e vídeos

A seção anterior discutiu muitos conceitos e nomenclaturas de vídeos analógicos. Como nosso interesse é manipulação de imagens e vídeos digitais em sistemas multimídia, esta seção discute a representação digital de imagens e vídeos.

### 2.4.1 Um Sistema Genérico de Vídeo Digital

Em um sistema de captura e apresentação de vídeo na forma analógica, uma câmera converte as cenas em um sinal de vídeo analógico. Este sinal é transmitido (ou armazenado para posterior transmissão) para um receptor para sua apresentação, como mostra a figura 4.



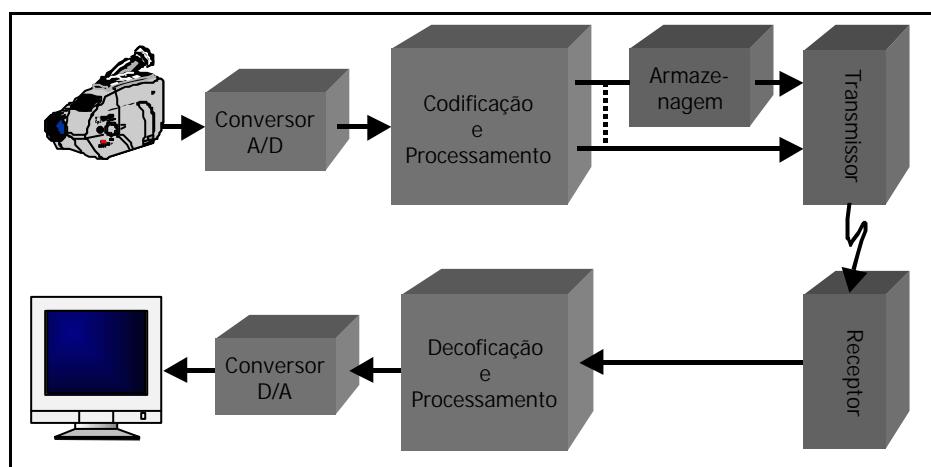
**Figura 13.** Um sistema de vídeo analógico

Em um sistema de vídeo digital, o sinal analógico do vídeo é convertido em uma forma digital (figura 5) imediatamente após a câmera e uma conversão digital/analógico é necessária imediatamente antes da apresentação do vídeo. Note que o vídeo está no formato analógico quando ele é capturado e reproduzido, o formato digital pode ser considerado um formato intermediário para facilitar o processamento e transmissão.

### Comparação de sistemas de vídeo analógico e digitais

Um sistema de vídeo analógico, tal como o sistema de televisão é muito difícil separar as imagens do vídeo, isto pois não há o armazenamento de imagens. Se nós desejarmos apresentar uma imagem de um vídeo durante por exemplo 5 minutos, esta imagem deve ser capturada, transmitida e apresentada repetidamente por 5 minutos. Ao contrário, em sistemas digitais há o armazenamento de imagens, se uma imagem deve ser apresentada durante um certo tempo, temos que capturar e enviar a imagem uma única vez.

Em sistemas analógicos, todos os componentes, câmera, transmissor, receptor e monitor, devem operar sincronamente. No caso de sistemas digitais estes componentes são independentes: a câmera registra uma imagem/cena, mas o transmissor pode transmitir seletivamente as imagens baseada no seu conteúdo, no poder de computação ou na largura de banda disponíveis; o monitor mostra aquilo que é disponível. A imagem mostrada não vai piscar se faltar quadros subsequentes pois a tela será atualizada constantemente com o conteúdo disponível.

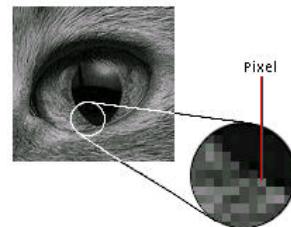


**Figura 14.** Um sistema de vídeo digital

Um dos problemas da utilização de sistemas de vídeo digitais está associado aos requisitos de largura de banda de transmissão. Assumindo a largura de banda de um vídeo analógico é  $B$  MHz, então a largura de banda mínima necessária para o canal de transmissão é  $B$  MHz. No caso de sistemas digitais, onde são transmitidos dados binários sobre um canal de  $B$  MHz, de acordo com o Teorema de Nyquist a taxa máxima de bits que o canal deve suportar é  $2B$  Mbits/s. De qualquer maneira, quando um vídeo é digitalizado no formato vermelho, verde e azul (visto mais adiante), temos uma taxa de bits de  $49B$  Mbits/s. Esta comparação sugere que a transmissão não compactada de vídeo digital requer um canal de transmissão com um significativo aumento de largura de banda comparado aos sistemas analógicos. Mas se a compressão de vídeo digital for utilizada numa taxa de compressão 24:1 ou maior, nós precisamos de um canal de transmissor de largura de banda mais baixa que os dos vídeos analógicos.

#### 2.4.2 Imagens Digitais

Imagens não são revisáveis porque seu formato não contém informações estruturais. Elas podem resultar de capturas do mundo real (via escaneamento de uma página impressa ou foto, câmeras digitais) ou elas podem ser sintetizadas pelo computador (via programas de *paint*, captura da tela, conversão de gráficos em imagens bitmap). Após digitalizadas, as imagens podem ser manipuladas com editores de imagens (por exemplo, Photoshop), que não produzem documentos que retêm a estrutura semântica.



##### *Formatos de Imagens*

Imagens no computador são representadas por *bitmaps*. Um bitmap é uma matriz bidimensional espacial de elementos de imagem chamados de pixels. Um pixel é o menor elemento de resolução da imagem, ele tem um valor numérico chamado de amplitude. O número de bits disponíveis para codificar um pixel é chamado de profundidade de amplitude (ou de pixel). Exemplos típicos de profundidade de pixel é 1 (para imagens preto&branco), 2, 4, 8, 12, 16 ou 24 bits. O valor numérico pode representar um ponto preto e branco, um nível de cinza, ou atributos de cor (3 valores) do elemento de imagem em imagens coloridas.

O número de linhas da matriz de pixels ( $m$ ) é chamado de resolução vertical da imagem, e o número de colunas ( $n$ ) é chamado de resolução horizontal. Denominamos resolução espacial, ou resolução geométrica, ao produto  $m \times n$  da resolução vertical pela resolução horizontal. A resolução espacial estabelece a freqüência de amostragem final da imagem. Dessa forma, quanto maior a resolução mais detalhe, isto é, altas freqüências, da imagem podem ser captadas na representação matricial. A resolução espacial dada em termos absolutos não fornece muita informação sobre a resolução real da imagem quando realizada em dispositivo físico. Isso ocorre porque ficamos na dependência do tamanho físico do pixel do dispositivo. Uma medida mais confiável de resolução é dada pela densidade de resolução da imagem que fornece o número de pixels por unidade linear de medida. Em geral se utiliza o número de pixels por polegada, ppi ("pixels per inch") também chamada de dpi ("dots per inch").

Formatos bitmap necessitam mais capacidade de armazenamento do que gráficos e textos. Como bitmaps ignoram a semântica, duas imagens de mesma dimensão (altura e largura) ocupam o mesmo espaço. Por exemplo, um quadrado ou uma foto digitalizada com dimensões idênticas ocupam o mesmo espaço. Os gráficos, como eles consideram a semântica, ocupam menos espaço.

#### 2.4.3 Imagens e Gráficos Animados

As imagens e os gráficos podem ser apresentados na tela do computador como uma sucessão de imagens/gráficos que podem criar a sensação de movimento.

## Quadro e Taxa de Quadro

Uma imagem ou gráfico individual de uma animação é chamado de quadro (ou *frame*). Para ser compreensível, os quadros que compõem a animação devem ser apresentados geralmente em uma taxa aproximadamente fixa. O número de quadros apresentados por segundo é definido como freqüência de quadros e é medido em termos de quadros por segundo (fps – frames per seconds).

A taxa de quadros de uma animação é determinada por 3 fatores:

- A taxa deve ser alta suficiente para produzir a sensação de movimento. Para isto, taxas maiores ou iguais a 25 fps devem ser utilizadas. A tabela abaixo resume as principais freqüências de quadro utilizadas atualmente.

Fps	Comentários
<10	Apresentação sucessiva de imagens
10 à 16	Impressão de movimento mas com sensação de arrancos
>16	Efeito do movimento começa
24	Cinema
30/25	Padrão de TV americana/européia
60	Padrão HDTV

- Maior a freqüência de quadros utilizada mais alto é a largura de banda necessária à transmissão. Isto pois maior a taxa, maior é o números de quadros que devem ser enviados. Portanto a rede utilizada pode ditar a freqüência de quadros a ser utilizada.
- Problema de freqüência de restauração (refreshing) de tela: a tela deve ser restaurada 50 vezes por segundo para evitar tremulações. Mas se a freqüência de quadros for 50 fps, a largura de banda necessária aumentará substancialmente. Para evitar problemas de tremulação utiliza-se vídeos entrelaçados, onde reduz-se pela metade o número de quadros requeridos por segundo, ou seja, 25 fps.

## Imagens Bitmap Animadas (Vídeo)

Na animação de imagens, cenas são registradas como um sucessão de quadros representados por imagens bitmap possivelmente compactadas. Estas imagens podem ser capturadas da vida real com câmeras ou criadas através do computador. A primeira técnica produz o que é chamado de **vídeo**.

Animação de imagens tem as mesmas características que as imagens: falta de uma descrição semântica e necessidade de uma grande capacidade de armazenamento.

## Gráficos Animados

O termo gráfico animado ou animação gráfica é utilizado para referenciar apresentação sucessiva de objetos visuais gerados pelo computador em uma taxa suficiente para dar a sensação de movimento e onde cada atualização é comutada de uma descrição abstrata em tempo de apresentação.

A principal vantagem das animações gráficas é que elas são mais compactas: elas são descritas por um conjunto de objetos com diretivas temporais (em outras palavras um programa a ser executado em tempo de apresentação). Outra vantagem é que animações gráficas são revisáveis. Existe uma desvantagem: é necessário um poder de processamento suficiente para apresentação.

## Vídeos Híbridos

Técnicas avançadas, incluindo reconhecimento de padrões, permitem formas híbridas combinando vídeos e animações gráficas. Tais aplicações são suportadas por programas avançados que necessitam de unidades de processamento poderosas.

Um exemplo é quando imagens bitmap individuais providas por um câmera de TV ao vivo ou videotape são analisados por programas de computadores e modificados de acordo com um critério predefinido. A apresentação de objetos reais ou pessoas podem ser modificados, ou em modo ao-vivo ou off-line.

## 2.4.4 Especificação da Cor

Imagens e vídeos coloridos são dois tipos de dados básicos em um sistema multimídia. Esta seção procura apresentar a forma de especificação da cor na representação de imagens e vídeos digitais.

### Propriedades da Cor

A luz visível é uma forma de radiação eletromecânica que tem um espectro de comprimento de onda variando aproximadamente de 400 nm a 780 nm. Uma luz de diferente comprimento de onda produz uma sensação de cor diferente (p.e. violeta de 380 a 450 nm, azul de 450 a 490 nm, verde de 490 a 560nm). As três propriedades físicas básicas da cor são: luminância (brilho), nuança (cor) e saturação (puridade).

### Sistema de especificação de cores

Para comunicar imagens e vídeos coloridos, a cor deve ser especificada usando certo método. A CIE, Comissão Internacional de Iluminação, é o órgão responsável pela padronização na área de fotometria e colorimetria. Existem vários padrões de cor estabelecidos pela CIE, entre eles o CIE-RGB e o CIE-XYZ.

## 2.4.5 Sistema RGB

No sistema RGB de representação de cor, uma cor é representada pela intensidade de três cores primárias (teoria Tristimulus): vermelho (Red), verde (Green) e azul (Blue), com cada valor variando de 0 a 255. Exemplos de cores familiares são apresentadas abaixo:

- Branco = 255,255,255
- Vermelho = 255,0,0
- Verde = 0,255,0
- Azul = 0,0,255
- Amarelo = 255,255,0
- Preto = 0,0,0

A representação de imagens coloridas pode ser feita através de cores por componente (*true color*), cores indexadas, ou cores fixas. Essa representação vai depender do propósito e dos dispositivos que vão ser usados para trabalhar com essas imagens.

### True Color

No True Color, cada pixel da imagem é representado por um vetor de 3 componentes de cores (RGB) com um certo número de bits para representar cada componente de cor (resolução de cor). Com isso, quanto maior for a resolução de cor mais qualidade teremos para representar as cores de cada pixel. Geralmente o número de bits para cada componente RGB é igual, ou seja quando temos um pixel sendo representado por 9 bits, usamos 3 bits para cada componente (3-3-3). Mas pode ser feito uma representação com diferentes valores para as componentes, por exemplo uma representação 8 bits/pixel, pode ser usado 3 para componentes R, 3 para G e 2 para B (3-3-2), tal representação em um byte é comumente usado e tira proveito do fato que a percepção humana da componente azul é menos sensível que as outras componentes.

O número de bits para representar cada componente fornece a quantidade de cores que podem ser representados por essa componente. Ou seja, se  $n$  é a resolução de cor então a quantidade de níveis possíveis é de  $2^n$  níveis. Por exemplo, uma imagem colorida representada por 12 bits/pixel, com 4 bits para cada componente RGB. Temos então:  $2^4=16$  níveis para cada componente de cor RGB, o que nos possibilita representar até 4.096 cores diferentes ( $16 \times 16 \times 16 = 4.096$ ) , o que é equivalente à  $2^{12}=4.096$ .

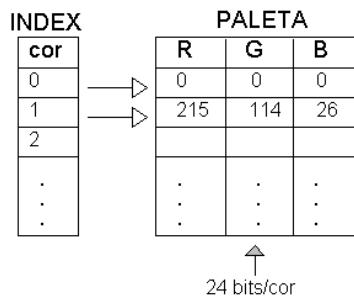
Temos alguns padrões de cores nesse formato que são:

Bits/pixel	Padrão	Componente de cor RGB	Máximo de cores
15 bits/pixel	High Color (15 bits)	5 bits/pixel, 32 níveis por componente	32.768 cores
16 bits/pixel	High Color (16 bits)	5/6 bits/pixel, 32/64 níveis por componente	65.536 cores
24 bits/pixel	True Color, (24 bits)	8 bits/pixel, 256 níveis por componente	16.777.216 cores

O padrão com 24 bits/pixel é o mais usado para representar com fidelidade as cores, pois o número de cores que podem ser representadas com essa resolução de cores é maior do que a visão humana pode reconhecer.

### Cores Indexadas

Nas cores indexadas, cada pixel é representado por um índice que aponta para uma tabela de cores (paleta) que contém as informações sobre as cores (Figura 15). Temos então um número de cores que podem ser representadas, que é o número de entradas na paleta. A paleta por sua vez, tem em geral 24 bits para representar cada cor no formato RGB. Dessa forma podemos representar  $n$  cores de um conjunto com mais de 16 milhões de cores. Nesse caso, para representar esse tipo de imagem, as informações das cores da paleta devem constar da estrutura além das dimensões e seqüência de índices.



**Figura 15.** Índice e paleta de cores

O número de cores e a resolução de cor da paleta podem variar. Os dois padrões mais usados são apresentados na tabela abaixo.

Bits/pixel	Padrão	Resolução de cor da paleta (RGB)
4 bits/pixel	16 cores indexadas	24 bits/cor
8 bits/pixel	256 cores indexadas	24 bits/cor

### Cores fixas

Nas cores fixas, cada pixel é representado por um índice que aponta para uma tabela de cores fixa. Esse sistema geralmente é usado quando o dispositivo não permite a representação de muitas cores, como no caso de placas de vídeo antigas ou padrões de cores (padrão de cores do MS Windows 3.x, 16 cores). O número de bits para representar um pixel depende do número de cores fixas, ou seja para representar por exemplo 16 cores, são necessários 4 bits/pixel.

### Imagens em Tons de Cinza

A representação de imagens em tons-de-cinza é feita discretizando a informação de luminância de cada ponto da imagem. Ou seja, cada pixel contém a intensidade de luminosidade representada em um certo número de bits. Assim uma imagem com resolução de cor de 8 bits, pode representar até 256 níveis de cinza, variando do preto ao branco.

Os padrões mais usados são de 16 e 256 tons-de-cinza, 4 e 8 bits/pixel respectivamente. Representações com mais que 256 tons-de-cinza não são percebidas pela vista

humana, ou seja representar uma imagem com 256 níveis é suficiente para a maioria das aplicações.

### **Imagens Binárias**

As imagens binárias são imagens com dois níveis, como preto e branco. São muito usadas por dispositivos de impressão e para representar imagens de documentos monocromáticos. Para representar um pixel de uma imagem binária, como o próprio nome diz, é necessário apenas 1 (um) bit. Essa informação é suficiente para representar cada pixel, ou seja temos uma representação de 1 bit/pixel. Em alguns casos, temos uma informação extra sobre a cor de cada informação, a cor para o bit com valor 0 (zero) e a cor para o bit de valor 1 (um). Essa informação de cor é geralmente é representada em 24 bits/cor no padrão RGB, podendo porém ser representada de outras formas.

#### **2.4.6 Sistema CIE 1931 XYZ**

A fim de permitir uma definição de cor independente do dispositivo de apresentação, o CIE definiu um sistema de representação de cores, chamado de CIE 1931 XYZ. Nele uma cor é definida por 3 valores XYZ, onde X, Y e Z podem variar de 0 a 1. Y identifica a luminância, X e Z a cor. Na prática, cores são raramente especificadas em termos XYZ, é mais comum o uso de coordenadas de cromaticidade que independem da luminância Y. As coordenadas de cromaticidade, x e y, são calculadas a partir dos valores de X, Y, Z usando as seguintes equações:  $x = X/(X+Y+Z)$ ,  $y = Y/(X+Y+Z)$

## **2.5 Principais características e Requisitos das Informações multimídia**

Esta seção resume as principais características das várias mídias apresentados neste capítulo e obtém a partir disto os principais requisitos de um sistema multimídia em rede.

#### **2.5.1 Requisitos de armazenamento e largura de banda**

Requisito de armazenamento é medido em termos de bytes ou Mbytes. No domínio digital, a largura de banda é medida como taxa de bits em bits/s ou Mbits/s. A unidade para armazenamento é byte e para largura de banda é bit.

### **Imagens**

Para imagens, o requisito de armazenamento pode ser calculado a partir do número de pixels (H) em cada linha, o número de linhas (V) na imagem e o (P) número de bits por pixel, da seguinte forma: **Requisito de armazenamento = HVP/8**. Por exemplo, uma imagem com 480 linhas, 600 pixels cada linha e um número de bits por linha igual a 24 necessita 864 Kbytes para representar a imagem.

A largura de banda necessário para a transmissão da imagem pode ser calculado a partir do seu requisito de armazenamento. Por exemplo, se a mensagem acima (864 Kbytes) deve ser transmitida em 2 segundos, então a largura de banda necessário é 3,456 Mbits/s. Em muitas aplicações, imagens devem ser apresentadas em sincronia com mídias contínuas tal como áudio. Neste caso, a transmissão de imagem impõem tempo restrito e requisitos de largura de banda.

### **Áudios e Vídeos**

Áudios e vídeos são mídias contínuas que são normalmente caracterizadas em bits/s ou Mbits/s. Para áudio este número é calculado baseado na taxa de amostragem e no número de bits por amostragem. Para vídeo este cálculo é baseado na quantidade de dados em cada quadro e no número de quadros por segundo. O resultado especifica a taxa de bits necessária ao canal de transmissão. Caso a duração do áudio ou do vídeo seja conhecido, o montante de requisito de armazenamento pode ser calculado.

A tabela abaixo apresenta os requisitos de largura de banda de áudios e vídeos de diferentes qualidades.

Aplicações	Taxa de transmissão (kbits/s)
CD-Audio	1.411,2
DAT	1.536
Telefone Digital	64
Radio digital, long play DAT	1.024
Vídeo de qualidade televisão	216.000
Vídeo de qualidade VHS	54.000
HDTV	864.000

A tabela abaixo mostra os requisitos de armazenamento para mídias estáticas e dinâmicas comuns de diferentes durações.

Aplicações	Requisitos de Armazenamento (MBytes)
Livro de 500 páginas	1
100 imagens monocromáticas	7
100 imagens coloridas	100
1h de áudio qual. telefone	28,8
1h de Áudio-CD	635
1h Vídeo qualidade VHS	24,3
1h TV	97000
1h HDTV	389000

## 2.5.2 Relações temporais e espaciais entre mídias

Em computação e comunicação multimídia, as diversas mídias estáticas e dinâmicas podem estar relacionadas em uma aplicação ou apresentação. Para obter o efeito desejado, a busca e a transmissão dos dados deve ser coordenada e apresentada de forma que as relações temporais sejam mantidas. Neste contexto o aparecimento temporalmente correto ou desejado de informações multimídia é chamado de **sincronização**. **Esquemas de Sincronização** definem os mecanismos usados para obter o grau requerido de sincronização.

Existem duas áreas de trabalho em sincronização multimídia:

- Desenvolvimento de mecanismos e ferramentas para deixar o autor especificar os requisitos temporais
- Garantia das relações temporais especificadas superando a natureza não determinística dos sistemas de comunicação.

## 2.5.3 Requisitos de atrasos e variações de atrasos (Jitter)

Para obter uma qualidade razoável na apresentação de áudios e vídeos, amostras de áudio e vídeo devem ser recebidas e apresentadas em intervalos regulares. Por exemplo, se uma peça de áudio é amostrada numa taxa de 8 kHz, ele deve ser apresentado a 8000 amostras por segundo. Como mídias contínuas têm essa dimensão temporal e os componentes do sistemas podem atuar assincronamente, suas correções dependem não apenas dos valores das amostras, mas também do tempo de apresentação das amostras.

**Atrasos fim-a-fim** é a soma de todos os atrasos em todos os componentes de um sistema multimídia, incluindo acesso a disco, conversão A/D, codificação, processamento no hospedeiro, acesso a rede, transmissão, buffering, decodificação e conversão D/A. O atraso aceitável é muito subjetivo e é dependente de aplicação:

- Aplicações de conversações ao vivo necessitam a manutenção da natureza interativa, para tal o atraso não pode ser superior a 300ms.
- Para aplicações de recuperação de informação o requisito de atraso não é muito forte desde que o usuário não aguarde muito pela resposta. Em muitas aplicações o atraso de alguns segundos é tolerável.

Para mídias contínuas a variações de atrasos deve ser pequena. Para voz com qualidade de telefone e vídeo com qualidade de televisão a variação de atraso deve ser inferior a 10ms. No caso de áudio de alta qualidade esta variação deve ser muito pequena (<1ms) isto pois nossa percepção do efeito estéreo é baseado nas diferenças de fase mínimas.

Tanto o atraso quanto a variação de atraso devem ser garantidos em toda a seção de comunicação. Isto não é suportado pelas redes, protocolos de transporte, sistemas operacionais usuais.

#### 2.5.4 Tolerância a erros e perdas em dados multimídia

Diferentes dos dados alfanuméricos, onde perdas e erros na transmissão são na sua grande maioria intoleráveis. Erros ou perdas em dados de áudio, vídeo e imagens podem ser tolerados. Isto pois estas perdas e erros de bits não são desastrosos e geralmente não são percebidos pelo usuário.

Para voz, nós podemos tolerar uma taxa de erros de bit de  $10^{-2}$ . Para imagens e vídeos pode-se tolerar uma taxa de bits de  $10^{-4}$  a  $10^{-6}$ . Outro parâmetro que mede o erro é a taxa de perda de pacotes. Os requisitos de taxa de perdas de pacote é mais forte que a de erros de bit, porque uma perda de pacote pode afetar a decodificação de uma imagem por exemplo. Quando técnicas de compressão são utilizadas a taxa de erro de bits deve ser pequena pois um erro de bit pode causar um erro de descompactação de muitos bits

Técnicas de recobrimento de erros podem ser empregadas para aumentar a qualidade de áudio e vídeo.

### 2.6 Qualidade de Serviço

Como discutido na seção anterior, dados multimídia impõem duros e diversos requisitos a sistemas multimídia. Eles requerem largura de banda, capacidade de armazenamento alto e alta taxa de transferência, limitação de atraso e variação, sincronizações espaciais e temporais. O conceito de **Qualidade de Serviço** (QoS) é utilizado para especificar o conjunto de parâmetros de requisitos. Não há um acordo universal deste conjunto, mas parâmetros de requisitos mais comuns cobrem os requisitos mencionados acima, incluindo largura de banda (taxa de transferência), limitações de atrasos e variações e requisitos de sincronização. Estes parâmetros são especificados em duas grades: qualidade preferível e qualidade aceitável.

QoS é um contrato negociado entre aplicações e sistemas multimídia (provedores de serviço). Quando uma aplicação necessita partir uma sessão ela submete um pedido com a QoS requerida para o sistema. O sistema rejeita ou aceita possivelmente com alguma negociação. Quando o sistema aceita o pedido, um contrato entre o sistema e a aplicação é assinado e o sistema deve prover a QoS requerida. Existem 3 formas de garantir este contrato:

- **Garantia determinística** (hard), onde a QoS deve ser satisfeita completamente.
- **Garantia estática** (soft), onde o sistema fornece uma garantia com uma certa probabilidade de  $p$ .
- **Melhor esforço**, onde há não há garantia que a QoS será respeitada.

Pesquisas em QoS são ainda muito novas e muitos temas de pesquisa existem. Exemplos destes temas de pesquisa são:

- Como converter parâmetros de QoS em requisitos de recursos.
- Como escalar sessões de forma que mais sessões possam ser suportadas dado um conjunto fixo de recursos.

## Capítulo 3 Compressão de Dados Multimídia

Como visto no capítulo anterior, áudios, imagens e vídeos necessitam de uma grande quantidades de dados para representar/armazenar e uma grande largura de banda para serem transmitidos. Assim a compressão de dados é essencial para que as informações ocupem espaço aceitáveis em disco e que possam ser transmitidas via rede em taxas razoáveis de transmissão. Não existiria multimídia hoje sem o drástico progresso que ocorreu nos últimos anos em algoritmos de compressão e suas implementações [Fluckiger, 95].

### 3.1 A Necessidade da Compressão

Compressão de dados é uma forma de codificar um certo conjunto de informações de maneira que o código gerado seja menor que o fonte. O uso de técnicas de compressão é claramente essencial para aplicações multimídia. As razões são as seguintes [Furht, 94]:

- O grande requisito de armazenamento de dados multimídia.
- A velocidade relativamente lenta dos dispositivos de armazenamento que não podem apresentar dados multimídia (principalmente vídeo) em tempo-real.
- A largura de banda da rede que não permite a transmissão de vídeo em tempo-real.

#### **Requisitos de armazenamento**

Caso nenhuma técnica de compressão for utilizada, 80 MBytes do disco de um PC seria ocupado por 8 minutos de som estereofônico qualidade CD, ou 3,5 segundos de vídeo de qualidade TV. No caso dos CD-ROMs, eles podem ocupar 72 minutos de música de alta fidelidade, mas apenas 30 segundos de vídeo de qualidade TV.

Uma aplicação multimídia típica contém mais de 30 minutos de vídeo, 2000 imagens e 40 minutos de som estéreo. Sendo assim, a aplicação necessitaria, caso não fosse aplicada nenhuma compressão, de aproximadamente 50 GBytes para armazenar o vídeo, 15 Gbytes para armazenar as imagens e 0,4 GBytes para armazenar o áudio. O que significa um total de 65,4 GBytes de armazenamento no disco. Assim é extremamente necessária a utilização de técnicas de compressão de dados multimídia para viabilizar o armazenamento destas informações.

#### **Velocidade de transferência dos dispositivos de armazenamento**

Mesmo que tenhamos enorme capacidade de armazenamento, nós não seríamos capazes de apresentar um vídeo em tempo-real devido a taxa de bits insuficiente dos dispositivos de armazenamento. Por exemplo, um dispositivo de armazenamento deveria ter uma taxa de 30 MBytes/s para apresentar um vídeo em tempo real com um quadro de 620x560 pixels a 24 bits por pixel de 30 fps. Mas a tecnologia de CD-ROM de hoje fornece uma taxa de transferência de até 300 Kbytes/s. No estado atual da tecnologia de armazenamento, a única solução é comprimir o dado antes de armazenar e descomprimir ele antes da apresentação.

#### **Largura de banda da rede**

Com relação a velocidade de transmissão, caso um som estereofônico de qualidade CD não compactado devesse ser transmitido, a rede deveria suportar uma taxa de 1,4 Mbits/s. Isto é possível em redes locais, mas em redes de media e longa distância este taxa torna-se atualmente inviável: a máxima taxa de transferência a ser compartilhada tipicamente varia de 10 Mbits/s (tecnologia Ethernet) a 100 Mbits/s (tecnologia fast Ethernet ou FDDI). No caso de um vídeo de qualidade PAL não compactado é necessário uma taxa de 160 Mbits/s, isto é incompatível com a maior parte das redes

locais. Sobre rede de longas distâncias, circuitos de cabos terrestres nesta taxa existem, mas o custo de dedicar a rede para uma único canal de vídeo torna esta transmissão proibitiva.

Técnicas de compressão modernas de imagem e vídeo reduzem tremendamente os requisitos de armazenamento e portanto os requisitos de largura de banda da rede e do dispositivo de armazenamento. Técnicas avançadas podem comprimir uma imagem típica em uma razão variando de 10:1 a 50:1 e para vídeo de até 2000:1.

## 3.2 Princípios de Compressão

Técnicas de compressão de dados multimídia exploram basicamente dois fatores: a redundância de dados e as propriedades da percepção humana.

### 3.2.1 Redundância de Dados

Recapitulando o capítulo anterior: um áudio digital é uma série de valores amostrados; uma imagem digital é uma matriz de valores amostrados (pixels); e um vídeo digital é uma seqüência de imagens apresentadas numa certa taxa. Geralmente amostras subsequentes de áudios e imagens (para vídeo) não são inteiramente diferentes. Valores vizinhos são geralmente de algum modo relacionados. Esta correlação é chamada **redundância**. A remoção desta redundância não altera o significado do dado, existe apenas uma eliminação da replicação de dados.

#### **Redundância em áudio digital**

Em muitos casos, amostragens de áudio adjacentes são similares. A amostra futura não é completamente diferente da passada, o próximo valor pode ser previsto baseado no valor atual. A técnica de compressão que se aproveita desta característica do áudio é chamada de **codificação predictiva**. Técnicas de compressão predictiva são baseadas no fato que nós podemos armazenar a amostra anterior e usar esta para ajudar a construir a próxima amostra.

No caso da voz digital há outro tipo de redundância: nós não falamos todo o tempo. Entre uma rajada e outra de informações há instantes de silêncio. Este período de silêncio pode ser suprimido sem a perda de informações, sabendo que este período é mantido. Esta técnicas de compressão é chamado da **Remoção de silêncio**.

#### **Redundância em imagem digital**

Em imagens digitais as amostras vizinhas em uma linha de escaneamento e as amostras vizinhas em linhas adjacentes são similares. Esta similaridade é chamada de **Redundância espacial**. Ela pode ser removida, por exemplo utilizando técnicas de codificação predictiva ou outras.

#### **Redundância em vídeo digital**

Vídeo digital é uma seqüência de imagens, portanto ele também tem redundância espacial. Além disso, imagens vizinhas em vídeos são geralmente similares. Esta redundância é chamada de **redundância temporal**. Ela pode também ser removida, por exemplo utilizando técnicas de codificação predictiva.

### 3.2.2 Propriedades Percepção Humana

O usuário final das aplicações multimídia são geralmente humanos. Humanos podem tolerar alguns erros de informação ou perdas sem afetar a efetividade da comunicação. Isto implica que a versão comprimida não necessita representar exatamente a informação original. Isto é bem diferente dos dados alfanuméricos que não se tolera qualquer erro ou perda (por exemplo, se ocorrer uma perda de uma informação em um programa ele pode não funcionar corretamente).

Como os sentidos humanos não são perfeitos, pequenas perdas e erros em áudios e vídeos não são percebidos. Além disso, algumas informações são mais importantes para a percepção humana que outras (por exemplo no caso de imagens, a intensidade luminosa é mais importante que a cor). Assim na hora de compactar uma certa informação, alguns dados de imagens, vídeos e sons podem ser ignorados pois suas apresentações ou não é completamente indiferente para os humanos.

### 3.3 Classificação das Técnicas de Compressão

Existem várias técnicas de compressão, elas podem ser classificadas de diversas maneiras: baseadas no algoritmo de compressão e no resultado das técnicas de compressão. Neste documento, elas serão classificadas quanto ao resultado. Esta classificação é mais útil para o usuário final e desenvolvedores de aplicações multimídia do que a classificação baseada em algoritmo.

#### 3.3.1 Técnicas de compressão com e sem perda

Se a informação, após sua compressão, pode ser exatamente reconstruída a técnica de compressão é dita sem perdas. Esta técnica deve ser utilizada obrigatoriamente para comprimir programas e documentos legais ou médicos. As técnicas de compressão sem perda não são idéias novas, elas são muito utilizadas. Estas técnicas exploram apenas estatísticas de dados (redundância de dados) e a taxa de compressão é normalmente baixa. Um exemplo deste tipo de compressão é substituir caracteres de espaços ou zeros sucessivos por um flag especial e o número de ocorrências. Como exemplo de técnicas sem perda temos: codificação aritmética, codificação Huffman e codificação Run-length.

Técnicas de compressão com perdas são utilizadas para compressão de áudio, imagens e vídeos, onde erros e perdas são toleráveis. Estas técnicas são baseadas normalmente em estatísticas de dados e propriedades da percepção humana. Com ela, altas taxas de compressão podem ser obtidas.

#### 3.3.2 Técnicas por Entropia, Codificação na Origem e Híbrida

A codificação por entropia (*Entropy encoding*) trata de cadeias de bits sem levar em conta seu significado [Tanembau, 97]. É uma técnica genérica, sem perda e totalmente reversível, que pode ser aplicada a todos os dados. Nesta apostila são apresentados alguns exemplos de técnicas por entropia, que são a codificação run-length e de Huffman.

Codificação na origem (*Source coding*) processa o dado original distinguindo o dado relevante e o irrelevante. Elas levam em consideração a semântica dos dados. Removendo os dados irrelevantes compõe o dado original. Como exemplo de técnicas de compressão da origem, nós temos: DPCM (*Differential pulse code modulation*), DCT (*discrete cosine transform*) e DWT (*Discrete wavelet transform*).

Codificação híbrida é a combinação de técnicas de compressão sem perdas e técnicas de codificação na origem. Normalmente, várias destas duas técnicas são agrupadas para formar uma nova técnica de codificação híbrida. Como exemplo deste tipo de técnica de compressão podemos citar os padrões H.261, H.263, JPEG, MPEG vídeo e áudio.

#### 3.3.3 Codificação com taxa de bits constantes e variáveis

Como apresentado no capítulo anterior, na conversão analógico para digital as amostragens são feitas em intervalos regulares e cada amostragem é representada através de um mesmo número de bits. Portanto a transmissão de áudios e vídeos digitais não compactados são formadas por fluxos de taxa de bits constantes. Algumas técnicas de compressão produzem fluxos de taxa de bits constantes e outras não: nós temos as técnicas de compressão com taxa de bits constantes (CBR) e as técnicas de compressão com taxa de bits variáveis (VBR).

É importante classificar se uma técnica é CBR ou VBR. Primeiro, se conteúdos são complexos, mais dados são necessários para representar. Isto é normalmente verdade para técnicas VBR e são normalmente mais eficientes e produzem áudios e vídeos de alta qualidade na mesma taxa de compressão. Segundo, técnicas VBR são difíceis de modelar e especificar, assim elas são difíceis de serem suportadas por sistemas de comunicação multimídia.

### 3.4 Medição do Desempenho de Compressão

No desenvolvimento de uma aplicação multimídia, os autores devem escolher que técnica de compressão utilizar. Esta escolha geralmente baseia-se nas classificações apresentadas anteriormente, nos parâmetros de desempenho da técnica e nos requisitos da aplicação. Os parâmetros de desempenho mais usados são:

- **Taxa de compressão:** razão entre o tamanho do dado original e o tamanho do dado após a compressão. No caso de técnicas sem perda, quanto maior a taxa de compressão melhor é a técnica de compressão. Para técnicas de compressão com perda deve-se considerar também a qualidade da mídia restituída.
- **Qualidade da mídia reconstituída:** medida em SNR (Razão Sinal/Ruído). Este parâmetro é aplicável apenas para técnicas com perda. Para a escolha de uma técnica de compressão com perdas, deve-se optar pelo compromisso entre uma alta taxa de compressão e a qualidade desejada para a aplicação em desenvolvimento.
- **Complexidade de implementação e velocidade de compressão:** geralmente quanto mais complexa a técnica menor é a velocidade de compressão. No caso de aplicações tempo-real, como videoconferência, estes parâmetros devem ser considerados. Isto pois a compressão/descompressão deve ser realizadas em tempo-real. No caso de aplicações do tipo obtenção e apresentação de informação a velocidade de compressão não é muito importante, mas a velocidade de descompressão é importante.

### 3.5 Técnicas de Compressão sem Perdas

Nesta seção são apresentadas algumas técnicas de compressão sem perda.

#### 3.5.1 Codificação Run-Length

Codificação run-length é uma codificação por entropia. Parte dos dados de imagem, áudio e vídeo amostrados podem ser comprimidos através da supressão de seqüências de mesmos bytes. Estas seqüências são substituídas por um número de ocorrências e um símbolo padrão (padrão de bits) para anotar a repetição em sim. Obviamente, o fator de compressão alcançável depende do dado de entrada. Usando uma marca de exclamação como flag especial para indicar a codificação run-length, o seguinte exemplo mostra como um fluxo de dados pode ser comprimidos substituindo a seqüência de seis caracteres "H" por "!6H":

- Dado original UHHHHHHIMMG1223
- Dado comprimido: U!6HIMMG1223

É claro que esta técnica não é utilizada para seqüências de caracteres iguais ou menores que quatro. Isto pois nenhuma compressão seria obtida neste caso. Por exemplo, substituindo a seqüência de dois caracteres "M" com o código run-length "!2M" aumentaria o tamanho do código em um byte. Se o flag especial no nosso exemplo ocorrer no dado, ele deve ser substituído por duas marcas de exclamação (*byte stuffing*).

O algoritmo apresentado acima pode ser facilmente otimizado; por exemplo, em vez de seqüências simples de caracteres, sentenças mais longas de diferentes caracteres podem também ser substituídas. Esta extensão requer que o tamanho da seqüência seja codificado ou pode-se utilizar um flag especial de fim. Existem diversas variações da codificação run-length.

Este método só traz ganhos relevantes se houver grandes agrupamentos de símbolos iguais. As principais aplicações do método de Run-Length, são em imagens binárias, imagens com grandes espaços envolvendo uma só cor e em imagens geradas por computador, onde os dados estão agrupados de forma mais geometricamente definida. Esse método é aplicado em formatos padrões como PCX, BMP(RLE).

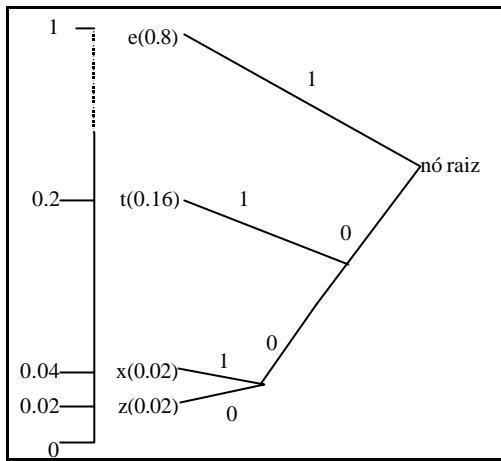
### 3.5.2 Codificação de Huffman

Neste método de compressão, é atribuído menos bits a símbolos que aparecem mais freqüentemente e mais bits para símbolos que aparecem menos. Assim, o tamanho em bits dos caracteres codificados serão diferentes. Codificação de Huffman é um exemplo de técnica de codificação estatística, que diz respeito ao uso de um código curto para representar símbolos comuns, e códigos longos para representar símbolos pouco freqüentes. Esse é o princípio do código Morse, em que E é • e Q é ---, e assim por diante.

Nós usaremos um exemplo para mostrar como a codificação de Huffman funciona. Suponha que nós temos um arquivo contendo 1000 caracteres, que são e, t, x e z. A probabilidade de ocorrência de e, t, x, e z são 0.8, 0.16, 0.02, e 0.02 respectivamente. Em um método de codificação normal, nós necessitamos 2 bits para representar cada um dos quatro caracteres. Assim, nós necessitamos de 2000 bits para representar o arquivo. Usando a codificação de Huffman, nós podemos usar quantidades de bits diferentes para representar estes caracteres. Nós usamos bit 1 para representar e, 01 para representar t, 001 para representar x e 000 para representar z. Neste caso, o número total de bits necessários para representar o arquivo é  $1000*(1*0.8+2*0.16+3*0.02+3*0.02)=1240$ . Assim, embora tenhamos utilizado mais bits para representar x e z, desde que seus aparições são mais raras, o número total de bits necessários para o arquivo é menor que o esquema de codificação uniforme. As regras para atribuir bits (códigos) aos símbolos é chamado um codebook. Codebooks são normalmente expressos em tabelas: w(e)=1, w(t)=01, w(x)=001, w(z)=000.

Agora vamos ver como os códigos Huffman são gerados. O procedimento é o seguinte (Figura 16):

- a) Coloque de todos os símbolos ao longo de uma linha de probabilidade acumulativa na seguinte ordem: probabilidade dos símbolos aumenta de baixo para cima. Se dois símbolos têm a mesma probabilidade, eles podem ser colocados em qualquer ordem.
- b) Junta-se os dois símbolos de menor probabilidade a um nó para formar dois ramos na árvore.
- c) A nova árvore formada é tratada como um símbolo único com a probabilidade igual a soma dos símbolos ramos.
- d) Repita os passos (b) e (c) até que todos os símbolos sejam inseridos na árvore. O último nó formado se chama o nó raiz.
- e) Partindo do nó raiz, atribua o bit 1 ao ramo de maior prioridade e bit 0 ao ramo de menor prioridade de cada nó.
- f) O código para cada símbolo é obtido montando códigos ao longo dos ramos do nó raiz para a posição do símbolo na linha de probabilidade. Por exemplo, lendo do nó raiz ao símbolo x, nós obtemos o código 001 para x.



**Figura 16.** Exemplo de árvore de codificação Huffman

A operação computacional mais custosa na determinação do código Huffman é a adição de floats, mais especificamente, a adição da probabilidade da ocorrência no processo de redução. Isto ocorre no lado do codificador. No lado do decodificador, ele tem que apenas realizar uma simples verificação na tabela. Portanto, o decodificador necessita da tabela Huffman usada no codificador. Esta tabela é parte do fluxo de dados ou já é conhecida pelo decodificador. Em áudio e vídeo, tabelas Huffman padrões são com muita freqüência utilizadas, isto é, tabelas são conhecidas pelo codificador e decodificador. A vantagem é a obtenção de uma codificação mais rápida pois as tabelas não precisam ser calculadas. A desvantagem é que tabelas Huffman padrões obtém um fator de compressão um pouco menor porque as tabelas não são necessariamente ótimas para o dado a ser codificado. Portanto, métodos de compressão executados em tempo-real usam normalmente tabelas padrões pois a codificação é mais rápida. Se alta qualidade é necessária mas tempo de codificação não é importante, tabelas Huffman otimizadas podem ser utilizadas.

Normalmente nem todos os caracteres tem uma representação codificada na tabela Huffman: apenas aqueles caracteres com alta probabilidade de ocorrência. Todos os outros são codificados diretamente e marcados com um flag especial. Esta técnica é útil quando um número de caracteres diferentes é muito grande mas apenas alguns deles tem uma alta probabilidade de ocorrência.

### 3.5.3 Codificação de Lempel-Ziv-Welch (LZW)

Codificação LZW é baseada na construção de um dicionário de frases (grupos de um ou mais caracteres) a partir do fluxo de entrada. Quando uma nova frase é encontrada, a máquina de compressão adicionada ao dicionário e um token que identifica a posição da frase no dicionário substitui a frase. Se a frase já foi registrada, ela é substituída pelo token de posição no dicionário. Esta técnica é boa para compressão de arquivos textos, onde temos uma grande repetição de frases, como por exemplo em português: "ela", "Contudo,", ", "onde, ", aparecem freqüentemente no texto.

O seguinte exemplo mostrará o poder da codificação LZW. Suponha que temos um arquivo de 10000 caracteres. Se nós representarmos o arquivo usando 8 bits por caractere, o arquivo requer 80000 bits para representá-lo. Usando o algoritmo LZW e assumindo que o arquivo tenha 2000 palavras ou frases das quais 500 são diferentes, então nós necessitamos 9 bits como token para identificar cada palavra ou frase distinta. Assim, nós precisamos de  $9 \times 2000$  bits para codificar o arquivo. Com isto nós obtemos uma taxa de compressão de 4,4. Na prática, o dicionário armazenando todas as frases únicas deve ser armazenado também, baixando a taxa de compressão obtida.

#### **Exemplo de Compressão**

O funcionamento básico de LZW será ilustrado através da compressão da cadeia de caracteres ABACABA:

- O primeiro passo consiste em inicializar uma tabela de códigos com todos os caracteres existentes na string que pretendemos comprimir: #0 = A, #1 = B, #2 = C, #3 = D.
- Em seguida, começamos a analisar a string e verificamos que o primeiro caracter é o A, que já existe na tabela. Nesta situação, atribuímos um prefixo ao A e vamos analisar o caracter seguinte, que é o B. Juntando o prefixo com o novo caracter, obtemos a sub-string AB que não se encontra na tabela. Neste caso, enviamos o código #0, correspondente ao A e adicionamos a nova sub-string à tabela, fazendo #4 = AB.
- Como próximo passo, tomamos para prefixo o caracter B e indo buscar o próximo caracter da string obtemos a sub-string BA, a qual também não se encontra na tabela. Repetindo o procedimento anterior, enviamos o código #1 e adicionamos a sub-string à tabela, fazendo #5 = BA.
- O novo prefixo é o A, o qual concatenado com o próximo caracter da string forma a sub-string AC que também não está na tabela. Assim, enviamos o código #0 e fazemos #6 = AC.
- O novo prefixo é o C, o qual concatenado com o próximo caracter da string forma a sub-string CA que também não se encontra na tabela. Enviamos o código #2, correspondente ao prefixo e fazemos #7 = CA.
- O novo prefixo é o A, o qual concatenado com o próximo caracter da string forma a sub-string AB, a qual já se encontra codificada na tabela. Desta forma, não enviamos nenhum código e tomamos para novo prefixo AB.
- Este prefixo concatenado com o próximo caracter da string forma a sub-string ABA que não se encontra na tabela. Enviamos o código #4, correspondente a AB, e fazemos #8 = ABA.
- O novo prefixo passa a ser A e como a string não tem mais caracteres, enviamos o código correspondente a este caracter, ou seja, o código #0.
- Desta forma obtemos a seguinte seqüência codificada: #0 #1 #0 #2 #4 #0.

### **Aplicação do LZW**

O algoritmo LZW pode ser encontrado em diversos descompactadores largamente utilizados atualmente. Dentro eles, citamos o PKUNZIP (para o DOS), ARJ, GUNZIP e o Uncompress (para o UNIX).

### **Formato de imagem GIF (*Graphical Interchange Format*)**

Uma das aplicações do LZW é a variação implementada no formato GIF, utilizado no armazenamento de imagens. Atualmente, este é um dos formatos de armazenamento de imagens sem perdas que oferece as melhores taxas de compressão. Sendo sem perdas, o GIF preserva todos os dados visuais na figura: nenhuma informação é descartada ou alterada durante a compressão/descompressão. Com isto, as taxas de compressão não são muito grandes, em geral 4:1.

O formato GIF apenas admite o tratamento de imagens com uma profundidade de cor de até 8 bits/pixel, ou seja, imagens com um máximo de 256 cores. Se possuirmos uma imagem true color, com 24 bits/pixel, ao convertermos esta imagem para o formato GIF, estamos perdendo grande parte da informação de cor. Neste caso, verifica-se que a qualidade da imagem obtida é bastante inferior à qualidade da imagem original, devido ao aparecimento de ruído em toda a imagem.

O algoritmo LZW é propriedade da Unisys. No início este algoritmo era de domínio público, no entanto, há relativamente pouco tempo, a Unisys resolveu passar a cobrar uma taxa pela sua utilização. Por este motivo, começou a pensar-se numa alternativa válida ao formato GIF, tendo surgido o formato PNG (Portable Network Graphics). PNG suporta até true color 48 bits por pixel ou 16 bits por pixel para escalas de cinza e não suporta animação. PNG usa os algoritmos de compressão Lempel-Ziv 77 (LZ77) e de Huffman.

## 3.6 Técnicas de Compressão de Áudio Digital

### 3.6.1 Quantificação não linear

Quando o mesmo tamanho de passo de quantificação é usado na conversão A/D sem olhar para a amplitude do sinal, o processo de conversão é dito uniforme. Este processo de conversão A/D é chamado de **modulação por pulso codificado** (PCM). Ele é simples mas não é eficiente: a quantificação linear resulta em uma mais elevada SNR na região de amplitude de sinal mais altas que na região de mais baixa amplitudes. Esta elevada SNR na região de amplitude mais altas não aumenta a qualidade percebida, isto pois nós somos mais sensíveis às componentes de amplitude mais baixas.

A fim de explorar este fato o tamanho de passo de quantificação que aumenta logaritmicamente com a amplitude do sinal é muito usado na quantificação de sinais de voz. Neste caso, os passos de quantificação são menores quando a amplitude é baixa. Esta técnica de compressão realizada uma transformação de um sinal linear em um sinal não linear.

Na prática, uma quantificação uniforme é aplicada a um sinal não linear transformado em vez de aplicar uma quantificação não uniforme ao sinal linear. Os resultados destas duas abordagens é o mesmo. O processo de transformação de um sinal linear em não linear é chamado de **companding**. A digitalização uniforme de um sinal **companded** é chamado de **companded PCM**. Esta é na realidade uma técnica de compressão analógica realizada antes da conversão A/D e expandida após a conversão D/A. Usando esta técnica, o sinal de 8 bits pode produzir um sinal de qualidade equivalente aquele sinal codificado PCM de 12 bit.

Na área da telefonia digital, utiliza-se um método de transformação de natureza logarítmica para comprimir áudios. Ele mapeia 13 ou 14 bits dos valores linearmente quantificados para códigos de 8 bits. O mapeamento de 13 para 8 é conhecido como transformação A-law, e o mapeamento de 14 para 8 é conhecido como transformação  $\mu$ -law. Usando esta transformação, a SNR da saída transformada é mais uniforme na faixa de amplitude do sinal de entrada. A transformação A-law é usada normalmente em redes ISDN (Redes Digitais de Serviços Integrados) na Europa, e  $\mu$ -law na América do Norte e Japão. A recomendação ITU G.711 (vista mais adiante), especifica as transformações A-law e  $\mu$ -law.

### 3.6.2 Codificação Predictiva

Na codificação predictiva, em vez de se transmitir uma amostra, a diferença entre uma previsão do valor da amostra e do valor real é codificada e transmitida. Esta diferença é chamada de erro de predição. Se esta diferença é quantificada e codificada, o esquema de codificação é chamado de PCM diferencial (DPCM), também chamado de modulação delta. Desde que o valor previsto por cada amostra é calculado apenas a partir do passado do sinal codificado, o valor previsto é disponível tanto para o codificador quanto para o decodificador. No momento da decodificação, o erro de predição é adicionado ao valor previsto da amostra.

A eficiência desta técnica é obtida explorando o fato que valores de amostras vizinhas são correlacionados e o erro de previsão será menor que o valor original da amostra. Em muitos casos, poucos bits por amostra necessitam ser transmitidos. Por exemplo, compactando um fluxo de áudio de 16 bits, utilizando 4 bits para representar o erro de predição, a qualidade pode ser mantida se o erro de predição for sempre menor que 16 passos de quantificação original.

A modulação delta apresenta alguns problemas. O variante mais prático é chamado de DPCM adaptativo (ADPCM). Para manipular sinais que mudam rapidamente tão bem quanto sinais que mudam lentamente, o tamanho passo de quantificação aumenta com o aumento da variação do sinal. Então se a forma de onda está trocando rapidamente, grandes passos de quantificação são utilizados e vice-versa. Este é o método usado em CD-I (compact disc-interative).

Esta técnica reduz a taxa de bits de áudio de alta qualidade de 1,4 Mbps para 32 kbps.

### 3.6.3 Sumário das recomendações ITU-TS para codificação de voz

A ITU-TS (International Telecommunication Union - Telecommunications Sector) recomenda uma série de esquemas de compressão de voz. Essas recomendações são resumidas na tabela abaixo.

Recomendação	Técnica de Compressão	Largura de banda da voz (kHz)	Taxa de amostragem (kHz)	Taxa de bit compactado (Kbits/s)
G.711	PCM (não compactado)	3,4	8	64
G.721	ADPCM	3,4	8	32
G.722	ADPCM sub-banda	7	16 (14 bits por amostra)	48, 56, 64
G.723	ADPCM	3,4	8	24
G.728	Quantificação vetorial	3,4	8	16

O G.711 define a representação de voz não compactada, usando modulação PCM. A recomendação G.721 converte um fluxo de bits de 64 Kbits/s em um fluxo de 32 Kbits/s. G.721 é baseado na técnica ADPCM e sua taxa de amostragem é de 8 kHz. Com exceção do G.722, 8 bits são usados para cada amostragem.

O G.722 fornece uma melhor qualidade de som que o esquema convencional G.711 PCM ou a técnica de compressão G.721. G.722 é baseado no método ADPCM Sub-banda, onde o sinal de voz é dividido em duas sub-bandas (bandas de alta e de baixa freqüência). O propósito desta subdivisão é que a sub-banda de baixa freqüência é mais importante, assim ela precisa ser codificada com maior precisão. A largura de banda de um sinal compactado G.722 varia de 50 Hz a 7 kHz.

G.723 é outro padrão de compressão sem perdas baseado no ADPCM que opera a 24 Kbits/s. A qualidade de som resultante é inferior ao padrão G.711 PCM não compactado e ao G.722. Este padrão tem pouco uso.

O G.728 objetiva baixas taxas de bits. Ele opera a 16 Kbits/s mas sua largura de banda é limitada a 3,4 kHz. A qualidade de som resultante é significativamente inferior ao G.711 e G.722. Ele usa uma técnica de vetor quantificação, onde um número de amostragens são agrupadas em um vetor e cada vetor é representado por um índice na entrada do dicionário.

### 3.6.4 MPEG Áudio

Até aqui foi apresentado algumas técnicas de codificação de áudio especialmente projetadas para a voz, com uma suposição de que a largura de banda do áudio está dentro de 3,4 kHz até 7 kHz. Esta seção introduz uma técnica de compressão de áudio genérico que pode ser usada para comprimir sons dentro da faixa dos sons audíveis (até 20 kHz).

#### Mascaragem

Um som pode tornar outro som impossível de ser ouvido, ou ele pode tornar outro som sem peso. Isto é chamado de mascaragem. A mascaragem de um som pode ser total ou parcial. Como sons mascarados não são audíveis eles podem ser descartados. Esta é a estratégia utilizada pelos padrões ISO MPEG-1 Audio e MPEG-2 Audio. Aqui nós veremos o padrão MPEG-1 Audio, por simplificação abreviado para MPEG-Audio.

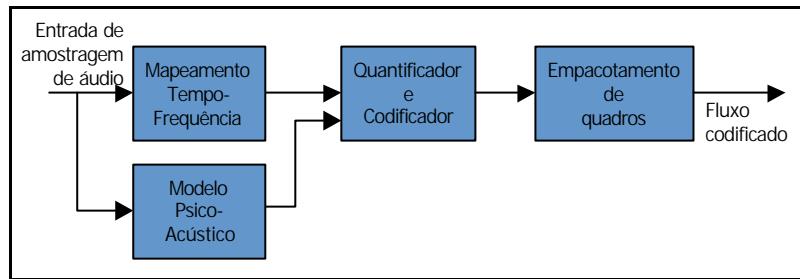
MPEG-Audio é um padrão de compressão de áudio genérico. Diferente de muitos outros codificadores especialmente projetados para sinais de voz, o codificador MPEG-Audio realiza a compressão sem a suposição acerca da natureza da fonte do áudio. Em vez disso, ele explora as limitações de percepção do sistema auditivo humano.

MPEG-Audio permite três freqüências de amostragens: 32, 44.1 ou 48 kHz. A seqüência de bits compactada pode suportar um ou dois canais de áudio. O fluxo comprimido pode ter uma das várias taxas de bits fixas e predefinidas variando de 32 a 244 Kbits/s. Dependendo da taxa de amostragem do áudio, o codificador MPEG-Audio pode ter uma razão de compressão variando de 2,7 a 24. Testes mostram que em taxas de compressão 6:1 ouvintes experientes não detectam diferenças.

### **Um codificador básico MPEG-Audio**

A Figura 17 mostra um diagrama de blocos de um codificador MPEG-Audio básico:

- Bloco mapeamento tempo-freqüência: divide a entrada em sub-bandas de freqüências múltiplas.
- Bloco modelo psico-acústico: cria um conjunto de dados para controlar a operação do bloco quantificador e codificador.
- Bloco quantificador e codificador: cria um conjunto de símbolos de código. As sub-bandas menos importantes e áudios inaudíveis são removidos.
- Bloco Empacotamento de quadros: monta e formata os símbolos de código e adiciona outras informações (tal como correção de erros) se necessário, para formar um fluxo de áudio codificado.



**Figura 17.** Codificador MPEG-Áudio Básico

O padrão MPEG-1-Audio define um fluxo de bits que pode suportar um ou dois canais de áudio: um canal único, dois canais independentes, ou um sinal estéreo. Os dois canais de um sinal estéreo podem ser processados independentemente ou como um estéreo conjunto (joined stereo) para explorar a redundância estéreo (correlações entre canais estéreo).

O padrão MPEG-2-Audio estende as funcionalidades do seu predecessor pela codificação multicanal com até cinco canais (esquerdo, direito, centro, e dois canais surround), mais um canal adicional de baixa freqüência, e/ou até sete canais multilíngues/comentários. Ele também estende a codificação estéreo e mono do MPEG-1-Audio com taxas de amostragens adicionais.

### **MP3**

MPEG Audio especifica uma família de 3 esquemas de codificação de áudio, chamadas de Layer-1, Layer-2 e Layer-3. O padrão especifica o formato do fluxo de bits e o decodificador para cada esquema de codificação. Ele não especifica o codificador em si para possibilitar avanços futuros.

De Layer-1 a Layer-3, a complexidade e desempenho (qualidade de som e taxa de bits) aumentam. Os três codificadores são compatíveis no modo hierárquico, isto é o decodificador Layer-N é capaz de decodificar um fluxo de bits fluxo codificado com codificador Layer-N e abaixo de N.

Dentre estes Layers, o esquema de codificação mais conhecido é o MPEG 1 Layer 3, conhecido como MP3. Este formato substituir o WAVE (arquivos de som do windows 95). Sua vantagem é a compressão, uma música de 4 min ocupando 45 MB em wave será reduzida a 4,5 MB no formato mp3. Existem várias aplicações para o MP3: para ouvir MP3 é necessários players especiais, como RealPlayer ou WinAmp; para criar MP3 é necessário codificadores de WAVE em MP3 e direto do CD para MP3.

O princípio de funcionamento básico do mp3 é buscar num sinal de áudio normal, como um arquivo wave, todas os sinais redundantes e irrelevantes que não sensibilizam nosso ouvido. Isto é relativo. O algoritmo de compactação do mp3 corta freqüências muito altas, acima dos 20kHz, que não são audíveis pelo ouvido humano. Só aí já são muitos bits economizados. Em qualquer música, se duas freqüências muito próximas foram "tocadas" aos mesmo tempo nosso ouvido somente ouvirá a mais forte, ou seja, o mp3 simplesmente diminui o número de bits desse sinal mais fraco e mantém os bits do sinal mais forte, diminuindo assim o tamanho final do arquivo na proporção 12:1 (qualidade semelhante ao CD).

Com a redução dos requisitos de armazenamento das músicas pensou-se em criar aparelhos digitais de música que tocassem o formato MP3: walkmans, portáteis e para carros, periféricos para computador, etc.

## 3.7 Técnicas de Compressão de Vídeo e Imagem Digital

Imagens e vídeos digitais puros são codificadas em PCM e eles são representados por vetores bi-dimensionais (possivelmente tri-dimensionais para imagens coloridas) de pixels. Reduzindo o conjunto de dados necessário à reprodução de imagens e vídeos (compressão) reduz requisitos de armazenamento, aumenta a velocidade de acesso, e é a única forma de obter vídeo digital em computadores pessoais.

Técnicas de compressão de imagens e vídeos digitais baseiam-se na alta redundância espacial e temporal destas mídias. No caso das imagens, geralmente certas áreas das figuras são uniformemente coloridas ou altamente correlatas (podendo formar padrões). Isto é chamado de redundância espacial ou correlação espacial. Esta redundância pode ser removida tanto quanto possível para uma certa qualidade de apresentação. No caso do vídeo, além da redundância espacial nós temos a redundância temporal ou correlação temporal, isto pois geralmente não existem grandes diferenças entre quadros de um vídeo. Para a compressão de vídeo, tanto a redundância temporal quanto a redundância espacial podem ser removidas e uma alta taxa de compressão pode ser obtida.

A seguir são apresentadas algumas técnicas de codificação de vídeo.

### 3.7.1 Técnica de Truncagem

Esta técnica de compressão é simples, ela consistem em truncar dados arbitrariamente baixando o número de bits por pixel. Isto é feito pela eliminação dos bits menos significativos de cada pixel. Esta técnica é atrativa pois ela é extremamente simples. Por exemplo, em imagens coloridas com 24 bits por pixel poderiam ser reduzidas para 16 bits por pixel.

### 3.7.2 Codificação de Sub-Amostragem Espacial e Temporal

Nesta técnica (também chamada de técnica de interpolativa), na codificação um pixel entre alguns pixels e um quadro entre alguns quadros são selecionados e transmitidos. No decodificador, os pixels e quadros faltantes são interpolados, baseado nas pixels e quadros recebidos, para gerar uma imagem ou uma seqüência de vídeo de baixa resolução. Como alternativa, o decodificador pode simplesmente decodificar e apresentar as imagens espacialmente sub-amostradas.

Esta técnica é utilizada em aplicações que não necessitam de alta resolução, como na videofonia. Ela é muito simples mas pouco eficiente. Por exemplo, quando um pixel entre quatro ou um quadro entre quatro são transmitidos, uma taxa de compressão 8:1 é obtida. Para aumentar a taxa de compressão, dados de imagens sub-amostradas temporalmente e espacialmente podem ser adicionalmente compactados usando outras técnicas.

### 3.7.3 Codificação Predictiva

Similar a codificação predictiva de áudio, esta técnica considera a estatística do sinal da imagem e o sistema visual humano para a compressão de áudios e vídeos. Em geral, os valores amostrados de elementos vizinhos de imagens são correlacionados e quadros de um vídeo têm correlação temporal entre os elementos de imagens em quadros sucessivos.

Correlação ou dependência estatisticamente linear indica que uma previsão linear dos valores amostrados pode ser feita. Esta previsão é baseada nos valores amostrados dos elementos vizinhos da figura. Geralmente esta previsão resulta em um erro de predição que tem uma variância menor que os valores amostrados.

Algoritmos de predição unidimensional usam correlação de elementos de imagem adjacentes dentro de uma linha. Outros esquemas mais complexos também exploram correlações linha-a-linha e quadro-a-quadro e são chamados de predição bidimensional e tridimensional, respectivamente.

#### DPCM

A forma mais simples de compressão predictiva opera a nível de pixel com uma técnica chamada de PCM diferencial (DPCM). No DPCM, nós comparamos pixels adjacentes e apenas as diferenças são transmitidas. Como pixels adjacentes são geralmente similares, o valor da diferença tem uma alta probabilidade de ser menor que o valor do novo pixel e ele pode ser expresso com uma quantidade menor de bits. Na descompressão, a informação da diferença é usado para modificar o pixel anterior para obter o novo pixel.

#### ADPCM

Há muitas maneiras de se implementar a técnica DPCM adaptativa (ADPCM), a mais comum é variar o tamanho de passo representado pelos bits diferenças. Por exemplo, se um passo preto-para-branco for detectado, pode-se aumentar o passo de quantificação antes do passo preto-para-branco chegar.

Tanto DPCM e ADPCM não são muito utilizados para compressão de vídeo, eles causam uma sobrecarga proibitiva quando se tenta uma taxa de compressão superior a 2:1. De qualquer modo, estas técnicas são utilizadas em conjunto com outras técnicas mais poderosas em algoritmos de compressão mais complexos.

### 3.7.4 Preenchimento Condicional

Neste esquema, a imagem é segmentada em áreas estacionárias e com movimento e apenas são transmitidos dados de áreas com movimento. Um detector de movimento localiza diferenças inter-quadros significantes.

### 3.7.5 Estimativa e Compensação de Movimento

Este esquema explora redundância temporal em vídeos. Animação de imagens usualmente implica que pixels na imagem anterior estão em diferentes posições que na imagem atual. Nesta técnica, cada imagem é dividida em blocos de tamanho fixo. Um casamento para cada bloco é procurado na imagem anterior. O deslocamento entre estes dois blocos é chamado vetor de movimento. Uma diferença de blocos é obtida calculando diferenças pixel a pixel. O vetor de movimento e a diferença de bloco é codificada e transmitida. É normalmente muito menor e mais eficiente transmitir o vetor de movimento mais a diferença que transmitir a descrição do bloco atual.

### 3.7.6 Técnica Transform Coding

Aqui o termo *transform* é o processo que converte um bloco de dados em uma forma substituta que é mais conveniente para algum propósito particular. Ele é um processo reversível, o dado original pode ser restaurado. No caso de compressão de imagem e vídeos, um bloco de dados é um grupo de pixels (usualmente um vetor bidimensional de

pixels para uma imagem). Uma transformação é feita para criar uma forma substituta que pode ser transmitida ou armazenada. No tempo de descompressão, uma transformação inversa é executada sob o dado para reproduzir o pixel original.

Nesta técnica, uma imagem é dividida em blocos ou sub-imagens. Uma transformada matemática unitária é aplicada a cada sub-imagem que transforma a sub-imagem do domínio do espaço para o domínio da freqüência. Exemplos de transformadas aplicáveis são: transformada de Karhunen-Loeve (KLT), transformada discreta de co-seno (DCT), transformada Walsh-Hadamard (WHT) e transformada de Fourier Discreta (DFT). Como resultado, muito da energia da imagem é concentrada em poucas amostragens na área de baixa freqüência. Note que dados não são perdidos aplicando uma transformada na imagem. Se os dados no domínio espacial são altamente correlacionados, o dado resultante no domínio da freqüência estará em uma forma desejável para a redução de dados com técnicas de compactação sem perdas.

### 3.8

## Padrões de Compressão Multimídia

Existem várias técnicas e produtos para compressão de imagens e vídeos. A definição e a utilização de padrões internacionais de compressão de dados multimídia promove a compatibilidade entre diferentes produtos. Os cinco principais padrões de compressão audiovisual mais utilizados são:

- ISO JBIG para compressão sem perda de imagens bi-níveis (1 bit/pixel) para transmissão fac-símile
- ISO/IEC JPEG para compressão de imagens;
- ITU-TS H.261 para videofonia e aplicações de teleconferências na taxa de bits múltiplos de 64 Kbps;
- ITU-TS H.263 para aplicações de videofonia na taxa abaixo de 64 Kbps;
- ISO MPEG para compressão de vídeo e áudio associado;

### 3.8.1 Padrão de Compressão de Imagens JPEG

O padrão JPEG (*Joint Photographic Expert Group*) foi desenvolvido, em 1992, pelo ISO/IEC em colaboração com a ITU-TS. Ele representa uma das melhores tecnologias de compressão de imagem. Dependendo da imagem, taxas de compressão podem alcançar 25 para 1 sem degradações notáveis. Em média, JPEG obtém uma taxa de compressão de 15:1.

O JPEG é usado hoje em dia por muitas aplicações envolvendo imagens. Ele pode ser implementado em software e hardware. Embora este padrão ter sido projetado inicialmente para imagens, codificação e decodificação JPEG tempo-real tem sido implementada para vídeo. Esta aplicação é chamada de *Motion JPEG* (MJPEG).

Um dos objetivos do JPEG é cobrir uma grande faixa de qualidades de imagens e permitir especificar o comportamento do codificador a partir de parâmetros. Assim, a relação entre a taxa de compressão e a qualidade resultante pode ser selecionada pelo usuário ou pelo software aplicativo que usa JPEG. Outro objetivo foi permitir que a compressão fosse possível em uma grande diversidade de computadores com diferentes poder de processamento. Esta meta levou a definição de quatro modos de operação:

- **Codificação seqüencial:** modo com perdas baseada em DCT. Cada componente de imagem é codificado em uma única varredura da esquerda para direita e de cima para baixo. Este modo é chamado de *baseline* e deve ser suportado por toda implementação JPEG.
- **Codificação progressiva:** com perdas baseada em DCT expandido. Fornece avanços ao modo *baseline*. Uma expansão importante é a codificação progressiva (varreduras sucessivas), em que a imagem é codificada em varreduras múltiplas para produzir uma imagem de maneira rápida e rústica quando a largura de banda de transmissão é baixa.

- **Codificação sem perda:** o processo de compressão é reversível, assim a reprodução é exata. Este modo sem perda não permite a obtenção de altos fatores de compressão, mas muitas aplicações necessitam armazenamento sem perda de imagens, como fotografias de raio X.
- **Codificação hierárquica:** a imagem é codificada em resoluções múltiplas. Neste esquema nós temos uma codificação piramidal de uma imagem em resoluções espaciais múltiplas. Assim, versões podem ser acessadas sem a necessidade de primeiro descompactar a imagem na resolução completa.

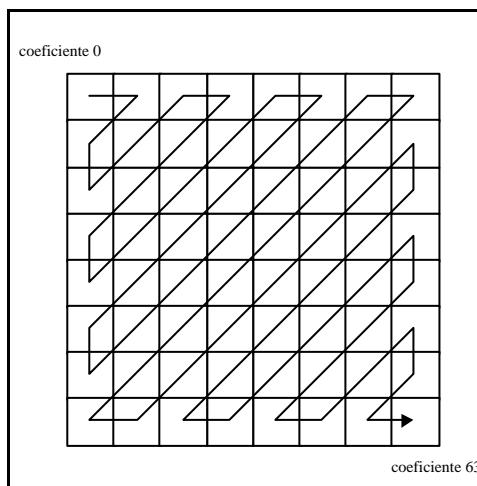
### Algoritmo de compressão JPEG

O algoritmo JPEG decompõe a imagem de entrada em blocos fonte de 8x8 pixels e então transforma estes blocos no domínio da freqüência usando a transformada discreta de co-seno (DCT). O DCT efetua uma separação dos componentes de baixa e de alta freqüência presentes numa imagem, permitindo que se faça uma seleção destas últimas, de acordo com a qualidade pretendida para a imagem comprimida.

O formato JPEG consegue alcançar boas taxas de compressão à custa da exploração das limitações da visão humana, a qual apresenta sensibilidades diferentes relativamente às componentes de freqüência presentes numa dada imagem. Como a visão humana é menos sensível às componentes de alta freqüência do que às de baixa freqüência, aquelas podem ser desprezadas sem que daí resultem grandes alterações no conteúdo dessa mesma imagem. O JPEG é parametrizável neste sentido, quanto maior é a compressão escolhida, menor é o número de componentes de alta freqüência desprezados.

O sinal discreto de 64 pontos (um para cada pixel) transformado é uma função de duas dimensões espaciais, x e y. Estes componentes são chamados coeficientes DCT ou freqüências espaciais. Para um bloco de imagem 8x8 típico, maior parte dos coeficientes DCT tem valores zero ou muito próximo de zero e não necessitam ser codificados. Isto é a base da compressão de dados. No próximo passo, todos os coeficientes DCT são quantificados usando valores de quantificação especificados em uma tabela de quantificação. Esta quantificação reduz a amplitude dos coeficientes que contribuem pouco ou nada para a qualidade da imagem, aumentando assim o número de coeficientes de valor zero.

Após a quantificação, os coeficientes DCT são ordenados em uma seqüência zig-zag, mostrada na Figura 18, para obter uma seqüência unidimensional de dados para ser usado na codificação por entropia. O propósito do escaneamento zig-zag é ordenar os coeficientes em uma ordem crescente de freqüências espectral: os coeficientes de alta freqüências (no canto direito inferior) tem valores mais próximos a zero que os coeficientes de baixa freqüência. Isto leva a uma maior eficiência da codificação por entropia.



**Figura 18.** Ordenamento zig-zag dos coeficientes DCT

Finalmente, o último passo do JPEG é a codificação por entropia. O padrão JPEG define dois métodos de codificação por entropia: codificação de Huffman e codificação aritmética. Esta etapa fornece uma compressão adicional. Codificação aritmética é normalmente 10% mais eficiente que a codificação de Huffman. Apenas a codificação Huffman é especificado no modo baseline. Para as demais, as duas técnicas podem ser usadas.

### Taxas de compressão obtidas

Como já referimos, quanto maior for a taxa de compressão maior será o número de componentes de alta freqüência desprezados. Para obtermos taxas de compressão muito elevadas, temos que deitar fora um número bastante significativo de componentes de alta freqüência, levando ao aparecimento do efeito de bloco, ou seja, perda de definição nos contornos das imagens. Geralmente obtém-se bons resultados com a utilização de taxas de compressão da ordem dos 10% a 50%. Comparando as taxas de compressão com a qualidade de imagem obtida, podemos afirmar que:

- Taxas de compressão de 10:1 a 20:1 – Alta qualidade de imagem;
- Taxas de compressão de 30:1 a 50:1 – Média qualidade de imagem;
- Taxas de compressão de 60:1 a 100:1 – Fraca qualidade de imagem.

### JPEG é para imagens fotográficas

Este formato apresenta ótimas taxas de compressão para imagens fotográficas naturais multi-tonais, diminuindo consideravelmente quando aplicado a imagens gráficas com contornos e áreas bem definidas de cor ou a imagens com texto, como é o caso dos logotipos. O JPEG introduz ruído nas zonas de imagem compostas por cores sólidas, o qual pode distorcer o aspecto geral da imagem. Comparado ao GIF, verifica-se que a imagem GIF comprime mais eficazmente que a JPEG e que apresenta uma melhor definição dos contornos do texto.

#### 3.8.2 CCITT H.261

Reconhecendo a necessidade de fornecer serviços de vídeo onipresentes na Rede Digital de Serviços Integrados (ISDN), o CCITT Study Group XV estabeleceu em 1984 um Grupo Especialista em Codificação para Telefonia Visual com o objetivo de recomendar uma codificação padrão de vídeo para transmissão a taxas, surgiu então a recomendação H.261.

H.261 é um dos padrões da família H.320 para videofonia e tele-conferência na taxa de 64 Kbps a 2 Mbps. Ele fornece poucos quadros por segundos com um resolução cerca de oito vezes mais baixa que a qualidade TV PAL/SECAM.

O padrão H.261, também chamado de px64, obtém grandes taxas de compressão para a transmissão de vídeo colorido tempo-real. O algoritmo combina codificação intraquadro e interquadro (redundância espacial e temporal) para fornecer um rápido processamento para compressão/descompressão tempo-real de vídeo, otimizado para aplicações tal como telecomunicações baseada em vídeo. Como estas aplicações usualmente não são a movimentos intensos, o algoritmo usa uma limitada estratégia de busca e estimativa de movimento para obter taxas de compressão mais altas. H.261 pode obter taxas de compressão de 100:1 a mais de 2000:1.

A recomendação H.261 define um padrão de codificação de vídeo para transmissão na taxa de  $p^*64\text{Kbps}$  ( $p=1,2,3,4\dots,30$ ) que cobre as capacidades do canal ISDN. Sendo que a aplicação alvo desta recomendação são a videofonia e a teleconferência, onde o algoritmo de compressão de vídeo opera em tempo-real com atraso mínimo. Para  $p = 1$  ou 2, devido a limitação de taxa de bits, apenas movimentos lentos, comunicação visual face-a-face (videofonia) são apropriados. Para  $p > 5$ , com uma maior taxa de bits disponível, imagens mais complexas podem ser transmitidas com melhor qualidade (videoconferência). Note que a máxima taxa de bits disponível é 1,92 Mbps ( $p=30$ ), que é suficiente para obter imagens de qualidade VHS [Raghavan, 98].

H.261 opera com dois formatos de imagem: CIF (*Common Intermediate Format*) e QCIF (quarter-CIF). CIF, de 320x288, permite usar um formato único dentro e entre regiões

usando padrões de TV de 625 e 525 linhas. QCIF, de tamanho 160x144, é mais útil em taxas de bit menores ( $p < 6$ ).

O algoritmo de codificação é um híbrido de predição inter-quadro, transform coding (DCT), similar ao JPEG, e compensação de movimento. A taxa de dados do algoritmo de codificação foi projetado para ser capaz de suportar 40 kbps e 2Mbps. A predição inter-quadro remove a redundância temporal. O transform coding remove a redundância espacial. Para remover redundâncias adicionais no bitstream a ser transmitido, uma codificação por entropia (normalmente codificação de Huffman) é utilizado para reduzir ainda mais o vídeo.

### 3.8.3 H.263

H.263 é um padrão de vídeo a baixa taxa de bits para aplicações de teleconferência que opera a taxas abaixo de 64 Kbps. A codificação de vídeo é uma extensão do H.261 e descreve um método de codificação DPCM/DCT.

Uma idéia interessante do H.263 é o quadro PB. Ele consiste de duas imagens codificadas em uma unidade. O nome PB é derivado da terminologia MPEG dos quadros P e B. Assim, um quadro PB consiste de um quadro P que é produzido a partir do último quadro P decodificado e um quadro B que é produzido a partir do último quadro P decodificado e do quadro P sendo decodificado.

H.263 suporta cinco resoluções. Além do QCIF e CIF que é suportado pelo H.261, existem o SQCIF, 4CIF e 16CIF. SQCIF é aproximadamente a metade da resolução do QCIF. 4CIF e 16CIF são aproximadamente 4 e 16 vezes a resolução do CIF. O suporte do 4CIF e 16CIF significa que o codec poderia então competir com outras codificações de mais altas taxas de bits como os padrões MPEG.

Testes atuais mostram que o H.263 tem um desempenho 1 a 2,5 melhor que o H.261. Isto significa que, dada uma qualidade de imagem, a taxa de bits H.261 é aproximadamente 2,4 vezes a gerada pelo H.263.

### 3.8.4 ISO/IEC MPEG (Motion Picture Expert Group)

O grupo da ISO/IEC MPEG foi estabelecido em 1988 para desenvolver padrões para representação codificada de vídeos, áudios associados, e suas combinações quando usados para armazenamento e recuperação em *Digital Storage Media* (DSM). O conceito DSM inclui os dispositivos de armazenamento convencionais, como CD-ROMs, drivers de fita, disco rígidos e canais de telecomunicação (ISDN e LAN).

MPEG usa a compressão interquadros (redundância temporal), obtendo taxas de compressão de até 200:1 pelo armazenamento apenas das diferenças entre quadros sucessivos. Especificações MPEG também incluem um algoritmo para compressão de áudio a taxas de 5:1 a 10:1.

#### **Grupos de Trabalho MPEG**

O grupo teve 3 itens de trabalho: codificação de vídeos e áudios associados sobre 1.5, 10 e 40 Mbps. Estes grupos são abreviados por MPEG-1, MPEG-2 e MPEG-3:

- A intenção do MPEG-1 (normalizado em 1993) é a codificação de vídeo com qualidade VHS: 360x280 pixels com 30 quadros por seg. na taxa de 1.5 Mbps (taxa dos drivers de CD-ROM da época). Quando se fala codificação MPEG é MPEG-1 que se está referenciando.
- MPEG-2 (normalizado em 1994) visa a codificação de vídeo com qualidade de televisão digital CCIR 601: 720x480 pixels com 30 quadros por seg. na taxa entre 2 a 10 Mbps.
- MPEG-3 visa a codificação de vídeo com qualidade HDTV na taxa de 40 Mbps. Estes trabalhos foram interrompidos em julho 1992.

Durante o processo de normalização, viu-se a necessidade da definição de codificação audiovisual a taxas de bits muito baixas, assim o surgiu o MPEG-4 (a ser aprovado em dezembro de 1999). A taxa de bits considerada aqui varia de 4,8 a 64 Kbps.

As atividades do MPEG cobre mais que a compressão de vídeo, desde que a compressão do áudio associado e a sincronização audiovisual não podem ser independente da compressão do vídeo.

### Partes do padrão MPEG

O padrão MPEG tem 3 partes principais:

- MPEG-Vídeo: trata da compressão de sinais de vídeo;
- MPEG-Áudio: trata da compressão de um sinal de áudio digital; e
- MPEG-Sistemas: trata da sincronização e multiplexação de bitstreams de áudio e vídeo compactados.

Além disso, uma quarta parte chamada Conformidade especifica o procedimento para determinar as características dos bitstreams codificados e para testar a conformância com os requisitos identificados no Áudio, Vídeo e Sistemas.

Anteriormente nesta apostila foi apresentada a compactação MPEG-1 Audio. Na seqüência será apresentada a compactação MPEG-1 Video.

### Compactação MPEG-1 Video

No vídeo, existem dois tipos de redundância: espacial e temporal. O MPEG-1 utiliza as duas. A redundância espacial pode ser explorada pela simples codificação em separado de cada quadro com o JPEG. Essa estratégia é utilizada às vezes, em especial quando há a necessidade de se acessar aleatoriamente cada quadro, como na edição de produções de vídeo.

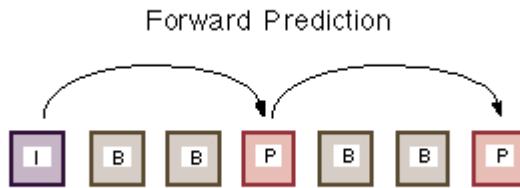
Uma compactação adicional pode ser obtida se nos beneficiamos do fato de que dois quadros consecutivos são, com freqüência, quase idêntico. Esta é a redundância temporal. Por exemplo, para cenas em que a câmera e o fundo da cena permanecem estáticos e um ou dois atores se movimentam lentamente, quase todos os pixels serão idênticos de um quadro para outro. Neste caso, subtrair cada quadro do seu antecessor e executar o JPEG na diferença seria suficiente. Entretanto, para cenas onde a câmera faz uma tomada panorâmica ou uma mudança rápida de plano, essa técnica é insuficiente. É necessária alguma forma de compensação de movimento. Isso é exatamente o que o MPEG faz; essa é, também, a diferença mais importante entre o MPEG e o JPEG.

A saída do MPEG-1 consiste de quatro tipos de quadros:

- Quadros I (Intracoded): imagens estáticas, independentes e codificadas com o JPEG.
- Quadro P (Predictive): Diferença bloco a bloco com o último quadro.
- Quadro B (Birectional): Diferença com o último quadro e com o quadro seguinte.
- Quadro D (DC-coded): Médias de bloco usadas para o avanço rápido (fast forward).

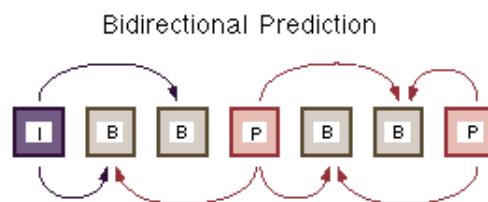
É necessário que quadros I apareçam periodicamente no fluxo de saída por três motivos: no caso de transmissão multicast, os receptores podem entrar no grupo em tempos distintos, requerendo um quadro I para começar a decodificação MPEG-1; se um quadro for recebido com erro, a decodificação não será mais possível; e sem quadros I, se houvesse um avanço rápido ou um retrocesso, o decodificador teria de calcular todos os quadros exibidos de modo a saber o valor total do quadro em que parou. Portanto, os quadros I são inseridos na saída uma ou duas vezes por segundo.

Os quadros P, ao contrário, codificam as diferenças entre os quadros. Elas se baseiam na idéia dos macroblocos, que cobre 16x16 pixels. Um macrobloco é codificado da seguinte forma: tentando-se localizá-lo, ou algo bem parecido com ele, no quadro anterior. Decodificar quadros P requer que o decodificador armazene o I ou P quadro anterior em um buffer e, depois, construa o novo quadro em um segundo buffer baseado em macroblocos completamente codificados e macroblocos contendo diferenças com o quadro anterior (Figura 19).



**Figura 19.** Decodificando quadros P

Quadros B são semelhantes aos quadros P, a diferença é que eles permitem que o macrobloco de referência esteja tanto no quadro I ou P anterior quanto no quadro seguinte (Figura 20). Essa liberdade adicional acarreta uma melhoria na compensação do movimento, e também é útil quando, no vídeo, objetos passam na frente ou por trás de outros objetos. Para realizar a codificação de quadros B, o codificador precisa manter três quadros decodificados na memória ao mesmo tempo: o quadro anterior, o atual e o próximo.



**Figura 20.** Decodificando quadros B

Os quadros D só são usados para possibilitar a apresentação de uma imagem de baixa resolução quando um avanço rápido, ou um retrocesso, estiver sendo realizado.

Uma seqüência de quadros codificados teria a seguinte forma: IBBPBBPBBPBBIBBPBBPBBPB.....

A codificação MPEG-2 é fundamentalmente semelhante à codificação MPEG-1, com quadros I, P e B. Os quadros D, entretanto, não são aceitos. Além disso, a transformação discreta de co-seno é de 10x10 em vez de 8x8, para proporcionar mais 50 por cento de coeficientes e, com isso, melhor qualidade.

### MPEG especifica apenas a sintaxe do bitstream

Uma das características mais importantes do MPEG é que ele especifica apenas a sintaxe dos bitstreams codificados para que decodificadores possam decodificar. Este padrão não especifica como gerar o bitstream. Esta escolha permite inovações no projeto e implementação de codificadores.

A fim de ilustrar a sintaxe dos bitstreams, a estrutura de um fluxo de vídeo MPEG-1 é apresentada na Figura 21. Ela é composta das seguintes camadas:

- Seqüência: composta de um cabeçalho de seqüência, seguido de um ou mais grupos de imagens e termina com um fim de seqüência.
- GOP (Grupo de imagens): fornece um ponto de acesso aleatório
- Camada de imagem: contém todas as informações codificadas de uma imagem. A este nível o cabeçalho contém a referência temporal de uma imagem, o tipo de codificação, etc.
- Camada Pedaços: as imagens são divididas em pedaços (*slices*), cada pedaço consiste de um número de macroblöcos de 16x16 pixels.
- Camada Macroblöco: um bloco é uma matriz 8x8 de valores de pixels tratados como unidades e entrada para o DCT.

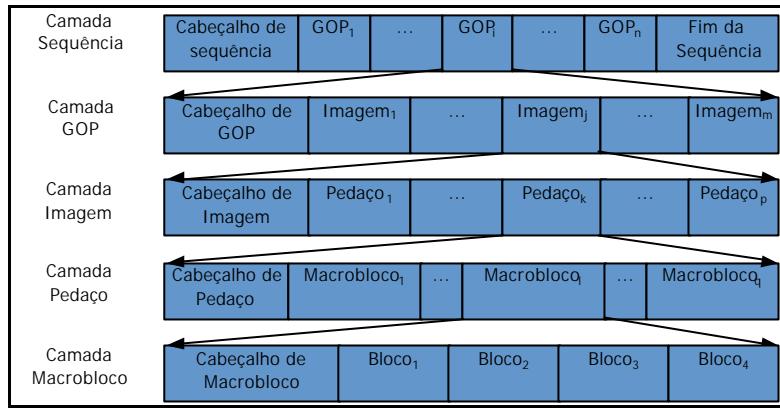


Figura 21. Estrutura do Fluxo de Vídeo MPEG-1

### MPEG-1 Sistemas

Enquanto MPEG-1 Vídeo e MPEG-1 Áudio especificam representação para áudios e vídeos, o MPEG-1 Sistemas define uma estrutura multiplexada para combinar fluxos elementares, incluindo áudio, vídeo, e outros fluxos de dados. Estes fluxos, chamados de fluxos MPEG, podem multiplexar até 32 fluxos de áudio MPEG, 16 fluxos de vídeo MPEG e 2 fluxos de dados de diferentes tipos. MPEG Sistemas também especifica o modo de representar as informações temporais necessárias para reprodução de seqüências sincronizadas em tempo real.

A especificação da codificação MPEG Sistemas fornece campos de dados para suportar as seguintes funções:

- sincronização de fluxos elementares,
- gerenciamento de buffer nos decodificadores,
- acesso aleatório,
- identificação do tempo absoluto do programa codificado.

### MPEG-4 [Vidal, 97]

No início do trabalho, o objetivo do MPEG-4 era a utilização em aplicações com baixas taxas de bits. Entretanto, MPEG adotou um plano de trabalho para as mudanças no ambiente áudio-visual e modificou suas finalidades consideravelmente.

Diferente da codificação linear de áudio e vídeo do MPEG-1/2, a codificação MPEG-4 é baseada em objetos, isto é, as cenas áudio-visuais são codificadas em termos de objetos. Um Objeto pode ser uma imagem ou um vídeo: um carro em movimento, uma fotografia de um cão. Também pode ser um objeto de áudio: um instrumento de uma orquestra, um latido de um cão.

Um novo conjunto de aplicações usarão MPEG-4, tais como vídeoconferência, comunicações móveis, acesso à vídeo de servidores remotos para aplicações multimídia, jogos, etc. Atualmente, o grupo MPEG-4 está direcionando os trabalhos para televisão digital, aplicações gráficas interativas e World Wide Web. As aplicações com baixas taxas de bits, as taxas serão de 5 a 64Kbits/s e para aplicações de TV, as taxas serão de 2Mbps.

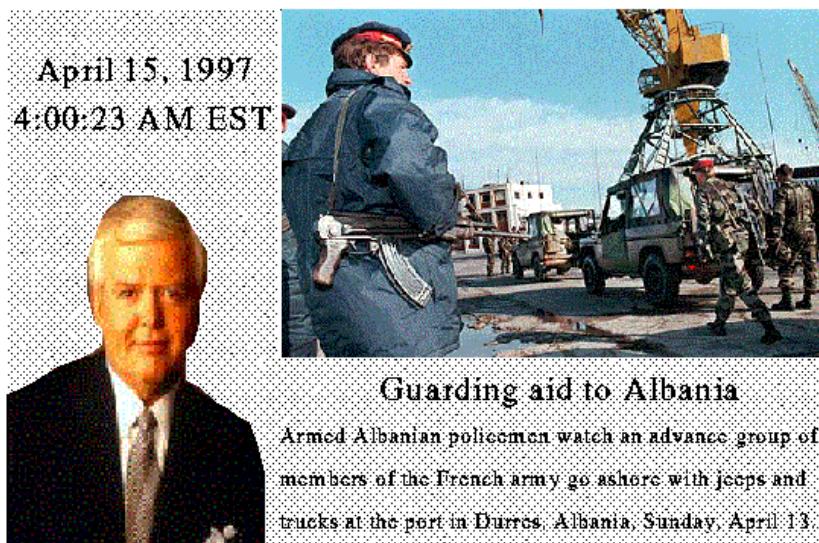
Os padrões correntes de compressão de vídeo transmitem um quadro inteiro de vídeo em um único fluxo de bits (bitstream). O MPEG-4 codificará objetos áudio-visuais em quadros separadamente. Os objetos serão compostos em um quadro no decodificador. Objetos codificados separadamente fornecem três benefícios:

- Reusabilidade - a abordagem orientada a objeto permite aos autores reusarem material áudio-visual mais rapidamente;
- Escalabilidade - objetos podem ser codificados usando diferentes resoluções espaciais e temporais. A largura de banda adicional pode ser alocada para objetos mais importantes. A resolução do objeto pode ser ajustada para casar com a capacidade do meio de transporte;

- Interatividade - por causa dos objetos áudio-visuais serem compostos em quadros no decodificador, o usuário pode controlar a saída.

A Figura 22 ilustra uma aplicação de difusão de notícias (*news broadcast*) [Huang, ?], que faz uso das três funcionalidades apresentadas acima. Os quatro objetos de vídeo incluídos na aplicação são: o vídeo da notícia, o vídeo do apresentador, o texto e um relógio. O usuário pode selecionar qual dos objetos serão usados para compor o quadro.

O espectador pode remover o vídeo do apresentador e usar somente o vídeo da notícia, ou pode desligar o áudio do apresentador e escutar o áudio da notícia enquanto lê o texto. O fluxo de bits codificado da notícia pode ser armazenado em uma biblioteca por outras organizações e reutilizado no futuro.



**Figura 22.** Aplicação MPEG-4

### 3.9 Implementando Algoritmos de Compressão

Na implementação de um algoritmo de compressão, a questão chave é como partitionar entre hardware e software afim de maximizar o desempenho e minimizar os custos [Furht, 94].

Nós podemos classificar as implementações de algoritmos de compressão em três categorias:

- Abordagem hardwared que maximiza o desempenho (por exemplo, C cube). Ela funciona bem, mas este tipo de hardware não é muito flexível. Por exemplo se uma placa MJPEG (Motion JPEG) for usada para codificação, decodificação e apresentação de vídeo em uma estação de trabalho, ela não podem decodificar e apresentar um vídeo MPEG, limitando o uso do sistema.
- Solução por software que enfatiza a flexibilidade com um processador de propósito geral. A implementação via software é flexível, mas ela é muito lenta para as CPUs atuais. Pesquisadores da UC Berkeley implementaram um codificador MPEG por software em C usando X Windows e analisaram seu desempenho em diferentes plataformas computacionais [Patel, 93]. O resultado mostrou que as estações RISC da época, tal como a HP750, podem codificar um vídeo 320x240 de 10 a 15 fps.
- Abordagem híbrida que usa processadores de vídeo especializados. Muitas implementações usam processadores de vídeo especializados e processadores DSPs (*Digital Signal Processors*) programáveis. Neste caso, o DSP pode ser programado para codificar e decodificar muitos tipos de fluxos, e muito mais rápido que a implementação por software apenas. Por outro lado, processadores RISC poderosos tem tornado satisfatórias as soluções por software apenas.

AT&T usa um abordagem híbrida para os codificadores AVP 4310E e o decodificador 4220D para os padrões H.261 e MPEG. O codificador aceita a entrada de vídeo a 30 fps e os dados de saída são gerados em uma taxa selecionada de 40 Kbytes/s a 4 Mbytes/s. O hardware implementa funções computacionalmente intensivas, tal como estimação de movimento e codificação de Huffman. O usuário pode programar parâmetros chave, tal como taxa de quadros, atraso, taxa de bits e resolução. Um processador RISC programável implementa funções menos estáveis.

## Capítulo 4 Requisitos de Hardware e Sistema Operacional para Sistemas Multimídia

Este capítulo apresenta os requisitos básicos de um terminal multimídia em termos de hardware e sistemas operacionais.

### 4.1 Introdução as Bases dos Sistemas Computacionais

Computadores são sistemas automatizados capazes de adquirir, processar, armazenar, e buscar informações. Neste seção são apresentados alguns conceitos de base de arquitetura de computadores.

#### 4.1.1 Unidade Central

A Unidade Central é o coração do computador. Ela é formada por duas partes: a Memória Central (MC), e a Unidade Central de Processamento (CPU). A MC contem o programa em execução e os resultados imediatos. Como programas e dados são representados na memória na forma binária, a capacidade de armazenamento da MC é expressa em MBytes (Megabytes). A memória central é composta em geral de vários integrados de memória, geralmente de tecnologia DRAM (*Dynamic Random Access Memory*), muito rápidas, geralmente com um tempo de acesso na ordem de 100 nanosegundos.

A CPU é o elemento responsável pela execução de programas. Seu desempenho é medido em termos de milhões de instruções que a CPU pode executar por segundo (MIPS). Mas nem todas as instruções são equivalentes.

CPU e MC se comunicam via barramento de alta velocidade. Por exemplo, em computadores Pentium, a taxa de bits é de 528 MBps (freqüência de 66 MHz e tamanho de palavra 64 bits).

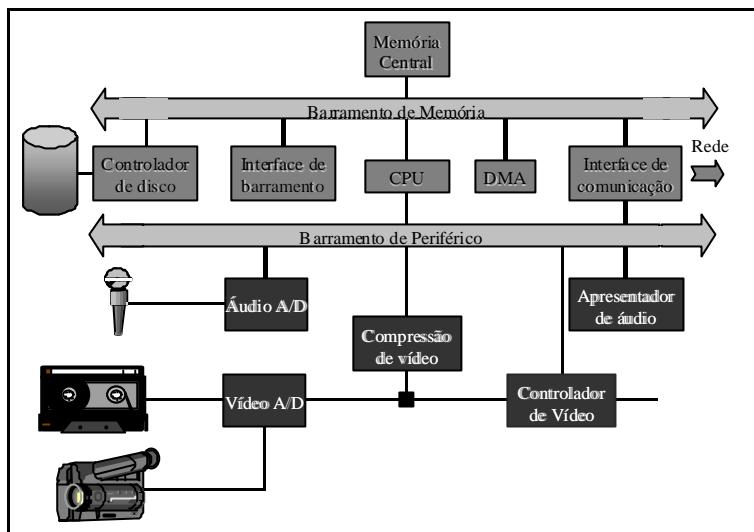
#### 4.1.2 Barramentos e Entrada/Saída

Periféricos devem se comunicar com a MC e a CPU. Para isto eles são conectados diretamente no barramento de memória (barramento do sistema) ou podem também ser conectados nos barramento de periférico (por exemplo, USB - *Universal Serial Bus*, Firewire/IEEE 1394) que por sua vez é conectado no barramento de memória.

Operações de entrada/saída de informações entre os periféricos e a MC são controladas pela CPU. Para descarregar a CPU desta tarefa, um mecanismo especial, chamado de DMA (*Direct Memory Access*), pode ser usado para realizar estas operações de entrada e saída.

### 4.2 Requisitos de hardware de terminais multimídia

Uma plataforma multimídia é um sistema capaz de suportar aplicações multimídia. De um ponto de vista simplificado, ela pode ser considerada como sendo um computador com dispositivos especiais de entrada e saída. Sua organização básica é mostrada na Figura 23.



**Figura 23.** Arquitetura simplificada de um computador multimídia

Os requisitos de hardware necessários a uma plataforma multimídia depende do tipo de utilização, por exemplo: hardware para apresentação de informações multimídia, para captura em tempo-real de informações multimídia, para autoria de documentos multimídia, para agir como servidor multimídia. Alguns requisitos de hardware para alguns tipos de utilização de plataformas multimídia são apresentados no próximo capítulo.

De modo geral, as aplicações multimídia impõem os seguintes requisitos na arquitetura de hardware:

- O hardware deveria ter um alto poder de processamento e uma alta taxa de transferência de dados (pois áudio e vídeo digital implicam numa grande quantidade de dados).
- Muitas aplicações exigem vários dispositivos de entrada e saída ao mesmo tempo. Assim um tipo de arquitetura paralela é preferida.
- A arquitetura de hardware deveria ser escalável para acomodar novos dispositivos de entrada e saída e aplicações.

#### 4.2.1 Dispositivos de Armazenamento Multimídia

Um dos requisitos dos terminais multimídia é a capacidade de armazenamento de informações multimídia. Os dispositivos de armazenamento devem ser baratos e hábeis de manter informações. Estes dispositivos têm tempo de acesso maior que a memória central. Os dispositivos de armazenamento mais convencionais utilizam um meio de registro magnético.

**Fitas magnéticas** registram seqüências de bits, mas 8 bits são registrados transversalmente. Cartuchos modernos têm capacidade de armazenamento de 200 a 800 MBytes. O acesso às fitas é seqüencial, isto é apropriado para apresentação de mídias contínuas individuais, mas o tempo de acesso à mídia é longo (tempo para posicionar a fita no início da mídia. Por exemplo, elas são usadas para armazenar som digital, o que é chamado de DATs (*Digital Audio Tapes*). A capacidade de armazenamento das fitas é um fator limitante para seu uso em sistemas multimídia.

**Discos magnéticos** registram informações digitais em trilhas concêntricas. Cada trilha é dividida em setores. Seu acesso semi-seqüencial: direto na trilha certa e seqüencial na trilha. O tempo de acesso é cerca de 50ms e a taxa de transferência é normalmente de 10 MBytes/s. A capacidade de armazenamento dos discos magnéticos normalmente variam de 80 MBytes a 8.4 GBytes.

**Discos óticos** não podem ser facilmente escritos por usuários, por isto eles são usados para distribuição de informações registradas, que é o caso dos CD-ROMs. CD-ROMs tem uma capacidade de 648 MBytes. Certos discos óticos podem ser escritos pelo usuário, tal como os discos WORM, que podem apenas ser escritos uma única vez. O

tempo de acesso é cerca de 300ms e a taxa de transferência é normalmente de 300 KBytes/s. Discos ópticos são apropriados para armazenamento multimídia, especialmente mídias contínuas.

#### 4.2.2 Sistemas de Captura e Apresentação Multimídia

Conforme a Figura 23, um computador multimídia tem muitos módulos funcionais similares aos computadores tradicionais, exceto que ele tem dois blocos funcionais adicionais:

- Bloco de entrada/saída de áudio: consistindo de um microfone, conversores A/D (chamados de digitalizadores de áudio) e D/A para captura e apresentação do som, dispositivos adicionais para compressão e descompressão e alto-falantes.
- Bloco de entrada/saída de vídeo: Dispositivos de captura de vídeo (câmera de vídeo), que digitaliza a entrada de vídeo analógico, dispositivos adicionais para compressão e descompressão e um monitor (sendo normalmente o monitor normal do computador).

Algumas funcionalidades dentro destes dois blocos podem ser implementados via hardware ou software.

Aplicações multimídia avançadas, como realidade virtual, necessitam de dispositivos de entrada e apresentação específicos.

#### 4.2.3 Multimídia em Rede

Um terminal multimídia pode atuar como um transmissor e um receptor de informações ao mesmo tempo, como na videofonia. Para a operação de transmissão, sob controle da CPU, o áudio e vídeo capturado é enviado para a interface de comunicação (de rede) para transmissão. Na recepção, a interface de rede passa eles para a CPU. Após certo processamento, a CPU passa para os dispositivos de saída de áudio e vídeo.

A compressão e a descompressão dos dados é realizadas nos terminais multimídia. Como apresentado no capítulo anterior, a compressão de dados é normalmente usado para reduzir o conjunto de dados a ser manipulado pelo terminal e pela rede. Mas a compressão e a descompressão são normalmente computacionalmente muito pesadas. A abordagem comum para a compressão e descompressão de áudio e vídeo é empregar um hardware específico. Ele funciona bem, mas este tipo de hardware não é muito flexível. Por exemplo se uma placa MJPEG (Motion JPEG) for usada para codagem, decodagem e apresentação de vídeo em uma estação de trabalho, ela não podem decodificar e apresentar um vídeo MPEG, limitando o uso do sistema. A implementação via software é flexível, mas ela é muito lenta para as CPUs atuais. Uma abordagem flexível é utilizar processadores de sinais digitais programáveis (DSP) para implementar compressões e descompressões de áudio e vídeo. Neste caso, o DSP pode ser programado para codificar e decodificar muitos tipos de fluxos, e muito mais rápido que a implementação por software apenas.

#### 4.2.4 Servidores Multimídia

Terminais multimídia podem interagir via uma rede de comunicação de maneira simétrica. Por exemplo, dois terminais providos de uma câmera, microfone, alto-falante para suportar videofonia bilateral. Neste caso os dois são fonte e destino de informações multimídia.

Várias aplicações multimídia são assimétricas: a tendência é compartilhar remotamente recursos em sistemas maiores ou especializados. Assim é possível uma economia de hardware e operação, onde potencialidades de armazenamento (discos e fitas magnéticos), softwares de aplicação (gestão de base de dados, CAD), poder de processamento podem ser compartilhados. Esta é a idéia do paradigma cliente-servidor.

No caso da multimídia, o servidor é um sistema computacional que oferece serviços multimídia para outros sistemas que agem como clientes multimídia. Servidores multimídia devem ter alta capacidade de armazenamento e funcionalidades de captura

de informações multimídia, e podem ser dedicados a um tipo de mídia específico (servidor de vídeo, de áudio,...). Abaixo são apresentados alguns exemplos de servidores multimídia:

- Servidor de Imagem: computador provido de uma alta capacidade de armazenamento (ótica ou magnética)
- Servidor de Vídeo: similar ao servidor de imagens, mas com um poder de processamento maior para permitir a saída múltipla de vários fluxos de vídeo.
- Servidor de Varredura: um ou um conjunto de estações de varredura (Scanners) compartilhadas.
- Servidor de TV: um *gateway* entre sinais de TV e a rede local.
- Servidor de Fax: um *gateway* entre um serviço de telefax público e a rede local.

#### 4.2.5 Apresentação de Áudio e Vídeo

Para a apresentação de um áudio ou vídeo armazenado em um disco, inicialmente ele deve ser transferido à MC (sob o controle do mecanismo DMA). Em seguida, no caso do fluxo de áudio, ele deve ser transferido para o conversor D/A para apresentação. No caso do fluxo de vídeo, as imagens devem ser transferidas para a memória de vídeo. Note que os dados de áudio e vídeo passam duas vezes pelo barramento de memória (e também pelo barramento de periférico). Neste caso, o elemento central que dita o desempenho do comportamento total é o barramento de memória (mais o barramento de periférico). E quando várias mídias contínuas estão envolvidas na aplicação, a largura de banda do barramento atinge seu limite, causando atraso. Para ter uma idéia, a largura de banda de memória é de 127,2 MBps para o PCI (33MHz e 32 bits) e 508,6 MBps para o 64bits PCI 2.1 (66MHz e 64 bits), e a do barramento de periférico é de 12 MBps para o USB e até de 400 MBps para o Firewire/IEEE 1394.

Pode-se pensar então porque não passar diretamente os dados do disco para o dispositivo de apresentação? Isto não é possível pois a transferência dos fluxos de áudio e vídeo para o conversor D/A e para a memória de vídeo deve ser feita sob o controle dos programas: a MC buferiza os fragmentos individuais de som e vídeo que são lidos do disco; esta buferização assegura que estes fragmentos sejam transferidos na cadência correta para apresentação, assegurando que as dependências temporais sejam respeitadas.

Além disso, é necessário notar que o barramento de memória (do sistema) é também compartilhado por todos os outros componentes do sistema, baixando assim a escalabilidade do sistema. Isto pois a capacidade do barramento limita o número de dispositivos e elementos de processamento conectados a ele.

#### Barramentos locais

Para solucionar estes problemas de limitação de barramento, algumas arquiteturas desenvolveram barramentos locais para conectar diretamente componentes de estações de trabalho, por exemplo um barramento conectando a interface de rede ao módulo de entrada/saída de vídeo e a memória principal. Isto reduz o gargalo do barramento do sistema, pois os dados são passados diretamente entre componentes da estação sem passar através do barramento compartilhado do sistema.

Infelizmente estes barramentos locais têm alguma limitação:

- O barramento do sistema (compartilhado) fornece um alto grau de flexibilidade e modularidade. A flexibilidade é perdida quando se utiliza barramentos locais, pois caminhos de comunicação dedicados não são flexíveis. Se por exemplo, um video codec (codificador/decodificador) libera um grande fluxo de dados diretamente para a saída gráfica, não existe a possibilidade de editar este fluxo, e a estação de trabalho é dedicada a apresentação do fluxo.
- Barramento locais são geralmente proprietários (não padronizados). Eles são soluções ad hoc para problemas específicos de aplicações específicas. Isto cria problemas de incompatibilidade de componentes de sistemas de diferentes

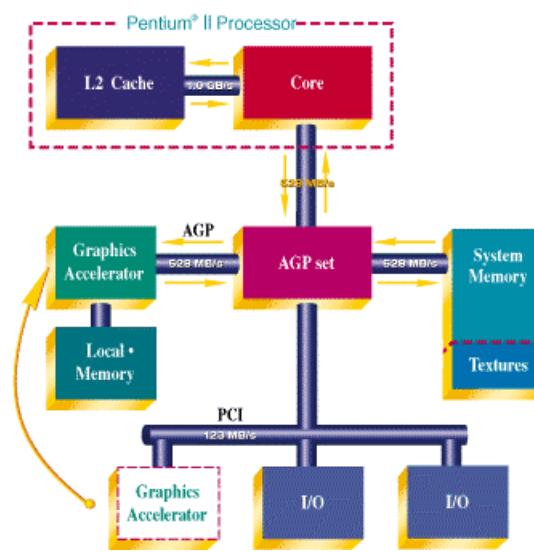
fabricantes, adicionando mais confusão à já confusa situação de muitos tipos de barramentos de sistema.

Um exemplo deste tipo de tecnologia é a interface AGP (*Accelerated Graphics Port*) desenvolvida pela Intel, projetada para aumentar as capacidades gráficas 3D e de vídeo do PC [Intel, 98]. A AGP introduz um canal ponto a ponto dedicado de modo que o controlador gráfico pode diretamente acessar a memória principal (Figura 24). O canal AGP é de 32 bits e roda externamente a 66 MHz.

AGP reduz a carga de vídeo no barramento PCI. Ele trabalha somente com 4 MB de memória de vídeo, emprestando a memória do sistema quando necessário. Esta tecnologia também acelera o mapeamento de textura<sup>3</sup>, que melhora a apresentação de jogos. Como resultado, esta tecnologia obtém uma apresentação gráfica mais suave e rápida. A diferença é notada principalmente em jogos 3D que necessitam que os jogadores se movimentem em um mundo virtual, como em Tomb Raider ou Quake.

Fisicamente separado do barramento PCI (Figura 24), AGP é um barramento de alta velocidade (528 MBps) entre o controlador gráfico e o chipset<sup>4</sup>. Em vez de realizar a pré-carga de texturas na memória gráfica local, a tecnologia AGP permite que o chip gráfico obtenha os mapas de texturas diretamente da memória do sistema. Isto permite que o chip gráfico processe dados de gráficos liberando o processador para realizar outras operações. Realizando o tráfego de gráficos no barramento AGP e fora do barramento PCI fornece às aplicações maior uso da largura de banda e menores atrasos. Isto libera o barramento PCI a usar adaptadores de rede a 100 Mbps, ultra DMA hard drives, e outros dispositivos PCI de alta velocidade. AGP também libera um slot PCI. Minimizando a necessidade de grande capacidade de memória gráfica local, o acesso direto do AGP à memória do sistema permite o uso de texturas mais ricas e maiores que aquelas que tipicamente muito grandes para a memória gráfica local. Armazenando dados de textura na memória principal do sistema fornece um maior realismo. Usuários podem então obter aplicações de 3D de alto desempenho a programas CAD para visualização sofisticada de dados e ferramentas de autoria. Inicialmente, AGP é disponível nos modos 1X (pico<sup>5</sup> de 264 MBps = 32bits\*66MHz/8) e 2X (pico de 528 MBps, operando a 66 MHz mas a sincronização é feita subida e na descida). A Intel está investigando futuros avanços para um AGP 4X (para 1 GBps em estações de trabalho de ponta).

A implementação AGP requer um chip gráfico compatível com AGP ou incorporado na placa mãe. Um novo conector foi projetado para a carga gráfica AGP. A Intel está incorporando suporte AGP em chipsets projetados especialmente para a família de processadores Pentium® II. Com a memória de sistema rápida para gráficos, o Intel 440LX AGPset fornece alto desempenho sem um aumento do custo para a video RAM.



**Figura 24. AGP**

<sup>3</sup> Em gráficos 3D, é a representação digital da superfície de um objeto.

<sup>4</sup> Conjunto de integrados que realizam as principais funções de uma placa mãe.

<sup>5</sup> De pico pois ele compartilha barramento de memória com o processador.

### Interfaces de rede multimídia dedicadas

Para solucionar as limitações dos barramentos locais, alguns pesquisadores propõem a construção de dispositivos de áudio e vídeo diretamente conectados na rede. A função da estação de trabalho normal é reduzida a controlar estes dispositivo. O princípio desta abordagem é similar ao uso de barramentos locais dedicados: deixar o dado multimídia bypassar o barramento de sistema da estação. Os dois projetos mais importantes adotando esta abordagem são o Pandora [Hopper, 90] da Olivetti Research e a Interface de Rede Multimídia (MNI) da Universidade de Lancaster [Blair, 93].

## 4.3 Requisitos de Sistema Operacional

Aplicações multimídia interagem diretamente com o sistema operacional. Um sistema operacional multimídia deveria satisfazer os seguintes requisitos:

- Deveria utilizar os recursos de hardware de maneira eficiente de modo que o uso destes recursos seja maximizado.
- Requisitos de Qualidade de Serviço (visto no capítulo II e detalhado mais adiante neste curso) deveriam ser garantidos usando gerenciamento de recursos e escalonamento de processos apropriados. Um dos maiores requisitos de QoS a nível de sistema operacional é o tempo de processamento garantido para cada tarefa.
- É preferível que o sistema operacional possa executar tanto aplicações multimídia quanto aplicações tradicionais. Isto tem duas implicações: a API (*Application Programming Interface*) tradicional deveria ser mantida; as aplicações tradicionais não deveriam ser privadas de recursos quando requisitos de QoS de aplicações multimídia são garantidos.

## Capítulo 5 Aplicações Multimídia

Como foi apresentado no primeiro capítulo desta apostila, a multimídia tem aplicações em diversas áreas. Para simplificar o estudo da problemática envolvida na computação e comunicação multimídia é necessário inicialmente classificar as diversas aplicações multimídia. Assim nós particionamos o mundo complexo da multimídia em grupos menores e analisaremos cada grupo. A ciência que trata das regras de classificação é chamada de Taxonomia.

Neste capítulo é apresentada inicialmente uma taxonomia de aplicações multimídia distribuídas proposta por [Fluckiger, 95] e outra recomendada pela ITU-TS. Em seguida, são apresentados exemplos de aplicações multimídia, onde são aplicadas a classificação proposta por [Fluckiger, 95] e também são apresentados os principais requisitos principalmente associados ao sistema de comunicação.

### 5.1 Aplicações Pessoa-a-Pessoa e Pessoa-a-Sistema

[Fluckiger, 95] classifica as aplicações multimídia em rede em duas grandes classes:

- **Aplicações Pessoa-a-Sistema** (ou pessoa-a-servidores de informações): onde indivíduos ou grupo de pessoas comunicam com sistemas remotos para acessar, receber, ou interagir com informações multimídia.
- **Aplicações Pessoa-a-Pessoa**: onde o objetivo principal é aumentar a comunicação entre humanos. O sujeito de comunicação varia de relações sociais privadas (p.e. em videofonia) até a comunicação de grupos em trabalho cooperativo. As aplicações pessoa-a-pessoa ainda podem ser subdivididas de várias maneiras, três critérios de subdivisão são:
  - aplicações privadas versus aplicações profissionais;
  - interpessoais (duas pessoas) versus aplicações orientada-a-grupos;
  - aplicações tempo-real ou síncronas versus aplicações assíncronas.

#### 5.1.1 Aplicações Multimídia Pessoa a Pessoa

As aplicações multimídia pessoa a pessoa podem ser subdivididas em aplicações síncronas e assíncronas.

##### **Aplicações síncronas (tempo-real)**

Aplicações multimídia síncronas ou tempo-real são aquelas onde a informação gerada em tempo-real pelo produtor ou que está armazenada no produtor, é transmitida por um produtor a um receptor que apresenta a informação na medida que ela está disponível. As aplicações nesta categoria incluem:

- **Aplicações Interpessoais (pessoa-a-pessoa)**: onde apenas dois indivíduos estão envolvidos, como na videofonia.
- **Aplicações de Distribuição (pessoa-para-grupo)**: onde informações multimídia, como áudio e vídeo, são transmitidos ao vivo de uma fonte para vários destinos (sem canal de retorno dos destinos para a fonte).
- **Teleconferência de Grupo (grupo-a-grupo)**: é o termo genérico para comunicação conversacional bidirecional entre dois ou mais grupos de pessoas.

##### **Aplicações Assíncronas**

As aplicações multimídia assíncronas são aquelas onde o instante de leitura da informação multimídia é diferente do tempo que esta informação foi transmitida pelo produtor (após é claro um atraso de transmissão). O momento da apresentação da informação é definida pelo usuário final. Nesta classe de aplicação multimídia nós temos:

- **E-Mail Multimídia:** semelhante a aplicação convencional de mensagem eletrônica, mas o documento trocado contém além de textos não formatados, textos formatados, áudio, vídeo, etc.
- **Conferência Multimídia Assíncrona:** são aplicações onde pessoas mantêm e seguem conversações assincronamente através de lista de tópicos e a mensagem trocada é multimídia. A técnica consiste em submeter ou buscar contribuições eletrônicas de um sistema computacional centralizado.

### 5.1.2 Aplicações Multimídia Pessoa-a-Sistema

O objetivo das aplicações pessoa-a-sistema é aumentar ou permitir modos inovadores de comunicação entre pessoas e fontes de informação (ou servidores de informação). Estas aplicações podem ser subdivididas em duas grandes categorias, dependendo do tipo de acesso ao servidor: aplicações interativas e de distribuição.

#### **Aplicações Interativas**

Neste caso, o iniciador da comunicação é o usuário final, que a qualquer momento pode pedir a informação a um servidor de informações. Existem dois tipos de aplicações interativas:

- **Aplicações orientada a busca (ou a recuperação):** neste tipo de aplicação o objetivo é localizar, acessar e apresentar informações multimídia.
- **Aplicações orientada transação:** onde a entrada do usuário é processada para outros propósitos que controlar a apresentação da informação.

Uma vez estabelecida a seção, a apresentação pode ser linear, como na vídeo sob-demanda ou interativa tal como em títulos multimídia.

#### **Aplicações de Distribuição**

Neste caso a distribuição da informação multimídia é iniciada pelo servidor. As aplicações de distribuição são de dois tipos:

- **Distribuição para Grupos Fechados:** onde somente usuários selecionados (grupos profissionais ou pagantes) recebem a informação.
- **Distribuição para Grupos Abertos:** onde nenhuma restrição existe. A única condição para receptores potenciais se juntarem ao grupo aberto é estar conectados a rede de distribuição e ter softwares e equipamentos apropriados.

## 5.2 Taxonomia Recomendada pela ITU

A ITU (*International Telecommunications Union*) propõe em sua recomendação I.211 uma taxonomia de serviços (aplicações) multimídia:

- **Serviços de conversação:** implica na interação entre um humano e outro humano ou sistema. Esta classe inclui serviços pessoa-a-pessoa tal como na videofonia e também serviços pessoa-a-sistemas, como telesegurança e telecompras.
- **Serviços de mensagem:** cobrem trocas de mensagem multimídia não tempo-real ou assíncronas entre mailboxes.
- **Serviços de recuperação (busca):** cobrem todos os tipos de acesso a servidores de informações multimídia. Tipicamente, o usuário envia um pedido para o servidor e a informação pedida é liberada para o usuário em tempo-real. Vídeo sob demanda é um exemplo deste tipo de serviço.
- **Serviços de distribuição:** cobrem serviços onde a informação é distribuída sob a iniciativa de um servidor. Um exemplo deste tipo de serviço é a transmissão de programas de TV. Esta classe ainda é subdividida em:
  - Serviços de Distribuição sem controle de apresentação pelo usuário, que é caso dos programas de TV, distribuição de jornais eletrônicos;
  - Serviços de Distribuição com controle de apresentação pelo usuário, que é o caso do tele-ensino e tele-anúncio.

## 5.3 Telefonia Assistida por Computador

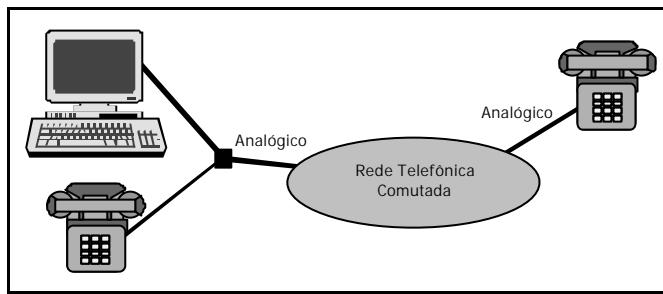
A telefonia assistida por computador trata-se de aplicações multimídia pessoa-a-pessoa, onde a informação multimídia transferida são áudios. Telefonia assistida por computador é o termo usado para descrever o uso de computadores na telefonia em geral. A idéia é auxiliar ou substituir o telefone convencional usando um computador pessoal.

### 5.3.1 Telefonia a Circuito Assistida por Computador

Neste caso, a rede usada é a rede de telefonia convencional, uma rede NISDN ou outros serviços a comutação de circuitos regional.

#### **Assistência a funções de chamada**

O computador pode agir como um auxílio, um complemento ao aparelho de telefone para algumas funções como discagem. Mas chamadas que chegam são recebidas pelo aparelho telefônico, o microfone e o auto-falante do computador podem ser usadas após a comunicação ser estabelecida. A Figura 25 ilustra este tipo de aplicação, nela considera-se que o sinal, disponível ao usuário final da rede telefônica, é analógico. Este é o caso, por exemplo, das companhias telefônicas brasileiras: embora a rede de transmissão possa ser digital, o ponto de acesso disponíveis aos usuários é analógico.

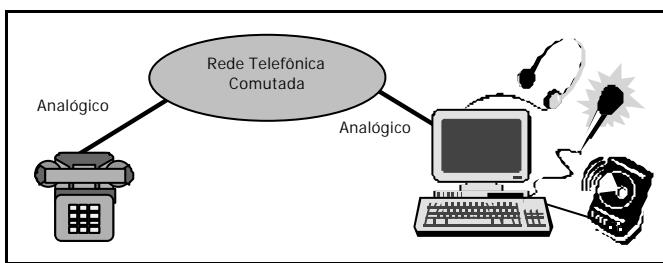


**Figura 25.** Assistente a funções de Chamada

Usando o computador, o usuário deste sistema pode ter um repertório das chamadas e números chamados: o computador funciona como um dispositivo facilitador de discagem, a partir da seleção de um número ou nome de uma lista. Além disso, o computador pode identificar a pessoa que está chamando, controlar os gastos com ligações telefônicas e gerar relatórios estatísticos sobre chamadas.

#### **Substituição do Aparelho Telefônico**

Neste caso, o computador repõe totalmente o aparelho telefônico, como ilustrado na Figura 26. Para receber chamadas, o computador deve estar permanentemente ligado. Além deste inconveniente, o computador deve ser provido de um sistema de cancelamento de ecos implementados em hardware ou software.



**Figura 26.** Substituição do aparelho telefônico

Neste esquema, como no caso anterior, o computador pode fornecer funções de repertório, estatísticas e de controle da conta telefônica. Além disso, outras funcionalidades são possíveis:

- Registro digital de chamadas podem ser realizadas, especialmente quando o usuário está ausente. Esta funcionalidade é similar às secretárias eletrônicas.

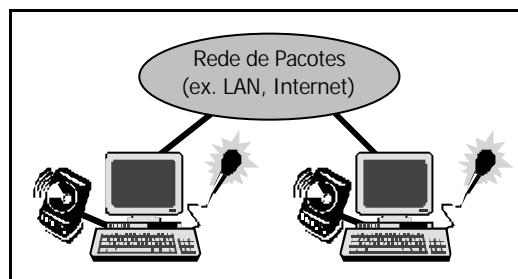
- Call-back direto, isto é análogo as funções de reply do correio eletrônico: após a apresentação de uma chamada registrada, um simples comando permite gerar uma chamada de retorno.
- Registro digital de mensagens enviadas, permitindo relembrar se uma dada chamada foi feita, a que momento, e o que exatamente foi dito.
- Chamadas automatizadas, permitindo arquivar uma mensagem e programar o momento que ela será levada ao destino (chamadas programadas). Caso o usuário não estiver no momento da chamada, o sistema pode automaticamente realizar várias chamadas em intervalos de tempo (chamadas repetitivas) até que a mensagem seja recebida. A mensagem pode ter múltiplos destinos, sendo o sistema responsável pela realização das várias chamadas e a apresentação da mensagem (chamadas múltiplas).

### Produtos

Muitas estações de trabalho e computadores pessoais são providos de hardware para funções de assistência ao telefone ou funções para substituição completa do telefone, como por exemplo placas FAX/Modem com voz. Uma dificuldade, muitas vezes não sanada, é o cancelamento de eco.

#### 5.3.2 Conversação por Pacotes de Voz

Esta classe de aplicação multimídia apresenta similaridades, em termos de funções, com a telefonia a circuito assistida por computador. Mas há uma grande diferença: a conversação a pacotes aproveita a infra-estrutura computacional de uma organização para servir também à comunicação entre o pessoal (Figura 27). Para isto, o computador deve ser provido de uma placa de som, um microfone e auto-falantes.



**Figura 27.** Telefonia a Pacotes de Voz

Do ponto de vista do usuário, o sistema comporta-se como se fosse telefonia assistida por computador. O usuário acessa um diretório de número freqüentemente chamados. A diferença é que a lista não contém números telefônicos, mas endereços de rede. Outra diferença é que um interlocutor deve sempre falar e em seguida aguardar que outro fale, não há conversação simultânea.

A conversação por pacotes de voz não pode ser considerada como um serviço universal como a telefonia convencional. Não há padrões de pacotes de voz como na telefonia convencional, além de existirem diversas tecnologias de rede e de endereçamento.

Em teoria, qualquer tecnologia suportando comunicação de pacotes de dados poderiam suportar pacotes de voz. Desde as redes sem conexão com meio compartilhado, como Ethernet, FDDI, Token Ring, DQDB, até as redes com conexão do tipo X.25, ATM, Frame Relay.

Para este tipo de aplicação é necessário que a rede suporte 64 Kbits/s, podendo esta taxa ser menor quando técnicas de compressão são utilizadas (requisitando por exemplo, 16 Kbits/s). Além disso, o áudio e especialmente a voz não é tolerante a perdas de pacote e a liberação fora de ordem.

Inicialmente a telefonia a pacotes ficou confinada às LANs, devido principalmente a taxa necessária, além do que não há padrões existentes. O uso em LAN facilita o acordo para um produto comum, de domínio público ou proprietário.

## Produtos

Existem vários softwares dedicados a telefonia a pacotes, como o produto comercial Internet Phone da VocalTec Inc, permitindo a comunicação pessoa-a-pessoa. Outras implementações permitem explorar a capacidade das redes a pacotes de endereçar múltiplos usuários. Em outras palavras, eles são ferramentas de conferência por pacotes de voz, projetados para trabalhar sobre Internet.

Existem vários softwares de domínio público que permitem a transferência de pacotes de voz sobre a Internet, uma lista não exaustiva inclui: vat - video-audio tool (<http://mice.ed.ac.uk/mice/archive/vat.html>), NEVOT - Network Voice Terminal (<http://www.fokus.gmd.de/step/nevot/>), IVS - INRIA Videoconferencing (<http://www.inria.fr/rodeo/ivs.html>), Free Phone (<http://www.inria.fr/rodeo/fphone/>), NVAT - Network Video Audio Tool (<http://www1.meshnet.or.jp/~mms-eizo/nvat/index.html>), RAT - Robust-Audio Tool ([http://www.cs.ucl.ac.uk/staff/vhardman/rat\\_project/](http://www.cs.ucl.ac.uk/staff/vhardman/rat_project/)) e Speak Freely (<http://www.fourmilab.ch/speakfree/unix/sfunix.html>). Netscape Communicator também dispõem de uma ferramenta de áudio-conferência, o Netscape Conference.

Uma outra consequência inevitável desta classe de aplicação é a integração total entre os assinantes das redes telefônicas e os utilizadores da Internet, o que permite efetuar chamadas de um PC para um telefone e vice-versa. Atualmente já é possível fazê-lo, graças gateways entre a Internet e a rede de telefonia. Net2Phone ([www.net2phone.com](http://www.net2phone.com)) e Internet Phone da Vocaltec ([www.vocaltec.com](http://www.vocaltec.com)) são aplicativos que fornecem uma interface com a rede de telefonia, permitindo desta forma efetuar chamadas de um PC para um telefone normal (e vice-versa para Internet Phone).

## 5.4 Videofonia

Esta seção trata da videofonia, que é a inclusão de vídeo na telefonia, também chamada de vídeo telefonia. Como no caso da telefonia, esta aplicação multimídia pertence a classe de aplicação interpessoal, projetada especialmente para aumentar a capacidade de comunicação entre dois participantes: o vídeo permite aumentar a transferência de informação emocional, como expressividade, via contato olho-a-olho.

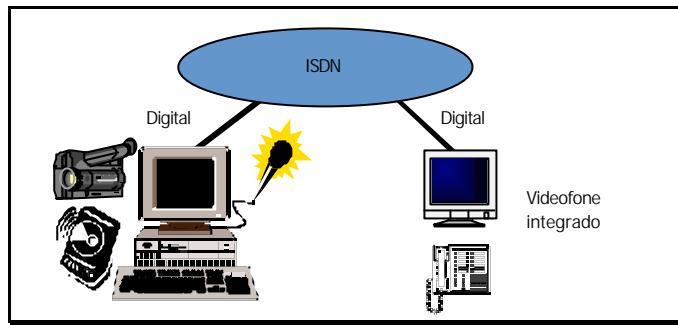
Neste tipo de aplicação, como nenhum detalhe necessita ser precisamente apresentado, a resolução da imagem pode ser média ou mesmo baixa. O movimento é mais importante que a resolução, isto pois o movimento aos solavancos impede a impressão de proximidade. Uma taxa de quadros de 6 a 12 fps é aceitável.

A videofonia pode ser usada para dar suporte ao trabalho cooperativo em conjunto com outras ferramentas de teleconferência que serão vistas mais adiante. Assim, pessoas podem se comunicar e apresentar o sujeito do trabalho cooperativo.

### 5.4.1 Videofonia a Circuito Integrada por Computador

Este tipo de videofonia pode ser denominada convencional, pois ela é baseada no uso de redes de telefonia de comutação de circuitos. Para a transmissão de vídeo compactado na qualidade desejada aqui, é necessário uma taxa de bits de 64 a 384 Kbits/s. Tal taxa pode ser obtida por exemplo usando redes NHSDN. A Figura 28 ilustra este tipo de aplicação multimídia.

Videofones integrados por computador, onde o computador suporta a videofonia e aplicações gerais, são disponíveis hoje em dia. A dificuldade está na interoperabilidade entre equipamentos de diferentes fabricantes. Um padrão ITU para codificação para telefonia visual e compressão, o H.320, e mais especialmente o seu componente de codificação de vídeo, H.261, tem grande aceitação, embora muitos padrões proprietários são utilizados.



**Figura 28.** Videofonia a Circuito sobre redes digitais

### Produtos

Já existem videofones a venda no mercado, como o ViaTV Phone VC 105 (<http://www.cti.com.br/video/vc105.html>), que transforma qualquer televisor e telefone de tom em um videofone. Ele usa as linhas de telefone padrão.

#### 5.4.2 Videofonia sob redes a pacote

A idéia é similar videofonia a circuitos, mas utilizando uma rede de comunicação de dados a pacotes. Quando implementada em estação de trabalho, ela é geralmente associada a outras ferramentas de teleconferência, e é na prática melhor adaptada para projeto cooperativo que produtos de telefonia a circuitos.

Em termos de taxa de bits, o canal de vídeo requer 80 a 200 Kbits/s mais 13 a 64 Kbits/s para o canal de áudio. Estes valores dependem da codificação e esquemas de compactação. Muitas LANs suportam esta carga, mas muito poucas comunicações podem estar ativa ao mesmo tempo. Esta é a grande barreira do uso extensivo de videofonia em organizações quando redes locais operam a 10 Mbits/s (no caso de redes Ethernet).

Hoje muitos vendedores de estações UNIX oferecem seus sistemas de videoconferência que operam sobre o IP. Embora eles são inicialmente projetados como ferramentas de comunicação geral entre dois ou mais grupos de pessoas, eles são mais adaptados a comunicação pessoa-a-pessoa. Infelizmente, muitos destes produtos não podem interoperar.

Um exemplo de produto comercial de videofonia é o Microsoft NetMeeting, fornecido juntamente com o Windows98. Existem muitos softwares de domínio público na comunidade da Internet em geral para audio-videoconferência, como vat - video-audio tool (<http://mice.ed.ac.uk/mice/archive/vat.html>), IVS - INRIA Videoconferencing (<http://www.inria.fr/rodeo/ivs.html>), NVAT - Network Video Audio Tool (<http://www1.meshnet.or.jp/~mms-eizo/nvat/index.html>).

## 5.5 Espaço Compartilhado para Trabalho Cooperativo Integrado por Computador (CSCW)

Nesta seção, nós veremos um novo tipo de aplicação pessoa-a-pessoa com funcionalidades que são geralmente usadas no modo um-a-um, mas cuja implementação também permite mais que duas pessoas participarem: o espaço de trabalho compartilhado. Espaço de trabalho compartilhado refere-se ao compartilhamento do monitor de vídeo do computador por todos os participantes envolvidos em uma tarefa comum e que colaboraram sem deixar seus espaços de trabalho.

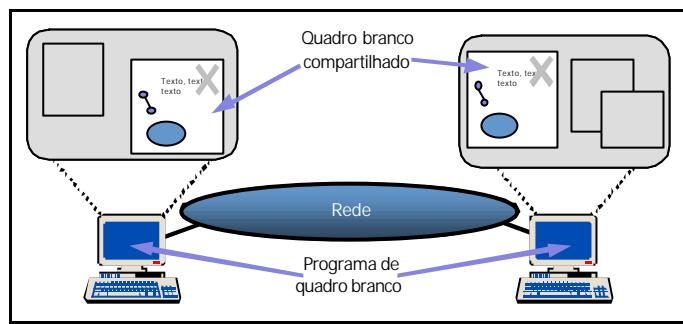
As várias aplicações de espaço de trabalho compartilhado apresentadas aqui são dedicadas a ambientes profissionais CSCW (*Computer-Supported Cooperative Work*). CSCW é o campo interessado no projeto de sistemas baseados em computador para suportar e aumentar o trabalho de grupos de usuários engajados em tarefas ou objetivos comuns, e o entendimento dos efeitos de usar tais sistemas.

Com relação a taxonomia apresentada, aplicações de espaço de trabalho compartilhado são do tipo pessoa-a-pessoa, para comunicações interpessoais ou de grupo. Ela opera no modo síncrono. Seu interesse é CSCW. O sistema, os produtos baseados em computador, suportando espaço compartilhado é chamado de **Groupware**.

A idéia de compartilhar espaço de trabalho é criar um compartilhamento a distância mediado por computador de informações efêmeras tão bem quanto compartilhar o produto real sobre a qual os indivíduos participantes estão trabalhando. Por informações efêmeras entende-se aquelas informações que dão suporte a apresentação de idéias, tipo rabiscar algo em uma folha em branco para demonstrar cálculos, ou um quadro branco para esboçar um conceito.

### 5.5.1 Ferramentas de Quadro Branco Compartilhado

As ferramentas de quadro branco compartilhado (Figura 29) emulam, na tela do computador, o quadro branco físico. Com esta ferramenta cada participante pode desenhar no quadro branco usando ferramentas de desenho para criar objetos geométricos e desenho a mão, ou digitar textos usando editores rudimentares. Tanto textos como objetos podem ser apagados, movidos, alterados os tamanhos, etc. Além disso, o quadro pode ser armazenado para uso posterior. Convenções simples, tipo uso de um código de cor para cada participante, podem identificar quem fez o que. Muitas ferramentas de quadro branco permitem que ao invés de se ter um fundo padrão branco, seja compartilhado uma imagem como fundo.



**Figura 29.** Princípios de quadro branco compartilhado

A Figura 30 apresenta o quadro branco do software Microsoft NetMeeting, chamado de quadro de comunicação.



**Figura 30.** Quadro de Comunicação do Microsoft NetMeeting

### Controle de Acesso (*Floor Control*)

Quadro branco compartilhado emula o quadro branco físico: quando duas pessoas num escritório estão trabalhando no mesmo projeto em um quadro branco físico, regras sociais e protocolos sociais governam o acesso ao quadro, assim eles não escrevem ao mesmo tempo, nem sobrescrevem ou apagam o que o outro fez. Tais regras de comportamento são difíceis de se produzir remotamente. A definição e implementação de políticas que regulam o acesso aos objetos compartilhados é um problema importante no espaço de trabalho compartilhado. Elas são chamadas políticas de controle de acesso.

Existem quatro abordagens básicas para controle de acesso:

- **Sem controle.** Neste caso o sistema deixa todo mundo acessar livremente a superfície compartilhada. Este esquema trabalha razoavelmente bem com duas pessoas, mas é impraticável quando o número de participantes aumenta.
- **Bloqueio implícito.** Cada vez que um participante entra com uma informação, este participante implicitamente toma o controle. O acesso ao quadro é automaticamente liberado num certo tempo (poucos segundos) após que a pessoa com o controle acabar sua entrada.
- **Bloqueio explícito.** É similar ao anterior, exceto que o usuário deve pedir e liberar explicitamente o acesso ao quadro via um botão.
- **Controle do moderador.** Um dos participantes é designado como moderador para a seção colaborativa. O moderador pode tomar o controle do quadro a qualquer instante. O moderador necessita de ferramentas para monitorar a lista de pedidos pendentes.

### Requisitos de Rede

As ferramentas de quadro branco compartilhado não exigem muito da rede. A taxa de transferência necessária é pequena, exceto quando imagens de fundo necessitam ser enviadas com freqüência que em prática é uma distorção do espírito do quadro branco compartilhado. Quando imagens não são transmitidas, os serviços de quadro branco compartilhado são mais sensíveis a atrasos do que a taxa de bits. O atraso máximo de transito deve estar na faixa de meio a um segundo. Como este tipo de aplicação exige multicast, se uma rede a comutação de circuitos tal como ISDN for utilizada, múltiplas conexões devem ser realizadas entre os participantes. Caso a rede use o *Internet Protocol* (IP), a funcionalidade de difusão (*multicasting*) pode ser utilizada.

### Usando quadro branco compartilhado com outras ferramentas de conversação

Quando se usa quadro brancos compartilhados, os participantes necessitam de canais diretos de comunicação. Isto ajuda a criar uma presença social útil no controle do quadro branco. Mas mais importante, este canal direto de comunicação é necessário em muitos casos para complementar o trabalho cooperativo feito com o quadro branco. Estes canais podem ser chamadas telefônicas ordinárias, troca de mensagens textuais, telefonia assistida por computador, ou videofonia.

### Produtos existentes

Existem muitas implementações de quadro branco compartilhado, incluindo o CoDraft do *IBM European Networking Center*, o JVTOS da *European Commission RACE II Initiative*. Produtos ou implementações são stand-alone ou empacotadas com ferramentas de áudio e vídeo. Exemplo de implementações stand-alone incluem WSCRAWL e wb (domínio público), e VENUE (DEC). Alguns produtos com ferramentas de áudio e vídeo são: InPerson (SGI), ShowMe (SUN), Person-to-Person (IBM), Communique! (InSoft). Netscape Communicator, Microsoft NetMeeting também dispõem de uma ferramenta de quadro branco.

#### 5.5.2 Ferramentas de aplicação compartilhada

As ferramentas de aplicação compartilhada, também chamadas de agentes de gestão de colaboração, permitem que múltiplos participantes compartilhem a apresentação e o

controle de qualquer aplicação interativa ordinária, por exemplo, um editor de texto ou gráfico. Se a aplicação é um editor de texto, qualquer participante pode rolar o texto apresentado ou entrar caracteres. Isto é chamado de edição colaborativa. Há duas idéias importantes nesta definição: a aplicação que é compartilhada não necessita ser projetada para suportar múltiplos usuários simultâneos e a ferramenta de compartilhamento torna possível o modo multi-usuário; além do compartilhamento da apresentação, há o seu controle.

Existem várias aplicações para este tipo de ferramenta, entre elas nós temos: revisão simultânea por muitos colaboradores, controle compartilhado de planilhas eletrônicas, desenvolvimento colaborativo de software ou assistência remota de software.

### **Problemas envolvidos**

As ferramentas de aplicação compartilhadas, são softwares que implementam o algoritmo necessário ao controle de acesso e o ordenamento de acesso aos recursos compartilhados, além disso protegendo recursos privados. Em geral, o controle de acesso do tipo bloqueio implícito é o mais utilizado.

Outra função necessária é a gestão de entrada e saída de participantes. Alguns sistemas permitem que usuários se juntem no início de uma sessão, sem controle adicional ou após a admissão por parte de um responsável. O sistema pode informar a todos a lista de participantes.

### **Requisitos de Rede**

Os requisitos de rede são um pouco mais severos que as ferramentas de quadro branco compartilhado. O nível de interação é maior e o serviço resultante é mais sensível a atrasos de transito na rede: o controle de acesso pode se tornar impraticável em redes que geram grandes atrasos. Muitas implementações atuais operam sobre o IP.

### **Produtos**

Existem várias ferramentas de aplicação compartilhadas de domínio público e alguns produtos comerciais. Elas são geralmente baseadas no X-Windows e rodam sobre o IP.

Shared X é um produto comercial da Hewlett-Packard que usa um algoritmo de controle de acesso do tipo bloqueio implícito. Nele, apenas o iniciador necessita explicitamente lançar a aplicação, os outros participantes são requisitados a se juntarem. O iniciador pode incluir novas aplicações.

XTV (*X Teleconferencing Viewing*) é uma implementação de domínio público onde o controle de acesso baseado em mediador. Todos devem lançar explicitamente a ferramenta XTV. Ela permite o compartilhamento de várias aplicações simultaneamente, sendo que o chairperson é quem pode incluir novas aplicações.

O Microsoft NetMeeting também permite o compartilhamento de qualquer aplicação Windows.

## **5.6 Distribuição de Áudio e Vídeo**

O objetivo deste tipo de aplicação é a disseminação de fluxos de áudio e vídeo para múltiplos indivíduos em localizações distintas. Com relação a taxonomia, as aplicações descritas aqui são do tipo pessoa-a-pessoa, de distribuição, unidirecional, e é endereçada à comunicação de grupos e não a de indivíduos. Usualmente ela opera no modo síncrono. Elas são muito usadas como groupware para suportar CSCW e para a difusão de eventos (como seminários, reuniões) sob a rede.

### **5.6.1 Campos de Aplicação**

A distribuição de áudio e vídeo no modo unidirecional é usado principalmente para a eventos ao vivo e armazenados, como a difusão de seminários, conferências, teses e reuniões. Podendo usar multicast para grupos fechados ou abertos, ou ainda

funcionalidades broadcast verdadeiras. O campo de aplicação inclui: educação, treinamento profissional, apresentações comerciais e de negócios, atividades promocionais públicas ou privadas, trabalho cooperativo e distribuição de eventos profissionais. A disseminação de seminários ou reuniões abertas podem ser classificada em duas categorias:

- **Distribuição institucional:** distribuição dentro de uma organização local, que implica o uso exclusivo de uma rede de transporte no local (por exemplo, em uma Intranet).
- **Distribuição externa:** para várias organizações sobre as WANs (por exemplo, a Internet).

Um exemplo de aplicação muito popular nos dias de hoje são os canais de TV e rádio sobre a Internet.

### 5.6.2 Características da distribuição de vídeo de apresentações

A seguir serão apresentados alguns requisitos de rede e qualidade de serviços das aplicações de distribuição de áudios e vídeos.

#### **Resolução/taxa de quadros de vídeo digital**

A menos que a taxa de bits disponível e o desempenho da rede permita uma qualidade de TV, vídeo digital é sempre uma questão de compromisso entre a taxa de quadros e a resolução. Com uma taxa de bits limitada, aumentando uma implica na redução da outra.

No caso de apresentações envolvendo apenas discussão, os requisitos são os mesmos da videofonia: é melhor reduzir solavancos do que ter alta resolução. No caso de apresentações com várias transparências ou uso intensivo do quadro é melhor ter alta resolução do que alta taxa de quadros. Por exemplo, a transmissão de vídeo na taxa de 1 fps com resolução de qualidade TV tem melhores resultados que 4 a 5 fps com baixa resolução.

Uma técnica para melhorar o desempenho global seria o controle, via operador, da resolução e da taxa de quadros. Assim, poderia-se escolher em melhorar a resolução quando uma certa transparência ou um quadro branco é focado, e aumentando a taxa de quadros quando a resolução não é importante. Infelizmente os sistemas atuais não suportam esta técnica.

#### **Qualidade do som**

O som é o componente essencial em aplicações de distribuição de áudio e vídeo. Diferente da videofonia ou mesmo da telefonia, usuários são mais tolerantes a distorções quando assistem passivamente a uma apresentação. Assim, uma pequena perda da semântica é aceitável.

### 5.6.3 Distribuição de Áudio e Vídeo Institucional

A distribuição de áudio e vídeo institucional refere-se a distribuição de eventos como seminários e reuniões abertas a nível de uma organização como um campus universitário ou uma companhia. Neste caso a rede de comunicação envolvida são as redes locais (LANs).

#### **Que computador usar?**

O recurso mínimo para um computador receber fluxos de áudio e vídeo é a instalação de módulos de software cliente de áudio e vídeo. Estes módulos descompactam os fluxos de áudio e vídeo e obviamente devem ser compatível com a fonte de emissão. Um módulo deste tipo em software é suficiente apenas para baixa taxa de quadros e baixa qualidade de resolução: com mais alta taxa de quadros a decompressão consumiria toda a capacidade de processamento da CPU.

A carga da CPU pode ser aliviada com o uso de cartas (placas) de decompressão. Assim a taxa de quadros pode ser alta, mas todavia a resolução não. Para isto o

computador deve ter também uma grande quantidade de memória de vídeo. Muitas cartas de compressão/descompressão substituem a memória de vídeo por buffers de quadro de alta resolução (24 bits) dedicados.

### **Que LAN usar?**

Redes locais baseadas em pacotes são mais onipresentes, e os requisitos são os similares a da videofonia a pacotes. A taxa de bits necessária é de no mínimo 200 Kbits/s. Esta taxa permite uma boa resolução, mas com movimentos aos trancos (até 2 fps). Esta taxa é compatível com LANs de meios compartilhados, tal como Ethernet ou Token Ring. Para obter uma qualidade de TV (25 ou 30 fps) é necessária uma taxa de 3 a 6 Mbits/s com MPEG-2. Redes Fast Ethernet e ATM suportam esta taxa.

### **Produtos existentes**

Poucos produtos existem para distribuição de áudio e vídeo para uso institucional. Muitas delas usam o protocolo IP sobre LANs e WANs a pacotes. Por exemplo, Uniflix/Paradise é um produto comercial projetado para Internet que roda sobre plataformas SUN e HP. Outras implementações fornecem comunicação bidirecional, com um canal de retorno, como os produtos comerciais InPerson/SGI e ShowMe/SUN e os softwares de domínio público vat, NEVOT (áudio apenas), nv e vic (vídeo apenas), e IVS.

Atualmente existem pouco uso da distribuições de eventos a nível institucional. A razão é a dificuldade de inteiramente controlar as técnicas multicast de alta taxa nas LANs atuais. É esperado que os avanços do wide-area broadcasting seja levado a serviços locais.

#### **5.6.4 Distribuição de Áudio e Vídeo em WANs**

O número de eventos que são transmitidos em redes de longa distância (WANs) vem aumentando a cada dia e são muito populares sobre a infraestrutura MBone (vista mais adiante). A maior parte da discussão apresentada na seção anterior se aplica a distribuição de áudio e vídeo em WANs (na maior parte das vezes sob a Internet). Nesta seção são apresentados apenas os aspectos que diferenciam a distribuição de áudio e vídeo institucional da distribuição em WANs.

Abaixo são apresentadas algumas aplicações desta classe de aplicações multimídia:

- *Apresentações comerciais*: apresentação de produtos comerciais.
- *Conferências Profissionais*: permitindo uma maior audiência. Um exemplo é a distribuição regular da IETF (*Internet Engineering Task Force*) sob a Internet.
- *Educação à Distância*: palestras são distribuídas a sítios distantes. Usado principalmente dentro do quadro de colaborações educacionais entre universidades.
- *Grandes projetos colaborativos distribuídos*: reuniões podem ser difundidas mundialmente.
- *Promoção de pesquisas ou agências governamentais*: agências públicas podem usar a rede para difundir eventos. Um exemplo de difusão ao vivo da missão a Marte pela Nasa, e a distribuição das palestras do SBRC'97 (Simpósio Brasileiro de Redes de Computadores), ambas sob a Internet.

### **Confidencialidade e Broadcasting**

Os problemas de confidencialidade ou privacidade é mais evidente sobre WANs que em redes fechadas a organizações (p.e. Intranet). Por exemplo, no caso do exemplo de difusão de reuniões de grandes projetos colaborativos, é interessante que apenas pessoas autorizadas pudessem assistir a reunião.

Existem duas formas (não excludentes) de se implementar alguma forma de confidencialidade para distribuição de serviço:

- **Criptografia**: neste caso apenas o receptor com uma chave de criptografia pode apresentar o fluxo recebido. Esta abordagem é usada, por exemplo, pelos canais

de TV pague-parte-ver. Progressos significativos estão ocorrendo no desenvolvimento algoritmos seguros e de pouco custo de CPU. Na maior parte das vezes, softwares de criptografia são instalados no emissor e nos receptores. A autoridade central distribui uma chave ao grupo fechado.

- **Multicast para grupos fechados:** no caso da criptografia não for possível, existem duas opções para obter um certo nível de confidencialidade: usar uma rede física que conecta exclusivamente os receptores autorizados (no caso de um ambiente corporativo); ou o fluxo é distribuído sobre uma rede aberta no modo multicast com um grupo fechado. Este último é difícil de implementar sobre o IP.

### Requisitos de Rede

A taxa de bits necessária à distribuição de áudio e vídeo sob WANs é a mesma da distribuição em redes locais: 200 Kbits/s para alta resolução e baixa taxa de quadros e 6 a 8 Mbits/s para qualidade de TV usando MPEG-2.

A dificuldade essencial das redes a pacotes é a taxa de bits disponíveis, como nem todos os pontos aceitam 200 Kbits/s adicionais na linha de acesso à WAN. Além disso, a WAN em si pode usualmente manipular apenas poucos fluxos broadcast simultâneos. IP, que é provido com uma funcionalidade broadcast emulada (obtida por multicast para grupos abertos), é a tecnologia de rede de escolha de hoje, e a Internet é a rede de transporte natural para suportar grandes distribuições. ST-II e IPv6 são tecnologias de comutação de pacotes que suporta multicast inerente. Diferente do IP usado atualmente (Ipv4), ST-II é provido de um mecanismo de reserva de recursos. Este limita o risco de perdas de pacotes que é inerente ao IP. Outro problema da rede IP é que ela é do tipo melhor esforço, tornando a qualidade altamente dependente da carga da rede. Uma solução para o problema de qualidade imprevisível é superdimensionar a taxa de bits da rede.

### Produtos

A maioria dos produtos de distribuição de áudio e vídeo em WAN usam o protocolo IP. Uniflix/Paradise é especialmente projetado para comunicação unidirecional, mas muitas implementações são bidirecionais, mais úteis para áudio-videoconferências. Como na videofonia a pacotes, a principal dificuldade está associada a interoperabilidade, parcialmente sanada pelo uso de softwares de domínio público, como vat, NEVOT, nv, Rendez-Vous e IVS.

Os produtos da RealNetworks ([www.real.com](http://www.real.com)), como RealPlayer (Figura 31) e RealEncoder vem sendo muito usado na Web. Além desse, existem os produtos VDOLive da VODNet ([www.vdo.net](http://www.vdo.net)), Vosaic ([www.vosaic.com](http://www.vosaic.com))

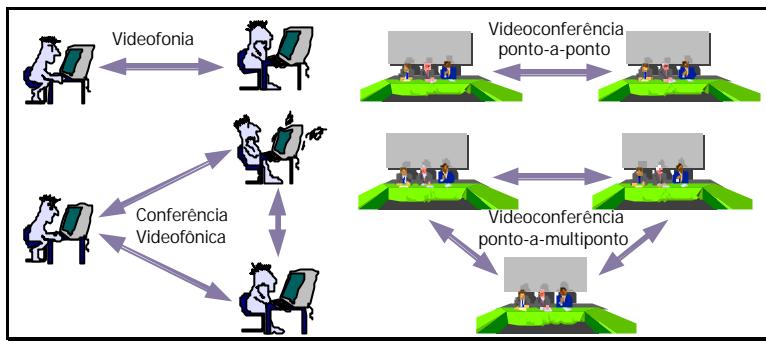


Figura 31. RealPlayer da RealNetworks

## 5.7 Áudio-Videoconferência

A Áudio-videoconferência ou simplesmente videoconferência, envolve vários indivíduos ou vários grupos de indivíduos engajados em diálogo. O objetivo não é manter uma simples conversação bilateral, mas suportar reuniões. Assim, a videoconferência pode

ser ponto-a-ponto como na videofonia, ou ponto-a-multiponto como na distribuição de vídeo, mas sempre implica numa comunicação bidirecional (ilustrado na Figura 32).



**Figura 32.** Videofonia e videoconferência

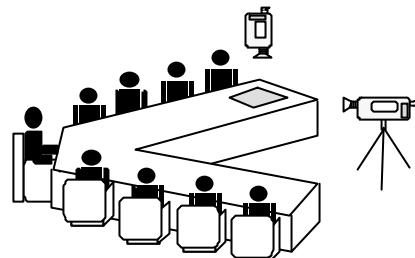
Segunda a taxonomia adotada neste capítulo, videoconferência é uma aplicação orientada a grupo, pessoa-a-pessoa e síncrona, orientada para ambientes profissionais e assim contribui ao CSCW.

### 5.7.1 Principais Características e Requisitos

A videoconferência necessita de vários equipamentos de suporte e módulos de software.

#### Manipulando Grupos

A videoconferência envolve ao menos um grupo de pessoas em uma das localizações. Assim é necessárias uma ou várias câmeras para registrar as cenas. As várias opções de registro de vídeo são:



- *uma câmera de TV fixa:* a câmera registra uma vista completa do grupo. Neste caso geralmente as pessoas se sentam em uma mesa na forma de um "V" ou na forma de um semi-círculo.
- *uma câmera de TV móvel:* a câmera filma apenas o interlocutor atual.
- *uma câmera fixa e outra móvel:* este caso é a junção dos dois anteriores, sendo que a visão é chaveada por um operador.

#### Manipulando Documentos

Geralmente documentos devem ser transferidos para dar suporte a videoconferência. Documentos podem ser:

- *Documentos impressos:* são capturados por câmeras verticais (chamadas câmeras documento). A dificuldade associada é a resolução. Muitos sistemas de videoconferência adotam como compromisso baixa resolução e taxa de quadros média, tornando documentos difíceis de serem lidos. Alguns produtos permitem a comutação, para captura de documentos apenas, para técnicas de compressão melhores adaptadas para imagens (JPEG). Outra opção para captura é utilizar scanners rápidos, permitindo uma melhor resolução.
- *Documentos projetados:* Reuniões geralmente contém projeções de slides, de transparências, que também exigem uma boa resolução para poderem ser lidos (implicando na necessidade de redução do número de quadros). Um bom resultado pode ser provido usando Motion JPEG.
- *Documentos eletrônicos:* transferência de documentos digitais, documentos scaneados antes da seção ou gerados por computador asseguram uma melhor qualidade.

#### Requisitos

Exceto quando documentos estão envolvidos, os requisitos da videoconferência são similares aqueles da videofonia: é importante reduzir o salto de movimento e uma

limitada resolução pode ser tolerada. Isto pois a baixa taxa de quadros altera a transmissão da informação emocional, essencial para emular a comunicação face-a-face. Portanto taxas de 8 a 12 fps são aceitáveis, assim a resolução é limitada geralmente a um quarto da qualidade de TV. Quando a videoconferência envolve a apresentação de documentos, ela necessita de uma resolução de média qualidade para que o documento seja legível.

Como apresentado anteriormente, usuários são mais tolerantes a distorções quando assistem passivamente. Assim uma pequena perda da semântica é aceitável, devendo o som ter uma qualidade suficiente para ser amplificado por alto-falantes.

### 5.7.2 Videoconferência baseada em circuitos e pacotes

Os primeiros sistemas de videoconferência eram baseados em difusão por satélite de televisão. A comunicação era assimétrica, com um sítio origem do vídeo e múltiplas localizações receptoras. O canal de retorno era normalmente feita via chamadas telefônicas. Desde então, os serviços de videoconferência transformaram-se em simétricos, e seu desenvolvimento resultou em duas abordagens tecnológicas:

- **Videoconferência baseada em circuito:** este tipo apareceu no início dos anos oitenta. Estes sistemas operam sobre redes com taxas de transferência garantidas - inicialmente conexões telefônicas e agora sobre ISDN.
- **Videoconferência baseada em pacotes:** este tipo apareceu no início dos anos 90, principalmente com o avanço das estações de trabalho e computadores pessoais.

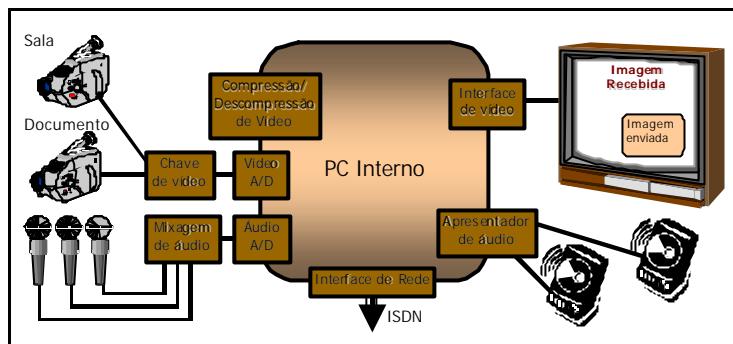
#### *Pequeno histórico da videoconferência baseada em circuito*

A videoconferência surgiu nos anos oitenta, puxada pelo desenvolvimento da tecnologia de TV digital e progressos nos algoritmos de compressão. Os primeiros serviços eram fornecidos pelos PTO (Operadoras Públicas de Telecomunicações), em salas de reuniões dedicadas (estúdios de vídeo), equipadas com dispositivos de áudio e vídeo analógicos, digitalizadores, e sistemas compressão/descompressão, e conexões aos sistemas de comutação a circuito.

A partir da metade dos anos oitenta surgiram os estudos privados e linhas continuamente alugadas conectando sítios privados. Apenas grandes organizações podiam dispor destes estúdios privados, isto devido aos seus altos custos dos circuitos de 1,5 ou 2 Mbits/s.

Nesta época apareceram produtos dedicados à videoconferência (Figura 33), compostos de pacotes com câmeras de TV, microfones, alto-falantes, monitores, e módulos para digitalização, compressão e descompressão, geralmente empacotadas em uma unidade física única. Estes produtos eram instalados em salas privadas, chamados de estúdio de videoconferência privados. A primeira característica destes produtos é que eles eram dispositivos dedicados (não de uso geral), projetados para operar sobre circuitos de taxa de transmissão garantidas. Estes sistemas geralmente produziam uma taxa de fluxo de bits constante. Tais sistemas eram chamados de Video-codecs, ou simplesmente codecs. Codecs é uma abreviação de codificador/decodificador usado no processamento de sinal. Eles são instalados em salas privadas conectadas a redes de comutação de circuitos, em geral ISDN, com velocidades variando de 112 ou 128 Kbits/s a 335 ou 384 Kbits/s.

A última geração de sistemas de videoconferência a circuito repousa-se sob o modo desktop onde o serviço é fornecido no escritório. Estes sistemas necessitam de um serviço comutado de domínio público, geralmente ISDN, no escritório. Muitos sistemas permitem dois circuitos simultaneamente, estes fornecendo uma taxa de bits de 112 ou 128 Kbits/s. Tipicamente, 16 Kbits/s é dedicado ao áudio e 96 a 112 Kbits/s ao canal de vídeo. Muitos sistemas oferecem funcionalidades de quadro branco compartilhado. Em geral, eles suportam apenas uma câmera de TV, mas podem oferecer uma câmera de vídeo documento opcional.



**Figura 33.** Anatomia de um sistema de videoconferência a circuito típico

### **Pequeno histórico da videoconferência baseada em pacotes**

A emergência de sistemas de videoconferência explorando redes a pacotes não pode ser separada do desenvolvimento da videofonia a pacotes, no início dos anos noventa. Ela resultou da exploração dos dispositivos (estações de trabalho e PC) e redes de dados (LAN) já disponível nos escritórios para suportar serviços de videoconferência. A videoconferência e a videofonia são suportadas pelos mesmos produtos, que são disponíveis na forma de kits de hardware e software para estações de trabalho e computadores pessoais.

Atualmente os produtos oferecem uma resolução de vídeo de alta qualidade, mas nos primeiros produtos, a apresentação e equipamentos de captura eram de qualidade média. O controle da câmera era também bastante limitado. Um dos maiores desafios destes sistemas é a integração de ferramentas de áudio e vídeo com outras ferramentas de teleconferência, tal como quadro branco e aplicações compartilhados, para fornecer uma plataforma de teleconferência multimídia integrada.

Produtos de videoconferência de fabricantes diferentes geralmente não interoperam. Softwares de domínio público são geralmente a única solução para interoperabilidade.

### **Campos de Aplicação**

Sistemas de videoconferência são essencialmente projetados para aumentar o trabalho cooperativo entre parceiros remotos, especialmente quando complementado por ferramentas de espaço de trabalho compartilhado e aplicações compartilhadas. Alguns campos de aplicação da videoconferência são:

- comunicação entre executivos
- gestão de projetos de desenvolvimento e comerciais distribuídos
- projeto colaborativo, engenharia conjunta, debates a distância
- assistência e consulta a distância
- tutoria remota na educação a distância
- telemedicina

### **5.7.3 Videoconferência a Pacotes com comunicação ponto-a-multiponto**

No capítulo 4, foi apresentado dois meios técnicos de obter funcionalidades multicast sobre uma rede de pacotes: múltiplas conexões ponto a ponto ou tecnologia multicast a pacotes.

#### **Múltiplas conexões ponto-a-ponto**

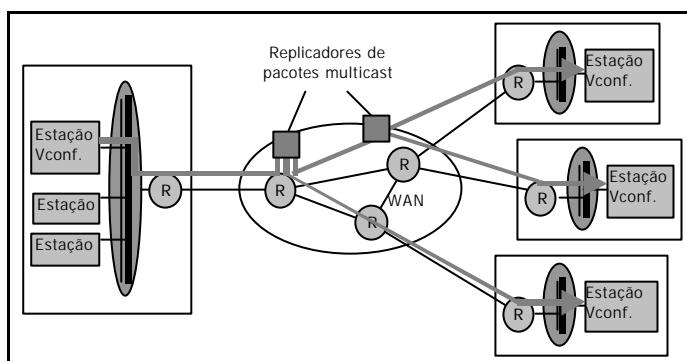
Neste caso um número de conexões ponto-a-ponto são estabelecidas entre os participantes da videoconferência. Diferentes dos sistemas de videoconferência a circuitos, a taxa de bits não é garantida pelas redes a pacotes. Exceto quando se utiliza redes de pacotes orientadas a conexões especiais tal como aquelas baseadas sobre o protocolo ST-II. Neste caso a rede básica não toma conhecimento destas conexões especiais, elas são convenções puras entre sistemas finais. Elas são ditas conexões a

nível de transporte. As convenções ou regras que os sistemas finais adotam em tais conexões são governadas pelos protocolos de transporte. Sobre a Internet normalmente é utilizado o protocolo RTP (*Real-Time Transport Protocol*) ou o TCP (*Transport Control Protocol*).

O maior problema deste esquema de implementação do multicast é que a fonte deve gerar múltiplos fluxos de dados idênticos, um para cada destino. Este esquema não é econômico em termos de processamento na fonte e de comunicação. A vantagem deste esquema é que a confidencialidade é mais alta, pois ela apenas envolve membros especificamente aceitos.

### **Multicast para grupos**

Suportar multicast para grupos em uma LAN é teoricamente simples: explora-se a capacidade multicast inerente às LANs de meios-compartilhados (Ethernet, Token Ring, FDDI). Sobre WANs, esta funcionalidade é disponível sobre vários tipos de rede sem conexão (IP, SMDS – Switched Multimegabit Data Service) ou orientada conexão (ST-II). Utilizando um serviço multicast, um fluxo de uma fonte chega a todos os membros do grupo. Ela é uma abordagem elegante, mas nem todas implementações asseguram confidencialidade. Quando utilizada para grupos abertos, a privacidade pode ser assegurada usando criptografia.



**Figura 34.** Anatomia de um sistema de videoconferência típico

#### **5.7.4 Requisitos de Rede**

Para suportar sistemas de videoconferência baseados em circuitos, as redes de escolha são ISDN ou Switched 56, sendo que estes sistemas estão projetados para operar sobre taxas de 56 a 384 Kbits/s. Este tipo de sistema é útil para videoconferência no modo sala, para o modo desktop ele seria muito caro. O serviço a circuito emulado oferecido pela tecnologia ATM é uma alternativa de alta velocidade ao ISDN, abrindo uma era de muito mais alta qualidade de áudio e vídeo. A tecnologia ATM é o sucessor natural do (N-)ISDN.

Os requisitos de taxa de bits da videoconferência baseada em pacotes são similares aqueles da distribuição de vídeo, embora o compromisso resolução/taxa de quadros seja diferente. Uma resolução VCR e 2 fps requer 100 Kbits/s, sendo que 13 a 64 Kbits/s devem ser adicionados para canal de áudio. Uma qualidade equivalente a TV requer 6 a 8 Mbits/s usando produtos JPEG-baseados. Diferente da distribuição de vídeo, que não é uma aplicação interativa, a videoconferência é sensível a atrasos. Atrasos introduzidos pelos sistemas finais são na ordem de segundos. Os atrasos introduzidos pela rede são muito menores, na ordem de décimos a poucas centenas de milisegundos.

Para videoconferências ponto-a-ponto sob longas distâncias com baixa taxa de quadros e media resolução pode-se utilizar qualquer rede, seja sem conexão (IP) ou orientada conexão (X.25, Frame Relay, ATM). Mas quando redes são sobrecarregadas, tecnologias providas de alguma forma de garantia de taxa de bits, tal como ST-II ou ATM, trabalham melhor.

Videconferências ponto-a-multiponto não trabalham muito bem sem multicasting a pacotes. Vários protocolos Internet suportam multicast, incluindo IP e ST-II. A versão inicial do IP multicast está evolutivamente sendo suportada por roteadores, mas eles não

suportam grupos fechados multicast. Multicast ST-II fornece uma melhor confidencialidade, em adição ao mecanismo de reserva de taxa de bits, mas esta tecnologia está sendo empregada apenas dentro de algumas comunidades.

Videoconferência sob redes a pacotes é uma extensão tecnológica da videofonia um-a-um que construtores de estações de trabalho propuseram no início dos anos noventa. O principal problema está associado a capacidade da rede a pacotes fornecer funções multicast, permitindo extensão dos membros da videoconferência com um custo adicional mínimo, e numa maior capacidade de estações incorporarem outras ferramentas multimídia de conferência.

### Produtos para Videoconferência a Pacotes

Existem vários fabricantes de componentes ou soluções integradas completas, entre eles estão: Bolt, Beraneck and Newman (Picture-Window), CLI (Cameo family), DEC (DECSpin), GTE (DTVS), InSoft (Communiqué!), IBM (Person-to-Person/2), Intel (ProShare), InVision Systems (inVision Videoconferencing) e Microsoft NetMeeting. Alguns destes produtos podem operar no modo pacote ou circuito. Várias implementações de domínio público são disponíveis. Elas incluem IVS, vat, NEVOT, nv, CU-SeeMe para áudio e/ou vídeo, e sd para anúncio e junção à conferências.

## 5.8 Aplicações Baseadas em Servidores Multimídia

Nesta seção serão apresentadas algumas aplicações multimídia cujo objetivo não é a comunicação entre pessoas, mas a comunicação pessoa-a-sistema. Desde que estes sistemas baseados em computadores podem servir a uma variedade de funções, eles são chamados de servidores. As aplicações multimídia de comunicação pessoa-a-sistema podem se dividir em duas classes:

- **Aplicações interativas:** onde o servidor age como um repertório de informações que o usuário acessa via rede, tais como vídeos e documentos multimídia. Eles oferecem interatividade remota para leitura de documentos, telecompras reservas de entradas, entretenimentos como video games e vídeo sob-demanda. Em todas as aplicações o iniciador da comunicação é o usuário final.
- **Aplicações de distribuição:** onde o servidor distribui espontaneamente informações para todos os usuários (*broadcast*) ou apenas para um subconjunto selecionado (*multicast*), e usuários individuais podem ou não ter alguma forma de controle individual sobre a apresentação local.

### 5.8.1 Acesso Interativo a Servidores Multimídia

Nesta seção nós exploraremos as aplicações pessoa-a-sistema do tipo interativas. Nesta classe de aplicação, o usuário interage com o servidor para pedir a informação desejada. Conceitualmente, as pessoas podem acessar servidores multimídia por dois motivos (Figura 35):

- **Aplicações orientadas a busca (ou a recuperação):** o objetivo é localizar, apresentar, e possivelmente registrar informações multimídia (figura 9a). Este caso engloba leitura de um documento multimídia, de um filme, etc.
- **Aplicações orientadas a transação:** neste caso há uma busca com processamento da entrada do usuário para outros propósitos que controlar a apresentação: aqui, a meta final não é acessar a informação multimídia, mas podendo ser a compra de itens, reserva de ingressos, armazenamento de exercícios educacionais, etc.

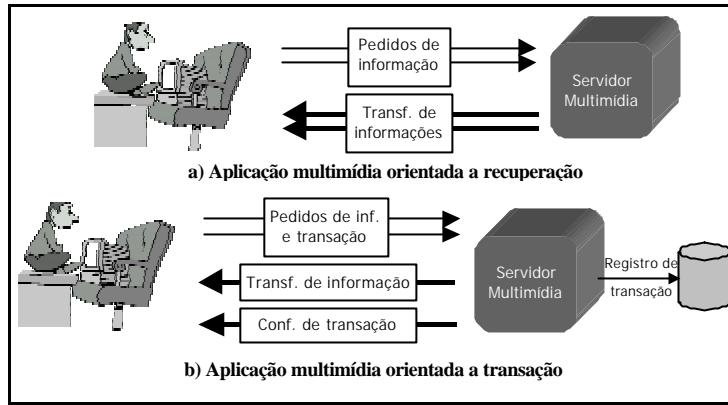


Figura 35. Aplicações de busca e de busca com processamento

### 5.8.2 Comunicação Assíncrona e Síncrona

Duas formas de transmissão são possíveis nesta classe de aplicação multimídia:

- **Transmissão Assíncrona** ou telecarga, armazenamento e apresentação: neste modo a informação, ou parte dela, é primeiro totalmente transferida e armazenada no receptor, para depois ser apresentada. A informação pode ser pedida pelo usuário ou espontaneamente distribuída pelo servidor. Este modo é mais simples de implementar que o modo de transmissão síncrona.
- **Transmissão Síncrona** ou tempo-real: neste modo parte ou toda a informação é transferida em tempo-real sobre a rede e apresentada continuamente no receptor. Este modo de comunicação exige um suporte de comunicação tempo-real.

Tomamos como exemplo uma enciclopédia eletrônica para entender que modo de sincronização utilizar na transferência de informações multimídia. Uma enciclopédia contém uma composição de textos, imagens, áudios e vídeos. É fora de questão telecarregar completamente uma enciclopédia no sistema cliente. Muitos sistemas cliente poderiam não ter a capacidade de armazenamento necessária. Assim, a enciclopédia é formada por partes que o usuário pode navegar através, e que podem ser transferidas independentemente.

Parte contendo texto apenas podem ser transferidas no modo assíncrono ou síncrono. Como textos não exigem muito da rede, o usuário praticamente não notará diferenças entre estes dois modos.

Para mídias contínuas do tipo áudio e vídeo, geralmente na leitura da enciclopédia o usuário é convidado, via uma interface gráfica, para ativar estas sequências clicando um botão. Na teoria, seguido o clique do botão a sequência de áudio ou vídeo pode ser transferida no modo síncrono ou assíncrono:

- *Modo assíncrono:* é usado normalmente para pequenas sequências que podem ser facilmente armazenadas no receptor (mesmo na memória central do sistema cliente). Neste caso a apresentação será realizada após a carga completa (Figura 36a). Geralmente, atrasos de 3 a 4 segundos no início da apresentação são toleráveis. Para sequências pequenas e boas condições de rede, o usuário terá a impressão de uma reação em tempo-real do servidor.
- *Modo síncrono:* é necessário para sequências de áudio e vídeo muito grandes, redes muito lentas, ou pouca capacidade de armazenamento no sistema cliente. Neste caso, o sistema não aguarda a carga completa da sequência (Figura 36b), há um pequeno atraso inicial e em seguida a sequência é apresentada na medida que pacotes chegam.

Muitos sistemas não suportam o modo de transferência síncrona. Como resultado, o conteúdo do documento deve ser estruturado de modo que longas sequências são cortadas em partes menores.

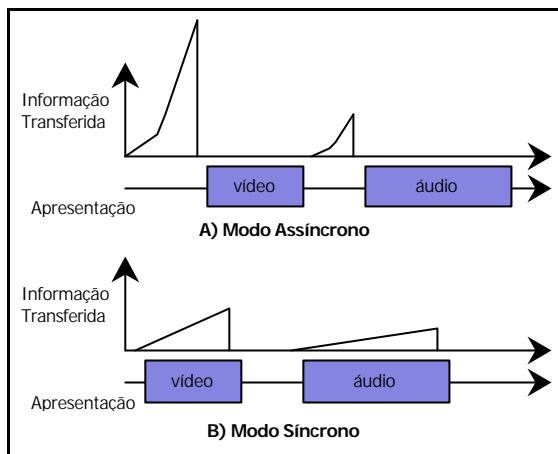


Figura 36. Modos de transmissão

### 5.8.3 Problemas na Transmissão Síncrona (Tempo-Real)

Neste capítulo, nós vimos que as informações multimídia transferidas podem ser compostas por mídias discretas (textos, gráficos), mídias contínuas (som e vídeos) e relações temporais podem existir entre estes elementos. Quando transferida sobre a rede as relações temporais podem ser alteradas devido a variação do atraso de trânsito de algumas redes, onde certos fragmentos podem ser atrasados em relação a outros. Além disso, certos fragmentos podem ser perdidos ou a ordem pode ser alterada.

Se alterações temporais ocorrerem, é necessário utilizar mecanismos no sistema cliente e servidor a fim de:

- restaurar as relações temporais dentro de cada fluxo transportando uma mídia contínua. Este processo é chamado de **Streaming** ou **sincronização intra-mídia**, que assegura que o fluxo recebido seja rearranjado para se parecer com o original. É necessário, por exemplo, garantir as dependências temporais intra-mídia entre os quadros de um vídeo.
- restaurar relações temporais entre os vários fluxos ou elementos (**Sincronização Intermídia**). Em outras palavras, restaurar dependências temporais inter-mídia.

### 5.8.4 Vídeo sob-demanda (VOD)

Vídeo sob demanda cobre todas as aplicações onde os usuários possam pedir acesso a servidores de imagens fixas e animadas numa base individual.

#### *Quasi vídeo sob-demanda (Q-VOD)*

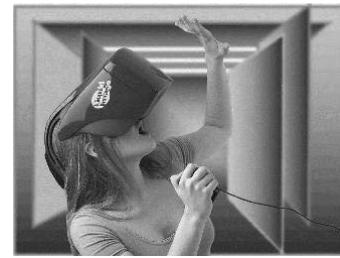
Verdadeiros VOD são extremamente custosos em termos de poder para acessar e ler o dispositivo de armazenamento, poder de processamento nos servidores e consumo de largura de banda na rede. Imagine que em uma hora de pico, pode haver vários pedidos de um vídeo popular ao mesmo tempo, cada um separados por intervalos de segundos. Verdadeiros VOD gerariam em um dado instante centenas de fluxos idênticos defasados apenas de poucos segundos. Diz-se que cada cliente tem uma comutação de fase independente no mesmo filme. QVOD agrupa pedidos idênticos que são servidos em intervalos regulares, isto afim de reduzir a comutação de fase e reduzir a carga do servidor. Esta técnica pode também reduzir a largura de banda, através do uso de técnicas multicast.

#### *Qualidade necessária e taxa de bits associada*

Uma qualidade equivalente ao videocassete VHS requer uma taxa de transmissão na ordem de 1,5 Mbits/s. De qualquer maneira, uma qualidade inferior aquela do NTSC e PAL/SECAM pode não ser aceita facilmente pelos clientes potenciais. Implementações existentes do padrão MPEG-2 opera em 6 Mbits/s com um qualidade um pouco superior ao de TV. Muitos operados estão planejando operar em uma infra-estrutura piloto sob uma taxa de 6 ou 7 Mbits/s.

## 5.9 Realidade Virtual

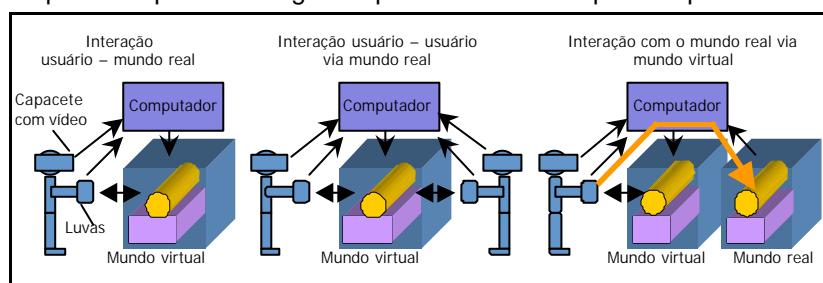
Realidade virtual é o último nível de sofisticação para a comunicação multimídia entre pessoas e entre pessoas e sistemas. O termo realidade virtual refere-se a uma aplicação baseada em computador que fornece uma interface homem-computador tal que o computador e seus dispositivos criam um ambiente sensorial que é dinamicamente controlado pelas ações do indivíduo, assim que o ambiente se assemelha a realidade para o usuário. A realidade virtual implica em uma comunicação direta entre o usuário e o computador, implicando em interatividade. Assim ela suporta aplicações síncronas.



O princípio fundamental é o fato que o computador cria um mundo sensorial que se parece com o mundo real que reage as entradas dos usuários. Nós chamamos este mundo sensorial de mundo virtual. Este mundo virtual não implica em um ambiente multisensorial (composto de várias mídias), ele pode ser apenas um mundo auditivo (*Virtual Auditory Worlds*) ou visual apenas.

Existem três grandes tipos de interação na realidade virtual, com diferentes metas (ilustrados na Figura 37):

- **Interação usuário-mundo virtual.** Esta interação é usada na exploração pelo usuário de um mundo virtual criado pelo computador, constituindo uma comunicação entre o usuário e o computador. Um exemplo de uso deste esquema é a exploração por arquitetos de uma simulação de suas futuras realizações.
- **Interação usuário-usuário via mundo virtual.** Esta interação permite vários usuários participando da mesma aplicação. Neste caso o computador age como mediador entre os usuários. Este esquema é utilizado em jogos tridimensional ou simulação de combate aéreo.
- **Interação com o mundo real via mundo virtual.** Esta interação permite realizar ações a distância no mundo real, através de uma representação visual do mundo. Neste caso computador age como mediador entre o usuário e o mundo real. É usado por exemplo em cirurgia tele presente assistida por computador.



**Figura 37.** Três tipos de interação em realidade virtual

# Capítulo 6 Documentos Hipermídia

## Multimídia e

Este capítulo é consagrado ao estudo da problemática de criação de documentos multimídia e hipermídia. Como apresentado no capítulo 1 desta apostila, um sistema multimídia pode ser definido como um sistema capaz de manipular ao menos um tipo de mídia discreta e um tipo de mídia contínua, as duas numa forma digital. Neste trabalho, nós consideramos que sistema hipermídia é um sistema multimídia na qual as informações monomídia e multimídia são acessadas e apresentadas com a ajuda de mecanismos de navegação baseadas em ligações (*Links*). Um documento multimídia/hipermídia define uma estrutura de informações multimídia.

Documentos multimídia/hipermídia existem em várias áreas e níveis. Tipicamente estas aplicações multimídia são aplicadas nas áreas de educação, treinamento profissional, quiosque de informações públicas, ou mercado varejista. Um exemplo de uma aplicação ou documento multimídia é um jornal de biologia que inclui a visualização de moléculas, funcionalidades para acessar e buscar artigos de pesquisas associados, funcionalidades de comunicação colaborativa, impressão, anotação, ou animação.

## 6.1 Definição de Documento Multimídia e Hipermídia

Hipermídia combina diferentes tipos de mídias de apresentação, oferecido pela multimídia, com a estrutura de informação oferecida pelo hipertexto [Hardman, 94].

### 6.1.1 Documentos Hipertexto

Um documento hipertexto é uma estrutura de informação organizada de maneira não linear: os dados são armazenados em uma rede de nós conectadas por ligações ou *links* (Figura 38):

- Os **nós** contém as unidades de informação compostas por texto e outras informações gráficas. Em geral, um nó representa um conceito ou uma idéia expressa de uma maneira textual ou gráfica.
- Os **links** definem as relações lógicas (ou semânticas) entre os *nós*, isto é eles definem relações entre conceitos e idéias. A noção de *âncora*, permite a especificação de uma parte da informação que será fonte ou destino de um *link*. Uma âncora hipertexto é descrita por uma seqüência de caracteres ou o *nó* inteiro. A Figura 38 apresenta uma correspondência entre as informações e as âncoras na tela, e os nós, os links e as âncoras na base de dados hipertexto.

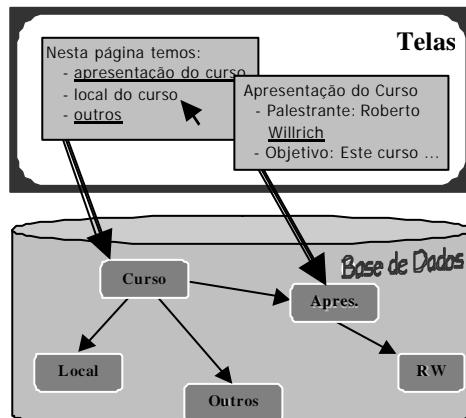


Figura 38. Base de dados hipertexto

Em geral, quando o usuário de uma aplicação hipertexto clica sobre uma âncora, o *link* é disparado (seguido), causando a apresentação da âncora destino. Dessa forma, o usuário pode “navegar” no documento através das âncoras e dos *links*. Algumas vezes um nó é definido como uma unidade de navegação.

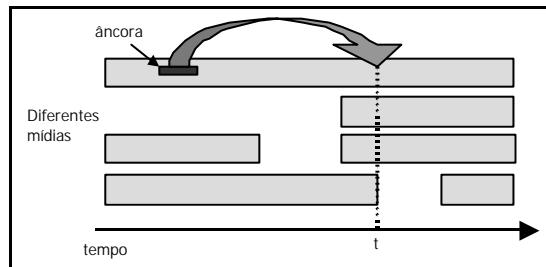
Existem outras formas de acesso às informações que a navegação via links hipertextos. Por exemplo, a partir de um mecanismo de alto nível, o leitor pode pedir uma busca por palavra-chaves, ir diretamente a um nó específico, ou percorrer os nós visitados anteriormente.

Como os textos e imagens são informações estáticas (as informações não evoluem no tempo), a noção de tempo não é utilizada quando da especificação de documentos hipertextos clássicos. No hipertexto, o tempo de apresentação determinado pelo leitor do documento.

### 6.1.2 Documentos Multimídia

Um documento multimídia pode ser visto como uma estrutura descrevendo a coordenação e o estilo de apresentação de uma coleção de componentes constituídos de mídias estáticas e dinâmicas. A Figura 39 apresenta um pequeno documento multimídia. Um exemplo de documento multimídia real pode ser a apresentação dos pontos turísticos de uma cidade, constituído de textos, imagens e vídeos de cada ponto turístico, apresentados em uma ordem seqüencial e possivelmente com alguns mecanismos de interação.

Quando o autor cria um documento multimídia, ele define a **orquestração** da apresentação dos componentes a partir da definição dos instantes de início e fim das apresentações e das relações condicionais e temporais entre e no interior dos componentes. Esta descrição da ordem temporal relativa de apresentação dos componentes é chamada de **cenário multimídia**.



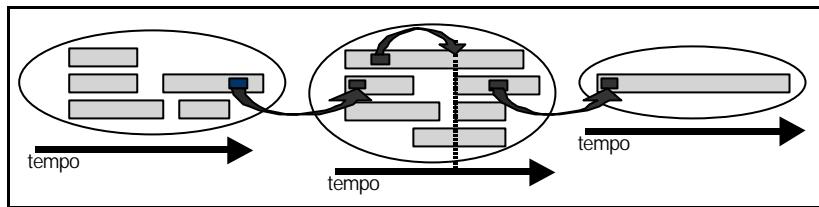
**Figura 39.** Documentos Multimídia [Hardman, 94]

Documentos multimídia podem ser interativos. Tipicamente, existem dois métodos de interação a partir do qual o leitor pode controlar a apresentação [Hardman, 94]:

- um método de controle de apresentação, que permite o ajuste da referência temporal utilizada pela apresentação. Com este mecanismo, o leitor pode interromper, repartir, fazer avanços ou retorno rápidos na apresentação do documento, isto a partir de uma interface similar aos controles de um videocassete.
- um método de interação similar a um link hipertexto rudimentar definido por uma âncora e um link. Seguindo o link, a apresentação salta para uma outra parte do documento. Este método é ilustrado na Figura 39: quando o leitor clica sobre a âncora, a apresentação salta para o ponto t.

### 6.1.3 Documentos Hipermídia

Um documento hipermídia é uma combinação de documentos hipertexto e multimídia (ilustrado na Figura 40). Esta classe de aplicação representa uma evolução natural do hipertexto, na qual as noções ou conceitos de nós hipertexto podem agora serem expressos por diferentes tipos de mídia de apresentação. A inclusão de dados multimídia aumenta o poder de expressão da informação contida em uma aplicação e torna a apresentação mais atrativa e realista.



**Figura 40.** Documentos Hipermídia

Como nos documentos hipertextos, as âncoras permitem a especificação de uma parte da apresentação de uma mídia que será fonte ou destino de um *link* hipermídia. Por exemplo, uma âncora pode ser uma seqüência quadros de um vídeo ou uma região de uma imagem.

Os links hipermídia podem ser temporizados. Este tipo de link é habilitado durante um determinado intervalo temporal e ele pode ser ativado automaticamente em função das suas restrições temporais. Por exemplo, em link temporizado pode levar a uma página de ajuda caso o leitor não ative nenhum outro link em 30 segundo. Este tipo de link introduz a noção de sistemas hipermídia ativos [Stotts, 90].

A inclusão de dados multimídia em estruturas hipertexto introduz um aspecto temporal na especificação de aplicações hipermídia. Dessa forma, é necessário que os modelos de especificação hipermídia forneçam mecanismos que permitam a integração temporal de uma variedade de mídias contínuas e discretas. Além disto, ele deve fornecer mecanismos de descrição das possíveis interações com o usuário do documento, isto é ele deve permitir a descrição dos link hipermídia, extensão dos link hipertextos (com um caráter temporal), e outros métodos de interação.

Em muitos trabalhos, não há a distinção entre documentos multimídia e hipermídia. Neste caso o termo documento multimídia engloba os documentos multimídia da forma apresentada na seção 6.1.2 e os documentos hipermídia da forma apresentada na seção 6.1.3. Nesta apostila, por simplificação, nós utilizaremos o termo documento multimídia para especificar os documentos multimídia e hipermídia.

## 6.2 Autoria de Documentos Multimídia

Ambientes de desenvolvimento multimídia facilitam e automatizam a autoria (criação) de documentos multimídia. Há uma grande variedade de tais ambientes, também chamados de sistemas de autoria. Estes sistemas de autoria são projetados para fornecer ferramentas de criação e organização de uma variedade de elementos de mídia como texto, gráficos, imagens, animações, áudio e vídeo a fim de produzir documentos multimídia. Os usuários deste tipo software, os autores dos documentos, são profissionais que desenvolvem apresentações educacionais, de marketing e artistas gráficos que fazem decisões acerca do layout gráfico e estilo de interação que o usuário final real (estudantes ou indivíduos no público) vêem e ouvem [Buford, 94].

Sistemas de autoria geralmente fornecem meios para gerar certos tipos de interações comuns através do uso de *templates* reutilizáveis ou módulos de software, que podem ser personalizados pelo projetor para um objetivo particular. Como exemplo, muitos programas de treinamento usam um formato de questão múltipla escolha.

Existem também pacotes de software de autoria dedicados a produção de aplicações de realidade virtual. Estes softwares de modelagem tridimensional permitem que o autor crie formas, luminosidade e regras de comportamento para o ambiente virtual.

### 6.2.1 Evolução histórica

Sistemas de autoria para educação e treinamento existem a vários anos. Os primeiros programas para criação de materiais educacionais surgiram na década de sessenta com o software Plato baseado em mainframe da Control Data Corporation. O sistema Plato resultou em numerosos descendentes para mainframes e mini-computadores.

A inclusão de vídeo interativo via videodiscos analógicos para cursos de treinamento foi uma importante contribuição do sistema TICCIT (*Time-shared Interactive Computer Controlled Information Television*), desenvolvido pela MITRE em conjunto com a Universidade Brigham Young nos inícios dos anos setenta.

A interface com o usuário para sistemas de autoria tem evoluído da mesma maneira que as interfaces com o usuário dos computadores pessoais nas duas décadas passadas. Inicialmente, os sistemas de autoria eram dirigidos por informações textuais. O processo de construção da estrutura e elementos do documento multimídia tornaram-se mais dirigidos a menu e então orientados a janelas até a metade dos anos oitenta quando vários sistemas pioneiros na abordagem baseadas em ícones surgiram, por exemplo Authorware e IconAuthor (vistos mais adiante).

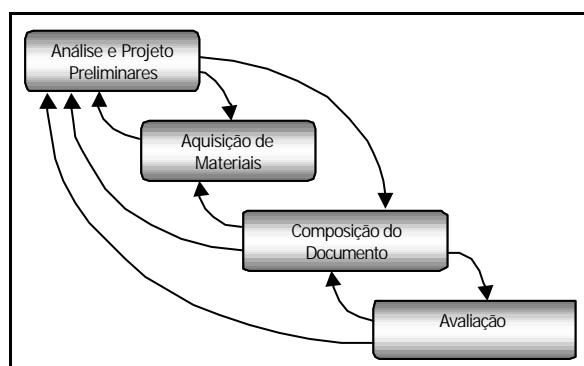
Existem hoje três grandes plataformas para a criação e apresentação de aplicações multimídia: compatíveis IBM PC rodando DOS, Microsoft Windows e Windows NT, ou OS/2; Apple Macintosh; e estações de trabalho com alguma variação do UNIX (Silicon Graphics, Sun, etc.). Estes sistemas de apresentação e autoria hoje fornecem uma gama de diferentes modelos para os autores criarem e apresentarem aplicações multimídia, chamados nesta apostila de **modelos multimídia**.

### 6.3 Criação de documentos multimídia

Embora diferentes documentos multimídia possam impor diferentes requisitos, há um conjunto de passos relativamente consistentes que o autor deve seguir quando ele desenvolve um documento. Estes passos, apresentados na Figura 41, guia o autor da inspiração inicial ao documento acabado. Como apresentado nesta figura, o processo de autoria pode ser dividido em quatro estágios:

- *análise e projeto preliminares*, em que os requisitos para o documento, seu conteúdo e suas interfaces são especificados;
- *aquisição de material*, em que os materiais que formarão o documento são coletados, criados ou digitalizados;
- *composição do documento*, em que é realizada a composição lógica (através de links), temporal e espacial dos componentes do documento;
- *avaliação e liberação*, em que o documento é testado, refinado e, finalmente, distribuído para sua audiência.

Observe na Figura 41 que há realimentações importantes no processo de autoria (p.e., na composição do documento pode-se verificar a necessidade da aquisição de novos materiais). Ainda nesta figura, estes quatro estágios não são sempre realizados em seqüência. Certos estágios podem ser realizados em paralelo (por exemplo, a aquisição de material pode ocorrer simultaneamente com a composição do documento). Além disso, a criação do documento pode ser feita via prototipagem, onde as etapas de aquisição de material, composição e avaliação podem ser repetidas até a conclusão do documento. Mesmo na prototipagem, a etapa de análise e projeto preliminar deve a mais completa que possível afim de aumentar as chances de se alcançar as metas definidas.



**Figura 41.** Passos na construção de documentos multimídia

Os passos individuais são discutidos a seguir.

### 6.3.1 Análise e Projeto Preliminares

O projeto de um documento multimídia começa com uma análise das metas e expectativas do produto resultante. Isto envolve considerar questões da autoria tradicional (baseada em texto) tão bem quanto questões similares ao projeto de programas de computadores. Estas questões incluem: as razões que levam o autor a criar o documento, o público alvo do documento, que e como materiais poderiam ser apresentados, os recursos (de pessoal, materiais, financeiros) disponíveis, as características especiais dos leitores do documento, e o conteúdo específico do documento.

O acesso ao perfil dos usuários potenciais é um ponto chave no projeto de documentos multimídia [Makedon, 94]. Um documento criado para novatos com material muito sofisticado é tão inútil quanto documentos criados para especialistas com material introdutório. Além disso, desde que um documento multimídia é uma coleção de informações e pacotes de software, os autores dos documentos devem avaliar os conhecimentos dos usuários do ponto de vista do conteúdo e da tecnologia de apresentação. Por exemplo, os especialistas em um campo particular podem não ser especialistas no uso de uma interface com o usuário complexa.

Como documentos multimídia podem ser feitos de uma grande diversidade de materiais, autores devem decidir qual é o tipo de mídia de apresentação que melhor informa a audiência e quando múltiplas formas de mídias são necessárias para segmentos particulares do material. Muitas questões podem aparecer neste processo de decisão. O custo e benefício de mídias não tradicionais e a maneira em que estas mídias serão incluídas deve ser pesados. O tempo estimado para produção do documento também afeta na decisão associadas a custo e benefício: um documento mais sofisticado exige um maior tempo de produção, assim tempos limitados para produção pode exigir a eliminação de algumas características desejáveis.

Os componentes do documento e a forma de organização destes componentes devem também ser decididos. Este é um dos principais resultados desta etapa do projeto de um documento multimídia: um relatório contendo a identificação dos componentes dos documentos e a estruturação destes componentes (por exemplo, na forma de capítulos, seções, etc.).

### 6.3.2 Criação e aquisição de materiais

Identificado todos os componentes desejáveis no documento, o passo seguinte é coletar ou criar materiais que comporão o documento e converter eles na forma digital. Estes materiais podem ser constituídos de dois tipos de informações: **dados mídia**: que podem ser visto pelo leitor (áudio, vídeo, texto, etc.); **dados não-mídia**: não podem ser vistos imediatamente (documentos HyperODA, HyTime).

Normalmente não é necessário criar todos os componentes do documento multimídia, alguns materiais já existem em base de dados existentes particulares ou de domínio público. A determinação de que materiais necessitam ser coletados especificamente para o documento e o que é necessário ser criado é um processo que consome tempo. Quando possível, é preferível obter materiais no formato eletrônico e em papel. O formato eletrônico facilita a transição para a forma digital; cópia em papel fornece um registro mestre para se usar como referência.

Autores de documentos multimídia/hipermídia necessitam de ferramentas para gerar dados originais ou importar de uma fonte externa para cada um dos tipos suportados pelo sistema de autoria. Na metade dos anos oitenta, vários desenvolvedores incluindo MacroMind (hoje Macromedia) promoveu o conceito de editores de domínio – aplicações para geração e captura de dados de diferentes tipos tal como texto, animações, áudio, etc. No passado, estas ferramentas de edição eram com freqüência proprietárias e podiam apenas ser usadas com um sistema de autoria particular. As interfaces com o usuário e características variavam muito de um vendedor para outro.

Hoje a tendência é que sistemas de autoria permitam ao autor criar e capturar dados usando pacotes de software especializados e importar o resultado na sua apresentação. Esta é uma abordagem interessante, permitindo que os autores utilizem ferramentas

mais familiares reduzindo o tempo de aprendizagem e promovem o compartilhamento de recursos. Por exemplo, texto pode ser editado via um processador de texto favorito, então aceito pelo sistema de autoria. Imagens podem ser escaneadas usando uma variedade de hardware e software e em seguida podem ser importadas em apresentações. Existem muitos pacotes de edição gráfica que podem ser usados para criar e importar imagens bitmaps da mesma maneira. Para isto, padrões de representação de informações monomídia devem ser adotados, permitindo uma maior reutilização deste tipo de informação entre sistemas de autoria.

A maioria dos sistemas de autoria comerciais, limitados a plataforma em que eles residem, acessam apenas mídias armazenadas localmente. O compartilhamento do conteúdo e equipamentos multimídia tem várias vantagens. MAEstro [Drapeau, 91] é um exemplo de um ambiente de autoria em que um conjunto de ferramentas de edição de mídia tem sido integrado e que pode acessar recursos localizados em qualquer parte da rede. A base das aplicações do MAEstro é um sistema de mensagens baseado na chamada de procedimento remoto (RPC), que permite o acesso distribuído. Outro exemplo é o GainMomentum, é um sistema de autoria comercial, que é construído sobre um sistema de gerência de base de dados distribuído.

Durante o projeto e desenvolvimento de um documento multimídia, é importante obter materiais adicionais para salva-guardas. Estes materiais adicionais podem ser incorporados no documento em possíveis futuras alterações, ou eles podem ser usados para corrigir materiais incluídos no documento. Por exemplo, mesmo se um documento multimídia inclui apenas o texto de uma fala, o áudio da fala pode ser usado durante o desenvolvimento afim de verificar se o que foi dito casa com o que foi digitado.

Após coletado e criado todo o material, eles devem ser convertidos na forma digital. Alguns materiais necessitarão ser digitalizados ou, se obtidos na forma eletrônica, eles podem necessitar ser convertidos para outros formatos. Por exemplo, documentos TeX (o formato empregado por muitos cientistas e matemáticos) pode ser convertido para PostScript, HTML ou ASCII.

Após a obtenção da representação digital da informação multimídia, o autor pode precisar editar a informação, na maioria das vezes para realizar uma simples limpeza dos dados multimídia. Por exemplo, pode-se remover "hum's" e "né's", pausas e "tiques" verbais do áudio. Isto torna o áudio muito mais agradável para apresentação, além de reduzir o tamanho. Pode também ser necessário retocar textos e figuras que não foram adequadamente digitalizados e reformatar alguns documentos para colocar de acordo com os requisitos da tela do computador. Algumas destas edições simples podem ser executadas automaticamente, mas muitas devem ser realizadas a mão para assegurar a qualidade do produto.

### 6.3.3 Composição de Documentos Multimídia

Desde que na forma digital, os materiais devem ser combinados para formar o documento multimídia. Isto envolve a composição lógica, temporal e espacial do material, onde o tipo correto de informação, no tipo correto de formato, deve aparecer no tempo correto e no lugar correto. Tal orquestração de informações multimídia é talvez o passo que mais consome tempo no processo de desenvolvimento.

No momento da composição do documento, algumas vezes é necessário editar novamente o material afim de ajustá-lo à orquestração do documento. Por exemplo, encurtar ou modificar a velocidade de apresentação de vídeos e áudios, alterar tamanhos de imagens, etc.

A tarefa de composição de documentos multimídia é uma das tarefas mais complexa no processo de criação de documentos multimídia. Este tema será retomado mais adiante.

### 6.3.4 Avaliação e refinamento

Neste ponto, o documento multimídia completo deve ser testado. Para isto pode-se selecionar um grupo de usuários para testar o documento, sendo que este grupo deve verificar a adequação do conteúdo, recursos e links. Quando possível, o documento

deve ser distribuído a usuários novatos e especialistas para obter opiniões deste dois grupos de usuários.

No caso de múltiplos autores, o primeiro grupo de teste deve ser composto pelos autores. Isto é muito importante, pois os autores têm o controle sobre seus materiais e eles ajudam na reparação de links semânticos incorretos.

Finalmente o documento multimídia pode ser liberado: (a) como um pacote de software disponível sobre a rede; (b) como um servidor de documento remoto na rede; (c) em um CD-ROM, ou (d) com um meio relacionado.

## 6.4 Requisitos para um Modelo Multimídia

Ferramentas de autoria de documentos multimídia são baseados em um modelo multimídia que induz uma técnica de descrição de documentos multimídia. Baseado na ISO ODA (*Office Document Architecture*) [ISO 8613] várias abordagens dividem a descrição de documentos em três partes:

- a **estrutura lógica**, que descreve as diferentes partes lógicas de um documento (ou componentes) e suas relações lógicas;
- a **estrutura de apresentação**, que descreve como, onde e quando os diferentes componentes serão apresentados;
- a **estrutura de conteúdo** descreve as informações que constituem os componentes.

[Fluckiger, 95] afirma que a partição da descrição do documento nestas três estruturas oferece várias vantagens. A separação entre a estrutura lógica e de apresentação, por exemplo, permite a reutilização da estrutura lógica quando da apresentação do documento em dispositivos diferentes, onde somente uma modificação simples ou adaptação da estrutura de apresentação deve ser realizada; e a separação entre a estrutura de apresentação e do conteúdo é necessária pois os documentos podem utilizar os mesmos dados em diferentes contextos [Schloss, 94] ou por diferentes documentos

A separação entre a semântica pura (a estrutura lógica) e a definição dos instantes de aparição dos componentes (o quando, definido pela estrutura de apresentação) é contestada por vários especialistas. Isto pois um documento multimídia não é somente aquilo que ele contém, mas também quando seus conteúdos são apresentados. Assim, baseado na arquitetura proposta por [Klas, 90], nós podemos considerar que um modelo multimídia ideal deveria permitir uma descrição multi-nível incluindo as estruturas seguintes:

- a **estrutura conceptual**, que descreve as diferentes partes lógicas do documento ou componentes, suas relações lógicas, e em que instante estes componentes serão apresentados;
- a **estrutura de apresentação**, que descreve como e onde os diferentes componentes serão apresentados;
- a **estrutura de conteúdo**, que descreve as informações que constituem os componentes.

A única diferença entre esta classificação e a primeira é que o “quando” da apresentação dos componentes é transferido da estrutura de apresentação para a estrutura lógica. Assim, contrariamente a afirmação de [Fluckiger, 95], quando da apresentação do documento em diferentes sistemas clientes, a estrutura lógica e o comportamento temporal (a estrutura conceptual) podem ser reutilizados. As apresentações (a estrutura de apresentação e do conteúdo) não suportados pelo sistema cliente devem ser adaptados e o “quando” da apresentação destes componentes deve ser respeitado. Por exemplo, se um componente constituído de uma apresentação de áudio e se o sistema cliente não suporta a apresentação deste tipo de mídia, o conteúdo deste componente deve ser substituído por uma informação textual. Mas os comportamentos temporais destas apresentações devem ser equivalentes.

Além da descrição dessas três estruturas, um modelo para essas aplicações deve permitir também a especificação dos possíveis métodos de interação com o usuário. Dessa forma, um modelo deve permitir a definição de âncoras, de *links* hipermídia e de outros métodos de interação de documentos multimídia.

#### 6.4.1 Estrutura de Conteúdo

Como apresentado na seção 6.3, uma das primeiras etapas da realização de um documento multimídia é a geração ou captura dos materiais (textos, imagens, áudios, vídeos, etc.) que vão ser utilizados para compor o documento. Por exemplo, o autor pode usar uma câmera de vídeo para registrar uma seqüência de vídeo, criar uma informação gráfica a partir de um editor. Nesta apostila, nós chamaremos estas informações de **dados primitivos**.

A estrutura do conteúdo é responsável pela descrição dos dados primitivos. Por exemplo, ela pode ser constituída por um conjunto de dados primitivos e seus descritores. O par dados primitivos e descritores destes dados é chamado de **objeto multimídia**.

Ao nível da estrutura do conteúdo, um modelo multimídia deve permitir, entre outros, a especificação das informações de acesso e de manipulação dos dados primitivos e os valores originais das características espaciais, sonoras e temporais de apresentação.

#### 6.4.2 Estrutura Conceptual

A estrutura conceptual especifica os componentes e os grupos de componentes de um documento, e a composição lógica e temporal destes componentes.

##### **Os componentes e grupos de componentes**

A estrutura conceptual contém as definições dos componentes semânticos de um documento, permitindo a divisão do documento em, por exemplo, capítulos e parágrafos, ou uma série de seqüências de vídeo e de títulos. Um exemplo desta estruturação é apresentado na Figura 42.

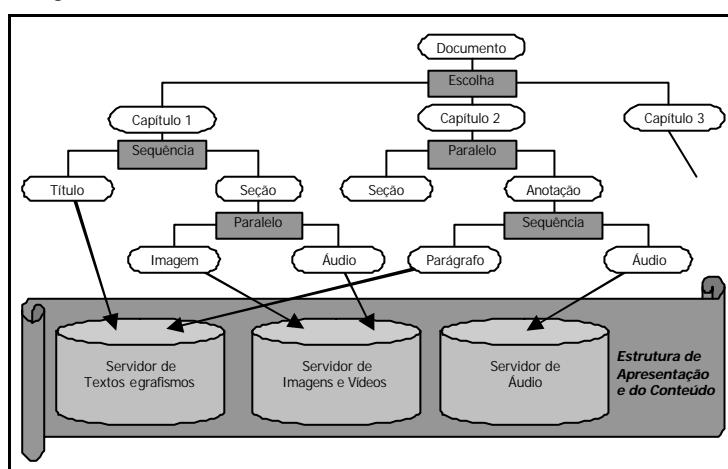


Figura 42. Componentes e grupos de componentes

A tarefa de especificação de documentos multimídia se torna delicada e complexa com o aumento de tamanho do documento. A estrutura conceptual é utilizada para a construção de apresentações complexas a partir de pequenos grupos. Assim, a descrição do documento pode ser dividida em seções, cada seção podendo ser criada de maneira independente das outras seções [Hardman, 95]: as apresentações que exprimem juntas uma idéia ou um conceito podem ser agrupadas para posterior reutilização. Além disso, estes mecanismos de estruturação permitem a composição do documento a partir de técnicas *top-down* e ou *bottom-up*. Eles permitem também a definição de relações lógicas e temporais entre estes grupos e a definição de restrições temporais associadas a um grupo de apresentações. Concluindo, a estrutura conceptual introduz o benefício da modularidade, da encapsulação e mecanismos de abstração.

### ***Independência dos componentes e seus conteúdos***

A definição do componente do documento deve ser realizada independentemente das informações acerca da representação do seu conteúdo [Klas, 90], que são descritos pelas estruturas de apresentação e do conteúdo (como ilustrado na Figura 42). Assim, os componentes da estrutura conceptual devem fazer unicamente menção às informações que estes componentes representam. Afim de simplificar a especificação do documento, esta menção deve ser independente do tipo de informação (mídia ou não-mídia) ou de sua localização geográfica (local ou remota).

Quando do armazenamento e da transmissão do documento, certas partes da estrutura conceptual e/ou do conteúdo dos componentes podem ser incluídas explicitamente ou implicitamente na estrutura do documento:

- **Inclusão explícita:** a estrutura do documento contém os dados primitivos. Assim, a estrutura do documento e os conteúdos incluídos explicitamente são armazenados no mesmo arquivo e transferidos conjuntamente.
- **Inclusão implícita:** a estrutura do documento faz referência aos dados primitivos ou outras partes da estrutura conceptual. Assim, no modo de transmissão em tempo real e de telecarga somente uma parte do documento é transferida. Quando da apresentação do documento o sistema cliente deve acessar de maneira remota as partes incluídas implicitamente no documento.

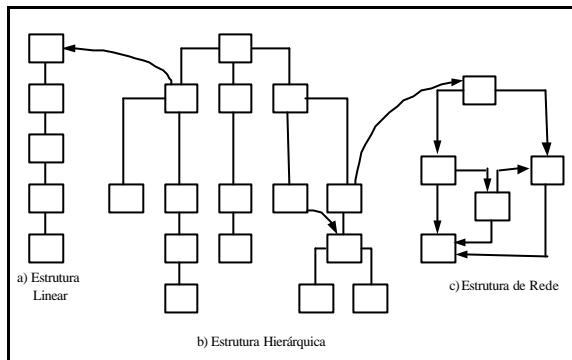
### ***Os caminhos de percurso e as estruturas de Informação***

Estrutura conceptual define os caminhos de percurso do documento. Estes caminhos de percursos são definidos pelos links. Fundamentalmente, existem dois tipos de links [Shackelford, 93]:

- **Link estrutural**, que fornece uma estrutura de base de um documento. Quando o usuário segue esses links, o discurso (ou a estrutura) de base é preservado como ele foi criado pelo autor do documento. Na Figura 43, os traços que ligam os componentes (representados por quadrados) são links estruturais.
- **Link hiper-estrutural**, que permite a definição das relações que transpassam a estrutura de base do documento. Estes links podem ser divididos em links associativos e referenciais [Ginige, 95]. Os **links associativos** conectam conceitos associados e os **links referenciais** conectam informações adicionais a um conceito. Na Figura 43, as flechas representam os links hiper-estruturais do documento.

Nós podemos distinguir três maneira de se estruturar a informação [Ginige, 95]:

- **Estruturas Lineares** (Figura 43a): ela é utilizada nos documentos do tipo “visita guiada”, onde o documento fornece um procedimento de ajuda com uma lista de ações que devem ser executadas em uma ordem particular. Por exemplo, as estruturas lineares estão presentes no material de treinamento, pois um instrutor normalmente ensina os módulos de uma maneira seqüencial. O autor pode incluir questões em vários pontos, caso o leitor der uma resposta errada, informações adicionais podem ser apresentadas via links referenciais.
- **Estrutura Hierárquica** (Figura 43b): ela está presente nos documentos comparáveis aos livros (organizados em capítulos e seções), e o autor pode utilizar, por exemplo, links referenciais para um glossário ou uma seção de referências.
- **Estruturas em Rede** (Figura 43c): contém ligações associativas. Estes links são de natureza semântica e pragmática, e são verdadeiramente não seqüenciais. A estrutura em rede é mais adaptada para a organização de informações na forma de enciclopédia.



**Figura 43.** Tipos de links e estruturas para navegação em um documento

Um documento multimídia pode conter ao mesmo tempo estes três tipos de estrutura; assim, um modelo hipermídia ideal deveria suportar todos estes tipos de estrutura.

### Composição temporal do documento

A estrutura conceptual define também a estrutura temporal de uma aplicação, que consiste na descrição dos instantes de partida e parada das apresentações dos componentes e de suas relações temporais e condicionais (ou causais). Essas relações são estabelecidas por eventos definidos no interior de diferentes apresentações. Existem dois tipos de eventos que podem ocorrer durante uma apresentação:

- os **eventos síncronos** (previsíveis): são os eventos cujas posições no tempo são determinadas previamente (e.g., início da apresentação de uma seqüência de áudio ou vídeo). A posição destes eventos pode ser determinada somente sob condições ideais (sem considerar os atrasos imprevisíveis, como sobrecarga na rede).
- os **eventos assíncronos** (imprevisíveis): são os eventos cujas posições no tempo não podem ser determinadas previamente. Como o instante em que a aplicação chega a um determinado estado ou a interação do usuário.

### As Relações Temporais

As relações temporais definem as posições temporais relativas entre e no interior dos componentes de um documento. Estas relações temporais definem sincronizações que podem ser definidas entre eventos definidos dentro de uma apresentação ou entre eventos definidos em apresentações distintas:

- **sincronização intramídia:** são as relações temporais entre os eventos ou intervalos definidos no interior de uma mídia contínua. Por exemplo, definidos entre os quadros de uma seqüência de vídeo.
- **sincronização intermédia:** são as relações temporais entre os eventos ou intervalos definidos em diferentes apresentações (mídias).

Em geral, as sincronizações intramídia são dependências temporais naturais que são definidas implicitamente quando da produção dos dados primitivos (na estrutura do conteúdo). O esquema de sincronização intramídia pode ser alterado pelo autor do documento quando da definição da estrutura de apresentação. Por exemplo, a sincronização entre dois quadros de uma seqüência de vídeo pode ser definida pela velocidade de apresentação desta seqüência. As sincronizações intermédia são geralmente dependências temporais artificiais especificadas explicitamente pelo autor da estrutura conceptual de um documento. Assim, em geral as sincronizações intermédia são descritas pela estrutura conceptual do documento, e as sincronizações intramídia fazem parte da estrutura de apresentação.

As relações temporais de um documento podem ser expressas de forma estática, que consiste em organizar explicitamente as relações temporais entre os componentes (orquestração); ou de forma dinâmica, em que as relações temporais denominadas relações *ao vivo* são definidas [Fluckiger, 95]. Um exemplo de relação *ao vivo* é a sincronização de um trabalho cooperativo. Um modelo multimídia trata unicamente da orquestração, já que as relações *ao vivo* não podem ser especificadas em avanço, no

momento da criação do documento. Assim, a estrutura conceptual define apenas os aspectos de orquestração (sincronização estática) da apresentação multimídia.

### Modelo de Composição

Para a orquestração de um documento é necessário um modelo de especificação das relações lógicas, temporais e condicionais. Um tal modelo pode definir a composição do documento de duas maneira:

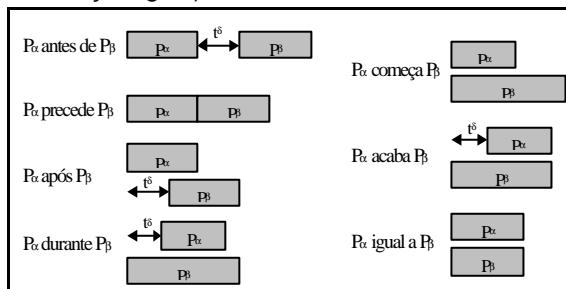
- **Descrição procedural:** descreve a estrutura de controle da apresentação do documento a partir da descrição ordenadas das ações.
- **Descrição declarativa:** descreve as restrições temporais entre as apresentações sem tomar em consideração a ordem cronológica das ações (um evento futuro pode influenciar uma ação no passado). Este caso requer um cálculo do escalonamento das apresentações que compõem o documento. Este cálculo pode ser realizado em tempo de compilação e/ou em tempo de execução [Blakowski, 96].

Na seção 6.5 retorna a este assunto, onde serão apresentados modelos básicos de composição de documentos multimídia.

### Modelo Temporal

Para a descrição das relações temporais, um modelo de composição requer um **modelo temporal**. Com relação a sua unidade elementar, duas classes de modelos temporais podem ser identificados [Wahl, 94]:

- **Modelos temporais baseados em pontos:** são os modelos cuja unidade temporal são os eventos. Existem três relações temporais que podem ser definidas entre dois eventos: um evento pode ocorrer antes, em simultâneo e após outro evento.
- **Modelos temporais baseados em intervalos:** são os modelos cuja unidade temporal são os intervalos. [Hamblin, 72] e [Allen, 83] definem que existem 13 relações temporais possíveis entre dois intervalos (Figura 44): antes, precede, após, durante, começa, acaba e igual, aos quais se ajunta as relações inversas (com exceção da relação igual).



**Figura 44.** Relações temporais básicas entre intervalos

Para a multimídia, modelos temporais baseados em intervalos são melhores que aqueles baseados em pontos. Isto pois em mais alto nível de abstração, as apresentações podem ser vistas como intervalos temporais, com um início, fim e uma duração. Isto simplifica a especificação das relações temporais e condicionais.

### Relações condicionais

Uma relação condicional é definida como uma condição associada a um conjunto de componentes, e ações que serão aplicadas a um conjunto de componentes quando esta condição é satisfeita. Por exemplo, “após o término da apresentação de A, se o link L está ativado, então apresentar B”, é uma relação condicional.

Um modelo hipermídia ideal deveria fornecer mecanismos para especificar este tipo de relação.

### **Sincronização em sistemas distribuídos**

Por causa do não determinismo da duração de tratamento das informações em sistemas multimídia distribuídos (causado pela rede de comunicação, atrasos de acesso a base de dados, etc.), as relações temporais desejadas podem não ser sempre garantidas. É por isso que um modelo multimídia deve permitir a especificação de **métodos de tolerância** de sincronização [Wynblatt, 95]. Assim, o autor pode expressar quais compromissos de sincronização são aceitáveis e os meios de tratar as exceções quando da violação.

#### **6.4.3 Estrutura de Apresentação**

A estrutura de apresentação de um documento descreve as características espaciais, sonoras e temporais de cada apresentação que compõem o documento<sup>6</sup>. A este nível, o autor deve especificar as características de apresentação de cada componente e a composição espacial destas apresentações em um instante dado. A especificação das características de apresentação inclui a descrição da maneira como o componente será visto (descrição das características espaciais) e/ou ouvido (descrição das características sonoras) pelo leitor do documento. A composição espacial descreve também as relações espaciais entre as apresentações.

#### **Especificação das características de Apresentação**

A estrutura conceptual define os componentes do documento e suas relações. A partir disto, o autor deve definir o conteúdo dos componentes (isto é, associar um objeto multimídia ao componente) e particularizar e/ou definir suas características de apresentação. Por exemplo, o autor pode trocar as características originais de apresentação dos objetos multimídia (como o volume sonoro, velocidade de apresentação, etc.), ou ele pode definir novas características, como a posição espacial de apresentação de uma imagem.

A fim de modelar uma apresentação, um modelo multimídia deve permitir a especificação das seguintes informações:

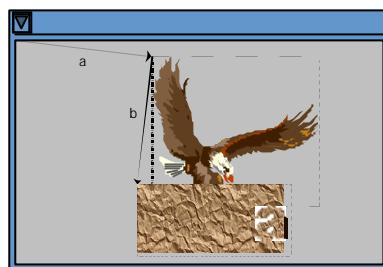
- Características temporais de apresentação das informações dinâmicas, como a velocidade, posição de início e de fim de um vídeo e o número de repetições.
- Características espaciais de apresentação de informações visuais, como o tamanho, a posição e o estilo de apresentação.
- Características das apresentações sonoras, como o volume de apresentação.
- Dispositivos de saída, chamados aqui de canais, na qual as informações serão apresentadas e vista pelo leitor (por exemplo, uma janela, um canal de áudio).
- Apresentações alternativas podem ser definidas afim de repor uma apresentação principal se ela não puder ser apresentada em um certo sistema (permitindo a criação de documentos adaptáveis aos recursos disponíveis), se existirem problemas de acesso, ou restrições temporais não satisfeitas.

#### **Composição espacial**

O método mais usual de posicionamento espacial da apresentação dos componentes na tela do computador é a definição da posição espacial absoluta (Figura 45a) ou relativa (Figura 45b) de cada apresentação em um sistemas de coordenadas virtuais.

---

6 Existem outros tipos de tratamento dos componentes que suas apresentações, como a preparação dos dados e a execução de executáveis. Algumas vezes, o autor pode personalizar o tratamento de informações. Por exemplo, um script pode conter parâmetros de entrada cujos valores podem ser definidos pelo autor. Neste trabalho, nós não consideraremos este tipo de caracterização de uma apresentação multimídia.



**Figura 45.** Composição Espacial

#### 6.4.4 Interações

Em aplicações interativas (p.e., multimídia interativa, hipertexto e hipermídia) o usuário deve dispor de um conjunto de mecanismos que permitam o controle da apresentação. existem basicamente quatro métodos de interação [Hardman, 95]:

- **Navegação:** este método permite a especificação de um conjunto de escolhas dado ao usuário a fim de que ele possa selecionar um contexto entre vários. Ela é geralmente definida através da criação de um *link* ou *script*<sup>7</sup> que liga as âncoras origens às âncoras destinos.
- **Controle da apresentação:** este método de interação é freqüentemente encontrado em documentos multimídia, onde o leitor pode parar, recomeçar, avançar ou retroceder a apresentação do componente multimídia.
- **Controle do ambiente:** este método permite a particularização do ambiente de apresentação do documento. Por exemplo, o leitor pode desativar o canal de áudio ou ainda pode alterar o tamanho de uma janela.
- **Interações da aplicação:** nos métodos anteriores, o autor cria o documento e o leitor apenas interage com ele. Existem aplicações que requerem mecanismos específicos, por exemplo, nas aplicações de tele-ensino, o modelo deve permitir a especificação da noção de acompanhamento dos alunos (por exemplo, histórico das atividades do aluno) e de avaliação (por exemplo, a partir de um campo de entrada de dados). Outro método de interação é a pesquisa por palavras-chave. Este tipo de mecanismo de interação é suportado por ferramentas especializadas.

Um modelo multimídia ideal deveria permitir a especificação de todos estes tipos de interação. Além disso, afim de simplificar a tarefa de estruturação conceptual de um documento multimídia, um modelo deveria fornecer uma abordagem uniforme de representação dos componentes, das relações lógicas e temporais, e dos mecanismos de interação.

### 6.5 Abordagens para Autoria de Documentos Multimídia

Existem várias ferramentas de autoria. Uma lista completa com links para as diversas ferramentas de autoria comerciais e de domínio público pode ser encontrada em [FAQ, 97]. Além destas, existem outras ferramentas que são encontradas na literatura acadêmica que exploram abordagens mais inovadoras. Esta apostila trata destas duas categorias de ferramentas de autoria.

Vários paradigmas básicos de autoria são suportados pelos sistemas de autoria multimídia. Esta apostila apresenta alguns destes paradigmas ou abordagens.

#### 6.5.1 Linguagens Scripting

O paradigma Scripting, ou baseada em linguagens, é o método de autoria no estilo da programação tradicional (Figura 46). Neste caso nós temos uma linguagem de programação que especifica elementos multimídia (por nome de arquivo),

<sup>7</sup> Refere-se a execução de um procedimento em linguagem de alto nível.

seqüenciamento, sincronização, etc. Esta forma de concepção de documento é adotada por [Gibbs, 91], [Herman, 94], [Klas, 90] e [Vazirgiannis, 93].

```
set win=main_win
set cursor=wait
clear win
put background "pastel.gif"
put text "heading1.txt" at 10,0
put picture "gables.gif" at 20,0
put picture "logo.gif" at 40,10
put text "contents.txt" at 20,10
set cursor=active
```

**Figura 46.** Exemplo de Script

Estes modelos têm um poder de expressão muito grande, mas a especificação da composição de um documento multimídia na forma textual é difícil de produzir e modificar ([Ackermann, 94], [Hudson, 93]). Além disso, a composição temporal dos componentes é difícil de identificar.

Os modelos gráficos, vistos mais adiante, têm a vantagem de ilustrar de maneira gráfica a semântica das relações temporais. Este tipo de modelo simplifica a especificação das restrições temporais e reduzem o tempo de criação do documento. Como deficiência, modelos gráficos têm um poder de expressão geralmente menor que os modelos orientados à linguagem. Para os sistemas de autoria, há portanto um dilema acerca de como balancear a facilidade de uso com poder e flexibilidade.

Fazer um software extremamente fácil para aprender e utilizar risca em restringir a ação de autores experimentado ou limitar as possibilidades interativas para o usuário final. Prover grande flexibilidade e poder risca em tornar o software de difícil manipulação. Uma solução tem sido combinar ferramentas de construção simples, similar aos programas de desenho dirigidos a menus com linguagens scripting. Scripting é a maneira de associar um script, um conjunto de comandos escritos numa forma semelhante a programa de computador, com um elemento interativo numa tela, tal como um botão. Alguns exemplos de linguagens de scripting são Apple HyperTalk para HyperCard, Lingo Macromedia para Director e Asymetrix OpenScript para Toolbook. Esta solução permite aos autores novatos a começar a trabalhar rapidamente em uma apresentação e permite aos autores mais avançados a criar comportamentos personalizados sofisticados.

### 6.5.2 Abordagem Baseada em Informação

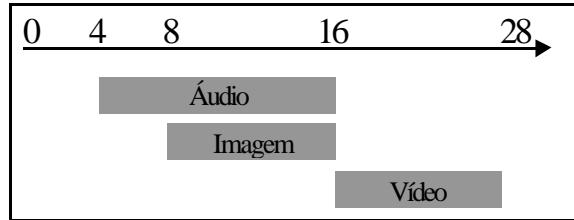
Nesta abordagem, também chamada de centrada na informação, o conteúdo é obtido de informações existentes, se disponível. Estas informações são então estruturadas e armazenadas em uma base de dados. A estruturação envolve a divisão da informação em nós e identificadores chaves e em seguida a ligação destes conceitos [Ginige, 95].

O WWW (World Wide Web) é um exemplo de um sistema centrado em informação. O autor primeiro cria o texto e outras mídias, então ele estrutura estes usando um editor HTML (*Hypertext Markup Language*). Ligações para outros documentos são embutidos no documento durante o processo de Markup. Usuários podem ver este documento usando um sistema de apresentação (Browser) tal como NCSA Mosaic, Nescape Navigator ou Internet Explorer.

Multimedia Viewer da Microsoft é outro exemplo de um sistema combinado de autoria e apresentação baseado na abordagem centrada na informação.

### 6.5.3 Linha temporal (Timeline)

A linha temporal permite o alinhamento das apresentações em um eixo temporal. A Figura 47 apresenta um exemplo de linha temporal. Ela indica que um áudio será apresentado de 4 até 16 unidades de tempo, uma imagem será apresentado de 8 até 16 unidades de tempo, e finalmente um vídeo será apresentado de 16 até 28 unidades de tempo.



**Figura 47.** Exemplo de Linha Temporal

Esta técnica oferece como principal vantagem uma grande simplicidade de expressão dos esquemas de sincronização. Além disso, o autor tem uma visão muito clara das informações que serão apresentadas e em que momento. Infelizmente ela apresenta várias limitações:

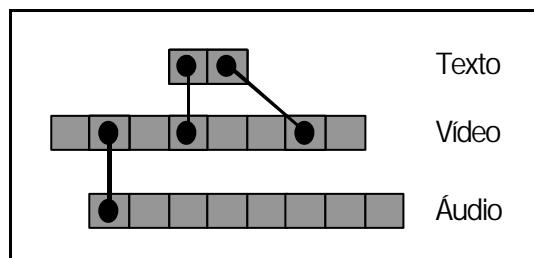
- Ela permite somente a especificação do alinhamento temporal ideal das apresentações, na medida que ela define pontos de partida e fim ideais das apresentações dos componentes do documento. Não é possível especificar métodos de tolerância da sincronização ou ações a serem realizadas.
- Ela requer o conhecimento exato da duração das apresentações. Geralmente, autor deve manipular os segmentos de informação de maneira manual afim de obter o comportamento temporal desejado. Isto através da edição dos dados ou pela mudança da velocidade de apresentação.
- Como esta abordagem não fornece mecanismos de estruturação, nem a representação das relações lógicas, ela não permite a definição da estrutura conceptual completa de documentos hipermídia e multimídia interativos.

Alguns modelos multimídia e hipermídia utilizam versões estendidas da abordagem linha temporal, permitindo a definição de relações lógicas e temporais entre apresentações (por exemplo, a abordagem proposta por [Hirzalla, 95]). Em todo caso, os modelos resultantes destas extensões tornam-se geralmente complexos e perdem a vantagem principal do modelo linha temporal de base que é a simplicidade de compreensão do comportamento temporal do documento.

O sistema MAEstro [Drapeau, 91], QuickTime [Apple, 91], Macromedia Director (Macintosh e Windows), Animation Works Interactive (Windows), MediaBlitz! (Windows), Producer (Macintosh e Windows), e a norma HyTime [Newcomb, 91] adotam este paradigma.

#### 6.5.4 Modelo de composição via pontos de referência

Na abordagem de composição via pontos de referência, as apresentações são vistas como seqüências de sub-unidades discretas [Blakowski, 92]. A posição de uma sub-unidade (p.e. um quadro de um vídeo ou uma amostragem de áudio) em um objeto é chamado de **ponto de referência**. As mídias discretas (p.e. texto) apresentam somente dois pontos de referência: início e fim de apresentação. A sincronização é definida por conexões entre pontos de referência. A Figura 48 ilustra a composição via pontos de referência.



**Figura 48.** Sincronização via Pontos de Referência

Uma vantagem do conceito de ponto de referência é a facilidade de mudança da velocidade de apresentação [Haindl, 96], porque não existe referência explícita ao tempo. As limitações desta abordagem são as seguintes:

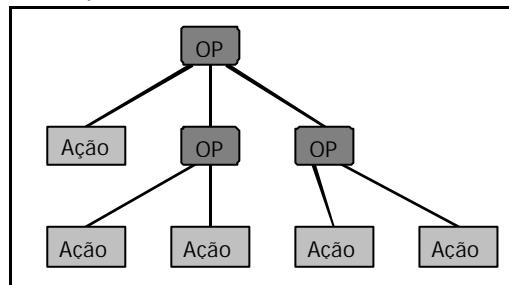
- As sincronizações são definidas por um conjunto de conexões entre pontos de referência. Quando da apresentação de um documento, as irregularidades de apresentação são ignoradas. Não existe a possibilidade de tolerância a este nível.
- A utilização única de uma abordagem de composição via pontos de referência não permite a especificação de retardos, nem a definição de sincronizações condicionais outras que as conexões entre os pontos de referência.
- Esta abordagem não é apropriada para a descrição da estrutura conceptual dos documentos multimídia. Isto pois ela não propõe nenhum mecanismo de estruturação e não há a distinção entre o componente e seu conteúdo.

### 6.5.5 Composição Hierárquica

Na composição hierárquica de documentos multimídia, os componentes de um documento são organizados na forma de uma árvore, como ilustrado na Figura 49. Nesta árvore, os nós internos representam relações temporais entre as sub-árvore de saída. Estas relações temporais são definidas por operadores de sincronização. Os principais operadores de sincronização são os operadores série e paralelo. Estes operadores definem que um conjunto de ações serão executados em série ou paralelo. Estas ações podem ser atômicas ou compostas [Blakowski, 92]: uma ação atômica manipula a apresentação de uma informação, a interação com o utilizador ou um atraso; as ações compostas são combinações de operadores de sincronização e ações atômicas.

Uma das vantagens desta abordagem é a possibilidade agrupar itens em "mini-apresentações" que podem ser manipuladas como um todo [Hardman, 95]. Esta abordagem apresenta também algumas limitações:

- Ela não permite a representação de relações temporais fora das fronteiras hierárquicas. Assim, os links hiper-estruturais não podem ser especificados.
- As relações lógicas e temporais não são representadas de uma maneira natural. Ao contrário da abordagem linha temporal, o autor não tem a visão do conjunto das informações apresentadas e suas datas de apresentação.
- Um conjunto de ações podem ser sincronizadas apenas com relação ao início ou fim de um conjunto de ações. Isto significa que, por exemplo, a apresentação de legendas em uma certa parte de um vídeo requer o corte da seqüência de imagens em vários componentes consecutivos.



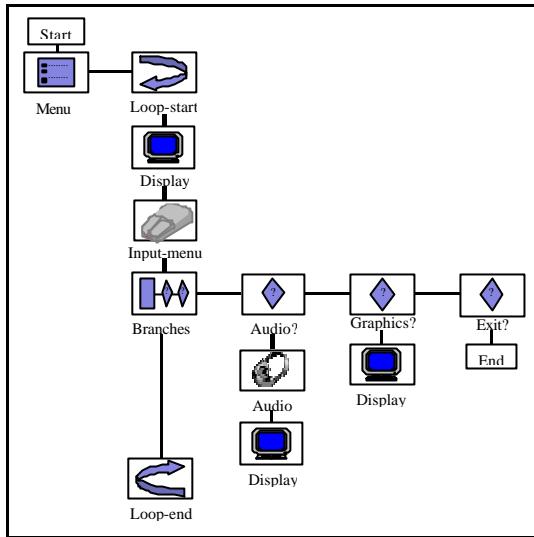
**Figura 49.** Composição Hierárquica

### 6.5.6 Modelos baseados em ícones

A criação de um documento multimídia pela utilização de uma abordagem baseada em ícones é similar a sua programação (por exemplo, utilizando a linguagem C), mas com a ajuda de uma interface gráfica. Esta interface fornece ícones de alto nível, afim de visualizar e manipular a estrutura do documento.

Um conjunto de ícones é arranjado em um grafo que especifica interações e caminhos de controle de apresentação de um documento multimídia. Em geral, as funcionalidades associadas a cada ícone podem ser modificadas utilizando menus e editores associados.

A Figura 50 apresenta um exemplo de utilização desta abordagem. Nesta aplicação, o leitor pode escolher, a partir de um menu, entre a apresentação de um áudio, a apresentação de uma imagem e o término da aplicação. Ao fim da apresentação do áudio ou da imagem, o menu é apresentado novamente.



**Figura 50.** Exemplo do uso de ícones para a autoria de documentos multimídia

Esta abordagem é de utilização simples para pequenas aplicações. Mas para aplicações complexas, a compreensão e a manipulação são dificultadas.

A abordagem de composição de documentos multimídia baseada em ícones é adotada, por exemplo, pelos softwares AimTech IconAuthor [AimTech, 93], Eyes M/M [Eyes, 92], Authorware Professional [Authorware, 89], mTropolis e HSC Interactive.

### 6.5.7 Modelos Baseados em Cartões ou Páginas

Neste paradigma, os elementos são organizados em páginas de um livro ou uma pilha de cartões. Ferramentas de autoria baseadas neste paradigma permitem que o autor ligue as páginas ou cartões formando uma estrutura de páginas ou cartões.

Sistemas de autoria baseados em cartões ou páginas fornecem um paradigma simples para organizar elementos multimídia. Estes tipos de ferramentas de autoria permitem que o autor organize os elementos em seqüências ou agrupamentos lógicos tal como capítulos e páginas de um livro ou cartões em um catálogo de cartões.

Sistemas de autoria baseados em páginas são orientados a objeto [Apple, 94]: objetos são botões, campos de texto, objetos gráficos, fundos, páginas e cartões, e mesmo o projeto em si. Cada objeto pode conter um script, ativado quando ocorre um evento (tal como um clique no mouse) relacionado ao objeto.

Exemplos de ferramentas de autoria adotando este paradigma incluem: HyperCard (Macintosh), SuperCard (Macintosh), ToolBook (Windows) e VisualBasic (Windows).

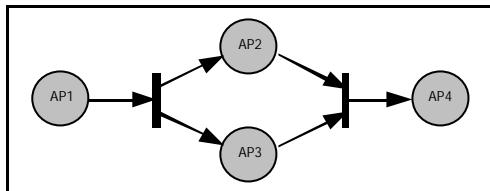
### 6.5.8 Redes de Petri

Rede de Petri é uma técnica de descrição formal muito utilizado na engenharia de protocolos, automação industrial e muitas outras áreas. Ela permite a realização de uma especificação formal de um sistema e permite também a aplicação de técnicas de análise sobre o sistema afim de verificar certas propriedades [Murata, 89].

Redes de Petri também se presta para a modelização de documentos multimídia, isto devido a sua representação gráfica, a facilidade de modelização dos esquemas de sincronização, e a possibilidade de analisar propriedades importantes do comportamento lógico e temporal do documento. Existem vários modelos multimídia baseados em redes de Petri. Uma lista não exaustiva inclui: OCPN [Little, 90], TSPN [Diaz, 93], RTSM [Yang, 96], Trellis [Stotts, 90], HTSPN [Sénac, 95], MORENA [Botafogo, 95], PHPN [Wang, 95]. Nestes modelos, geralmente um lugar da rede de Petri representa uma apresentação, as transições e arcos definem relações lógicas e temporais.

A Figura 51 ilustra a utilização de redes de Petri para especificar um cenário multimídia. Esta especificação indica que inicialmente AP1 será apresentada; terminada esta

apresentação, AP2 e AP3 serão apresentadas em paralelo; tão logo estas duas apresentações terminem, AP4 será apresentada.



**Figura 51.** Composição usando Redes de Petri

Esta abordagem apresenta algumas desvantagens:

- De maneira similar a composição hierárquica, na maior parte dos modelos baseados em redes de Petri, um conjunto de ações podem ser sincronizadas com relação ao início ou fim de um conjunto de ações.
- Esta abordagem não contém os conceitos apropriados para a representação das sincronizações condicionais mais complexas nem para os métodos de interação de aplicação (seção 6.4.4).
- Em [Hudson, 93], os autores afirmam que uma outra limitação desta abordagem é sua complexidade: este formalismo tende a ser muito potente e de muito baixo nível para sua utilização direta pelo autor do documento.

A grande vantagem do uso de um modelo de autoria baseado em Redes de Petri é o caráter formal desta técnica de descrição. A especificação de documentos multimídia utilizando modelos informais ou semi-formais podem ser fontes de ambigüidades e geralmente não permite a utilização de ferramentas de análise avançadas. Ao contrário do anterior, modelos multimídia baseados em métodos formais, como aqueles baseados em Redes de Petri, permite a construção de uma semântica precisa e não ambígua de um documento e eles permitem a utilização de técnicas de análise sofisticadas, como a simulação, validação, verificação do comportamento do documento.

Nem todos os modelos ditos baseados em redes de Petri apresentados na literatura podem ser considerados modelos formais. Assim, alguns destes modelos não permitem a aplicação de técnicas de análise para validar o comportamento do documento multimídia.

## 6.6 Apresentação de Documentos Multimídia

A partir da descrição completa do documento, um sistema pode utilizar duas abordagens diferentes para apresentá-lo: via uma interpretação direta da descrição lógica do documento; ou via uma tradução dessa descrição lógica em uma representação física que será em seguida armazenada, transferida e apresentada. As representações físicas podem ser classificadas em dois tipos:

- Dedicadas a um certo sistema de apresentação. Neste caso o sistema de apresentação possui um formato particular para representar o documento. Este é o caso do sistema de autoria IconAuthor.
- Padronizadas de fato (por exemplo, os ISO MHEG [ISO 13522], HyTime [Newcomb, 91], HyperODA [Appelt, 93]) ou de direito (por exemplo, Java [Sun, 95]).

Para esta última abordagem, a adoção de normas internacionais de representação de informações multimídia e hipermídia como modelo físico é recomendado. Existem basicamente 3 normas de representação de documentos que são as normas ISO MHEG [ISO 13522], ISO HyTime [Newcomb, 94] e HyperODA [Appelt, 93]. Estas normas serão vistas mais adiante. A adoção da representação ISO MHEG como representação física do documento, por exemplo, permite que documentos possam ser transferidos em um sistema aberto e apresentados via interpretadores da representação normalizada ou traduzidos em uma representação específica. Além destas representações normalizadas por órgãos de normalização internacionais, existe uma representação portável que vem

sendo muito utilizada nos dias de hoje: a linguagem Java. Ela pode ser considerada uma normalização de fato para representar aplicações multimídia.

Na primeira abordagem, a utilização de representações proprietárias implica na utilização de sistemas de apresentação proprietários (uma solução ad-hoc). Neste caso, afim de permitir a apresentação de um documento em um sistema aberto, este sistema deve suportar a tradução da representação proprietária do documento para um formato normalizado de representação de informações multimídia e hipermídia, ou então, o sistema de apresentação deve ser portável.

## 6.7 Padronização de Formatos de Documentos Multimídia

A maioria dos sistemas multimídia/hipermídia são baseados em modelos e conceitos diferentes. Assim, informações multimídia e hipermídia de um sistema não são facilmente reutilizadas por outros sistemas. Estes sistemas são chamados de sistemas fechados.

As informações utilizadas, armazenadas e transferidas por aplicações multimídia representam um investimento importante, e portanto é vital que estas informações sejam utilizáveis em um mundo em rápida evolução. Para isto é necessária a definição de uma linguagem comum para todos os sistemas de comunicação multimídia/hipermídia. Portanto é necessária a definição e utilização de padrões internacionais de formatos de representação e de transferência de informações multimídia e hipermídia. A adoção de normas internacionais permitem a criação de sistemas abertos, que podem se comunicar com outros sistemas abertos. Neste caso estas informações multimídia/hipermídia são ditas portáveis, podendo ser tratada por diversos sistemas.

A padronização ao nível monomídia (apresentada no capítulo 3 desta apostila), como JPEG para imagens fixas e MPEG para áudio e vídeo, não é suficiente para possibilitar a portabilidade, isto pois as aplicações multimídia necessitam informações adicionais para a apresentação destes dados (como a identificação de algoritmos de codagem, atributos a utilizar quando da apresentação, etc.) e informações sobre as interrelações entre os dados multimídia. Para definir estas informações adicionais foram estabelecidos novos padrões do tipo ISO SGML, ODA, MHEG e HyTime. Estas normas são ditas padrões de direito, que são estabelecida por um órgão internacional de padronização, o ISO (*International Standardization Organization*).

Como em muitas outras áreas, na multimídia também surgiram algumas formas de representação de informações multimídia e hipermídia instituídas por empresas ou grupos de empresas que ao longo do tempo foram adotadas por muitos consumidores e empresas. Estes são chamados padrões de uso, entre eles temos a linguagem Java.

### 6.7.1 Standard Generalized Markup Language (SGML)

ISO SGML (ISO 8879) é uma linguagem para a descrição da estrutura conceptual inicialmente de documentos textuais. Ela é uma linguagem aberta que define uma sintaxe geral para desempenhar o *mark-up* e permite a definição de tipos de documentos. SGML usa *mark-up* permitindo que o usuário anote documentos através de *tags* que definem o formato do documento, sua composição em termos de capítulos, seções, índices, referências, etc. Este *mark-up* especifica que parte do documento segue a que componentes do documento, ou elementos SGML, e é atribuído valores a variáveis, ou atributos SGML, que cada um destes elementos tem (veja Figura 52). Ela também pode atribuir um elemento a um identificador único de forma que ele possa ser referenciado em qualquer parte no documento.

SGML é uma meta-linguagem não utilizada diretamente pelo usuário, mas permitindo a definição da linguagem que será realmente empregada pelo usuário. Cada documento SGML usa uma definição de tipo de documento (DTD) que declara que tipos de elementos podem existir no documento, que atributos cada um destes tipos de elemento

podem ter, e como instâncias destes tipos de elementos podem ser hierarquicamente relacionadas. Tipicamente, um DTD define uma classe completa de documentos, e muitas instâncias de documentos compartilham um DTD comum. Um tipo particular de DTD é o HTML (*HyperText Markup Language*), para a composição de páginas Web.

```
Exemplo de SGML DTD
<!element book      - o (chapter*)>
<!element book title cdata #required>
<!element chapter   - o (title, section)>
<!element title     - o (#pcdata)>
<!element section    - o (header, paragraf*)>
<!element header    - o (#pcdata)>
<!element paragraf   - o (#pcdata) +(newterm)>
<!element newterm    - - (#pcdata)>

Documento SGML
<!doctype book system "book.dtd">
<book title = "How to build a sink">
  <chapter>
    <title>Building the faucet
    <section>
      <header>Building the handles
      <paragraf>The handles are hard to build.
      <paragraf>You must use <newterm>washers</newterm>.
```

**Figura 52.** Exemplo de Documento SGML e DTD

### 6.7.2 Office Document Architecture (ODA)

A arquitetura de Documento de Escritório e Formato de Transferência (ODA) é um padrão de documento definido pela ISO sob a referência ISO 8613. ODA é também endossado pela ITU (*International Telecommunication Union*) sob a referência T.410. ODA foi definido no contexto de aplicações de escritório. Como resultado, ela objetivou inicialmente a transferência de documentos de escritório, do tipo cartas, relatórios, memorandos, recibos, e mais geralmente documentos mistos (textos e gráficos) produzidos por processadores de texto.

O padrão define uma estrutura conceptual do documento – que é a organização em capítulos, seções – tão bem quanto a estrutura de apresentação do documento. Em sua forma inicial, o padrão ODA apenas define as mídias texto, gráfico e imagem. Informações do tipo som, vídeo não eram suportadas. Trabalhos foram iniciados em 1993 para estender o ODA (chamado HyperODA), que resultou em 1995 no suporte de sons, referências externas, relações temporais entre componentes do documento.

### 6.7.3 Hypermedia/Time-based Structuring Language (HyTime)

O HyTime [Newcomb, 91] é um padrão ISO (ISO 10744) para a representação de documentos multimídia. Este padrão especifica como certos conceitos universais de documento multimídia podem ser representados usando SGML. Ela é portanto uma linguagem *mark-up*, mais precisamente uma meta-linguagem para a definição de tipos de documentos multimídia. HyTime é uma extensão do SGML. Os tipos de documentos são definidos pela Definições de Tipos de Dados (DTD). Por exemplo, uma DTD pode definir a estruturação do documento em capítulos, seções. Cada DTD requer um programa de apresentação da aplicação HyTime.

HyTime permite a especificação da estrutura conceptual do documento: suas relações espaciais, relações temporais (usando um modelo temporal baseado em timeline), estrutura lógica, e apontadores entre elementos dentro e fora do documento. HyTime não permite a definição das características de apresentação nem dos modelos de interação. Estas funcionalidades podem ser descritas via construtores não HyTime, sendo que o software de apresentação deve conhecer estes construtores.

### 6.7.4 Multimedia and Hypermedia Expert Group (MHEG)

MHEG é um padrão ISO (ISO 13522) que define um formato padrão de transferência de informações multimídia e hipermídia. Este padrão permite representar a estrutura do conteúdo, de apresentação e conceptual de um documento multimídia/hipermídia para fins de transferência do documento. Portanto ele permite a transferência de vários tipos de mídias, a transferência das estruturas representando a composição temporal e espacial do documento, e as interações quando da apresentação do documento.

MHEG é baseado em uma abordagem orientada à objetos: as informações são representadas por objetos MHEG. Estes objetos podem conter ou fazer referência aos dados codificados (JPEG, MPEG, etc), podem exprimir relações entre outros objetos MHEG, exprimir o comportamento dinâmico dos objetos e exprimir as informações para otimizar as manipulações tempo real dos objetos

Este padrão é dividido atualmente em 6 partes:

- Parte 1: utiliza ASN.1 (*Abstract Syntax Notation One*) para representar os objetos MHEG. Esta representação não vem alcançando o sucesso esperando, sendo que existem muito poucos trabalhos associados.
- Parte 3: define a representação, codificação e semântica dos objetos Script, para serem intercambiados como objetos parte de uma aplicação MHEG.
- Parte 4: define os procedimentos de registros de objetos e formatos suportados por MHEG.
- Parte 5: uma versão simplificada de MHEG-1 para aplicações simples, possibilitando a criação de um interpretador da representação MHEG-5 com requisitos mínimos de recursos, para ser utilizado em uma *set-top-box* (sistema caracterizado por pequena quantidade de memória e capacidade limitada de processamento).
- Parte 6: estende o padrão MHEG-5 para dar suporte a aplicações interativas avançadas. Para tanto, MHEG-6 representa programas e *applets* usando a codificação bytecode da linguagem Java como formato de intercâmbio de programas.
- Parte 7: especifica testes de interoperabilidade e conformância.

A parte 5 deste padrão, MHEG-5, vem sendo muito utilizado atualmente na área de TV interativa, a tecnologia que nos próximos anos irá substituir os sistemas de TV atuais por sistemas digitais.

## 6.8 Exemplos de Modelos Multimídia e Sistemas de Autoria

Nesta seção são apresentados alguns modelos multimídia e ferramentas de autoria de documentos multimídia.

### 6.8.1 Modelo de Referência Hipertexto Dexter

O modelo de referência hipertexto<sup>8</sup> Dexter [Halasz, 90], definido por um consórcio de fabricantes, busca capturar abstrações importantes encontradas na maior parte dos sistemas hipertexto (incluindo hipermídia).

Dexter é um modelo a 3 camadas (Figura 53):

- A camada *Dentro do Componente* descreve a estrutura interna dos dados primitivos. A definição desta camada não faz parte do modelo Dexter, mas de modelos de referência concebidos especificamente para modelizar a estrutura de aplicações particulares, de documentos ou de tipos de dados.

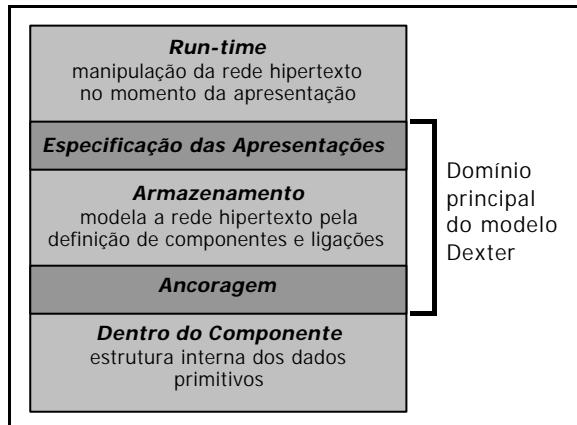
---

<sup>8</sup> No modelo Dexter, os documentos hipertextos englobam os documentos hipermídia.

- A camada *Armazenamento* modeliza a rede hipertexto pela definição de componentes do documento e de seus links. O mecanismo de âncora define a separação entre esta camada e a camada *Dentro do Componente*.
- A camada *Run-time* é responsável pela manipulação da rede hipertexto no momento da apresentação do documento. A interface entre esta camada e a camada *Armazenamento* é realizada pela especificação das características de apresentação dos componentes.

O domínio principal do modelo Dexter é a camada de *Armazenamento*. Nesta camada, a entidade fundamental e a unidade de base de endereçamento é o conceito de componente. Existem três tipos de componentes:

- **Componente Atômico** representa dados primitivos. Cada componente atômico contém um identificador único, o conteúdo (por inclusão explícita) e as informações sobre o componente. Estas informações descrevem as propriedades do componente: uma lista de âncoras que podem ser fonte ou destino de um link; especificação de apresentação, que contém as informações originais de apresentação do componente; e um conjunto de atributos/valores utilizados, por exemplo, para a definição de palavras-chaves do componente.
- **Componente Composto** fornece um meio de estruturação do documento, sendo que um componente composto é uma coleção de componentes atômicos, compostos e link.
- **Links** são as entidades que representam as relações entre os componentes. Um link é definido por um lista de especificadores. Este último especifica o identificador do componente e a âncora que é um ponto final do link, e a direção especificando se o ponto definido é um ponto de origem, destino ou os dois.



**Figura 53.** As camadas do modelo Dexter

A descrição multi-camadas do modelo Dexter não apresenta os mesmo níveis de abstração que a arquitetura de descrição de documentos multimídia/hipermídia apresentados na seção 6.4:

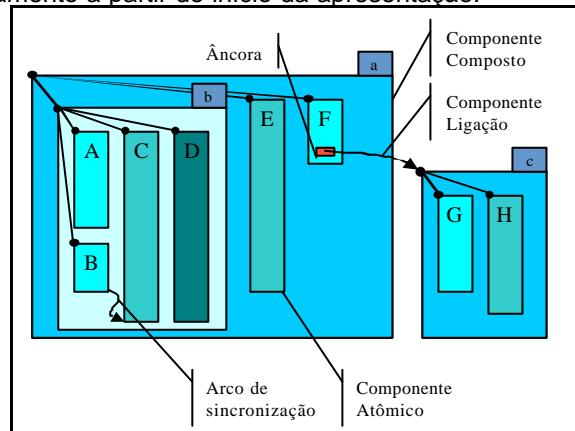
- Os componentes descrevem e contém os dados primitivos, não existe distinção entre a estrutura conceptual e o conteúdo. Assim, os objetos multimídia não podem ser reutilizados.
- A noção de componente composto não fornece um mecanismo de abstração adequado para a representação da estrutura do documento, onde um conjunto de componentes deveria ser tratado como um objeto único. Isto não é o caso do modelo Dexter que não permite a definição de uma âncora preta a um composto (um conjunto de componentes) [Grenbaek, 94].
- A camada de *Armazenamento* do modelo Dexter não fornece os meios de definição das relações temporais e condicionais entre os componentes do documento.

### 6.8.2 Amsterdam Hypermedia Model (AHM)

O AHM [Hardman, 93] é um modelo hipermídia obtido a partir de uma extensão do modelo de referência Dexter e do modelo multimídia CMIF (*CWI<sup>9</sup> Multimedia Interchange Format*) [Bulterman, 91]. Este modelo resolve alguns problemas do modelo Dexter, que são a ausência da noção de tempo na camada *Armazenamento* (impossibilitando a representação das relações temporais entre as apresentações), a ausência de meios de especificar a informação de apresentação de alto nível, e a ausência da noção de contexto para as âncoras.

Contrário ao modelo Dexter, AHM permite a representação das relações temporais entre o conjunto de apresentações que compõem um documento hipermídia. Um exemplo de composição temporal é apresentado na Figura 54. Abaixo são apresentados os elementos utilizados para a descrição desta composição:

- Componente atômico modela os dados, sua duração de apresentação e o canal<sup>10</sup> onde os dados serão apresentados. Assim, no modelo AHM um componente atômico representa o par dados primitivos/apresentação. Além disso, no AHM um componente atômico pode conter os dados codificados (inclusão explícita) ou lhes fazer referência (inclusão implícita). Esta última é utilizada no caso onde existem várias apresentações utilizando a mesma informação, permitindo assim a reutilização de dados primitivos.
- Componente composto representa uma composição lógica e temporal. Os atributos de um componente composto são:
  - o tipo de composição, podendo ser paralelo (todos os componentes são apresentados) ou escolha (um só componente é apresentado);
  - o tempo de partida dos componentes (relativo ao início de apresentação do composto);
  - os arcos de sincronizações que permitem a representação de relações temporais entre dois componentes, definidas por intervalos de sincronização contendo um tempo ideal e uma variação aceitável e o tipo de sincronização. Este tipo indica se o intervalo é relativo ao início ou ao fim da apresentação de um componente, ou ainda se ele se trata de um deslocamento a partir do início da apresentação.



**Figura 54.** Relações temporais no modelo AHM

- Aos links é adicionada a noção de contexto [Hardman, 94]: um contexto fonte para um link é a parte de uma apresentação hipermídia afetada pela ativação de um link, um contexto destino é a parte que é apresentada na chegada ao destino do link. AHM define três de links contexto:
  - Continua, a apresentação fonte não é encerrada quando o link é ativado;
  - Pausa, a apresentação fonte é interrompida, mas quando da retomada da apresentação, ela se faz a partir do ponto de interrupção;

<sup>9</sup> CWI: Centrum voor Wiskunde en Informatica (Holanda)

<sup>10</sup> No AHM, um canal é um dispositivo de saída abstrato utilizado para a apresentação dos componentes. Um canal especifica as características de apresentação por default dos dados apresentados (por exemplo, o formato dos caracteres, o volume).

- Reposição, a apresentação fonte se interrompe, e quando de uma retomada de sua apresentação, ela se fará desde seu ponto inicial.

Como o modelo Dexter, o AHM não segue o modelo multi-nível apresentado na seção 6.4. Os componentes fazem unicamente referência aos dados primitivos. Sendo que apenas os dados primitivos podem ser reutilizados, a descrição destes dados não são reutilizáveis. Além disso, este modelo não é baseado em métodos formais, não existindo um método de análise para verificar a consistência lógica e temporal do documento.

### 6.8.3 Modelo de Composição Hierárquico de [Hudson, 93]

Em [Hudson, 93], os autores propõem uma abordagem hierárquica de composição de documentos multimídia. Nesta abordagem os componentes do documento são representados por objetos primitivos ou compostos (construídos a partir de objetos primitivos e compostos). Estes objetos são representados em quatro partes: o conteúdo, os recursos necessários, uma máquina temporal que controla o tempo de apresentação e um drive de apresentação responsável pela manipulação e apresentação do conteúdo. Estes objetos têm um conjunto de propriedades comuns que são a duração, a reductibilidade (capacidade e método de modificação da duração de apresentação), e a capacidade de controle de fluxo (indica a capacidade de interromper, reformar, acabar, de troca de direção e modificação de velocidade de apresentação).

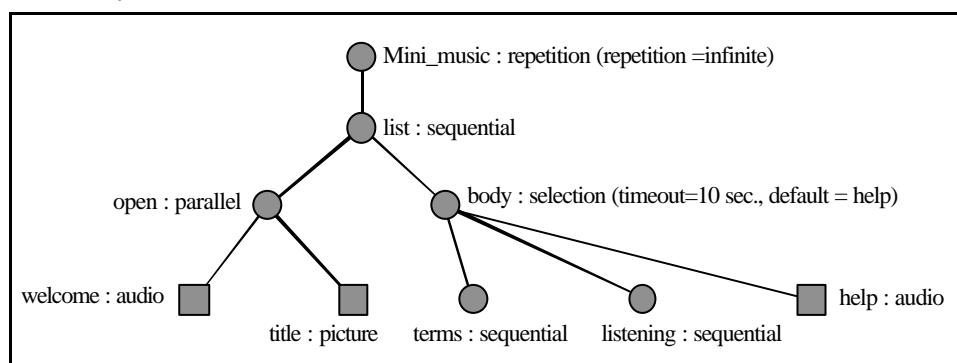
[Hudson, 93] utiliza uma abordagem hierárquica de composição temporal dos objetos que compõem um objeto composto. Este último é estruturado como uma árvore com nós externos ocupados por objetos primitivos e nós internos ocupados por operadores (seqüencial, paralelo, linha temporal, escape, sincronização contínua, condicional, seleção, repetição).

A Figura 55 apresenta um exemplo da utilização desta abordagem. Nesta aplicação, chamada *mini\_music*, inicialmente um áudio e uma imagem são apresentadas em paralelo. Em seguida, o utilizador pode escolher entre estudar alguns termos musicais, escutar músicas ou pedir ajuda. Se o leitor não fizer nada durante 10 segundos, um áudio é apresentado.

Para esta abordagem, apesar de não ser baseada em métodos formais, foram desenvolvidos métodos de análise específicos para garantir a consistência do documento: permitindo realizar uma detecção de conflitos de recursos compartilhados e a verificação das restrições de sincronização.

Esta abordagem permite a descrição das estruturas de conteúdo, conceptual e de apresentação do documento. Mas ela não é multi-nível. Ela não faz distinção entre os objetos multimídia (a estrutura do conteúdo) e os componentes (a estrutura conceptual), e deste fato a reutilização de objetos não é possível.

Esta abordagem força a utilização de uma estrutura unicamente hierárquica (e linear) do documento. Ela não permite a definição das relações temporais e lógicas entre os elementos de diferentes níveis da estrutura. Assim, os links hiper-estruturais não podem ser representados por esta abordagem. Além disso, ela não propõe métodos de tolerância de sincronização necessário quando os dados são armazenados em servidores dispersos em rede.



**Figura 55.** Abordagem de composição proposta por [Hudson, 93]

### 6.8.4 Sistema Firefly

O Firefly ([Buchanan, 92], [Buchanan, 93]) é um sistema utilizado para a construção de documentos hipermídia. Este sistema baseia-se em uma abordagem de composição via pontos de referência. Ele utiliza uma representação gráfica a partir do qual o leitor pode criar documentos hipermídia. Por exemplo, a Figura 56 descreve um curso de eletricidade e magnetismo. Esta representação gráfica é baseada sobre as linhas temporais. Cada item do documento tem sua própria linha temporal. Os itens com uma duração imprevisível são representadas por linhas pontilhadas. Os nós do grafo representam eventos: os retângulos representam os eventos de início e fim (assim cada item é representado por dois retângulos conectados); os círculos representam eventos internos. As restrições temporais entre estes eventos são representados por arcos tipados que ligam os eventos.

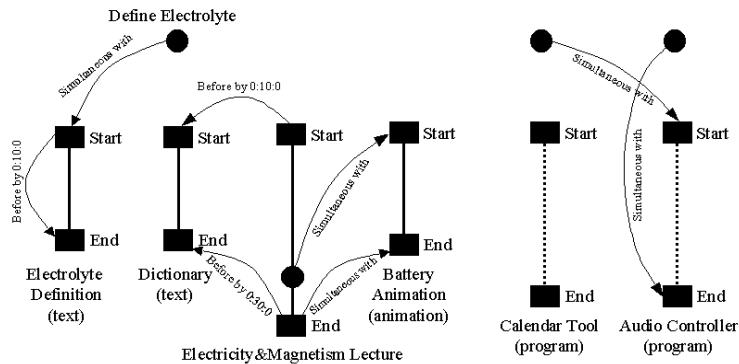


Figura 56. Abordagem de composição proposta por [Hudson, 93]

### 6.8.5 Asymetrix Toolbook

No sistema de autoria Asymetrix Toolbook, o autor cria um “livro” que consiste de uma série de páginas. Cada página pode ter um ou mais objetos, tal como botões, gráficos, e texto (Figura 57). A interatividade é fornecida pela associação de um script com um objeto que o leitor pode selecionar. A linguagem scripting Toolbook, chamada OpenScript, é dita orientada a objetos no sentido em que scripts são associados a objetos e são ativados utilizando passagem de mensagens. De fato, OpenScript não é orientada a objetos (não fornece classe, herança, encapsulação, etc.).

Toolbook opera de duas maneiras: *author* e *view*. Quando no modo *author*, objetos podem ser criados e editados usando menus e ferramentas Toolbook. Enquanto no modo *view*, o autor pode executar a aplicação e testar seu comportamento.

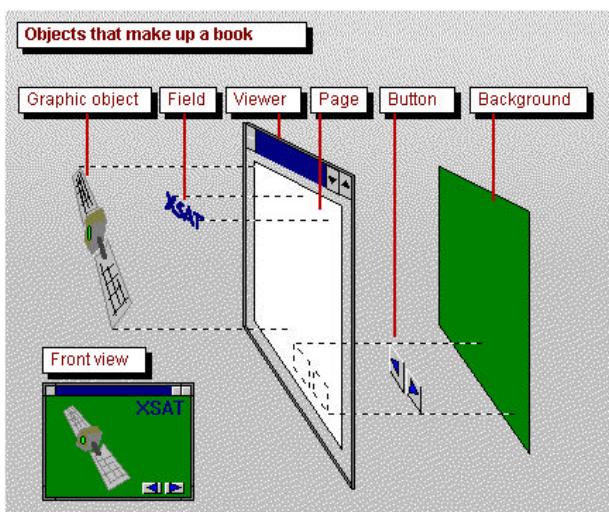


Figura 57. Asymetrix Toolbook

## 6.8.6 Sistemas Baseados em Ícones IconAuthor e Authorware

### **Asymetrix IconAuthor**

No sistema de autoria Asymetrix IconAuthor [Assimetix, 99], a área principal de trabalho, apresentada na Figura 58 contém um simples grafo de ícones. Os ícones neste grafo são instâncias dos ícones do catálogo de ícones disponíveis ao autor (no lado esquerdo). Os parâmetros associados com os ícones são editáveis via caixas de diálogo, que são apresentados quando o ícone é selecionado. Existem editores para a criação de textos, gráficos, e animações.

Este ambiente não apresenta linguagem scripting em si, mas usa uma abordagem de programação visual para definir interações e fluxo de controle. O conjunto de ícones corresponde encontradas nas linguagens de programação convencionais, com ícones adicionais para apresentação de diferentes dados multimídia e mecanismos de interação. Portanto IconAuthor é interessante para não programadores, que não estão habituados com linguagens de programação. Além disso, neste sistema de autoria o autor define de forma procedural a aplicação, e não de forma orientada a eventos como no ToolBook. Isto torna a autoria mais simples para não programadores.

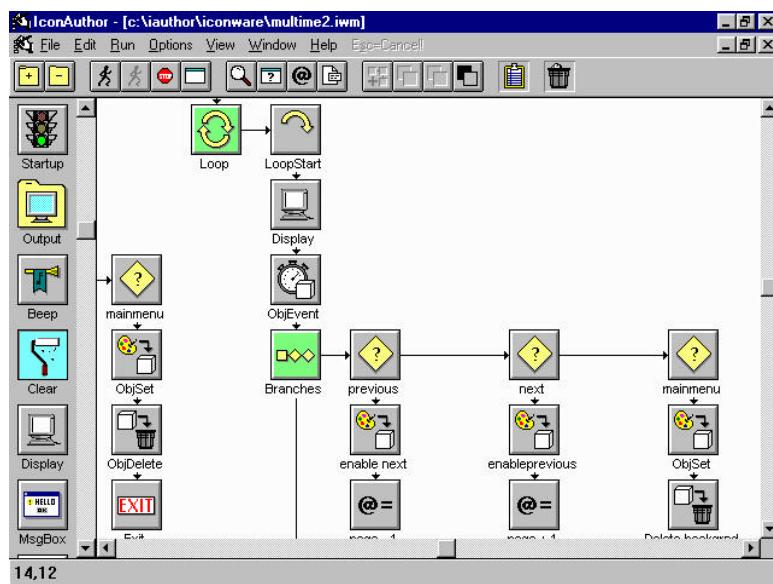


Figura 58. AimTech IconAuthor

### **Authorware Professional**

Authorware Professional [Authorware, 89] usa também uma interface de programação visual baseada em ícones (Figura 59) caracterizada por um pequeno mas poderoso conjunto de ícones. Este software controla a complexidade do grafo pelo fornecimento de um pequeno número de tipos de ícones e limitando o número de ícones que aparecem em uma janela. Esta última força o autor a construir composições hierarquicamente. O grafo estruturado resultante pode ser facilmente navegado e entendido do mesmo modo que um grafo plano.

Uma comparação detalhada entre os sistemas de autoria IconAuthor e Authorware é feita em [Koegel, 93].

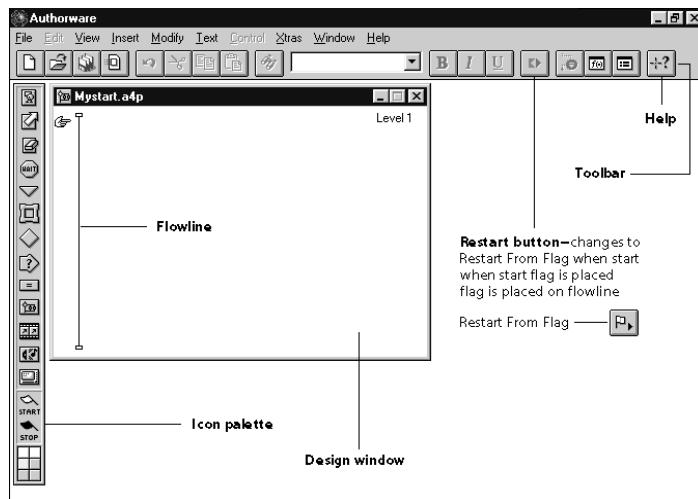


Figura 59. Authorware

### 6.8.7 Macromedia Director

O Director usa a linha temporal (*timeline*) como abordagem de autoria, que no Director é conhecida como “filme”. A metáfora filme consiste de:

- Membros do elenco (*cast members*): que atuam no filme, p.e. gráficos, animações, vídeos, textos, áudios, botões, etc. São gerenciados a partir da Janela Cast, mostrada na Figura 60, que cria um banco de dados dos membros. Uma vez definidos os membros podem ser colocados no score.

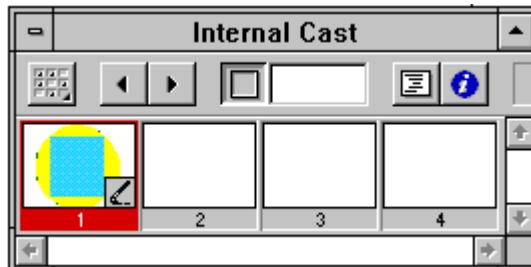


Figura 60. Janela Cast

- Cenário (*stage*): é a área onde a apresentação visual do filme acontece. Membros do elenco podem ser arrastados e posicionados no Cenário.
- Score: dita o movimento geral dos membros do elenco no tempo. Ele é descrito por uma janela score (Figura 61), que contém uma matriz de células. As colunas da matriz representam frames do filme. Cada frame representa um instante de tempo de uma duração definida. As linhas representam camadas no cenário onde os membros do elenco podem aparecer. Uma célula pode conter scripts, efeitos especiais, instruções temporais, paleta de cores, controle de som.

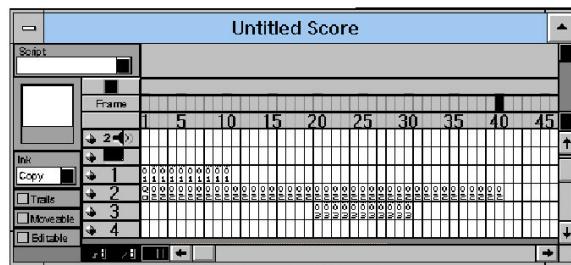


Figura 61. Director 6.0

Director usa uma linguagem scripting orientada a objeto, chamada de Lingo, para aumentar o poder da metáfora filme. Cada membro do elenco pode ter um script associado. Objetos podem capturar eventos e modificar o fluxo de controle da aplicação através de um salto para um quadro particular.

### 6.8.8 HTSPN (Hierarchical Time Stream Petri Nets)

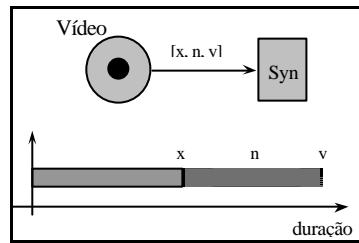
O Modelo HTSPN [Sénac, 95] repousa sobre os conceitos estabelecidos pelo modelo de referência Dexter [Halasz, 94] e sobre o modelo TSPN [Diaz, 93] para a especificação da estrutura conceptual de documentos hipermídia. Ou seja, ele permite a especificação formal dos componentes e de suas relações lógicas e temporais e a análise do comportamento.

Como HTSPN é um modelo formal, ele permite a aplicação das seguintes técnicas de análise [Sénac, 96]:

- a verificação da consistência temporal das partes mais críticas dos documentos hipermídia, que são os cenários multimídia;
- a simulação das atividades dinâmicas do documento hipermídia modelado, isto graças ao conceito de ficha de redes de Petri;
- como uma HTSPN pode ser traduzida numa rede de Petri temporal (com intervalos temporais nas transições) todas as técnicas de análise deste último modelo podem ser utilizadas afim de validar a especificação.

Este modelo é na realidade é uma extensão do modelo rede de Petri a arco temporizado, onde (Figura 62):

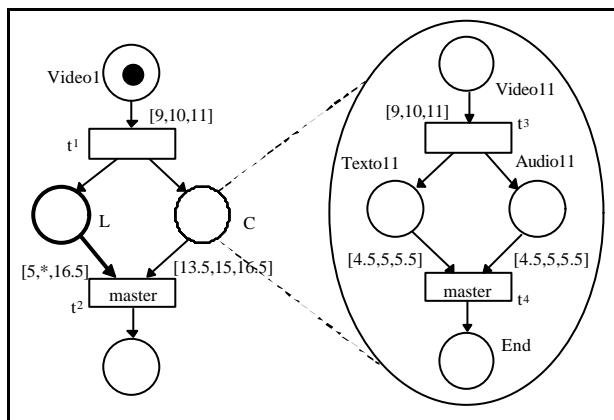
- Lugares representam os componentes do documento,
- Intervalos de validade temporais (IVT) representam as restrições temporais admissíveis dos tratamentos,
- Transições tipadas definem as regras de sincronização.



**Figura 62.** Principais elementos do modelo HTSPN

Os lugares do modelo HTSPN também são tipados (Figura 63):

- **Atômicos:** arcos associados a IVT e lugares do tipo atômico,
- **Ligações:** arcos temporizados ( $L, t_2$ ) onde  $L$  é um lugar do tipo ligação,
- **Compostos:** lugares abstratos do tipo composto, associado a uma sub-redes.



**Figura 63.** Tipos de Componentes

Neste modelo as transições também são tipadas permitindo a definição de estratégias de sincronização que envolvem um conjunto de apresentações (Figura 64).

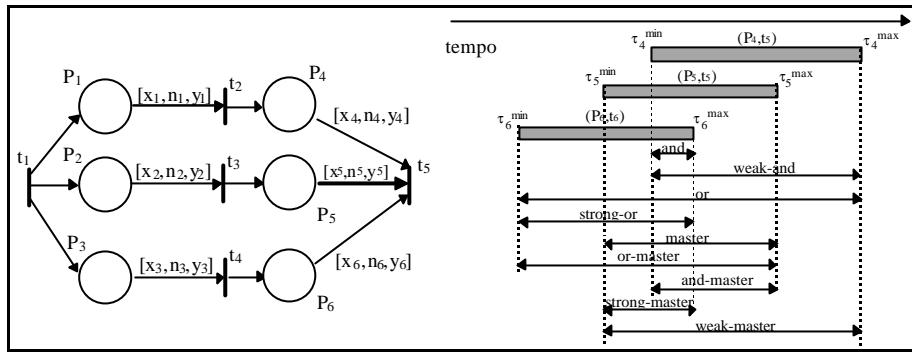


Figura 64. Semântica de sincronização do modelo HTSPN

### Modelo I-HTSPN

Em [Willrich, 96a], os autores propõem uma extensão ao modelo HTSPN afim de permitir a especificação completa de documentos hipermídia a partir da especificação das estruturas conceptuais, de apresentação e do conteúdo. Esta extensão é uma interpretação do modelo HTSPN, chamada de I-HTSPN (*Interpreted-HTSPN*). Este modelo é interpretado no sentido que são associadas semânticas particulares ao lugares do modelo HTSPN: são associados a eles informações de acesso e manipulação dos dados que constituem o componente (estrutura do conteúdo) e são associadas as características de apresentação do componente (estrutura de apresentação).

O modelo I-HTSPN permite a especificação multinível de um documento multimídia/hipermídia, incluindo as estruturas do conteúdo, de apresentação e conceptual. A Figura 65 apresenta a estrutura multi-nível do modelo I-HTSPN e as interfaces entre estes níveis. Nesta estrutura, a estrutura conceptual é especificada utilizando o modelo HTSPN (uma rede de Petri estruturada), a estrutura do conteúdo é especificado por um conjunto de objetos da classe *Data* (que especifica informações de acesso e manipulação dos dados), e a estrutura de apresentação é especificada por um conjunto de objetos *Presentation* (especifica as características de apresentação dos componentes) e *Channel* (especifica o dispositivo de saída onde será feita a apresentação).

Em [Willrich, 96a], os autores propuseram uma versão interpretada do modelo HTSPN orientada para a geração de representações MHEG. Os documentos hipermídia assim representados podem ser transferidos em um sistema aberto e apresentados a partir de um interpretador MHEG. Este último é um software que interpreta (apresenta) representações MHEG.

Em [Willrich, 97], os autores propõem uma nova extensão do modelo HTSPN, chamada Java IHTSPN, permitindo a concepção formal de aplicações Java. Nesta nova extensão, somente os atributos dos objetos *Data* e *Presentation* foram modificados afim de que as informações de acesso aos dados e as características de apresentação definidas por estes objetos tenham elementos correspondentes na linguagem Java.

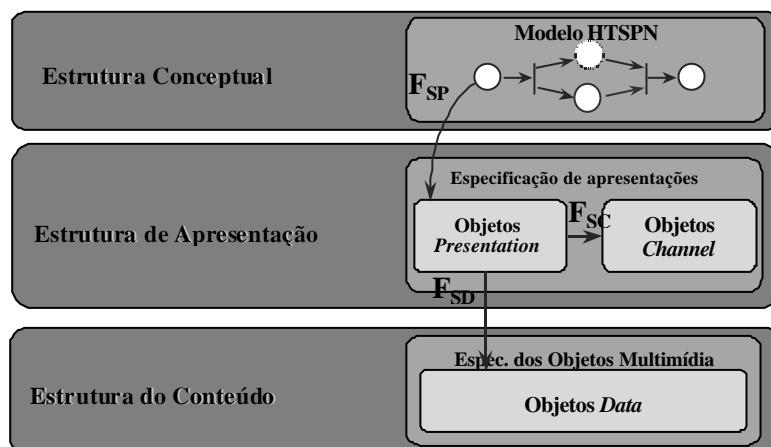
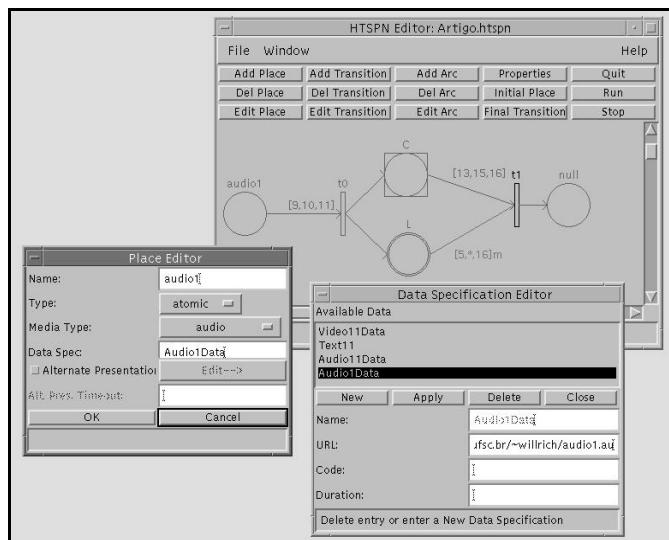


Figura 65. Modelo Multi-nível I-HTSPN

Um ambiente Java, chamado *ambiente Java I-HTSPN*, está sendo implementado afim de auxiliar o processo de criação, de análise e de apresentação de documentos multimídia interativos. Baseado no modelo Java I-HTSPN, este ambiente implementa uma metodologia de desenvolvimento de documentos multimídia comportando as seguintes etapas:

- Utilização do modelo Java I-HTSPN para a especificação formal do documento multimídia interativo. Este modelo fornece os meios para uma descrição precisa e não ambígua das estruturas conceitual, de apresentação e do conteúdo do documento. A Figura 66 mostra o Editor Java IHTSPN, implementado em Java, que fornece ao autor os meios de construir, de maneira gráfica, o documento multimídia interativo.
- Simulação da especificação HTSPN, afim de verificar a correção do comportamento lógico e temporal do documento. Se o comportamento não é aquele desejado, o autor deve retornar à primeira etapa;
- Aplicação de técnicas de verificação suportadas pelo modelo I-HTSPN. Estas técnicas são: a verificação da consistência temporal dos cenários multimídia modeladas e a aplicação das técnicas desenvolvidas para o modelo redes de Petri temporais (após a tradução da HTSPN em uma rede de Petri temporal). Se existem erros, o autor deve retornar à primeira etapa;
- Teste do documento multimídia pela interpretação de sua especificação Java + HTSPN analisada. Nesta fase, o autor pode visualizar e interagir com documento multimídia em desenvolvimento. Se o comportamento desta apresentação não é aquele desejado, o autor deve retornar à primeira etapa da metodologia.
- Produção automática de uma aplicação multimídia Java implementando o documento multimídia interativo.



**Figura 66.** Interface do Editor I-HTSPN

## Capítulo 7 Requisitos de Comunicação para Sistemas Multimídia Distribuídos

Neste capítulo, nós identificaremos os principais requisitos de rede de comunicação para transmissão de áudio e vídeo. Inicialmente, na seção 7.1, serão definidos alguns parâmetros de desempenho de redes de computadores que são importantes para a comunicação multimídia. Em seguida, a seção 7.2 apresenta uma caracterização das fontes de tráfego multimídia. Na seção 7.3, os principais requisitos de rede para a comunicação de áudio e vídeo serão identificados.

### 7.1 Parâmetros de Desempenho de Redes

Como nós vimos nos capítulos iniciais, existem parâmetros de desempenho chaves para a comunicação multimídia. Nós vimos que a taxa de bits de uma rede é uma característica de rede crucial. Nesta seção nós definiremos outros parâmetros importantes para a multimídia.

#### 7.1.1 Velocidade da Rede, Largura de Banda e Taxa de bits

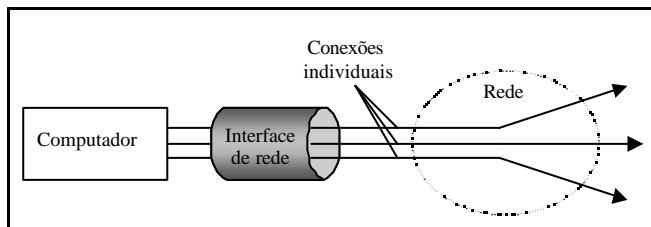
Embora a palavra mais usada para se referenciar a capacidade de transmissão da rede seja *velocidade*, o termo mais apropriado é *largura de banda*. A velocidade da luz em uma fibra não é muito diferente da velocidade de propagação dos elétrons em cabos de cobre. A diferença está nos intervalos de tempo representando cada bit: as redes de alta velocidade tem bits menores que redes de baixa velocidade. Assim, estritamente falando o termo largura de banda é mais apropriado que velocidade.

A largura de banda é determinada pelo meio de transmissão utilizado, protocolos, distância entre nós intermediários e velocidade de comutação nos nós intermediários. Entre par trançado, cabo coaxial e fibra ótica, esta última é a que oferece maior largura de banda. Em teoria, um única fibra ótica poderia suportar alguns terabits/s. Ela é a melhor escolha para redes de longa distância.

Outro termo utilizado para expressar a largura de banda da rede é a *taxa de bits*. A taxa de bits entre dois sistemas comunicantes é o número de dígitos binários que a rede é capaz de transportar por unidade de tempo (expresso em Kbps, Mbps, Gbps, etc.).

#### 7.1.2 Taxa de bits agregada e individual

A interface entre o computador e a rede a que ele é conectado pode dar suporte a uma ou várias comunicações com sistemas remotos ao mesmo tempo. Por exemplo, sob uma interface entre o computador e uma Ethernet, centenas de comunicações simultâneas podem coexistir. Neste caso, podemos dizer que a comunicação é multiplexada. Como resultado, deve-se distinguir a taxa de bits de uma comunicação individual da taxa de bits agregada (Figura 67)



**Figura 67.** Taxa de bits agregada e individual

### 7.1.3 Velocidade de acesso e taxa de bits

Existem duas noções associadas à taxa na interface entre o computador e a rede: a velocidade de acesso e a taxa de bits.

A velocidade de acesso refere-se a freqüência na qual os bits podem ser enviados e recebidos na interface de rede. Esta freqüência é normalmente definida pela tecnologia usada pela rede, ou pela subscrição de um cliente feita a um serviço de rede. Por exemplo, a velocidade de acesso de uma LAN Ethernet convencional é de 10 Mbps.

Infelizmente, nem todas as redes são capazes de transmitir dados na velocidade de acesso fornecida pela interface de rede. Vários tipos de redes não aceitam dados durante certos intervalos de tempo devido a congestão interna, falta de capacidades, ou pois o usuário se subscreveu a uma taxa de bits menor que a taxa de acesso.

Na prática, redes baseadas em pacotes normalmente não podem manter a taxa de bits igual a velocidade de acesso da interface, a menos que ela esteja totalmente descarregada. O princípio destas redes é que componentes de comutação e transmissão são compartilhados por muitas comunicações, cada uma obtendo uma fração da largura de banda total.

### 7.1.4 Vazão (Throughput)

A vazão de uma rede é sua taxa de bits efetiva, ou a largura de banda efetiva. Assim, nós definimos a vazão como sendo a diferença entre a taxa de bits da ligação e os vários overheads (sobrecarga) associados a tecnologia de transmissão empregada. Por exemplo, a tecnologia ATM sobre o sistema de transmissão de fibra ótica SONET (*Synchronous Optical NETwork*) a uma taxa de bits de 155,52 Mbps tem como principais sobrecarga aproximadamente 3% para o SONET e 9,43% para ATM. Assim, a vazão máxima desta rede é cerca de 136 Mbps.

A vazão da maioria das redes, seja rede local ou de longa distância, varia com o tempo. Algumas vezes a vazão pode mudar muito rapidamente devido a falhas nos nós da rede ou linhas ou devido a congestão quando grandes fluxos de dado são introduzidos na rede. Alguns fatores que afetam a vazão da rede são: falha de nós e ligações; congestão (devido a sobrecarga ou gargalos); gargalos no sistema (por exemplo, quando ligações via satélite TransAtlantic a 128 Kbps são usadas); capacidade do buffer nos sistemas finais e na interface de rede; e controle de fluxo feito por protocolo fim-a-fim que limita a taxa de transferência.

Muitas vezes, a sobrecarga é considerada implícita, e a vazão é simplesmente igual a taxa de bits do sistema.

### 7.1.5 Taxa de Erro

Outro parâmetro importante para redes multimídia é a taxa de erros. Este parâmetro pode ser definido de diversas maneiras. Uma é a taxa de erro de bits (BER – Bit Error Rate), que é a razão entre o número médio de bits corrompidos ou errados e o número total de bits que são transmitidos. Outro é a taxa de erro de pacote (PER) definido como o anterior, substituindo bits por pacotes. Um terceiro parâmetro é a taxa de erro de quadro (FER), que é aplicada a redes ATM, definida como o número de quadros (células) errados sob o total de quadros transmitidos.

Erros ocorrem mais em redes a comutação de pacotes. Eles ocorrem quando: bits individuais em pacotes são invertidos ou perdidos, pacotes são perdidos no trânsito, pacotes são cortados ou atrasados, ou quando pacotes chegam fora de ordem.

Nas redes de hoje, as taxas de erro de bits são muito baixas. Por exemplo, em transmissão sob fibras óticas, o BER varia de  $10^{-9}$  a  $10^{-12}$ . Em sistemas de transmissão via satélite, o BER é na ordem de  $10^{-7}$ . Isto significa que em média, há um bit errado em um arquivo de 10 milhões de bits, para o sistema de transmissão via satélite. Considerando que um único quadro de vídeo pode consistir de muitos milhões de bits, tal BER implica que pode haver aproximadamente um bit errado por quadro em uma transmissão de vídeo digital.

Em redes orientadas a conexão, quando pacotes tem bits errados, perdidos ou cortados, o sistema receptor é normalmente capaz de detectar tal situação e informar o problema ao emissor. O sistema receptor nem sempre tem a informação precisa acerca de quais pacotes contêm erros ou que pacotes foram perdidos ou cortados. Uma abordagem padrão para o problema é fazer o emissor transmitir novamente a maior parte dos pacotes recentes. Nas redes onde a abordagem correção de erro em avanço (FEC - *Forward Error Correction*) é empregada, o receptor pode normalmente detectar e corrigir erros de bits em pacotes sem exigir a retransmissão. FEC não trabalha muito bem quanto os pacotes inteiros são perdidos ou cortados na transmissão.

No caso de redes sem conexão, os pacotes perdidos ou cortados são difíceis, senão impossíveis, de serem detectados. A primeira razão para pacotes serem cortados ou perdidos em redes de alta velocidade é o espaço de buffer insuficiente no receptor devido a congestão na rede.

Poderia-se pensar que se tais erros são sempre desastrosos. Em algumas aplicações tal como transmissão de vídeo, um único erro em um quadro normalmente não pode ser visto pelo olho humano. Em outros, tais como transferência bancária de fundos, um único bit errado poderia ser desastroso.

### 7.1.6 Atraso Fim-a-Fim

Um dos principais parâmetros de desempenho de rede é o atraso. Ele pode ser definido de várias maneiras. Nós consideramos inicialmente o **atraso fim-a-fim**, que significa o tempo levado para transmitir um bloco de dados de um emissor a um receptor. O atraso fim-a-fim é composto dos seguintes componentes:

- **Atraso de trânsito**, que é o parâmetro denotando o tempo de propagação necessário para enviar um bit de um local a outro, limitado pela velocidade da luz. Esta parâmetro é dependente apenas da distância percorrida e é significante apenas quando ligações de satélite são usadas.
- **Atraso de transmissão**, que é definido como o tempo necessário para transmitir um bloco de dados fim-a-fim. Este parâmetro é dependente apenas da taxa de bits da rede e do tempo de processamento dos nós intermediários, tal como roteadores e bufferização.
- **Atraso na interface**, que é definido como o atraso ocorrido entre o tempo em que o dado está pronto para ser transmitido e o tempo em que a rede está pronta para transmitir o dado. Este parâmetro é importante para redes orientadas a conexão, do tipo X.25 em que um circuito fim-a-fim deve ser estabelecido antes da transmissão do dado. Este parâmetro é relevante também em redes token-ring quando a transmissão não pode ser feita até um token livre chegar.

Alguns atrasos de rede são inevitáveis, tão imutáveis quanto as leis da física. Se dois terminais estão comunicando via satélite, então o atraso de transito é aproximadamente de 0,25 segundos. Desde que o satélite está a 36.000 Km da terra, o tempo de subida na velocidade da luz é cerca de um quarto de segundo. Outros atrasos são devido a taxa de bits na ligação: maior é a largura de banda, menor é o atraso.

### 7.1.7 Atraso Ida-e-Volta

Este parâmetro tem sentido em redes orientadas a conexão, quando reconhecimentos fim-a-fim são necessários. Atraso ida-e-volta é definido como o tempo total necessário para um emissor enviar um bloco de dados pela rede e receber um reconhecimento de que o bloco foi recebido corretamente. Este parâmetro é importante, por exemplo, em redes TCP executando sob o protocolo IP sem conexão.

Quando as redes estão muito congestionadas, este parâmetro fornece uma melhor imagem do desempenho da rede do que o atraso fim-a-fim.

### 7.1.8 Variação de atraso (Jitter)

Na transmissão de vídeo digital, o fluxo de vídeo e de áudio são normalmente enviados separadamente. Em redes a pacotes, estes fluxos são adicionalmente divididos em

blocos de dados, e cada bloco é transmitido em seqüência. Se a rede é capaz de enviar todos os blocos com uma latência uniforme, então cada bloco deveria chegar no destino após um atraso uniforme. Muitas redes hoje em dia não garantem um atraso uniforme para seus usuários. Variações em atrasos são comuns. Esta variação de atrasos na transmissão é causada por muitos fatores, tal como diferenças de tempo de processamento dos pacotes, diferenças de tempo de acesso à rede e diferenças de tempo de enfileiramento. Se as variações nos atrasos são devido as imperfeições do sistema na rede (software ou hardware), ou devido as condições de tráfego dentro da rede, estas variações são normalmente chamadas de ginga (*jitter*). No projeto de uma rede multimídia, é importante colocar um limite superior na ginga admissível.

## 7.2 Caracterização do Tráfego Multimídia

Tráfego multimídia é composto de cinco categorias: áudio, vídeo, dado, imagens bitmap e gráficos. Esta seção trata unicamente das fontes de áudio e vídeo tempo-real, sendo estas as mídias que impõem maiores requisitos aos sistemas de computação e comunicação. Neste tipo de tráfego, mesmo estes fluxos de áudio e vídeo sendo quebrados em pacotes ou quadros para transporte na rede, é importante manter a integridade destes dados, e isto implica em algumas restrições quanto aos parâmetros de desempenho da rede.

### 7.2.1 Tipos de transferência: Assíncrona e Síncrona

Como visto no capítulo 5, existem duas formas de transmissão de áudio e vídeo de uma fonte a um destino:

- **Transmissão Assíncrona ou Telecarga:** neste modo a informação é primeiro totalmente transferida e armazenada no receptor, para depois ser apresentada.
- **Transmissão Síncrona ou Tempo-Real:** neste modo a informação é transferida em tempo-real sobre a rede e apresentada continuamente no receptor.

No caso de áudio e vídeo armazenados (gerados off-line), a transmissão pode ser assíncrona ou síncrona:

- No caso de uma transferência assíncrona, o usuário final da informação terá que aguardar a transferência completa do arquivo da fonte para o destino, além de exigir uma capacidade de armazenamento no destino suficiente para armazenar todo o arquivo. Como visto no capítulo 5, este tipo de transferência é viável apenas para seqüências de áudio e vídeo muito pequenas. Caso contrário, o atraso no início de apresentação seria muito grande e os requisitos de armazenamento no destino seriam muito alto (p.e. um vídeo MPEG de 1 minuto codificado MPEG a 2Mbps consumiria 15 MBytes e levaria 1 minuto para ser transferido caso a vazão da rede fosse de 2Mbps).
- No caso de uma transferência síncrona, o usuário final não aguardará a carga completa do arquivo, a informação deve ser transferida da fonte em uma taxa muito próxima a de apresentação e, após um pequeno atraso de transmissão, deve ser apresentada no destino. Neste tipo de transferência, os requisitos de armazenamento no destino são reduzidas, pois ela deve manter apenas uma pequena parcela da informação. Em compensação, este tipo de transferência impõe duros requisitos ao sistema de comunicação. Estes requisitos serão detalhados neste capítulo.

O restante deste capítulo trata unicamente da transferência síncrona, ou tempo-real, de áudio e vídeo.

### 7.2.2 Variação de vazão com o tempo

O tráfego multimídia pode ser caracterizado como taxa de bits constante (CBR) ou taxa de bits variável (VBR).

### Tráfego a taxa de bits constante

Muitas aplicações multimídia, tal como aplicações CD-ROM, geram saída a taxa de bits constante (*CBR - Constant Bit Rate*). Outro exemplo é o áudio PCM. Por exemplo, no caso de um áudio qualidade telefone, amostras de 8 bits são produzidas a intervalos fixo de 125  $\mu$ s (mais ou menos uma pequena variação). Para aplicações tempo-real envolvendo fluxos de dado CBR, é importante que a rede transporte estes fluxos de dado a uma taxa de bits constante também. Caso contrário, é necessário a realização de uma bufferização custosa em cada sistema final. É importante notar aqui que em muitas redes tal como ISDN é natural transportar dados CBR.

### Tráfego a taxa de bits variável

Tráfego a taxa de bits variável (*VBR - Variable Bit Rate*) tem uma taxa de bits que varia com o tempo. Tal tráfego normalmente ocorre em rajadas, caracterizados por períodos aleatórios de relativa inatividade quebrados com rajadas de dados. Uma fonte de tráfego em rajada gera uma variação do conjunto de dados em diferentes intervalos de tempo. Uma boa medida deste tipo de tráfego é dada pela relação entre o pico da taxa de bits pela taxa de tráfego média em um dado período de tempo.

Áudio e vídeo são compactados geralmente geram tráfego a taxa de bits variável. Neste caso o tráfego geralmente é caracterizado por uma taxa média e uma taxa de pico.

#### 7.2.3 Dependência temporal

Quando pessoas estão envolvidas na comunicação, o atraso total deveria ser abaixo de um nível de tolerância, permitindo assim um certo nível de interatividade. Por exemplo, na videofonia, o atraso total de transmissão das imagens e da voz de um interlocutor da fonte para o destino deve ser pequena. Caso contrário, a conversação perde em interatividade. Na videoconferência, a experiência tem mostrado que o atraso deve ser de no máximo 150 ms a fim de que os participantes não percebam seus efeitos. Em outras aplicações, tal como e-mail multimídia, o tráfego gerado não é tempo-real.

#### 7.2.4 Continuidade Temporal

No caso de mídias contínuas, como áudio e vídeo, embora a compressão reduza o tamanho dos dados, o requisito de continuidade temporal existe tanto para fluxos compactados ou não. Ou seja, as amostragens de áudio ou quadros de vídeo, mesmo que compactados, devem ser amostrados e apresentados em intervalos regulares, senão a qualidade percebida será inaceitavelmente baixa. Esta propriedade é chamada de **isocronia** ou **sincronização intramídia**. Por exemplo, a voz de telefonia digital é codificada na forma de amostras de 8bits feita a todo 125  $\mu$ s. Para uma boa qualidade de apresentação, estas amostras devem ser apresentadas em intervalos de 125  $\mu$ s mais ou menos uma pequena variação. Caso uma amostra não possa ser apresentada no instante correto, geralmente ela deve ser descartada.

### 7.3 Requisitos para Transmissão de Áudio e Vídeo

É relativamente fácil garantir desempenho para comunicação multimídia se são usados computadores dedicados e redes a comutação de circuitos. De qualquer maneira, por razões econômicas, os sistemas multimídia mais interessantes e potencialmente úteis são distribuídos, compartilhados entre vários usuários e usam um tipo de rede a comutação de pacotes em vez de redes a comutação de circuitos dedicados [Lu, 96].

A natureza síncrona das mídias contínuas impõe duros requisitos em termos de largura de banda, atrasos, variação de atrasos e outros. Como apresentado no capítulo 2, seqüências de áudio e vídeo, mesmo compactadas, necessitam grandes capacidades de armazenamento e alta largura de banda de transmissão (p.e. um vídeo MPEG de alta qualidade requer uma taxa de bits de vários Mbit/s).

Esta seção identifica os principais requisitos que a transmissão de áudio e vídeo impõem às redes de comunicação. Estes requisitos serão expressos em termos de características de desempenho da rede tal como vazão, confiabilidade e atraso. Outros requisitos, tal como comunicação multicast também são discutidos.

### 7.3.1 Sistemas multimídia versus sistemas tempo-real críticos

Sistemas multimídia são freqüentemente chamados de sistemas tempo-real devido aos seus requisitos de atraso e de variação de atraso. Mas eles se diferenciam de outros tipos de sistemas tempo-real, chamados de *sistemas tempo-real crítico*, em que o sistema é completamente programado para fornecer um serviço único, tal como controle de tráfego aéreo ou guia de foguete. Sistemas tempo real críticos têm duras restrições temporais e desastres podem ocorrer se elas não forem respeitadas. Ao contrário, falhas em sistemas multimídia geralmente não são desastrosas.

Sistemas tempo-real críticos executam em sistemas dedicados, onde todos os recursos são dedicados à aplicação. Assim, quando recursos suficientes são disponíveis, as restrições da aplicação são quase sempre satisfeitas. Já em sistemas multimídia, é assumido que os recursos do sistema (memória, CPU, acesso ao disco, rede, etc.) são compartilhados por várias aplicações. Neste caso é difícil prever a disponibilidade de recursos. Isto é um dos problemas no projeto de sistemas multimídia.

Embora a falha de desempenho das aplicações multimídia não seja desastrosa, uma garantia deveria ser provida em uma alta percentagem para ter uma boa qualidade de apresentação. Existem duas formas de obter boa qualidade de apresentação. O primeiro é satisfazer inteiramente os requisitos de desempenho da aplicação. O segundo é que no caso de falha na satisfação dos requisitos de desempenho da aplicação, dados menos importantes possam ser sacrificados em favor dos dados mais importantes e técnicas de cancelamento de erros são usadas para obter uma alta qualidade percebida da apresentação. Esta última técnica é chamada de degradação progressiva da qualidade.

### 7.3.2 Requisitos de vazão

Os requisitos multimídia associados à vazão são discutidos abaixo:

#### **Requisito de grande largura de banda de transmissão**

Uma grande largura de banda é um requisito básico para aplicações multimídia, sem a qual a rede é definitivamente inapropriada para multimídia. Geralmente em aplicações multimídia, cada usuário necessita de alguns Mbits/s (com compactação).

Os requisitos de largura de banda das aplicações multimídia são muito dependentes da qualidade escolhida para os áudios e vídeos transmitidos e técnica de compressão utilizada (como apresentado no capítulo 3). [Fluckiger, 95] apresenta várias qualidades de apresentação de áudio e vídeo e seus requisitos em termos de taxa de bits. Dois padrões de compressão de vídeo são particularmente relevantes: ISO MPEG e ITU H.261. Em termos de largura de banda, eles necessitam 1,2 a 80 Mbps para MPEG e MPEG-2 e 64 Kbps a 2 Mbps para H.261. Baseado em experiências práticas, um total de 1,4 Mbps para áudio e vídeo é muito interessante pois fornece uma boa qualidade de vídeo e permite o uso de equipamentos audiovisuais comerciais (CD player) e transmissão sob linhas T1 (1,5 Mbps). Com relação ao custo de transmissão em WANs, H.261 usando 6\*64 Kbps, isto é 384 Kbps, é uma alternativa atrativa. Implementações H.261 existentes mostram que 64 Kbps é aceitável apenas em alguns vídeos estáticos (vídeo mostrando apenas a cabeça da pessoa que fala), enquanto 384 Kbps é interessante mesmo para vídeos mostrando cenas normais. Assim, nós podemos concluir que para as aplicações multimídia atuais é necessário uma vazão entre 0,4 a 1,4 Mbps. Esta vazão é necessária para fluxos unidirecionais (pois o tráfego multimídia é normalmente de natureza altamente assimétrica).

### **Requisito de grande largura de banda de armazenamento**

Em redes de alta vazão, é importante que o sistema receptor tenha capacidades de armazenamento suficientes para receber o tráfego multimídia que chega. Além disso, é necessário que a taxa de entrada do buffer seja alta suficiente para acomodar o fluxo de dado que chega da rede. Esta taxa de dados é algumas vezes chamada de largura de banda do buffer de armazenamento. Este de fato não é um requisito de rede, mas um requisito de terminal multimídia.

### **Requisito de continuidade temporal**

Uma rede multimídia deve ser capaz de manipular grandes fluxos de dados, tal como aqueles gerados por fontes de áudio e/ou vídeo. Isto significa que a rede deve ter uma vazão suficiente para assegurar a disponibilidade dos canais de alta largura de banda por grandes períodos de tempo. Por exemplo, não é suficiente para a rede oferecer ao usuário um espaço de tempo de 5 segundo a 1,5 Mbps se o usuário necessita enviar um fluxo de 30 Mbits. A rede satisfaz os requisitos de continuidade temporal quando ela pode oferecer a disponibilidade contínua de um canal de 1,5 Mbps para o usuário. Se existem vários fluxos na rede ao mesmo tempo, a rede deve ter uma capacidade de vazão igual ou maior que a taxa de bits agregada dos fluxos.

#### **7.3.3 Requisitos de confiabilidade (controle de erro)**

É difícil precisar os requisitos de controle de erro para redes multimídia pois as aplicações multimídia são, de certo modo, tolerantes a erros de transmissão. Parte da razão desta tolerância é devido aos limites da percepção sensorial humana, como apresentado no capítulo 3 desta apostila.

Requisitos de controle de erro são também difíceis de se quantificar pois em muitos casos os requisitos de controle de erro e requisitos de atraso fim-a-fim são contraditórios. Esta contradição ocorre pois muitos esquemas de controle de erro envolvem a detecção e retransmissão do pacote com erros ou perda. Algumas vezes, a transmissão deve ser realizada na base fim-a-fim, que significa um aumento no atraso. Para transmissão tempo-real de áudio e vídeo, o atraso é mais importante que a taxa de erros, assim em muitos casos, é preferível ignorar o erro e trabalhar simplesmente com o fluxo de dado recebido.

#### **7.3.4 Requisitos de atraso e variação de atraso**

Em todos os sistemas multimídia distribuídos sempre existe um atraso entre a captura/leitura de uma informação (p.e. um vídeo) em uma fonte e sua apresentação em um destino. Este atraso, chamado de atraso fim-a-fim, é gerado pelo processamento da informação na fonte, sistema de transmissão e processamento no destino.

Em redes a comutação de pacotes, os pacotes de dados não chegam ao destino em intervalos fixos como necessário para transmissão de mídias contínuas. Por causa desta variação de atrasos, pacotes de áudio e vídeo que chegam não podem ser imediatamente apresentados. Caso contrário, teríamos a apresentação de vídeos aos trancos e apresentação de áudios de má qualidade. A nível de percepção humana, a variação de atrasos na transmissão de pacotes de voz é um problema crítico, podendo tornar a fala incompreensível.

A abordagem mais utilizada para a remoção desta variação de atraso é o uso de buffers do tipo FIFO (*First-In First-Out*) no destino antes da apresentação. Esta técnica é chamada de **técnica de bufferização** (Figura 68). Nela, na medida que os pacotes chegam (em uma taxa variada) eles são colocados no buffer; o dispositivo de apresentação retira amostragens do buffer em uma taxa fixa.

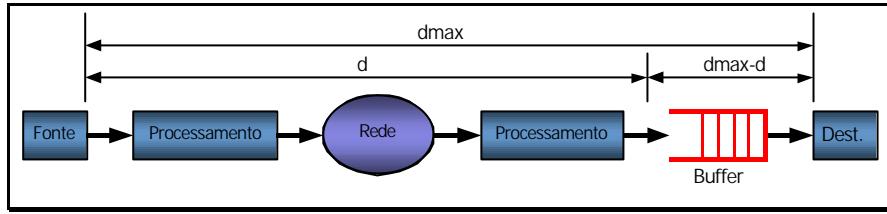


Figura 68. Técnica de buferização [Lu, 96]

O princípio da técnica de bufferização é adicionar um valor de atraso variável a cada pacote de tal forma que o atraso total de cada pacote seja o mesmo. Por esta razão, este buffer é chamado de buffer de uniformização de atrasos. A Figura 68 ilustra todas as operações realizadas nos sistemas finais para a transmissão de uma mídia contínua. Supondo que um pacote pode atrasar (incluindo tempo de bufferização) de um tempo mínimo de atraso  $d_{min}$  e um tempo máximo de atraso  $d_{max}$ . Se um pacote com atraso de  $d$  é bufferizado durante  $(d_{max}-d)$ , todos os pacotes terão um atraso fixo de  $d_{max}$ . O destino partirá a apresentação do dado  $d_{max}$  segundos após ele ter sido enviado. Portanto cada pacote será apresentado em tempo (assumindo que a taxa de apresentação é a mesma que a taxa de geração do dado).

Neste esquema, o tempo máximo de bufferização é  $(d_{max}-d_{min})$ , que é a maior variação de atraso. Maior este valor, maior é o tamanho do buffer necessário. Para satisfazer os requisitos de comunicação multimídia, o buffer não deve sofrer sobrecarga ou subutilização. Em contrapartida, o tamanho do buffer não dever ser muito grande: um buffer grande significa que o sistema é custoso e o atraso fim-a-fim é muito grande.

[Lu, 96] utiliza o modelo produtor/consumidor para analisar os requisitos de largura de banda de transmissão, atraso e variação de atraso. Por simplicidade, nesta análise é assumido que a informação multimídia é codificada a uma taxa de bits constante, embora o princípio discutido também se aplique a fluxos codificado a taxa de bits variável. Codificação a taxa de bits constante significa que o destino consome dados a uma taxa constante. [Lu, 96] introduz a função de dados que chegam  $A(t)$  e função de consumo de dados  $C(t)$ .  $A(t)$  indica o conjunto de dados que chegam no cliente dentro do intervalo temporal 0 a  $t$ .  $C(t)$  indica o conjunto de dados consumidos dentro do intervalo 0 a  $t$ . As funções  $A(t)$  e  $C(t)$  são funções não decrescentes.  $C(t)$  aumenta com o tempo em uma taxa constante.  $A(t)$  normalmente não aumenta a taxa fixa devido a variações de atraso. Assumindo que o tempo de envio do primeiro pacote é 0, o tempo de chegada do primeiro pacote de dados no cliente é  $t_1$  e o cliente apresenta o primeiro pacote em  $t_2$ , então temos a função de chegada  $A(t-t_1)$  e a função de consumo  $C(t-t_2)$ , como mostra a Figura 69. Para satisfazer os *requisitos de continuidade*,  $A(t-t_1)$  dever ser igual ou maior que  $C(t-t_2)$ . A diferença é bufferizada no cliente e representa a ocupação do buffer.

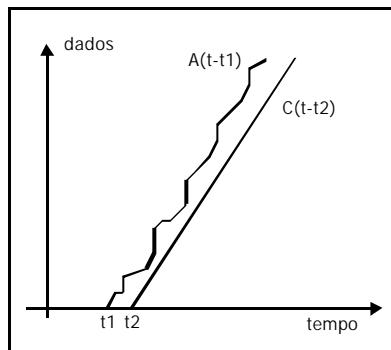
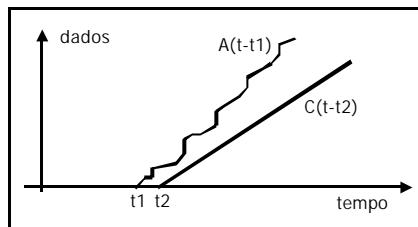


Figura 69. Taxa de chegada próxima a taxa de consumo

### Requisitos de Largura de Banda

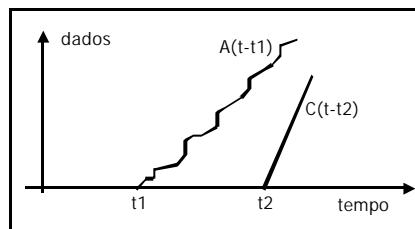
A inclinação de  $A(t-t_1)$  representa a taxa de chegada de dados. O valor médio desta taxa deve ser igual ou próximo a taxa de consumo, como é ilustrado na Figura 69. Se a taxa de consumo é menor, a diferença de  $A(t-t_1)$  e  $C(t-t_2)$  que representa a ocupação do buffer, aumenta com o tempo (como ilustrado na Figura 70). Isto significa que para o

sucesso da apresentação do fluxo o tamanho do buffer é infinito ou a apresentação do fluxo pode apenas se manter durante um tempo limitado (determinado pelo tamanho do buffer). Senão ocorrerá a sobrecarga do buffer. Para prevenir isto, um controle da taxa de transmissão deve ser usado de modo que a taxa de transmissão seja próxima a taxa de consumo no receptor.



**Figura 70.** Taxa de chegada maior que taxa de consumo

Por outro lado, se a taxa de consumo é maior que a taxa média de chegada, para satisfazer o requisito que  $A(t-t_1)-C(t-t_2) \geq 0$ ,  $t_2$  deve ser menor (Figura 71). Isto significa que um atraso inicial é menor. Consequentemente, o tempo de resposta se torna curto, necessitando de um tamanho de buffer menor. Como mostrado na Figura 71, menor o fluxo a ser apresentado, menor é o atraso inicial e menor os requisitos do buffer, que não são desejáveis nem praticáveis. Portanto a taxa de chegada média deveria ser igual à taxa de consumo para permitir a apresentação com sucesso de fluxos contínuos. Para obter isto, o transmissor deveria enviar na taxa de consumo, e a largura de banda de transmissão fim-a-fim deve ser ao menos igual a taxa de consumo.



**Figura 71.** Taxa de chegada menor que taxa de consumo

### Requisitos de Atraso e Variação de Atraso

É óbvio que em sistemas multimídia o atraso fim-a-fim deve ser pequeno. Esta necessidade pode ser vista nas figuras 2 a 4:  $t_1$  representa o atraso entre o instante do envio do primeiro pacote ao instante de recepção deste pacote pelo cliente; e  $t_2$  é o atraso fim-a-fim do primeiro pacote. Se o atraso fim-a-fim não é limitado, o tempo de resposta do sistema também não é limitado. Isto é indesejável especialmente em aplicações multimídia interativas. Por exemplo, experiências mostram que um atraso fim-a-fim abaixo de 0,3 segundos é necessário para telefonia, e para aplicações interativas de vídeo sugere-se um atraso fim-a-fim máximo de até 150 ms.

### Requisito Tamanho do Buffer de Uniformização

O requisito de tamanho do buffer é igual ao tempo de bufferização multiplicado pela taxa de chegada de dados, sendo que o tempo máximo de bufferização é igual a máxima variação de atraso de chegada dos pacotes. Assim, quanto maior a variação de atraso, maior é o tamanho de buffer requerido. Portanto, para limitar o tamanho do buffer requerido, a variação de atrasos deve ser pequena.

### 7.3.5 Compartilhamento eficiente de recursos de rede

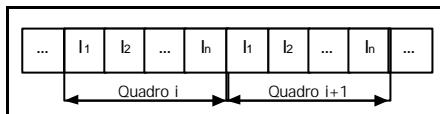
Dados multimídia são geralmente transmitidos em rajadas, especialmente após a compactação. Portanto, se o usuário reservar uma largura de banda igual a seu pico de taxa de transmissão, parte da largura de banda é desperdiçada quando a taxa de bits não estiver no máximo. A melhor abordagem para uso eficiente da rede é o princípio de *largura de banda sob demanda* ou *multiplexação estatística*: uma aplicação pode usar

tanta largura de banda quanto necessária sujeito a um valor máximo. Quando uma aplicação não usa toda a sua largura de banda outra aplicação pode usar.

A tecnologia comutação de circuitos não é apropriada, isto pois é reservada uma largura de banda fixa para o circuito entre a fonte e o destino mesmo que a largura de banda não seja completamente usada. Igualmente, comutação de pacotes a Multiplexação por Divisão de Tempo Síncrono (STDM) não é apropriada pois a largura de banda de cada canal é fixado sem olhar seu uso. Portanto uma forma de redes a comutação de pacotes é mais adaptada ao eficiente uso dos recursos da rede.

Multiplexação aqui refere-se ao compartilhamento do meio de transmissão por várias conexões distintas (lógicas ou virtuais). Uma técnica de multiplexação muito utilizada é a chamada *multiplexação por divisão de tempo* (TDM) em que o tempo de transmissão do meio é compartilhado entre várias conexões ativas. Nela, alguns bits são agrupados em um intervalo temporal que pode ser usado por uma única conexão.

Na multiplexação síncrona, o tempo é dividido em quadros de tamanho fixo que por sua vez são divididos em intervalos de tamanho fixo (Figura 72). Por exemplo, assuma que todo quadro de transmissão é dividido em 10 intervalos e eles são numerados de 1 a 10. Se o intervalo 1 é atribuído a uma conexão, o emissor pode transmitir dados sob esta conexão apenas no intervalo 1. Caso ele tenha mais dados a transmitir após este intervalo, o transmissor deve aguardar o próximo quadro. Se ele não usa este intervalo temporal, nenhuma outra conexão pode utilizá-lo. Este tipo de multiplexação é chamada *multiplexação por divisão de tempo síncrona* (STDM).



**Figura 72.** STDM: Multiplexação por divisão de tempo síncrona

A comutação a circuito e STDM são capazes de fornecer garantias de desempenho *hard*. Mas para muitas aplicações multimídia a garantia estatística é suficiente. *Multiplexação estatística*, compartilhamento da largura de banda baseada no princípio largura de banda sob demanda, pode suportar muitas aplicações dada a mesma largura de banda. Multiplexação estatística é uma técnica que multiplexa vários fluxos de dados independentes em um canal de alta largura de banda. Como as mídias contínuas são transmitidas em rajadas e geralmente são independentes, provavelmente enquanto alguns fluxos serão em suas taxas baixas de transmissão outros serão transmitidos em suas taxas altas. Assim, a largura de banda agregada é menor que a soma dos picos de taxa de bits dos fluxos. Assim multiplexação estatística usa a largura de banda eficientemente.

Por exemplo, assumindo que nós temos 50 fluxos de vídeo independentes codificados a taxa de bits variável, com um pico da taxa de bits é 8 Mbits/s e uma taxa média é igual a 2 Mbits/s. Usando STDM, necessita-se de uma ligação de largura de banda de 400 Mbits/s para transmitir estes 50 fluxos. Usando multiplexação estatística é necessária uma ligação com uma largura de banda em torno de 100 Mbits/s. Por simplicidade, este valor foi calculado a partir da taxa de bits média (2 Mbits/s). Na prática, este cálculo é baseado na distribuição estatística de cada fluxo e depende da taxa de perda de pacotes e atrasos.

O tamanho de pacotes também afeta o uso eficiente da largura de banda. Quando pacotes são muito grandes, muitos pacotes não são completos, especialmente o último pacote de uma mensagem, desperdiçando largura de banda. Quando pacotes são muito pequenos, a maior parte da largura de banda é desperdiçada pelo uso de informações de controle.

A retransmissão de dados perdidos também afeta o uso eficiente de recursos da rede. Para apresentações ao vivo de áudio e vídeo, a retransmissão não é desejável pois causa atrasos extra na transmissão de pacotes sucessivos e além disso perdas e erros em apresentações multimídia são toleráveis. Em geral, a retransmissão desperdiça largura de banda e causa atrasos extra para pacotes subsequentes.

Uma boa abordagem para reduzir efeitos da perda de pacotes na qualidade de apresentação é usar prioridades de tráfego. O meio físico de transmissão, especialmente fibras ópticas, são muito confiáveis. As perdas geralmente ocorrem na escassez de buffer nas trocas de rede. Se prioridades são atribuídas ao dado de acordo com sua importância, os dados de baixa prioridade serão descartados na escassez de buffer.

### 7.3.6 Garantias de Desempenho

Para garantir o desempenho, a rede deveria garantir que um pacote possa acessar a rede em um tempo especificado e que quando na rede, o pacote deveria ser liberado dentro de um tempo fixo. Existem duas possibilidades da não garantia de desempenho:

- Protocolo de controle de acesso ao meio de transmissão (MAC) não garantindo o tempo de acesso à rede. Neste caso, o pacote pode levar muito tempo para ter acesso à rede. Por exemplo, CSMA/CD usado na Ethernet faz com que as estações aguardem até que a rede esteja livre; após a transmissão, se colisões ocorrerem, o dado é retransmitido em um tempo indeterminado. Outra característica indesejável do CSMA/CD é que o tamanho mínimo de pacote é determinado pela velocidade de transmissão e pelo comprimento do cabo: em uma rede Ethernet de 10 Mbits/s, o tamanho mínimo de pacote é de 64 bits para um cabo de 5 Km; mas se uma rede Gigabit Ethernet (1 Gbits/s) for usada, o tamanho mínimo de pacote é 512 bits, que é muito grande para o uso eficiente da largura de banda e para a comunicação de mídias contínuas (o tempo de montagem de cada pacote é muito grande).
- Falha de garantia de desempenho nos nós intermediários ou comutadores da rede. Uma vez o pacote na rede, o meio de transmissão enviará o pacote na velocidade da luz ou de elétrons da fonte ao destino diretamente ou para um nó intermediário da rede ou comutador. Se a fonte estiver conectada diretamente ao destino o pacote será transmitido sem atrasos adicionais. Mas se existirem um ou mais comutadores entre a fonte e o destino, atrasos extras e perdas de pacotes podem ocorrer. Porque em cada nó intermediário da rede, o pacote deve ser bufferizado, o canal de saída para o pacote é determinado, e apenas quando o canal é disponível o pacote é transmitido. Este processo armazenar-e-retransmitir é indeterminístico. O buffer no comutador pode estar cheio, causando perdas de pacotes. O processador pode estar ocupado, atrasando a decisão de comutação. Canais de longa distância podem estar ocupados, causando atrasos extra.

Este segundo problema pode ser resolvido através de um controle de admissão e gestão de recursos de rede, especialmente filas em comutadores. Se o desempenho de um canal não é garantido, ou sua aceitação afetar outros canais existentes, este canal não deveria ser admitido.

As seguintes técnicas podem ser empregadas para permitir o uso eficiente de recursos de rede e garantir o desempenho da aplicação:

- Determinação das características dos diferentes tipos de tráfegos em termos de requisitos de pico da taxa de dados, intervalos de rajada, atrasos, variação de atrasos. Estas informações de tráfego deveriam ser enviados à rede quando a conexão é pedida e são usadas para a decisão de aceitação da conexão e para checar se o canal viola estes requisitos durante a sessão.
- Tempo de acesso à rede deveria ser garantido.
- Recursos de rede (largura de banda e buffers de fila) deveriam ser eficientemente gerenciados de maneira que tantas aplicações quanto possível sejam suportadas com garantias de desempenho.

### 7.3.7 Adaptabilidade da Rede

Existem 3 tipos de adaptabilidade:

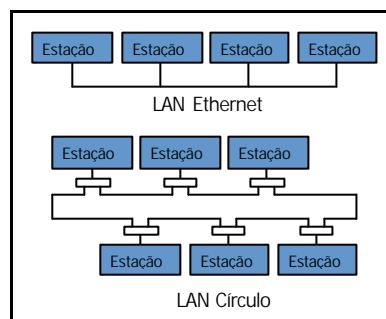
- **Distância:** a mesma arquitetura de rede e protocolo deveria operar em redes locais (LAN) e em redes de longa distância (WAN). Isto facilita a interconexão entre estas redes.

- **Largura de banda:** a largura de banda da rede deveria ser capaz de aumentar com o aumento do pedido da largura de banda do usuário sem mudar o protocolo de rede.
- **Número de usuários:** a rede deveria ser capaz de suportar um grande número de usuários e a largura de banda disponível para cada um não deveria ser afetada pelo número de estações ativas conectadas na rede. Muitas LANs usam meio compartilhado, assim todas as estações são conectadas ao meio de transmissão e compartilham a quantidade fixa de largura de banda total. A largura de banda compartilhada para cada estação diminui quando o número de estações ativas aumenta. Isto não é desejável.

### 7.3.8 Capacidade de Multicasting

Várias aplicações multimídia necessitam distribuir fluxos para vários destinos, como na distribuição de áudio e vídeo em geral. É muito lento e dispendioso enviar uma cópia da informação para cada destino um-a-um. É lento pois o acesso à rede e a transmissão tomam tempo. É dispendioso pois a mesma informação pode ser transmitida sobre a mesma ligação de rede muitas vezes. Uma solução é fazer com que a fonte envie o dado apenas uma vez e a rede seja responsável pela transmissão do dado a múltiplos destinos. Esta técnica é chamada de *Multicasting*.

Em muitas LANs, todas as estações compartilham o meio de transmissão (Figura 73). Portanto, o dado transmitidos por uma estação são recebidos por todas as outras. Quando um endereço especial é usado, o dado é recebido por todas as estações. Este modo de transmissão é chamado de *Broadcasting*. Neste tipo de rede é fácil implementar o *multicasting*: estações esperando receber o dado *multicasting* pode obter uma cópia quando o dado passa por ela.



**Figura 73.** LANs: meio compartilhado

Em WANs, é difícil para uma rede a comutação de circuito conectar dinamicamente uma ligação de entrada a múltiplas ligações de saída à longa distância. Mas uma comutação de pacotes pode enviar um pacote para múltiplos destino facilmente. Técnicas *multicasting* tem sido desenvolvida para redes de comutação de pacotes. O princípio básico das técnicas *multicasting* é o seguinte:

- estações interessadas em receber fluxos *multicasting* formam um grupo *multicasting* com um endereço único.
- todos os comutadores de rede envolvidos são informados do grupo e seu endereço.
- quando uma estação deseja enviar dados para este grupo, o endereço de grupo é usado como endereço de destino.
- quando um comutador recebe um pacote com o endereço de grupo ele envia o pacote sobre as ligações que conduzem às estações que pertencem a este grupo, sendo que um pacote não será enviado duas vezes sob a mesma ligação.

Note que os receptores em um grupo *multicast* podem ter capacidades e requisitos de QoS diferentes. Para o uso eficiente da rede: apenas o dado necessário e útil deveria ser enviado aos receptores individuais. Isto significa que alguns dispositivos ou filtros deveriam ser instalados na rede para liberar o conjunto de dados certo para receptores individuais. Este tema será retomado mais adiante.

## Capítulo 8 Garantias de QoS Fim-a-Fim para Comunicação de Áudio e Vídeo

Neste capítulo será analisada a necessidade de garantias de desempenho fim-a-fim para a transmissão tempo-real de áudio e vídeo. Para tanto, inicialmente serão analisados os principais componentes das aplicações multimídia e suas influências no desempenho do sistema.

### 8.1 Componentes dos sistemas multimídia e suas influências no desempenho

Como apresentado no capítulo 5, as aplicações multimídia podem ser classificadas em aplicações pessoa-a-pessoa e pessoa-a-sistema. Esta seção apresenta os principais componentes que compõem estes dois tipos de aplicação e analisa suas influências no desempenho total do sistema.

#### 8.1.1 Principais componentes das aplicações pessoa-a-pessoa

Aplicações multimídia pessoa-a-pessoa (Figura 74), por exemplo videofonia, envolvem a comunicação tempo-real entre pessoas usando áudio, vídeo e outros dados. Neste tipo de sistema, áudio e vídeo são gerados em tempo-real por microfones e câmeras de vídeo. Como apresentado no capítulo 2, estes sinais são convertidos na forma digital através de CAD (Conversores Analógico para Digital). Estas informações são então codificadas e passadas através de uma pilha de protocolos para transmissão via rede. No destino, os dados recebidos são passados através de uma pilha de protocolos e decodificados. Os dados decodificados são convertidos em sinais analógicos para apresentação no monitor ou alto-falante.

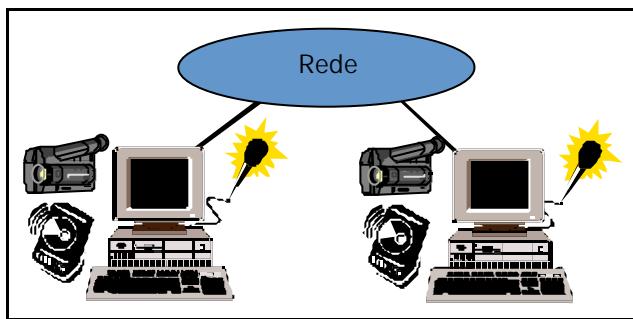


Figura 74. Sistema pessoa-a-pessoa

Os principais subsistemas que afetam a taxa de bits, atraso e variação de atraso são os seguintes [Lu, 96]:

- **Codificadores:** devido a requisitos de tempo-real neste tipo de aplicação, os codificadores são geralmente implementados em hardware. Este subsistema, onde dados são compactados em tempo-real, gera atrasos na ordem de milisegundos. Nenhuma variação de atraso é adicionada. Dados são emitidos quando ele são compactados, assim o hardware de codificação não é um gargalo na largura de banda de transmissão.
- **Pilhas de Transporte no Transmissor e Receptor:** após codificados, os dados são enviados à pilha de transporte no transmissor, à pilha de protocolos no receptor e ao decodificador sob o controle de software. Muitas pilhas de transporte são implementadas por software. Devido a natureza de compartilhamento do tempo dos sistemas finais, o tempo de processamento varia de acordo com a

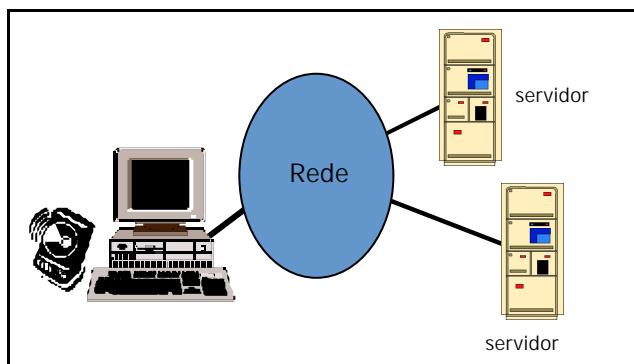
carga do sistema final. A pilha de protocolos causa atrasos e variação de atrasos significantes.

- **Acesso à Rede:** diferentes redes tem diferentes protocolos de controle de acesso. Algumas redes garantem que um pacote seja capaz de acessar o meio dentro de um certo tempo, em outras este tempo não pode ser garantido. Obviamente este último protocolo de acesso à rede não é ideal para aplicações multimídia.
- **Rede de transmissão:** um dado na rede pode passar por vários nós intermediários antes de atingir o destino. Estes nós intermediários podem ter recursos insuficientes para retransmitir o dado. Assim estes dados podem ser descartados ou atrasados. Isto depende dos protocolos de transporte e gerência de rede usados.
- **Decodificador:** decodificadores de dados podem ser implementados em hardware ou software. Decodificadores em hardware tem pequeno efeito na taxa de bits, atrasos e variação de atrasos. Mas um codificador em software pode ter problemas pelas mesmas razões das pilhas de protocolo.

### 8.1.2 Principais componentes das aplicações pessoa-a-sistema

Aplicações multimídia pessoa-a-sistema (Figura 75) têm subsistemas similares às aplicações pessoa-a-pessoa, exceto as seguintes diferenças [Lu, 96]:

- A informação multimídia é codificada e composta off-line e armazenada no dispositivo de armazenamento.
- A informação multimídia é buscada do dispositivo de armazenamento (e não gerada em tempo-real pelos codificadores). Isto causa sérios problemas, porque o tempo de acesso e taxa de transferência do dispositivo não pode ser determinado pois o dispositivo pode servir a múltiplas aplicações ao mesmo tempo. Para garantias de desempenho, servidores multimídia têm sido projetados e desenvolvidos.
- Quando grandes quantidades de informações são armazenadas no servidor, o sistema pode levar muito tempo para encontrar a informação e o tempo pode variar muito de pedido a pedido.
- Os requisitos de atrasos são menores que as aplicações pessoa-a-pessoa.
- As informações podem ser oriundas de diferentes servidores e devem ser sincronizadas para apresentação no cliente. Este ponto será analisado em detalhe no capítulo 12.



**Figura 75.** Sistema pessoa-a-sistema

## 8.2 Gerenciamento de Qualidade de Serviço

Para comunicações multimídia é necessário garantias de desempenho fim-a-fim, desde a fonte da informação até o destino. Para fornecer uma única abordagem para que diferentes aplicações especifiquem as garantias de desempenho necessárias e para que sistemas forneçam as garantias requeridas, foi introduzido o conceito de Qualidade de Serviço (QoS).

### 8.2.1 Definição de QoS

Não existe uma definição de QoS universalmente aceita atualmente. [Lu, 96] define QoS como segue:

**QoS É UMA ESPECIFICAÇÃO QUALITATIVA E QUANTITATIVA DOS REQUISITOS DE UMA APLICAÇÃO QUE UM SISTEMA MULTIMÍDIA DEVERIA SATISFAZER AFIM DE OBTER A QUALIDADE DESEJADA.**

Baseada nesta definição, existem dois aspectos para QoS: aplicações especificam os requisitos de QoS e o sistema fornece as garantias de QoS. A QoS é normalmente especificada por um conjunto de parâmetros, principalmente a taxa de bits, taxa de erros, limites de atraso e de variação de atraso. Um ou mais valores são associados a cada parâmetro. Eles especificam o valor desejado ou um intervalo admissível.

A noção de QoS foi inicialmente usada em comunicações de dados para caracterizar o desempenho da transmissão de dados em termos de confiabilidade, atraso e vazão. Por exemplo, o modelo de referência OSI tem alguns parâmetros de QoS descrevendo a velocidade e confiabilidade da transmissão, tal como vazão, atraso de trânsito, taxa de erro e probabilidade de falha de estabelecimento de conexão. Estes parâmetros são especificados na camada de transporte e não tem seus significados diretamente observáveis ou verificáveis pela aplicação. Estes parâmetros não cobrem todos os requisitos da comunicação multimídia e são usados a nível de transporte apenas. Além disso, nenhum mecanismo é especificado no modelo de referência OSI para garantir os requisitos de QoS especificados. Para comunicações multimídia, a QoS deve ser especificada e garantida fim-a-fim em todos os níveis. Portanto, aplicações multimídia requerem um novo modelo de QoS.

### 8.2.2 Estrutura Geral de QoS

[Lu, 96] propõe um modelo simplificado de operações de QoS de um sistema de comunicação multimídia, onde:

- A aplicação especifica seus requisitos de QoS e submete ao sistema.
- A partir da especificação da QoS requerida, o sistema determina se ele tem recursos suficientes para satisfazer os requisitos. Em caso afirmativo, ele aceita a aplicação e aloca os recursos necessários. Caso contrário, o sistema pode rejeitar a aplicação ou sugerir uma QoS mais baixa.

Baseado neste modelo de operação, os seguintes elementos são necessários para fornecer garantias de QoS:

- Um mecanismo de especificação da QoS para aplicações especificarem seus requisitos.
- Um processo de negociação de QoS permitindo admissão de várias aplicações.
- Controle de admissão para determinar se novas aplicações possam ser admitidas sem afetar a QoS das aplicações atuais.
- Alocação de recursos e escalonamento para satisfazer os requisitos.
- Policiamento de tráfego para tornar seguro que aplicações geram o correto conjunto de dados conforme a especificação aceita.
- Mecanismos de renegociação é necessário de modo que as aplicações possam mudar suas especificações de QoS iniciais.
- Mecanismos de monitoramento da sessão afim de que na ocorrência de problemas no provimento da QoS negociada sejam tomadas as devidas providências.
- Técnicas de degradação gradativa da qualidade e outras deveriam ser usadas juntamente com o mecanismo anterior para fornecer serviços satisfatórios para aplicações multimídia.

Todos os subsistemas devem prover os mecanismos acima e cooperar para fornecer garantias de QoS fim-a-fim. Para facilitar o desenvolvimento de uma interface QoS configurável e controle dirigido a QoS e mecanismos de gerenciamento em todas as camadas arquiteturais, vários arquiteturas de QoS foram desenvolvidas [Campbell,

94][Campbell, 96]. Estas arquiteturas fornecem uma abordagem semântica para fornecer garantias de QoS fim-a-fim.

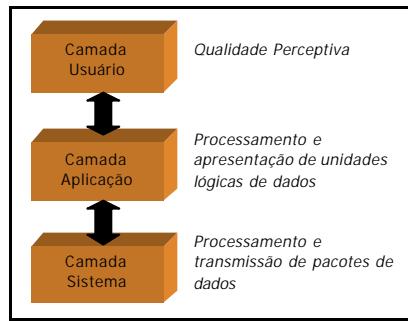
Na sequência, serão descritos alguns elementos importantes e problemas do gerenciamento de QoS.

### 8.2.3 Especificação da QoS

Para fornecer especificações e garantias de QoS geralmente uma sessão orientada a conexão é utilizada. Antes do estabelecimento da conexão, parâmetros de QoS são especificados e negociados com todos os subsistemas interessados.

Diferentes camadas do sistema manipulam diferentes dados. Por exemplo, na camada de aplicação de vídeo, o dado manipulado é uma imagem ou quadro; na camada rede o dado é um pacote; e na camada física o dado é um bit. Consequentemente, diferentes parâmetros devem ser usados por diferentes níveis dos sistema, sendo que os requisitos das camadas mais altas devem ser mapeados em requisitos de camadas mais baixas.

[Lu, 96] define um modelo de QoS considerando três camadas conceituais, como mostrado na Figura 76.



**Figura 76.** Um modelo conceptual de QoS [Lu, 96]

#### Camada Usuário

A camada mais alta é a camada usuário. Desde que em muitos casos os receptores finais são usuários finais, o resultado da QoS é a qualidade percebida pelo usuário. Normalmente, o usuário final também é o iniciador da QoS. A este nível, a qualidade é normalmente medida qualitativamente, tal como excelente, bom, aceitável, não aceitável, muito pobre. O usuário pode especificar tipos de aplicação, a qualidade desejada e requisitos de segurança. A qualidade percebida é então algo subjetivo.

A qualidade escolhida pelo usuário implica diretamente na carga de serviço: maior a qualidade requerida, maior a carga. Este fato desencorajará usuários a sempre escolher a melhor qualidade.

Os usuários serão então o ponto de partida para uma consideração global de QoS. Assim, a fonte primária dos requisitos de QoS é o usuário e uma interface apropriada deve ser fornecida para facilitar a escolha dos parâmetros. Neste nível, muitos parâmetros de QoS poderiam ser sem sentido para o usuário e deveriam ser ocultados. Uma melhor abordagem é apresentar escolhas a partir de exemplos de diferentes qualidades, tal como vídeo de qualidade TV normal ou HDTV, ou áudio de qualidade telefone ou CD. A escolha do usuário é automaticamente mapeada em parâmetros de aplicação.

#### Camada Aplicação

Na camada aplicação, as escolhas feitas pelo usuário são mapeadas (transladadas) em um conjunto de parâmetros que o nível de aplicação deve satisfazer para cumprir os requisitos do usuário. Os parâmetros neste nível são associados às unidades lógicas de dados, tal como quadro de vídeo e amostras de áudio. Para vídeo, alguns parâmetros típicos são: tamanho de imagem, bits por pixel e taxa de imagens. Para áudio, alguns parâmetros típicos são: taxa de amostragem, bits por amostragem e volume. Além disso, relacionamentos entre áudios, vídeos e outras mídias devem também ser especificados

quando duas ou mais mídias relacionadas são usadas. A tabela abaixo mostra algumas aplicações, qualidade especificada pelo usuário e parâmetros de QoS correspondentes no nível de aplicação.

Especificação do Usuário	Parâmetros de Aplicação	Parâmetros de Sistema
Qualidade de Voz Telefone	Taxa de amostragem = 8 kHz 8 bits por amostra	Taxa de Bit = 64 Kbits/s (s/ compactação) Taxa de Bit = 16 Kbits/s (c/ compactação) Atraso fim-a-fim < 150ms Taxa de perdas de pacote < 1%
Áudio CD	Taxa de amostragem = 44.1 kHz 12 bits por amostra 2 canais	Taxa de Bit = 1.41 Mbits/s (s/ compactação) Taxa de Bit = 128 Kbits/s (c/ compactação) Atraso fim-a-fim < 150ms Taxa de perdas de pacote < 0,01% Distorção entre 2 canais < 11 $\mu$ s
Vídeo NTSC	30 fps Resolução 720x480	Taxa de Bit = 200 Mbits/s (s/ compactação) Taxa de Bit = 2 Mbits/s (c/ compactação)
HDTV	30 fps Resolução 1440x1152	Taxa de Bit = 800 Mbits/s (s/ compactação) Taxa de Bit = 10 Mbits/s (c/ compactação)
Sincronização labial	Distorção intermídia < 400 ms	Variação de Atraso Requisitos de Buffer

### Camada Sistema

Na camada sistema, os parâmetros de QoS devem estar associados com as propriedades dos pacotes ou bits, tal como taxa de bits, taxa de pacotes e atraso de pacote. A coluna 3 da tabela acima mostra alguns exemplos de parâmetros a este nível. Sistemas devem satisfazer estes parâmetros para cumprir os requisitos da aplicação. Esta é uma camada composta que inclui o sistema operacional, protocolo de transporte, armazenamento secundário e rede básica. Portanto os parâmetros podem ser adicionalmente decompostos. Parâmetros nestas sub-camadas podem ser diferentes e devem ser transladadas de uma camada para outra. Por exemplo, parâmetros na camada de transporte são associadas com unidades de dado de protocolo de transporte; a camada de rede é interessada com pacotes de rede; e o armazenamento em disco manipula dados em blocos.

Na camada sistema, parâmetros fim-a-fim, tal como atraso e variação de atraso fim-a-fim, são especificados. Durante a execução, estes parâmetros devem ser divididos em sub-requisitos que devem ser satisfeitos por subsistemas individuais. Por exemplo, se o atraso total fim-a-fim para uma aplicação é 100ms, então este atraso deve ser dividido em 30ms para um sistema-final, 40ms para a rede, e 30ms para o outro sistema final. Estes atrasos devem ser adicionalmente divididos tanto quanto necessário (p.e. atraso de rede de 40ms pode ser dividido em 10ms para tempo de acesso à rede e 30ms para atrasos de enfileiramento nos comutadores). Esta subdivisão dos parâmetros fim-a-fim é um problema complicado e deve ser feito durante o processo de negociação.

Geralmente os mapeamentos entre camadas arquiteturais não são um-a-um. Alguns parâmetros são mutualmente dependentes ou contraditórios. Por exemplo, reduzindo a taxa de erros pela permissão de retransmissão aumenta o atraso de trânsito médio. Além disso, na prática, valores de QoS requeridos não correspondem a pontos bem definidos mas a regiões no espaço do parâmetro. O ponto de trabalho instantâneo dentro desta região pode mudar no tempo.

#### 8.2.4 Negociação e Renegociação de QoS

Durante o processo de negociação, os seguintes passos são realizados:

- Os parâmetros de QoS são transladados e negociados de uma camada (ou um subsistema) para outro.
- Cada camada ou subsistema deve determinar se ele pode suportar o serviço requerido. Em caso afirmativo, certos recursos são reservados para a sessão. Quando todos os subsistemas aceitam os parâmetros de QoS a sessão é estabelecida. Senão a sessão é rejeitada. Neste último caso, sistemas sofisticados

podem indicar ao usuário qual nível de QoS pode ser suportado. Se o usuário está contente com o nível de qualidade sugerida a sessão é estabelecida.

As comunicações multimídia não são estáticas. Durante uma sessão ativa, trocas na QoS podem ser necessárias por várias razões:

- O usuário decide reduzir a qualidade de apresentação ou eliminar certos canais.
- O usuário decide aumentar a qualidade de apresentação.
- Necessidade de um canal extra para acessar informações multimídia adicionais.

Portanto, é necessário fornecer mecanismos de renegociação para satisfazer trocas de requisitos de comunicações multimídia. Algumas vezes não é possível satisfazer requisitos para aumentar a QoS porque os recursos requeridos podem não estar disponível.

### **Escalamento de Mídia ou Degradção Suave da Qualidade**

Outro problema com um contrato de QoS estático entre os emissores e os receptores de um fluxo multicast é a variância de muitos parâmetros durante uma transmissão nos nós intermediários e dentro da rede. Poderia ser interessante ajustar os parâmetros de QoS durante a conexão multimídia. Quando aplicado ao fluxo de dados multimídia, isto é chamado de escalamento de mídia ou degradação suave da qualidade.

Os protocolos orientados a conexão tradicionais tem sempre suportado um método simples de escalamento de mídia, chamado controle de fluxo, prevenindo um emissor rápido de sobrecarregar um receptor lento. Note que o único parâmetro controlado por estes algoritmos é a taxa de dados. Isto é razoável para conexão não tempo-real (por exemplo, e-mail, transferência de arquivo, etc.), mas ele obviamente dificulta a transferência isócrona de fluxos tempo-real.

Escalamento de mídia permite o controle de outros parâmetros além da taxa de dados. Em um fluxo de vídeo, a qualidade da imagem é um candidato óbvio.

Para implementar escalamento de mídia, a interface entre a aplicação e a rede deve ser estendida para passar informações de controle para cima e para baixo. Por exemplo, se a rede sinaliza sobrecarga, um codificador de vídeo MPEG poderia ajustar o tamanho de passo de quantificação: quanto maiores os passos de quantificação menor será a taxa de dados e a qualidade da imagem.

Muitos projetos de pesquisa estudam o escalamento de mídia ([Käppner, 94], [McCanne, 96], [Wittig, 94]), mas não há padrões ainda para interface aplicação/rede ou para algoritmos e protocolos dentro da rede.

#### **8.2.5 Outros usos da declaração de QoS**

Nas seções anteriores, os parâmetros de QoS foram usados pelo sistema como critério para aceitar ou recusar (possivelmente com renegociações) comunicações. Em outros esquemas, pode-se usar estes parâmetros para outros fins [Fluckiger, 95]:

- Requisitos Hard (Obrigatórios): onde certos requisitos são obrigatórios. Isto pode ser o caso para uma transmissão de vídeo tempo-real que não poderia iniciar se uma taxa de bits não possa ser fornecida pela rede.
- Requisitos QoS usados pela rede para otimizar recursos internos: os parâmetros de QoS são usados pela rede como uma pura indicação para otimizar recursos internos, escolhendo rotas menos carregadas. Algumas redes usam a QoS neste sentido.

Como a definição e funcionamento de gerenciamento de QoS não é ainda universalmente bem estabelecida, outras formas de operação podem existir.

#### **8.2.6 Diferentes Níveis de Garantia**

Até aqui nós falamos que a QoS deveria ser garantida. Na prática, o usuário pode especificar um grau (ou nível) de garantia. Em geral, existem três níveis de garantia:

- **Garantia Determinista ou Hard:** a QoS especificada pelo usuário é garantida a 100%. Esta garantia é mais custosa em termos de recursos, pois os recursos são alocados na base pior caso, e eles não podem ser usados por outras aplicações mesmo quando não estão sendo usados, resultando num baixo uso dos recursos. Por exemplo, um vídeo codificado a taxa de bits variável com uma variação na saída de 200kbit/s a 4Mbit/s, os recursos deveriam ser reservados baseado em 4Mbit/s (pior caso). Uma vantagem é que esta garantia é de fácil implementação, pois recursos são reservados estaticamente.
- **Garantia Estatística ou Soft:** a QoS especificada pelo usuário é garantido em uma certa percentagem. Esta garantia é mais adequada para mídias contínuas pois elas não necessitam precisão de 100% na apresentação. Além disso, o uso de recursos é mais eficiente. Ele é baseado em multiplexação estatística: recursos não usados por uma aplicação podem ser usados por outras. Este é um modo desejado para comunicação multimídia, mas ele é de difícil implementação devido à natureza dinâmica do tráfego e uso dos recursos.
- **Melhor Esforço:** neste caso nenhuma garantia é fornecida e a aplicação é executada com os recursos disponíveis. Os sistemas computacionais tradicionais operam neste modo.

### 8.2.7 Fornecendo Garantias de QoS

QoS pode ser garantida apenas quando recursos suficientes são disponíveis e o escalonamento de processos é adequadamente implementado. A prevenção de sobrecargas requer **controle de admissão** e a prevenção de que aplicações não utilizem mais recursos do que aquele alocado requer **mecanismos de policiamento**.

Desde que QoS fim-a-fim é necessária, cada subsistema deveria ter funções de gerenciamento de QoS, incluindo cinco elementos: especificação de QoS, controle de admissão, negociação e renegociação, alocação de recursos e policiamento de tráfego. Várias arquiteturas QoS para gerenciar estes subsistemas tem sido propostas. Estas arquiteturas fornecem uma abordagem de trabalho na qual todos os subsistemas devem cooperar para fornecer as garantias de QoS.

### 8.2.8 Um exemplo de manipulação de QoS

Um exemplo simples ilustra como os conceitos de QoS apresentados anteriormente são usados na prática. Suponha que um cliente deseja obter e apresentar uma peça de áudio de qualidade telefone de um servidor remoto. Os seguintes passos estabelecem a sessão para o usuário:

- O usuário seleciona o nome do arquivo de áudio e a qualidade de telefone através de uma interface com o usuário.
- A aplicação traduz o requisito do usuário nos seguintes parâmetros: taxa de amostragem - 8 kHz (possivelmente com uma pequena variação), bits por amostra = 8 (uma amostra deve ocorrer a cada 125 µs).
- A aplicação passa o pedido ao sistema operacional cliente, que determina se ele pode processar um byte todo 125 µs com a carga atual no cliente. Se não, a sessão é rejeitada.
- O sistema operacional passa o pedido ao sistema de transporte incluindo protocolo de transporte e todas as camadas mais baixas, que determina se ele pode suportar uma taxa de bits de cerca de 64 kbit/s. Se não, o pedido de sessão é rejeitado.
- O servidor passa o pedido ao controlador de disco para determinar se ele pode suportar uma taxa de transferência de 64 kbit/s sem afetar a QoS das sessões existentes. Se não, o pedido de sessão é rejeitado.
- A sessão é estabelecida com sucesso e o usuário ouve a peça de áudio pedida.

## 8.3 Reserva de Recursos em Avanço

Nós dissemos que recursos devem ser reservados para fornecer garantias de QoS. Até aqui nós assumimos que a reserva de recursos inicia com a tentativa de reserva e termina em um tempo não especificado. Quando os recursos não são disponíveis, a sessão é rejeitada. Isto é indesejável para aplicações importantes com hora marcada, tal como videoconferências importantes. Mas também não é aceitável interromper aplicações correntes devido a sessões importantes. Para resolver este problema, reserva de recursos em avanço é utilizada.

Na reserva de recursos em avanço, os usuários podem reservar recursos em avanço especificando o tempo de partida e a duração da sessão. A aceitação do pedido é sujeito a disponibilidade de recursos pedidos no tempo requerido. Um problema do projeto deste esquema é como dividir o conjunto total de recursos entre sessões reservadas em avanço e sessões convencionais, como especificar precisamente a duração da reserva e como informar aos usuários que tem reserva no caso de falha do sistema. Existem duas soluções para o primeiro problema: partição dos recursos em duas partes ou todas as sessões compartilham os recursos. O segundo problema é mais difícil de resolver.

## 8.4 Aplicações Adaptativas

Alguns pesquisadores propõem uma abordagem muito radical à QoS e escalonamento de mídia (seção 8.2.4): eles se opõem a qualquer espécie de reserva de recurso dentro da rede e argumentam que as aplicações deveriam ser capaz de se adaptar às mudanças arbitrárias nas condições de rede [Bolot, 94] [Diot, 95]. Os principais argumentos são:

- Reserva de recurso é muito cara em termos de mensagens de protocolo.
- Reservas excessivas leva a um uso ineficiente de recursos.
- Reservas Hard (a 100%) não podem adaptar a fluxos com variação de QoS durante sua vida.
- Se a rede é realmente muito congestionada, você pode sempre lançar largura de banda ao problema.

Assim, a idéia é transmitir fluxos multimídia sob redes tradicionais do tipo melhor esforço, tal como a Internet. O emissor pode variar sua taxa de pacotes a qualquer momento. Se a rede sinaliza uma sobrecarga, as aplicações terão mecanismos específicos para reagir, tal como a modificação do parâmetro de quantização em MPEG ou a escolha de um outro algoritmo de compressão de áudio. Aplicações adaptativas usam escalamento de mídia como sua única maneira de ajustar a QoS; não há protocolo de reserva ou contrato de QoS com a rede.

## Capítulo 9 Suportes de Rede para Multimídia

Neste capítulo, nós analisaremos como diversas tecnologias de rede satisfazem os requisitos de aplicações multimídia distribuídas. Inicialmente, alguns dos requisitos de redes para a comunicação apresentados no capítulo 7 serão quantificados. Estes números servirão como métrica entre as diferentes tecnologias de rede.

O sucesso de uma tecnologia de rede depende não apenas dos aspectos funcionais, mas muito mais dos custos na introdução nas infra-estruturas existentes, com relação a integração do sistema-final, cabeamento, interconexão e gerenciamento. Portanto, nós investigaremos a evolução das infra-estruturas de rede, para entender melhor o papel que as diferentes tecnologias podem ter.

### 9.1 Requisitos de Rede para Comunicação Multimídia

Para avaliar se as principais tecnologias de rede satisfazem os requisitos identificados no capítulo 7, nós analisaremos as seguintes categorias como uma métrica:

- **Vazão:** ao menos 1,4 Mbps por direção. Como apresentado no capítulo 7, este valor tem mostrado muito interessante pois fornece uma boa qualidade de vídeo e permite o uso de equipamentos audiovisuais comerciais (CD player). Portanto, a largura de banda da rede deveria ser de ao menos 100 Mbps para LANs e maior para WANs.
- **Atraso de transmissão:** um máximo de aproximadamente 10 a 15 ms. Este valor não é preciso, mas uma aproximação que pode servir como referência para a discussão posterior. Aqui nós não tratamos variação de atraso separadamente, pois um esquema de bufferização de equalização de atraso pode ser usado para obter o comportamento síncrono fim-a-fim, e o atraso adicional é somado ao atraso fim-a-fim. Mas para limitar este atraso fim-a-fim, a variação de atraso deve ser limitada.
- **Ser baseada em comutação de pacotes a multiplexação estatística** em vez de circuitos dedicados para compartilhamento eficiente de recursos da rede.
- **Comunicação multicast:** como diversos tipos de aplicações multimídia exigem a difusão, é necessário que a rede tenha suportes de multicasting.
- **Confiabilidade:** mecanismos de controle de erro ou recobrimento integrados à rede. LANs suportam um mecanismo em hardware de checksum e corta quadros corrompidos dentro de sua camada MAC. Dado a confiabilidade inerente de muitas LANs de hoje, isto é uma estratégia aceitável para dados tão bem quanto para fluxos multimídia. Diferentes estratégias são principalmente encontradas em WANs, assim nós discutiremos a confiabilidade apenas a nível de WAN.

### 9.2 Alguns Conceitos Importantes

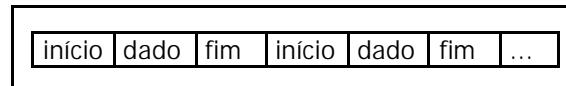
Antes de analisarmos se as principais tecnologias de rede satisfazem os requisitos de comunicação multimídia, é importante introduzir alguns conceitos importantes usados na discussão de redes de alta velocidade.

#### 9.2.1 Transmissão de Dados Assíncrona e Síncrona

No mais baixo nível de comunicação, dados são transmitidos bit a bit sobre um meio de transmissão. Nós usamos o termo assíncrono ou síncrono para descrever como os bits de dados são transmitidos:

- Em *transmissão assíncrona*, os relógios no transmissor e no receptor não são sincronizados. Caso não haja dados a transmitir, o meio de transmissão se

mantém em um estado “ocioso”. Se o transmissor tem dados para transmitir, ele envia antes dos dados um sinal de partida (um padrão de bits) para notificar o receptor do envio de dados (*início* na Figura 77). Uma vez partida, os bits são enviados em uma taxa de bits fixa. Por fim, o transmissor envia ao receptor um sinal de fim para informar o final da transmissão de dados. O tempo de transmissão não é estruturado: o transmissor pode partir a transmissão de dados a qualquer tempo. O termo assíncrono refere-se a este caráter aleatório do tempo de transmissão de dados: a transmissão de dados pode começar a qualquer momento.



**Figura 77.** Parte da unidade de informação na transmissão assíncrona

- Na transmissão sincronizada, os relógios no transmissor e no receptor estão sincronizados. O tempo é dividido em intervalos de tamanho fixo. Um intervalo corresponde a um bit. Bits de dados são transmitidos continuamente sobre o meio de transmissão sem qualquer sinal de início e fim. O termo síncrono refere-se a este intervalo fixo de bit.

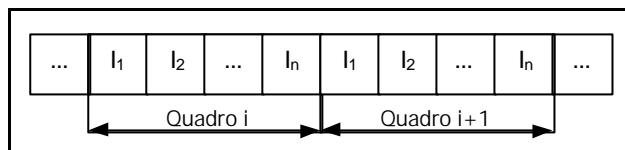
As vantagens de transmissão assíncrona são: geração de caracteres por meio de dispositivos eletro-mecânicos e transmissão de caracteres irregularmente espaçados no tempo. As desvantagens são: uma parte considerável do que é transmitido não transporta informação útil: a sincronização depende dos *início/fim* que podem não ser detectados por causa de distorções do sinal. Uma utilização da transmissão assíncrona é quando não se necessita de transmissão freqüente de informação.

A transmissão síncrona apresenta como vantagem uma melhor proteção contra erros, pois ao término de cada bloco uma configuração de *bits* para detecção de erros pode ser enviada; é mais eficiente pois a proporção de mensagem transmitida como informação em relação à configuração de sincronização é maior que na transmissão assíncrona; não é tão sensível à distorção e opera a velocidades bem mais altas que no modo assíncrono. As desvantagens são: caso haja erro de sincronização, todo bloco é perdido; os caracteres são enviados em blocos e não antes destes poderem ser formados, obrigando que os equipamentos sejam dotados de memória de armazenamento para a coleta dos caracteres, até que se forme o bloco com o comprimento usado pelo equipamento. Memória, nesse caso, são *buffers*, o que encarece seu custo.

### 9.2.2 Multiplexação assíncrona e síncrona

Multiplexação aqui refere-se ao compartilhamento do meio de transmissão por várias conexões distintas (lógicas ou virtuais). Uma técnica de multiplexação muito utilizada é a chamada *multiplexação por divisão de tempo* (TDM) em que o tempo de transmissão do meio é compartilhado entre várias conexões ativas.

- Na multiplexação síncrona, o tempo é dividido em quadros de tamanho fixo que por sua vez são divididos em intervalos de tamanho fixo (Figura 72). Por exemplo, assuma que todo quadro de transmissão é dividido em 10 intervalos e eles são numerados de 1 a 10. Se o intervalo 1 é atribuído a uma conexão, o emissor pode transmitir dados sob esta conexão apenas no intervalo 1. Caso ele tenha mais dados a transmitir após este intervalo, o transmissor deve aguardar o próximo quadro. Se ele não usa este intervalo temporal, nenhuma outra conexão pode utilizá-lo. Este tipo de multiplexação é chamada *multiplexação por divisão de tempo síncrona* (STDM).
- Na *multiplexação por divisão de tempo assíncrona* (ATDM), os intervalos temporais podem ter tamanho fixo ou variável e estes intervalos não são atribuídos a nenhuma conexão. Uma conexão pode usar qualquer intervalo de tempo se ele não está sendo utilizado por outra conexão. Uma ATDM específica é ATM em que os intervalos temporais têm intervalos temporais fixos e pequenos.



**Figura 78.** STDM: Multiplexação por divisão de tempo síncrona

É claro que STDM deve ser usada para transmissão síncrona a nível de bits. Por outro lado, ATDM pode usar transmissão síncrona ou assíncrona a nível de bits. ATM é assíncrona a nível de multiplexação mas síncrona a nível de bits.

### 9.2.3 Comunicações Isócronas

Comunicação isócrona refere-se ao serviço de rede que fornece garantias de uma largura de banda fixa e sem variação de atraso para uma conexão. Este tipo de serviço é normalmente provido por uma rede a comutação a circuito.

Áudio digital e vídeo descompactado geram dados isocronamente, isto é, em uma taxa fixa. Serviço isócrono pode suportar a comunicação de áudio e vídeo, mas isto não é um requisito necessário pelas seguintes razões:

- áudio e vídeo são normalmente compactados e o fluxo compactado tem taxas de bits variáveis. Usando um serviço isócrono significa um desperdício da largura de banda.
- comunicações de áudio e vídeo podem tolerar uma pequena variação de atraso. Um buffer de equalização de atraso pode ser usado para remover esta variação de atraso em muitas aplicações.

## 9.3 ISDN

A ISDN (*Integrated Services Digital Network*) ou RDSI (Rede Digital de Serviços Integrados) é uma rede de longa distância definida pelo CCITT (*International Telegraph And Telephone Consultative Committee*), hoje chamado de Setor de Padrões de Telecomunicações da ITU (*International Telecommunications Union*). A rede ISDN é considerada uma evolução das antigas redes telefônicas analógicas, fornecendo um suporte a uma grande variedade de diferentes serviços, incluindo serviços de voz e outros (dados, imagens, vídeos, etc). Ao contrário da maioria das redes telefônicas em uso hoje no Brasil, onde a rede pode ser digital, mas o acesso doméstico é analógico, a rede ISDN fornece uma conectividade digital fim-a-fim.

A rede ISDN utiliza a infra-estrutura das antigas redes telefônicas analógicas: os cabos são reutilizados, mas um sinal digital, em vez do sinal analógico é transmitido. Os comutadores das companhia telefônica devem ser suportar conexões digitais.

### **ISDN é serviços integrados**

A integração de diferentes serviços é a marca da ISDN. No passado, vídeo, áudio, voz e serviços de dados necessitavam de redes separadas:

- Vídeos eram distribuídos em linhas coaxiais
- Áudio sobre linhas balanceadas,
- Voz sobre pares de cabos de cobre
- Dados necessitavam de cabos coaxiais ou par trançados.

Este ambiente de plantas múltiplas era de instalação e manutenção caras. ISDN é diferente. Ela integra voz, vídeo, áudio e dados sobre uma mesma rede. Estes serviços avançados são disponível, em grande parte, pois ISDN é digital.

### **ISDN é Digital**

Uma das grandes desvantagens das velhas redes telefônicas é que elas eram analógicas. Isto é particularmente um problema para os sistemas digitais. Computadores são dispositivos digitais, realizar a comunicação de dados sobre linhas analógicas é

complicado. Um processo chamado de modulação é usado para tomar os uns e zeros binários dos computadores e converter-los em tonalidades analógicas. E no lado do receptor, um processo chamado demodulação converte estas tonalidades em seus equivalentes digitais. A partir do processo de modulação demodulação, nós obtemos o nome do dispositivo que realiza esta conversão, modem. Os modems são muito inefficientes, tonalidades podem ser corrompidas por ruídos ecos e pontas de linha e o tamanho de banda é limitado. Modems atualmente disponíveis tem velocidades de 33,6 e 56 Kbps, mas são limitadas pela qualidade da conexão analógica e geralmente não chegam a 26,4 ou 28,8 Kbps. Usando redes ISDN, o usuário tem acesso diretamente a um sinal digital.

### **Canais ISDN**

ISDN foi projetado principalmente sobre a noção de canais de 64 Kbps, síncronos, comutados a circuitos, que podem ser usados para tráfegos contínuos orientados a fluxo de bits ou para comunicação em pacotes. O número 64 Kbps surgiu do fato que este é essencialmente a taxa em que linhas analógicas são amostradas por companhias telefônicas digitais (8000 amostras por segundo, 8 bits por amostra).

ISDN é essencialmente combinações destes canais de 64 Kbps e também canais de 16 ou mesmo de 64 Kbps usado para sinalização. Os canais de 64kbps são chamados de canais B (*Bearer* = Transportador), e os canais de sinalização são chamados de canais D (Delta ou *Data*).

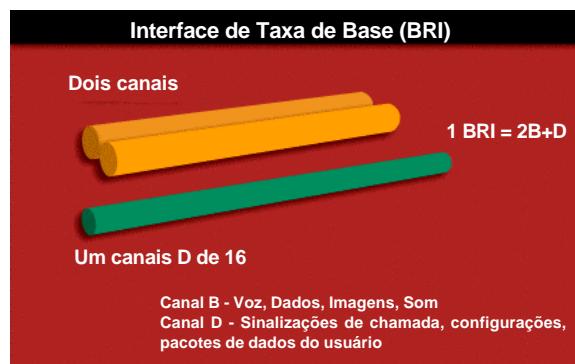
- O bloco básico de construção do ISDN é o canal B. Ele é projetado para voz, vídeo e dados. Ele suporta serviços síncronos e assíncronos numa taxa de até 64 kbps. Canais B podem ser agregados para aplicações de maior largura de banda. Cada canal B é tratado independentemente pela rede, permitindo a conexão simultânea de voz, dados, etc.
- O canal D é um canal de 16 ou 64 kbps (dependendo do tipo de serviço) usado para sinalização (comunicação entre a chave telefônica e o dispositivo ISDN). Sinalização é a troca de informação para chamadas de configuração e controle. Sinalização indica uma chamada telefônica, identifica o chamador, o tipo de chamada (voz/dado), e o número que foi chamado. Assim, em vez da companhia telefônica enviar um sinal elétrico para tocar o telefone, ela envia um pacote digital neste canal. O canal D fornece uma partição de 9,6 Kbps para dados do usuário.

Assim, ISDN permite múltiplos canais digitais operarem simultaneamente sobre o mesmo cabo da linha telefônica regular. Antes era necessário ter uma linha para cada dispositivo que se desejava usar simultaneamente: uma linha para um telefone, um fax, um computador, etc. Uma conexão ISDN residencial suporta até 2 conexões de voz sobre uma mesma linha física ou fax ou conversação de PC.

### **Interface de Taxa de Base BRI**

A grande vantagem da ISDN é que ela pode ser construída em vários tamanhos. A porção de tamanho simples é chamada de Interface de Taxa de Base BRI (Figura 79). Uma BRI tem dois canais B de 64 Kbps e um canal D de 16 Kbps (2B+D), fazendo uma taxa total de 144 Kbps. Este serviço de base satisfaz as necessidades da maioria dos usuários individuais, incluindo residencial e micro-pequenas.

Alguns sistemas utilizando BRI fornecem a facilidade de usar canais B adicionais para permitir taxas de transmissão mais rápidas. Isto é conhecido como agregação de canal ou multiplexação inversa. Por exemplo, um sistema de videoconferência poderia agragar 3 conexões 2B+D (6 canais B) para fornecer uma largura de banda de 384 Kbps. Esta capacidade pode ser implementada em hardware, mas softwares padrões estão em desenvolvimento para fornecer a mesma funcionalidade.



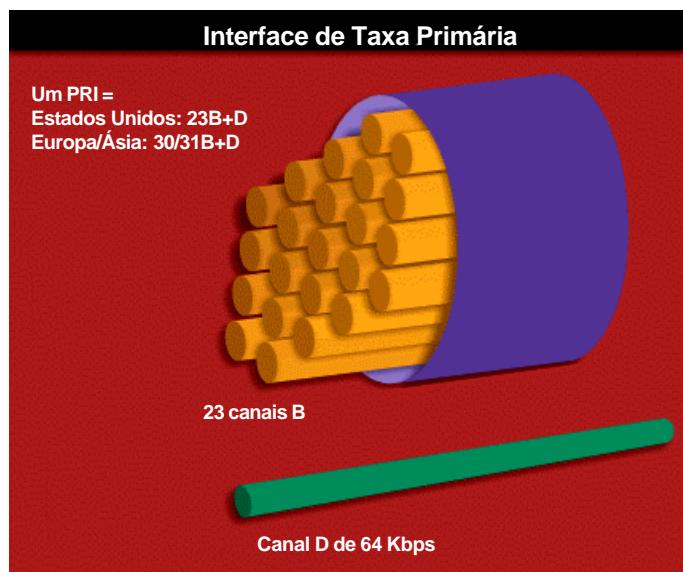
**Figura 79.** Interface de Taxa de Base (BRI)

Em países que dispõem de redes ISDN, para acessar serviços BRI, é necessário subscrever-se a uma linha telefônica ISDN. Clientes necessitam de equipamentos especiais para comunicar-se com as chaves da companhia telefônica, e com outros dispositivos ISDN.

Na prática, nos EUA, o serviço ISDN disponível inclui apenas um canal do tipo B e outro do tipo D. Um segundo canal B pode ser alugado com um custo extra. Além disso, algumas empresas oferecem canais B de apenas 56 Kbps.

### Interface de Taxa Primária PRI

A interface de taxa primária ou PRI nos Estados Unidos consiste até 23 canais B de 64 Kbps e um canal D de 64 Kbps, ou uma conexão 23B+D. Isto perfaz um total de banda passante de 1,544 Mbps através do padrão de tronco norte americano T-1. Na Europa, como os padrões de transmissão diferem um pouco, a PRI é fornecida com 30/31B+D.



**Figura 80.** Interface de Taxa Primária

O serviço primário é fornecido para usuários com maior requisitos de capacidade, como escritórios com PBX ou uma rede local.

Com uma PRI, pode-se optar pela combinação de vários canais B em um maior canal chamado canal H. Eles são implementados como:

- $H_0 = 384 \text{ Kbps}$  (6 canais B)
- $H_{10} = 1472 \text{ Kbps}$  (23 canais B)
- $H_{11} = 1536 \text{ Kbps}$  (24 canais B)
- $H_{12} = 1920 \text{ Kbps}$  (30 canais B) para a Europa/Ásia apenas.

## ISDN e a Multimídia

Olhando as redes de longa distância, a ISDN (Integrated Services Digital Network) ou RDSI (Rede Digital de Serviços Integrados) foi projetada para dar suporte a uma grande variedade de diferentes serviços, incluindo dados alfanuméricos, voz e vídeo. Ela fornece canais de 64 Kbps, síncronos, comutados a circuitos, que podem ser usados para tráfegos contínuos orientados a fluxo de bits ou para comunicação em pacotes.

Existem várias vantagens da ISDN comparada a outras redes de longa distância para a comunicação multimídia:

- Grande disponibilidade;
- Escalabilidade de largura de banda na interface a taxa primária, em passos de 64 Kbps;
- Características isócronas;
- Suporte a taxa de bits constante tão bem quanto tráfegos em pacotes.

ISDN não suporta qualquer função multicast dentro da rede. Mas, para fornecer conexões ponto-a-multiponto, em particular para aplicações de conferência, equipamentos externos ao ISDN, chamados Multicast Control Units (MCUs), estão sendo utilizados. Cada estação participando de uma conferência ISDN constrói uma conexão ponto-a-ponto com o MCU. O MCU então mixa todos os canais de áudio de entrada e distribui o sinal combinado sobre as ligações ponto-a-ponto para todas as estações participantes. Embora o conceito de mixagem é útil para áudio, ele não é diretamente aplicável para vídeo ou outros fluxos de dados. Quando se utiliza MCU para vídeo ou dado, a aplicação deve selecionar um fluxo de entrada para distribuição e não mixar os vários fluxos de entrada.

ISDN fornece uma largura de banda limitada de até 2 Mbps com características isócronas. A ISDN é atualmente a única rede popular (na Europa e EUA) disponível para comunicações multimídia interativas de longa distância, excetuando as linhas alugadas (de mais alto custo). A falta de serviços multicast requer o uso de um equipamento especial (MCUs) para configurar conferências multiponto ou serviços de distribuição.

## 9.4 Ethernet

A rede Ethernet a 10 Mbps baseada em CSMA/CD (IEEE 802.3) é a tecnologia de rede locais mais utilizada. Segundo a IDC (*International Data Corporation*), mais de 85% de todas as redes instaladas até o fim de 1997 eram Ethernet. Isto representa mais de 118 milhões de PCs, estações de trabalho e servidores conectados. Todos os sistemas operacionais e aplicações populares são compatíveis com Ethernet, como são os protocolos da camada acima dele, como o TCP/IP (*Transmission Control/Internet Protocol*), IPX<sup>11</sup>, NetBEUI<sup>12</sup> e DECnet<sup>13</sup>.

Vários fatores contribuíram para tornar Ethernet a tecnologia de rede mais popular:

- **Confiabilidade:** a confiabilidade da rede é uma característica crítica para o sucesso de uma empresa, assim a tecnologia de escolha deve ser de fácil instalação e suporte. Desde a introdução em 1986 dos concentradores centrais ou hubs<sup>14</sup> 10BASE-T, o cabeamento tem continuado a evoluir e hubs e comutadores tiveram suas confiabilidades aumentadas.

<sup>11</sup> *Internetwork Packet eXchange*, um protocolo NetWare (sistema operacional de rede desenvolvido pela Novell) similar ao IP (Internet Protocol).

<sup>12</sup> Protocolo Microsoft para seus produtos Windows NT e LAN Manager.

<sup>13</sup> Arquitetura de rede proprietária da DEC (Digital Equipment Corporation), um sistema para redes de computadores. Ela opera sob redes ponto-a-ponto, X.25 e Ethernet.

<sup>14</sup> Hubs são com freqüência utilizados para conectar dois ou mais segmentos Ethernet de qualquer mídia de transmissão. Um hub contém portas múltiplas. Quando um pacote chega em um porto, ele é copiado para outro porto de modo que todos os segmentos da LAN possam ver todos os pacotes. A construção dos *hubs* teve uma evolução contínua no sentido de que os mesmos não implementem somente a utilização do meio compartilhado, mas também possibilitem a troca de mensagens entre várias estações simultaneamente. Desta forma as estações podem obter para si taxas efetivas de transmissão bem maiores. Esse tipo de elemento, também central, é denominado comutador (*switch*).

- **Disponibilidade de Ferramentas de gestão e diagnóstico:** ferramentas de gestão para Ethernet, possíveis graças a adoção de padrões de gestão incluindo o protocolo SNMP (*Simple Network Management Protocol*) e seus sucessores, permite a um administrador ver o estado de todos os computadores e elementos de rede. Ferramentas de diagnóstico Ethernet suportam vários níveis funcionais, desde uma simples luz de indicação de ligação a analisadores de rede sofisticados.
- **Extensibilidade:** o padrão Fast Ethernet, aprovado em 1995, estabeleceu Ethernet como uma tecnologia extensível. Hoje em dia, o desenvolvimento da Gigabit Ethernet, aprovado em 1998, ampliou a extensibilidade da Ethernet ainda mais. Agora as escalas Ethernet vão de 10, 100 e 1000 Mbps.
- **Baixo custo:** o preço por porta Ethernet está reduzindo a cada dia.

Assumindo que alguma largura de banda é necessária deixar para tráfego de dados e de controle e que Ethernets não poderiam ser mais carregadas que 70% a 80% para manter as colisões a um nível aceitável, então apenas 5 a 6 Mbps são realmente disponíveis para fluxos multimídia. Assim, não mais que quatro fluxos de vídeo compactados em paralelo.

Outro problema da Ethernet é o comportamento não determinista do método de acesso CSMA/CD (*Carrier Sense Multiple Access with Collision Detection*), já discutido no capítulo 7. Em situações de alta carga, o CSMA/CD não permite o controle do tempo de acesso e da largura de banda para muitas aplicações. Se uma aplicação tradicional, tal como acesso a arquivo remoto, tentar utilizar uma grande porcentagem da largura de banda disponível, nenhum mecanismo existe para assegurar uma distribuição equalitária da largura de banda. Além disso, Ethernet não fornece mecanismos de prioridade, e assim não se pode dar um tratamento diferenciado para tráfego tempo-real sobre dados convencionais.

Embora na teoria Ethernet possa distinguir até  $2^6$  endereços multicast, na prática um adaptador Ethernet gerencia um número limitado de endereços de grupo multicast.

Apesar dos seus problemas, muitas aplicações multimídia experimentais de hoje usam Ethernet como mecanismo de transporte, geralmente em um ambiente controlado e protegido. Sem mais que três estações ativas em um segmento Ethernet participando em uma aplicação de conferência, não há uma real competição da largura de banda. Nesta espécie de ambiente, Ethernet é perfeitamente desejável como transporte de rede.

Concluindo, devido a falta de garantias de atraso, Ethernet não é uma boa rede para multimídia distribuída. Mas, ela fornece uma largura de banda suficiente para alguns fluxos mais uma função multicast, que torna-a desejável para aplicações experimentais com um número limitado de estações.

## 9.5 Isochronous Ethernet

Uma variante do Ethernet, chamada Isochronous Ethernet (Iso-Ethernet), tem sido proposta pelo padrão IEEE 802.9A Integrated Voice Data LAN (IVD LAN). O objetivo deste padrão é fornecer um serviço de pacotes MAC parecido com IEEE 802 juntamente com canais isócronos tal como ISDN em um par trançado UTP único para um terminal. Existem poucos produtos comerciais disponível.

Uma técnica de codagem mais eficiente adiciona 6 Mbps ao cabo padrão Ethernet, fornecendo 96 canais B isócronos (64 kbps) e um canal Ethernet 10 Mbps padrão.

Iso-Ethernet é uma abordagem a meio compartilhado com uma largura de banda relativamente limitada e sem suporte para multicasting. Ela fornece tráfego verdadeiramente isócrono, o que significa uma característica de atraso ótima. Sua estrutura em canal tal como ISDN é projetada para transmissão de áudio ou vídeo codificado com H.261, mas ele falha para fluxos codificados MPEG. Assim, nós podemos ver este como uma extensão de rede local ISDN sob uma rede de base existente (Ethernet) melhor que uma solução geral a comunicação multimídia integrada.

## 9.6 Fast Ethernet (100Base-T)

Organizações modernas dependem de suas redes locais (LANs) para prover a conectividade de um número crescente de computadores executando aplicações cada vez mais complexas e críticas ao funcionamento da organização. Dentre estas novas aplicações podemos citar gráficos de alta resolução, vídeo e outros tipos de mídia. Como o volume de tráfego da rede tende a aumentar a cada ano, a largura de banda oferecida pelas redes Ethernet típicas, 10 Mbps, se tornou rapidamente inadequada principalmente para manter um desempenho aceitável apesar de um número crescente de computadores conectados a rede.

Dentre as tecnologias de LAN a alta velocidade disponíveis hoje, Fast Ethernet, ou 100BASE-T, se tornou um líder de escolha. Construída a partir da Ethernet 10BASE-T, usando o mesmo protocolo de acesso CSMA/CD, a tecnologia Fast Ethernet fornece uma evolução razoável de velocidade, chegando a 100 Mbps.

Em termos de velocidade a Fast Ethernet é relativamente rápida, mas ainda não é satisfatória devido ao protocolo MAC CSMA/CD e compartilha as mesmas limitações da Ethernet 10 Mbps com relação as características de atraso de acesso.

Como o padrão Ethernet, a máxima faixa de utilização da largura de banda varia de 50% a 90%, dependendo da configuração a tamanhos dos quadros. Esta taxa fornece uma vazão suficiente para um grande número de fluxos multimídia.

Concluindo, Ethernet 100Base-T fornece uma vazão suficiente para vários fluxos multimídia em paralelo. Mas garantias de atraso não podem ser dadas, em particular, qualquer estação na rede pode quebrar o fluxo multimídia através de um tráfego pesado. Multicasting é disponível. Portanto, 100Base-T é uma escolha aceitável para pequenas a médias configurações, mas não é uma alternativa ideal pois ela produz uma subutilização das capacidades da rede 100Base-T e o bom comportamento de todas as estações.

## 9.7 Gigabit Ethernet

Apesar da evolução em taxa de bits do 100BASE-T, já existe hoje uma clara necessidade de uma nova tecnologia de rede de mais alta velocidade a nível de backbone e servidores. Idealmente, esta nova tecnologia deveria também ser um caminho de atualização suave, não ser de custo proibitivo e não requerer novos treinamentos dos usuários e gerenciadores de rede.

A solução de escolha da IEEE para o problema anterior é a rede Gigabit Ethernet. Gigabit Ethernet fornece uma largura de banda de 1 Gbps para redes a nível de campus com a simplicidade da Ethernet de baixo custo comparada as outras tecnologias de mesma velocidade. Ela oferece um caminho de atualização (*upgrade*) natural para as atuais instalações Ethernet.

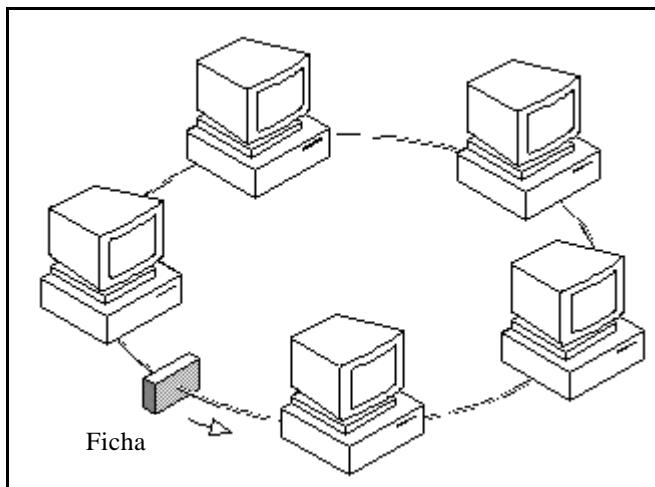
Gigabit Ethernet emprega o mesmo protocolo CSMA/CD (*Carrier Sense Multiple Access with Collision Detection*), o mesmo formato de quadro e mesmo tamanho de quadro de seus predecessores (Ethernet e Fast Ethernet). Para a vasta maioria de usuários da rede, isto significa que os investimentos feito nas redes já instaladas não serão perdidos e estas redes instaladas podem ser estendidas para velocidades gigabit com um custo razoável. Além disso, sem a necessidade de reeducar suas equipes de suporte e usuários.

Gigabit Ethernet completa a Ethernet oferecendo conexões de alta velocidade para servidores e um backbone extensão natural para bases instaladas de Ethernet e Fast Ethernet. Um dos maiores desentendimentos acerca do Gigabit Ethernet é que muitas pessoas na indústria dizem que o suporte de serviços tempo real é um problema de protocolo. Isto não é completamente verdadeiro: a camada física de uma rede que fornece garantias de qualidade de serviço deveria oferecer controle de admissão de conexão e chegada de pacotes previsível. Gigabit Ethernet é uma tecnologia sem conexão que transmite pacotes de tamanho variável. Como tal, ela simplesmente não

pode garantir que os pacotes tempo-real tenham o tratamento que eles exigem. Concluindo, Gigabit Ethernet não fornece uma estrutura de comunicação ótima para multimídia.

## 9.8 Token Ring

Token-Ring pode operar a 4 ou 16 Mbps e todas as estações são conectadas em um anel lógico (Figura 81). Nesta rede local, uma mensagem especial, chamada de ficha, circula no anel se todas as estações estão em estado de espera. Quando uma estação deseja transmitir um quadro, ela deve aguardar a chegada da ficha e remove ela do anel antes da transmissão do quadro para outra estação. Quando o receptor obtém o quadro, ele seta um flag no quadro confirmando a recepção e libera o quadro para trás no anel. O originador detectando que o quadro foi recebido (ou não) libera uma nova ficha para permitir que outros sistemas tenham acesso ao anel.



**Figura 81.** Token Ring

Note que o token-ring tem um comportamento muito previsível, ele garante que todo sistema na LAN tenha oportunidade de transmitir e as fichas e os quadros de dados circulam de maneira temporalmente determinista. Nesta metodologia, cada estação tem um acesso igual à ficha, nenhum sistema tem prioridade sobre outro.

Uma opção em muitas implementações Token-Ring é o uso de diferentes prioridades de acesso ao anel. Elas podem ser usadas para separar dados tempo-real (alta prioridade) e dados normais (baixa prioridade). Existe um campo no quadro de dados indicando a prioridade. Quando uma estação recebe um quadro, ela verifica se a prioridade do dado a transmitir é superior à prioridade do quadro. Em caso afirmativo, a estação altera o campo prioridade do quadro para conter a sua prioridade. Quando o originador restaura a ficha, ele seta a prioridade da ficha igual a prioridade do quadro (a maior prioridade). Apenas as estações com maior ou igual prioridade poderão capturar a ficha e transmitir os dados. O Token Ring tem mecanismos para quebrar o monopólio de uma estação.

### **Token Ring e o Tráfego Multimídia**

O protocolo de acesso Token Ring é muito mais interessante que Ethernet para suportar dados multimídia. Uma razão é a maior largura de banda do Token Ring, que é de 16 Mbps (e Ethernet é de 10 Mbps).

A disponibilidade de prioridades a nível de MAC, juntamente com o atraso máximo de propagação de um quadro fixado, torna possível a transmissão de mídias contínuas garantidas. A principal vantagem é a natureza previsível da estação transmitir. No pior caso, ou o maior período de tempo que uma estação espera até obter uma ficha, é determinístico e calculado [Steinmetz, 95].

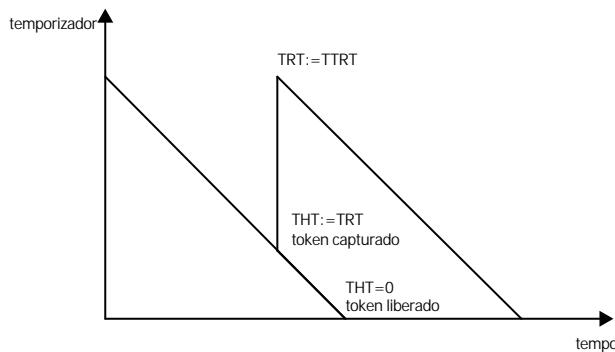
Token Ring fornece uma largura de banda suficiente para um número limitado de fluxos multimídia. Além disso, multicasting é disponível. Assim 16 Mbps Token Ring é uma rede viável para transporte de (poucos) fluxos multimídia.

## 9.9 FDDI

O protocolo FDDI (*Fiber Distributed Data Interface*) é uma extensão do padrão Token Ring. Ele é um padrão de rede local operando a 100 Mbps a fibra ótica e passagem de token.

O padrão FDDI especifica uma topologia em anel duplos, com cada anel operando a uma taxa de 100 Mbps. O uso do anel duplo aumenta a confiabilidade. Desde que em operação normal os anéis operam independentemente, nós precisamos considerar apenas a operação de um único anel a 100 Mbps. O acesso ao anel é controlado pela passagem de um token de permissão a volta do anel. Um nó contendo o token pode transmitir pacotes de dados por um período de tempo determinado pelas regras de contenção do token, e então passa o token para o próximo nó no anel. O token é passado imediatamente após a transmissão do último pacote de dados.

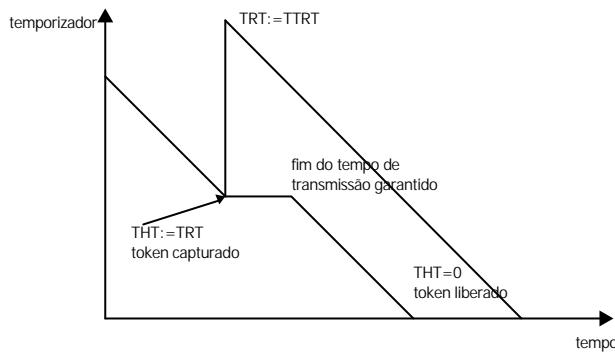
As regras de contenção do token são ditadas por um protocolo de token temporizado. Inicialmente, um Tempo Alvo de Rotação do Token (TTRT - *Target Token Rotation Time*) é negociado por todas as estações. Após isto, toda estação armazena o mesmo TTRT negociado. Além do TTRT, cada nó tem mais duas variáveis: temporizador de rotação do token (TRT - *Token Rotation Timer*) e temporizador de contenção do token (THT - *Token Holding Timer*). Em cada nó, TRT é iniciado com TTRT. TRT mede o tempo (descrescente) que passou desde a última recepção do token. Quando um nó recebe o token, ele copia o valor do TRT no temporizador THT e reseta TRT para TTRT. Então a estação podem enviar dados até o temporizador THT expirar. A Figura 82 ilustra o protocolo FDDI.



**Figura 82.** Protocolo token temporizado FDDI

Este esquema garante que cada estação terá acesso ao menos uma vez dentro de TTRT.

Além deste esquema, FDDI também suporta fluxos multimídia fornecendo uma classe de tráfego síncrono. A operação da classe de tráfego síncrono é mostrado na Figura 83. A cada estação é atribuída um intervalo de TTRT/n (n é o número de estações síncronas no anel) como largura de banda de transmissão síncrona garantida. Quando uma estação capture um token, ela pode enviar o tráfego da classe síncrona por TTRT/n unidades de tempo. Após este intervalo de tempo síncrono tenha passado, a estação pode transmitir os dados de tráfego assíncrono de acordo com o protocolo de token temporizado normal apresentado anteriormente. O tempo de rotação do token é portanto aumentado para 2xTTRT.



**Figura 83.** Classe de tráfego síncrono FDDI

Este esquema permite o tráfego síncrono, i.e. atraso limitado, com um limite de atraso ( $2 \times TTRT$ ) configurável no tempo de inicialização do anel. Infelizmente, um limite baixo de atraso leva a uma redução da utilização do anel. Na prática, o limite de atraso está entre 5 a 50 ms. Devido a alta largura de banda do FDDI, o atraso de transmissão é principalmente governado pelo atraso de acesso (tempo de transmissão para um pacote de 4 Kbytes é menos que 0,5 ms). Note que embora a classe de tráfego síncrono tenha sido construída no protocolo FDDI, poucas implementações de FDDI fornecem acesso a este tipo de tráfego.

FDDI suporta bem a comunicação multimídia graças a sua grande largura de banda, sua classe de tráfego síncrono e a possibilidade de multicasting. Quando o tráfego síncrono não é suportado por um adaptador, um esquema similar ao Token Ring, baseado em prioridade de tráfego e gerenciamento da largura de banda, pode ser usado.

FDDI não é capaz de manter uma conexão a taxa de dados contínua entre duas estações. A classe de tráfego síncrono do FDDI garante apenas uma taxa de dados mínima sustentável e atraso limitado, ele não fornece um fluxo de dados uniforme sem variabilidade. O atraso limitado é igual a  $2 \times TTRT$  e a mínima largura de banda é determinada por  $TTRT/n$ . Portanto, a escolha destes parâmetros é muito importante. Quando selecionados apropriadamente, o serviço síncrono pode suportar a comunicação de áudio e vídeo. Neste caso, o transmissor e o receptor deveriam ter buffers para tratar com a variação de atraso. No transmissor, a disponibilidade de dado a ser transmitido pode não coincidir com a chegada do token, assim o dado deve ser bufferizado. O tempo de bufferização é determinado por  $TTRT$  e as características de rajada do dado a ser transmitido. No receptor, o buffer é necessário para equalizar a variação de atraso.

Outra restrição de se usar FDDI para aplicações multimídia é que todas as estações compartilham um tamanho fixo de largura de banda, limitando o número total de sessões que podem ser suportadas.

Para fornecer este tipo de serviço isócrono, o FDDI básico foi estendido para FDDI-II. Visto a seguir.

## 9.10 FDDI-II

FDDI-II fornece um serviço a circuito comutado (emulado) e mantém o serviço comutado a pacote controlado por token do FDDI original. Com o FDDI-II é possível configurar e manter uma conexão a taxa de dados constante entre duas estações. Esta isocronia é provida pelo uso de quadros cíclicos a uma taxa de 8 kHz. A largura de banda total de 100 Mbps é cortada em até 16 canais de 6,144 Mbps isócronos (WBCs - Wide-Band Channels) e largura de banda FDDI assíncrona. A distribuição entre largura de banda isócrona e assíncrona é dinâmica. O esquema de gerenciamento necessário para alocar dinamicamente a largura de banda entre dados assíncronos e isócronos não está ainda definido.

Cada WBC é composto de 96 canais (B-channels) de 64 Kbps. Este esquema fornece uma boa interoperabilidade com o ISDN e Iso-Ethernet. De fato, FDDI-II pode ser visto como uma rede backbone para Iso-Ethernets.

Embora FDDI-II seja derivada do FDDI e indica que é compatível com FDDI, esta indicação é apenas parcialmente correta. Toda estação FDDI-II pode suportar o tráfego FDDI normal. Mas, uma estação FDDI normal não pode manipular o formato de quadro FDDI-II. Assim, quando apenas uma estação FDDI está no anel, as vantagens do FDDI-II não podem ser exploradas. Devido a este conflito, nenhum caminho de migração leva FDDI a FDDI-II, exceto o reuso das fibras óticas. Assim, FDDI-II não é concorrente muito forte pois é duvidoso que exista espaço suficiente para uma nova, complexa, proposta de rede de alta velocidade com o ATM já disponível.

FDDI-II foi projetado para suportar tráfego a taxa de bits constante. Ele fornece canais isócronos verdadeiros com atrasos na ordem de poucos milisegundos. A largura de banda é suficiente para um número limitado de estações, e multicasting é disponível. Assim, ele fornece os ingredientes para a comunicação multimídia. Mas quando a taxa de bits de áudio e vídeo são em rajadas, parte da largura de banda reservada é desperdiçada pois a largura de banda é normalmente reservada na taxa de pico do fluxo, e quando o fluxo não é na sua taxa de pico, outras estações não podem usar a largura de banda não utilizada. Também existem várias desvantagens em termos de complexidade e compatibilidade.

Outra restrição de se usar FDDI-II para aplicações multimídia é que todas as estações compartilham um tamanho fixo de largura de banda, limitando o número total de sessões que podem ser suportadas.

## 9.11 ATM (Asynchronous Transfer Mode)

ATM foi originalmente desenvolvido pela AT&T (em 1980) como uma técnica rápida de comutação de pacotes para fornecer a possibilidade de mixar voz e transmissão digital sobre uma única rede digital. Em 1988, ATM foi adotada pelo ITU-T (*International Telecommunication Union*) como o mecanismo de multiplexação e comutação para B-ISDN (*Broadband Integrated Service Digital Network*).

Progressos no desenvolvimento dos padrões ATM e tecnologias foram relativamente lentos até 1991, quando o ATM Forum, fundado por fabricantes de equipamentos, se focalizaram na definição de redes ATM privadas para redes locais e longas distâncias. Atualmente o ATM Forum com seus 800 membros, incluindo todos os principais provedores de equipamentos de dados e telecomunicações, e a ITU cooperam na padronização ATM, de maneira que os dois conjuntos de especificações são suficientemente similares.

A rede B-ISDN suporta um grande número de serviços, incluindo serviços de voz e outros (dados, imagens, vídeos, etc.). Ela é baseada no uso das fibras óticas, que podem acomodar 100 Gbits/s e tem uma taxa de erros bem menores que os pares trançados de cobre ( $10^{-12}$  contra  $10^{-6}$ ). Na B-ISDN, a taxa máxima de transferência é de magnitude de 155,52 Mbits/s ou 622,08 Mbits/s. A razão que estas taxas de transferências são necessárias é que comunicações comerciais e de automação de escritórios através de estações de trabalho e LANs são esperados realizar vários pedidos na rede de comunicação. Em particular, a B-ISDN com sua tecnologia ATM é um modo eficiente para implementar a comunicação multimídia, com informações de vídeo adicionada as informações convencionais de voz e dados.

### Deficiências do STDM tradicional

O termo *modo de transferência* de ATM refere-se ao mecanismo de multiplexação e comutação de uma rede. A palavra chave em ATM é assíncrono. Assincronismo saliente a diferença entre ATM e a STDM tradicional.

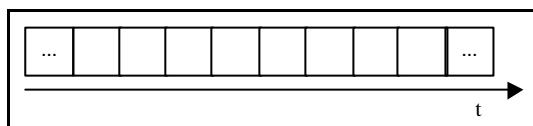
Como já apresentado na seção 9.2.3, na multiplexação STDM, uma conexão pode apenas usar o intervalo temporal de cada quadro dedicada a ela (Figura 72), e portanto a transmissão de dados ocorre sincronamente com o tempo. A multiplexação STDM é feita por reserva. Se a fonte não tem dados a transmitir durante o intervalo atribuído a sua conexão, o intervalo é perdido e não pode ser usado por outra conexão. Portanto, um intervalo de tempo pode apenas ser usado pela conexão, ou canal, que o reservou

durante o seu estabelecimento. No caso do transmissor ter mais dados a transmitir, ele deve aguardar o próximo quadro, ou deve reservar mais que um intervalo em cada quadro. Se cada intervalo corresponde a 64 Kbps (ISDN padrão), então a conexão pode apenas ter um largura de banda múltiplo de 64 Kbps. Se a conexão necessita apenas de 16 Kbps, um intervalo de tempo deve ser reservado, assim 48 Kbps são perdidos. Se uma conexão necessita de 70 Kbps, dois intervalos (128 Kbps) em cada quadro deve ser reservado e 58 Kbps são desperdiçados.

### **ATM - Divisão do tempo em pequenos intervalos**

Em ATM, o tempo em uma ligação é dividido em pequenos intervalos de tempo fixos (Figura 84). Mas ao contrário da STDM, os intervalos temporais não são reservados para uma determinada conexão. O transmissor pode utilizar qualquer intervalo disponível para transmitir dados sob uma conexão. Assim, a transmissão de dados de cada canal é assíncrono no tempo.

ATM é portanto um esquema de multiplexação por divisão de tempo em que intervalos temporais não são atribuídos a uma conexão particular. Por esta razão, ATM é também chamada de multiplexação por divisão de tempo assíncrona (ATDM). Note que estes intervalos são fixos e síncronos no tempo; assíncrono significa que estes intervalos podem ser usados independentemente por qualquer canal. A principal vantagem do ATM é largura de banda sob demanda: cada conexão pode usar tanta largura de banda quanto necessário sujeito a um limite máximo. Intervalos de tempo não utilizados podem ser usado por qualquer outro canal. Desta forma, a largura de banda da rede é compartilhada eficientemente entre aplicações.



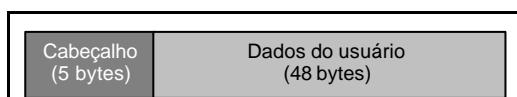
**Figura 84.** ATM: Tempo é dividido em intervalos fixos usados por qualquer canal

### **Noção de células**

Em ATM, todos os dados são enviados a rede pelas ligações de acesso em pequenos pacotes encaixados nos intervalos temporais. Estes pequenos pacotes de tamanho fixo são chamados de **células**. O equipamento do usuário é responsável por transformar o dado gerado pelas aplicações neste formato comum. Nesta abordagem, grandes blocos de dados são segmentados e transmitidos via rede como uma seqüência de células; estes blocos de dados são então remontados a partir das células que chegam no receptor.

O equipamento terminal controla sua largura de banda efetiva por meio da freqüência com que as células são geradas, usando mais ou menos os intervalos de tempo de tamanho fixo que aparecem nas ligações de acesso. A única diferença que distingue uma aplicação de baixa largura de banda de aplicações de alta largura de banda é a freqüência em que as células são geradas, e usuários podem acessar a largura de banda sob demanda a qualquer taxa de dados efetiva até o máximo da velocidade permitida pela ligação de acesso.

A ITU-T definiu que as células ATM para a B-ISDN têm o tamanho de 53 bytes (Figura 85): 5 bytes são o cabeçalho e 48 bytes são dados do usuário (*payload*). Todos os tipos de tráfego apresentados pelo usuário final são transformados em células e transmitidos para o destino apropriado baseado nas informações do cabeçalho. Assim, ATM é muito flexível - um usuário necessita de um ponto de acesso e pode transmitir e receber todos os tipos de informações através deste ponto.



**Figura 85.** Formato da célula ATM

Na rede, todas as células são manipuladas autonomamente. A rede opera apenas no cabeçalho da célula que contém informações necessárias à rede para transferir a célula para seu receptor. A rede não considera a natureza do conteúdo transportado em uma célula, que pode ser informações de voz, imagem, dados alfanuméricos ou vídeo.

### 9.11.1 Como o tamanho da célula foi determinado

Muitos fatores foram considerados para a determinação do tamanho da célula pela ITU-T (53 bytes). Dentre estes fatores incluem a sobrecarga (*overhead*) do cabeçalho em cada célula, o atraso na montagem da célula, o atraso na multiplexação e o atraso na retransmissão.

#### **Sobrecarga**

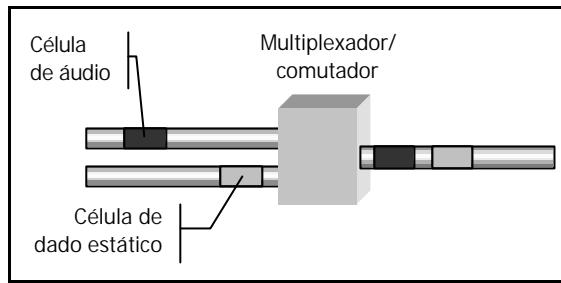
Cada célula deve ter um cabeçalho contendo informações necessárias ao transporte da célula. Este cabeçalho tem um tamanho de 5 bytes. Estes bytes são considerados uma sobrecarga aos dados do usuário afim de possibilitar a transmissão destes. Se o tamanho da célula é muito pequeno, a sobrecarga é relativamente grande comparado ao dado do usuário e largura de banda efetiva da rede baixa significativamente. Por exemplo, se o tamanho da célula fosse de 16 bytes e o cabeçalho de 5 bytes, então a sobrecarga seria de 23,81% (23,81% da largura de banda é usada para transportar outras informações que os dados do usuário). Assim, para o uso eficiente da largura de banda da rede, o tamanho da célula deveria ser grande. Na ATM, como o tamanho da célula é de 53 bytes e o cabeçalho é de 5 bytes, a sobrecarga gerada é de 9,43%.

#### **Atraso de montagem das células**

Outro fator importante na determinação do tamanho da célula é o atraso na montagem de cada célula, que refere-se ao tempo de compor uma célula. Isto é importante para transmissões em tempo-real (por exemplo, na transmissão de áudio e vídeo tempo-real). Maior o tamanho de célula, maior é tempo de montagem. Assim, para reduzir o atraso na montagem da célula, o tamanho da célula deveria ser pequeno.

#### **Atraso na multiplexação**

Uma das metas do ATM é permitir que todos os tipos de tráfego (dados alfanuméricos, mídias estáticas do tipo imagens e texto, e mídias contínuas do tipo áudio e vídeo) sejam transmitidos sobre a rede. Portanto em uma rede ATM, algumas células contêm informações sensíveis ao tempo tão bem quanto informações estáticas. O tamanho da célula afeta o atraso da célula quando múltiplos fluxos são multiplexados em uma única linha. Por exemplo, considere dois fluxos multiplexados em uma única ligação (Figura 86). Suponha que a célula transportando dado estático chega no multiplexador mais cedo que uma célula transportando áudio. Consequentemente, a célula de áudio deve esperar até que a célula de dado estático seja transmitida. Se a célula é grande, a célula de áudio deve esperar muito tempo. O atraso também depende na velocidade da ligação. Por exemplo, se o tamanho da célula fosse de 1 Kbytes e a ligação tivesse uma velocidade de 1 Mbps, o atraso máximo na multiplexação seria de 8 ms. Quando a velocidade da linha aumenta, este atraso diminui proporcionalmente. Quando a velocidade de ligação é de 10 Mbps, o atraso é reduzido para 0,8 ms (se torna irrisório). Mas nós devemos nos lembrar que a célula pode atravessar diversos comutadores no seu caminho da fonte ao destino e atrasos em todos os nós são adicionados. Portanto, um pequeno tamanho de célula deveria ser usado para reduzir o atraso na multiplexação.



**Figura 86.** Multiplexação de células de dado estático e de áudio

### Atraso na retransmissão

Atraso na retransmissão é similar ao atraso na montagem das células, mas ele ocorre nos comutadores de rede. Atraso de retransmissão é o período de tempo entre o instante em que o primeiro byte da célula chega no comutador e o instante em que o comutador retransmite a célula para a ligação de saída. O método mais comum usado nos comutadores é armazenar-e-retransmitir: o comutador armazena os bytes que chegam até ter a célula completa antes de retransmí-la à ligação de saída. Se a ligação de saída estiver ocupada, a célula sofre um atraso extra de multiplexação como discutido anteriormente. O atraso na retransmissão é o tempo de receber uma célula inteira. A maior diferença entre o atraso de montagem e atraso de retransmissão é que no atraso de montagem a velocidade de geração de dados é determinado pelo processo de aplicação (por exemplo, pela taxa de amostragem na conversão analógico para digital em aplicações de áudio e vídeo), enquanto no atraso na retransmissão a velocidade dos dados que chegam é determinado pela velocidade da ligação. A velocidade da ligação é normalmente muito mais alta que a taxa de geração dos dados na aplicação, assim cada atraso de retransmissão é muito menor que o atraso de montagem. Mas é importante notar que o atraso na montagem apenas ocorre uma vez na fonte da informação, enquanto o atraso na retransmissão ocorre a cada comutação no caminho da célula da fonte ao destino. Portanto, para reduzir atrasos na retransmissão, o tamanho da célula deve ser pequeno.

### Considerações sobre o tamanho da célula

Concluindo as considerações acima, nós temos que ter um tamanho de célula pequeno para reduzir o atraso total de transmissão da célula. Mas para aumentar o uso da largura de banda, o tamanho da célula deveria ser grande. Portanto, o tamanho da célula deve ser determinado por um compromisso entre os fatores acima descritos.

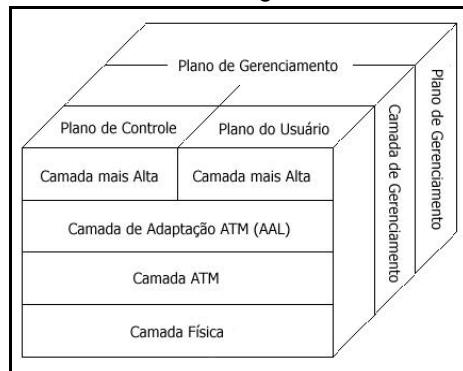
O tamanho destinado aos dados do usuário de 48 bytes do ATM é um valor que o comitê de padronização determinou para este compromisso. Este comitê dividiu-se em dois grupos no momento desta determinação: um desejando 128 bytes e outro desejando 16 bytes de dado por célula. O maior tamanho de célula foi considerado inicialmente como sendo mais apropriado para tráfego de dados estáticos. O menor tamanho de célula foi otimizado para comunicação de voz. Após algumas negociações, os dois lados revisaram suas propostas para 64 e 32 bytes. O comitê determinou a média aritmética destas duas propostas e adotou o tamanho de 48 bytes para a área destinada aos dados do usuário. Esta média pode ser considerada um compromisso entre o tamanho ideal para transmissão de dados estáticos e dinâmicos.

#### 9.11.2 Modelo de Referência B-ISDN

B-ISDN é uma arquitetura em camadas fornecendo múltiplos serviços às aplicações, tal como voz, dados, e vídeo, a serem mixados em uma rede. A Figura 87 ilustra o modelo de Referência para Protocolos (*Protocol Reference Model - PRM*) definido pela recomendação I.321 do ITU-T. O PRM constitui-se num formato tridimensional composto por três planos:

- **Plano usuário**, se refere a transmissão dos dados dos usuários.
- **Plano de Controle**, como o próprio nome diz, executa funções de controle como manter e desativar conexões.

- **Plano de Gerenciamento**, utilizado no gerenciamento dos planos e das camadas.



**Figura 87.** Modelo de Referência para Protocolos - PRM

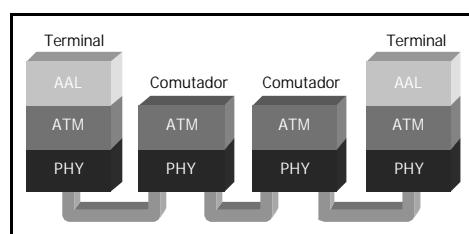
O plano de gerenciamento se divide conforme as entidades que se pretende gerenciar. O plano do usuário e o plano de controle são compostos por quatro camadas (Figura 88):

- A Camada mais Alta, que no plano usuário (camada de aplicação) são responsáveis pela produção e apresentação de vários tipos de dados, e no plano controle realiza serviços de controle.
- A camada de Adaptação ATM (AAL) assegura que os serviços tenham as características apropriadas para as aplicações e divide todos os tipos de dados em 48 bytes que formarão a célula ATM.
- A Camada ATM toma os dados a serem enviados e adiciona o cabeçalho da célula (5 bytes) que assegura que a célula seja enviada na conexão correta. Ela é uma camada de comutação e multiplexação independente da camada física.
- A Camada Física (PHY) define as características elétricas e interfaces de rede.

Camadas	Funções
Camadas mais Alta	Produção e apresentação de vários tipos de dados
Camada de Adaptação ATM (AAL)	Segmentação, remontagem, funções dependentes de aplicação (p.e. controle de erro)
Camada ATM	Controle de fluxo genérico, geração/extracção de cabeçalho de célula, translacão VPI/VCI, multiplexação/demultiplexação de célula
Camada Física	Transmissão de bits sobre o meio de transmissão

**Figura 88.** Camadas e funções da arquitetura B-ISDN para o plano usuário

AALs existem apenas na fonte e no destino. Diferentes AALs podem ser usadas para diferentes tipos de tráfego. Desde que uma célula está na rede, os 48 bytes de dados do usuário não são tocados. Os comutadores de rede rotearão a célula de acordo com as informações do cabeçalho. O cabeçalho pode ser alterado de *hop* (caminho de transmissão direto) a *hop* como será discutido mais adiante. A arquitetura conceptual ATM é mostrado na Figura 89.



**Figura 89.** Arquitetura conceptual do ATM

### 9.11.3 Camada Física

A camada física define o mapeamento das células ATM no meio físico e os parâmetros da transmissão física. Ela é responsável pela transmissão dos bits sobre o meio de transmissão. É esta camada que determina a taxa de transmissão de bits. Taxas de bits de 155 Mbps e 622 Mbps foram inicialmente propostas para transmissão em fibra ótica,

mas outras taxas de transmissão também foram definidas para outros meios de transmissão. Hoje, ATM não fixa o tipo específico de transporte físico. Mas o meio mais comum considerado para longas distâncias é fibra ótica usando o protocolo SONET (*Synchronous Optical Network*). O padrão SONET foi desenvolvido pela Bell Communications Research e é agora um padrão ANSI (American National Standards Institute).

SONET especifica como sinais digitais síncronos podem ser transportados por redes de fibra ótica. Ela apresenta como unidade básica de transporte um formato de quadro composto de 810 bytes que se repete a cada 125µs, tendo assim uma taxa de bits de 51,84 Mbps. A taxa 51,84 Mbps é chamada de STS-1 (*Synchronous Transport Signal level 1*) ou OC-1 (Optical Carrier Level 1). Uma taxa OC-n é meramente o equivalente ótico de um sinal elétrico STS-n.

A ITU-T desenvolveu um padrão similar a SONET, chamado de SDH (*Synchronous Digital Hierarchy*), a partir de três sinais STS-1 concatenados, denominando o novo sinal de STM-1 (*Synchronous Transport Module level 1* - 155,52 Mbps) e adotou este como o sinal básico para a interface NNI e UNI. O padrão SONET foi posteriormente também aprovado como sinal básico.

A hierarquia de transmissão ótica SDH é baseada em uma taxa básica de 155,52 Mbps, chamada STM-1 (*Synchronous Transport Model Level 1*). Interfaces de alta velocidade podem ser criadas simplesmente por multiplexação (usando byte *interleaving*) de canais de mais baixa velocidade. Isto simplifica o processo de criação da faixa de velocidades de transmissão oferecidas por um serviço particular. Algumas velocidades de interfaces comuns atualmente reconhecidas por SONET/SDH são dados na tabela abaixo.

SONET Optical carrier (OC) level	SDH STM level	Line rate (Mbps)
OC-1		51,84
OC-3	STM-1	155,52
OC-12	STM-4	622,08
OC-24	STM-8	1244,16
OC-48	STM-16	2488,32

A ligação máxima de 2,4 Gbits/s da tabela acima não é a mais alta taxa obtida por SONET/SDH, ela simplesmente representa a maior taxa possível com os produtos atuais. A hierarquia de velocidade de transmissão SONET/SDH estende-se indefinitamente, limitada apenas pelos avanços na tecnologia para suportar transmissão de mais alta velocidade. O padrão ITU-T SDH concentra-se nas velocidades de transmissão de 155 Mbps e acima.

O ATM Forum definiu primeiramente o uso do padrão de transmissão física FDDI a 100 Mbps para redes locais ATM. Outras opções incluem cabos de par trançado (UTP-3) a 25.6 ou 155 Mbps, ou linhas síncronas tal como T3 (45 Mbps) ou E3 (34 Mbps). A tabela abaixo apresenta a lista de camadas físicas suportadas atualmente [Kuo, 98]. Esta lista ainda está crescendo.

Em termos de vazão, mesmo 25 Mbps de largura de banda dedicada excede suficientemente aos 1,4 Mbps requerido para fluxos de áudio e vídeo. Atualmente, nós vemos as primeiras interfaces ATM a 622 Mbps, e mesmo 2.4 Gbps não estão longe de surgirem.

Taxa	Nome e sincronização de quadros	Definido por
622 Mbps	SDH, STM4/SONET STS-12 sob SMF	ITU
155 Mbps	SDH STM-1/SONET STS-3C sob SMF, MMF, STP, UTP-5	ITU/ATM Forum
155 Mbps	Baseado em célula sob MMF, STP, UTP-5	ATM Forum
100 Mbps	Baseado em célula sob MMF (TAXI)	ATM Forum
51 Mbps	UTP-3, MMF, SMF	ATM Forum
45 Mbps	G.804/T3	ATM Forum/ANSI
34 Mbps	G.804/E3	ATM Forum/ANSI

25.6 Mbps	STP, UTP-3, UTP-5 <sup>15</sup>	ATM Forum
2 Mbps	E1	ATM Forum/ETSI
15 Mbps	T1	ATM Forum/ANSI
SDH: Synchronous Digital Hierarchy (ITU)		
STM: Synchronous Transfer Module (ITU)		
SONET: Synchronous Optical Network (ANSI)		
SMF: Single-Mode Fiber <sup>16</sup>		
MMF: Multi-mode Fiber <sup>17</sup>		
TAXI: Transmitter/receiver interface- physical FDDI interface		
STP: Shielded Twisted Pair		
UTP: Unshielded Twisted Pair		

### 9.11.4 Elementos básicos ATM

#### ATM uma rede hierárquica

Uma rede ATM é hierárquica (Figura 90). Os terminais (sistemas finais) são conectados a comutadores diretamente através de pontos de acesso. Um comutador é constituído por várias portas que se associam às linhas físicas da rede. Um comutador deve receber as células que chegam pelas portas de entrada e retransmiti-las através das portas de saída, mantendo a ordem original das células em cada conexão. É possível que terminais sejam ligados a uma LAN, que é conectada a uma rede ATM através de um ponto de acesso. A velocidade da ligação entre um ponto de acesso e um comutador é chamado de *velocidade de acesso* e é dedicada ao ponto de acesso. Comutadores agregam células vindas pelas ligações de entrada e retransmitem elas para outros comutadores ou terminais. A largura de banda entre comutadores é chamado de *largura de banda agregada*. Ela é normalmente maior que a velocidade de acesso, mas ela não precisa ser a soma das taxas de pico de todas as suas ligações de entrada pois ATM usa multiplexação estatística. A interface entre o usuário (terminal) e a rede é chamada *interface usuário-rede* (UNI - User-Network Interface). A interface entre os comutadores de rede é chamada *interface rede-rede* (NNI - Network-Network Interface). O formato da célula é ligeiramente diferente na UNI e na NNI (discutido na próxima seção).

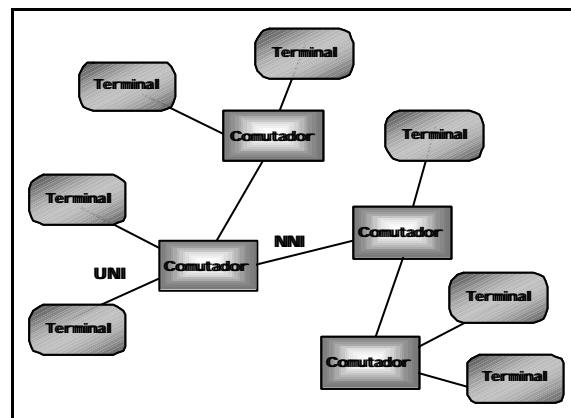


Figura 90. Configuração hierárquica da rede ATM

#### ATM é orientada a conexão

ATM é um serviço orientado a conexão usando conexões virtuais comutadas e permanentes. Em ATM, um *Canal Virtual* (*Virtual Channel* - VC) é uma conexão lógica entre dois comutadores ATM ou entre um terminal ATM e um comutador ATM. A rota que as células seguem consiste de uma seqüência de VCs entre terminais ATM. A

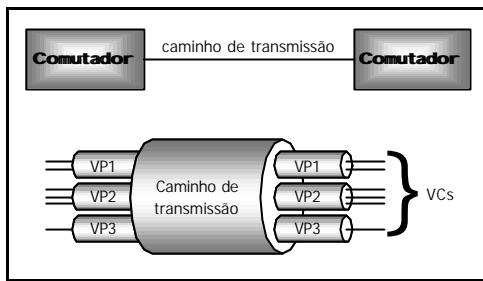
<sup>15</sup> Padrão ISO e TIA (Telecommunications Industries Association) com quatro pares encapados em um plástico termal isolador, trançado um no outro e encapsulado em um polímero retardador de chama. Ele apresenta uma freqüência de operação máxima de 100 MHz.

<sup>16</sup> Cabo constituído de uma única fibra, possui uma largura de banda maior que o modo múltiplo, mas o seu fino condutor torna muito difícil o encaixe sem ferramentas especiais e conhecimento técnico. Além disso, este modo requer um laser (em vez do LED) como fonte de sinal, o que é de maior custo.

<sup>17</sup> Cabo constituído de várias fibras, possui uma largura de banda menor que o modo mono, mas é muito mais fácil de encaixar.

seqüência de VCs que forma uma rota associada com uma chamada virtual é conhecida como uma conexão de canal virtual (*Virtual Channel Connection - VCC*).

Cada canal virtual pertence a um caminho virtual (*Virtual Path - VP*) entre comutadores. Um caminho virtual é um agrupamento (em uma ligação física) de vários VCs em uma única entidade lógica como ilustrado na Figura 91. Enquanto os caminhos virtuais são persistentes, os canais virtuais são estabelecidos quando a conexão é iniciada e associada a um caminho virtual. O cabeçalho das células ATM contém informações necessárias para comutar a célula de um porta de entrada de um comutador ATM à porta de destino de acordo com as informações na tabela de roteamento do comutador (visto mais adiante).



**Figura 91.** Canal Virtual (VC) e Caminho Virtual (VP)

#### 9.11.5 Formato das Células ATM

Como já mencionado, as células ATM tem o tamanho de 53 bytes. Os primeiros 5 bytes formam o cabeçalho, que é usado pelos comutadores para rotear a célula. Os demais 48 bytes contém dados do usuário, que é formatado em um dos formatos da camada AAL.

O cabeçalho da célula ATM tem dois formatos:

- um para células transmitidas entre terminais e comutadores (na interface UNI);
- outro para células transmitidas entre comutadores (na interface NNI).

O cabeçalho altera quando uma célula é movida da UNI para a NNI.

#### Cabeçalho de célula ATM na NNI

O cabeçalho da célula NNI tem cinco campos (Figura 92):

- *Virtual Path Identifier (VPI)* e *Virtual Channel Identifier (VCI)*: VPI ocupa 12 bits e VCI ocupa 16 bits. Uma conexão ATM é unicamente identificada por um endereço de 28 bits formado pela combinação de VPI e VCI.
- *Payload type field*: este campo de 3 bits indica se os dados contidos na célula são dados do usuário ou operação de rede, administração, e dado de gerência (OAM). Quando o primeiro bit do campo é 1, então a célula é uma célula OAM, senão ela contém dados do usuário.
- *Cell Loss Priority (CLP)*: este campo de 1 bit indica a prioridade da célula. Quando ele é setada (1), a célula é de baixa prioridade. Caso o comutador tenha que descartar algumas células devido ao preenchimento completo do buffer, células de baixa prioridade serão descartadas. Este campo é utilizado para dar prioridade a certos tipos de informação ou evitar usuários abusivos. Nesta última aplicação deste campo, caso o terminal transmitir a uma taxa de bits maior que a taxa estabelecida durante a conexão, a rede marcará as células extras como não prioritárias. Caso haja congestionamento na rede, estas células serão descartadas primeiro, protegendo a qualidade de serviço das outras conexões.
- *Header CRC*: este campo é usado para permitir a detecção de erro e correção para o cabeçalho. Como o cabeçalho é alterado em cada hop, o CRC deve ser checado e recalculado a cada hop na rede ATM. Como é esperado que ATM execute sob um meio de transmissão confiável, a ocorrência de erros devem ser raras. Devido a isto, nenhuma detecção de erro nos dados é provida pela camada ATM.

Bits:	8	7	6	5	4	3	2	1
Byte 1								Virtual path identifier (VPI)
Byte 2								
Byte 3								Virtual channel identifier (VCI)
Byte 4						Payload type		CLP
Byte 5						Cyclic redundancy check (CRC)		

**Figura 92.** Cabeçalho da célula no NNI**Cabeçalho de célula ATM na UNI**

O cabeçalho de célula ATM na UNI é ligeiramente diferente do cabeçalho na NMI. Os campos que diferem são (Figura 93) a adição do campo GFC e a redução do campo VPI para 8 bits. O campo GFC é adicionado para reconhecer que o terminal ATM pode ser conectado a redes de acesso compartilhado tal como DQDB, que é conectada a redes ATM. Para este tipo de terminal, o método de multiplexação deve ser negociado entre eles. GFC é usado para indicar as prioridades de células destes terminais para acessar a rede ATM. Ele não é inteiramente definido e este campo é normalmente setado a 0.

Bits:	8	7	6	5	4	3	2	1
Byte 1								Generic flow control (GFC)
Byte 2								Virtual path identifier (VPI)
Byte 3								Virtual channel identifier (VCI)
Byte 4						Payload type		CLP
Byte 5						Cyclic redundancy check (CRC)		

**Figura 93.** Cabeçalho da célula no UNI**9.11.6 Procedimento de estabelecimento de chamada**

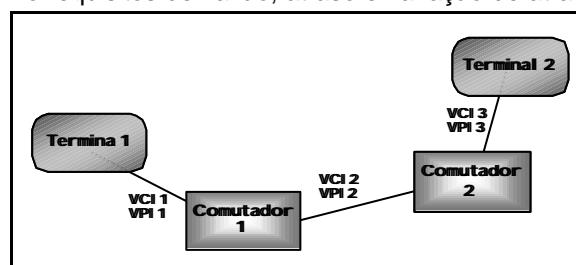
Como já apresentado, ATM fornece serviços orientados a conexão. Antes de uma célula ser transmitida entre transmissor e receptor, uma conexão deve ser estabelecida. As duas principais funções executadas durante o estabelecimento da conexão são:

- Executar um teste de admissão e negociação da Qualidade de Serviço (QoS) entre os terminais e a rede;
- Atribuir um VPI e um VCI para a conexão se os parâmetros de QoS e o pedido de conexão são aceitos.

Note que estas funções são as ideais. Algumas funções, tal como especificação de QoS e controle de admissão, não têm sido implementadas em muitas redes ATM.

Como VPIS e VCIs são alocados dinamicamente durante o estabelecimento da conexão, esta operação é complexa. O processo de estabelecimento de conexão usa o conjunto de procedimentos do padrão de sinalização Q.2931 (antigo Q.93B). Estes procedimentos não são inteiramente especificados. Neste curso, nós usamos um exemplo para apresentar as bases dos procedimentos para estabelecer uma chamada entre um terminal 1 e 2 (ilustrado na Figura 94). Estes passos são os seguintes:

- Terminal 1 envia um pedido de chamada ao comutador 1. O pedido de chamada contém o nome ou endereço do destino (terminal 2) e um conjunto de parâmetros de QoS tal como requisitos de vazão, atraso e variação de atraso.

**Figura 94.** Exemplo de cenário de comunicação

- Comutador 1 determina que linha de saída deve ser usada baseada na informação do destinatário. Ele então desempenha um teste de admissão para ver se o receptor fornece a QoS requerida pela chamada baseado nos recursos

disponíveis (velocidade de comutação, tamanho de buffer e largura de banda da linha de saída).

- Se o comutador não pode suportar a QoS requerida, ele deve enviar uma mensagem de rejeição de conexão ao terminal 1. Opcionalmente, o comutador pode sugerir um novo conjunto de QoS ao terminal. Se o terminal 1 aceita a nova QoS, a conexão é estabelecida. Senão a conexão é rejeitada.
- Se o comutador 1 pode suportar a QoS, ele escolherá o próximo valor VCI disponível como o VCI para a conexão referente ao segmento entre terminal 1 e comutador 1. Cada comutador tem um número de VCIs e VPIs a sua disposição. O número é determinado pelo número de bits atribuído a VCI e VPI no cabeçalho da célula. Por exemplo, se 8 bits são atribuídos ao campo VCI, VCI pode ser um número entre 0 a 255. Um comutador aloca números VCI para conexões seqüencialmente a partir desta faixa. Quando uma conexão é fechada, o VCI (e possivelmente o VPI) é liberado para ser usado por novas conexões. Considere neste exemplo, que o VCI alocado para a conexão deste segmento é VCI 1. Se não há conexões existentes entre terminal 1 e comutador 1, um VPI é alocado para a conexão para este segmento da mesma forma que o VCI é alocado. Se existem conexões, o VPI das conexões existentes é usado para esta conexão também. Neste exemplo, consideramos que o VPI usado para a conexão é VPI 1. VPI 1 e VCI 1 são usados para a ligação entre terminal 1 e comutador 1.
- O comutador 1 envia o pedido de conexão ao comutador 2. Comutador 2 checa se o terminal 2 é ativo (ligado). Se ele não está, a chamada é rejeitada. Se o terminal 2 estiver ativo, o comutador 2 executa um teste de admissão para ver se a QoS requerida pode ser suportada. Se não, a chamada é rejeitada, ou uma negociação será feita entre o terminal 1, comutador 1 e 2.
- Se a QoS é aceitável, o comutador 2 alocará um VPI para a conexão para o segmento entre comutador 1 e comutador 2 se não há conexão existente entre eles. Se existirem conexões entre eles, o VPI das conexões existentes é usada para esta conexão também. Neste exemplo, chamaremos a VPI para este segmento de VPI 2. Comutador 2 também aloca um VCI para a conexão. Nós chamaremos neste exemplo de VCI 2.
- O Comutador 2 envia o pedido ao terminal 2, se o terminal não aceita a chamada, o pedido é rejeitado ou os parâmetros de QoS são negociados. Por outro lado, a conexão é estabelecida com sucesso.
- O Comutador 2 aloca o VPI e VCI para o segmento entre comutador 2 e terminal 2. Chamaremos estes de VPI 3 e VCI 3. Relembrando que um novo VPI é usado apenas se não há conexões entre as duas entidades de comunicação.
- O Comutador 2 notifica o comutador 1 com uma confirmação da conexão contendo VPI 2 e VCI 2 a serem usados pelo comutador 1 para enviar células ao comutador 2. Comutador 1 notifica o terminal 1 a confirmação de conexão contendo VPI 1 e VCI 1 a serem usados pelo terminal 1 para enviar células ao comutador 1 para esta conexão.
- Recebendo a confirmação de conexão, o terminal 1 pode enviar células sobre a conexão com VPI 1 e VCI 1 no cabeçalho.

O apresentado acima é o procedimento de estabelecimento de conexão em uma direção. Os VPIs e VCIs atribuídos são usados para transmissão do terminal 1 para o terminal 2. Se mais que uma conexão é necessária entre terminal 1 e 2, como normalmente é o caso, eles diferem apenas nos VCIs e o VPI é compartilhado por todas as conexões. Um terminal pode ter múltiplas conexões para outros terminais. Se o terminal 2 desejar enviar dados ao terminal 1, um procedimento de estabelecimento de conexão similar é executado. Note que Q.93B suporta o estabelecimento de conexões unidirecional e bidirecionais.

Diferentes tipos de conexões ATM podem ser estabelecidas:

- Conexões ponto a ponto: as quais conectam dois terminais ATM. Tais conexões podem ser unidireccionais ou bidirecionais.
- Conexões ponto a multiponto: as quais conectam uma simples fonte (conhecida como o nó raiz) a múltiplos destinatários (conhecidos como folhas). As células são

replicadas dentro da rede pelos comutadores ATM. Tais conexões são unidirecionais, permitindo a fonte transmitir para vários destinatários, mas os folhas não podem transmitir para o raiz, nem para outros folhas, na mesma conexão.

### 9.11.7 Roteamento de células

O roteamento de células é relativamente simples após o estabelecimento da conexão. Em cada comutador, uma tabela de roteamento é configurada durante o processo de estabelecimento de conexão. Usando o exemplo acima, nós temos entradas na tabela no comutador 1 e 2 como mostrado na Figura 95.

Tabela de roteamento comutador 1		Tabela de roteamento comutador 2	
Entrada	Saída	Entrada	Saída
ligação h: VPI 1, VCI 1	ligação i: VPI 2, VCI 2	ligação j: VPI 2, VCI 2	ligação k: VPI 3, VCI 3
...	...	...	...

Figura 95. Tabela de roteamento nos comutadores 1 e 2

Cada comutador pode ter várias ligações de entrada e de saída. Ligações entre terminais e comutadores são normalmente fixadas, não há alternativas. Mas para ligações entre comutadores, diferentes rotas podem ser usadas, assim as ligações físicas usadas para a conexão virtual estabelecida em diferentes instantes podem ser diferentes. Numa conexão estabelecida, as ligações físicas usadas para a conexão são fixas. Nas tabelas acima, nós chamamos as ligações usadas para a conexão do terminal 1 para o terminal 2 de ligações h, i, j e k.

Quando o terminal 1 desejar enviar uma célula ao terminal 2 sobre a conexão estabelecida, VPI 1 e VCI 1 são usados no cabeçalho da célula. A célula chega ao comutador 1 na ligação h. O comutador 1 verifica a tabela de roteamento, que indica que a célula deve ser enviada sobre a ligação i com VPI 2 e VCI 2 no cabeçalho. Assim o comutador 1 muda o cabeçalho da célula e a envia na ligação i. Quando o comutador 2 recebe a célula na ligação j com VPI 2 e VCI 2, ele olha em sua tabela de roteamento. Ele então envia a célula sob a ligação k com VPI 3 e VCI 3 no cabeçalho da célula. Quando o terminal 2 recebe a célula, ele sabe que conexão a célula pertence baseado na combinação de VPI 3 e VCI 3.

O par VPI e VCI é único para cada hop-by-hop. Para cada segmento de conexão, os dois comutadores em cada ponta entram em acordo com o par VPI e VCI para servir como o identificador do hop para cada conexão. Quando uma célula passa através de um comutador, o comutador altera o par VPI e VCI do par usado para identificar o canal no último hop para o par usado para identificar o canal no próximo hop. Não há confusão neste esquema. Cada par de comutadores conhece que pares VPI e VCI estão sendo usados em uma dada ligação e não dá um novo canal o mesmo VPI e VCI de um canal existente.

VPI e VCI são pequenos e em localizações fixas em cada célula. Assim eles podem ser usados eficientemente como índice para tabelas de roteamento, permitindo o desenvolvimento de comutadores ATM rápidos.

A maneira na qual as tabelas de roteamento são estabelecidas determinam os dois tipos fundamentais de conexão ATM:

- Conexões Virtuais Permanentes (*Permanent Virtual Connections - PVC*): Uma PVC é uma conexão estabelecida por algum mecanismo externo, tipicamente gerenciadores de rede, na qual um conjunto de comutadores entre uma fonte ATM e um destino ATM são programados com os rótulos VPI/VCI apropriados.
- Conexões Virtuais Comutadas (*Switched Virtual Connections - SVC*): uma SVC é uma conexão que é estabelecida automaticamente (ou dinamicamente) através de um protocolo de sinalização. SVCs não requerem uma interação manual, necessárias para as PVCs, fazendo com que as SVCs sejam muito mais largamente utilizadas.

### 9.11.8 Camadas de Adaptação ATM

ATM envia células sob uma conexão. O transmissor deve segmentar fluxos de dados em células, e no receptor os dados nas células são remontados em fluxos de dados novamente. A camada AAL representa uma interface entre os protocolos de alto nível e a camada ATM. Quando ela recebe dados das camadas superiores, ela realiza a segmentação de dados em células ATM que são passadas para a camada ATM. No sentido inverso (quando recebe células ATM da camada ATM) a camada AAL deve remontar em fluxos de dados que as camadas superiores podem entender.

O termo camada de adaptação é devido a sua função que é a de adaptar o dado em uma forma desejável para ATM. O propósito da AAL é empacotar eficientemente as várias espécies de dados de alto nível, tal como arquivos, amostragens de áudio, quadros de vídeo, em uma série de células que podem ser enviadas sob conexões ATM e reconstruídas em um formato apropriado no receptor.

AALs são necessárias de maneira a que aplicações do mesmo tipo possam comunicar com outras. O primeiro passo realizado pelo ITU-T quanto a camada de adaptação, foi a definição das classes de serviço oferecidas pela AAL através da recomendação I.362. Logo após, o ITU-T definiu os tipos de AAL que suportam estas classes, através da recomendação I.363. A ITU-T determinou os seguintes classes de aplicações:

- **Classe A:** Aplicações sensíveis ao tempo com taxa de bits constantes (Constant Bit Rate - CBR). Estas aplicações enviam e recebem dados a taxas constantes. Elas também necessitam que o atraso da fonte ao destino seja limitado. Este tipo de aplicação inclui comunicação de áudio e vídeo codificados a taxa de bits constantes.
- **Classe B:** Aplicações sensíveis ao tempo com taxa de bits variáveis (Variable Bit Rate - VBR). Estas aplicações enviam dados a taxa de bits variáveis, mas requerem atrasos limitados. Exemplos incluem comunicação de áudio e vídeo codificados a taxa de bits variáveis.
- **Classe C:** Aplicações de dados orientadas a conexão. Esta classe suporta aplicações que eram historicamente usadas em serviços de rede tal como X.25.
- **Classe D:** Aplicações de dados sem conexão. Aplicações nesta classe enviam e recebem dados usando datagramas sem conexão.

As principais características destas classes são resumidas na Figura 96.

	Classe A	Classe B	Classe C	Classe D
Tempo na fonte e no destino	Relacionado		Sem relação	
Taxa de geração de bits	Constante (CBR)	Variável (VBR)		
Modo de conexão	Orientada à conexão		Sem conexão	

**Figura 96.** Classes de Serviço definidas na recomendação I.362 do ITU-T

A fim de suportar estas diferentes classes de serviços, a ITU-T propôs três diferentes camadas de aplicação: AAL1, AAL2 e AAL3/4. O ATM Forum propôs uma nova camada AAL5 e investiga uma nova camada, AAL6.

- AAL1 é para o serviço classe A e emprega funções de empacotamento/desempacotamento para converter fluxos a taxa constante de bits em células no emissor e montar as células em fluxos a taxa constante de bits no receptor. Um fluxo de bit inteiramente sincronizado é transferido ao receptor, requerendo pequeno atraso e controle de variação de atraso na rede.
- AAL2 fornece serviço classe B. Devido a taxa de bits variável, é difícil reservar recursos para esta espécie de tráfego e assim é difícil implementar a AAL2. Ou o pico da largura de banda é reservada - não muito eficiente - ou garantias de atraso e perdas não podem ser dadas. Atualmente, a ITU está definindo uma primeira versão da AAL2 com um foco ligeiramente diferente. Ela é voltada particularmente a agregação de vários fluxos de voz compactados a taxa variável em um circuito virtual ATM (*trunking*), que é um serviço classe B importante para redes de longa

distância. De qualquer maneira, produtos não estão sendo esperados em um futuro próximo.

- AAL3/4 implementa serviços classes C e D. Como ATM é inherentemente orientada a conexão, serviços sem conexão necessitam ser providos por servidores sem conexão (CLS - Connection-Less Servers), que são acessados por comunicação orientada conexão. Assim, não é necessário distinguir serviços classes C e D a este nível. Este fato foi apenas descoberto durante o processo de padronização, e assim os tipos 3 e 4 anteriores são agora juntados em uma AAL 3/4 comum. A principal função da AAL3/4 é segmentação e remontagem de mensagens em células, e vice-versa. Além disso, AAL 3/4 fornece um campo identificador de mensagem (MID), permitindo um transporte intercalado de diferentes mensagens sob o mesmo VC. Esta característica é útil no contexto de serviços multicast ou sem conexão. Isto permite ao servidor retransmitir células individuais de uma mensagem sem remontagem no servidor.
- AAL5 também fornece serviços classes C e D (como proposto por construtores de computadores no Forum ATM). Ela é mais simples e mais eficiente que AAL3/4. Ela não permite o intercalamento (*interleaving*) de mensagens, assim não havendo necessidade do campo MID na célula. Consequentemente, AAL5 fornece uma melhor utilização da largura de banda disponível. Cada célula de 53 bytes do AAL3/4 tem 5 bytes de cabeçalho e 4 bytes adicionais de sobrecarga de segmentação. Na AAL5, a sobrecarga de segmentação de 4 bytes por célula pode ser reduzida a 6 bytes por mensagem, e não por célula (campo de tamanho de 2 bytes, mais CRC de 4 bytes). A desvantagem é que uma célula corrompida sempre provocará o descarte da mensagem na AAL5 e a AAL4/3 fornece meios para localizar erros de bits em células individuais. Esta característica é interessante no contexto da comunicação multimídia, desde que alguns fluxos multimídia podem ser capazes de usar mensagens parcialmente corretas.
- AAL6 está em estágio de discussão. O ATM Forum está investigando uma AAL adaptada ao empacotamento de fluxos de dados multimídia, em particular para vídeo MPEG e MPEG-II. Discussões incluem o uso de técnicas FEC (*forward error-detection*) para aumentar a confiabilidade na comunicação ao nível onde nenhum recobrimento de erro extra é necessário e para suportar requisitos de sincronização MPEG.

Note que AALs fazem mais que segmentação e remontagem de dados, elas efetivamente implementam protocolos de transporte adaptadas aos diferentes tipos de aplicação.

Atualmente, a tendência vai cada vez mais para AAL1 para interligação com ISDN e outros serviços síncronos, e AAL5 para quase todo o resto. Na perspectiva multimídia, a vantagem do AAL3/4 é limitada. O mecanismo de manipulação de erro do AAL3/4 é útil, mas a capacidade de multiplexação tem pouco valor, assumindo a disponibilidade de VCIs a um número suficiente em primeiro lugar. Assim, eficiência e disponibilidade são provavelmente mais importantes que funções incrementadas. Embora inicialmente não antecipada pela ITU, há uma demanda do uso de AAL3/4 e AAL5 em serviços atraso limitado a fim de suportar tráfego multimídia interativo em pacotes. Nas primeiras redes ATM pilotos, que tendem a ser menor e menos carregadas, não há problemas pois não há contenção (disputa). No futuro, um gerenciamento de tráfego avançado é necessário para garantir a vazão e os atrasos requeridos.

### 9.11.9 Gerenciamento de tráfego

O modelo seguido pelo gerenciamento de tráfego ATM é um contrato negociado pelo usuário e a rede. Usuário e a função de controle de admissão da rede negociam uma descrição de tráfego e de qualidade de serviço.

O descritor de tráfego ATM é baseado nas seguintes características do fluxo:

- PCR: Taxa de Pico de célula (células/s)
- SCR: Taxa de Célula Sustentável (células/s)
- MBS: Máximo tamanho da rajada (células).

- BT: Tolerância de rajada =  $(MBS-1)/(1/SCR-1/PCR)$
- MCR: Taxa Mínima de Célula (apenas para tráfego ABR)

É responsabilidade do usuário tornar seguro que ele não exceda as características de tráfego acordados. O UPC (*Usage Parameter Control*) ou função de policiamento da rede controla as células que chegam e pode descartar qualquer célula não conformante.

Esta descrição de tráfego baseado nas taxas de células pode ser útil para permitir que o controle de admissão faça os cálculos necessários para ganhar vantagem da multiplexação estatística pela combinação de vários fluxos sem exceder a capacidade da linha ou do comutador, tão bem quanto garantir os atrasos de bufferização baixos. Isto é, de qualquer maneira, muito difícil para um usuário ou uma aplicação precisamente fornecer estes parâmetros desde que não há um relacionamento direto entre a taxa de quadros do vídeo ou resolução de vídeo com a taxa de células característica de um fluxo de vídeo compactado.

Para as características de tráfego especificados, a rede e o usuário entram em acordo com uma certa qualidade de serviço, que a rede fornece para um tráfego de entrada aceitável. Os parâmetros de QoS usados são:

- CLR: Taxa de Perdas de Célula (número de células perdidas/ números de células transmitidas).
- CTD: Atraso de Trânsito de Célula
- CDV: Variação de Atraso de Célula

Dependendo como os parâmetros do descritor de tráfego são especificados, o tráfego ATM pode ser classificado em cinco categorias de serviços diferentes ou classes de tráfego, listado abaixo e resumidas na tabela que segue:

- CBR: tráfego a taxa de bits contínua com atraso fixo e taxa de células para serviços síncronos tal como emulação de ISDN ou canais de áudio/vídeo síncronos.
- Real-time VBR: tráfego a taxa de bits variável para fluxos de dados de áudio e vídeo compactados com características tempo-real (CDV especificado).
- Non-real-time VBR: tráfego a taxa de bits variável para fluxos de dados de áudio e vídeo compactados com características tempo-real (CDV não especificado).
- ABR (available bit rate): tráfego a taxa de bits disponível para comunicação de dados com perda de células desprezível.
- UBR (unspecified bit rate): taxa de bits não especificada para tráfego sem características conhecidas.

	CBR	rt-/nrt-VBR	ABR	UBR
CLR	especificado		especificado	não especificado
CTD	especificado		não especificado	
CDV	especificado	opcional	sem efeito	
PCR		especificado		
SCR/BT	sem efeito	especificado	sem efeito	
MCR	sem efeito		especificado	sem efeito

De um ponto de vista multimídia, principalmente CBR e VBR são de interesse, CBR sendo desejável para qualquer fluxo síncrono a taxa de bits constante, e VBR para fluxos tempo-real compactados a taxa variável. Mas, a provisão de um bom gerenciamento de tráfego rt-VBR com garantias de desempenho aceitáveis é uma empreitada cara em termos de recursos de rede.

### 9.11.10 ATM satisfaz os requisitos de comunicação multimídia

ATM satisfaz quase todos os requisitos e é potencialmente satisfatória para comunicações multimídia. A lista abaixo discute como ATM satisfaz alguns dos requisitos de comunicação multimídia:

- **Largura de Banda:** ATM fornece um acesso de alta velocidade para cada sistema final. Duas velocidades de acesso de 155 e 622 Mbps tem sido padronizadas, embora velocidades mais altas e mais baixas são possíveis. Estas velocidades de acesso são suficiente para qualquer tipo de aplicação multimídia com a tecnologia de compressão atual.
- **Flexibilidade e garantias de QoS:** redes ATM podem suportar aplicações com diferentes características de tráfego e requisitos de comunicação, que inclui largura de banda, restrição de atraso e sensibilidade a erro. ATM permite a alocação dinâmica de largura de banda para cada aplicação, enquanto o serviço orientado a conexão fornecido pela ATM pode reservar recursos para garantir a QoS de cada aplicação.
- **Integração:** o protocolo ATM pode simultaneamente suportar múltiplas aplicações com diferentes características e requisitos pela mesma UNI. Isto é crucial para suportar aplicações multimídia, pois múltiplos circuitos virtuais podem ser requeridos concorrentemente.
- **Arquitetura:** a topologia física em mecha ponto-a-ponto empregada pela B-ISDN fornece uma largura de banda dedicada muito maior para cada usuário comparado às redes usando meio compartilhado tal como FDDI e DQDB. Além disso, isto permite que a mesma arquitetura pode ser empregada tanto como solução de rede pública e privada, as duas executando essencialmente o mesmo protocolo.
- **Eficiência:** Desde que ATM emprega a multiplexação estatística para compartilhamento da largura de banda de transmissão, o ganho estatístico é significante quando aplicações com alto pico para a taxa de largura de banda média (i.e., tráfego em rajada) são multiplexadas. Desde que várias aplicações multimídia e colaborativas tendem a ser muito em rajadas, o ganho estatístico resultante implica em um compartilhamento muito eficiente do custo de transmissão.
- **Multicasting:** ATM permite a implementação de multicasting pela alteração de VCI e VPI em comutadores.
- **Versatibilidade:** ATM usa um mecanismo uniforme de multiplexação e comutação independente da taxa de bits, tamanho da rede e meio de transmissão. ATM é um padrão para serviços de rede locais, campus/backbone, pública e privada. Esta uniformidade simplifica o gerenciamento da rede usando a mesma tecnologia para todos os níveis de rede.

Embora a arquitetura de protocolo ATM fornece os ingredientes necessários para suportar os requisitos das aplicações multimídia, maior parte desta arquitetura ainda não foi implementada. A tecnologia ATM está sendo estudada, testada e padronizada pela ITU-TS e pelo ATM Forum. Sua completa potencialidade ainda não foi implementada. Atualmente, ATM tem sido principalmente usada em LANs e WANs para comunicação de dados convencionais. Embora áudio e vídeo tenham sido transmitidos sobre estas redes, a largura de banda e conexões são alocadas estaticamente. As principais razões para isto são:

- A sinalização ATM ainda não foi concluída e implementada.
- Muitas questões acerca das garantias de desempenho com multiplexação estática, tal como especificação QoS, controle de admissão e gerenciamento de recursos, não estão ainda resolvidas.

A primeira onda de produtos ATM de redes locais focalizam a comunicação de dados do tipo “melhor esforço”, usando LAN emulation ou RFC 1577 (Classical IP and ARP over ATM) [Laubach, 94] [Delgrossi, 93], que atualmente esconde o controle de QoS do usuário. Embora várias questões de gerenciamento de tráfego permaneçam não resolvidas, redes ATM fornecem ampla largura de banda para fluxos multimídia. Atraso e variação de atraso são suficientemente pequenos sob carga normal. Controle de atraso e confiabilidade terá de ser analisado sob duras condições de carga, e um bom gerenciamento de tráfego está ainda para ser desenvolvido. Além disso, ATM fornece comunicação multicast, de modo que esta tecnologia satisfaz os requisitos multimídia listados acima.

## 9.12 Outras tecnologias

Além das tecnologias de rede apresentadas acima, existem outras classes de tecnologia de rede especificamente para residências. A característica básica desta classe de tecnologia é o uso das linhas telefônicas existentes e cabos de TV para fornecer serviços digitais, tal como vídeo sob-demanda, usando técnicas de modulação avançadas.

Companhias telefônicas, companhias de TV a cabo e novos participantes estão todos convencidos de que quem chegar lá primeiro será o grande vencedor; portanto, estamos vendo uma grande proliferação de tecnologias sendo instaladas [Tanembau, 97]. Um exemplo desta tecnologia é o ADSL (*Asymmetric Digital Subscriber Line*). ADSL faz uso dos pares de fios de cobre trançados das companhias telefônicas e se beneficia dos avanços no processamento do sistema digital para eliminar ecos e ruídos de linha eletronicamente. Cada assinante tem um unidade interna ADSL permitindo receber fluxos a 1,536 Mbps (downstream) e transmitir a 16 Kbps (upstream), mais uma comunicação analógica a 4 KHz.

## Capítulo 10 Protocolos de Transporte para Comunicação Multimídia

A função básica de qualquer protocolo de transporte é fornecer funções e serviços necessários às aplicações pelo uso dos protocolos de níveis mais baixo e a rede física. No caso de aplicações multimídia, um protocolo de transporte deve fornecer funções apropriadas para estas aplicações. Este capítulo analisa os principais requisitos para protocolos de transporte multimídia e a partir deste estudo analisa alguns protocolos.

### 10.1 Requisitos de Protocolos de Transporte Multimídia

Protocolos de transporte multimídia diferem dos protocolos de transporte de dados convencionais no sentido que estes primeiros devem suportar garantias de QoS para aplicações multimídia. A função do protocolo de transporte multimídia é então estabelecer e manter uma conexão com garantias de QoS e fornecer uma interface para as aplicações. Os três principais requisitos são alta vazão, suporte *multicast* e a interface deveria fornecer mecanismos para a especificação da QoS que deve ser garantida pelos protocolos das camadas mais baixas.

#### 10.1.1 Alta Vazão

Dados multimídia, especialmente vídeo, necessitam de uma grande largura de banda. Por exemplo, um vídeo de alta qualidade compactado necessita cerca de 1.4 Mbps (como apresentado no capítulo 9). Todos os dados passam pela pilha de transporte, assim o protocolo de transporte deve ser rápido suficiente para suportar este requisito de grande largura de banda. Do ponto de vista da aplicação, como uma aplicação pode envolver vários fluxos de dado, a velocidade do protocolo de transporte deve ser maior que a largura de banda agregada destes fluxos.

Outra maneira de ver o requisito de vazão de um protocolo de transporte é a partir do sistema de comunicação. A vazão do protocolo de transporte deveria ser maior ou próximo a velocidade de acesso a rede. Senão a largura de banda fornecido pelos pontos de acesso à rede não poderiam ser inteiramente usados e o protocolo de transporte seria então o gargalo no sistema de comunicação.

Protocolos otimizados a velocidade mas não suportando garantias de QoS e/ou *multicast* são chamados *protocolos de transporte de alta velocidade* ou *lightweight* (leves).

#### 10.1.2 Capacidades Multicast

Muitas aplicações multimídia envolvem múltiplos participantes, sendo assim elas necessitam de capacidades *multicast* do sistema de transporte. *Multicast* é normalmente implementado na camada de rede. Muitos sistemas de transporte multimídia usam o algoritmo IP *multicast* ou assumem a existência de certos algoritmos de roteamento *multicast*.

#### 10.1.3 Especificação e Garantias de QoS

Fluxos de dados multimídia necessitam de garantias fim-a-fim de QoS, em termos de largura de banda, atrasos e variação de atrasos. Para satisfazer estes requisitos, o sistema de transporte deve fornecer mecanismos para as aplicações poderem especificar e negociar parâmetros de QoS.

Uma aplicação pode especificar as características de tráfego usando um conjunto de parâmetros diretamente à camada de Rede, ou usando um identificador de perfil tal como “vídeo qualidade de TV”, “áudio qualidade telefone”. Neste último caso, o protocolo de transporte deve transladar o identificador de perfil em um conjunto de parâmetros de

QoS. Não há um acordo geral sob que método é o mais preferido, o protocolo de transporte deve suportar estas duas possibilidades.

Os parâmetros de QoS especificados ao protocolo de transporte são passados para o protocolo da camada de rede. Um protocolo da camada de rede, chamado *protocolo de reserva*, propaga os requisitos e reserva os recursos necessários ao longo da conexão de rede. Esta conexão é normalmente uma conexão multicast em aplicações multimídia.

Fornecer garantias de QoS necessita a cooperação de todos os subsistemas do sistema de transporte, incluindo execução da pilha de transporte, gestão de recursos (teste de admissão, alocação de recursos e policiamento), controle de acesso à rede e gestão de filas em comutadores. Para fornecer garantias de desempenho fim-a-fim, o desempenho da execução da pilha de transporte deveria ser garantida. Parte da pilha de transporte (incluindo protocolo de transporte, protocolo de rede, e outros protocolos das camadas mais baixas) é implementada em software. Como a execução deste software é controlada pelo sistema operacional, para garantir o desempenho da execução da pilha de transporte, é necessário um sistema operacional que possa fornecer garantias de QoS para as aplicações multimídia.

## 10.2 Protocolos de transporte tradicionais

Os protocolos de transporte mais comuns são o TCP da série de protocolos Internet e o TP4 da série de protocolo OSI. Estes protocolos foram projetados para a comunicação de dados confiável em redes de baixa largura de banda e altas taxas de erro. Como resultado, eles não são otimizados para operações de alta velocidade, além de não fornecer suporte de QoS e *multicast*. Portanto eles são indesejáveis para comunicações multimídia. Os principais aspectos indesejáveis destes protocolos para aplicações multimídia são: cópia de dados, multiplexação em camadas, controle de fluxo, controle de erro e não fornecimento de suportes de QoS.

### 10.2.1 Cópia de Dados

Os dois maiores objetivos das séries RM-OSI (Modelo de Referência OSI) e TCP/IP foram quebrar o problema de comunicação em camadas simples e permitir a interoperabilidade entre os diferentes níveis. O alto desempenho não foi claramente o objetivo destes protocolos.

A estrutura em camadas das séries RM-OSI e TCP/IP representa um gargalo para a comunicação. Nesta estrutura em camadas, os dados movem-se de uma camada para outra, onde em cada camada o dado é processado novamente, copiado para um novo espaço de memória e novos dados de controle são adicionados. Esta cópia desnecessária contribui à latência de rede. Ela consome tempo porque o acesso à memória é um processo relativamente lento comparado com a velocidade da CPU. A solução é reduzir as cópias de dados através da passagem de ponteiros entre camadas em vez de copiar o dado em si.

### 10.2.2 Multiplexação em camadas

Multiplexação pode ser vista como uma função em uma camada que mapeia um número de conexões da camada superior em uma conexão na camada inferior. O principal efeito da multiplexação é o compartilhamento dos recursos das camadas inferiores. Algumas multiplexagens são justificadas, como para o compartilhamento do meio de transmissão. Mas multiplexação em muitas camadas reduz o desempenho. Isto pois ela aumenta a complexidade dos protocolos, aumenta os atrasos e variações, e conexões diferentes com diferentes requisitos de QoS são tratadas da mesma forma, resultando em um desperdício de recursos ou falha na satisfação da QoS.

### 10.2.3 Controle de Fluxo

O mecanismo de controle de fluxo mais comum é o controle de fluxo por janela de escorregamento, que permite um número fixo de bytes (uma janela de bytes) serem

transmitidos sem necessidade de confirmação ou reconhecimento pelo receptor. Após o envio de uma janela de bytes, o transmissor deve aguardar um reconhecimento. O tamanho da janela de escorregamento do TCP é 64 Kbytes. Para redes lentas, este tamanho é muito grande. Por exemplo, uma rede de 64 Kbps leva 8s para transmitir 64 Kbytes. O atraso de ida-e-volta normal é muito menor que 8 segundos, assim o transmissor receberá um reconhecimento antes de acabar o envio dos bits de uma janela.

Para transmissões em alta velocidade, o mecanismo de controle de fluxo por janela de escorregamento do TCP não é satisfatório. Isto pois o tamanho da janela é muito pequeno e o transmissor aguardará muito para receber a permissão de transmissão. Assim, a largura de banda não é inteiramente utilizada. Por exemplo, o transmissor enviará 64 Kbytes em 50 ms na velocidade de 10 Mbps. Para uma WAN o atraso ida-e-volta é normalmente muito maior que 50 ms (estamos nos dirigindo para um atraso fim-a-fim de cerca de 150 ms). Uma solução parcial seria aumentar o tamanho da janela.

Este mecanismo de controle de fluxo não é adequado para multimídia. Ele assume que a taxa de transmissão pode se adaptar às condições da rede e do receptor. Isto não é possível para mídias contínuas que deveria ser enviada numa taxa intrínseca. Por exemplo, se um sinal de áudio é amostrado a 8000 amostras por segundo com 8 bits por amostra, 8000 amostras deveriam ser transmitidas e recebidas todo segundo (com uma pequena variação) para uma apresentação de boa qualidade.

O controle de fluxo mais apropriado para transmissão de dados contínuos em alta velocidade é o controle baseado em taxa. O transmissor, rede e receptor entram em acordo com a taxa de transmissão em termos de média, taxa de pico e assim por diante antes da transmissão do dado. Quando a taxa é aceita por todos as partes, o transmissor envia nesta taxa, a rede fornece certas garantias para este tráfego, e o receptor aloca recursos suficientes para receber e apresentar o dado na taxa especificada.

#### 10.2.4 Controle de Erro

O TCP e o TP4 fornecem uma comunicação de dados confiável. Quando um pacote é perdido ou corrompido ele é retransmitido. Esta estratégia de retransmissão não é ideal para comunicação multimídia por várias razões:

- A implementação de estratégias de retransmissão envolve temporizadores e buffers grandes e tornam o protocolo complicado e lento.
- Dados multimídia toleram algum erro ou perda, sendo assim o controle de erro é desnecessário para as tecnologias de rede atuais.
- Retransmissão causa atrasos para dados subsequentes, resultando geralmente em mais dados sem utilidade no receptor (pois dados multimídia que chegam atrasados normalmente são descartados).

Para comunicações multimídia, somente a detecção de erros deve ser fornecida. Na detecção de um erro, a aplicação deveria ser notificada e é ela que deveria decidir se a retransmissão é necessária ou não. Se necessária, apenas os pacotes perdidos são retransmitidos (retransmissão seletiva). Outra solução é a codificação *forward error correction* (FEC) em que informações extras são enviadas para permitir correções de erro no receptor sem necessidade de retransmissão. O problema desta solução é o consumo adicional de largura de banda.

#### 10.2.5 Falta de Suporte de QoS

As questões anteriores estão associadas com a eficiência e vazão de protocolos de transporte. Estas questões podem ser resolvidas até certo ponto e protocolos de transporte de alta velocidade podem ser obtidos por otimizações, paralelismo e implementação por hardware. Por exemplo, existem otimizações de TCP que podem chegar a velocidades de gigabits/s.

Alta velocidade é necessária, mas outros fatores, como garantias de QoS e *multicast*, também são necessários para comunicações multimídia. Em TCP não existe a noção de

QoS. TP4 permite negociação de parâmetros de QoS, mas os parâmetros são orientados a comunicação de dados e não são suficientes ou desejáveis para comunicações multimídia.

## 10.3 Xpress Transfer Protocol (XTP)

O XTP é um protocolo de comunicação da camada de transferência confiável e leve. A camada de transferência é formada pela integração das camadas ISO transporte e de rede. Ele foi desenvolvido nos anos oitenta para satisfazer requisitos de diversas aplicações distribuídas e tomar vantagem de redes de alta velocidade.

O XTP é um protocolo importante pelas seguintes razões: incorpora as melhores idéias e lições ensinadas pelos protocolos existentes (como TCP, TP4, NETBLT, VMTP e Delta-t); implementado por diferentes organizações e muito usado na área de controle tempo-real e comunicações de dados de alta velocidade *multicast*, e influenciou no projeto de outros protocolos (incluindo protocolos multimídia).

As principais características do protocolo XTP são:

- **Tamanho do pipeline de dados:** para trabalhar com alto produto largura de banda/atraso, XTP permite uma janela de deslizamento de 4 gigabits usando um número de seqüência de 32-bits. Neste caso, o transmissor pode enviar um grande conjunto de dados sem esperar pelo reconhecimento de recepção.
- **Controle de fluxo:** XTP fornece várias opções para o controle de fluxo para prevenir congestionamento da rede e estouro de buffer do receptor:
  - *Controle de taxa:* durante o estabelecimento da conexão são negociados a taxa máxima (bytes/s) e a rajada máxima (bytes/rajada). O transmissor não pode enviar dados numa taxa maior que a taxa máxima e não pode continuamente enviar mais dados que a rajada máxima.
  - *Tradicional:* usando a janela de deslizamento para dados normais.
  - *Reserva:* que permite transmissões apenas quando se conhece a disponibilidade de bufferização no receptor.
  - *Sem fluxo:* desabilita o controle de fluxo de modo que o transmissor possa enviar dados na medida que ele gera. Tal fluxo livre ou modo *streaming* de operação é desejável para comunicações multimídia.
- **Prioridades de mensagem:** XTP permite ao usuário especificar prioridades às suas mensagens (32-bits) e todos os pacotes que resultam da segmentação da mensagem herdam a mesma prioridade. As mensagens de mais alta prioridade são servidas primeiro nos nós da rede. Esta é uma característica essencial para suportar aplicações tempo-real e multimídia.
- **Dados fora de faixa (Out-of-Band):** algumas vezes é útil enviar informações acerca do fluxo de dados sem embutir isto dentro do fluxo de dados. Como opção, XTP pode transportar com cada pacote de dados até 8 bytes de dados etiquetados. Dados etiquetados são passados do transmissor ao receptor e sua presença é indicada ao receptor, mas nada é interpretado pelo XTP. Dados etiquetados são úteis para protocolos da camada superior para fornecer informações semânticas acerca do dado (p.e. para propósito de sincronização para a camada de apresentação). Dados etiquetados fornecem também uma maneira para transmitir marcas temporais.
- **Controle de erro selecionável:** XTP fornece três opções de controle de erro: *controle de erro tradicional*, que usa *checksums* de transporte para transmissão confiável; *reconhecimento negativo rápido*, na qual o receptor pode gerar um reconhecimento negativo rápido para acelerar o processo de retransmissão em casos tal como a chegada de pacotes fora de seqüência (útil em ambientes como LANs onde fora de seqüência implica na perda do dado e não dado atrasado); *sem erro*, que suspende o esquema de retransmissão normal e onde dados recebidos corretamente são passados para a aplicação. Este último é útil para a transmissão de áudio e vídeo digitais.

- **Política contra Mecanismo:** XTP separa política de mecanismo. Ele fornece habilidades e mecanismos (tal como o erro selecionável e opções de controle de fluxo), o usuário ou aplicação decide que mecanismo usar baseado em seus requisitos. Devido a esta flexibilidade, XTP pode ser usado por muitas aplicações diferentes.
- **Multicast:** XTP *multicast* permite que qualquer transmissor envie dados para um grupo de receptores. Esta transmissão é confiável, onde todos os membros ativos do grupo recebe o dado e qualquer erro é recoberto transparentemente pelo protocolo. A gestão de grupo *multicast* é fora da definição XTP, assim operações como juntar, deixar, voltar em uma comunicação *multicast* são manipuladas separadamente.

### **Conveniência para comunicações multimídia**

XTP tem muitas características desejáveis para multimídia, tal como alta vazão, *multicast*, controle de erro selecionável, controle de fluxo e retransmissão. A maior deficiência do XTP é que ele não fornece gestão de QoS. Não há mecanismos para fornecer mecanismos para especificação de QoS (exceto a taxa máxima e a especificação de rajada) e garantias de QoS. Portanto, XTP não é inteiramente apropriado para comunicação multimídia.

## 10.4 Protocolos de Reserva de Recurso

Para fornecer garantias de QoS, técnicas de gestão de recursos devem ser usadas. Sistemas multimídia não podem fornecer QoS confiável aos usuários sem a gerência de recursos (ciclos de processamento de CPU, largura de banda da rede, espaço em buffer nos comutadores e receptores) nos sistemas finais, rede e comutadores. Sem a reserva de recursos, atrasos ou corte de pacotes devido a não disponibilidade de recursos necessários podem acontecer na transmissão de dados multimídia.

Para possibilitar esta reserva de recursos é necessário a existência de um protocolo de reserva de recurso na camada de Rede. Este tipo de protocolo na realidade não executa a reserva do recurso em si, ele é apenas o veículo para transferir informações acerca dos requisitos de recursos e usado para negociar os valores de QoS que o usuário deseja para suas aplicações. Para a reserva de recursos, os subsistemas devem prover funções de administração de recurso que forçam e escalonam acessos a recursos durante a fase de transmissão de dados.

Na seqüência serão apresentados e comparados dois protocolos de reserva de recursos: os protocolos da família Internet ST-II e RSVP.

### 10.4.1 Protocolo ST-II

O ST-II (*Stream Protocol Version II*) [Topolcic, 90] é um protocolo de comunicação, derivado do IP, especialmente projetado para suportar fluxos de mídias contínuas. ST-II foi projetado em 1990 como um sucessor do protocolo ST que foi especificado em 1979. Este protocolo não alcançou o sucesso esperado por seus promotores.

O protocolo ST-II está no mesmo nível do IP, nível rede, e geralmente coexiste com o IP nos roteadores. Ele fornece um serviço orientado a conexão, suporta *multicast* e negociação de qualidade de serviço.

#### **Estabelecimento de Conexão**

Dados podem ser transmitidos apenas após uma conexão um para vários (chamada fluxo) seja estabelecido com sucesso. Estas conexões criam árvores de conexões com largura de banda reservada de uma origem para um conjunto de destinos (Figura 97). O processo de estabelecimento da conexão é o seguinte:

- O estabelecimento da conexão é iniciado quando um agente ST fonte gera um pacote *Connect*. Este pacote contém a especificação de fluxo (parâmetros de QoS) e o conjunto inicial de participantes.

- Cada agente ST intermediário processa o pacote *Connect* e determina os próximos caminhos para todos os receptores, instala o estado de retransmissão *multicast* e reserva recursos ao longo de cada caminho. Cada roteador e host implementando o protocolo ST-II tem um Gerenciador de Recursos responsável pela administração da QoS. Quando o pacote *Connect* chega no nó, o Gerenciador de Recursos checa se há recursos locais suficiente para acomodar a nova conexão sem a violação das garantias de conexões existentes. Se há menos recursos que o necessário, a especificação de fluxo é atualizada.
- Na recepção de uma indicação *Connect*, o receptor deve determinar se ele aceita ou não entrar no grupo, enviando uma mensagem *Accept* ou *Refuse* para a fonte do fluxo.
- A fonte do fluxo deve esperar por uma aceitação ou recusa de cada receptor antes da transmissão de dados. Se ele recebe uma aceitação com uma especificação de fluxo reduzida ele deve adaptar-se a baixa da QoS para o fluxo inteiro ou rejeitar a participação no grupo do receptor específico enviando uma mensagem *Disconnect*. Portanto todos os participantes recebem dados com a mesma especificação de fluxo, embora as capacidades de apresentação podem diferir de um receptor para outro.

Após estabelecida a conexão é criado uma ligação virtual que geralmente segue uma rota fixa sobre a rede. Em condições normais, onde nenhuma falha do roteamento ocorre, os fluxos estabelecidos podem ser servidos com uma QoS garantida (tal como largura de banda e buffer). O ST-II apenas fornece mecanismos para passar e negociar a QoS. A reserva de recurso e escalonamento não é especificada dentro do ST-II. Isto tem que ser implementado por cada agente ST-II.

### **Especificação do Fluxo**

O primeiro passo para obter garantias de QoS é a especificação do fluxo. Alguns parâmetros especificados são:

- *LimitOnPDUBytes* x *LimitOnPDURate*: largura de banda mínima.
- *LimitOnDelay*: limite superior do atraso.
- *AccDelayVariance*: indica a variação de atraso esperado.
- *ErrorRate*: máxima taxa de erro de bits.
- *Reliability*: taxa de perdas de pacote.

### **Troca de Informação**

ST-II foi projetado para transportar fluxos de mídias contínuas. Desde que a ligação é estabelecida, o fluxo segue a ligação virtual unidirecional da fonte aos destinos. Esta ligação recebe um identificador global único SID. Na fase de transferência de dados, todos os pacotes gerados pelo emissor para este fluxo transporta o SID. Um canal de retorno transporta mensagens de controle na outra direção, sem consumir os recursos reservados (Figura 97).

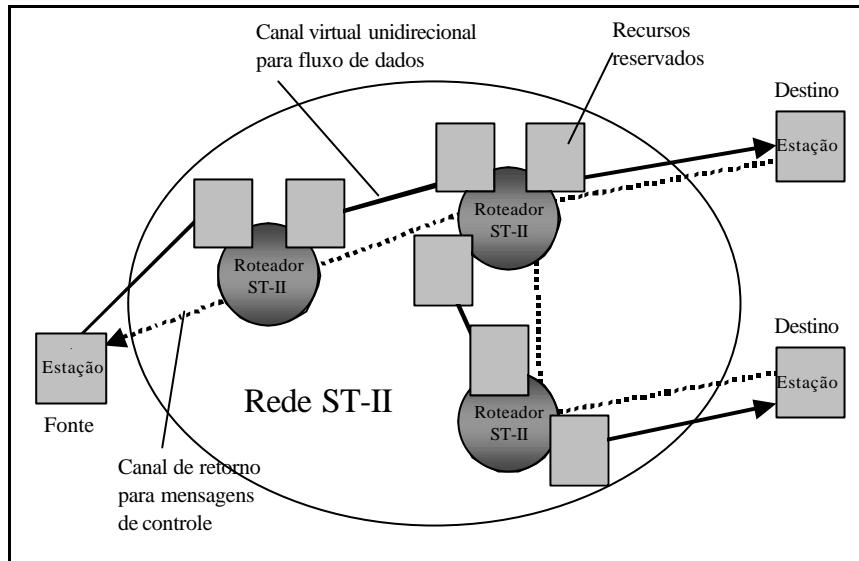


Figura 97. Rede ST-II

### ST-II+

Uma versão revisada do protocolo ST-II surgiu com o ST-II+ [Delgrossi,95]. Além do protocolo de transmissão de dados, também foi definido um protocolo de controle chamado SCMP (*Stream Control Message Protocol*). Ele é usado para configurar conexões multicast orientadas a emissor, isto é, o emissor especifica o grupo de receptores para o multicast no estabelecimento da conexão. A suposição é que a árvore de roteamento é feita uma única vez, e a árvore permanece a mesma durante toda a conexão multicast.

No ST-II+, é possível definir grupos de árvores multicast. Esta característica é útil em comunicações de grupo com modificações de emissores. A observação é que os recursos de rede poderiam ser compartilhados pelas árvores do grupo. Por exemplo, em uma conferência de áudio, nós podemos estabelecer a regra que apenas um membro do grupo pode enviar em qualquer tempo. Neste caso, os recursos reservados para os fluxos de áudio nas ligações podem ser reutilizados por várias das árvores de roteamento baseada em emissor.

### Protocolo ST-II para Multimídia

ST-II tem demonstrado um bom desempenho no suporte de fluxos de pacotes de áudio e vídeo tempo-real. Como ST2 é orientado a emissor, não é fácil a entrada de um novo membro em um grupo multicast existente. No protocolo original ST2, o único modo é enviar um pedido de junção como um pacote IP regular para o emissor, e o emissor pode então reiniciar o processo de construção da árvore multicast. Uma vantagem do multicast orientado a emissor é que o emissor sempre conhece quem são os ouvintes e pode explicitamente decidir se ele deseja aceitar o novo membro. Isto é desejável para muitas aplicações, tal como videoconferência, mas indesejável para outras, tal como uma difusão pública de um evento. Note que os algoritmos multicast baseados em emissor não trabalham bem para grupos grandes e dinâmicos, sendo que o emissor se torna um severo gargalo [Kuo, 98], [Schulzrinne, 95].

A mais nova versão ST-II+ do protocolo também permite de um certo modo ir contra o espírito de ST-II e permitir que os receptores se juntem ao grupo multicast sem informar o emissor.

#### 10.4.2 Protocolo RSVP

O RSVP (*ReSource ReserVation Protocol*) [Zhang, 94] é um protocolo projetado para aumentar o suporte para aplicações tempo-real em redes IP. Ele permite a reserva de recursos em um caminho, mas a transmissão de dados é de responsabilidade do IP.

Neste sentido, ele deve ser visto como um protocolo companheiro do IP. RSVP adotou a abordagem sem conexão. Assim, RSVP permanece compatível com a filosofia IP.

RSVP é similar ao ST-II no sentido que é definido uma árvore de conexões com qualidade de serviço garantida. A inovação essencial do RSVP é que a qualidade de serviço não é especificada para a rede pelo emissor da informação, mas pelo receptor. A idéia é que o receptor está melhor colocado que o emissor para saber que qualidade de serviço é necessária. Por exemplo, não há necessidade de que o emissor envie um fluxo de vídeo a 6Mbps para o receptor se ele não tem poder de processamento para decodificar e descompactar mais que 3Mbps. A implicação desta diferença é que RSVP é mais eficiente no uso de recursos da rede, pois é reservado o estritamente utilizável, além de permitir requisitos de receptor heterogêneos.

Resumidamente, o mecanismo de reserva do RSVP trabalha da seguinte forma:

- As aplicações fonte enviam regularmente mensagens especiais chamadas *Path* para um endereço *multicast*. Estas mensagens contém a especificação do fluxo. Esta mensagem estabelece o estado *Path* nos agentes RSVP intermediários que é usado na propagação dos pedidos de reserva (feita pelos destinatários) para uma fonte específica.
- Na recepção da mensagem *Path*, cada receptor usa informações desta mensagem e informações locais (recursos computacionais, requisitos da aplicação, restrições de custo) para determinar a QoS. Em seguida, ele responde a mensagem *path* por uma mensagem *Reservation* especificando a qualidade de serviço requerida.
- A rede reserva recurso no caminho de retorno da mensagem *Reservation* para a aplicação fonte. Na passagem da mensagem *Reservation*, os agentes intermediários reservam recursos de rede ao longo do caminho e usam o estado *Path* estabelecido para propagar o pedido para o grupo emissor. A propagação da mensagem *Reservation* termina quando o caminho emenda em uma árvore de distribuição com recursos alocados suficientes para satisfazer os requisitos pedidos.

Quando a qualidade de serviço exigida por um receptor difere do fluxo emitido pela fonte, filtros de tráfego são usados para reduzir o requisitos de QoS nos agentes RSVP apropriados. Por exemplo, se um receptor é apenas capaz de apresentar imagens preto&branco e a fonte libera dados de imagens coloridas, um filtro será usado para remover os componentes de cor. Portanto, o estilo de reserva iniciado pelo receptor acomoda requisitos heterogêneos dos receptores. Além de propósito da filtragem é preservar a largura de banda da rede.

Observe que RSVP não transfere o dado em si, um mecanismo para associar a rota reservada e a transmissão de dados é necessário. Não é muito claro como isto pode ser feito de maneira eficiente para garantir a qualidade de serviço na transmissão.

A fim de suportar uma grande variedade de aplicações, o RSVP permite um agrupamento de reservas muito flexível. Em uma primeira opção, um receptor pode especificar no seu pedido se a reserva deveria ser *Distinct*, isto é apenas para pacotes originários de um emissor particular, ou *Shared*, para fluxos de pacotes originários de todos os emissores. Em uma segunda opção, ele pode restringir a reserva para uma lista de emissores (*Explicit*) ou permitir que todos os emissores possam usar os recursos (*Wildcard Filter - WF*). A combinação destas opções é mostrada na tabela abaixo. Por exemplo, em uma videoconferência, o receptor poderia usar o estilo WF de reserva para um fluxo de áudio; qualquer áudio enviado ao grupo poderia usar os mesmos recursos, sob a condição que apenas uma pessoa poderia falar ao mesmo tempo. Ele poderia usar o estilo FF, com uma reserva por emissor, para assegurar que janelas individuais de vídeo de alta qualidade pudessem mostrar todos os participantes da conferência.

Reserva	Distinct	Shared
Escolha do emissor		
Explicit	Estilo Fixed-Filter (FF)	Estilo Shared-Explicit (SE)
Wildcard	Não especificado	Estilo Wildcard Filter (WF)

O protocolo RSVP permite o *tunelamento*: se um nó IP intermediário entre dois nós RSVP não implementa o RSVP, ele pode retransmitir as mensagens RSVP. Mas neste caso, nenhum recurso será reservado no nó intermediário; assim, o caminho fim-a-fim terá uma ligação melhor esforço, e as garantias fim-a-fim determinística não será mais possível.

### **Problema do comportamento anti-social dos receptores**

Toda rede usando um mecanismo onde as reservas de recursos são feitas via declaração dos usuários dos sistemas sofrem da mesma fraqueza potencial: como assegurar que os receptores não peça recursos que eles não precisam (problema de honestidade). Além disso, como evitar que usuários peçam sistematicamente a maior qualidade. Estes problemas são referenciados como problemas de comportamento anti-social dos receptores.

Existem três maneiras de tratar este problema:

- *Taxação*: se há um custo associado à reserva de recursos o gestor do sistema terá cuidado para não pedir recursos desnecessários.
- *Configuração autoritária*: em ambiente fechados (como Intranets) a configuração do software de rede do usuário final pode ser determinado e forçado por uma autoridade central.
- *Pressão social por usuário da comunidade*: esta é a última alternativa: apontar os sistemas e gestores anti-sociais. Isto pressupõe a detecção de tal comportamento, que não é uma tarefa trivial.

### **10.4.3 Comparação de ST-II e RSVP**

#### **Abordagem de Iniciação da Reserva**

ST-II e RSVP têm diferentes abordagens para reserva de recursos. Em ST-II a reserva é iniciada pelo fonte e todos os receptores recebem o dado com a mesma especificação de fluxo sem olhar suas capacidades. Ao contrário, em RSVP a reserva é iniciada pelos destinatários que recebem o conjunto de dados necessários. Portanto RSVP usa recursos de rede de maneira mais eficiente, especialmente quando o grupo *multicast* é grande. De toda maneira, muitas questões ainda devem ser resolvidas pelo IETF e outros pesquisadores. Como o protocolo RSVP não transfere o dado em si, um mecanismo para associar a rota reservada e a transmissão de dado é necessária. Não é claro como isto pode ser feito eficientemente para garantir a qualidade de serviço da transmissão de dados.

#### **Conexão e Sem Conexão**

Outra diferença entre estes dois protocolos é que ST-II é orientado a conexão e o RSVP é sem conexão. Mais especificamente, o fluxo em ST-II é uma conexão um-para-muitos partindo do fonte. Se existir mais que um transmissor, cada fonte requer uma conexão um-para-muitos. Assim cada receptor pode ter várias conexões, um para cada fonte. Estas conexões reservam recursos que podem não serem usados a todos os instantes, desperdiçando recursos. Em RSVP, um fluxo é uma conexão muitos-para-muitos. Cada receptor tem somente uma conexão para receber dados de emissores múltiplos. Esta conexão pode ser usada por diferentes emissores em turnos ou em combinação. Para aplicações como videoconferência, nem todos os emissores estão ativos em qualquer tempo. Portanto RSVP permite um uso eficiente dos recursos da rede.

#### **Uso de filtros nos nós da rede**

A eficiência do RSVP é baseada na existência dos filtros nos nós da rede. A implementação de filtros não é simples. Primeiro, existem várias formas de representação de mídias, assim muitos tipos de filtros são necessários. Existem sugestões onde os filtros devem ser fornecidos pelos emissores ou receptores durante a fase de reserva de recurso, mas o mecanismo exato não é claro. Segundo, para aplicações tempo-real, filtros devem executar suas funções em tempo-real. Isto pode ser

difícil pois algumas funções de filtragem necessitam de um tratamento computacional muito grande e os comutadores geralmente já funcionam sobrecarregados.

### Posição na pilha de protocolos

Outra diferença entre ST-II e RSVP é com relação as suas posições ou regras na pilha de protocolos. ST-II é um protocolo de comunicação completo que contem funções de manipulação e de reserva de recursos. Por exemplo, na Internet ele substitui o IP. Já o RSVP é um protocolo companheiro para um protocolo de transmissão de dados, por exemplo IP. A vantagem da abordagem RSVP é que ele pode trabalhar com diferentes protocolos de transmissão de dados e roteamento. Mas o problema é como assegurar que protocolos, como IP, possam usar recursos reservados pelo RSVP. Por exemplo, como IP é sem conexão, diferentes pacotes podem ser roteados através de diferentes caminhos. Se este é o caso, a QoS destes pacotes IP não pode ser garantida.

### Gerenciamento de informação de estado dos fluxos

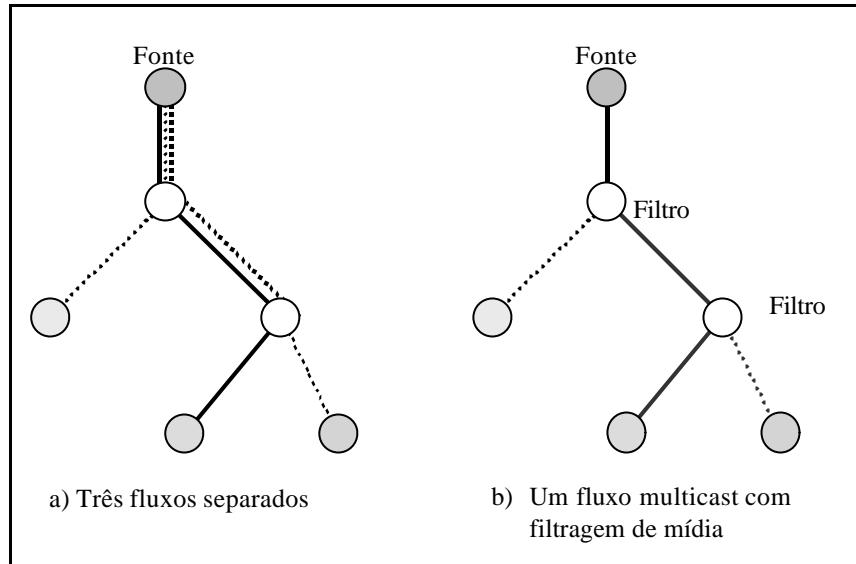
Outra diferença é o modo em que as informações de estado da conexão são gerenciadas. ST-II é um protocolo orientado a conexão que necessita de informações de estado para assegurar as conexões para cada nó participante. Similar ao ST-II, RSVP armazena em cada nó participante informações acerca de fluxos existentes. Estas informações são chamadas de *soft-state*, que significa que as informações de estados são guardadas na base de *timeout* e necessitam ser periodicamente atualizada.

#### 10.4.4 Filtragem de mídia

Muitos dos algoritmos que nós temos visto assume que uma aplicação conhece seus requisitos de QoS em avanço. Quando uma conexão é estabelecida, ou um fluxo iniciado, os recursos requeridos para manter a QoS são reservados ao longo do caminho do emissor aos receptores. A rede garante que os recursos serão disponível durante a conexão. Mas, em um cenário multicast, nem todos os receptores têm os mesmos requisitos de QoS. Por exemplo, um PC conectado via uma linha telefônica não é capaz de receber um vídeo na mesma taxa de uma estação de trabalho UNIX de última geração conectada via ATM. Uma solução para este problema é a filtragem de mídia.

A idéia básica é adicionar nós internos a rede mais inteligentes; se eles entendem o conteúdo de um fluxo de dado, eles podem filtrar pacotes que descem na árvore multicast. A Figura 98 mostra um exemplo, Os nós folhas são os receptores de um vídeo multicast. Os diferentes receptores tem diferentes requisitos de QoS baseado nos seus poderes de processamento locais, capacidade da linha, e preferências do usuário final. Se a rede não for capaz de filtrar o fluxo de mídia, o emissor tem que configurar três conexões separadas, desperdiçando largura de banda nas ligações. Por outro lado, se os nós internos da rede implementam a filtragem de mídia, o emissor cria apenas um fluxo satisfazendo a máxima QoS, e uma considerável largura de banda será salva.

Filtragem de mídia pode trabalhar apenas se um fluxo de dado multimídia pode ser decomposto em sub-fluxos com diferentes parâmetros (resolução, velocidade, etc.). Por exemplo, um áudio codificado PCM não pode ser reduzido omitindo três pacotes de cada quatro. Codificação hierárquica de fluxo é necessário, como aquele presente em MPEG-1 e MPEG-2: um fluxo de vídeo MPEG pode ser empacotado de tal modo que abandonando  $m$  pacotes de um quadro consistindo de  $n$  pacotes, nós obtemos uma imagem completa com qualidade mais baixa. Isto é chamado de técnica LRM (*Layered Receiver-media encoding*). Com isto podemos entender a relação entre técnicas de codificação de mídia e algoritmos e protocolos na rede.

**Figura 98.** Filtragem de Mídia

O nível de inteligência que deveria ser posta na rede é ainda uma questão aberta para pesquisas. Alguns pesquisadores consideram que os nós internos de uma rede deveriam ser tanto quanto possível eficiente e independente da aplicação; o grande crescimento da Internet se deu principalmente devido a este princípio de projeto. Outros propõem a carga remota de programas filtro escrito pelo usuário em nós internos da rede; isto poderia oferecer máxima flexibilidade mas introduz o perigo do cavalo de Tróia. Muitos operadores de rede poderiam provavelmente recusar o download de código escrito pelo usuário em seus nós da rede. Uma solução intermediária razoável poderia ser a escrita de filtros controlados por parâmetros para um menor conjunto de codificações de fluxos mais usados e fornecer estes como parte do código executando nos nós da rede, similar aos plug-ins dos navegadores WWW. Quando um fluxo é configurado, os parâmetros dos filtros são setados pelos receptores ou pelo emissor para otimizar a filtragem da mídia. O Protocolo RSVP é perfeitamente adequado para esta abordagem.

## 10.5 Protocolo RTP

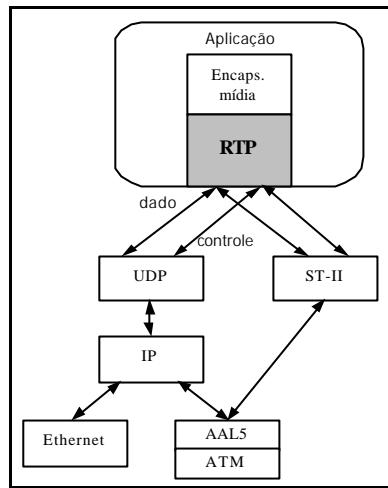
Mais recentemente, vários grupos de trabalho Internet tem sido formados para estudar os requisitos da comunicação multimídia e tempo-real. O Grupo de Trabalho Transporte de Áudio e Vídeo está desenvolvendo um protocolo que forneça um serviço de transporte fm-a-fim para dados com características tempo-real como áudios e vídeo, o protocolo RTP (*Real-Time Transport Protocol*) [Schulzrinne, 97]. O objetivo principal do RTP é satisfazer as necessidades da conferência multimídia multi-usuários, mas seu uso não é limitado a isto.

RTP não contém praticamente nenhuma suposição específica sobre as capacidades das camadas inferiores. Ele não contém nenhum endereço da camada de rede de forma que RTP não é afetado por mudanças de endereçamento. Além disso, qualquer capacidade adicional da camada inferior como segurança ou garantias de qualidade de serviço pode ser usada obviamente por aplicações que empregam RTP. Inclusive RTP suporta transferência de dados para destinos múltiplos usando distribuição *multicast* se provido pelas camadas inferiores.

Embora RTP seja um protocolo ainda muito novo, várias aplicações Internet já implementam RTP, em particular, as ferramentas MBone vic e vat (capítulo 12). Os algoritmos RTP não são usualmente implementados como uma camada separada mas são parte do código da aplicação. É esperado que a nova geração de navegadores WWW também usarão RTP para fluxos de áudio e vídeo.

A Figura 99 mostra as relações entre vários protocolos. Nesta estrutura, o objetivo do RTP é fornecer uma comunicação fim-a-fim, tempo-real sobre a Internet. Note que ele fornece uma função de transporte de dados e é chamado protocolo de transporte, mas

ele é realmente usado com freqüência no topo do UDP, que é um protocolo de transporte sem conexão. Para fornecer garantias de qualidade de serviço, RTP deveria operar no topo de um protocolo de reserva de recursos tal como o ST-II.



**Figura 99.** Pilha de protocolo típica para mídias contínuas usando RTP

RTP consiste de duas partes, uma parte de dados e outra de controle. O último é chamado de RTCP (*RTP Control Protocol*).

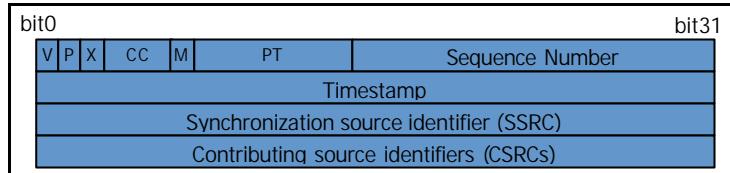
### 10.5.1 Parte de dados do RTP

A parte de dados do RTP é um protocolo fornecendo suporte para aplicações com propriedades tempo-real tal como mídias contínuas. Aplicações enviam dados em unidades de dado de protocolo RTP. A Figura 100 apresenta o cabeçalho de uma RTP PDU. Os significados dos campos são os seguintes:

- *Version* (V, 2 bits): identifica a versão utilizada do padrão RTP. A versão atual é identificada por 2, o valor 1 é usado para a primeira versão do draft e o valor 0 é usado pelo protocolo inicialmente implementado pela ferramenta de áudio vat.
- *Padding* (P, 1 bit): se este bit estiver a um, o pacote contém um ou mais bytes adicionais no final (p.e. necessários por exemplo para criptografia).
- *Extension* (X, 1 bit): quando a um, uma extensão do cabeçalho segue o cabeçalho (contém informações dependentes de aplicação).
- *CSRC count* (CC, 4 bits): contém o número de identificadores CSRC que seguem o cabeçalho fixo.
- *Marker* (M, 1 bit): a interpretação deste bit é definido pelo perfil de mídia que define as características de diferentes mídias. Este campo é usado por exemplo para definir eventos como fronteiras de um quadro de vídeo.
- *Payload type* (PT, 7 bits): identifica o formato do dado transportado no pacote, por exemplo, vídeo MPEG, H.261 ou Motion-JPEG. O mapeamento do código ao formato do fluxo é normalizado.
- *Sequence Number* (16 bits): contém o contador de número de seqüência dos pacotes RTP que é incrementado por um para cada pacote enviado.
- *Timestamp* (32 bits): esta marca temporal é incrementada com a freqüência de relógio e determinada pelo formato do dado transportado no pacote. Por exemplo, para um áudio a taxa fixa, a marca temporal poderia ser incrementada por um para cada amostra. Vários pacotes RTP consecutivos podem ter marcas temporais iguais se eles são gerados logicamente no mesmo instante (p.e., pertença ao mesmo quadro de vídeo).
- *SSRC* (32 bits): identifica a fonte do fluxo de pacotes RTP (fonte de sincronização). Todos os pacotes de uma fonte de sincronização fazem parte um mesmo espaço de número de seqüência e de tempo de maneira que o receptor pode montar o dado de uma mesma fonte para a apresentação. Este identificador é escolhido aleatoriamente (existem detecção de conflito a este nível). Exemplos

de fontes de sincronização incluem o emissor de um fluxo de pacotes derivados de uma fonte de sinal tal como microfone ou câmera, ou um mixer RTP<sup>18</sup>.

- CSRC list (até 15 itens, 32 bits cada): identifica as fontes que contribuíram para o fluxo combinado produzido por um mixer RTP. O número de CSRCs é dado pelo CC. Identificadores CSRC são inseridos por mixers, usando os identificadores SSRC das fontes contribuintes. Uma aplicação exemplo é a audio-conferência onde um mixer indica todos os interlocutores cuja fala foi combinada para produzir o pacote, permitindo que o receptor indique o interlocutor atual, mesmo que todos os pacotes de áudio contém o mesmo identificar SSRC (que é do mixer).



**Figura 100.** Cabeçalho de uma PDU RTP

As seguintes propriedades do RTP são interessantes para multimídia:

- Reconstrução temporal: o protocolo RTP não fornece nenhum mecanismo para garantir tempo de entrega, nem efetua reordenação de pacotes. Porém, com as marcas temporais e número de seqüência incluídas no cabeçalho da mensagem, a aplicação recebedora dos datagramas pode reconstruir a temporização da origem e a seqüência de pacotes enviados. Para mídias a taxa de bits constante (CBR), como áudio PCM, os fluxos podem ser sincronizados pelos números de seqüência, mas para mídias a taxa de bits variável (VBR) tal como vídeo codificado MPEG são necessárias marcas temporais.
- Detecção de perdas pode ser feita a partir da numeração dos pacotes com um número de seqüência definido no cabeçalho da mensagem. Com a numeração de seqüência, é possível também que seja determinada a correta localização de um pacote sem necessariamente decodificar os pacotes em seqüência, como em decodificações de vídeo. Além disso, um *checksum* pode ser adicionado em extensões do cabeçalho dependentes de aplicação.
- Identificação de conteúdo, graças a um código no cabeçalho da mensagem (por exemplo, se é um vídeo JPEG ou áudio GSM).
- A marca temporal e um identificador da fonte de sincronização (um microfone, um mixer, ou uma câmera) podem ser utilizados para realizar sincronizações intermídia e intramídia para aplicações multimídia. O receptor é então capaz de identificar a fonte de cada pacote podendo colocar no contexto apropriado e apresentar em um tempo apropriado para cada fonte.
- Mídias múltiplas podem ser multiplexadas sobre uma conexão e demultiplexadas no destino, isto graças a identificação das fontes contribuintes contida nos pacotes.

Dados de mídias contínuas são transportadas em pacotes de dados RTP. Se pacotes RTP são transportados em datagramas UDP (como ilustrado na Figura 99), pacotes de dados e de controle usam portos separados. Se outros protocolos servem o RTP (como RTP diretamente sobre ATM AAL5), é possível transportar os dois uma unidade de dado de protocolo única, com controle seguido do dado.

### 10.5.2 Parte de controle do RTP

O protocolo de controle do RTP, chamado de RTCP, fornece suporte para conferência tempo real de grupos de qualquer tamanho dentro da Internet. Este suporte inclui:

- Monitorização de QoS e Controle de Congestionamento: os membros da sessão emitem periodicamente relatórios (aproximadamente a cada 5s) para todos os

<sup>18</sup> Um sistema intermediário que recebe pacotes RTP de uma ou mais fontes (possivelmente modificando o formato do dado) e combina os estes pacotes e os envia em um novo pacote RTP. Desde que o tempo entre tempo entre fontes de entradas não serão geralmente sincronizadas, o mixer gerará seu próprio tempo para o fluxo combinado. Assim, todos os pacotes de dados originários do mixer serão identificados como tendo o mixer como fonte de sincronização.

emissores contendo o número de seqüência do último dado recebido, número de pacotes perdidos, e uma medida da variação temporal entre recepções e marcas temporais (necessários à estimativa do atraso de transmissão entre emissores e receptores).

- Sincronização intermídia: aplicações que enviaram dados recentemente emitem um relatório que contém informações úteis para a sincronização intermídia: o tempo real e uma correspondente marca temporal. Estes dois valores permitem a sincronização de mídias diferentes, como a sincronização labial.
- Identificação da fonte: pacotes de dado RTP não identificam a fonte, mensagem RTCP contém o e-mail do usuário emissor e opcionalmente outras informações (nome, endereço, telefone).

### 10.5.3 RTP e a comunicação multimídia

O protocolo RTP tem algumas das características desejáveis para a comunicação multimídia. Algumas delas são:

- Ele é um protocolo leve, assim uma alta vazão pode ser obtida. Não há mecanismos de controle de fluxo e de erro, embora o emissor pode ajustar a taxa de transmissão baseado na informação de realimentação provida pelos pacotes de controle.
- Ele transporta informações usadas para realizar sincronizações intramídia e intermídia para aplicações multimídia.
- *Multicast* é possível se as camadas baixas fornecem roteadores *multicast*.
- RTP especifica ligações com QoS através de um perfil de tráfego. Ele assume que existe um número predefinido de perfis de tráfego que definem requisitos de QoS para aplicações típicas. Por exemplo, se o perfil especificado é "áudio telefone", os parâmetros armazenados no perfil são: taxa de bit = 64kbps, máximo atraso = 100ms e máxima taxa de erro = 0,01. A gestão de recursos e garantias da QoS são de responsabilidade das camadas mais baixas. Por exemplo, ST-II pode ser usado para fornecer garantias de QoS (para reservar recursos ao longo da rota de pacotes de dados).

RTP, apesar do nome, não assegura a comunicação tempo-real. Nenhum protocolo fim-a-fim, incluindo RTP, pode assegurar a comunicação em tempo. Isto requer suporte das camadas inferiores que realmente tem controle sobre a rede nos comutadores e roteadores.

Outro problema é que o cabeçalho do pacote RTP é ligeiramente grande (12 bytes no mínimo) na qual outros cabeçalhos podem ser adicionados, tal como o cabeçalho ST-II. Isto resulta num baixo uso da largura de banda da rede se pacotes menores são usados.

A comunicação do RTP é mais flexível, mas incorre numa sobrecarga na pilha de protocolos, notável particularmente na voz com seus tamanhos de pacote menores. Parece ser impossível satisfazer os requisitos de atraso de 24ms para voz sem cancelamento de eco usando RTP [Lu, 96].

## 10.6 RTSP (Real-Time Streaming Protocol)

O RTSP (*Real Time Streaming Protocol*) (RFC 2326) é uma proposta de padrão IETF (*Internet Engineering Task Force*) para o controle de fluxo de mídia sobre redes IP (incluindo Internet). Ele foi submetido juntamente para a IETF em outubro de 1996 pela RealNetworks e Netscape Communications Corporation e com o suporte de 40 empresas que trabalham em multimídia. O draft foi desenvolvido pela RealNetworks, Netscape, Columbia University, e o Grupo de trabalho IETF MMUSIC, e foi publicado como uma proposta de padrão IETF em Abril de 1998 [Schulzrinne, 98].

O RTSP é um protocolo cliente-servidor a nível de aplicação que permite o controle de fluxos multimídia transferidos, por exemplo, via RTP e outros (UDP, *multicast* UDP, TCP). Ele fornece métodos para realizar comandos similares às funcionalidades providas por um leitor de CD ou um videocassete, tal como partir apresentação, avanço rápido,

pausa, parada e registrar. Em outras palavras, RTP age como um "controle remoto de rede" para servidores multimídia. A fonte do dado pode incluir dados ao vivo ou armazenados.

RTSP compartilha muitas propriedades com o protocolo HTTP (usado atualmente na Web). A motivação é para reutilizar a tecnologia que foi desenvolvida para o HTTP (*caching* de conteúdo, autenticação, criptografia, PICS - Plataforma para seleção de conteúdo Internet) quando acessando conteúdos multimídia tempo-real.

Este protocolo suporta as seguintes operações [Schulzrinne, 98]:

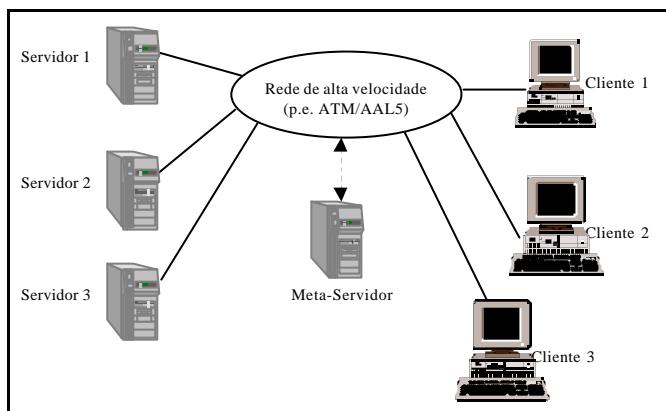
- Acesso de mídias de servidores de mídia: O cliente pode pedir uma descrição de apresentação via HTTP ou algum outro método. Se a apresentação é *multicast*, a descrição de apresentação contém os endereços *multicast* e os portos a serem usados pelas mídias contínuas. Se a apresentação é para ser enviada apenas para um cliente via *unicast*, o cliente fornece o destino por razões de segurança.
- Convite de um servidor de mídia para uma conferência: um servidor de mídia pode ser convidado a se juntar a uma conferência existente para apresentar mídias ou registrar parte ou toda a apresentação. Este modo é útil para aplicações de ensino distribuídas. As várias partes na conferência podem ativar os "botões de controle remoto".
- Adição de mídia a uma apresentação existente: Particularmente para apresentações ao vivo, isto é útil se o servidor pode informar ao cliente informações sobre novas mídias que tornaram-se disponíveis.

Outras informações sobre este protocolo pode ser encontrado no URL: <http://www.cs.columbia.edu/~hgs/rtsp/>.

## Capítulo 11 Servidores Multimídia

Servidores multimídia são componentes importantes em aplicações baseadas em servidores do tipo mídia sob-demanda. Este capítulo apresenta alguns requisitos para este tipo de sistema, e em particular nós consideremos em detalhes os servidores de vídeo (que armazenam vídeos e áudios associados).

Um sistema servidor multimídia consiste de clientes e servidores conectados a uma rede de alta velocidade. Além disso, meta-servidores podem também ser parte de um cenário de servidor multimídia [Bernhardt, 95]. A Figura 101 mostra um modelo simplificado de um sistema multimídia distribuído em que vários clientes e servidores são interconectados por uma rede de alta velocidade. O cliente apresenta um interface ao usuário. Pedidos dos usuários são enviados para servidores apropriados pelo cliente. Quando os servidores recebem estes pedidos, eles recuperam o dado pedido do seu dispositivo de armazenamento e os enviam para os clientes para serem apresentados aos usuários.



**Figura 101.** Arquitetura de Servidor de Vídeo

O meta-servidor fornece funções para os clientes. Por exemplo, um cliente pode perguntar ao meta-servidor os nomes e endereços dos servidores necessários para obter o fluxo de vídeo. O meta-servidor pode também fornecer informações para o cliente tal como nome de arquivos, tamanho de arquivos, taxa de quadros, esquema de compressão ou descrição do conteúdo do vídeo. Dependendo da informação recebida do meta-servidor, o cliente seleciona um servidor de vídeo apropriado.

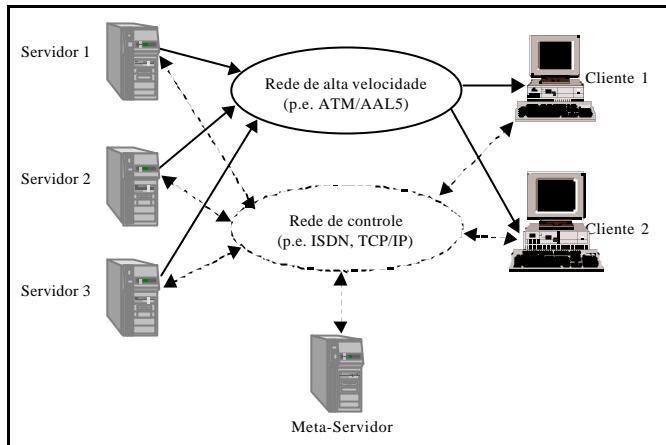
O meta-servidor fornece funções de controle para gerenciar o sistema servidor de vídeo global. Ele pode configurar os servidores e gerenciar o sistema de armazenamento, por exemplo, vídeos populares poderiam ter cópias armazenadas em vários servidores. O meta-servidor pode também fornecer funções de controle de admissão e coletar dados necessários para faturamento, tal como o número e a duração de acesso a vídeos. Estatísticas de acesso colecionadas pelo meta-servidor podem ser usadas para otimizar o desempenho do sistema global.

### 11.1 Separação do controle e transferência de dados

Em aplicações de mídia sob-demanda, os requisitos de comunicação para mensagens de controle e transferência de mídias contínuas são diferentes. Dados de vídeo e áudio provavelmente consumirão mais largura de banda que informações de controle, mas a confiabilidade requerida para dados de controle é muito maior que para dados de vídeo. Além disso, dados de controle são trocados bidirecionalmente, enquanto fluxos de vídeo são em uma direção apenas (do servidor ao cliente).

Portanto, faz sentido usar diferentes sistemas de comunicação para trocar informações de controle e dados de vídeo e áudio [Kuo, 98]. Por exemplo, como ilustrado na Figura

103, redes ATM poderiam ser usadas para transferir dados de vídeo, enquanto ligações de baixa velocidade tal como ISDN ou linhas telefônicas fornecem capacidades suficientes para dados de controle. Para transferência de vídeo, um serviço AAL5/ATM não confiável pode ser tolerável na medida que a perda de pacotes é abaixo de um certo limite, mas confiabilidade de 100%, como provida por um serviço TCP/IP, é necessária para troca de informações de controle. A abordagem de servidor de vídeo descrita em [IBM, 97] usa esta abordagem: AAL5/ATM para transferência de vídeo e TCP/IP para informações de controle.



**Figura 102.** Redes Distintas para controle e mídias contínuas

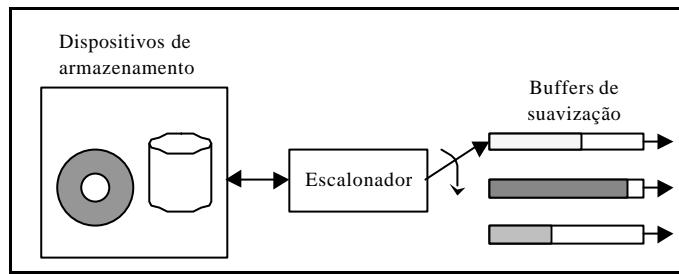
## 11.2 Requisitos de Servidores Multimídia

Abaixo são listados alguns requisitos de servidores multimídia identificados por [Lu, 96]. Estes requisitos são impostos pelas características dos dados multimídia que são basicamente um grande volume de informações e tempo-real.

- A capacidade de armazenamento e a taxa de transferência de dados deveria ser suficientemente alta para suportar vários clientes simultaneamente afim de tornar o sistema econômico. Vídeo e áudio digital são normalmente compactados para reduzir os requisitos de armazenamento e largura de banda de transferência. Mas mesmo após compressão, estes dados continuam a exigir uma capacidade de armazenamento e largura de banda consideráveis.
- Servidores multimídia deveriam fornecer garantias de QoS para os fluxos multimídia. Para satisfazer isto, o servidor deveria implementar controle de admissão e escalonamento tempo-real. Para fluxos a taxa de bits constante, os requisitos de QoS são relativamente fáceis de especificar, e assim o controle de admissão e escalonamento podem ser implementados com uma relativa facilidade para satisfazer os requisitos. Agora quando o fluxo é a taxa de bits variável, pode-se reservar recursos para a taxa de bits de pico para uma garantia a 100% (hard). Como visto no capítulo 7, isto levaria a um baixo uso dos recursos do servidor. Para aumentar o uso dos recursos, garantias estatísticas ou soft deveriam ser usadas e os clientes deveriam permitir um escalonamento da mídia ou degradação suave da qualidade (seção 8.2.4). O problema com a garantia estatística por degradação suave da qualidade é que a especificação da QoS, o gerenciamento dos recursos e o projeto da aplicação se tornam complicados.
- A arquitetura e técnica usada em um servidor deveria ser escalável e capaz de suportar uma grande população de usuários.
- Muitas aplicações multimídia são interativas, assim o servidor deveria ser capaz de suportar vários tipos de interações com o usuário tal como pausa, avanço e retrocesso rápidos.
- O servidor deveria fornecer capacidades de busca.

### **Modelo de Servidor Multimídia**

A Figura 103 apresenta o modelo de servidor multimídia proposto por [Lu, 96]. Os principais componentes são os dispositivos de armazenamento, um escalonador e um buffer de suavização. O escalonador determina qual fluxo é o próximo a ser servido (caso existam vários fluxos sendo transmitidos simultaneamente, ou seja existam vários pedidos de vídeo que estão sendo atendidos pelo servidor). Como a leitura de dados dos dispositivos de armazenamento é realizada em rajada para fluxos individuais, buffers de suavização são usados para transmitir fluxos de mídias contínuas para aplicações (no caso de o servidor e o cliente estando em uma mesma máquina) ou para um sistema de transporte. Desta forma, o dado será transmitido para a rede na mesma taxa que ele deve ser apresentado no cliente. A meta do servidor multimídia é servir simultaneamente tantos fluxos quanto possível satisfazendo os requisitos de continuidade dos fluxos e guardando os requisitos de bufferização ao mínimo. Para obter isto, dispositivos de armazenamento apropriados, posicionamento dos dados nestes dispositivos e técnicas de escalonamento devem ser usados.



**Figura 103.** Um modelo de servidor multimídia [Lu, 96]

Alguns pesquisadores propõem esquemas onde um cliente obtém e armazena um arquivo multimídia completo antes da apresentação do arquivo. Estes esquemas têm a vantagem de que eles não impõem requisitos temporais no servidor e na rede. Mas eles impõem altos requisitos de armazenamento no cliente e o tempo de resposta do sistema também é muito longo quando o arquivo é grande. Além disso, o usuário pode descobrir que o arquivo transferido não é aquele que ele desejava no momento da apresentação, assim o tempo e largura de banda usados para transferir o arquivo completo é perdido.

### **Reserva de Recursos em Servidores Multimídia**

Aplicações de servidor multimídia e servidores de vídeo em particular são normalmente baseadas em reserva de largura de banda afim de fornecer um filme com boa qualidade. Recursos podem ser reservados em avanço, definindo o intervalo de tempo desejado para apresentação. O meta-servidor então calcula o melhor tempo de partida de uma transmissão de vídeo considerando todos os outros pedidos conhecidos e seus requisitos de largura de banda.

O servidor de vídeo tem alguns conhecimentos a priori acerca do vídeo armazenado. Estes conhecimentos podem ser usados para gerenciar eficientemente os recursos de rede necessários, sendo útil para reserva de recursos para transmissão de vídeo compactado em rajadas. Como já visto neste curso, neste tipo de tráfego a reserva da taxa de pico pode desperdiçar largura de banda desde que a taxa de pico é requerida apenas para algumas poucas cenas de um filme. Uma abordagem para evitar desperdício de largura de banda é dividir o fluxo de vídeo em diferentes cenas com diferentes valores de pico de taxa de bits. Para cada cena, uma nova reserva é realizada.

### **Multicast otimiza uso dos recursos**

Se um filme popular é pedido por vários clientes simultaneamente, estes clientes podem ser juntar a um único grupo multicast. O servidor de vídeo pode então fazer o multicast do fluxo de vídeo sob a rede de alta velocidade e usar todas as vantagens da transmissão multicast.

Nas próximas seções, nós discutiremos algumas questões e técnicas para satisfazer os requisitos apresentados nesta seção.

## 11.3 Dispositivos de armazenamento

Nesta seção veremos os requisitos de armazenamento e largura de banda para um servidor multimídia de propósito geral. Serão analisados os vários tipos de dispositivos de armazenamento e as técnicas para aumentar a capacidade de armazenamento e largura de banda de transferência.

### 11.3.1 Requisitos de capacidade de armazenamento e Largura de banda de transferência

Assuma que um servidor armazena 2000 filmes de 100 minutos cada. Estes filmes são compactados em MPEG-2 a 8 Mbps com qualidade de TV digital, e o servidor deveria servir 1000 clientes simultaneamente. Então a capacidade de armazenamento requerida é 12 TB e a taxa de transferência de 8 Gbps. Estes números indicam grosseiramente quanto um dispositivo de armazenamento ou uma combinação de dispositivos de armazenamento de um servidor típico deveria fornecer como capacidade de armazenamento e largura de banda de transferência. Note que a largura de banda de transferência é a largura de banda percebida pelos usuários e não a largura de banda de pico do dispositivo de armazenamento. A largura de banda real pode se aproximar a largura de banda de pico apenas quando o tempo de acesso<sup>19</sup> é minimizado. Minimização do tempo de acesso pode ser obtido por um posicionamento de dados e escalonamento de disco apropriados, como discutido a seguir.

### 11.3.2 Tipos de dispositivos de armazenamento

Os dispositivos de armazenamento mais comuns são discos magnéticos, discos óticos e tapes (por exemplo, fitas DAT<sup>20</sup>). Características típicas destes dispositivos são resumidas na tabela abaixo [Lu, 96]. Um jukebox<sup>21</sup> consistindo de vários discos óticos ou tapes pode fornecer maiores capacidades.

Característica	Disco magnético	Disco ótico <sup>22</sup>	Low-end tape	High-end tape
Capacidade	9 GB	200 GB	500 GB	10 TB
Modo de acesso	Randômico, 50 ms	Randômico, 200 ms	Seqüencial	Seqüencial
Taxa de transf.	10 MB/s	300 KB/s	100 KB/s	1 MB/s
Custo	\$4000	\$50000	\$50000	\$500000
Custo por GB	\$444	\$250	\$100	\$50

Discos magnéticos tem as características desejáveis para suportar aplicações multimídia: eles permitem um acesso randômico rápido e tem altas taxas de transferência. Mas eles são relativamente caros comparado a outros dispositivos de armazenamento. Discos óticos têm mais alta capacidade que discos magnéticos e também permitem o acesso randômico, mas o tempo de acesso é grande e a taxa de transferência é baixa. Tapes têm a mais alta capacidade de armazenamento, mas não podem ser acessados randômicamente (para ler um bloco de dados particular, é necessário ler todos os blocos precedentes) e a taxa de transferência é baixa.

A partir das características dos vários dispositivos de armazenamento apresentados na tabela acima, pode-se ver que não é possível para um dispositivo único satisfazer a capacidade de armazenamento e largura de banda de transferência requerida descrita na seção 11.3.1. A solução é usar vários dispositivos de armazenamento em um servidor. Nós descrevemos três abordagens na seqüência:

<sup>19</sup> O tempo de acesso é o tempo necessário desde o pedido de leitura de dados no dispositivo de armazenamento ao início da transferência destes dados no dispositivo.

<sup>20</sup> Cartuchos Digital Audio Tape (DAT) tem uma capacidade de 2 GB a 24 GB e necessitam de drives relativamente caros. Eles também têm uma taxa de transferência relativamente lenta.

<sup>21</sup> Um jukebox é um dispositivo que armazena vários CD-ROMs (até 500) e usa um braço mecânico, carrossel ou outro dispositivo para ligar o disco à estação ótica para leitura e escrita

<sup>22</sup> Discos óticos podem ser de três tipos: CD-ROM (somente leitura), WORM (para leitura e escrita) e Eraseable optical EO (que podem sofrer escrita, serem lidos e apagados).

- Uso de vários discos para formar um vetores de disco para aumentar a capacidade de armazenamento e largura de banda de transferência.
- Junção de vários dispositivos de armazenamento para formar uma hierarquia de armazenamento de baixo custo para formar um servidor multimídia de grande capacidade.
- Utilizar vários servidores para formar um vetor de servidores.

### 11.3.3 Vetores de Disco e RAID

Como visto anteriormente, o disco magnético é o tipo de dispositivo de armazenamento mais interessante para aplicações multimídia. Mas um simples disco não é capaz de armazenar muitos arquivos multimídia e suportar vários usuários simultaneamente. Uma solução natural é usar vários discos para formar um vetor. Formando vetores de disco aumenta-se a capacidade de armazenamento e a taxa de transferência. O aumento da taxa de transferência é possível pois um arquivo tem seus dados espalhados em diversos discos rígidos e estes dados podem ser lidos de uma maneira paralela, balanceando a carga entre diversos drives.

A tecnologia de vetor de disco muito usada em produtos comerciais e protótipos de pesquisa são os RAIDs (*Redundant Array of Inexpensive Disk*). Hoje, RAID são mais referenciados como Vetor Redundante de Discos Independentes (*Redundant Array of Independents Disk*). RAID são redundantes pois discos extra são usados para armazenar informações redundantes que são usadas para recuperar a informação original quando um disco falha. Comparado com discos grandes e caros, os RAIDs têm melhor desempenho, confiabilidade, menor consumo e escalabilidade, eles são muito usados como dispositivo de armazenamento de dados multimídia.

### 11.3.4 Hierarquia de armazenamento

Discos magnéticos são mais adequados para armazenamento multimídia mas eles são relativamente caros. Para usar a alta capacidade de discos óticos e tapes, uma hierarquia de armazenamento contendo vários tipos de dispositivos de armazenamento é usada para implementar servidores de baixo custo. Existem várias abordagens para construir tal hierarquia de armazenamento.

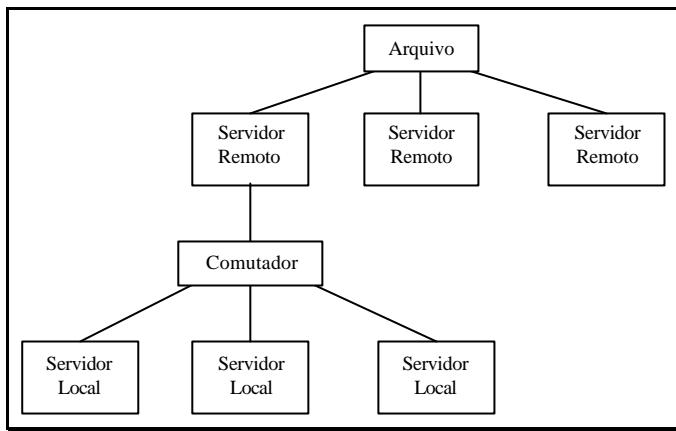
Uma abordagem é usar tapes e discos óticos como arquivos de armazenamento e usar discos magnéticos apenas para armazenar os segmentos iniciais dos arquivos multimídia. Estes segmentos podem reduzir a latência de inicialização e assegurar transições suaves na apresentação. Mas esta abordagem não suporta muitos fluxos ao mesmo tempo devido a taxa de transferência relativamente baixa dos discos óticos e tapes. Além disso, quando tapes são usados, o acesso randômico não é possível.

Em uma segunda abordagem, um arquivo completo é movido dos tapes para os discos magnéticos quando pedido. A deficiência desta abordagem é que o atraso inicial associado a carga do arquivo completo será muito grande para arquivos grandes tal como filmes. Este problema pode ser resolvido explorando os dois fatores seguintes:

- Primeiro, não há muitos arquivos ou filmes populares ao mesmo tempo, assim nós podemos carregar os arquivos mais populares no disco antes de eles serem pedidos ou nas suas primeiras requisições. Os discos são usados como cache para arquivos mais populares.
- Segundo, padrões de usuário são previsíveis em muitas aplicações. Por exemplo, quando um usuário requisita o episódio um de uma série, é claro que ele acessará o episódio dois quando o episódio um terminar. Assim, nós podemos carregar os arquivos previsíveis no disco antes de eles serem pedidos.

O conceito de hierarquia de armazenamento pode ser estendido para uma hierarquia de armazenamento distribuída. Uma hierarquia de armazenamento distribuída consiste de um número de níveis de servidores incluindo servidores locais, servidores remotos e arquivos remotos, como mostrado na Figura 104. Um arranjo possível destes servidores é que os servidores locais são colocados em LANs, servidores remotos em MANs ou redes de campus, e arquivos em WANs. Os servidores locais e remotos usam discos magnéticos como dispositivo de armazenamento, enquanto servidores arquivo usam

discos ópticos ou tapes (ou jukebox) como dispositivo de armazenamento. A vantagem desta organização é a boa escalabilidade.



**Figura 104.** Uma hierarquia de armazenamento distribuída

### 11.3.5 Vetor de Servidores

Um ponto importante em sistemas servidores de vídeo é a QoS. Degradação da qualidade pode ocorrer devido a desempenho insuficiente do servidor ou da rede. Um único servidor pode normalmente não suportar um grande número de clientes com um grande número de fluxos de vídeo diferentes. Vetores de servidores é necessário. Existem duas formas de implementar vetores de servidores [Kuo, 98].

- A primeira abordagem distribui apenas filmes completos em servidores. O cliente deve perguntar ao meta-servidor o servidor de vídeo que pode ser acessado para obter o filme desejado. Esta abordagem adapta-se bem ao acréscimo do número de pedidos enquanto os servidores permanecerem descarregados e capazes de satisfazer certo número de pedidos. Um meta-servidor deveria ter o conhecimento acerca da popularidade dos filmes e configurar um número apropriado de servidores de vídeo para armazenar os filmes mais populares.
- Outra abordagem é implementar um mecanismo de faixas (*striping*) [Bernhardt, 95]. Neste caso, todos os filmes são divididos em subconjuntos não sobrepostos de quadros de vídeo, e cada subconjunto de quadros é armazenado em um servidor separado. Quando um cliente pede um filme, cada servidor transmite seu subconjunto de quadros como um subfluxo. O cliente recebe os diferentes subfluxos e reconstrói eles em um completo fluxo de vídeo. Por exemplo, é possível armazenar todos os quadros de vídeo com os números de seqüência  $3n-2$  no servidor de vídeo 1, quadros  $3n-1$  no servidor 2, e quadros  $3n$  no servidor 3 ( $n \in \mathbb{N}$ ). A abordagem tem a vantagem de que a carga do servidor é bem distribuída entre os servidores e que cada servidor deve armazenar o mesmo montante de dados. De qualquer maneira, os diferentes servidores devem partir suas transmissões aproximadamente no mesmo instante, e a sincronização de fluxo deve ser realizada a fim de evitar grandes tempos de bufferização nos clientes.

## 11.4 Posicionamento de dados no disco

A meta de servidores multimídia é satisfazer restrições temporais dos fluxos multimídia e servir simultaneamente tantos fluxos quanto possível. Para obter isto, o posicionamento apropriado de dados no disco, escalonamento de disco e controle de admissão são essenciais. Esta seção trata do primeiro problema. Em muitos casos, estes três problemas são diretamente relacionados e dependentes: uma certa disciplina de escalonamento de disco requer um tipo particular de posicionamento de dados e baseia-se em um algoritmo de controle de admissão particular.

Um arquivo normalmente é quebrado em um conjunto de blocos de armazenamento para leitura e escrita. Existem dois métodos gerais para posicionar estes blocos: eles são

colocados continuamente em um disco ou espalhados ao redor do disco. Para aumentar o desempenho, variações destes dois métodos foram propostos. Na seqüência, nós discutiremos estes métodos de posicionamento de dados. Mas antes disso será dada uma visão geral do funcionamento dos discos rígidos.

### O que são discos rígidos e como funcional

Em termos gerais, um disco rígido usa discos achatados chamados pratos, revestido nos dois lados por um material magnético projetado para armazenar informações. Os pratos são montados em uma pilha. Estes pratos (o disco) giram a uma rotação constante (3600 a 7200 rpm) desde que o computador é ligado. Dispositivos especiais de leitura/escrita, chamados de cabeçotes, são usados para escrever ou ler informações no/do disco, sendo que sua posição no disco é controlada por um braço atuador. Cada prato contém dois cabeçotes, um na parte superior do prato e outro na parte inferior. Assim, um disco rígido com dois pratos tem quatro cabeçotes. Todos os cabeçotes são presos a um único braço atuador, eles não se movem individualmente.

Dados são organizados no disco em cilindros, trilhas e setores (Figura 105). Cilindros são trilhas concêntricas na superfície dos discos. Hoje, existem aproximadamente 3000 trilhas em cada lado de um prato de 3,5 polegadas. Uma trilha é dividida em setores. Cada setor tem o tamanho de 512 bytes. Em uma operação de leitura de um setor, o controlador de disco interpreta o endereço do dado e move os cabeçotes para o cilindro que contém os dados. Quando os cabeçotes estão na posição correta, o controlador ativa o cabeçote específico para ler a trilha que contém o dado. O cabeçote então lê a trilha procurando o setor que contém o dado para leitura. A placa controladora do disco coordena o fluxo de informação vinda do disco rígido em uma área de armazenamento temporária (cache). Ela então envia a informação pela interface do disco rígido.

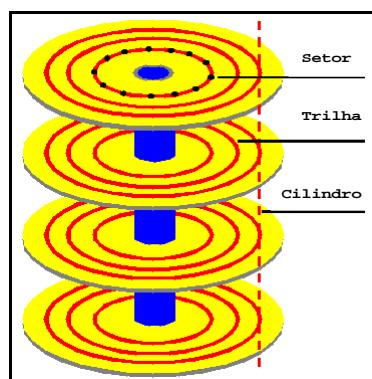


Figura 105. Disco Rígido

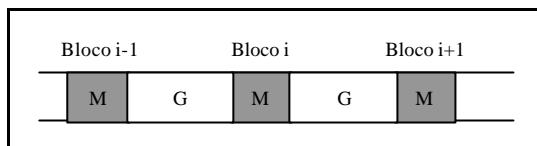
#### 11.4.1 Posicionamento de dados de maneira contínua

No posicionamento contínuo, os blocos de dados de um arquivo são armazenados no disco continuamente. Não há espaços vazios dentro de um arquivo. Arquivos contínuos são simples de implementar. Na leitura de um arquivo contínuo, apenas uma busca (posicionamento do cabeçote do disco na trilha) é necessária para posicionar o cabeçote do disco no ponto de partida do arquivo (se o cabeçote do disco é dedicado a leitura de um arquivo).

Este posicionamento de dados tem desvantagem para edição de dados. Primeiro, uma grande sobrecarga de cópia de dados durante as inserções ou eliminações de arquivo é necessária para manter a continuidade: quando um bloco de dados necessita ser inserido em um ponto do arquivo, todos os blocos de dado abaixo deste ponto deve ser movidos um bloco afrente para remover o espaço livre. Segundo, posicionamento contínuo pode causar fragmentação de disco: quando um arquivo é apagado, o lugar deixado pode apenas ser usado para armazenar outro arquivo com o mesmo ou menor tamanho. Portanto, arquivos contínuos são apenas apropriados para aplicações onde os dados são escritos uma única vez e lidos muitas vezes. Mídia sob-demanda é um destes tipos de aplicação.

### Posicionamento Restrito

Uma variação do posicionamento contínuo de dados é chamado posicionamento restrito onde um espaçamento restrito é colocado entre blocos consecutivos, como mostra a Figura 106. Este posicionamento usa o fato que a taxa de leitura do cabeçote do disco é muito maior que a taxa de apresentação de um fluxo. Assuma que M seja o tamanho do dado dentro de cada bloco e o espaçamento corresponda a um tamanho G de dados. A taxa de transferência para cada cabeçote do disco seja r e a apresentação dos M bytes de dados demora T segundos. Então, o arquivo pode ser lido continuamente se a seguinte condição for satisfeita:  $T \geq (M+G)/r$ .



**Figura 106.** Posicionamento de dados restrito

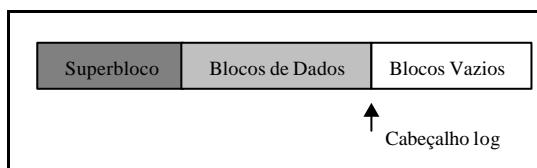
Blocos de outros arquivos podem ser postos nos espaçamentos e a continuidade destes arquivos pode também ser satisfeitas ajustando os tamanhos dos blocos. A condição acima é a condição de admissão quando a disciplina de serviço round-robin é usada (vista mais adiante).

#### 11.4.2 Posicionamento espalhado de dados

Posicionamento espalhado espalha os blocos de dados de um arquivo ao redor do disco e não tem limitações de posicionamento contínuo. Este posicionamento de dados tem várias limitações. Primeiro, alguns mecanismos são necessários para rastrear os blocos de um arquivo (p.e. lista ligada, FAT – *File Allocation Table* no DOS, Inode do UNIX). Segundo, quando da leitura de vários blocos em um arquivo espalhado, uma busca deve ser realizada para a leitura de cada bloco.

#### 11.4.3 Posicionamento de dados log-structured

Posicionamento de dados log-structured consiste de um superbloco e muitos blocos de dados, como mostra a Figura 107 [Lougher, 93]. O Superbloco armazena informações de gerenciamento tal como o ponteiro do cabeçalho log (o endereço do primeiro bloco vazio disponível) e os endereços dos blocos. Arquivos de dados são continuamente armazenados em Blocos de Dados. Dados a serem inseridos são colocados continuamente partindo do cabeçalho log. Buracos deixados após a eliminação de dados não são preenchidos com novos dados. Desta modo, a estrutura log é interessante para leitura e edição.



**Figura 107.** Posicionamento de dados log-structured

Se não há edições, o posicionamento de dados log-structured mantém localidade temporal; que é, o dado que são temporalmente relacionados são colocados em blocos contínuos no disco. Isto ajuda a satisfazer os requisitos e reduz o tempo de busca. Mídias relacionadas, tal como vídeo e áudio associado, podem ser entrelaçado para ajudar a sincronização intermídia.

#### 11.4.4 Posicionamento de dados em vetores de discos

Quando vetores de disco são utilizados, existem várias opções para armazenar o arquivo. Um método simples é armazenar um arquivo completo em um disco. Mas neste caso, o número de acessos concorrentes a este arquivo é limitado a vazão de um único disco. Por exemplo, se a vazão do disco é 4 Mbytes/s e a taxa de apresentação do fluxo

é 1 Mbyte/s, então o número máximo de acessos concorrentes é 4. Isto não é suficientemente grande para um filme popular. Para resolver esta limitação, os filmes populares podem ser copiados em vários discos, mas isto requer espaço adicional em disco. Um método mais econômico é espalhar os blocos de um arquivo em vários discos. Deste modo mais usuários podem acessar simultaneamente sem a necessidade de duplicar o arquivo.

Existem duas técnicas para espalhar os blocos de um arquivo em vários discos:

- *Data stripping* (corte dos dados): onde os blocos são organizados para acesso em grupo (em paralelo), os mesmos setores em todos os discos são acessados ao mesmo tempo.
- *Data interleaving* (entrelaçamento de dados): onde os blocos armazenados em discos diferentes são acessados independentemente.

É preferível que um bloco de armazenamento corresponda a uma unidade lógica de mídia. Por vídeo, um bloco pode corresponder a um quadro ou um grupo de imagens (quando MPEG é usado). Para áudio, a unidade de dados lógica é uma amostra ou grupo de amostras. É comumente necessário sincronizar áudio e vídeo, assim é conveniente fazer o tamanho do bloco de áudio corresponder ao número de amostras em um intervalo de um quadro de vídeo.

Em toda a discussão acima, nós assumimos que os fluxos são codificados a taxa de bits constante. Quando nós temos codificação a taxa de bits variável, posicionamento de dados, escalonamento e controle de admissão será complicado se nós esperamos fazer uso eficiente dos recursos.

## 11.5 Escalonamento de Disco e Controle de Admissão

Como visto anteriormente, para comunicações multimídia de qualidade é necessário que todos os subsistemas envolvidos garantam o desempenho fim-a-fim. No caso de aplicações baseadas em servidor, o servidor é apenas um destes subsistemas: ele deve transmitir o dado para a rede na mesma taxa que ele deveria ser apresentado no cliente.

Como operações de disco são indeterminísticas devido ao tempo de acesso randômico e o compartilhamento dos recursos do servidor com outras aplicações, disciplinas de escalonamento de disco são necessárias para manter a continuidade de dados multimídia contínuos. Como em outros subsistemas, garantias de desempenho podem apenas serem obtidas quando o sistema não está sobrecarregado. Portanto, um controle de admissão é necessário para prevenir a sobrecarga do sistema.

Para evitar a falta de dados (*data starvation*), o acesso aos dados no disco deveria ser feita antes do tempo de transmissão do dado (assumindo que a taxa de transmissão é a mesma que a taxa de apresentação no receptor). O dado excedente deve ser bufferizado. O tamanho deste buffer não deve ser muito grande pois um buffer grande significa que o sistema é caro e o dado sofrerá grande atrasos. Portanto, a tarefa do servidor é prevenir a falta de dados minimizando os requisitos do buffer e atraso.

### Controle de admissão

Técnicas de controle de admissão são normalmente associadas às disciplinas de escalonamento de disco. O controle de admissão no servidor deveria ser baseado no seguinte critério: a largura de banda total (LBT) de todos os fluxos pedidos deveria ser mais baixa que a taxa de transferência ( $T$ ) do disco. Na prática, LBT deveria ser muito mais baixa que  $T$ . Isto pois não há tempo de busca<sup>23</sup> zero quando a cabeça do disco termina um serviço e move-se para outro lugar para servir o próximo pedido.

---

<sup>23</sup> Tempo de busca é o tempo levado pelo braço atuador para mover o cabeçote entre trilhas. Hoje, existe aproximadamente 3000 trilhas em cada lado de um prato de 2,5 polegadas. Assim, para ir da trilha atual para a trilha contendo o dado a ler pode implicar no movimento de uma trilha ou até 2999 trilhas. O tempo de busca de uma trilha é de aproximadamente 2 ms, e o máximo tempo de busca (movimento de 2999 trilhas) consome cerca de 20ms. O tempo de busca média dos drives atuais é de 8 a 14 ms.

### 11.5.1 Algoritmos de escalonamento de disco tradicionais

Algoritmos de escalonamento de disco são usados mesmo em sistemas de arquivo tradicionais, mas seu propósito é reduzir o tempo de busca e a latência de rotação<sup>24</sup>, obtendo mais alta vazão ou fornecendo acesso igual para cada aplicação. Os algoritmos de escalonamento de disco mais comuns são:

- FCFS (*First-Come-First-Served*), onde pedidos são servidos de acordo com sua ordem de chegada. Ele ignora o movimento e a localização do cabeçote do disco, assim o tempo médio de busca é alto.
- SSTF (*Shortest Seek Time First*), que tenta minimizar o tempo de busca servindo o pedido cujo dado está mais próximo da localização atual do cabeçote. SSTF favoriza o pedido no meio de um disco. Quando o servidor é muito carregado, os pedidos de transferência de dados associados às trilhas mais internas e mais externas algumas vezes não são servidos.
- Scan: também tenta minimizar o tempo de busca servindo pedidos na ordem do movimento dos cabeçotes do disco. Ele serve primeiro todos os pedidos em uma direção até que todos os pedidos sejam servidos nesta direção. O movimento dos cabeçotes são então invertido e serve os pedidos nesta direção.

Estes algoritmos de escalonamento não levam em consideração a temporização de cada fluxo. Assim, eles não podem ser diretamente utilizados para escalonamento de servidores multimídia (a não ser que seja limitado o número de pedidos).

### 11.5.2 Algoritmo EDF (*Earliest Deadline First*)

EDF é o algoritmo mais comum para escalonamento tempo-real de tarefas com deadlines. EDF escalona para primeiro o bloco de mídia com deadline mais próximo. A limitação do EDF aplicado ao escalonamento de disco é que ele ignora a posição do cabeçote, produzindo um grande tempo de busca e latência de rotação. Desde que um tempo não relevante é desperdiçado no movimento do cabeçote no disco sem realmente ler dados úteis, assim a vazão não é inteiramente usada.

Suponha que o servidor está servindo  $n-1$  pedidos com garantias hard. A condição suficiente para aceitar um novo pedido  $n$  é:  $(b_1 + b_2 + \dots + b_n) + (n-1)*S*r \leq r$ , onde  $b$  é a taxa de dados máxima do fluxo  $i$ ,  $S$  é o tempo de busca máximo do disco e  $r$  é a vazão do disco. Como EDF não impõe nenhum posicionamento de dados específico e o tempo de busca é indeterminístico (pode variar de 0 ao tempo de busca máximo), é difícil desenvolver um algoritmo de controle de admissão que possa satisfazer todos os requisitos de continuidade dos fluxos e ao mesmo tempo usar recursos eficientemente.

### 11.5.3 Algoritmo Scan-EDF

Scan-EDF combina Scan e EDF para reduzir o tempo de busca médio do EDF [Reddy, 94]. Scan-EDF serve o pedido com deadline mais próximo como EDF, mas quando vários pedidos tem o mesmo deadline, seus respectivos blocos são acessados com o algoritmo Scan. É claro que quanto mais pedidos tiver o mesmo deadline, mais eficiente é o Scan-EDF. Quando todos os pedidos têm um deadline diferente, Scan-EDF se reduz a EDF apenas.

Para aumentar a eficiência do Scan-EDF, algumas técnicas foram propostas para aumentar o número de pedidos com o mesmo deadline. Por exemplo, pode-se forçar que todos os pedidos tenham deadlines que são múltiplos de um período de  $p$ . Neste caso, os deadlines são servidos em blocos.

O controle de admissão do Scan-EDF é similar ao do EDF. Mas em vez de usar o tempo de busca máximo para cada chaveamento de serviço, o tempo de busca máximo é usado apenas quando do chaveamento de um pedido com um deadline para outro pedido com um deadline diferente. Assim, o scan-EDF pode servir mais pedidos simultaneamente (ou ao menos igual ao número de pedidos) que o EDF.

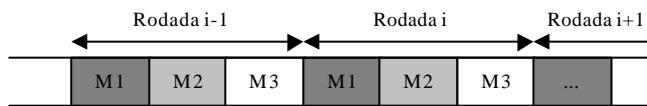
<sup>24</sup> Desde que o cabeçote está posicionado na trilha apropriada, ela deve aguardar para que o drive rode o prato para o setor correto. Este período de espera, chamado de latência de rotação, depende da velocidade de rotação do drive. Um drive de 3600 rpm tem uma latência de rotação média de 8,3 ms e um drive rápido de 7200 a latência média é de 4,2 ms.

### 11.5.4 Algoritmo Round-Robin

No Scan, pedidos são servidos em rodadas, ou seja em cada rodada é lido parte de cada fluxo em transferência e a rodada se repete ciclicamente. Mas os pedidos são servidos em diferentes ordens em diferentes rodadas dependendo da posição do dado. O intervalo de tempo máximo entre dois serviços sucessivos para um fluxo é igual ao tempo de duas rodadas. Isto pois um fluxo pode ser servido primeiro em uma rodada (uma direção de movimento do cabeçote) e o último em outra rodada (outra direção do cabeçote). No EDF, não há rodadas distintas, fluxos com pedidos freqüentes podem ser servidos várias vezes enquanto outros são servidos apenas uma vez.

No escalonamento round-robin, fluxos são servidos em rodadas distintas e a ordem de serviço de cada fluxo é fixa em cada rodada. Assim, o intervalo de tempo entre dois serviços sucessivos para um fluxo é limitado a duração da rodada. Se a duração da rodada é constante, então o intervalo também é constante. Desde que a ordem de serviço para fluxos é fixada, o posicionamento de dados é importante para determinar o tempo de busca.

Quando round-robin é usado com posicionamento de dado restrito, o tempo de busca é mínimo e a continuidade de cada fluxo é mantida. Por exemplo, na Figura 108, três fluxos podem ser obtidos continuamente. Comparando a Figura 108 e a Figura 106, o espaçamento G para o fluxo 1 é  $(M_2 + M_3)$ . Durante a implementação, M e G podem ser ajustados para satisfazer o requisito de continuidade. Por exemplo, se um vídeo HDTV é compactado a 0.5 Mbits/quadro (62500 bytes/quadro) e registrado a 60 fps em um disco com uma taxa de transferência de 25600 setores/s (cada setor é igual a 512 bytes). Escolhendo que cada bloco é um bloco, então o bloco ocuparia 123 setores ( $62500/512$ ). Então  $M=123$ . Usando a condição do posicionamento restrito ( $T \geq (M+G)/r$ ), então  $G \leq (25600/60) - 123$ , ou seja  $G \leq 303$  setores.



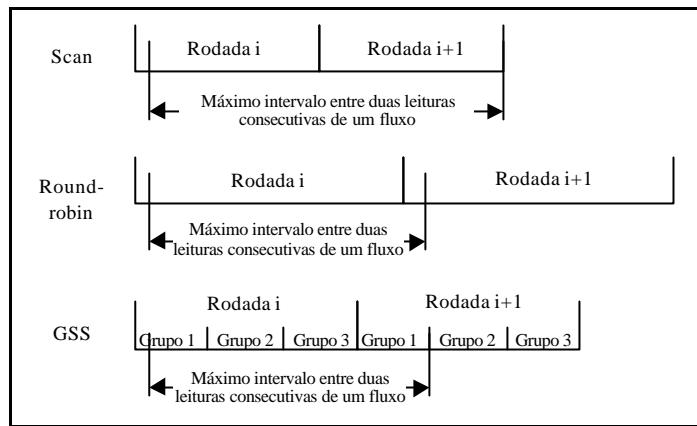
**Figura 108.** Posicionamento de dados restrito

O intervalo de tempo entre acessos sucessivos de um fluxo de dado tem várias implicações nos requisitos de bufferização e o atraso da partida da apresentação. Para round-robin, a apresentação de um fluxo pode ser iniciada após o acesso do primeiro bloco do fluxo (assumindo que a apresentação é na mesma máquina que o servidor). Entretanto, com Scan a apresentação deve esperar até o fim da primeira rodada (pois a posição do serviço na rodada varia). Para prevenir falta de dados no dispositivo de saída, round-robin necessita um espaço de buffer suficiente para satisfazer o consumo de dado para uma rodada, enquanto Scan necessita espaço suficiente para satisfazer o consumo de aproximadamente duas rodadas. Mas em geral, o tamanho da rodada para Scan é mais curto que aquele do round-robin como o tempo de busca é normalmente mais curto para Scan (exceto no caso de posicionamento restrito).

### 11.5.5 Escalonamento GSS (Group Sweeping Scheduling)

No escalonamento GSS, cada rodada é partitionada em grupos. Cada fluxo é associado a um grupo, e os grupos são servidos em uma ordem fixa em cada rodada, tal como em round-robin. Dentro de cada grupo, Scan é usado para servir fluxos do grupo. Scan reduz o tamanho da rodada e round-robin reduz o intervalo entre serviços sucessivos. Assim GSS é um compromisso entre o tamanho da rodada e o intervalo de serviço sucessivo. A Figura 109 compara os tamanhos das rodadas e os intervalos máximos de serviços sucessivos para Scan, round-robin e GSS. O tamanho da rodada é mais curto para Scan, mas o intervalo de serviço máximo é o dobro do tamanho da rodada. O tamanho da rodada do round-robin é o mais longo, mas o intervalo de serviço máximo é igual ao tamanho da rodada. O tamanho da rodada do GSS está entre os tamanhos do Scan e round-robin. Seu intervalo de serviço máximo é um pouco maior que seu

comprimento de rodada. Assim, em muitos casos, o intervalo de serviço do GSS é mais curto.



**Figura 109.** Comparação dos tamanhos de rodada e intervalos máximos entre duas leituras consecutivas de um fluxo para Scan, round-robin e GSS

## 11.6 Interações com o usuário

Um servidor multimídia distribuído deve fornecer uma interface para que os clientes possam pedir informações. Além disso, o sistema deveria suportar interações com o usuário tal como pausa, retomada e avanço rápido.

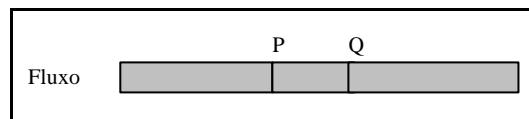
Um servidor pode fornecer dois tipos de interface para o cliente:

- Sistema orientado a arquivo, onde o dado multimídia é tratado como arquivos normais. O cliente utiliza operações típicas de arquivo para acessar o dado multimídia, tal como abrir, fechar e ler. O servidor pode usar a operação abrir para forçar o controle de admissão. Certa quantidade de dados é lida em cada operação de leitura. Usando esta interface, o cliente pode facilmente implementar operações de pausa simplesmente pela parada de emissão de comandos de leitura e a operação retomada pela retomada da emissão de comandos de leitura. Mas neste esquema é difícil implementar avanço e retrocesso rápidos. Além disso, é difícil manter a continuidade do fluxo usando este tipo de interface pois o comando de leitura pode sofrer atrasados indeterminados na rede.
- Sistema orientado a fluxo, onde o cliente emite um comando de partida e o servidor emite periodicamente dados em uma taxa predefinida sem pedidos adicionais de leitura. Este é o modo de operação preferido pois com este é fácil manter a continuidade do fluxo.

A seguir será descrita a implementação de operações de comando semelhantes ao videocassete, como pausa, retomada, avanço rápido e retrocesso rápido. Além destas operações, também é necessário implementar operações de busca e acesso baseado no pedido do usuário. Por exemplo, o usuário pode querer encontrar todos os vídeo-clips similares ao que ele está assistindo.

### 11.6.1 Pausa e Retomada

Se um usuário está acessando um fluxo independente de outros usuários, é relativamente fácil implementar os comandos de pausa e retomada simplesmente pela emissão dos respectivos comandos para os servidores. Mas mesmo neste caso simples, cuidados devem ser tomados para tratar possíveis dados extras enviados ao cliente após a emissão do comando pausa. Isto ocorre pois o comando pausa leva algum tempo para chegar ao servidor, e durante este tempo o servidor pode enviar dados ao cliente. Como mostra a Figura 110, o usuário emite o comando pausa no instante P, mas antes do servidor receber este comando, ele envia dados até o instante Q do fluxo. O que fazer com estes dados recebidos entre P e Q?



**Figura 110.** Comando pausa é emitido no ponto P mas o servidor para de enviar dados em Q

Existem três opções para tratar estes dados extras:

- A primeira opção é não parar a apresentação do fluxo imediatamente após o pedido de parada do usuário. Em vez disso, a apresentação continua até o ponto Q e para quando ele recebe uma confirmação de pausa do servidor. Quando o usuário emite um comando de retomada, o fluxo é retomado no ponto Q.
- Na segunda opção, o cliente para a apresentação imediatamente no pedido do comando de pausa (P) e bufferiza o dado entre P e Q. Quando o usuário emite o comando retomada, a apresentação se inicia em P usando o dado bufferizado e o servidor emite dados a partir do ponto Q.
- Uma terceira opção é parar a apresentação no instante P e descartar os dados recebidos entre P e Q. Quando o usuário emite um comando de retomada, o cliente pede ao servidor o acesso e transmissão de dados a partir do ponto P. Após um certo atraso inicial, a apresentação se inicia em P.

As opções 1 e 2 são mais eficientes pois elas usam os dados recebidos entre P e Q sem que o servidor necessite enviar novamente estes dados. Mas infelizmente é difícil obter sincronização intramídia e intermídia com estas duas opções. Isto pois o tempo entre P e Q são diferentes para fluxos diferentes devido a atrasos aleatórios na rede sofridos pela transmissão do comando pausa do cliente aos diferentes servidores. Além disso, é difícil na segunda opção tornar segura a apresentação suave no ponto Q devido a variações de atraso da rede. Portanto, a opção 3 é preferida a despeito do desperdício de largura de banda devido a retransmissão de dados entre P e Q.

Quando o servidor lê os dados uma única vez e difunde (broadcast/multicast) estes dados a vários clientes ao mesmo tempo, a implementação dos comandos de pausa e retomada é complicada devido ao compartilhamento da informação por vários clientes e é necessário que a interação de um cliente não interfira na apresentação da informação para os outros clientes. Existem duas soluções para este problema:

- Em uma primeira solução, o sistema impõe que o usuário possa apenas parar a apresentação por um certo intervalo. A abordagem para suportar tal cenário é transmitir o filme como N diferentes fluxos com tempos de partidas defasados, por exemplo, fluxo n tem um atraso de 5 minutos comparado ao fluxo n-1. Este tempo de partida defasado é chamado de diferença de fase. Esta abordagem permite ao cliente interromper a transmissão por 5 minutos. Antes da parada, o cliente recebe o fluxo n-1; após a parada ele receberá o fluxo n.
- Em uma segunda opção, um buffer grande e um gerenciamento de buffer complicado são necessários para fornecer a interação independente de clientes individuais [Wong, 95].

### 11.6.2 Avanço e retrocesso rápidos

Existem dois modos de implementar avanço rápido:

- O primeiro modo é apresentar a mídia em uma taxa mais rápida que a taxa original.
- O segundo modo é apresentar a mídia na taxa normal, mas saltando alguns quadros. Por exemplo, se quadros alternativos são saltados, o vídeo avança com uma velocidade duplicada.

A primeira opção é a mais fácil de implementar, mas ela requer grande largura de banda e poder de processamento. A rede e o cliente podem não ser capazes de satisfazerem os requisitos de largura de banda e computacional. Os requisitos podem ser reduzidos usando técnicas de codificação de vídeo escaláveis. Quando o vídeo é compactado com técnicas de subbanda (escalável), apenas a primeira banda com vídeo de baixa qualidade pode ser transmitido para operações rápidas.

O retrocesso rápido pode ser implementado de maneira similar aquelas do avanço rápido exceto que os dados são lidos para trás.

A implementação de avanço e retrocesso rápidos é complicada devido a duas características de compactação: codificação interquadro e taxa de bits variável.

- Quando o vídeo é codificado a taxa de bits variável, diferentes quadros tem diferentes quantidades de dados. Um bloco de armazenamento pode conter diferentes números de quadro. Portanto para implementar avanço e retrocesso rápidos, cada quadro deveria ser indexado individualmente.
- Quando o vídeo é codificado interquadro (explorando a redundância temporal dos quadros), como é o caso do MPEG, a decodificação de alguns quadros depende na disponibilidade de outros quadros. Uma opção para este caso é marcar quadros desejáveis para avanço e retrocesso rápidos e durante o avanço e retrocesso rápidos apenas estes quadros são usados.

### 11.6.3 QoS e interações com o usuário

Uma questão de QoS muito importante e pouco estudada é que quando o usuário altera o modo de apresentação do modo normal para outro modo (pausa, avanço e retrocesso rápidos), os requisitos de QoS para esta sessão alteram.

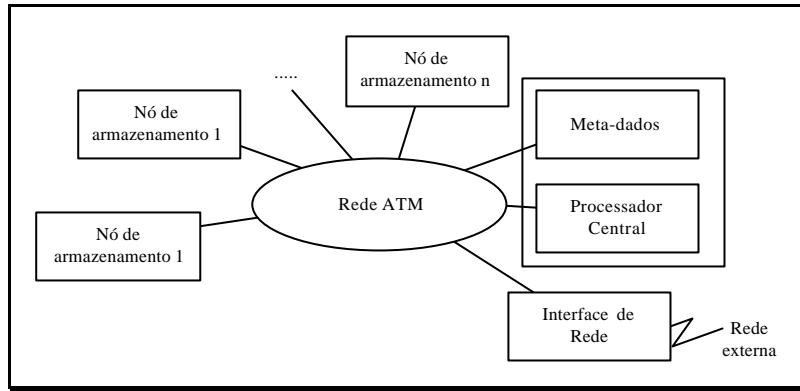
Durante a pausa, a sessão não é ativa por um tempo indeterminado. Se os recursos originalmente reservados para a sessão são chaveados para servir outra sessão, o usuário pode não ter recursos suficientes quando ele emite um comando de retomada e a apresentação não poderá ser retomada. Mas se os recursos são mantidos reservados durante a pausa, o usuário pode pagar por estes recursos mesmo se ele não está assistindo o filme.

Requisitos de QoS são normalmente aumentados durante operações de avanço e retrocesso rápidos, como visto anteriormente. Portanto, quando o usuário emite um comando de avanço ou retrocesso rápido, ele pode não ter recursos suficientes para executar a operação. Isto não é desejável, mas se cada sessão reservar recursos extras prevendo operações rápidas, estes recursos são desperdiçados na maior parte do tempo e o usuário paga mais por isto. Isto também não é interessante. A melhor solução é comprimir o vídeo e colocar os dados compactados de tal modo que nenhum requisito de QoS extra é necessário durante a operação rápida.

## 11.7 Configuração de Servidor e Conexão a Rede

O problema estudado nesta seção é como interconectar dispositivos de armazenamento em um servidor e como conectar o servidor na rede de tal modo que o servidor possa transmitir tantos fluxos quanto possível ao mesmo tempo.

Existem várias propostas para interconectar dispositivos de armazenamento e conectar estes dispositivos na rede. O mais comum é usar uma pequena rede ATM para realizar a conexão entre o processador e os dispositivos de armazenamento. A Figura 111 mostra a arquitetura comum de um servidor multimídia. Há um processador central que tem acesso a meta-dados de fluxos armazenados nos outros dispositivos de armazenamento. Meta-dados contém informações sobre os dados multimídia armazenados, tal como título, localização, relacionamento entre mídias e palavras-chave. Dispositivos de armazenamento são agrupados em nós de armazenamento. Estes nós podem ser independentes ou relacionados pelo uso de *data stripping* (seção 11.4.4). Cada nó é capaz de enviar dados diretamente à rede. Dependendo das necessidade de largura de banda, mais que uma ligação de acesso a rede pode ser necessário para um servidor. A função do processador central é receber pedidos dos clientes, realizar testes de admissão, e configurar conexões para os pedidos aceitos.



**Figura 111.** Arquitetura de um servidor multimídia [Lu, 96]

A arquitetura de servidor acima é escalável para algumas extensões. Mais dispositivos de armazenamento podem ser adicionados quando necessário. Mas a interface de rede pode se tornar o gargalo do desempenho. Uma solução parcial é usar várias interfaces de rede e ligações de acesso à rede.

Mais recentemente, uma arquitetura de servidor chamada SCAM (SCAlable Multimedia Server) foi proposta [Lougher, 96]. Em SCAM, servidores são distribuídos em uma WAN e são logicamente organizados em vários níveis, como discutido na seção 11.3. Arquivos multimídia são replicados ou distribuídos nestes servidores usando um mecanismo de faixas (seção 11.3.5). Durante a operação estes servidores cooperam para servir muitos usuários simultaneamente. Esta arquitetura hierárquica tem duas vantagens. Primeiro, a carga de tráfego é balanceada entre os vários servidores de modo que um servidor individual não se torne o gargalo de desempenho. Segundo, a arquitetura é inteiramente escalável em termos do número de usuários e em termos do número de arquivos multimídia.

## Capítulo 12 Sincronização Multimídia

O objetivo principal das aplicações multimídia é apresentar informações multimídia ao usuário de forma satisfatória, sendo que estas informações podem ser oriundas de fontes ao vivo, como câmeras de vídeo e microfones, ou originária de servidores distribuídos. Uma das principais problemáticas de sistemas multimídia é a **sincronização multimídia**, especialmente em sistemas distribuídos. Neste contexto, sincronização pode ser definida como o aparecimento (apresentação) temporal correto e desejado dos componentes multimídia de uma aplicação, e um esquema de sincronização define os mecanismos usados para obter a sincronização requerida.

O aparecimento temporal correto e desejado de componentes multimídia em uma aplicação tem três significados quando usado em diferentes situações [Lu, 96]:

- Quando usado para um fluxo contínuo, o aparecimento temporal correto e desejado significa que as amostras de um áudio e quadros de um vídeo devem ser apresentados em intervalos fixos. Como já introduzido no capítulo 6, este tipo de sincronização é chamado de **sincronização intramídia**. Como nós vimos no capítulo anterior, esta sincronização está relacionada com a sincronização entre a taxa de consumo do dado no receptor e a taxa na qual o dado é gerado no transmissor.
- Quando usado para descrever os relacionamentos temporais entre componentes multimídia, o aparecimento temporal correto e desejado significa que os relacionamentos temporais desejados entre os componentes devem ser mantidos. Como já introduzido no capítulo 6, este tipo de sincronização é chamada de **sincronização intermídia**, que é relacionada com a manutenção das relações temporais entre componentes envolvidos em uma aplicação.
- Quando usado em aplicações interativas, o aparecimento temporal correto e desejado significa que a resposta correta deveria ser fornecida em um tempo relativamente curto para obter um boa interação. Este tipo de sincronização é chamada de **sincronização de interação**, que é relacionada com a manutenção de que o correto evento (resposta) ocorra em um tempo relativamente curto.

Existem basicamente duas categorias de trabalhos em sincronização multimídia: trabalhos na área de especificação das relações temporais entre componentes; e trabalhos que buscam a definição de mecanismos para satisfazer as relações temporais especificadas. A primeira categoria, sincronização multimídia, já foi discutida na seção 6.5 desta apostila, onde foram apresentados os paradigmas básicos suportados pelos sistemas de autoria multimídia (linguagem scripting, linha temporal, redes de Petri, etc.). Este capítulo trata unicamente da segunda categoria, ou seja, nós veremos alguns dos mecanismos necessários à satisfação das relações temporais especificadas.

### 12.1 Requisitos de Sincronização

Antes de vermos os mecanismos necessários para o fornecimento da sincronização multimídia, é necessário conhecer os requisitos de sincronização. Esta seção descreve as diferentes aplicações multimídia e seus requisitos de sincronização intramídia e intermídia, as medidas do nível de sincronismo e finalmente as causas da perda de sincronização multimídia.

#### 12.1.1 Classes de aplicação e a requisitos de sincronização

Diferentes classes de aplicação multimídia têm diferentes requisitos de sincronização. Como apresentado no capítulo, [Fluckiger, 95] classifica as aplicações multimídia em rede em duas grandes classes: aplicações pessoa-a-sistema (ou baseadas em servidores), onde indivíduos ou grupo de pessoas comunicam com sistemas remotos para acessar, receber, ou interagir com informações multimídia; e aplicações pessoa-a-

pessoa (ou conversacionais): onde o objetivo principal é aumentar a comunicação entre humanos. As aplicações pessoa-a-pessoa ainda podem ser subdivididas aplicações tempo-real (como videofonia, teleconferência) e aplicações assíncronas (como E-mail multimídia). As aplicações pessoa-a-sistema podem ser subdivididas em aplicações interativas (como vídeo sob-demanda) e de distribuição (como difusão de áudio e vídeo).

Em aplicações assíncronas, como E-mail multimídia, e em sistemas multimídia *standalone*s, a especificação da sincronização multimídia é importante, mas na maior parte das vezes elas não têm que tratar com problemas tal como variações de atraso da rede e diferenças temporais entre sistemas (causas da perda de sincronização). Na autoria deste tipo de aplicação, o autor deve dispor de um modelo de autoria para especificar os componentes de uma aplicação e suas relações lógicas e temporais. Este modelo deve garantir a consistência lógica e temporal da aplicação e o sistema deve prover os recursos necessários à apresentação da aplicação, respeitando a Qualidade de Serviço (QoS) requerida.

A fim de analisar os requisitos de sincronização das demais classes de aplicações multimídia, é interessante ainda classificá-las de acordo com a topologia das fontes e destinos das informações:

- Configuração um-a-um: um ou mais fluxos de dado são transmitidos de um fonte para um destino.
- Configuração um-para-muitos: um ou mais fluxos de dado são transmitidos de um fonte para múltiplos destinos.
- Configuração muitos-para-um: fluxos de dado são transmitidos de muitas fontes para um destino.
- Configuração muitos-para-muitos: visto como a combinação das configurações um-para-muitos e muitos-para-um.

Em todas estas configurações a informação deveria ser apresentada para os usuários nos destinos de maneira síncrona. Embora a sincronização entre apresentações em diferentes destinos não é normalmente necessária. Como veremos mais adiante, estas diferentes topologias de comunicação tem diferentes requisitos de sincronização.

Existem três diferenças principais nos requisitos de sincronização entre aplicações conversacionais tempo-real e baseadas em servidores [Lu, 96]:

- Em aplicações conversacionais, dados são transmitidos na medida em que eles são gerados das fontes ao vivo, tal como câmeras de vídeo e microfones. Não há muito controle requerido no transmissor. Ao contrário, em aplicações baseadas em servidores, as relações temporais entre componentes multimídia têm que ser especificadas e mecanismos de escalonamento são necessários para garantir que os dados sejam acessados e retransmitidos sincronamente.
- Requisitos de atraso fim-a-fim são mais duros em aplicações conversacionais que em aplicações baseadas em servidores (devido interatividade entre os participantes da conversação). Isto significa, como já apresentado no capítulo anterior, que a bufferização excessiva não poderia ser usada em aplicações conversacionais.
- É muito raro que componentes multimídia originários de duas ou mais localizações diferentes necessitam ser sincronizadas em aplicações conversacionais. Mas em aplicações baseadas em servidores, isto é um requisito muito comum para acessar informações de mais que um servidor e apresentá-las de maneira sincronizada.

### 12.1.2 Medidas do nível de sincronismo

De acordo com a definição de sincronização como o aparecimento temporal correto e desejado, a sincronização pode ser medida pelos seguintes parâmetros:

- **Atraso**, mede o atraso fim-a-fim nas aplicações conversacionais e o tempo de resposta nas aplicações baseada em servidores. O atraso deve ser pequeno nas duas situações, mas ele é mais crítico em aplicações conversacionais tempo-real.

- **Variação de atraso**, mede o desvio do tempo de apresentação de amostras de mídias contínuas do seu tempo de apresentação natural (e fixo).
- **Distorção intermídia**, mede a diferença entre o tempo efetivo da apresentação de um componente e o tempo ideal definido na relação temporal especificada. O valor aceitável para a distorção intermídia é dos tipos de mídia relacionadas.
- **Taxa de erro tolerável**, mede a taxa de erro de bits e a taxa de erro de pacotes permitidas para um fluxo particular em uma aplicação. Como já apresentado nesta apostila, para mídias contínuas a transferência livre de erro não é essencial para obter uma qualidade de comunicação aceitável.

Diferentes aplicações multimídia necessitam diferentes tipos de sincronização com diferentes precisões. Na concepção de sistemas multimídia é importante reconhecer os requisitos destas aplicações de maneira que estes requisitos possam ser satisfeitos e recursos do sistema possam ser usados de maneira eficiente. Os requisitos de sincronização são uma parte importante dos requisitos de QoS das aplicações e deveriam ser especificados implicitamente ou explicitamente na partida da aplicação.

Atrasos e variações de atraso já foram vistos no capítulo anterior. A seguir estudaremos a distorção intermídia e taxa de erro tolerável de diferentes aplicações.

### **Distorção Intermídia**

A composição temporal de componentes compostos por diferentes tipos de mídia tem diferentes tipos de requisitos de distorção. [Steinmetz, 92] apresenta algumas medidas de tolerâncias de distorção intermídia e seus resultados são mostradas na tabela abaixo. Como o desvio intermídia tolerável é subjetivo e dependente de aplicação, diferentes experimentos podem resultar em diferentes valores. Portanto, os valores apresentados na tabela abaixo servem apenas como indicação dos requisitos de sincronização intermídia. Por exemplo, de acordo com [Steinmetz, 92] a distorção na sincronização labial de estar em torno de 80ms (do áudio com relação ao vídeo e vice-versa), já [Bulterman, 91] indica que esta distorção deve estar entre 10 a 100ms e [Leydekkers, 91] relata que a percepção humana requer que o desvio entre vídeo e áudio esteja entre 20 a 40ms. O requisito de desvio para voz e legenda e para voz e imagem são similares, eles deveriam estar entre 0.1 a 1s [Bulterman, 91]. Comentário de áudio e teleponteiros deveria ser sincronizados dentro de 0.5s [Rothermel, 92].

Mídias envolvidas	Modo ou Aplicação	Distorção intermídia permitida
Vídeo e animação	correlacionados	+/- 120ms
Vídeo e áudio	sincronização labial	+/- 80ms
Vídeo e imagem	superposição	+/- 240ms
Vídeo e imagem	sem superposição	+/- 500ms
Vídeo e texto	superposição	+/- 240ms
Vídeo e texto	sem superposição	+/- 500ms
Áudio e animação	correlacionados	+/- 80ms
Áudio e áudio	estritamente relacionados (estéreo)	+/- 11μs
Áudio e áudio	fracamente relacionados	+/- 120ms
Áudio e áudio	fracamente relacionados (música de fundo)	+/- 500ms
Áudio e imagem	fortemente relacionados (música com notas)	+/- 5ms
Áudio e imagem	fracamente relacionadas (apresentação de slides)	+/- 500ms
Áudio e texto	anotação de texto	+/- 240ms
Áudio e ponteiro	áudio relacionado para mostrar item	- 500ms a 750ms

### **Taxa de Erro Tolerável**

Como já visto, informações multimídia toleram certa quantidade de erros. Existem dois tipos de erros: *erros de bit* e a *perda de pacotes*. Eles têm diferentes efeitos na qualidade de apresentação. Por exemplo, certo erro de bit pode não afetar muito a qualidade de uma imagem, mas a perda de pacote pode afetar a apresentação dos pacotes subsequentes. Na prática, taxa tolerável de erro de bit ou perda de pacote é altamente dependente do método de compressão usado para compressão de áudio, vídeo e imagem. Para voz descompactada, a taxa de erro de bit deveria ser menor que 0.1

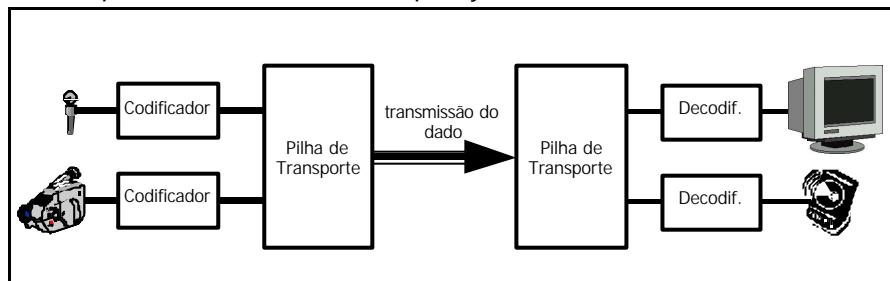
[Hehmann, 90]. Para vídeo de qualidade TV não compactado, a taxa de erro de bit deveria ser menor que 0.01. Para imagens, a taxa de erros de bit permitida é próximo a 0.0001. Para áudio e vídeo compactado, os requisitos são geralmente mais duros pois o erro de 1 bit pode afetar a decodificação de uma sequência de dado. Portanto, devido ao uso de muitos tipos de técnicas de esquemas de compressão e dependência do conteúdo do áudio e vídeo, é difícil generalizar a taxa de erro aceitável para dados multimídia compactados.

### 12.1.3 Causas da perda de sincronização multimídia

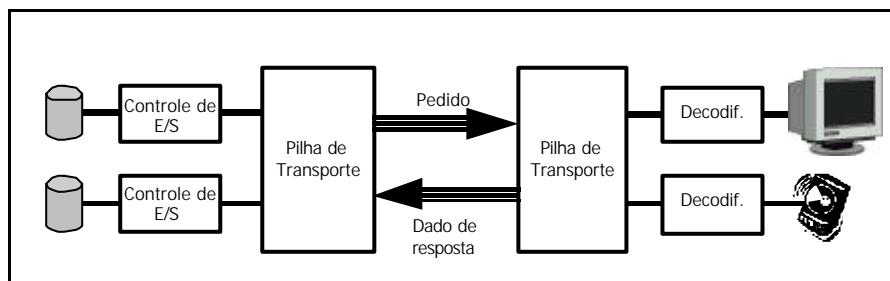
Antes de olharmos os mecanismos de sincronização, nós examinaremos a causa da perda de sincronização e as razões que tornam a sincronização difícil de obter.

As figuras 1 e 2 mostram os processos de comunicação fim-a-fim para aplicações conversacionais e baseadas em servidores:

- Nas aplicações conversacionais, os dados de áudio e vídeo são capturados pelo microfone e câmera de vídeo. Estes dados são compactados antes de serem enviados para a pilha de transporte para empacotamento, processamento do protocolo e colocação na rede para transmissão. No lado do receptor, estes dados são passados para a pilha de protocolos para desempacotamento antes deles serem decodificados para apresentação.
- Nas apresentações baseadas em servidores, o receptor (cliente) envia o pedido de informação para as respectivas fontes da informação (servidores). A informação pedida é acessada do armazenamento secundário sobre o controle de um controlador de E/S. Esta informação normalmente está na forma compactada. O resto do processo é o mesmo das aplicações ao vivo.



**Figura 112.** Processo de comunicação fim-a-fim para aplicação ao vivo



**Figura 113.** Processo de comunicação fim-a-fim para aplicação baseada em servidor

No lado das fontes de informação nós temos as seguintes considerações com relação a sincronização:

- No caso das aplicações conversacionais (Figura 112), o áudio e o vídeo são capturados diretamente via microfone e câmera de vídeo. Idealmente, cada amostragem de áudio e quadro de vídeo gerados ao mesmo tempo deveriam ser reproduzidos no mesmo tempo no receptor. A ocorrência desta situação ideal é dificultada por vários fatores que causam a distorção intermídia. Um destes fatores no lado da fonte é o tempo de compressão: o processo de compressão de áudios e vídeos digitais toma diferentes intervalos de tempo (normalmente a compressão de vídeo leva mais tempo). Além disso, quando a codificação de áudio e vídeo é feita via software, o tempo de decodificação para um fluxo particular pode variar de tempos a tempos, causando também variação de atrasos.

- Em aplicações baseadas em servidores (Figura 113), os pacotes de pedido de informação podem sofrer diferentes atrasos no caminho entre o cliente e os servidores, especialmente quando os servidores estão em localizações diferentes. Isto causa dificuldade na coordenação do tempo de transmissão para diferentes fluxos de mídias, pois o pacote de pedido de dado pode sofrer variações de atraso consideráveis, afetando o tempo de acesso ao dado e tempo de transmissão [Lu, 96]. Usualmente, os controladores de E/S e armazenamento secundário são compartilhados entre vários processos. Alguns mecanismos de escalonamento são necessário afim de habilitar o acesso requerido em intervalos fixos.

O próximo passo é a pilha de transporte, incluindo empacotamento e processamento do protocolo. O tempo de processamento para estes dois são variados, causando variações de atrasos e distorção intermídia.

O próximo elemento a ser considerado é a rede de transmissão. Este estágio implica em tempo de acesso à rede, tempo de transmissão do pacote e tempo de bufferização nos comutadores e gateways. Todos estes tempos variam de pacote a pacote e de tempo a tempo, causando variação de atraso entre um fluxo e distorção entre fluxos.

No lado do receptor (cliente), o tempo de desempacotamento e tempo de processamento do protocolo variam para pacotes diferentes. Além disso, o tempo de decodificação para diferentes mídias são diferentes. Se um decodificador em software for usado, o tempo de decodificação varia de tempo a tempo mesmo para o mesmo fluxo. Tudo isto causa variação de atrasos dentro de um fluxo e a distorção entre fluxos.

Além das apresentadas acima, existem outras causas possíveis para perda de sincronização:

- As taxas de relógio no transmissor e no receptor podem ser diferentes. Se a taxa de relógio é mais rápida no transmissor que no receptor, o transmissor envia mais dados que o consumidor consome, causando sobrecarga de dados no receptor. Por outro lado, se a taxa de relógio é mais lento no transmissor que no receptor, o receptor terá o problema de falta de dado. Estes dois casos não são desejáveis, porque eles afetarão a qualidade das mídias contínuas.
- Quando protocolos de transporte sem conexão são usados, os pacotes compõe um fluxo podem chegar no receptor fora de ordem.
- A existência de fontes múltiplas dificulta a coordenação dos tempos de transmissão do fluxo nas diferentes localizações, causando problemas de sincronização intermídia nos receptor(es).
- A sincronização é complicada pelas interações com os usuários. É importante manter a sincronização durante e após estas interações. Por exemplo, quando uma apresentação é retomada após uma pausa de um certo tempo, a apresentação deveria ser retornada sincronamente.

## 12.2 Mecanismos para Sincronização Multimídia

### 12.2.1 Medidas de contenção de variações de atraso na rede

A variação de atrasos na rede, se não uniformizadas (por exemplo, via bufferização), afeta a sincronização intramídia e intermídia. A variação de atrasos é uma das características intrínsecas das redes a comutação de pacotes. Medidas para conter variação de atrasos podem ser divididas em dois grupos: *corretiva* e *preventiva*.

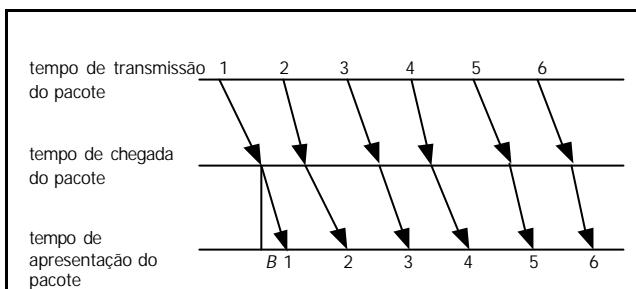
#### **Medidas corretivas**

Medidas corretivas tentam uniformizar o atraso no destino antes do dado ser apresentado ao usuário. Como já apresentado no capítulo anterior, esta uniformização é obtida via utilização de buffers de uniformização de atraso no receptor. Uma questão crítica nesta técnica é quanto tempo cada pacote deveria ser bufferizado para obter uma certa qualidade de apresentação.

Em teoria, caso o tempo de bufferização (chamado atraso intencional) adotado é grande suficiente, todo o atraso é uniformizado. Mas este tempo não pode ser muito grande, pois ele contribui para o atraso fim-a-fim que não pode ser muito grande principalmente em aplicações conversacionais. Além disso, um grande tempo de bufferização implica na necessidade de um imenso tamanho de buffer.

A seguir, nós discutiremos três esquemas de bufferização, que são classificados de acordo como o tempo de bufferização é determinado [Barberis, 80]: dispositivo *null timing information* (NTI), dispositivo *incomplete timing information* (ITI) e dispositivo *complete timing information* (CTI).

No esquema NTI, o primeiro pacote do fluxo é bufferizado por um período de tempo de  $B$  segundos antes de ser apresentado (ilustrado no diagrama temporal da Figura 114). O pacote seguinte é apresentado numa taxa fixa se ele é disponível. Quando a variação de atraso não é muito grande e  $B$  é apropriadamente selecionado, a variação de atraso da rede pode ser removida eficientemente. Mas este esquema não considera o atraso real do pacote. Assim, mesmo se o primeiro pacote sofrer o atraso máximo da rede, ele é atrasado de  $B$  segundos, causando atraso extra desnecessário.



**Figura 114.** Diagrama temporal mostrando o efeito da buferização do esquema NTI

No esquema ITI, o atraso sofrido por cada pacote é estimado baseado na informação de atraso dos pacotes anteriores. Se o atraso estimado é abaixo de um limiar prefixado  $T$ , chamado **tempo de controle**, então este pacote é bufferizado por um tempo igual a  $T$  menos o atraso estimado. Uma questão importante é como estimar o atraso do pacote baseado na informação de atraso. Esta questão foi estudada por [Naylor, 82] no contexto da busca do tempo de bufferização adaptativo e por [Mill, 91] no contexto de obtenção de um relógio de rede síncrono.

No esquema CTI, é assumido que o atraso sofrido por cada pacote é conhecido. Se o atraso sofrido por um pacote é inferior a um limiar prefixado  $T$ , então o pacote é bufferizado por um período de tempo igual a  $T$  menos o atraso do pacote. Neste caso, cada pacote terá o mesmo atraso fim-a-fim (igual a  $T$ ). Se o atraso de rede é limitado,  $T$  é o limite superior do atraso da rede, por outro lado,  $T$  é determinado baseado na estatística de atraso da rede. Para usar o esquema CTI, um relógio de rede síncrono é normalmente necessário para calcular o atraso do pacote baseado na marca temporal taxada ao pacote. [Barberis, 80] apresenta um esquema que permite que o relógio lido do receptor seja próximo ao do lido no transmissor. Um relógio de rede síncrono pode não ser necessário em aplicações baseadas em servidores onde várias fontes de informações (servidores) compartilham um destino comum (cliente). O atraso de rede ida-e-volta pode ser calculado baseado na informação do tempo de pedido no cliente e o tempo de chegada do pacote no cliente. Neste caso,  $T$  é determinado baseado no atraso de rede ida-e-volta estatístico e não no atraso de um caminho simples de um cliente para o servidor.

Quando o caminho de transmissão é longo e muitos nós são usados, a variação de atraso é grande. Assim, um buffer grande deveria ser provido no receptor. Uma alternativa é usar buffers nos nós ao longo do caminho para uniformizar a variação de atraso ocorrida antes do nó. Este mecanismo torna a distribuição dos requisitos de espaço de bufferização mais uniforme sobre uma rota e reduz o tamanho do buffer requerido em cada receptor.

Até aqui a discussão foi focalizada na sincronização intramídia. Quando uma aplicação envolve mais de um fluxo multimídia, para obter a sincronização apesar da existência da

variação de atraso, um circuito virtual multimídia (MVC) [Leung, 90] ou um marcador de sincronização (SM) [Shepherd, 90] podem ser usados.

No MVC, todos os fluxos são multiplexados e compartilham o mesmo canal. Desta maneira, o dado transmitido no mesmo tempo chega no receptor no mesmo instante para apresentação, obtendo assim a sincronização intermídia. Neste esquema, a camada de Sessão invoca a camada de Transporte para criar um circuito virtual (VC) e adicionar, apagar, ou modificar canais no VC. Os requisitos de serviço do canal (classe de serviço, prioridade e características de controle de fluxo/erro) são também passados para a camada de Transporte como parâmetros. A camada de Transporte multiplexa todos os dados multimídia (canais) em um único VC e assegura que os requisitos dos serviços são satisfeitos. A maior vantagem deste esquema é que ele é fácil de implementar. A sua principal fraqueza é que mesmo quando diferentes mídias têm características muito diferentes (e portanto propriedades de QoS diferentes), eles devem ser transmitidos sobre a mesma conexão de rede com um conjunto único de parâmetros de QoS. Esta multiplexação em uma única conexão leva um uso inefficiente dos recursos devido à incapacidade de usar informações específicas dos fluxos. Além disso, este esquema não pode ser estendido para fornecer sincronização entre mídias de diferentes fontes.

No SM, diferentes fluxos usam diferentes conexões e pacotes de dados de diferentes fluxos transmitidos no mesmo tempo são marcados com a mesma marca temporal. No receptor, aquele que chega mais cedo é bufferizado até que todos os pacotes com a mesma marca temporal tenham chego. Este esquema é simples e de fácil implementação, mas ele tem várias desvantagens. Primeiro, pode ser necessário um grande tamanho de buffer no receptor (dependendo das variações de atraso e velocidades de transmissão envolvidas). Segundo, quando diferentes fluxos são oriundos de diferentes servidores, é difícil de inserir SMs sincronamente. Finalmente, este esquema não suporta construções de sincronização que são mais complexas que *sincronizar todos em paralelo*.

Em resumo, a limitação destes dois esquemas é que eles não são aplicáveis no caso onde existam mais que uma fonte de informação. Para isto é necessário uma base temporal comum. Uma solução é ter um relógio de rede sincronizado [Escobar, 92]. Outra solução é se basear no relógio do destino todos os tempos de transmissão de pacote e de apresentação [Lu, 94].

### **Medidas Preventivas: Suporte de Rede e de Transporte para Mídias Contínuas**

Medidas preventivas tentam minimizar atraso e variação de atrasos melhorando o protocolo de transporte e a rede como discutido nos capítulos 9 e 10 desta apostila.

Quando a variação de atrasos é muito grande, medidas corretivas não são muito úteis. Isto pois elas necessitam de um tamanho de buffer imenso e longo tempo de bufferização, que não é apropriado para muitas aplicações. Uma solução geral é usar mecanismos de suporte de protocolo de transporte e rede como apresentado nos dois capítulos anteriores. Com estes mecanismos, atrasos e variações de atrasos são limitados. A variação de atraso dentro de uma fronteira pode ser uniformizada simplesmente usando um pequeno buffer de maneira que o atraso fim-a-fim total de cada pacote seja igual ao máximo atraso admissível.

#### **12.2.2 Medidas de contenção de desvios de processamento**

O desvio do tempo de processamento de dois fluxos pode ser reduzido com a bufferização do fluxo com menor tempo de processamento por um período igual a diferença dos tempos de processamento antes da apresentação.

#### **12.2.3 Medidas de contenção de diferenças de taxa de relógio**

Um relógio sincronizado de rede é necessário pelas seguintes razões:

1. Taxas de consumo e de transmissão de dados são normalmente baseadas nas taxas dos relógios locais. Diferenças de taxa de relógio entre o transmissor e

receptor causa a sobrecarga ou a falta de dados no receptor (se a diferença de taxa de relógio é significante). Estes defeitos dependem do tamanho do fluxo multimídia pois o efeito é acumulativo. Quando o tamanho é relativamente curto, o efeito da taxa de relógio não é significante. Por exemplo, se a diferença de taxa é 50 partes por milhão (ppm), então para uma comunicação de 5 min, o desvio acumulado é de 15ms, que pode não afetar a qualidade da apresentação. De toda maneira, se a duração é de 2 horas, o tamanho de um filme típico, então o desvio acumulado é de 360ms (tempo equivalente a cerca de 10 quadros de vídeo), que causará sobrecarga ou falta de dados dependendo de qual relógio é mais rápido. Uma solução para reduzir esta discrepância é pular ou repetir alguns pacotes.

2. Em aplicações muitos-para-um, se as taxas de relógio das fontes são muito diferentes, o receptor terá problemas para sincronizar os fluxos vindos destas fontes. Uma solução para reduzir este problema é a leitura pela parte do receptor da taxa de relógio do transmissor que emite o fluxo mais sensível ao tempo, tal como áudio. O dado dos outros fluxos são descartados ou repetidos para guardar a sincronização intermédia com o fluxo mais sensível ao tempo. Em [Ramanathan, 93], os autores sugerem a utilização da técnica de realimentação: receptores retornam ao transmissor pacotes leves indicando o instante que o pacote é apresentado. Baseada na informação de realimentação, o transmissor determina se há qualquer assincronismo entre estes receptores. Se existe, algumas operações são tomadas, tal como reduzir a taxa de transmissão de um fluxo.
3. Em uma aplicação muitos-para-um, para sincronizar fluxos de múltiplas fontes, é normalmente necessário coordenar os tempos de transmissão destes fluxos. Esta coordenação requer uma base temporal comum.
4. Marcas temporais são normalmente usadas para medir atrasos de transmissão, como usado nas técnicas de uniformização da variação de atrasos da rede. Para comparar o tempo de transmissão do pacote com o tempo de chegada, uma base temporal comum é necessária.

Para resolver os problemas 1 e 2, nós necessitamos que as taxas de relógio das fontes e destinos sejam a mesma ou muito próximas. Para resolver os problemas 3 e 4, nós necessitamos que as leituras dos relógios e as suas taxas sejam a mesma entre fontes e destinos. O mais conhecido protocolo de sincronização de tempo em rede é o NTP (*Network Time Protocol*) [Mills, 91]. Em NTP, existem dois tipos de servidores de tempo: o servidor de tempo primário e secundário. O tempo nos servidores primários é diretamente sincronizado por fontes de referência externa tal como receptores *timecode*, que recebem informações de tempo padrão de estações de difusão de rádio ou satélite. Os servidores secundários sincronizam com os servidores primários trocando mensagens entre eles. Algoritmos especiais são usados para minimizar o efeito da variação de atraso. É possível hoje sincronizar relógios dentro de alguns milisegundos.

#### 12.2.4 Medidas de contenção dos problemas de perdas de pacote

É mais interessante utilizar serviços orientados à conexão de maneira que os parâmetros de QoS requeridos possam ser negociados antes de iniciar a transmissão. Mas devido à falta de disponibilidade desta espécie de protocolo de transporte, muitas pesquisas e projetos atuais usam serviços sem conexão (tal como UDP). Neste caso, quando serviços sem conexão não confiáveis são usados, pacotes podem chegar fora de ordem ou duplicados. A abordagem mais comum para este problema é anexar a cada pacote um número de seqüência. No receptor, pacotes são reordenados de acordo com seus números de seqüência e pacotes duplicados são ignorados.

#### 12.2.5 Medidas para coordenar múltiplas fontes para transmissão síncrona

Quando aplicações necessitam acessar informações oriundas de múltiplas fontes distribuídas para a apresentação para um usuário, o tempo de acesso e transmissão das informações multimídia devem ser coordenados. Existem duas abordagens:

- A primeira abordagem é baseada em relógio síncrono de rede. Uma estação age como coordenador para controlar os tempos de acesso e transmissão. Este método é utilizado em Sincronização de Fluxo [Escobar, 92] e pelo esquema proposto por [Little, 91].
- A segunda abordagem não requer relógio síncrono de rede. O cliente inicia o acesso e transmissão de diferentes mídias baseado no escalonamento de apresentação [Lu, 94]. Todos os tempos de acesso e transmissão são relativos a um relógio cliente comum. A variação de atraso ida-e-volta é compensado pelo uso do tempo de controle.

### 12.2.6 Suporte da estação de trabalho e servidor para mídias contínuas

Suporte de estação de trabalho considera aspectos da arquitetura de hardware da estação de trabalho, sistema operacional, e gestão de armazenamento. O principal aspecto é especificar QoS ou requisitos de sincronização neste nível e a estação de trabalho ou servidor deveria satisfazer os requisitos.

### 12.2.7 Medidas para fornecer sincronização de interações com o usuário

Em sincronização de interação, os principais interesses são o tempo de resposta, interatividade e manutenção da sincronização após a interação (p.e., como retornar a apresentação síncrona após a pausa).

O tempo de resposta do sistema é a soma dos atrasos de todos os subsistemas. Este atraso total deveria ser especificado na partida e satisfeito por todos os subsistemas.

Interatividade depende dos modelos de especificação de sincronização. Para fornecer uma boa interatividade, a informação deve ser indexada e organizada de maneira adequada [Rowe, 92] e idealmente a informação deveria ser acessada em suas unidades lógicas [Lougher, 93].

### 12.2.8 Técnicas de apresentação para obter efeitos de apresentação ótimos

Após tomar todas as medidas anteriores, existem chances de que o receptor tenha problemas de sobrecarga, falta de dados ou dificuldades para manter a sincronização intermídia por razões tais como má estimativa do tempo de controle. A solução mais comum para este problema é saltar ou repetir algumas unidades de dado de maneira que a qualidade de apresentação é degradada suavemente. Desde que a percepção humana é mais sensível a variação de áudio, a prática mais comum é apresentar o áudio de maneira mais cadenciada possível e saltar ou repetir quadros de vídeo para guardar a sincronização intermídia.

Outra solução é modificar os requisitos de QoS para trocar o ritmo da transmissão do dado de maneira que a sincronização seja mantida ou realizar uma degradação progressiva da QoS.

## Capítulo 13 Internet e a Multimídia

Este capítulo apresenta um estudo sobre os recursos e viabilidade do uso da rede Internet como suporte de comunicação para aplicações multimídia. Além disso, ele apresenta várias aplicações multimídia disponíveis hoje sob a Internet/MBone.

### 13.1 Introdução à Rede Internet

É pela ARPA (*U.S Defense Department's Advanced Research Projects Agency*) que tudo começou nos anos 60, através de um projeto de interconexão dos computadores das principais instituições de pesquisa, ensino e governamentais dos Estados Unidos. O objetivo da operação era, em caso de ataque nuclear, encontrar um sistema de rede de informação que seja capaz de se auto-configurar caso uma das malhas venha a não funcionar. O sistema foi chamado de ARPAnet (isto é rede da ARPA). Esta rede fornecia apenas serviços básicos de correio eletrônico e transferência de arquivos.

Colocada no domínio público, ela foi pega pelos universitários que viram a ocasião de fazer conferências por meio de correio eletrônico. Desde então, a ARPAnet começou sua expansão, não mais pelo exército americano, mas pelas universidades mundiais, laboratórios de pesquisas e grandes empresas, recebendo desde o fim da década de 80, a denominação Internet (*Inter networking*).

#### 13.1.1 Arquitetura Internet

A arquitetura Internet está organizada em quatro níveis: Físico, Rede, Transporte e Aplicação. A Figura 115 ilustra estes quatro níveis.

##### Nível Físico

Este nível não define um padrão próprio de protocolo, uma vez que o objetivo da Internet é justamente acomodar os diversos tipos de rede existentes. Isto significa que, neste nível, é possível utilizar padrões de redes locais como aqueles definidos no IEEE (IEEE 802.2, 802.3, 802.4, etc.), padrões como o HDLC (norma X.25), ou mesmo protocolos proprietários para redes de longa distância (SDLC, BDLC, etc.).



**Figura 115.** Arquitetura Internet

##### Nível de Rede

Os serviços e protocolos implementados a este nível asseguram o poder de conectividade da Internet, sendo a interconexão de diversas redes a função básica deste nível.

Neste nível foi adotado o protocolo IP (*Internet Protocol*) que implementa um serviço de comunicação sem conexão, baseado em comutação de mensagens. O IP implementa um mecanismo de roteamento das mensagens que permite que um programa de

aplicação troque informações com outro, mesmo que eles estejam executando em estações conectadas a redes completamente diferentes.

### Nível de Transporte

Este nível oferece um serviço confiável de transferência de dados fim-a-fim entre aplicações. Os serviços provados a 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, seqüenciamento, etc. Além disso, dada a sua orientação para um conjunto diversificado de aplicações, ele deve dar suporte para o controle de vários canais de comunicação entre as aplicações, simultaneamente.

Os principais protocolos definidos para este nível da Internet são o TCP (*Transmission Control Protocol*) e o UDP (*User Datagram Protocol*). O IP é um protocolo de rede que opera no modo sem conexão, enquanto o TCP é um protocolo de transporte orientado à conexão. Desta forma, 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, que opera no modo sem conexão e possui funcionalidades bem mais simplificadas do que o TCP.

### Nível de Aplicação

Este nível oferece ao usuário o acesso à Internet, implementando um conjunto de protocolos e serviços padronizados de comunicação para as tarefas mais frequentemente realizadas na rede: o correio eletrônico (protocolo SMTP - *Simple Mail Transfer Protocol*), a conexão remota (TELNET) e a transferência de arquivo (o protocolo FTP – *File Transfert Protocol*), entre outros.

#### 13.1.2 Protocolo IP

Existem várias versões do protocolo IP definidos. A versão utilizada na Internet atual é a IPv4, que será abreviada IP nesta apostila. IPv6 (visto mais adiante) é uma evolução do IPv4, chamada IPng (*Internet Protocol Next Generation*). A Internet baseada IPv4 deve migrar para uma nova Internet principalmente baseada no IPv6.

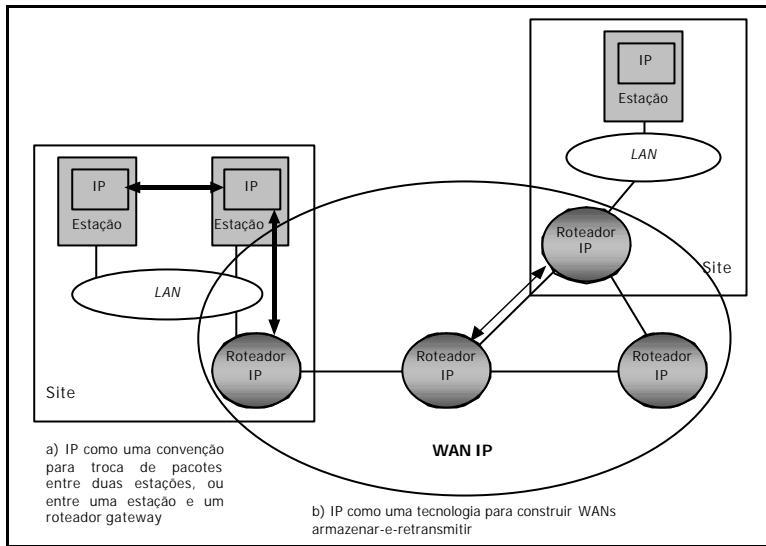
Existem dois modos de emprego do protocolo IP (Figura 116) [Fluckiger, 95]:

- Como um mecanismo de pacote armazenar-e-retransmitir, usado para criar redes IP formadas por nós IP, chamados **roteadores**, e de ligações de vários tipos entre estes nós (Figura 116b).
- Como um formato usado pelos computadores para estruturar seus fluxos de informação em blocos, os **pacotes IP**, que são transferidos sobre qualquer tipo de rede e não necessariamente sobre uma malha de roteadores IP (Figura 116a).

Para o primeiro caso (Figura 116a), uma rede IP é constituída resumidamente por sub-redes conectadas através de elementos denominados *roteadores*. Os *roteadores*, através do protocolo IP, são responsáveis pelo encaminhamento dos blocos de dados, denominado *datagramas*, do host de origem ao host de destino, que são identificados por endereços IP. O protocolo também fornece o serviço de fragmentação e remontagem de datagramas longos que precisam ser separados em pacotes de tamanho menor pela limitação do tamanho do pacote imposta por algumas redes por onde o datagrama passa até chegar ao destino. Nesta camada é realizado também o mapeamento de endereços IP em endereços físicos, e vice-versa.

Abaixo são resumidas as principais características das redes IP:

- IP é um protocolo sem conexão. Isto significa que a rede não tem conhecimento das comunicações entre computadores. A rede apenas vê e transporta pacotes independentes: os *datagramas*. Não existem mecanismos de controle de admissão usados pela rede para regular o fluxo de dados emitidos. Isto produz variações de atraso e perda de pacotes.



**Figura 116.** IP para comunicação entre estações sobre LANs e WANs

- A rede resultante fornece um serviço de transmissão de pacotes do tipo melhor esforço. Nenhum recurso explícito é reservado para uma dada comunicação entre computadores. O IP transmite aquilo que o serviço de enlace pode fornecer (menos uma pequena fração de sobrecarga). Como resultado, os pacotes são liberados com atrasos de trânsito variáveis.
- No caso de sobrecarga, a rede pode descartar pacotes, sendo que as estações (sistemas finais) não são notificados dos pacotes descartados. Geralmente a perda de pacotes ocorre nas filas dos roteadores IP.
- Os roteadores, através do protocolo IP, são responsáveis pelo encaminhamento dos datagramas do *host* (estação conectada a rede) de origem ao *host* de destino, que são identificados por endereços IP. Existem várias protocolos de roteamento nas redes IP; que são mecanismos para decidir qual é a rota mais adequada para tomar quando um pacote é submetido por uma fonte para o transporte até um destino remoto. Alguns protocolos são estáticos, no sentido que as rotas apenas mudam no caso de falhas de um componente ao longo do caminho. Assim, na ausência de falhas, pacotes são liberados em seqüência. Outros protocolos tentam balancear a carga sobre a rede, estimando a carga instantânea de cada rota. Pacotes podem assim serem liberados fora de seqüência.
- O meio normal de interconexão de uma rede local conectando estações (que internamente usam IP) à uma WAN IP é usando um roteador IP como gateway. Um roteador IP gateway tem ao menos duas portas: uma para a rede local e uma para conexões de longa distância. Roteadores normalmente suportam praticamente todas as tecnologias de conexão longa distância (incluindo um protocolo ponto-a-ponto dedicado chamado PPP, X.25, a rede de telefone comutada, N-ISDN, Frame Relay, SMDS, ou ATM).
- IP é provado com mecanismos multicast, embora nem todos os roteadores os implementam.

### 13.1.3 IP Multicast

O conceito de endereçamento com multicast (difusão seletiva) para comunicações de grupo foi inventado por Steve Deering em sua tese de doutorado na *Stanford University*, USA, e depois desenvolvido na Xerox PARC. O projeto e desenvolvimento do MBone foi ajudado por Van Jacobson e seu grupo no *Lawrence Berkeley Labs* (LBL) e Steve Casner no *Information Science Institute* (ISI), além de muitos outros engenheiros [Kumar, 95].

IP multicast é uma extensão do IPv4 que permite que um datagrama IP seja transmitido para um conjunto de máquinas que formam um grupo multicast identificado por um endereço IP único.

A rede diferencia pacotes normais e pacotes multicast olhando no formato do endereço destino contido no pacote. Na versão 4 do padrão, o endereço IP é de 32 bits; se os 4 primeiros tem o valor "1101", isto significa que o pacote é multicast (da classe D), o resto do endereço serve para identificar o grupo multicast.

### Roteadores IP Multicast

O IP multicast pode ser utilizado em uma rede física única ou através da Internet. Neste último caso, os datagramas são transmitidos por roteadores com capacidade multicasting denominados roteadores multicast, que podem ser os mesmos ou diferentes dos roteadores Internet. No caso de uma rede física única, uma estação transmite um datagrama multicast na rede local, que atingirá todos os membros do grupo multicast pertencentes a rede local.

Para operar no modo multicast, os roteadores IP devem ser estendidos para entender endereços de grupo, duplicação de pacotes que chegam e retransmissão dos pacotes às ligações corretas. Para fazer isto de maneira eficiente, apenas as funções de duplicação de pacotes e roteamento são executados em um roteador IP multicast em tempo de execução; o cálculo das rotas multicast ótimas, entrada e saída de membros, etc., são manipulados por um protocolo separado chamado IGMP (*Internet Group Management Protocol*).

### Grupos Multicast

IP multicast é uma extensão do IPv4 que permite que um datagrama IP seja transmitido para um conjunto de máquinas que formam um grupo multicast. Um grupo multicast é identificado por um endereço IP único e são formados por um conjunto de zero ou mais máquinas que podem estar espalhadas ao longo de redes físicas separadas.

Os membros de um grupo são dinâmicos: estações podem entrar ou deixar grupos a qualquer momento. Não há autoridade central para que pedidos têm que ser feitos. Assim não há restrições para o número de membros de um grupo, e uma estação pode participar de mais de um grupo simultaneamente. Pertencer a um grupo determina se o host receberá os datagramas enviados para o grupo multicast, sendo permitido que uma estação envie datagramas para um grupo mesmo sem pertencer a ele.

Um grupo pode ser permanente ou transitório. Grupos permanentes são aqueles que possuem um endereço conhecido e fixo. Apenas o endereço é permanente. Os membros deste grupo não são permanentes, e num dado momento um grupo permanente pode ter qualquer número de membros, inclusive zero. Os endereços IP multicast que não são reservados para nenhum grupo permanente estão disponíveis para atribuição dinâmica de grupos temporários que existem somente enquanto possuem membros. Estes grupos são criados quando necessário, e descartados quando o número de membros atinge zero ou seu tempo de vida terminar.

Existem 3 diferentes tipos de membros de um grupo IP multicast: uma estação da rede (*host*) pode apenas enviar (não pode fazer parte do grupo), apenas receber ou enviar e receber pacotes multicast.

### Criação de um grupo multicast

Antes de um multicast ser iniciado no nível IP, o emissor tem que reservar um endereço de grupo disponível, e os receptores têm que encontrar este endereço. Na Internet, os protocolos SAP (*Session Announcement Protocol*) e SDP (*Session Description Protocol*) foram definidos para estes propósitos. Eles informam aos roteadores multicast a existência do grupo de maneira que o mapeamento do endereço do grupo para as funções de roteamento e duplicação de pacotes possam ser feito.

### Entrada em um grupo multicast

Um dado host desejando se juntar ou deixar um grupo não propaga ele mesmo sua informação de membro pela rede. Ele faz isto indiretamente, via um roteador multicast que propaga a informação do membro para outros roteadores multicast. O mecanismo de entrada no grupo é portanto um procedimento iniciado pelo receptor.

Um hospedeiro pode simultaneamente fazer parte de vários grupos. Isto é particularmente importante para multimídia, como na distribuição de áudio e vídeo baseado em IP, o som é distribuído para um dado grupo - ou para vários grupos com diferentes graus de qualidade - e o vídeo é em um grupo diferente.

### **Características de confiabilidade do IP Multicast**

A entrega de um datagrama multicast é realizada com as mesmas características de confiabilidade que os datagramas *unicast* regulares IP, o que significa que não há garantia contra perda, retardo, ou entrega fora de ordem para nenhum dos membros do grupo.

### **IP Multicast está sendo usado?**

Atualmente, IP multicast é ainda considerado um experimento. Mas desde que IP multicast será completamente integrado em IPv6, os protocolos companheiros serão também inteiramente integrados na próxima geração de hosts e roteadores Internet.

A fim de testar o IP multicast, uma rede dentro da rede Internet, chamada MBone (Multicast Backbone) foi criada. Os nós participantes desta rede são roteadores IP Multicast. A rede MBone será vista em detalhes na seção 13.3.

#### **13.1.4 IPv6**

IP versão 6 (ou IPv6) é uma nova versão do protocolo Internet IP. O IETF decidiu em 1992 desenvolver uma nova versão do IP pois o espaço de endereçamento disponível do IPv4 provavelmente terminaria no início do século 21. Atualmente, IPv4 está em uso por todo o mundo. IPv6 foi projetado para ser um passo evolucionário do IPv4, extensões para fluxos de dados multimídia são a principal razão para esta nova versão, mas não esta não é a única. Outras principais são o aumento do espaço de endereçamento, autenticação e criptografia.

Uma meta importante de projeto do IPv6 é a compatibilidade com IPv4. Em uma rede grande e heterogênea como a Internet, seria impossível a migração de todos os nós para uma nova versão de protocolo no mesmo tempo. Novos hosts e roteadores executando IPv6 serão capazes de coexistirem com velhos hosts IPv4, habilitando assim uma migração gradativa da Internet.

IPv6 é baseado nos principais paradigmas das velhas versões do IP: sem conexão (um protocolo datagrama), sem controle de erro e de fluxo na camada de rede. Os novos objetivos que devem ser alcançados com esta nova versão do protocolo IP são [Tanenbaum, 96] [Kuo, 98]:

- Suporte a bilhões de *hosts*, através da expansão do espaço de endereçamento e uma hierarquia mais versátil (mais níveis). Sendo o espaço de endereçamento ampliado para 128 bits (são 23 bits no IPv4);
- Suporta o *multicast*. O campo *scope* dentro de um endereço multicast limita o domínio de validade deste endereço (por exemplo, para uma Intranet em uma empresa);
- Os novos campos *flow label*, *traffic class* e *flow ID* no cabeçalho permitem a identificação de todos os pacotes de um mesmo fluxo de dados (chamado um fluxo no IP). Um fluxo é uma seqüência de pacotes enviados por um host para um endereço unicast ou multicast. Assim, todos os roteadores no caminho podem identificar os pacotes de um fluxo e tratar eles de um modo específico ao fluxo. O campo *traffic class* é muito similar às classes de tráfego ATM apresentadas no capítulo 10. A classe de tráfego para fluxos contínuos (p.e. vídeo e áudio) terão mais alta prioridade nos roteadores que a classe de tráfego para fluxos de dados tradicionais. O *flow label* é uma das características chaves do IPv6 para reserva de recursos e QoS no nível IP na Internet. Nas velhas versões do IP não foi possível identificar pacotes pertencendo a um fluxo multimídia particular (os endereços fonte e destino não são obviamente suficientes), e assim a reserva de recursos e garantias de QoS eram impossíveis de implementar.

- Redução da tabela de roteamento e melhorias no roteamento, inclusive no que tange a *hosts* móveis;
- Protocolo passível de expansão, através do uso de cabeçalhos de extensão;
- Simplificação do cabeçalho do protocolo, diminuindo o tempo de processamento na análise dos cabeçalhos, por parte de roteadores e *hosts*;
- Garantia de mais segurança (autenticação, integridade e criptografia) em relação à versão atual;
- Permissão de máquinas *wireless* mudarem fisicamente de lugar sem mudança em seus endereços IP;
- Habilitação de máquinas se autoconfigurarem (número IP, servidor de nome...) ao serem ligadas na rede (operação "plug and play");
- Um novo tipo de endereço chamado *anycast*, conceitualmente uma "cruz" entre *unicast* e *multicast*: esse tipo de endereço identifica um conjunto de nós, onde um pacote enviado para um endereço *anycast* será entregue a um destes nós;
- Coexistência das duas versões do protocolo por um bom tempo, pois não se pode determinar uma data específica para que todas as máquinas no mundo troquem seus softwares.

### **IPv6 e a multimídia**

O IPv6, como o IPv4, sozinho fornece apenas serviço melhor esforço para as camadas mais altas. Como já visto nesta apostila, muitas aplicações multimídia necessitam mais que isto; o IETF definiu então novos serviços para complementar os serviços do tipo melhor esforço. Estes fornecem reserva de processamento, memória, roteadores e rede. Para isto, RSVP transfere informações de reserva de recurso entre sistemas IP (roteadores e sistemas finais). Os protocolos de transporte no topo do IP, TCP e UDP, não são suficientes para suportar aplicações tempo-real tal como transferência de dados de áudio e vídeo na Internet. Eles não suportam sincronização dentro de um único fluxo e entre dois ou mais fluxos. Para isto foi definido o protocolo RTP que suporta transferência de dados tempo-real. IPv6, RSVP e RTP fornecem um suporte significante para aplicações multimídia [Braun, 97b].

## **13.2 TCP/IP e a Multimídia**

Como um protocolo da camada de rede, IP não impede o uso eficiente de ligações de comunicação de alta-velocidade [Kuo, 98]. Mas as redes IP oferecem um serviço de transporte de datagrama do tipo melhor esforço. Portanto, as redes IP de hoje não fornece isocronismo. Além disso, a ausência do modo orientado a conexão torna a reserva de recursos pela rede difícil. Portanto, a taxa de bits não pode ser garantida e os atrasos de trânsito podem variar na ordem de segundos. Neste modelo, a mais alta garantia da rede é a transferência confiável usando protocolos como TCP, sendo que a retransmissão TCP causa variação de atraso adicional. Isto é desnecessário, pois o transporte de mensagens confiáveis não são necessárias para áudio e vídeo, pois eles toleram a perdas de quadros (perdas são raramente fatais, embora reduzem a qualidade de imagens e sons). Portanto, o suporte de aplicações com mídias contínuas tempo-real é portanto dificultado nas redes TCP/IP.

UDP evita esta deficiência para multimídia, com um serviço orientado datagrama simples sem confiabilidade. Aplicações podem rodar no topo do UDP com funções adicionais integradas nas aplicações, delegando-se às estações o recobrimento das dificuldades que a rede tem quanto a garantias de serviço. Para suportar aplicações necessitando transmissões tempo-real, técnicas de recobrimento usadas pelas estações incluem técnicas de bufferização, o uso de protocolos de transporte melhores adaptados, como o RTP, e mecanismos de admissão pelo qual as estações tentam estimar antes do lançamento de uma aplicação se esta tem alguma chance de ser satisfatoriamente suportada.

A integração na aplicação de estratégias de controle de fluxo ou retransmissão corresponde ao conceito *Application Level Framing* (ALF) [Clark, 90]. Integrar funções do

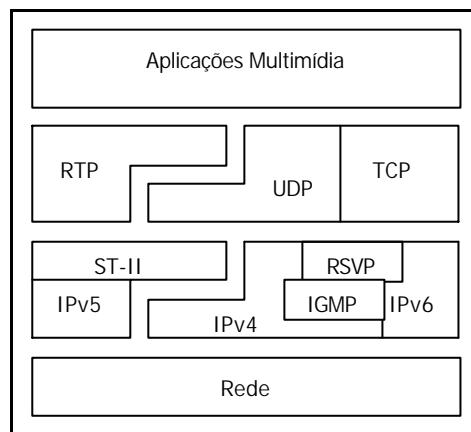
protocolo de transporte tradicional tal como suporte de confiabilidade requer que as aplicações controlem o tamanho do pacote de rede. ALF propõe um tamanho de pacote único para todas as funções da pilha de comunicação, que aumenta o desempenho do sistema e permite a implementação mais eficiente e avançada. Para suportar a interoperabilidade, os protocolos ALF são definidos pela especificação principalmente dos formatos da unidade de dados de protocolo. As abordagens de protocolo ALF não definem algoritmos para controle de fluxo e retransmissão, estes são baseados nos requisitos das aplicações. O Protocolo RTP é um bom exemplo de projeto de protocolo baseado no ALF.

### **Outros Protocolos para suprir necessidades específicas**

No sentido de resolver os problemas do IP, a IETF (*Internet Engineering Task Force*) tem atualmente ampliado seu conjunto de protocolos como ilustrado na Figura 117. Na camada de rede, o IETF definiu duas abordagens distintas: arquiteturas sem e com conexão. As duas abordagens suportam multicast, reserva de recursos, gestão de largura de banda, qualidade de serviço, etc.

- O modelo sem conexão é baseado no IPv4 (presente na Internet atual) ou sua evolução IPv6. Ela é uma camada de rede orientada a pacotes que necessita do protocolo RSVP para gestão de recursos para QoS. IPv4 necessita também de extensão multicast, provida através do protocolo IGMP (*Internet Group Management Protocol*) e protocolos de roteamento multicast. A Internet baseada IPv4 deve migrar para uma nova Internet principalmente baseada no IPv6, que integra o multicast de maneira nativa.
- O modelo com conexão é baseado no IPv5 (*Internet Protocol version 5*), um protocolo de roteamento orientado a circuito que é definido para coexistir com o IP e o protocolo experimental ST-II para gestão de reservas e garantias de QoS.

Na camada de transporte estão os protocolos RTP, UDP e TCP que usam os protocolos de rede.



**Figura 117.** Estrutura de Serviço Multimídia IETF

### **Outros protocolos ALF não padronizados**

Além dos protocolos apresentados na Figura 117, existem outros protocolos baseados no conceito ALF destinados a satisfazer requisitos das aplicações.

O modelo LWS (*Lightweight Session*), baseado no ALF, generaliza o IP multicast para suportar aplicações colaborativas tal como conferências audiovisuais e quadros brancos compartilhados. O bloco de construção LWS inclui IP multicast, recobrimento de temporização via adaptação do receptor, e camadas de transporte leves de acordo com o conceito ALF. Isto adiciona escalabilidade, tolerância a falhas e robustez ao IP multicast.

A abordagem SRM (*Scalable Reliable Multicast*) [Floyd, 95], também baseada no ALF, propõe algoritmos que aumentam a confiabilidade sobre sistemas de comunicação UDP/IP não confiáveis. SRM basicamente estende o conceito LWS para suportar transferência multicast confiável de dados. Embora ainda não padronizado, o formato de

dado RTP estendido por SRM poderia incluir os elementos de informação requeridos para pedidos de retransmissão e retransmissões.

### **Conclusões gerais**

Na prática, a Internet é a rede multimídia de longa distância mais usada. A pobre qualidade das transmissões de áudio e vídeo não é surpresa, dada a falta geral de largura de banda em grande áreas. Além disso, variações de atraso ocorrem devido as razões apresentadas anteriormente. Todavia, devido a disponibilidade universal da Internet, o IP tem grandes méritos de promover o desenvolvimento de aplicações multimídia distribuídas.

## **13.3 Qualidade de Serviço na Internet**

Como visto nas seções anteriores, a Internet de hoje fornece um serviço do tipo melhor esforço: o tráfego é tratado tão rápido quanto possível, mas não há garantias temporais ou limites de erro. O termo serviço é usado para descrever algo oferecido aos usuários fim da rede, como comunicações fim-a-fim ou aplicações cliente-servidor.

Com a rápida transformação da Internet em uma infra-estrutura comercial, o fornecimento de qualidade de serviço está sendo considerada cada vez mais um requisito essencial. O casamento dos termos qualidade e serviço pode ser visto como a capacidade para diferenciar entre tráfego ou tipos de serviço, de forma que o sistema possa tratar uma ou mais classes de tráfego diferentemente de outros. Neste sentido, seria interessante que um Provedor de Serviços Internet (ISP – *Internet Service Provider*) pudesse fornecer a seus clientes diversas possibilidades em termos de qualidade de serviço, garantido de um certo modo a taxa de bits e atraso. É claro, quando o cliente optasse por um serviço de qualidade, ele pagaria um custo maior do que aquele pago por um tipo de serviço melhor esforço.

Para isto, o cliente de um ISP optará por uma das várias classes de serviço (garantindo diferentes qualidades). Por exemplo, uma classe de serviço forneceria serviços Internet “previsíveis” para companhias que fazem negócios na Web. Tais companhias têm interesse em pagar um certo preço para tornar seus serviços confiáveis e fornecer a seus usuários um acesso rápido a seus sites Web. Esta classe de serviço poderia conter um serviço único ou poderia conter *Serviços Ouro, Prata e Bronze*, que reduz em qualidade. Outra classe de serviço forneceria serviços de pequeno atraso e baixa variação de atraso para aplicações tal como telefonia Internet e videoconferência. Companhias despejarão pagar um preço para executar uma videoconferência de alta qualidade para reduzir custos e tempo de trabalho. Finalmente, o Serviço Melhor Esforço permanecerá para clientes que requerem apenas conectividade.

### **Qualidade de Serviço e o aumento da largura de banda**

A necessidade de fornecer mecanismos na rede para garantir a qualidade de serviço é um ponto muito debatido. Um grupo de pesquisadores considera que com o aumento da largura de banda (por exemplo, usando fibras óticas), a qualidade de serviço será automaticamente satisfeita. Mas outro grupo afirma que mesmo com o aumento da largura de banda, novas aplicações serão inventadas para fazer uso desta largura de banda. Portanto, mecanismos serão necessários para fornecer garantias de QoS.

### **Trabalhos da IETF relacionados com garantias de QoS**

A *Internet Engineering Task Force* (IETF) tem proposto vários modelos de serviço e mecanismos para satisfazer a necessidade de QoS na Internet, proporcionando um melhor controle sobre o tráfego na Internet, na forma de priorização de certas aplicações (com certas restrições temporais) em detrimento do restante (tráfego essencialmente melhor esforço). Entre estes trabalhos estão o modelo Serviços Integrados/RSPV [Branden, 94] e o modelo Serviços Diferenciados [Black, 98]. Eles são descritos nas duas próximas seções.

### 13.3.1 Serviços Integrados/RSVP

Os serviços Integrados (*Integrated Services ou IntServ*) [Branden, 94] foram projetados para prover um conjunto de extensões ao modelo de entrega de tráfego de melhor esforço atualmente utilizado na Internet. Em essência, eles foram projetados para dar tratamento especial para certos tipos de tráfego e prover um mecanismo para que as aplicações possam escolher entre múltiplos níveis de serviços de entrega para seu tráfego.

IntServ é baseada na reserva de recursos. Para aplicações tempo real, antes dos dados serem transmitidos, as aplicações devem primeiro configurar caminhos e reservar recursos. RSVP (visto anteriormente) é um protocolo de sinalização para configurar os caminhos e reservar recursos.

O modelo de Serviços Integrados propõe duas classes de serviço em adição ao Serviço Melhor Esforço:

- **Serviço Garantido (Guaranteed Service)** [RFC 2212]: fornece limites firmes (matematicamente prováveis) em termos de atrasos de enfileiramento que os pacotes sofrerão nos roteadores. Ele garante tanto o atraso quanto a taxa de bits. Basicamente uma sessão requisitando Serviço Garantido está requerendo que os bits em seus pacotes tenham uma taxa de bits garantida. Note que este serviço não tenta minimizar a variação de atraso, ele controla o atraso máximo de enfileiramento. Para este tipo de serviço, todos os nós intermediários devem implementar os serviços garantidos. Este serviço pode ser útil para aplicações requerendo fronteiras de atraso fixa, tal como aplicações tempo-real e aplicações de áudio e vídeo tempo-real.
- **Serviço de Carga Controlada (Controlled Load Service)** [RFC 2211]: uma sessão requerendo tal serviço receberá uma qualidade de serviço muito próxima da qualidade que um fluxo poderia receber de uma rede não sobrecarregada. Em outras palavras, a sessão pode assumir que uma “percentagem muito alta” de seus pacotes passará com sucesso através do roteador sem ser cortados e com um atraso de enfileiramento muito próximo a zero. Note que o Serviço de Carga Controlada não fornece garantias quantitativas acerca do desempenho – ele não especifica o que constitui uma “percentagem muito alta” de pacotes nem que qualidade de serviço aproximada será fornecida por um elemento de rede não sobrecarregado. Este tipo de serviço é dirigido para aplicações tempo-real adaptativas que estão sendo desenvolvidas hoje na Internet. Estas aplicações executam razoavelmente bem quando a rede não é sobrecarregada, mas elas se degradam rapidamente quando a rede se torna congestionada.

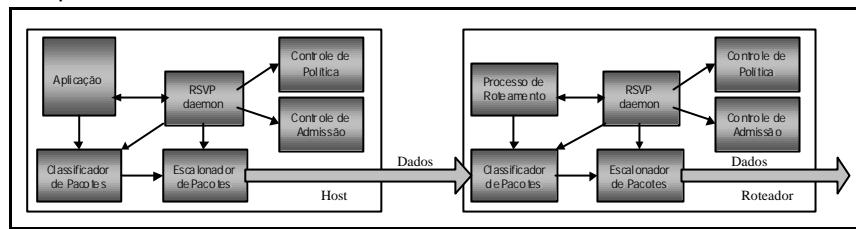
No modelo IntServ, os roteadores devem ser capazes de reservar recursos afim de fornecerem QoS especial para fluxos de pacotes específicos do usuário. Neste caso, o estado específico dos fluxos devem ser mantidos pelos roteadores. Este modelo é implementado por quatro componentes (Figura 118):

- **Protocolo de Sinalização:** aplicações necessitando de Serviço Garantido ou Serviço de Carga Controlada devem configurar um caminhos e reservar recursos antes da transmissão de seus dados. Para isto elas devem usar um protocolo de sinalização (p.e. RSVP).
- **Rotina de Controle de Admissão:** decidem se um pedido de alocação de recursos pode ser garantido. Ela implementa o algoritmo de decisão que um roteador ou host usa para determinar se um novo fluxo pode ter sua QoS garantida sem afetar fluxos anteriormente garantidos.
- **Classificador:** quando um roteador recebe um pacote, o classificador executará uma classificação Multi-Campo (MF – Multi-Field) e colocará o pacote em uma fila específica baseado no resultado da classificação. A classificação MF é o processo de classificar pacotes baseado no conteúdo dos seus campos tal como endereço fonte, endereço destino, byte TOS<sup>25</sup> (*Type of Service*) e identificador do

<sup>25</sup> Campo de 8 bits do cabeçalho do datagrama IP: 3 bits mais significativos indicam a natureza e prioridade do datagrama; TOS (4 bits) especifica o tipo de serviço: minimiza atraso, maximiza vazão, maximiza confiabilidade, minimiza custos monetários, serviço normal; e o bit menos significativo é reservado. Geralmente este campo é ignorado pelos roteadores.

protocolo<sup>26</sup>. Cada pacote que chega deve ser mapeado em alguma classe; todos os pacotes na mesma classe obtém o mesmo tratamento do escalonador. Uma classe pode corresponder a uma grande categoria de fluxos, por exemplo, todos os fluxos de vídeo ou todos os fluxos atribuível a uma organização particular. Por outro lado, uma classe pode manter apenas um único fluxo. Uma classe é uma abstração que pode ser local a um roteador particular; o mesmo pacote pode ser classificado diferentemente por diferentes roteadores ao longo do caminho. Por exemplo, roteadores backbone podem escolher o mapeamento de muitos fluxos em poucas classes agragadas, enquanto roteadores periféricos, onde existem menos agregação, podem usar uma classe separada para cada fluxo.

- **Escalonador de Pacotes:** após a classificação, o escalonador de pacotes escalona o pacote de modo a satisfazer os requisitos de QoS. O escalonador de pacotes gerencia a retransmissão dos diferentes pacotes usando um conjunto de filas e possivelmente outros mecanismos tal como timers.



**Figura 118.** RSVP nos hosts e Roteadores

Na Internet de hoje, a retransmissão IP é completamente igualitária: todos os pacotes recebem a mesma qualidade de serviço, e os pacotes são retransmitidos usando uma fila FIFO. Para IntServ, um roteador deve implementar uma QoS apropriada para cada fluxo, de acordo com o modelo de serviço. A função do roteador que cria diferentes qualidades de serviço é chamada de controle de tráfego. Controle de tráfego é implementado pelos componentes escalonador de pacote, classificador e controle de admissão, visto acima.

A arquitetura Serviços Integrados/RSVP representa uma mudança fundamental na arquitetura atual da Internet, que é fundada no conceito que todas as informações de estado relacionadas aos fluxos deveriam estar nos sistemas finais. Neste sentido, existem alguns problemas com a arquitetura Serviços Integrados [Xiao, 99]:

- O montante de informações de estado aumenta proporcionalmente ao número de fluxos. Isto causa uma sobrecarga de armazenamento e processamento nos roteadores. Portanto esta arquitetura não é escalável.
- Os requisitos nos roteadores são altos: todos os roteadores devem implementar RSVP, controle de admissão, classificação MF e escalonamento de pacotes.
- Para Serviço Garantido, toda a rede deve suportar IntServ. Uma instalação gradativa de Serviço de Carga Controlada é possível pelo emprego de funcionalidades RSVP e Serviço de Carga Controlada nos nós gargalos de um domínio e tunelando as mensagens RSVP para outras partes do domínio.
- IntServ/RSVP não são muito aplicáveis a aplicações do tipo navegadores WWW, onde a duração de um fluxo típico é apenas de poucos pacotes. A sobrecarga causada pela sinalização RSVP poderia facilmente deteriorar o desempenho da rede percebida pela aplicação.

### 13.3.2 Serviços Diferenciados

Devido as dificuldades de implementar e utilizar Serviços Integrados/RSVP, os Serviços Diferenciados (DS - *Differentiated Services ou DiffServ*) [Black, 98] foram introduzidos. Eles têm sido escolhido para ser implementado na Internet2. Neste modelo, os pacotes são marcados diferentemente para criar várias classes de pacotes. Pacotes de classes diferentes recebem diferentes serviços. O campo TOS (*Type Of Service*) do cabeçalho do pacote IPv4 e o campo *Class* do cabeçalho do pacote IPv6 podem ser setados pelas

<sup>26</sup> Especifica qual protocolo de alto nível foi usado para criar a mensagem que está sendo transportada na área de dados do datagrama.

aplicações para indicar a necessidade de serviço de pequeno atraso e alta vazão ou baixa taxa de perdas. Mas as escolhas são limitadas.

### **DS é um esquema de prioridades**

A meta do DiffServ é definir métodos relativamente simples (comparados a IntServ) para prover classes diferenciadas de serviço para o tráfego na Internet. A ideologia é não pregar uma nova e completa arquitetura onipresente, mas produzir um pequeno e bem definido conjunto de blocos de construção dos quais uma variedade de serviços podem ser construídos. O mecanismo é que um pequeno padrão de bits, no campo TOS do IPv4 ou Class do IPv6, é usado para marcar um pacote para que ele receba um tratamento de encaminhamento particular, ou PHBs (*Per-Hop Behaviors*), em cada nó da rede. PHB é o comportamento observável externamente de um pacote em um roteador suportando DS. O enfoque do DiffServ é padronizar uma estrutura comum a ser usado para o campo TOS do IPv4 ou Class do IPv6, agora chamado de DS (*Differentiated Services*). Modificando os formatos definidos anteriormente pela IETF, este campo é definido em [Nichols, 98]:

- Seis bits do campo DS são usado como codepoint DSCP (*Differentiated Service CodePoint*) para selecionar o PHB que o pacote terá em cada nó. Este campo é tratado como um índice de uma tabela que é usada para selecionar um mecanismo de manipulação de pacotes implementado em cada dispositivo. Este campo é definido como um campo não estruturado para facilitar a definição de futuros PHBs.
- Um campo de dois bits é reservado (são ignorados por nós DS-conformantes).

Marcando os campos DS dos pacotes diferentemente, e manipulando pacotes baseados nos seus campos DS, várias classes de Serviços Diferenciados podem ser criadas. Portanto, Serviços Diferenciados é essencialmente um esquema de prioridades.

Quando aos serviços oferecidos por um domínio DS-conforme, devemos notar:

- Serviços DS são todos para tráfego unidirecional apenas.
- Serviços DS são para tráfegos agregados, não fluxos individuais.

[Black, 98] define um Serviço como o tratamento global de um subconjunto do tráfego do cliente dentro de um domínio DS-conforme ou fim-a-fim. O tráfego na rede hoje geralmente atravessa uma concatenação de redes que podem incluir hosts, redes residenciais e de escritório, redes corporativas/campus e várias redes de longa distância. Redes residenciais e de escritório são normalmente clientes de redes campus ou corporativas, que são por sua vez clientes de redes longa distância. Note que existem várias fronteiras cliente/provedor em que o conceito de serviço se aplica.

A fim de que os clientes recebam Serviços Diferenciados de seu Provedores de Serviço Internet (*ISP – Internet Service Provider*), eles devem firmar um Acordo de Nível de Serviço (*SLA – Service Level Agreement*) com seu ISP. Vários aspectos dos SLAs (como termos de pagamento) são fora do escopo de padronização; é a Especificação do Nível de Serviço (*SLS – Service Level Specification*) é que especifica as classes de serviço suportados e o montante de tráfego permitido em cada classe. Um SLA pode ser estático ou dinâmico. SLA estáticos são negociados mensalmente, anualmente, etc. Clientes com SLA dinâmicos devem usar um protocolo de sinalização (p.e. RSVP) para pedir por serviços sob demanda.

Os clientes podem marcar os campos DS de pacotes para indicar o serviço desejado ou estes campos são marcados pelo roteador que liga o cliente a rede ISP (leaf router) baseado na classificação MF (multicampo).

No ingresso às redes ISP, os pacotes são classificados, policiados e possivelmente atrasados para torná-los conformes a algum perfil de tráfego pré-instalado. As regras de classificação, policiamento e retardos usadas nos roteadores de ingresso são derivados a partir dos SLAs. O montante de espaço de bufferização necessário para estas operações também é derivado dos SLAs.

Um exemplo simples de perfil de tráfego poderia ser: medir o fluxo de pacotes do endereço IP a.b.c.d e se sua taxa fica abaixo de 200 kbps, sete o byte-DS para o valor X,

senão sete o byte-DS para o valor Y. Se a taxa excede 600 kbps, corte os bytes excedentes. Os perfis são configurados pelo operador de acordo com o SLAs. Como os perfis são fornecidos (configuração manual ou sinalização) é fora do escopo do dffserv. Dentro da rede (nos roteadores internos ao domínio), o byte DS é usado para determinar como os pacotes são tratados. O tratamento, também chamado de PHB ou comportamento agregado, pode incluir diferentes prioridades envolvendo atraso de enfileiramento (escalonamento), diferentes prioridades na decisão de descarte na sobrecarga de filas (gerenciamento de fila), seleção de rota, etc. De  $2^6$  possíveis significados (dados pelo campo DSCP), o grupo de trabalho DiffServ especificará (padronizará) alguns PHBs globalmente aplicáveis, e deixará o resto para uso experimental. Se os experimentos indicarem que um certo PHB não padronizado é claramente útil, ele pode ser padronizado posteriormente. Isto ainda está sob debate.

É de responsabilidade das ISPs decidir que serviços fornecer. Os seguintes serviços poderiam ser fornecidos:

- Serviço *Premium*, para aplicações requerendo serviço de pequeno atraso e pequena variação de atraso. Neste caso, o usuário negocia com seu ISP a máxima largura de banda para enviar pacotes através da rede e as alocações são feitas em termos de taxa de pico. Uma desvantagem é o fraco suporte a tráfegos em rajada e o fato de que o usuário paga mesmo quando não usa completamente a largura de banda.
- Serviço *Assegurado*, para aplicações requerendo melhor confiabilidade que Serviço Melhor Esforço. Este serviço não garante a largura de banda como o Serviço Premium, mas fornece uma alta probabilidade de que o ISP transfere os pacotes marcados com alta prioridade confiavelmente. Ele não foi completamente definido, mas oferece um serviço equiparável ao Serviço de Carga Controlada do IntServ;
- Serviço *Olympic*, que fornece três tipos de serviços: Ouro, Prata e Bronze, que reduz em qualidade.

Serviços Diferenciados é significativamente diferente de Serviços Integrados:

- Primeiro, há apenas um número limitado de classes de serviço indicados no campo DS. Desde que o serviço é alocado na granularidade de uma classe, o conjunto de informações de estado é proporcional apenas ao número de classes e não proporcional ao número de fluxos. Serviços Diferenciados é portanto mais escalável do que Serviços Integrados.
- Segundo, as operações de classificação, marcação, policiamento e retardo são apenas necessárias nas fronteiras das redes. Roteadores ISP internos (core) necessitam apenas implementar a classificação Comportamento Agregado (BA - *Behavior Aggregate*), que é uma classificação baseada apenas no byte DS. Portanto, Serviços Diferenciados é mais fácil de implementar e usar.

No modelo Serviços Diferenciados, um serviço assegurado pode ser fornecido por um sistema que suporta parcialmente os Serviços Diferenciados. Roteadores que não suportam Serviços Diferenciados simplesmente ignoram os campos DS dos pacotes e fornecem a pacotes serviço assegurado o Serviço Melhor Esforço. Desde que pacotes serviço assegurado têm menos probabilidade de serem cortados em roteadores compatíveis com DS, o desempenho total do tráfego serviço assegurado será melhor que o tráfego Melhor Esforço.

## 13.4 Rede MBone

MBone é a abreviação de *Virtual Internet Backbone for Multicast IP* ou *Multicast Backbone*. Ela é uma rede virtual dentro da camada física da Internet construída como uma estrutura de teste do IP multicast. O MBone foi criado como uma rede experimental do protocolo IP Multicast.

A MBone iniciou em março de 1992, durante o encontro do *Internet Engineering Task Force* (IETF) em San Diego, USA, quando muitos eventos do encontro foram transmitidas em áudio utilizando transmissões de pacotes de difusão seletiva do site

IETF para participantes de 20 sites em três continentes, abrangendo 16 zonas horárias [Casner, 92] [Casner, 95]. Foram transmitidas todas as sessões gerais do encontro e também algumas sessões de grupos de trabalho. Os participantes remotos tinham a possibilidade de falar, participando das discussões e fazendo perguntas, e embora a transmissão não fosse perfeita, foi razoável em ambas as direções. No site do IETF e na maioria dos sites remotos foi utilizada a ferramenta VAT, *Visual Audio Tool* (vista mais adiante), desenvolvida por Van Jacobson e Steve McCanne, do *Lawrence Berkeley Laboratory*. Foi usado a velocidade de 64 Kbps para áudio codificado com PCM.

A transmissão do evento foi possível graças a utilização de IP multicast implementado nos roteadores e *hosts*, onde uma cópia simples é enviada de cada pacote, sendo replicada apenas quando houver um braço na árvore lógica dos destinos. Se fosse necessário enviar uma cópia de cada pacote para cada destino, a largura de banda e processamento requerido seriam proibitivos.

Atualmente, a MBone tem mais de 1000 sub-redes, sua popularidade e importância é devido principalmente a necessidade da comunicação humana natural e tempo-real entre humanos através das redes de computadores. Desde que na MBone, o usuário pode enviar e receber áudio e vídeo tempo-real a partir de um computador, através de um conjunto de ferramentas (uma lista completa destas ferramentas pode ser encontrada em <http://www.merit.net/net-research/mbone/index/titles.html>).

Como já citado, as duas principais aplicações da rede MBone são a distribuição de áudio e vídeo, para a difusão dos chamados eventos. Alguns eventos são constantes, estão sempre acontecendo, como o MBone RTP Audio e o Radio Free Vat, por exemplo, que são grupos como estações de rádio do MBone onde todos podem ser *disk jockeys*. Já outros eventos acontecem temporariamente, transmitindo congressos, apresentações, ou eventos que estão ocorrendo temporariamente.

### 13.4.1 Arquitetura MBone

Como já apresentado, a MBone é uma rede virtual que utiliza parte da rede física da Internet para suportar roteamento de pacotes IP multicast. Ela consiste de subredes ilhas que diretamente suportam IP multicast como LANs multicast como Ethernet. Estas ilhas são unidas por ligações virtuais ponto-a-ponto chamadas de túneis. Estes túneis são formados por ligações e roteadores IP unicast tradicionais. Nele os pacotes IP multicast são encapsulados como pacotes unicast normais. Os pontos finais dos túneis são tipicamente estações executando daemon multicast. Estes roteadores multicast (mrouters) replicam os pacotes da saída nas interfaces múltipla e túneis tanto quanto necessário ao longo da árvore de distribuição.

A Figura 119 apresenta os principais roteadores e ligações do MBone em 1994. Naquele momento, os continentes eram conectados por um único túnel, os EUA tem vários túneis entre seus estados. A maioria dos túneis são roteados dentro de linhas T1(1.5 Mpbs) ou E1 (2Mbps).

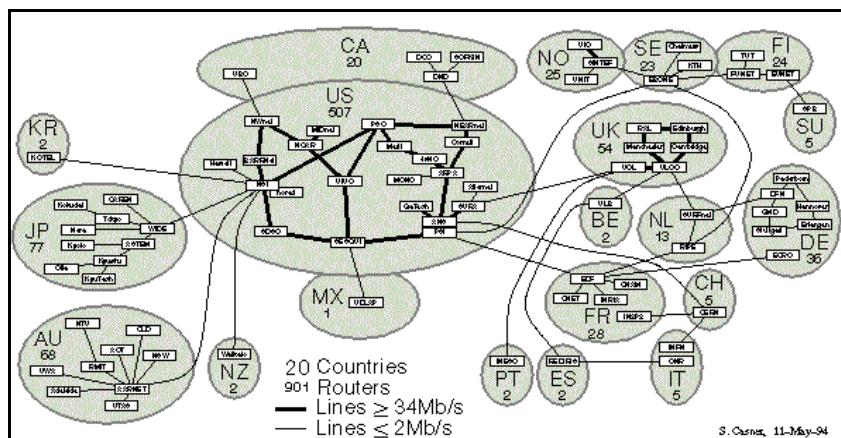


Figura 119. Principais roteadores e ligações MBone

### **Mroutes**

mrouters são na maioria das vezes apenas uma estação de trabalho Unix que executa um software especial em nível do usuário. Os mrouters têm a responsabilidade de replicar e distribuir os quadros de dados de difusão seletiva para os túneis que conduzem a estações dos participantes do grupo e para a rede local, caso exista um membro daquele grupo nela. Quando um datagrama é enviado por uma estação, que o coloca na rede local, ele é apanhado pelo mrouter da sub-rede. O roteador consulta sua tabela de roteamento e retransmite o datagrama para os túneis que conduzem a membros do grupo. No outro lado do túnel, o outro roteador recebe os datagramas e consulta sua tabela de roteamento para decidir se o pacote deve ser enviado para algum outro túnel e se há alguma estação em sua sub-rede que está inscrita neste endereço, colocando-o então na sub-rede local para ser recebido pela estação.

### **Aplicações MBone usam UDP**

O MBone utiliza o protocolo *User Datagram Protocol* (UDP) na camada de transporte, diferente do tradicionalmente usado para aplicações interativas *Transmission Control Protocol* (TCP). Como visto na seção 13.2, o motivo para o uso do UDP é que o controle de fluxo e a confiabilidade oferecidos pelo TCP não são adequados para, por exemplo, transmissão de áudio ao vivo. Além disso, o protocolo TCP não pode ser facilmente utilizado para realizar difusões IP.

O uso do UDP, protocolo não confiável, leva a não garantia de ordenação e confiabilidade na entrega. Para isto, o protocolo RTP é utilizado pela maioria das aplicações MBone para se efetuar ordenação dos pacotes e controle da transmissão.

### **Como MBone garante a transmissão de áudio e vídeo?**

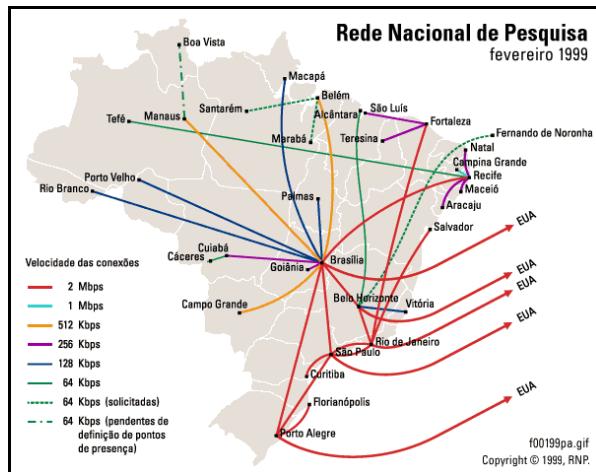
No MBone não há reserva de recursos para suportar a transmissão de áudio e vídeo tempo-real. A principal característica é a habilidade multicast. A transmissão de áudio e vídeo é possível graças a três técnicas e ferramentas [Fluckiger, 95]:

- A primeira técnica é que o túnel é normalmente constituído em ligações não congestionadas. Se todos as ligações possíveis são congestionadas, a largura de banda disponível ao túnel é muito pequena e a qualidade de áudio e vídeo é muito baixa e provavelmente não aceitável.
- A segunda técnica é que o UDP é usado em vez do TCP para evitar retransmissões para pacotes de áudio e vídeo e reduzir atrasos. No topo do UDP, RTP é usado para fornecer informações de tempo e sequenciamento aos receptores.
- O terceiro fator habilitador está no projeto de transmissor e receptor de áudio e vídeo. Primeiro áudio e vídeo são transmitidos apenas na qualidade mínima aceitável. Segundo, um buffer de apresentação adaptativo é usado para reduzir as variações de atraso da rede para apresentações contínua, mas guardando o atraso total ao mínimo para boa interatividade.

#### **13.4.2 Situação atual**

A qualidade de áudio do MBone é normalmente muito boa. Por limitações da largura de banda, canais de vídeo são geralmente configurados a 128 Kbps. Isto nos dá um a quatro quadros por segundo, com uma resolução que é geralmente mais baixa que em sistemas de videoconferência a circuito. Como resultado, o descompactação do fluxo de vídeo é geralmente feita via software.

A responsabilidade do uso diário de uma rede MBone consiste basicamente de se assegurar da não sobrecarga da largura de banda de sua rede local ou a regional (veja as taxas de transferências do backbone da RNP na Figura 120). Esta questão pode ter maior impacto no desempenho da rede, por exemplo um fluxo de vídeo (de 14 fps) consome perto de 128Kbps de largura de banda, que é aproximadamente 6,4% de uma ligação a 2 Mbps, muitas sessões simultâneas facilmente saturam a rede, suas ligações e roteadores.



**Figura 120.** Backbone da rede nacional de pesquisa

Devido a falta de garantias da QoS, o áudio e vídeo sofrem com o congestionamento da rede causando perda de pacotes e atrasos. Para fornecer áudio e vídeo garantidos sobre a Internet, controle de QoS, protocolos de transporte tempo-real, sistemas finais potentes são necessários. Além disso, as redes básicas também necessitam ser atualizadas de modo que elas possam garantir a QoS.

Como já apresentado, atualmente existem várias atividades de pesquisa para fornecer garantias de QoS sobre a Internet. O resultado provável é que RSVP e/ou ST-II estendido seja usado para estabelecimento de conexão. Um modo efetivo para reservar e gerenciar recursos é ainda necessário. Sem uma solução da especificação e garantia de QoS, não podemos apenas desenvolver aplicações usando a abordagem melhor esforço (aquele fornecido pelo IP).

### 13.4.3 Ferramentas de Anúncio de Sessão

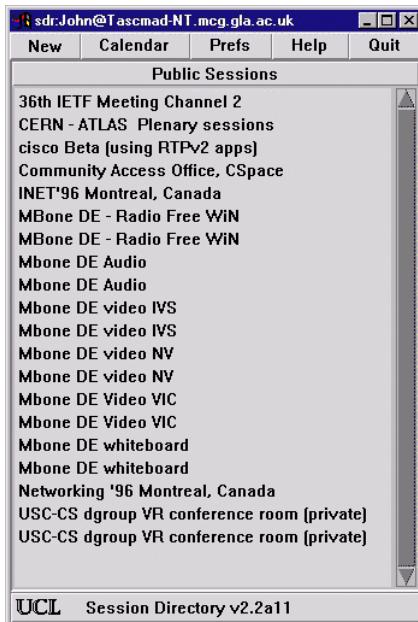
As ferramentas de anúncio de sessão são utilizadas para criar, reservar e anunciar uma sessão multicast que será utilizada para transmitir um evento. No momento da criação, é reservado um ou mais endereços multicast para o evento. Estes endereços ficam reservados para o evento por um período de tempo determinado no momento da criação, sendo liberados quando o período de transmissão do evento se encerra.

Além de serem utilizadas pelos criadores de uma sessão, estas ferramentas também são utilizadas pelos usuários que desejam participar de grupos específicos. A ferramenta anuncia todos os eventos que estão sendo transmitidos no momento, identificando-os pelos seus nomes, e fornece informações sobre eles, apresentando uma descrição do evento, o período da transmissão, os tipos de mídia utilizados pelo evento e os endereços e portas reservados para cada um destes. De posse destas informações, o usuário pode escolher a quais grupos ele deseja participar, bastando solicitar que os grupos das mídias desejadas sejam ativadas.

Existem diversas ferramentas deste tipo para a MBone, como o *Session Directory* (SDR) e a *Multimedia Conference Control* (MMCC).

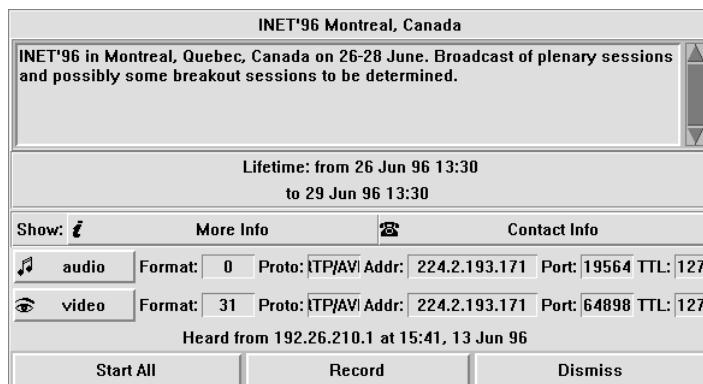
SDR é uma ferramenta de sessão projetada para permitir o anúncio e uso de conferência *multicasting* sobre o MBone. Assim, ela permite que sejam reservados canais de difusão seletiva de áudio, vídeo e quadro branco para conferências, e que se participe nos diversos grupos anunciados. SDR é disponível em várias plataformas, incluindo Windows95, Solaris, SunOS e AIX.

A Figura 121, ilustra a janela principal da SDR. Ela mostra quais as sessões estão sendo transmitidas no momento, fornecendo informações sobre elas e executando, caso desejado, as ferramentas de áudio, vídeo, quadro branco e texto nos grupos vinculados ao evento automaticamente.



**Figura 121.** Janela principal da Session Directory

A Figura 122, apresenta todas as informações sobre uma sessão selecionada, incluindo os tipos de mídia (áudio e vídeo) sendo transmitidos, o endereço utilizado por cada uma destas mídias, o TTL inicial dos datagramas<sup>27</sup>, etc. Para associação em qualquer um destes, basta escolher os botões audio, video ou Start All, que automaticamente acionarão as aplicações de áudio e vídeo para este grupo.

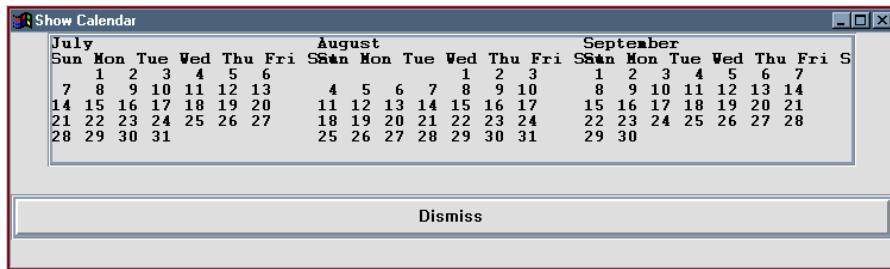


**Figura 122.** Janela de informação de uma sessão da Session Directory

A SDR pode ser usada também para criação de uma sessão. O processo é simples, bastando selecionar, a partir da janela principal, o botão New, que acionará uma janela onde devem ser informadas quais mídias serão utilizadas, o nome e descrição do evento, o responsável, o escopo do alcance da sessão (definido pelo TTL), o dia, o horário e a duração. A ferramenta alocará um endereço livre por aquele período de tempo, e a sessão será criada.

A ferramenta apresenta também um calendário dos eventos da MBone já anunciados (Figura 123), mostrando os eventos agendados para cada dia. O calendário é acessado a partir do botão Calendar na janela principal. Existe também um site na Internet, denominado MBone Agenda [MBone, 99], onde podem ser anunciados previamente os eventos que serão transmitidos em datas futuras, e podem também ser convertidos automaticamente os horários de uma transmissão para os horários locais, já que os horários são anunciados segundo a hora local de cada lugar de origem.

<sup>27</sup> TTL (Time To Live) define a alcançabilidade do evento: local, regional, internacional.



**Figura 123.** Janela de informações dos eventos já anunciados

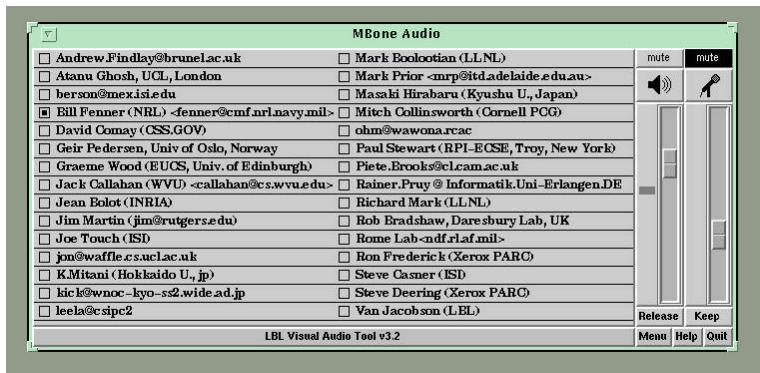
### 13.4.4 Ferramentas de Áudio

As ferramentas de áudio do MBone podem ser utilizadas facilmente, necessitando apenas de alto-falante para serem utilizadas em estações receptoras. Para transmissão, elas requerem um microfone conectado a estação transmissora. Entre as ferramentas de áudio podemos citar VAT (<http://www-nrg.ee.lbl.gov/vat/>), FreePhone (<http://www-sop.inria.fr/rodeo/fphone/>) e o Network Voice Terminal (Nevot - <http://www.fokus.gmd.de/step/nevot/>).

#### VAT

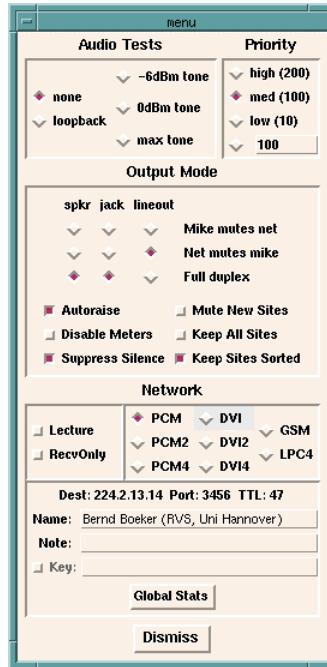
*Visual Audio Tool* (VAT) é a ferramenta de audioconferência mais usada na MBone, permitindo aos usuários realizar áudioconferências sobre a Internet de um computador para outro computador (ponto-a-ponto) usando endereços unicast padrões e também entre vários computadores (multicast). Foi desenvolvida por Van Jacobson e Steve McCanne, no *Lawrence Berkeley Labs* da *University of California*, em Berkeley [Kumar, 95]. Vat é disponível em várias plataformas, incluindo Microsoft Windows, Solaris e SunOS. Esta ferramenta utiliza o protocolo RTP para encapsulamento dos pacotes de áudio.

Através de sua janela principal (Figura 124) esta ferramenta permite que sejam identificados os membros participantes do grupo e o membro transmissor (identificado por um sinal na esquerda do membro) e que seja controlada a ativação do microfone e auto-falante, e permitindo regular o volume para ambos. Esta ferramenta oferece ainda informações como nome, endereço eletrônico e última transmissão enviada dos membros participantes, obtidas através de um clique sobre o membro. Informações de controle são trocadas por pacotes de controle RTCP do protocolo RTP.



**Figura 124.** Janela principal da Visual Audio Tool

Esta ferramenta oferece uma série de opções de transmissão, permitindo que sejam selecionados o padrão de compressão de dados (PCM, PCM2, PCM4, DVI, DVI2, DVI4, GSM - General Special Mobile e LPC - Linear Predictive Coding), que a sessão seja criptografada e acessada apenas pelos usuários que possuam sua senha, etc. Estas opções são configuradas na janela Menu, apresentada na Figura 125.



**Figura 125.** Janela menu da Visual Audio Tool

O vat implementa a sincronização intramídia no receptor utilizando as informações adicionais providas pelo protocolo RTP. Como a Internet pode introduzir atrasos variados, o receptor vat tenta eliminar os efeitos da variação de atraso através de um buffer de apresentação (seção 7.3.4): pacotes são armazenados no buffer de apresentação no destino e são liberados para a aplicação no tempo correto. O tamanho requerido do buffer de apresentação depende da variação de atraso dos pacotes recebidos. Se todos os pacotes têm aproximadamente o mesmo atraso, o buffer de apresentação pode ser pequeno. Mas se a variação de atraso é grande, o buffer de apresentação deve ser aumentado. Neste caso, como foi visto na seção 7.3, o atraso fim-a-fim aumenta. Portanto, é interessante ter um buffer grande apenas se os atrasos de pacotes diferem muito. Vat tenta minimizar o atraso adaptando o tempo de apresentação às características de atraso da rede. Desde que a adaptação dos tempos de apresentação para cada pacote poderia resultar em buracos audíveis, o tempo de apresentação é adaptado na partida de uma rajada de fala apenas pela mudança do tamanho do intervalo de silêncio.

Vat também fornece características estatísticas permitindo ao usuário monitorar as estatísticas da rede tal como análise de atraso e detecção de perda.

### FreePhone

FreePhone (Figura 126) é uma aplicação de audioconferência baseada no protocolo RTP desenvolvida pelo INRIA (*Institut National de Recherche en Informatique et en Automatique*). Ela é disponível em várias plataformas, incluindo Windows95, SunOS e Linux.

Esta ferramenta gerencia várias sessões unicast e multicast, assim não há necessidade de se executar simultaneamente várias instâncias do software. Esta ferramenta combina controle de taxa e correção FEC (*Forward Error Correction*) para fornecer uma qualidade de áudio estável na Internet. Além disso, dependendo dos relatórios de estado do RTCP, a ferramenta adapta a taxa de transmissão e o conjunto de informações redundante adicionados aos pacotes de áudio. FreePhone também inclui um mecanismo para adaptar o buffer de apresentação dependendo das estatísticas de QoS.

FreePhone fornece as codificações PCM (*Pulse Code Modulation*), ADM (*Adaptative Delta Modulation*) de 16 Kbps (ADM2) até 48 Kbps (ADM6), GSM (13 Kbps) e LPC (4,8 Kbps).



**Figura 126.** Janela Principal do FreePhone

### 13.4.5 Ferramentas de Vídeo

As ferramentas de vídeo não necessitam de nenhum equipamento adicional na estação para receber vídeo. Elas utilizam vários algoritmos de compressão, que comprimem os quadros de vídeo de forma bastante significativa numa taxa de até 20:1 [Trentin, 96].

A ferramenta nv (Network Video - - <http://mice.ed.ac.uk/mice/archive/nv.html>) desenvolvida pela Xerox PARC e a IVS (INRIA video conferencing system - <http://www.inria.fr/rodeo/ivs.html>) são as primeiras ferramentas de videoconferência Internet capazes de realizar multicast. IVS também suporta a transferência de áudio. Baseada na experiência obtida pelo NV e IVS, o LBL criou a ferramenta vic (VideoConference - <http://mice.ed.ac.uk/mice/archive/vic.html>). O INRIA desenvolveu o sucessor do IVS, chamado de Rendez-Vous (<http://zenon.inria.fr/rodeo/rv/>). Além dessas ferramentas, novos produtos comerciais para teleconferência usando IP Multicast começaram a emergir, entre outras podemos citar o Picture Window, da Bolt Beranek InPerson da Silicon Graphics, e o ShowMe da Sun Microsystems.

Muitas dessas ferramentas são baseadas no RTP e usam o formato de pacote padronizado para encapsular diferentes esquemas de codificação de áudio e vídeo e transmitem eles sob UDP e IP multicast. Graças a adoção do protocolo RTP, estas ferramentas podem interoperar. RTP é também a base de muitos produtos de áudio/vídeo-conferência disponíveis comercialmente hoje.

### IVS

A ferramenta IVS (INRIA Video Conferencing System) (Figura 127) é um dos primeiros softwares de videoconferência da Internet (primeira versão surgiu em julho de 92). Ela usa a compressão H.261 para funcionar bem em ligações de baixa largura de banda. Para codificação de áudio ela utiliza PCM e ADPCM. Codificações de áudio e vídeo são realizadas via software.

IVS implementa um algoritmo de controle de congestionamento adaptando sua taxa de saída às condições de rede. O mecanismo de controle de congestionamento é baseado em dois componentes:

- Um **sensor de rede**, que mede a perda de pacotes pelos relatórios de QoS RTP gerados pelos receptores após cada 100 pacotes. Estes relatórios de QoS periódicos são mais eficientes que enviar reconhecimentos negativos para perda de pacotes desde que a taxa de perdas é maior que 1%, isto é, se mais que 1 pacotes a cada 100 é perdido, que é geralmente o caso na Internet.
- Um **controlador de vazão**, que ajusta a taxa de saída máxima do codificador de modo que a taxa de perda média para todos os receptores mantenha-se abaixo do valor tolerável. A taxa de saída do codificador pode ser ajustada alterando a taxa

de quadros, os fatores de quantificação, ou o domínio de pesquisa do vetor de movimento.



**Figura 127.** Janela principal do IVS

IVS permite a seleção de dois modos diferentes: privilégio a qualidade (PQ) e privilégio a taxa de quadros (PFR). PQ é apropriado para aplicações exigindo alta precisão. PFR é apropriado quando a percepção do movimento é um importante fator de qualidade.

As seguintes plataformas são usadas:

- Estações Silicon Graphic com IndigoVideo, GalileoVideo e VinoVideo;
- PC/Linux com SCREENMACHINE II;
- Estações DEC 5000 com VIDEOTX;
- Estações DEC ALPHA sem placas de captura de vídeo;
- Estações PC/FreeBSD2.0 sem placas de captura de vídeo;
- Estações HP com VideoLive.

### Rendez-Vous

O sucessor do IVS, chamado de Rendez-Vous, está atualmente sendo desenvolvido no INRIA [Lyonnet]. Ela já suporta JPEG e H.261, e H.263<sup>28</sup> está sendo implementado também. Arquivos de vídeo MPEG-1 e MPEG-2 podem ser lidos pela ferramenta e transcodificados para um dos formatos mencionados. Procedimentos de escalonamento otimizados aumenta o uso dos recursos do sistema, resultando em um melhor comportamento em tempo de execução e melhor qualidade fornecida ao usuário.

Rendez-Vous é disponível nas plataformas Windows95, Solaris, Linux e outras.

### VIC

Videoconference (VIC) é hoje a ferramenta de videoconferência mais usada na MBone. Esta ferramenta foi desenvolvida com uma arquitetura flexível e extensível para suportar ambientes e configurações heterogêneas. Ela é baseada no protocolo RTP e está disponível para a maioria das plataformas Unix, incluindo PCs rodando Linux e BSD/386<sup>29</sup>. Para Windows95 é disponível apenas uma aplicação para a recepção de vídeos.

A Figura 128, apresenta a janela principal do vic, onde são fornecidas informações sobre o evento, como originador, taxa de quadros sendo transmitidos e velocidade. Ela

<sup>28</sup> H.263 foi projetado para codificação de vídeo para comunicações de baixa taxa de bits (inferiores a 64 Kbps), mas atualmente esta limitação foi removida. A codificação h.263 é similar aquela usada no H.261, mas com melhorias e trocas para aumentar o desempenho e correção de erros.

<sup>29</sup> BSD (Berkeley System Distribution)/386 é uma versão de Unix para 386

apresenta uma versão reduzida do vídeo, caso o usuário clique sobre este vídeo, será aberto uma janela com dimensões maiores.



**Figura 128.** Janela principal da VideoConference

A ferramenta oferece diferentes esquemas de compressão, utilizando no modo padrão (sem alterações) o esquema de compressão H.261 e o RTP como o mecanismo de transporte da aplicação. Oferece ainda controle da velocidade de transmissão e possibilidade de criptografia de acordo com a especificação RTP. Estas alterações são feitas na janela Menu, que é acessada com um clique no botão Menu da janela principal.

Vic pode executar sobre diferentes camadas de redes tal como UDP/IP, a pilha de protocolos Tenet [Banerjea, 96] ou AAL5/ATM. Em [Braun, 97a] é proposta uma extensão do vic para operar sob o protocolo RSVP (isto pois o código fonte do vic é disponível livremente).

Se IVS e nv suportam a compressão/descompressão em software apenas, vic foi projetada para também tomar vantagem da compressão/descompressão por hardware.

## NV

A ferramenta de videoconferência *Network Video* (NV) permite a usuários transmitir e receber vídeo a baixa taxa de quadros via UDP/IP sobre a Internet. Ela utiliza um algoritmo de compressão de vídeo desenvolvido especialmente para atingir baixa vazão de dados e alta vazão dos quadros, e o protocolo RTP versão 1 no nível de transporte da aplicação. Fluxos de vídeo podem ser enviados ponto-a-ponto ou para vários destinos simultaneamente usando IP multicast. NV é disponível nas plataformas Solaris, Linux, SunOS e várias outras.

O vídeo transmitido é uma imagem 320x240 para NTSC ou 384x288 para PAL. A taxa de quadros varia de acordo com o movimento e a largura de banda disponível. Taxas de quadro de 3-5 fps são valores típicos para a largura de banda de 128kbps.

Assim como na VIC, a janela principal do NV (Figura 129) apresenta uma imagem reduzida do vídeo e informações sobre a sessão, tais como nome do originador e TTL inicial dos datagramas. Com um clique sobre a imagem, a janela aumentada da aplicação é disparada.

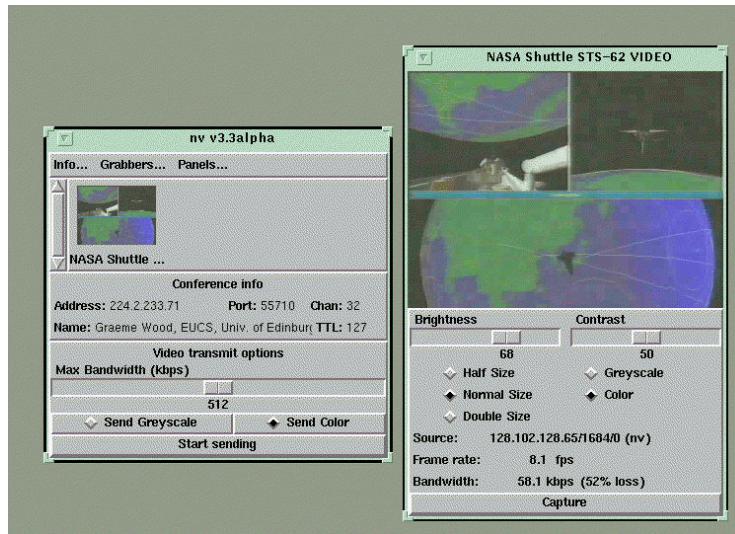


Figura 129. Janela Principal do NV

### 13.4.6 MBone VCR: um servidor de vídeo para a MBone

O MBone VCR (*Video Conference Recorder*) [Hoffelder, 95] (Figura 130) fornece funcionalidades para registrar e apresentar sessões MBone ao vivo com múltiplos fluxos de dados multimídia oriundos de diferentes aplicações. O MBone VCR permite aos usuários assistir e escutar conferências originárias de localizações de zonas de tempo muito diferentes. Por exemplo, um usuário brasileiro pode registrar uma transmissão MBone de uma conferência nos EUA durante a noite e assistir no próximo dia.



Figura 130. Interface com o usuário do MBone VCR

Uma sessão registrada pode consistir de quantos canais multicast que o usuário desejar registrar. Os canais podem ser originários de diferentes localizações e aplicações.

Durante o registro, o MBone VCR sincroniza os fluxos de dados multimídia, baseado em informações providas pelo RTP. Para apresentar os fluxos de dados, o MBone VCR transmite o dado registrado, usando a temporização formato dos pacotes originais. Os receptores devem apenas executar uma aplicação tal como vic ou vat para assistir e ouvir o dado registrado.

O MBone VCR fornece características interessantes, tal como indexação do dado registrado. O usuário pode facilmente saltar certas fases, por exemplo, um certo palestrante, de uma sessão. Índices permitem o usuário saltar de um lado para outro entre cenas marcadas. Outras características, tal como avanço e retrocesso rápido, são similares aos vídeos cassete. A interface do usuário é também similar aquela dos vídeo cassete. Sessões a serem registradas podem ser selecionadas a partir do sd (session directory), que fornece ao usuário uma lista dos eventos MBones atuais e futuros.

A nova versão do VCR, chamada de MBone VCR on Demand Service (MVoD) [Hoffelder, 97], disponível em <http://www.informatik.uni-mannheim.de/informatik/pi4/projects/MVoD/>, (Figura 131) é agora implementada em Java.

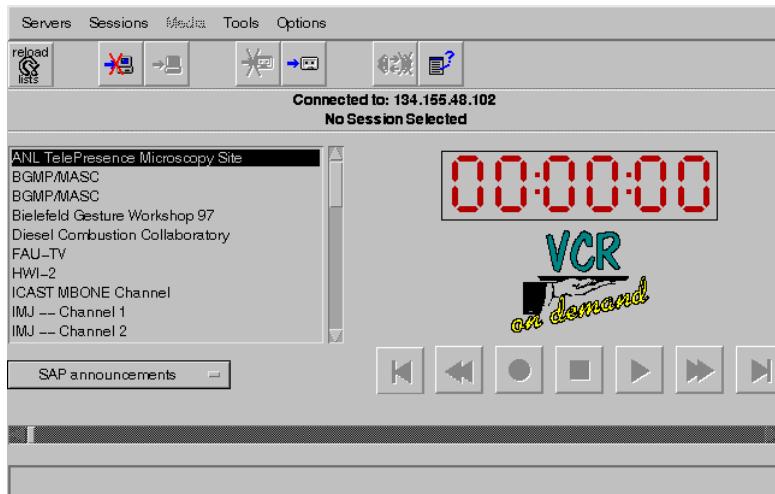


Figura 131. MBone VCR on Demand Service (MVoD)

### 13.4.7 Ferramentas de Quadro Branco Compartilhados

As ferramentas de quadro branco compartilhado permitem que os membros de uma sessão compartilhem documentos em tempo real, podendo ser efetuadas anotações sobre estes documentos a qualquer momento. Elas são especialmente indicadas para transmissões de conferências e aulas de ensino a distância, já que podem ser utilizados como um retroprojetor virtual, onde cópias de transparências de uma apresentação para o público local também podem ser apresentadas para os membros virtuais [Trentin, 96]. Entre as principais ferramentas desta classe podem ser citadas a WhiteBoard (WB - <ftp://ftp.ee.lbl.gov/conferencing/wb/wb-1.60-solaris24.tar.gz>) e o TeleDraw (<http://www.uni-stuttgart.de/RUS/Projects/MERCI/MERCI/TeleDraw/Info.html>).

#### WB

A aplicação *WhiteBoard* foi desenvolvida por Van Jacobson e Steve MacCanne no *Lawrence Berkely Labs* da *University of California*, em Berkeley e é a ferramenta de quadro branco mais usado na MBone. Ela permite que qualquer participante da conferência crie uma página e edite qualquer página. Nestas páginas podem ser compartilhados documentos em tempo real, podendo ser utilizados documentos em texto ASCII, desenhos, anotações a mão livre e páginas em PostScript.

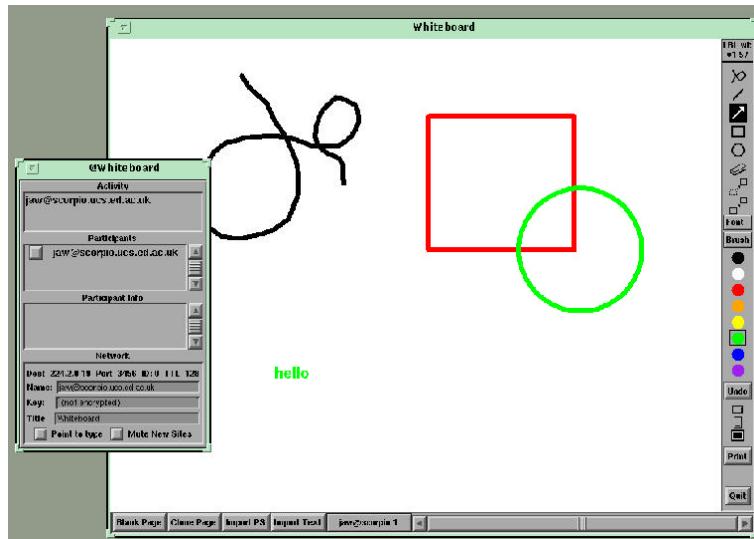
WB é disponível nos sistemas Solaris, SunOS, Linux, AIX e outros.

A janela principal desta ferramenta apresentada Figura 132 mostra um exemplo de documento da sessão. Esta figura apresenta também uma janela adicional que fornece informações sobre o evento tais como participantes, membro que está interagindo no momento, informações sobre os participantes, etc.

WB é destinado a discussão e anotações de documentos, ele não é uma ferramenta de desenho completa. Um conjunto básico de operações de desenho, tal como linhas, ponteiros, caixas, círculos, funcionalidades de apagador, desenhos a mão livre e anotadores de texto são disponíveis. Um conjunto de cores diferentes pode ser usado para diferenciar participantes e indicar que usuário realizou uma operação de desenho ou anotação.

O controle da sessão é informal, onde todos os participantes podem fazer alterações digitando texto, desenhando gravuras ou fazendo anotações a mão livre sobre a sessão compartilhada. Existem porém características especiais da ferramenta que fornecem algum controle sobre a sessão, como a capacidade de criar uma sessão no modo de conferência (*Lecture mode*), que garante que nenhuma das alterações efetuadas por estações diferentes da transmissora serão transmitidas para os outros participantes.

WB é uma ferramenta que pode ser executada mesmo em condições de baixa largura de banda da Internet. Um valor típico de largura de banda gerada é 64 Kbps, que é a mesma ordem de magnitude de uma áudio-conferência.



**Figura 132.** Janela de informações da WhiteBoard

Em contraste com as aplicações de áudio e vídeo, wb requer um alta confiabilidade (de 100%). Retransmissões são geradas por um algoritmo chamado de multicast confiável escalonável (SRM - Scalable Reliable Multicast) [Floyd, 95]. É oferecido também esquema de criptografia para a sessão. Para isto, todos os participantes com permissão de participarem de uma sessão devem conhecer a chave para descriptografar o dado criptografado.

O sucessor do wb, chamado de MediaBoard (<http://www-mash.CS.Berkeley.EDU/mash/>), está atualmente sendo desenvolvido na Universidade da Califórnia em Berkeley. Como o wb, MediaBoard baseia-se na abordagem SRM para multicast confiável. Além das operações convencionais de desenho tal como linhas, retângulos, e elipses, ele também suporta texto, imagens (GIF e PostScript), animações, e fluxo de vídeo embutidos. MediaBoard é disponível em várias plataformas Unix e windows95.

### 13.4.8 Editor de texto compartilhado

A ferramenta NTE (Network Text Editor) [Handley, 97] (<http://mice.ed.ac.uk/mice/archive/nt.html>) é um editor compartilhado projetado para a MBone. Os participantes podem editar o mesmo simultaneamente. Os usuários devem sincronizar seus trabalhos para evitar conflitos quando dois ou mais usuários editam o mesmo bloco de texto simultaneamente. Esta sincronização pode ser feita por conversação de áudio durante a conferência. Blocos de texto podem também ser protegidos bloqueando edições indesejáveis.

NTE foi projetado para obter um bom desempenho e escalabilidade para muitos usuários. Devido à necessidade de desempenho, inconsistências temporais, isto é diferentes visões do texto compartilhado por diferentes usuários, são toleradas. Inconsistências podem resultar devido a perdas de pacotes, posterior junção de membros, e trocas simultâneas por diferentes usuários no mesmo objeto de texto. Estas causas da inconsistência ocorrem pois o NTE é baseado no serviço IP multicast não confiável.

Vários mecanismos foram implementados para detectar visões inconsistentes. Periodicamente, cada membro do grupo da sessão realiza a difusão de mensagens de sessão indicando a lista de membros da conferência e contendo uma marca temporal da modificação mais recente e um checksum de todos os dados. A taxa das mensagens de sessão depende do número de membros da sessão a fim de evitar muito tráfego de controle. Se a marca temporal de outro sítio é posterior a marca temporal indicada, o receptor pode pedir todas as trocas do intervalo faltante. Os checksums são usados para detectar que blocos diferem entre dois ou mais sítios.

Este editor é disponível para Solaris, SunOS, Linux e outras plataformas UNIX.

### 13.4.9 Ferramenta MultiTalk

A aplicação talk UNIX permite que dois usuários Internet se falem via uma interface baseada em texto. A tela do terminal é dividida em duas partes, uma parte para cada participante. Se o primeiro usuário digita um texto, este texto aparece na parte superior da tela, o texto vindo do segundo usuário aparece na parte inferior. Este cenário pode ser facilmente estendido para mais que dois usuários para formar uma aplicação chat. Neste caso, existe apenas uma janela, e o texto de todos os usuários é apresentado na ordem em que eles chegam em um servidor. Tradicionalmente, estas aplicações chat são baseadas em TCP. Um servidor central estabelece uma conexão TCP ponto-a-ponto para cada cliente e transfere o dado. Este método significa que o servidor deve copiar o dado n vezes para distribuir este para n participantes via n conexões TCP.

A aplicação MultiTalk [Anderson] (<http://pipkin.lut.ac.uk/~ben/multitalk/>) permite conversações baseadas em texto entre vários usuários, baseado no IP multicast. Similar as ferramentas de áudio e videoconferência, dois portos são configurados, um para enviar dados e outro para transferir dados de controle. Como o IP multicast é baseado no serviço de transporte UDP não confiável, mecanismos para garantir a troca de texto confiável deveria ser implementado com a aplicação MultiTalk. MultiTalk atualmente não fornece um serviço confiável; sua confiabilidade portanto depende da confiabilidade da rede.

## 13.5 WWW

A ferramenta que tornou popular a Internet é incontestavelmente o WWW, o *World Wide Web*, em uma palavra a Web. A palavra Web designa em inglês uma teia de aranha e *World Wide Web* designa então uma teia cobrindo o mundo inteiro. Pela Web você pode visitar uma exposição, ler jornais, aprender inglês, pedir uma Pizza. Além de que todos os conectados podem ter suas páginas Web pessoais.

Os softwares de leitura da Web são chamados de navegadores ou browsers. O primeiro leitor gráfico da web conhecido pelo público foi Mosaic de NCSA (*National Center for Supercomputing Applications*) da Universidade de Illinois. Os mais utilizados hoje em dia são o Netscape Navigator e o Microsoft Internet Explorer. Certos navegadores propõem a possibilidade de ajuntar componentes de software, chamados Plug-Ins, capazes de efetuar tarefas evoluídas. Além disso, os navegadores reconhecem hoje as linguagens Java e VRML.

O WWW é um sistema mídia sob-demanda: nós temos os clientes, dotados de navegadores, e os servidores WWW. Nós encontramos hoje navegadores um pouco multimídia, que permitem a apresentação até de seqüências de vídeo. Mas no momento, a maioria dos navegadores não suportam mídias contínuas tempo-real.

### **Clientes e servidores no WWW**

A linguagem permitindo escrever as páginas Web é o HTML (*HyperText Markup Language*). De um ponto de vista técnico, o WWW liga servidores que enviam páginas HTML a computadores dotados de navegadores. O protocolo de comunicação entre os navegadores e os servidores é o HTTP (*HyperText Transfert Protocol*).

O HTTP é um protocolo a nível de camada de aplicação que roda sob TCP. HTTP usa um mecanismo cliente-servidor: um cliente tal como um navegador WWW pergunta informações a um servidor WWW usando o método *GET request*, *HEAD* obtém meta-informações tal como datas de modificação. *PUT* transfere informações do cliente ao servidor.

Concluindo, o WWW pode ser definido como uma tela de aranha de servidores de informações ligados uns aos outros por ligações físicas (a rede física) e por ligações lógicas (os links hipertextos). Estes links hipertextos permitem viajar de um servidor a outro sobre a rede Internet.

### 13.5.1 Problemas da Web para multimídia

O HTTP é um protocolo que utiliza o TCP/IP. Em outras palavras, o HTTP garante a conectividade lógica da comunicação, mas não fornece garantias temporais. Além desta, o WWW hoje apresenta algumas outras limitações.

Primeiro, mídias contínuas, tal como vídeo e áudio, são usadas de maneira muito limitada. O WWW trata cada mídia ou peça de informação como um arquivo estático. Cada arquivo é buscado do servidor e armazenado inteiramente no cliente local antes de ser apresentado ao usuário (exceto que nas novas versões de browsers suportam apresentação progressiva limitadas para texto e imagens). Para áudio e vídeo, apresentadores externos são chamados quando o arquivo de áudio ou vídeo são obtidos e armazenados localmente. Por causa deste método de operação, um grande conjunto de espaço de armazenamento no cliente é necessário, e o tempo de resposta do sistema é quase inaceitável. Além disso, o usuário não é capaz de buscar e ver arquivos de áudio e vídeo cujo tamanho é maior que o espaço de disco alocado no cliente (a transmissão será interrompida quando o espaço alocado é completo). Por exemplo, um espaço de armazenamento de 120 MBytes e um atraso (tempo de resposta) de 480 segundo são necessários para uma seqüência de vídeo de 10 minutos compactado a 1,6 Mbps assumindo que a taxa de transmissão da rede seja de 2 Mbits/s.

Segundo, não é possível obter e apresentar múltiplas mídias simultaneamente e sincronamente (exceto texto com imagens em linha). Assim o que o usuário vê são imagens sem som ou sons sem imagens.

Terceiro, os apresentadores externos geralmente apresentam vídeo em uma taxa determinada pela carga e poder da CPU da máquina cliente, ao invés de na taxa natural de quadros do vídeo.

Por causa das limitações apresentadas, o potencial dos documentos multimídia com sons, imagens, gráficos, texto, etc. não é completamente utilizado. Para isto o WWW deveria satisfazer os requisitos para acesso, transmissão, e apresentação de documentos multimídia. Os requisitos mais importantes são:

- A latência do sistema (tempo de resposta) deveria ser curto. Isto significa que o dado deveria ser apresentado enquanto ele está chegando no cliente (transferência tempo-real).
- A continuidade e suavidade das mídias contínuas deveria ser mantida.
- Mídias múltiplas deveriam ser acessadas, transmitidas, e apresentadas simultaneamente e suas relações temporais e espaciais deveriam ser mantidas.

Para fazer com que o WWW suporte o acesso, transmissão e apresentação de documentos multimídia, os seguintes pontos devem ser estudados:

- O HTML atual deve ser estendido e/ou um script (como JavaScript ou VisualBasic) deve ser provido de maneira que um link possa fazer referência a múltiplos fluxos de mídias e suas relações temporais e espaciais possam ser especificadas. Atualmente a inclusão de Applets Java em páginas HTML resolve parcialmente este problema. De qualquer maneira estes scripts têm desvantagens bem conhecidas: conteúdo baseados em script é difícil de produzir e manter; é difícil construir máquinas de busca, ferramentas de convenção e outras ferramentas automatizadas para linguagens scripting.
- O browser deveria ser estendido para suportar a apresentação simultânea de fluxos de mídia simultaneamente. A apresentação deveria ser escalonada de acordo com as relações temporais e espaciais especificadas para cada documento multimídia. Para obter isto, a arquitetura de hardware e sistema operacional deve suportar multimídia.
- Os servidores deveriam acessar do dispositivo de armazenamento e liberar dados na mesma taxa que a taxa de apresentação no cliente.
- O HTTP atual é baseado no TCP/IP, que não é ideal para comunicações de mídias contínuas. Um sistema de transporte tempo-real ou multimídia deveria ser usado para substituir o TCP/IP. Quando os browsers suportarem a nova API JMF

(Java Media Framework) [Sun, 97], áudios e vídeos poderão ser transmitidos utilizando RTP.

### 13.5.2 Suportes de áudio e vídeo na Web

Atualmente, a Internet não suporta muito bem comunicações multimídia devido aos tipos de rede e protocolos de transporte usados. Mas o RTP e o RSVP estão ganhando popularidade, e existem esforços para fornecer áudio e vídeo tempo-real no WWW. Nesta seção serão apresentados alguns exemplos de tentativa de inserir áudio e vídeo na Web.

#### **Video Mosaic (Vosaic)**

Um dos esforço da inserção de áudio e vídeo nos documentos Web é o Vosaic [Chen] (<http://www.vosaic.com>), abreviação de Video Mosaic. Vosaic é uma ferramenta que estende a arquitetura do NCSA Mosaic para o suporte de áudio e vídeo tempo-real. Sendo assim, Vosaic incorpora áudio e vídeo em tempo-real em páginas Web e o vídeo é apresentado no lugar.

Nesta ferramenta, a transferência de áudio e vídeo é feita em tempo-real, assim não existe a latência de acesso a informação. Para permitir a inclusão de áudios e vídeos nas páginas Web, o HTML e URL foram estendidos. A especificação URL estendida suporta protocolos de transmissão MBone usando a chave mbone (p.e. mbone://224.2.252.51:4739:127:nv<sup>30</sup>) e protocolos para mídia contínuas sob-demanda usando a chave cm (p.e. cm://showtime.ncsa.uiuc.edu:8080:mpegvideo:puffer.mpg). A incorporação em linha de áudio e vídeo em HTML necessita a adição de dois construtores na sintaxe HTML. Segmentos de áudio e vídeo em linha são especificados da seguinte forma: <video src = "address:port/filepath option=cyclic|control"> e <audio src = "address:port/filepath option=cyclic|control">. Src especifica o servidor de informação incluindo endereço e número do porto.

Documentos HTML contendo áudio e vídeo integrados são caracterizados por uma variedade de protocolos de transmissão, formatos de codagem de dados, e mecanismos de controle de dispositivo (p.e. monitor gráfico, controle de dispositivo de áudio, controle de placa de vídeo). Na atual implementação do Vosaic, a transmissão de textos e imagens são suportadas por TCP, apresentação tempo-real de áudio e vídeo usam VDP, e RTP é usado para permitir a compatibilidade com as transmissões de conferência da MBone. VDP [Chen] é um protocolo RTP aumentado com tolerância a falhas embutidas para a transmissão de áudio e vídeo sobre o WWW.

#### **XML**

XML (*Extensible Markup Language*) [Bray, 98] [FAQ, 98] é uma linguagem projetada para permitir o uso do SGML no WWW. Ela foi desenvolvida por um grupo de trabalho XML formado pelo W3C (*WWW Consortium*) (<http://www.w3.org>) em 1996. Ela se chama expansível pois ela não tem um formato fixo como o HTML.

Uma linguagem markup regular define um modo de descrever informações em certas classes de documento. HTML é um destes tipos de documento: ele define um único e fixo tipo de documento com markup que deixa você descrever uma classe comum de relatório simples, com títulos, parágrafos, listas, ilustrações e outros além de fornecer provisões para hipertexto e multimídia.

XML em si não é uma simples linguagem de markup, XML é uma meta-linguagem permitindo que você crie sua própria linguagem de acordo com SGML. XML deixa você particularizar linguagens markup para muitas classes de documento. Ela pode fazer isto pois ela é escrita em SGML. SGML é considerada uma língua mãe usada para descrever centenas de diferentes tipos de documentos em muitos campos da atividade humana. XML é projetada para facilitar o uso da SGML diretamente na Web: para facilitar a definição de documentos, para facilitar a criação e gestão de documentos definidos em SGML, e para facilitar a transmissão e compartilhamento sobre a Web. XML é portanto

---

<sup>30</sup> Codifica uma transmissão MBone no endereço 224.2.252.51, no porto 4738, com fator de tempo de vida de 127, usando o formato de vídeo nv.

uma versão abreviada do SGML que é completamente descrito na especificação XML [Bray, 98]. Ela omite as partes mais complexas e menos usadas do SGML facilitando a escrita e entendimento, e mais adaptada para transporte e interoperabilidade na Web.

Em resumo, a meta do XML é habilitar o SGML genérico a ser servido, recebido e processado na Web da mesma maneira que é possível com HTML hoje em dia. Por esta razão, XML está sendo projetado para facilitar a implementação e para interoperabilidade com o SGML e HTML.

O XML não foi desenvolvido para ser um HTML++, mas como sendo um SGML. XML e HTML devem co-existir. Você pode se perguntar porque não foi feito uma extensão do HTML? HTML já está sobrecarregado com dezenas de invenções interesses de diferentes companhias mas frequentemente incompatíveis, pois HTML fornece apenas uma maneira de descrever sua informação. XML permite que grupos de pessoas ou organizações criem suas próprias linguagens markup para transferência de informação no seu domínio (música, química, eletrônica, finanças, lingüística, matemática, história, engenharia, etc.)

HTML está no limite de sua utilizada como um meio de descrever informações, e por enquanto ela continua a ser importante para quantidade de informação que ela representa atualmente. Várias novas aplicações necessitam uma infraestrutura mais robusta e flexível.

## **SMIL**

Após um workshop W3C sobre “Multimídia Tempo-Real e a Web” em outubro de 1996, o W3C estabeleceu um grupo de trabalho em multimídia sincronizada em março de 1997. Este grupo trabalha no projeto de uma linguagem declarativa para descrever apresentações multimídia na Web usando XML. Esta linguagem foi chamada de SMIL (pronunciada *smile*) [W3C, 99] abreviação de Linguagem de Integração de Multimídia Sincronizada. Um primeiro draft (primeiro estágio de uma normalização) desta linguagem foi liberado em novembro de 1997. Documentos SMIL são documentos XML.

A idéia básica da linguagem SMIL é nomear componentes do documento como texto, imagens, áudio e vídeo como URIs (*Uniform Resource Identification* - um termo mais genérico que URL usado no momento na Web) e escalar suas apresentações ou em paralelo ou em seqüência. Uma apresentação SMIL típica tem as seguintes características:

- A apresentação é composta de vários componentes que são acessíveis via URIs, p.e. arquivos armazenados em um servidor Web.
- Os componentes têm diferentes tipos de mídia de apresentação, tal como áudio, vídeo, imagem, texto.
- Os instantes de início e fim dos diferentes componentes são especificados relativo a eventos em outros componentes. Por exemplo, na apresentação de slides, um slide particular é apresentado quando um narrador inicia a fala acerca do slide.
- Controle familiares via botões tal como parar, avanço e retrocesso rápidos e “rebobinar” permite ao usuário controlar a apresentação.
- Funções adicionais para acesso randômico (a apresentação pode ser partida de qualquer lugar) e apresentação “câmera lenta” (apresentação em velocidade reduzida).
- O usuário pode seguir hiper-links embutidas na apresentação.

SMIL tem sido projetado de maneira a facilitar a autoria das apresentações com um editor de texto. A chave do sucesso do HTML foi que documentos hipertextos poderiam ser criados sem a necessidade de ferramentas de autoria sofisticadas. SMIL permite o mesmo para hipermídia sincronizada.

O atual protocolo de transporte usado pela Web, o HTTP, é utilizável para carga completa de arquivos. Além da versão atual do HTTP, multimídia sincronizada no W3C é baseado nos protocolos RTP e RTSP para transferência e controle de mídias contínuas.

### 13.5.3 Compartilhamento multicast de páginas Web

Em um cenário convencional de navegação na web, um cliente conecta-se a um servidor e recupera páginas web, incluindo texto e gráficos, pelo estabelecimento de uma seqüência de conexões TCP unicast.

Em uma videoconferência poderia-se imaginar o uso de navegadores web como um suporte de apresentação. Apresentações ou discussões interativas poderiam ser baseadas em uma distribuição multicast de páginas web em um grupo multicast. Em tal cenário, o palestrante navega através de um conjunto de páginas web de apresentação, e todos os outros participantes vêem a mesma página web como o palestrante. Vários projetos de pesquisa estão atualmente estudando este problema.

Três abordagens diferentes estão sendo usadas para permitir o compartilhamento multicast de páginas Web:

- A primeira abordagem modifica um dado navegador e adiciona funcionalidades multicast.
- Em uma segunda abordagem, applets Java são usadas para realizar a difusão de dados e para controlar a apresentação do browser.
- A terceira abordagem não é limitada aos navegadores WWW e permite o compartilhamento de aplicações X baseado na distribuição multicast de dados.

#### ***Extensão multicast de navegadores Web***

Shared Mosaic [Kumar, 95] ([http://www.eit.com/goodies/software/share\\_mosaic/](http://www.eit.com/goodies/software/share_mosaic/)) é uma extensão do navegador WWW XMosaic da NCSA. Vários participantes podem surfar juntos através da web. Uma sessão Mosaic compartilhada é fracamente controlada. Qualquer um pode dar um tour aos sites preferidos e compartilhar páginas web com outros. Qualquer usuário pode clicar em um link ou abrir páginas web via hotlists, e os conteúdos das páginas web são distribuídas para os outros participantes através do IP multicast. O fraco controle pode levar a problemas de sincronização. A sincronização deve portanto ser efetivada por outros meios tal como uma ferramenta de áudio permitindo que os participantes negoçiem quem controla a sessão Mosaic compartilhado.

mMosaic [Dauphin] (<ftp://sig.enst.fr/pub/multicast/mMosaic/mMosaic-2.7b4m3.tar.gz>) também é um navegador derivado do NCSA XMosaic. Qualquer membro do grupo multicast pode distribuir ou receber páginas web. Cada membro do grupo tem uma janela de controle mostrando todos os membros da sessão. Se um membro desejar ver a página atual do outro membro, ele simplesmente clica no nome daquele membro da sessão.

#### ***Applets Java para Distribuição de dados multicast e controle de navegadores***

Enquanto ferramentas de distribuição de páginas WWW descritas acima são fortemente ligadas a um navegador específico tal como XMosaic, a abordagem implementada por mWeb [Parnes, 97] e WebCanal [Liao, 97] (<http://monet.inria.fr/>) é independente de navegador. As duas ferramentas são implementadas como applets Java e são baseadas em um protocolo multicast confiável para distribuir páginas web para grupos multicast.

mWeb age como uma espécie de gateway entre um navegador WWW e a MBone. Ele fornece funções para a distribuição de páginas WWW e precaching de arquivos a serem usados em uma sessão. mWeb salva todas as páginas recebidas no disco local. Quando uma apresentação baseada em páginas WWW é distribuída sob a MBone, o apresentador envia mensagens de apresentação dedicadas para os membros do grupo, indicando quais páginas cached o navegador deveria apresentar. Se uma página é encontrada na cache, ela é pedida pelo cliente bcast mWeb. mWeb é baseado em um protocolo multicast confiável parecido com o SRM.

A aplicação WebCanal roda como um servidor de proxy HTTP para qualquer navegador WWW. WebCanal carrega páginas web quando um navegador emite um pedido para carregar uma página web. O documento é então trazido pela aplicação WebCanal como um servidor de proxy regular, mas o documento é então enviado para um grupo

multicast. Quando uma página web é recebida, WebCanal pede ao navegador para apresentar a página web mais recentemente recebida ou uma lista de páginas recebidas da qual o usuário pode selecionar uma página.

### Compartilhamento de aplicação no Sistema IRI

O sistema IRI (*Interactive Remote Instruction*) [Abdel-Wahab, 96] (Figura 133) foi projetado na Old Dominion University em Norfolk para ensino a distância. Ele integra áudio, vídeo, aplicações compartilhadas e vários utilitários colaborativos multiusuário com uma única interface com o usuário.

IRI distribui áudio, vídeo e apresentação do professor aos estudantes via IP multicast. Estudantes podem fazer notas da apresentação com utilitários especiais, apresentar cursos registrados, fazer questões, e apresentar a si próprio. Estudantes podem também falar privadamente com o professor ou em grupos menores de estudantes.



**Figura 133.** Sistema IRI

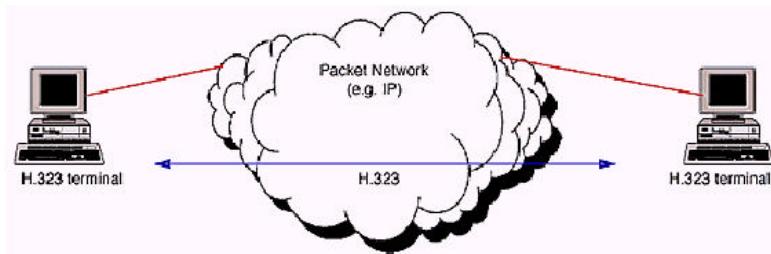
Uma característica interessante do sistema IRI é o componente compartilhamento de aplicação, que faz uso do IP multicast para fornecer uma boa escalabilidade. O processo XTV (*X Tools sharing process*) do sistema IRI permite o compartilhamento de aplicações X arbitrárias entre o professor e estudantes, por exemplo um navegador usado como uma ferramenta de apresentação. XTV intercepta os pedidos da aplicação X do navegador WWW e envia estes ao servidor local X e aos servidores X dos estudantes. Os estudantes têm exatamente o mesmo navegador visto pelo professor. Apenas uma pessoa, o professor ou um dos estudantes, mantém uma ficha. Apenas o mantenedor da ficha pode interagir com o navegador, por exemplo carregando páginas web ou seguindo links.

Os processos XTV trocam pacotes X, usando IP multicast. Mas, as aplicações X tal como navegadores WWW necessitam sistemas de comunicação confiáveis. Para isto é utilizado um protocolo multicast confiável (RMP) [Whetten, 94] entre o processo XTV e o serviço multicast baseado em UDP/IP não confiável [Abdel-Wahab, 96]. Isto é uma diferença significante de outras abordagens implementando aplicações X compartilhadas que executam sobre o confiável, mas não multicast, protocolo TCP. Em contraste, comunicações de áudio e vídeo não precisam de confiabilidade, portanto nesta ferramenta áudio e vídeo são difundidos sob UDP/IP.

## Capítulo 14 Padrão de Videoconferência H.323

### 14.1 Introdução

O H.323 é uma recomendação da União Internacional de Telecomunicações (ITU) que fornece um framework (componentes, protocolos, e procedimentos) de comunicação para áudio, vídeo e dados através de redes a comutação de pacotes. Redes baseadas em pacotes inclui redes locais, corporativas, metropolitanas e mundiais baseadas em IP (incluindo a Internet) ou IPX (*Internet Packet Exchange*).



**Figura 134.** Terminais H.323 em redes a pacotes

O H.323 possibilita uma variada gama de aplicações interativas multimídia tais como:

- Internet fone;
- Videoconferência em desktop;
- Computação colaborativa;
- Conferência de negócios;
- Ensino a distância;
- Voz sobre IP (VoIP);
- Aplicações de suporte e Help Desk (localização em uma organização onde chamadas podem ser recebidas para resolução de problemas e/ou um conhecimento básico das informações acerca dos produtos/serviços de uma companhia).
- Compra/venda interativas;
- Outras.

O padrão H.323 é especificado pelo Grupo de Estudos 16 do ITU. A versão 1 da recomendação H.323 – sistemas de telefonia visual e equipamentos para LANs que não fornecem garantias de QoS - foi aprovada em 1996. Ela foi, como o nome sugere, duramente voltada para a comunicação multimídia em um ambiente de rede local e não fornece garantias de QoS.

A emergência das aplicações de voz-sobre-IP (VoIP) levou a uma revisão da especificação H.323. A ausência de um padrão para VoIP resultou em produtos que não eram compatíveis. Com o desenvolvimento do VoIP, novos requisitos surgiram, tal como fornecer comunicação entre telefones baseados em PC e telefones tradicionais em redes a comutação de circuito. Tais requisitos forçaram a necessidade de um padrão para a telefonia IP. Uma nova versão do H.323, chamada de versão 2 – sistemas de comunicação multimídia baseadas em pacotes - foi definida para acomodar estes requisitos adicionais e foi aprovada em janeiro de 1998. Novas características estão sendo adicionadas ao padrão H.323, que deve resultar na sua versão 3. Entre estas novas características estão incluídas fax sobre redes a pacote, comunicação gatekeeper-gatekeeper (visto mais adiante) e mecanismos de conexão rápida.

O H.323 é parte de uma série maior de padrões de comunicação que habilitam videoconferência através de uma gama de redes. Conhecido como H.32X, esta série inclui o H.320 e o H.324, que endereçam comunicações ISDN e PSTN (*Public Switched*

*Telephone Network* - Rede Comutada de Telefonia Pública), respectivamente. Os padrões associados são os seguintes:

- H.320 – o padrão de videoconferência ISDN original.
- H.323 – uma extensão do H.320 para videoconferência em redes locais, compostas dos seguintes padrões:
  - H.225 – protocolo de controle de chamada: RAS (*Registration, Admission and Status*). Especifica mensagens para controle de chamada incluindo sinalização, registro e admissão, empacotamento e sincronização de fluxos de mídias.
  - H.245 – protocolo de controle de mídia. Especifica mensagem para abertura e fechamento de canais para fluxos de mídias e outros comandos.
  - H.261 – codificação de vídeo a taxa  $\geq 64$  kbps.
  - H.263 - codificação de vídeo a taxa  $<64$  kbps.
  - G.711 – codificação de áudio PCM 56/64 kbps
  - G.722 - codificação de áudio para 7 KHz a 48/56/64 kbps
  - G.723 - codificação de voz para 5.3 e 6.4 kbps
  - G.728 - codificação de voz para 16 kbps
  - G.729 - codificação de voz para 8/13 kbps
- H.324 – uma extensão do H.320 para videoconferência sobre linhas PSTN.
- H.322 – serviço de comunicação multimídia sobre LAN com suporte de QoS.
- Q.931 – serviço de sinalização de chamadas para estabelecimento e terminação de chamadas.
- T.120 – protocolo de conferência de dados tempo-real.

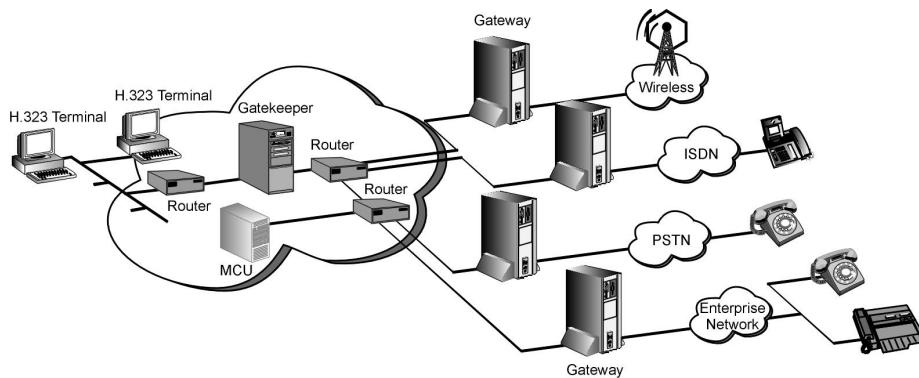
A adoção do padrão H.323 trás uma série de benefícios. Os principais são:

- **Codificadores/Decodificadores padrões**: ele estabelece padrões para compressão de descompressão de fluxos de áudio e vídeo, assegurando que equipamentos de diferentes fabricantes tenham alguma área de suporte comum.
- **Interoperabilidade**: ele permite aos usuários fazerem uma conferência sem se preocupar com a compatibilidade dos vários pontos finais. Além de assegurar que o receptor possa decodificar a informação, o H.323 estabelece métodos para clientes receptores comunicarem suas capacidades ao remetente.
- **Independência de rede**: ele foi projetado para executar sobre arquiteturas de rede comuns. Na medida que evolui a tecnologia de rede e melhoram as técnicas de administração da largura de banda, soluções baseadas em H.323 poderão tirar proveito dessas capacidades aumentadas.
- **Independência de Plataforma e Aplicação**: ele não é dependente de hardware ou sistema operacional. Plataformas aderentes ao H.323 estão disponíveis em muitos tamanhos e formas, elas incluem computadores pessoais com capacidades de áudio e vídeo, plataformas dedicadas, telefones com IP habilitado e adaptadores para TV a cabo.
- **Suporte multiponto**: permite a realização de conferências multiponto.
- **Administração de Largura de Banda**: os tráfegos de áudio e vídeo geralmente geram uma grande taxa de bits e podem congestionar a rede corporativa. O H.323 trata este assunto provendo administração de largura de banda: os gerentes de rede podem limitar o número de conexões H.323 simultâneas dentro da rede ou a largura de banda disponível para aplicações H.323. Estes limites asseguram que o tráfego crítico não será prejudicado.
- **Supporte Multicast**: ele suporta transporte multicast em conferências multiponto, permitindo o uso mais eficiente da largura de banda uma vez que todas as estações no grupo multicast lêem um único fluxo de dados.

## 14.2 Arquitetura H.323

A recomendação H.323 cobre os requisitos técnicos para serviços de comunicação de áudio e vídeo em LANs que não fornece garantias de Qualidade de Serviço (QoS). H.323 referencia a especificação T.120 para conferência de dados. O escopo do H.323 não inclui a LAN em si ou a camada de transporte que pode ser usada para conectar as várias LANs. Apenas os elementos necessários para a interconexão com a rede a comutação de circuitos (*Switched Circuit Network - SCN*) estão dentro do escopo do H.323.

O H.323 especifica quatro componentes principais: Terminais, Gateways, Gatekeepers e MCUs (*Multipoint Control Units*) como mostra a Figura 135. Terminais, Gateways e MCUs são classificados como dispositivos terminais (*endpoints*) – são dispositivos que podem iniciar e receber chamadas. Outras tecnologias importantes associadas ao H.323 são os CODECs usados para codificar e decodificar transmissões de vídeo e áudio.



**Figura 135.** Componentes principais da arquitetura H.323

### 14.2.1 Terminais

Terminais H.323 são clientes (*endpoints*) que provem comunicação multimídia bidirecional em tempo real, executando a pilha H.323 e as aplicações multimídia. Ele suporta a comunicação de áudio e pode opcionalmente suportar a comunicação de vídeo e dados (prove funcionalidades tais como “chat” em texto, Quadro Branco - “white boarding - WB” compartilhado). Como o serviço básico fornecido por um terminal H.323 é a comunicação de áudio, um terminal H.323 é um elemento importante nos serviços de telefonia IP.

A principal meta do H.323 é a interoperabilidade com outros terminais multimídia. Terminais H.323 são compatíveis com: terminais H.324 das redes PSTN e redes sem fio; terminais H.310 na B-ISDN ou ISDN; terminais H.321 em B-ISDN; e terminais H.322 em redes locais com garantias de QoS.

Terminais podem ser implementados como softwares baseados em PCs ou em dispositivos stand-alone tais como vídeo fones, ou Internet fones. Atualmente, a grande maioria dos terminais são implementados como software de PCs. Como estes terminais não são rigidamente especificados pela recomendação H.323, terminais baseados em PCs normalmente usam uma placa de som, microfone e alto-falantes ou fones de ouvido. Existem muitos fabricantes que fornecem placas de PC onde pode ser conectado um telefone padrão.

### 14.2.2 Gateways

Um gateway H.323 é um *endpoint* que provê comunicação bidirecional em tempo-real entre terminais H.323 e outros terminais padrão ITU (isto é telefones), ou outro Gateway H.323. Eles executam a tradução de controle de chamada e de conteúdo necessário nas chamadas para converter uma chamada da rede de comutação de pacotes formato H.323 para o formato das redes de comutação a circuito PSTN e vice-versa.

Gateways são componentes opcionais na rede. Eles são necessários quando se está conectando outros tipos de terminais tais como telefones comuns ou terminais H.320 (Vídeo conferência ISDN).

Um uso comum para gateways é para transporte de tráfego de telefone de longa distância sobre uma rede IP. As empresas podem reduzir seus custos de chamadas de longa distância usando gateways para transporte de voz sobre as redes de dados existentes. Neste tipo de aplicação, o usuário discia um número de acesso local para conectar-se ao gateway, e discia o número de destino. O Gateway local faz a conexão IP para o outro gateway localizado na área de destino chamada. O Gateway remoto completa a chamada discando o número local. A Figura 136 ilustra uma chamada de gateway-para-gateway.



**Figura 136.** Chamada de Gateway-para-Gateway

Gateways são usados também para fornecer uma interface entre clientes H.323, tal como Microsoft NetMeeting e a rede PSTN, como visto na Figura 137. Esta configuração pode ser usada por um “call center” para permitir que usuários on-line possam contatar um atendente a partir do web site.

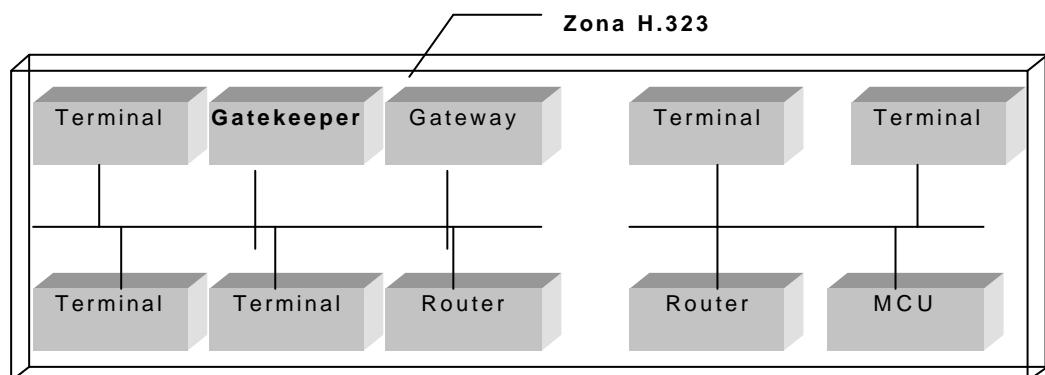


**Figura 137.** Chamada de Terminal-para-Gateway

#### 14.2.3 Gatekeepers

Um gatekeeper pode ser considerado o cérebro da rede H.323. Ele é ponto de foco para todas as chamadas dentro da rede H.323. Embora eles não são necessários, os gatekeepers fornecem serviços importantes tal como endereçamento, autorização e autenticação de terminais e gateways; gerenciamento de largura de banda; contabilidade; faturamento; e debitação.

O gatekeeper (Figura 138) é um componente opcional do padrão H.323. Ele provê diversos serviços importantes e deverá fazer parte da maioria das redes H.323 no futuro.



**Figura 138.** Uma rede H.323 com Gatekeeper

A seguir serão apresentados os serviços fornecidos por um gatekeeper.

### ***Gerenciamento de Zona***

Gatekeepers introduzem o conceito de **zonas**. Uma zona é uma coleção de todos os terminais, gateways e MCUs gerenciadas por um gatekeeper. Uma zona deve incluir pelo menos um terminal, e pode ou não incluir gateways ou MCUs. Uma zona tem um e somente um gatekeeper. Ele é independente da topologia da rede e pode incluir múltiplos seguimentos de LAN.

### ***Controle de admissão***

O gatekeeper concede ou nega admissão à zona para endpoints H.323. Isto pode estar baseado na disponibilidade de largura de banda ou algum outro critério.

### ***Controle de Largura de Banda***

O gatekeeper também controla todos os pedidos para mudanças de largura de banda. Isto pode estar baseado na contenção de largura de banda na zona, em uma limitação de largura de banda específica estabelecida pelo administrador da zona, ou algum outro critério.

### ***Tradução de endereços***

O gatekeeper mantém um banco de dados dos endpoints associados a sua zona. Em casos de pedido de endpoints, ele traduz pseudônimos de endpoint (por exemplo, hostname, endereço de email) para endereços de rede (por exemplo, endereços IP) e traduz endereços externos (número de telefone formato E.164) para endereços de rede. O endpoint usa então o endereço de rede para iniciar uma chamada a outro endpoint.

### ***Autorização de Chamada***

O gatekeeper pode escolher rejeitar chamadas de um endpoint devido à falha de autorização (por exemplo, você escolhe não aceitar chamadas de H.323 de um concorrente).

### ***Administração de Largura de Banda***

O gatekeeper pode limitar o número de endpoints H.323 que podem acessar a rede simultaneamente, rejeitando novas chamadas devido a limitações de largura de banda. Administração de largura de banda é permite controlar a banda de rede limitando o número de novas chamadas. Isto difere de controle de largura de banda, definido acima, o qual refere-se a pedidos de mudanças de largura de banda em chamadas existentes.

### ***Administração de Chamadas***

O gatekeeper mantém um banco de dados de chamadas de H.323 em andamento. Estes dados são usados para localizar o estado de endpoints para uma variedade de propósitos, inclusive localização de chamada e administração de largura de banda.

#### **14.2.4 Multipoint Control Units**

MCUs fornecem suporte para conferências de 3 ou mais terminais H.323. Todos os terminais participantes na conferência estabelecem uma conexão com o MCU. O MCU gerencia recursos de conferência, negocia entre terminais para solicitar o CODEC (codificador/decodificador) de áudio e vídeo a ser usado, e pode também manipular o fluxo multimídia.

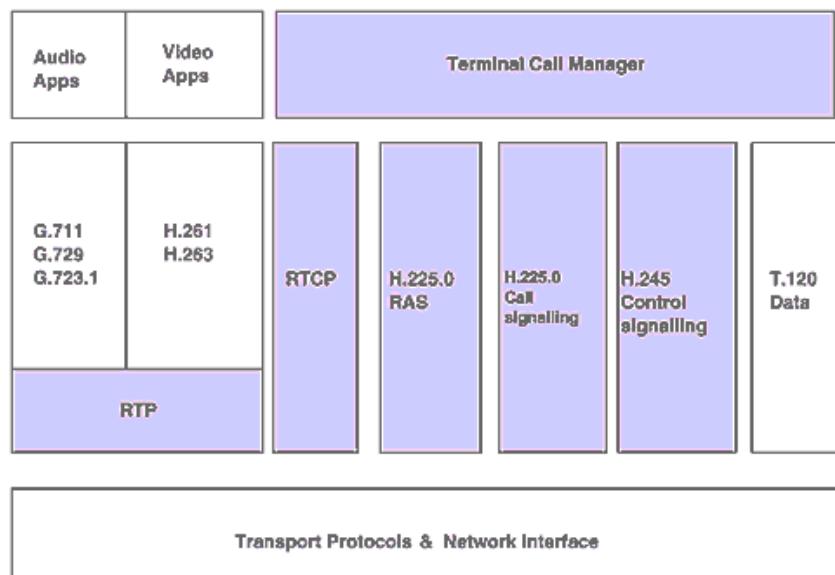
Os gatekeepers, gateways e MCUs são componentes logicamente separados do padrão H.323, mas podem ser implementados como um dispositivo físico único.

## **14.3 Protocolos Definidos pelo H.323**

Os protocolos especificados pelo H.323 são listados abaixo (Figura 139):

- CODECs de áudio;

- CODECs de vídeo;
- H.225 registro, admissão e status (RAS);
- H.225 sinalização de chamada;
- H.245 controle de sinalização;
- RTP (Real-Time Transfer Protocol);
- RTCP (Real-Time Control Protocol);



**Figura 139.** Protocolos H.323

#### 14.3.1 CODEC de Áudio

Um CODEC de áudio codifica o sinal oriundo de um microfone para transmissão por parte de um terminal H.323 e decodifica o código de áudio recebido e o envia ao alto-falante no terminal H.323 receptor.

Os algoritmos de compressão suportados pelo H.323 são todos padrões aprovados pelo ITU. Terminais H.323 devem suportar o padrão G.711 para compressão de voz. Suporte para os outros padrões de voz ITU é opcional.

As diferentes recomendações ITU para digitalização e compressão de fala refletem uma relação de compromisso entre qualidade da fala, taxa de bits, capacidade de computação e retardo do sinal. O G.711 geralmente transmite voz a 56 ou 64 kbps, bem dentro da largura de banda disponível na LAN, porém ele foi projetado originalmente para rede de taxa contínua de bits. Pelo fato de o padrão G.723 operar em taxa de bits muito baixa (5.3 e 6.3 kbps), ele está sendo altamente considerado como um CODEC obrigatório e tende a ser o CODEC predominante em aplicações H.323.

#### 14.3.2 CODEC de Vídeo

Um CODEC de vídeo codifica o sinal de vídeo oriundo de uma câmera para transmissão no lado do terminal H.323 emissor e decodifica o código de vídeo que é enviado para o display do terminal H.323 receptor.

Embora a capacidade de vídeo seja opcional, qualquer terminal H.323 que possibilite vídeo deve suportar o CODEC H.261; o suporte para H.263 é opcional. A informação de vídeo é transmitida a uma taxa não superior aquela selecionada durante a negociação de capacidades.

### 14.3.3 H.225 Layer

O H.225 layer descreve a camada mais baixa que encapsula e desencapsula os vários fluxos de dados para dentro de pacotes para transmissão sobre uma LAN. O H.225 também manipula o controle de chamadas para iniciar e finalizar chamadas entre terminais, gateways, MCUs e gatekeepers.

A camada H.225 inclui o Q.931, RAS e o RTP/RTCP.

### 14.3.4 H.225 Registro, Admissão e Status

Registro, admissão e status (RAS) é o protocolo entre os endpoints (terminais e gateways) e gatekeepers. O RAS é usado, entre outros, para:

- Descoberta do gatekeeper (GRQ). O processo de descoberta do gatekeeper é usado pelos endpoint H.323 para determinar o gatekeeper na qual o endpoint deve registrar. Esta descoberta pode ser feita estaticamente ou dinamicamente. Na descoberta estática, o endpoint conhece o endereço de transporte do seu gatekeeper a priori. No método dinâmico, o endpoint realiza um multicast da mensagem GRQ ("Who is my gatekeeper?") no endereço multicast de descoberta de gatekeepers. Um ou mais gatekeepers podem responder a com uma mensagem GCF ("I can be your gatekeeper").
- Registro do endpoint. Registro é o processo usado pelos endpoints para se juntar a uma zona e informar o gatekeeper dos endereços de alias e de transporte.
- Localidade do endpoint. É o processo pelo qual o endereço de transporte de um endpoint é determinado e dado seu nome alias ou endereço E.164.
- Controle de admissão, para restringir a entrada de um endpoint na zona.
- Controle da largura de banda.
- Desengajamento, onde o endpoint é desassociado de um gatekeeper e de sua zona.

As mensagens RAS são transportadas por um canal RAS que é não confiável (p.e. UDP). A troca de mensagens RAS deve ser associada com timeouts e numeradas.

### 14.3.5 H.225 Sinalização de Chamada

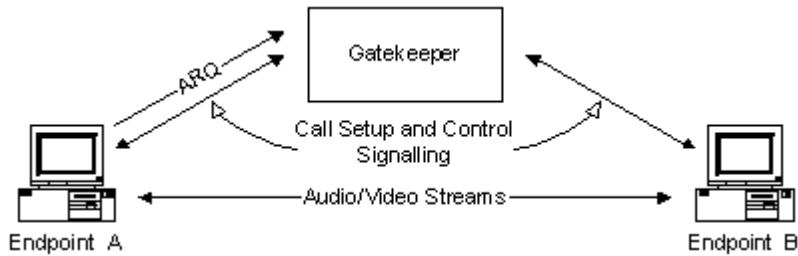
O H.225 sinalização de chamada é usado para estabelecer um canal entre dois endpoints H.323, na qual dados em tempo-real podem ser transportados. Isto é obtido pela troca de mensagens H.225 no canal de sinalização de chamada confiável (p.e. TCP em uma rede baseada em IP). O canal de sinalização de chamada é aberto entre dois pontos finais H.323 ou entre um ponto final e o gatekeeper.

O Canal de Sinalização de Estabelecimento de Chamada usa o Q.931 para estabelecer uma conexão entre dois terminais.

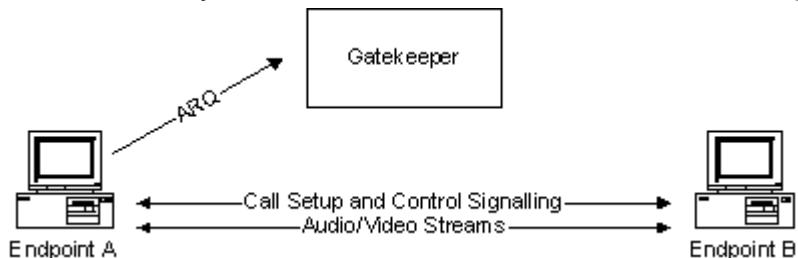
As mensagens H.225 são trocadas entre endpoints se não existam gatekeepers na rede H.323.

Quando um gatekeeper existir na rede, as mensagens H.225 são trocadas diretamente entre endpoints ou entre os endpoints após serem roteadas pelo gatekeeper:

- No modelo de chamada roteado, o endpoint originador da chamada (A) (Figura 140) envia um pedido admissão de zona ao gatekeeper via o canal RAS. O gatekeeper recebe as mensagens de sinalização de chamada sobre os canais de sinalização de chamada vindo de um endpoint e roteia elas para outro endpoint no canal de sinalização de chamada do outro endpoint. Isto permite ao gatekeeper monitorar o estado e controlar todas as chamadas dentro de sua zona.

**Figura 140.** Modelo de Chamada Roteada

- No modelo de chamada direta, o endpoint originador da chamada (A) (Figura 141) envia um pedido de admissão de zona (ARQ) para o gatekeeper. Durante a confirmação da admissão, o gatekeeper indica que os endpoints podem trocar mensagens de sinalização de chamada diretamente. O estabelecimento da chamada e a sinalização de controle são feitos diretamente entre os endpoints.

**Figura 141.** Modelo de Chamada Direta

#### 14.3.6 H.245 Sinalização de Controle

H.245 sinalização de controle é usado para trocar mensagens de controle fim-a-fim para governar a operação do ponto final H.323. Elas são transportadas sobre os canais de controle H.245. Este canal de controle é o canal lógico 0 e é permanentemente aberto, diferente dos canais de mídia.

Estes mensagens transportam informações relacionadas a:

- Troca de potencialidades: para que os terminais possam enviar aos seus pares a capacidade de receber e processar fluxos de mídia.
- Abertura e fechamento dos canais lógicos usados para transportar fluxos de mídia;
- Mensagens de controle de fluxo;
- Comandos e indicações gerais.

#### 14.3.7 RTP (*Real-Time Transport Protocol*)

Como visto anteriormente nesta apostila, o RTP fornece serviços de liberação de áudio e vídeo tempo-real fim a fim. Como já apresentado, o RTP fornece a identificação do tipo de dado transportado, numeração da seqüência, marcas temporais, e monitoração.

#### 14.3.8 RTCP (*Real-Time Transport Control Protocol*)

Como visto anteriormente nesta apostila, o RTCP é o protocolo parceiro do RTP e fornece serviços de controle. A principal função do RTCP é fornecer uma realimentação sobre a qualidade da distribuição dos dados. Outras funções RTCP incluem um identificador da fonte RTP.

#### 14.3.9 Dados

Data-conferência é uma capacidade opcional. Quando suportada, possibilita a colaboração através de aplicações tais como: Quadro branco, compartilhamento de aplicações e transferência de arquivos.

O H.323 suporta data-conferência através da especificação T.120. O padrão ITU, T.120 endereça conferências ponto-a-ponto e multiponto. Ele provê interoperabilidade no nível de aplicação, rede e transporte. Um sistema H.323 pode suportar dados incorporando capacidades T.120 nos clientes e unidades de controle multiponto. A MCU pode controlar e multiplexar as informações de conferência de dados.

A recomendação para suporte multicast no T.120, conhecida como T.125 Anexo A ou Protocolo de Adaptação Multicast, foi aprovado pelo ITU em janeiro de 1998.

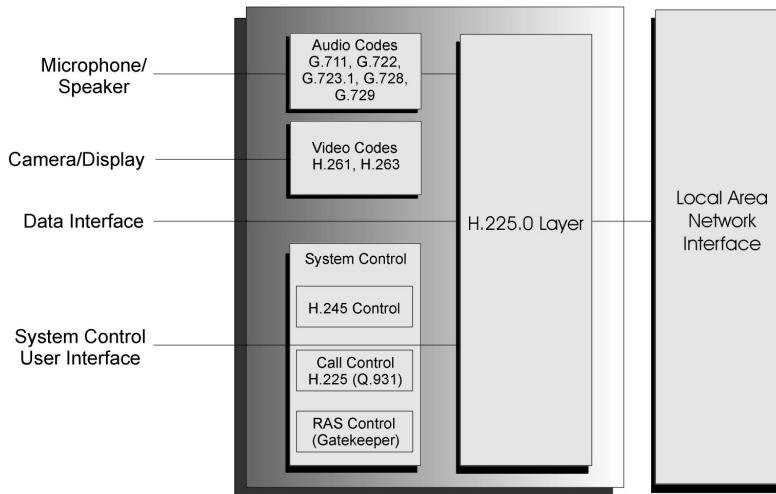
## 14.4 Características dos Terminais H.323

Terminais H.323 devem suportar os seguintes componentes:

- H.245 para troca de capacidades do terminal e criação de canais de mídia.
- H.225 para sinalização de chamada e configuração de chamada.
- RAS para registro e outros controles de admissão com o gatekeeper.
- RTP/RTCP para sequenciamento de pacotes de áudio e vídeo.

Terminais H.323 devem também suportar o CODEC de áudio G.711. Componentes opcionais em um terminal H.323 são CODECs de vídeo, protocolos de conferência de dados T.120 e MCU.

Um exemplo de um terminal é mostrado na Figura 142. O diagrama mostra a interface do equipamento do usuário, codec de vídeo e de áudio, equipamento de telemática, H.225.0 layer, funções de controle do sistema e a interface com a rede baseada em pacotes.



**Figura 142.** Terminal H.323

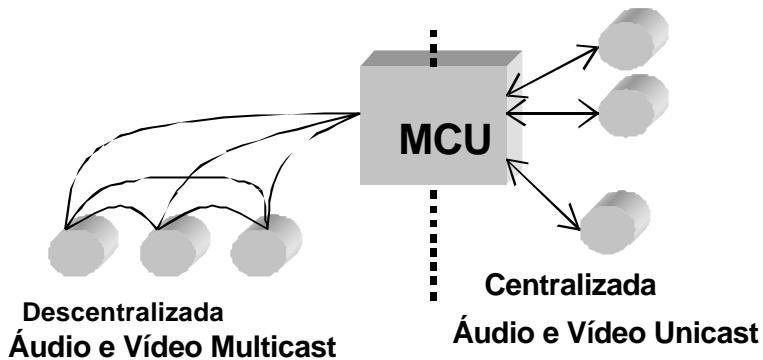
## 14.5 Unidades de Controle Multiponto (MCUs)

Uma MCU consiste de pelo menos um Controlador Multiponto (MC) e de Processador(es) Multiponto (MP) opcionais:

- O MC provê as funcionalidades de controle de chamada necessárias para conferência multiponto, incluindo negociação de facilidade comuns entre os endpoints.
- O MP, se existir, provê processamento (multiplexação ou chaveamento) de fluxos de diferentes mídias (áudio, vídeo, e/ou dados).

MCs e MPs funcionalmente podem ser incorporados em outras entidades do H.323 – terminais, gateways ou gatekeepers.

No H.323, conferências multiponto são controladas de diversos modos e configurações. A recomendação utiliza os conceitos de conferência centralizada e descentralizada, como descrito Figura 143.



**Figura 143.** MCU – Conferência Centralizada, Conferência Descentralizada.

A conferência centralizada multiponto requer a existência de uma MCU. Todos os terminais enviam áudio, vídeo, dados e fluxos de controle para a MCU no modo ponto-a-ponto. O MC de forma centralizada gerencia a conferência utilizando funções de controle do H.245 que também define as capacidades de cada terminal. O MP faz a multiplexação de áudio, distribuição de dados, e funções de seleção de vídeo tipicamente executados em conferências multiponto e envia o fluxo resultante para os terminais participantes. O MP pode também fornecer conversão entre diferentes CODECs e taxas de bits e pode também usar multicast para distribuir o vídeo processado.

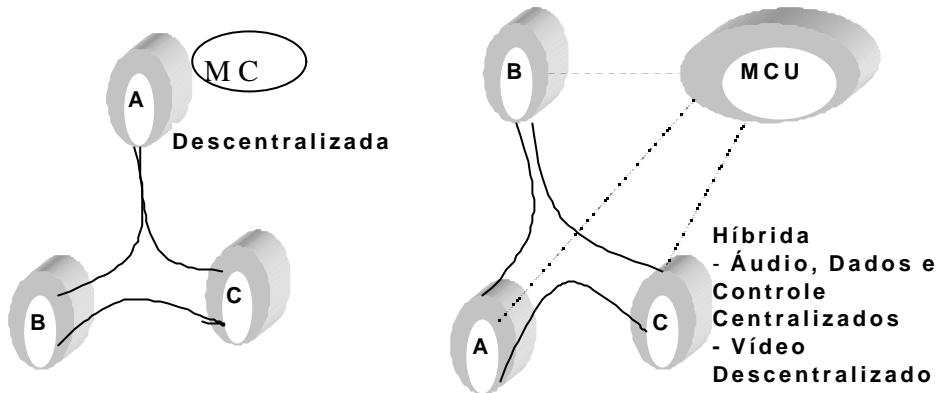
Conferências multiponto descentralizadas podem fazer uso da tecnologia multicast. Os terminais H.323 enviam áudio e vídeo multicast para outros participantes sem enviar dados para uma MCU. Note que o controle de dados multiponto continua sendo processado centralizadamente pela MCU, e as informações de controle H.245 são transmitidas no modo ponto-a-ponto para um MC. Os terminais receptores são responsáveis pelo processamento de múltiplos fluxos de áudio e vídeo. Os terminais usam o canal de controle H.245 para indicar ao MC quantos fluxos simultâneos de áudio e vídeo eles podem decodificar. O MP pode ainda prover seleção de vídeo e multiplexação de áudio em conferência multiponto descentralizada.

Conferências multiponto híbridas usam uma combinação de características de conferências centralizadas e descentralizadas. Sinais H.245 e fluxos de áudio e vídeo são processados através de mensagens ponto-a-ponto para a MCU. E a partir desta, os sinais de áudio e vídeo são transmitidos para os terminais H.323 participantes através de multicast. Uma vantagem da conferência centralizada é que todos os terminais H.323 suportam comunicação ponto-a-ponto. A MCU pode fazer múltiplas transmissões unicast para os participantes e não são requeridas capacidades especiais na rede. Por outro lado, a MCU pode receber múltiplos fluxos, multiplexar áudio e vídeo e transmitir fluxos multicast, conservando largura de banda da rede.

O suporte H.323 para conferências multiponto híbridas permite que alguns terminais estejam em uma conferência centralizada e outros em uma conferência descentralizada. Neste caso o MCU provê uma ponte entre os dois tipos. O terminal não está atento à natureza mista da conferência, somente ao modo de conferência na qual ele envia e recebe. Suportando ambas as abordagens multicast e unicast, o H.323 é adequado tanto às tecnologias de rede atuais quanto às futuras gerações de redes. O suporte a multicast faz uso mais eficiente da largura de banda, mas exige maior poder computacional dos terminais que devem multiplexar e selecionar seu próprio fluxo de recepção de áudio e vídeo. Adicionalmente, suporte a multicast é requerido nos roteadores e switches da rede.

Considere um exemplo onde uma conferência multiponto é realizada entre três clientes (Figura 144). Um terminal cliente (Cliente B) executa uma função MC. Todos os terminais podem usar multicast para participar em uma conferência descentralizada. Uma função MP em cada nó irá multiplexar e apresentar os sinais de áudio e vídeo destinados ao

usuário. Esta abordagem minimiza a necessidade de recursos especializados na rede. Entretanto, a rede deve ser configurada para suportar multicast.

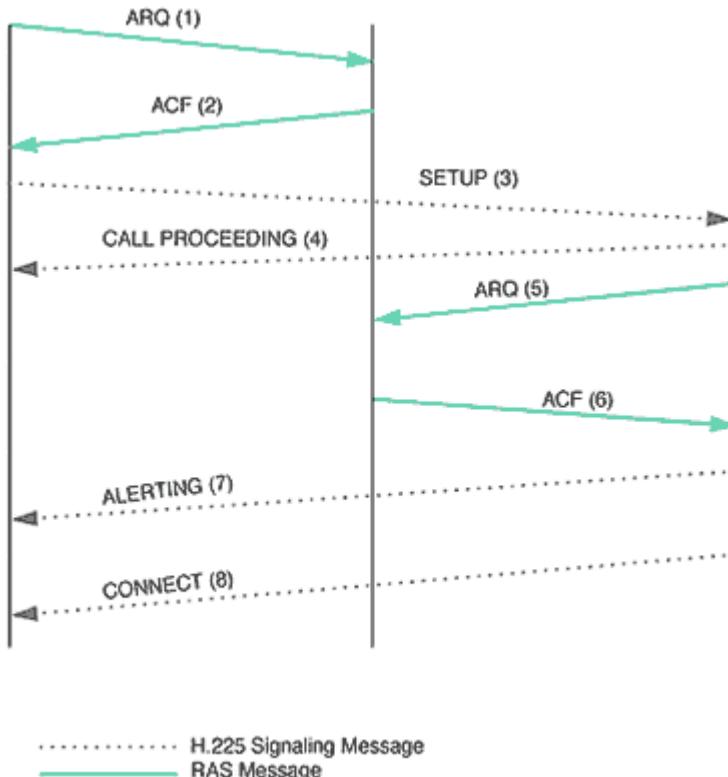


**Figura 144.** Conferência Descentralizada, Conferência Híbrida

Uma MCU pode ser usada para manipular somente áudio, dados e funções de controle. Nesta configuração o vídeo pode ser multicast conservando largura de banda da rede. O MCU pode ser um sistema dedicado ou um terminal com capacidades extras.

## 14.6 Procedimentos de Conexão

Esta seção descreve os passos envolvidos na criação de uma chamada H.323, a comunicação multimídia e a liberação da chamada. Este exemplo contém dois terminais H.323 (T1 e T2) conectados a um gatekeeper. Neste exemplo considera-se o uso modelo de chamada direta e que os fluxos multimídia utilizam o RTP. A Figura 145 ilustra o estabelecimento da chamada H.323.

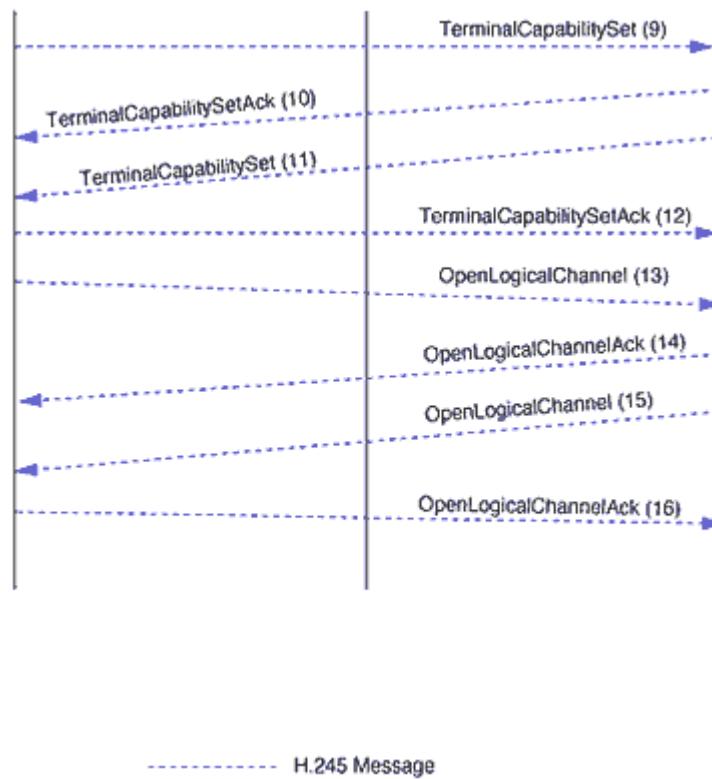


**Figura 145.** Estabelecimento da chamada H.323

O procedimento do estabelecimento da chamada H.323 é o seguinte:

- (1) T1 envia uma mensagem RAS ARQ no canal RAS para o gatekeeper para registro. T1 solicita o uso do modelo de chamada direta.
- (2) O gatekeeper confirma a admissão de T1 pelo envio de ACF para T1. O gatekeeper indica no ACF que T1 pode usar o modelo de chamada direta.
- (3) T1 envia uma mensagem de configuração de sinalização de chamada H.225 para T2 solicitando uma conexão.
- (4) T2 responde com uma mensagem H.225 call proceeding para T1.
- (5) T2 envia uma mensagem RAS ARQ para se registrar no gatekeeper através do canal RAS.
- (6) O gatekeeper confirma o registro pelo envio de uma mensagem RAS ACF para T2.
- (7) T2 alerta T1 do estabelecimento da conexão pelo envio de uma mensagem de alerta H.225.
- (8) Então T2 confirma o estabelecimento da conexão pelo envio de uma mensagem conexão H.225, e a chamada é estabelecida.

A Figura 146 ilustra o fluxo de sinalização de controle H.323.

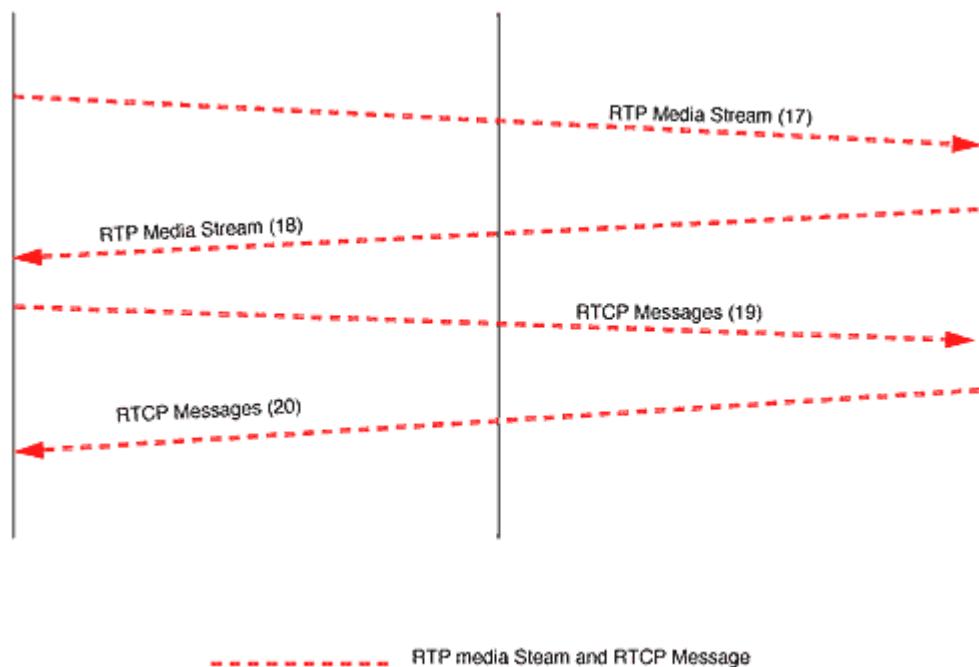


**Figura 146.** Fluxos da Sinalização de Controle H.323

- (9) O canal de controle H.245 é estabelecido entre T1 e T2. T1 envia uma mensagem H.245 TerminalCapabilitySet para T2 para informar suas capacidades.
- (10) T2 reconhece as capacidades de T1 pelo envio de uma mensagem H.245 TerminalCapabilitySetAck.
- (11) T2 informa suas capacidades para T1 pelo envio de uma mensagem H.245 TerminalCapabilitySet.
- (12) T1 reconhece as capacidades de T2 pelo envio de uma mensagem H.245 TerminalCapabilitySetAck.
- (13) T1 abre um canal de mídia com T2 pelo envio de uma mensagem H.245 openLogicalChannel. O endereço de transporte do canal RTPC é incluído nesta mensagem.

- (14) T2 reconhece o estabelecimento do canal lógico unidirecional de T1 para T2 pelo envio da mensagem H.245 openLogicalChannelAck. Incluído nesta mensagem está o endereço de transporte RTP alocado pelo T2 para ser usado pelo T1 para envio de fluxos de mídia RTP e o endereço RTCP recebido do T1 anteriormente.
- (15) Então, T2 abre um canal de mídia com T1 pelo envio de uma mensagem H.245 openLogicalChannel. O endereço de transporte do canal RTCP é incluído na mensagem.
- (16) T1 reconhece o estabelecimento do canal lógico unidirecional de T2 para T1 pelo envio da mensagem H.245 openLogicalChannelAck. Incluído nesta mensagem está o endereço de transporte RTP alocado pelo T1 para ser usado pelo T2 para envio de fluxos de mídia RTP e o endereço RTCP recebido do T2 anteriormente. Agora a comunicação bidirecional de fluxos de mídia está estabelecida entre T1 e T2.

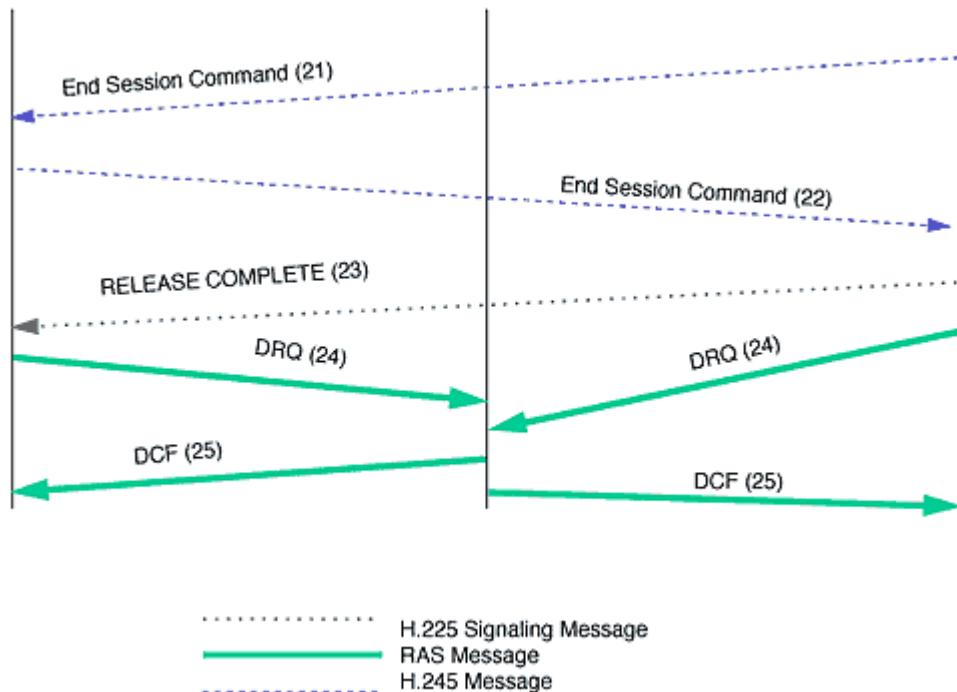
A Figura 147 ilustra envio do fluxo de mídia H.323 e o fluxo de controle de mídia.



**Figura 147.** Fluxos de mídia H.323 e fluxos de controle de mídia

- (17) T1 envia o fluxo de mídia encapsulado em pacote RTP para T2.  
 (18) T2 envia o fluxo de mídia encapsulado em pacote RTP para T1.  
 (19) T1 envia mensagens RTCP para T2.  
 (20) T2 envia mensagens RTCP para T1.

A Figura 148 ilustra a liberação da chamada H.323.

**Figura 148.** Liberação da chamada H.323

- (21) T2 inicia a liberação da chamada pelo envio de uma mensagem H.245 EndSessionCommand para T1.
- (22) T1 libera a chamada e confirma a liberação pelo envio da mensagem H.245 EndSessionCommand para T2.
- (23) T2 completa a liberação da chamada pelo envio da mensagem H.225 release complete para T1.
- (24) T1 e T2 desligam-se do gatekeeper pelo envio de uma mensagem RAS DRQ para o gatekeeper.
- (25) O gatekeeper libera T1 e T2 e confirma pelo envio das mensagens DCF para T1 e T2.

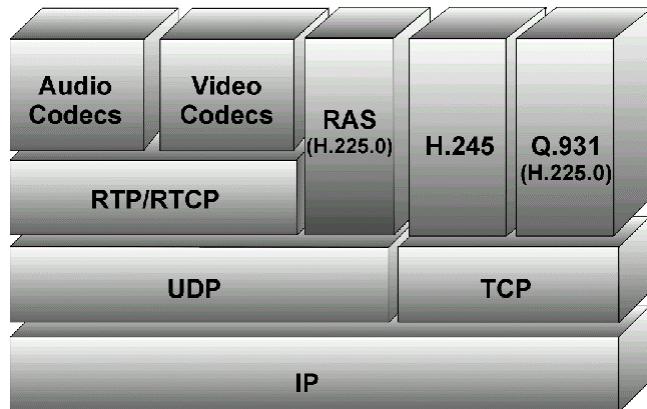
## 14.7 H.323 em Redes IP

O H.323 usa comunicações seguras/confiáveis e incertas/não confiáveis. Sinais de controle e dados requerem transporte seguro porque os sinais devem ser recebidos na ordem na qual foram enviados e não podem ser perdidos. Fluxos de áudio e vídeo perdem o valor com o tempo. Se um pacote de áudio chegar atrasado, pode não ter relevância ao usuário final. Sinais de áudio e vídeo usam transporte incerto/não confiável de modo mais eficiente.

A transmissão segura de mensagens usa o modo orientado a conexão para transmissão de dados. Na pilha IP, este tipo de transmissão é realizado com o TCP. A transmissão segura garante a seqüência, correção de erros, controle de fluxo na transmissão de pacotes, mas pode atrasar a transmissão e reduzir o throughput (vazão).

O H.323, cuja pilha de protocolos na rede IP é mostrada na Figura 149, usa os serviços seguros de conexão de fim-a-fim do (TCP) para o Canal de Controle H.245, para o Canais de Dados T.120 , e para o Canal de Sinalização da Chamada.

Dentro da pilha de IP, serviços incertos/não confiáveis são providos pelo protocolo UDP. Transmissão incerta/não confiável é um modo sem conexões que não garante nada mais que "melhor esforço" para entrega dos pacotes. O UDP oferece informação de controle mínima. O H.323 usa UDP para áudio, vídeo, e para o Canal RAS.



**Figura 149.** Pilha de Protocolos H.323

Em conferências com múltiplos fluxos de áudio e vídeo, o transporte não confiável via UDP usa IP Multicast e o Real-Time Protocol (RTP), desenvolvido pelo Internet Engineering Task Force (IETF) para manipular os fluxos de áudio e vídeo.

O Real-Time Control Protocol (RTCP) é usado para controlar o RTP. O RTCP monitora a qualidade de serviço, transporta informação sobre os participantes da sessão, e periodicamente distribui pacotes de controle que contêm informação de qualidade a todos os participantes da sessão através dos mesmos mecanismos de distribuição dos pacotes de dados.

Largura de banda suficiente para uma chamada multimídia é um recurso crítico e difícil de assegurar em redes de pacotes do tipo da Internet ou intranets corporativas. Um outro protocolo, o Resource Reservation Protocol (RSVP), permite ao receptor requisitar uma montante específico de largura de banda para um fluxo de dados em particular e receber uma resposta indicando quando a requisição foi garantida. Embora o RSVP não seja parte integrante do padrão H.323, alguns produtos H.323 podem suportá-lo. O RTP necessita ser suportado por Terminais, Gateways, e MCUs com Processadores Multiponto. O RSVP pode também ser suportado pelos mesmos componentes e pelos roteadores e switches intermediários.

## 14.8 Ferramentas H.323

Esta seção apresenta um levantamento das aplicações (terminais clientes, MCUs, Gatekeepers e Gateways) que implementam o padrão H.323.

### 14.8.1 Servidores H.323

Chama-se servidor H.323 um endpoint com capacidade de fazer funções de MCU, de Gatekeeper ou de Gateway ou qualquer combinação entre estes três tipos. Este item lista as características de alguns desses servidores disponíveis no mercado.

#### **Servidores H.323 – MCUs**

Existem alguns servidores H.323 que implementam a funcionalidade de unidade de controle multiponto MCU:

- RADVision MCU-323 (<http://www.radvision.com>)
- Ezenia Encounter NetServer (<http://www.ezenia.com>)
- White Pine MeetingPoint Conference Server (<http://www.wpine.com>)

#### **Servidores H.323 – Gatekeepers**

Segundo o padrão H.323, o gatekeeper é definido como um componente opcional, porém ele está se tornando elemento fundamental no centro de controle de redes convergentes para voz, dados e vídeo. A seguir são listados alguns produtos:

- RADVision NGK-100 (<http://www.radvision.com>);
- Ezenia Encounter Gatekeeper (<http://www.ezenia.com>)
- White Pine MeetingPoint Conference Server (<http://www.wpine.com>)

### Servidores H.323 – Gateways

Os gateways H.323 permitem a troca de mensagens entre sistemas de conferência que utilizam este padrão e sistema que implementam outro padrão, p.e H.320 ou H.324. A seguir são listados alguns produtos:

- RADVision H.323/H.320 Gateway – L2W-323P (<http://www.radvision.com>)
- **Ezenia Encounter Netgate** (<http://www.radvision.com>)

### 14.8.2 Clientes H.323 – Características Descritas

Clientes H.323 são os “endpoints” do tipo terminal que implementam as funcionalidades mínimas previstas no padrão ITU-H.323. A tabela abaixo mostra os principais produtos existentes no mercado atualmente bem como um resumo de suas características. Esta tabela permite ao usuário selecionar os produtos com características que mais adequadas as suas necessidades. Após esta pré-seleção um estudo mais detalhado pode ser conduzido buscando identificar principalmente aspectos relacionados a facilidades de uso e estabilidade da ferramenta.

Uma boa fonte recomenda para obtenção de informações sobre clientes H.323 é <http://www.openh323.org>.

Produto	AU	VI	WB	CH	FT	AS	AM	MC	Fax	GK	LDAP	AV	Plat.
Microsoft NetMeeting	S	S	S	S	S	S	N	S	N	S	S	S	9X/NT
Intel VideoPhone	S	S	N	N	N	N	N	S	N	N		S	9X
VocalTec Internet Phone	S	S	S	S	N					N		S <sup>1</sup>	9X/NT/mac
Voxware Voxphone	S												
Netscape Conference	S												9X/NT
Netspeak Webphnone	S	S	N	N	N	N	S	4	N	N		S <sup>2</sup>	9X/NT
QuickNet LineJack/ PoneJack	S <sup>3</sup>												
Paradise Simplicity H323	S <sup>3</sup>												
Sorenson Envision	S	S	S	S	S	S	S	N	N	S	S	N	9X/NT
Selsius-Phone	S												
SmithMicro CommSuite	S	S					S				S		9X/NT
White Pine CUSeeMe Pro	S	S	S		S								9X/NT
DC-Share	S	S											SUN/SGI/HP
Macchina Pty Ltd	S	S	N	N	N	N	S	S	N	N			9X/NT

Notas:

- 1 Versão demonstração expira em 2 semanas
- 2 Versão demonstração permite fazer chamadas de até 3 minutos
- 3 Equipamento Stand-alone

**Legenda:**

- AU = Áudio – Implementa facilidade de áudio;
- VI = Vídeo - Implementa facilidade de vídeo;
- WB = White Board - Implementa facilidade de quadro branco compartilhado
- CH = Chat - Implementa facilidade de troca de mensagens em tempo real;
- FT = File Transfer – Implementa facilidade de transferência de arquivos;
- AS = Application Sharing = Implementa facilidade de compartilhamento de aplicações;
- AM = Answering Machine - Implementa facilidade de resposta automática
- MC = Multiple Calls – Suporta executar múltiplas chamadas;
- Fax = Implementa facilidade de fax;
- GK = Suporta a utilização de Gatekeeper;
- ILS = Internet Locator Server - Facilidade de localização na Internet baseada em LDAP
- LDAP = Facilidade de uso de diretório LDAP local;
- AV = Cópia de avaliação disponível para download;
- Plat. = Plataformas suportadas.

## Capítulo 15 Bibliografia

---

- [Abdel-Wahab, 96] H. Abdel-Wahab, K. Maly, E. Stoica. Multimedia Integration Into a Distance Learning Environment. 3<sup>rd</sup> International Conference on Multimedia Modelling, Toulouse, November, 1996.
- [Ackermann, 94] P. Ackermann. Direct Manipulation of Temporal Structures in a Multimedia Application Framework. In proc. of the ACM Multimedia 94.
- [Asymetrix, 99] Asymetrix Learning Systems, Inc. URL: <http://www.asymetrix.com/>.
- [Allen, 83] J.F. Allen. Maintaining Knowledge about Temporal Intervals. Communications of the ACM 26(11):832-843, 1983.
- [Anderson] B. Anderson. MultiTalk. URL: <http://pipkin.lut.ac.uk/~bem/multitalk>.
- [Appelt, 93] W. Appelt. HyperODA - Extensions for Temporal Relationships. Proposed Draft Amendment (version 2) ISO/IEC JTC 1/SC 18/WG 3/N 2516, 1993.
- [Apple, 91] Apple Computer Inc. QuickTime Developer's Guide. Developer technical publications, 1991.
- [Apple, 94] Multimedia Demystified - A Guide to the World of Multimedia from Apple Computer, Inc. Random House/NewMedia Series, 1994.
- [Authorware, 89] Inc. Authorware Professional Manual. Redwood City, CA, 1989.
- [Barberis, 80] G. Barberis. D. Pazzaglia. Analysis and Optimal Design of a Packet-Voice Receiver. IEEE Transactions on Communications 28(2):217-227, 1980.
- [Benarjea, 96] A. Benarjea, D. Ferrari, B. Mah, M. Moran, D. Verma, H. Zhang. The Tenet Real-time Protocol Suite: Design, Implementation, and Experience. IEEE/ACM Transaction on networking. February 1996:1-10.
- [Bernhardt, 95] C. Bernhardt, E. Biersack. A Scalable Video Server: Architecture, Design, and Implementation. In Realtime Systems Conference, Paris, pp. 63-72, January 1995.
- [Black, 98] D. Black, S. Blake, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Services", Internet Draft, May 1998.
- [Blair, 93] G. Blair, et al. A network Interface Unit to Support Continuous Media. Journal of Selected Area in Communications 11(2):264-275, 1993.
- [Intel, 98] [http://channel.intel.com/business/ia\\_chan/micinfo/agp.htm](http://channel.intel.com/business/ia_chan/micinfo/agp.htm).
- [Blakowski, 92] G. Blakowski, J. Jens Hübel, U. Langrehr, M. Mühlhäuser. Tool Support for the Synchronization and Presentation of Distributed Multimedia. Computer Communication, 15(10):611-618. December, 1992.
- [Blakowski, 96] G. Blakowski and R. Steinmetz. *A Media Synchronization Survey: Reference Model, Specification, and Case Studies*. IEEE Journal on Selected Areas in Communications, 14(1):5-35, January, 1996.
- [Bolot, 94] J.C. Bolot, T. Turletti, I. Wakeman. Scalable Feedback Control for Multicast Video Distribution in the Internet. Proc. ACM SIGCOMM pp. 58-67, 1994.
- [Botafogo, 95] R. Botafogo, D. Mossé. The MORENA Model for Hypermedia Authoring and Browsing. In proc. of the IEEE Int. Conf. on Multimedia Computing and Systems, pp. 42-49, Washington, DC, May 1995.
- [Branden, 94] R. Braden, D. Clark & S. Shenker. "Integrated Services in the Internet Architecture: an Overview", RFC 1633, June 1994.
- [Braun, 97a] T. Braun, H. Stütten. Implementation of an Internet Video Conferencing Application over ATM. Presented at IEEE ATM'97 Workshop, May 1997.
- [Braun, 97b] T. Braun. Internet Protocols for Multimedia Communications. Part I: IPng The Foundation of Internet Protocols. IEEE Multimedia 4(3):85-90, 1997. Part II: Resource Reservation, Transport, and Application Protocols. IEEE Multimedia 4(4):74-82, 1997.
- [Bray, 98] T. Bray, J. Paoli, C.M. Sperberg-MacQueen. Extensible Markup Language (XML) 1.0. URL: <<http://www.w3.org/TR/REC-xml/>>. Fevereiro de 1998.
- [Buchanan, 92] M.C. Buchanan, P.T. Zellweger. Specifying Temporal Behavior in Hypermedia Documents. Multimedia Systems Journal, 1993.
- [Buchanan, 93] M.C. Buchanan, P.T. Zellweger. Automatically Generating Consistent Schedules for Multimedia Documents. Multimedia Systems Journal, 1993.

- [Buford, 94] Multimedia Systems. John F. Koegel Buford. ACM Press - SIGGRAPH Series - New York, New York, 1994.
- [Bulterman, 91] D.C.A. Bulterman, G. van Rossun, R. van Liere. A Structure for Transportable, Dynamic Multimedia Documentos. In proc Summer 1991 Usenix Conference, Hashville TN, June 1991.
- [Bulterman, 91] D.C.A. Bulterman and R.V. Liere. Multimedia Synchronization and UNIX. Proceedings of Second International Workshop on Network and Operating System Support for Digital Audio and Video, pp. 108-119, 1991.
- [Campbell, 94] A. Campbell et al. A Quality of Service Architecture. Computer Communication Review 24(3):6-27,1994.
- [Campbell, 96] A. Campbell, C. Aurrecoechea, L. Hauw. A Review of QoS Architectures. ACM Multimedia System Journal, 1996. Também no URL <ftp://ftp.ctr.columbia.edu/CTR-Research/comet/public/papers/96/>.
- [Casner, 92] Casner, S.; Deering, S. First IETF Internet Audiocast, ACM SIGCOMM Computer Communications Review. San Diego – Califórnia, jul. 1992, pp. 92-97. Documento também disponível em ftp://venera.isi.edu/pub/MBone/ietf-audiocast-article.ps.
- [Casner, 95] Casner, S. Frequently Asked Question (FAQ) on *Multicast Backbone* (MBONE). Dez. 1995. Documento disponível em <http://www.reserach.att.com/MBone-faq.html>.
- [Chen] Z. Chen, S.G. Tan, R.H. Campbell, Y. Li. Real Time Video and Audio in the World Wide Web. <http://choices.cs.uiuc.edu/srg/stan/vosaic/Overview.html>.
- [Clark, 90] D. Clark, D. Tenenhouse. Architectural Considerations for a New Generation of Protocols. Proc. ACM SIGCOMM 90, ACM Press, New York, 1990, pp. 200-208.
- [DataBeam, 98] DataBeam Corporation, (1998), A Primer on The H.323 Series Standard, <http://www.databeam.com>
- [Dauphin] G. Dauphin. mMosaic Informations. URL: <http://sig.enst.fr/~dauphin/mMosaic/index.html>
- [Delgrossi, 93] L. Delgrossi, D. Hehmann, R.G. Herrtwich, C. Halstrick, O. Krone, J. Sandvoss, C. Vogt. Media Scaling for Audiovisual Communication with the Heidelberg Transport System. Proceedings of the 1st ACM Multimedia Conference, Anaheim 1993.
- [Diaz, 93] M. Diaz, P. Sénac. Time Streams Petri Nets, a Model for Multimedia Streams Synchronization. In proc. Of the First International Conference on Multi-media Modeling - Vol. 1, Singapore, World Scientific, Tat-Seng Chua & T. L. Kunii (Eds.), pp. 257-273, November 1993.
- [Diot, 95] Ch. Diot. Adaptative Applications and QoS Guarantees. Proc. IEEE International Conference on Multimedia and Networking. IEEE Computer Society Press pp 99-106, Aizu, Japan, 1995.
- [Dodd, 98] Dodd, Annabel Z., (1998): *The Essential Guide to Telecommunications*, [Prentice-Hall, Inc.](#)
- [Drapeau, 91] G.D. Drapeau, H. Greengiled, H. MAEstro - A Distributed Multimedia Authoring Environment. In proc. of the USENIX Summer'91 Conference, pp. 315-328. Nashville TN, 1991.
- [Escobar, 92] J. Escobar, D. Deutsh, C. Partridge. Flow Synchronization Protocol. Proceedings of IEEE Globecom Conference, Orlando, Florida, Dec. 1992.
- [Eyes, 92] Eyes 2.0 Reference Manual. Center for Productivity Enhancement. University of Massachusetts-Lowell. 1992.
- [FAQ, 97] Multimedia Authoring Systems FAQ. Version 2.13. <http://www.tiac.net/users/jalisglar/MMASFAQ.HTML>. June 1997.
- [FAQ, 98] Frequently Asked Questions about the Extensible Markup Language. URL: <<http://www.ucc.ie/xml/>>.
- [Floyd, 95] S. Floyd, V. Jacobson, S. McCanne. A Reliable Multicast Framework for Lightweight Sessions and Application Level Framing. ACM SIGCOMM'95, pp. 342-356, 1995.
- [Fluckiger, 95] Understanding Networked Multimedia: Applications and Technology. Prentice Hall International (UK) Limited, 1995.

- [França Neto, 98] L.R. França Neto. Um Ambiente para Processamento de Grandes Acervos de Imagens. Dissertação de Mestrado do Departamento de Informática da Universidade Federal de Pernambuco.
- [Furht, 94] Borko Furht. Multimedia Systems: An Overview. IEEE Multimedia Spring 1994, pp. 47-59.
- [Gibbs, 91] S. Gibbs. Composite Multimedia and Active Objects. In proc. of the OOPSLA'91. Phoenix, Arizona. October, 1991.
- [Ginige, 95] A. Ginige, D.B. Lowe, J. Robertson. Hypermedia Authoring. IEEE Multimedia 2(4):24-35, 1995.
- [Gronbaek, 94] K. Gronbaek, R.H. Trigg. Design Issues for a Dexter-Based Hypermedia System. Communication of the ACM 37(2):40-49, 1994.
- [Gumbrich, 97] S. Gumbrich, H. Emgrunt, T. Braun. Dynamic Bandwidth Allocation for Stored VBR Video in ATM End System. Presented at 7<sup>th</sup> International IFIP Conference on High-Performance Networking, White Plains, USA, April 1997.
- [Haindl, 96] M. Haindl. Multimedia Synchronization. Computer Science/Department of Interactive Systems. IEEE Journal on Selected Areas in Communication 14(1):73-83, 1996.
- [Halasz, 90] F. Halasz, M. Schwartz. The Dexter Hypertext Reference Model. NIST Hypertext Standardization Workshop, Gaithersburg, MD, January, 1990.
- [Halasz, 94] F. Halasz, M. Schwartz. The Dexter Hypertext Reference Model. Communication of the ACM 37(2):29-39, February 1994.
- [Halsall, 88] Fred Halsall. Data Communications, Computer Networks and OSI. Addison Wesley, 1988.
- [Hamblin, 72] C.L. Hamblin. Instants and Intervals. In proc. of 1st Conf. Int. Soc. For the Study of Time, J.T. Fraser et al. (Eds.), Springer-Verlag, pp. 324-331, New York, 1972.
- [Handley, 97] M. Handley, J. Crowcroft. Network Text Editor (NTE) A Scalable Shared Text Editor for the MBone. Presented at ACM SIGCOMM, Cannes, September 1997.
- [Hardman, 93] L. Hardman, D.C.A. Bulterman, G. van Rossum. The Amsterdam Hypermedia Model: Extending Hypertext to Support Real Multimedia. Hypermedia Journal 5(1), May, 1993.
- [Hardman, 94] L. Hardman, D.C.A. Bulterman, G Van Rossum. The Amsterdam Hypermedia Model: Adding Time, Structure and Context to Hypertext. Communication of the ACM 37(20):50-62. February 1994.
- [Hardman, 95] L. Hardman, D.C.A. Bulterman. Authoring Support for Durable Interactive Multimedia Presentations. STAR Report in Eurographics'95.
- [Herman, 94] I. Herman, G.J. Reynolds. MADE: A Multimedia Application Development Environment. In proc. of the Int. Conf. on Multimedia Communication and Systems, pp. 184-193, 1994.
- [Hirzalla, 95] N. Hirzalla, B. Falchuk and A. Karmouch. *A Temporal Model for Interactive Multimedia Scenarios*. IEEE Multimedia, pp. 24-31, Fall 1995.
- [Hudson, 93] S.E. Hudson, C.H. Hsi. A Framework for Low Level Analysis and Synthesis to Support High Level Authoring of Multimedia Documents. Graphics Visualization and Usability Center Technical Report GVU-93-14. Georgia Institute of Technology, 1993.
- [Hehmann, 90] D. Hehmann, M. Salmony, H. Stuttgen. Transport Services for Multimedia Applications on Broadband Networks. Computer Communications 13(4):197-203, 1990.
- [Holfelder, 95] W. Holfelder. MBone VCR - Video Conference Recording on the MBone. ACM Multimedia'95, pp. 237-238, 1995.
- [Holfelder, 97] W. Holfelder. <http://www.informatik.uni-mannheim.de/informatik/pi4/projects/MVoD>.
- [Hopper, 90] A. Hopper. Pandora – An Experimental System for Multimedia Applications. ACM Operating System Review 24(2):19-34, 1990.
- [Huang, ?] T. S. Huang, S. M. Kang, J. Stroming, MPEG-4 Project, Universidade de Illinois, EUA. <http://uivlsci.csl.uiuc.edu/stroming/mpeg4/>
- [IBM, 97] IBM International Technical Support Organization. IBM Networked Video Solution Over ATM Implementation. IBM Redbook SG24-4958, January, 1997.

- [IMTC, 98] International Multimedia Teleconferencing Consortium, Inc (IMTC) (1998), Multimedia Teleconferencing Standards, <http://www.imtc.org>
- [ISO 8613] ISO 8613 Information Processing - Text and Office Systems - Office Document Architecture (ODA) and Interchange Format (ODIF), 1998.
- [ISO 13522] Committee Draft ISO/IEC 13522-1 Information technology - Coding of Multimedia and Hypermedia - Part 1: MHEG Object Representation - Base Notation (ASN.1), October, 1994.
- [Käppner, 94] T. Käppner, L.C. Wolf. Media Scaling in Distributed Multimedia Object Services. In Multimedia: Advanced Teleservices and High-Speed Communication Architectures, R. Steinmetz (Ed.). Springer LNCS 868, pp. 34-43, 1994.
- [Klas, 90] W. Klas. E.J. Neuhold, M. Schrefl. Using an Object-Oriented Approach to Model Multimedia Data. Computer Communication 13(4):204-216, 1990.
- [Koegel, 93] J.K. Koegel, J. Heines. Improving Visual Programming Languages for Multimedia Authoring. In proc. of the ED-MEDIA'93, World Conf. on Educational Multimedia and Hypermedia, pp. 286-293, 1993.
- [Kotha, 98] Kotha S., (1998), Deploying H.323 Applications in Cisco Networks, Cisco System Inc., <http://www.cisco.com>.
- [Kumar, 95] Kumar, V. MBone: Interactive Multimedia on the Internet. Indianapolis: New Riders Publishing, 232p. 1995.
- [Kuo, 98] Franklin Kuo, Wolfgang Effelsberg, J.J. Garcia-Luna-Aceves. Multimedia Communications: Protocols and Applications. Prentice Hall PTR, 1998.
- [Laubach, 94] M. Laubach. Classical IP and ARP over ATM, IETF RFC 1577, 1994.
- [Leung, 90] W.H.F. Leung et al. A Software Architecture for Workstations Supporting Multimedia Conferencing in Packet Switching Networks. IEEE Journal on Selected Areas in Communications 8(3):380-390, 1990.
- [Leydekkers, 91] P. Leydekkers, B. Teunissen. Synchronization of Multimedia Data Streams in Open Distributed Environments. Proceedings of Second International Workshop on Network and Operating System Support for Digital Audio and Video, pp. 94-103, 1991.
- [Liao, 97] T. Liao. WebCanal: a Multicast Web Application. Presented at 6<sup>th</sup> International on World Wide Web Conference, Santa Clara, CA, April 7-11, 1997.
- [Little, 90] T.C.D. Little, A. Ghafoor. Synchronization and Storage Models for Multimedia Objects. IEEE Journal on Selected Areas in Communication 8(3):413-427, April 1990.
- [Little, 91] T.C.D. Little and A. Ghafoor. Spatio-Temporal Composition of Distributed Multimedia Objects for Value-Added Networks. IEEE Computer, pp. 42-50, Oct. 1991.
- [Louher, 93] P.K. Louher. The Design of a Storage Server for Continuous Media, PhD Thesis, Department of Computing, Lancaster University, 1993.
- [Louher, 96] P. Louher, R. Louher, D. Shepherd, D. Pegler. A Scalable Hierarchical Video Storage Architecture. Proc. Of Multimedia Computing and Networking 1996, pp. 18-29, 1996.
- [Lu, 94] G. Lu et al. Temporal Synchronization Support for Distributed Multimedia Information Systems. Computer Communications, 17(12):852-862, 1994.
- [Lu, 96] G. Lu. Communication and Computing for Distributed Multimedia Systems. Guojun Lu. Artech House Inc., 1996.
- [Lyonnet] F. Lyonnet. Rendez-Vous, the next generation videoconferencing tool. URL: <http://www.inria.fr/rodeo/personnel/Frank.Lyonnet/IVSStng/ivstng.html>.
- [Makedon, 94] F. Makedon et al. Multimedia Authoring, Development Environments, and Digital Video Editing. Technical Report TR94-231, Dartmouth Experimental Visualization Lab., Dartmouth College, Hanover, New Hampshire, 1994.
- [MBone, 99] MBone Agenda. Disponível em <http://www.cilea.it/MBone/agenda.html>.
- [McCanne, 96] S. McCanne. Scalable Compression and Transmission of Internet Multicast Video. Ph.D dissertation, University of California, Berkeley, 1996.
- [Melchior, 97] C. Melchior, L.M.R. Tarouco. Uma Análise do Mbone. XV Simpósio Brasileiro de Redes de Computadores, pp. 313-324, 1997.

- [Mills, 91] D.L. Mills. Internet Time Synchronization: The Network Time Protocol. IEEE Transaction on Communications 39(10):1482-1492, 1991.
- [Morgam, 97] Morgam, Edward B., (1997), Voice Over Packet White Paper, Tely Networks Inc., <http://www.tely.com>.
- [Murata, 89] T. Murata. Petri Nets: Properties, Analysis and Applications. IEEE 77(4):541-580. April, 1989.
- [Naylor, 82] W.E. Naylor, L. Kleinrock. Stream Traffic Communication in Packet Switched Networks: Destination Buffering Considerations. IEEE Transaction on Communications 30(12):2527-2534, 1982.
- [Newcomb, 91] S. Newcomb et al. The HyTime Hypermedia/Time-based Document Structuring Language. Communication of the ACM 34(11):159-166, 1991.
- [Nichols, 98] Nichols, K., Blake, S. Differentiated Services Operational Model and Definitions. Internet Draft, February 1998.
- [Open H323, 99] Open H323 Project, (1999), H.323 Clients, <http://www.openh323.org>.
- [Parnes, 97] P.Parnes, M. Mattsson, K. Symmes, D. Schefström. The mWeb Presentation Framework. Presented at 6<sup>th</sup> International World Wide Web Conference, Santa Clara, CA, April 7-11, 1997.
- [Raghavan, 98] V. Raghavan, S.K. Tripathi. Networked Multimedia Systems: Concepts, Architecture, and Design. Prentice Hall, 1998.
- [Ramannathan, 93] S. Ramannathan, P.V. Rangan. Feedback Techniques for Intra-Media Continuity and Inter-Media Synchronization in Distributed Multimedia Systems. The Computer Journal, Special Issue on Distributed Multimedia Systems, Mar. 1993.
- [Patel, 93] K. Patel, B.C. Smith, L.A. Rowe. Performance of a Software MPEG Video Decoder. Proc. ACM Multimedia'93m ACM Press, New York, 1993, pp. 75-82.
- [Reddy, 94] A.L.N. Reddy, J.C. Wyllie. I/O Issues in a Multimedia System. Computer 7(3):69-74, Mar. 1994.
- [RFC 2211] J. Wroclawski, "Specification of the Controlled-Load Network Element Service". RFC 2211, Sept. 1997.
- [RFC 2212] S. Shenker, C. Partridge, R. Guerin, "Specification of Guaranteed Quality of Service", RFC 2212, September 1997.
- [Rothermel, 92] K. Rothermel, G. Dermler. Synchronization in Joint-Viewing Environments. Proceedings of third International Workshop on Network and Operating System Support for Digital Audio and Video, pp. 106-118, 1992.
- [Rowe, 92] A.L. Rowe et al. MPEG Video in Software: Representation, Transmission, And Playback. Proceedings of SPIE, vol. 2188, pp. 134-144.
- [Schloss, 94] G.A. Schloss, M.J. Wynblatt. Building Temporal Structures in a Layered Multimedia Data Model. In proc. ACM Multimedia'94, pp. 271-278, 1994.
- [Sénac, 95] P. Sénac, R. Willrich, P. de Saqui-Sannes. Hierarchical Time Stream Petri Nets: A Model for Hypermedia Systems. 16th International Conference on Application and Theory of Petri Nets. In Application and Theory of Petri Nets 1995, Lecture Notes in Computer Science no. 935, G. De Michelis and M. Diaz (Eds.), Springer, pp. 451-470.
- [Sénac, 96] P. Sénac, M. Diaz, A Léger, P. de Saqui-Sannes. Modeling Logical and Temporal Synchronization in Hypermedia Systems. IEEE Journal on Selected Areas in Communication 14(1):84-103, 1996.
- [Shackelford, 93] D.E. Shackelford, J.B. Smith, F.D. Smith. The Architecture and Implementation of a Distributed Hypermedia Storage System. Technical Report 93-013, Department of Computer Science, The University of North Carolina at Chapel Hill, 1993.
- [Shepherd, 90] D. Shepherd, M. Salmony. Extending OSI to Support Synchronization Required by Multimedia Applications. Computer Communications 13(7):399-406, 1990.
- [Steinmetz, 92] R. Steinmetz. Synchronization Properties in Multimedia Systems. IEEE Journal on Selected Areas in Communications 8(3):410-412, 1990.
- [Steinmetz, 95] R. Steinmetz, K. Nahrstedt. Multimedia: Computing, Communications e Applications. Prentice Hall Series in Innovative Technology. 1995.
- [Stotts, 90] P.D. Stotts, R. Furuta. Temporal Hyperprogramming. Journal of Visual Languages and Computing 1:237-253, 1990.

- [Schulzrinne, 95] H. Schulzrinne. Internet Services: from Electronic Mail to Real-Time Multimedia. Proc. of KIVS (Kommunikation in Verteilten Systemen) (Klaus Franke, Uwe Hübler, and Winfried Kalfa, eds.), Informatik aktuell, (Chemnitz, Germany), pp. 21-34, Gesellschaft für Informatik, Springer Verlag, Feb. 1995.
- [Schulzrinne, 97] H. Schulzrinne et al. RTP: A Transport Protocol for Real-Time Applications. Internet Draft draft-ietf-avt-rtp-02.new-01.ps, 1997.
- [Schulzrinne, 98] H. Schulzrinne et al. RTP: Real Time Streaming Protocol (RTSP). Internet Draft, URL: <ftp://ftp.isi.edu/in-notes/rfc2326.txt>. Abril, 1998.
- [Sun, 95] Sun Microsystems. The Java Language Specification, 1995.
- [Sun, 97] Sun Microsystems, Inc. Java Media Framework 1.0 beta. URL: <http://java.sun.com/products/java-media/>.
- [Tanenbaum, 97] A. Tanenbaum. Computer Networks. Third Edition, Prentice Hall, 1997.
- [Topolcic, 90] C. Topolcic. Experimental Internet Stream Protocol, Version 2 (ST-II), Request for Comments RFC 1190, Internet Engineering Task Force, Oct. 1990.
- [Trentin, 96] M. Trentin, L.M.R. Tarouco. Suporte Multimídia para Educação a Distância. I Workshop de Educação à Distância (I WEAD), XIV SBRC, 1996.
- [Vazirgiannis, 93] M. Vazirgiannis, M. Hatzopoulos. A Script Based Approach for Interative Multimedia Applications. In proc. of Multimedia Modelling'93, pp. 129-143, Singapore, November, 1993.
- [Vidal, 97] P.C.S. Vidal. Evolução do Padrão MPEG. <http://www.gta.ufrj.br/~vidal/mpeg/mpeg.html>
- [Xiao, 99] X. Xiao, L.M. Ni. Internet QoS: A Big Picture. URL: <http://www>.
- [Yang, 96] C.C. Yang, J.H. Huang. A Multimedia Synchronization Model and Its Implementation in Transport Protocols. IEEE Journal on Selected Areas in Communications 14(1):212-255, 1996.
- [W3C, 99] W3C Architecture domain: Synchronized Multimedia. URL: <<http://www.w3.org/AudioVideo/Activity.html>>.
- [Wang, 95] H.K. Wang, J.L.C. Wu. Interactive Hypermedia Applications: A Model and Its Implementation. Software-Practice and Experience 25(9):1045-1063, 95.
- [Wahl, 94] T. Wahl, K. Rothermel. Representing Time in Multimedia Systems. In proc. of the International Conference on Multimedia Computing and Systems (ICMCS'94), pp. 538-543, Boston, 1994.
- [Whetten, 94] B. Whetten, T. Montgomery, S. Kaplan. A High Performance Totally Ordered Multicast Protocol. Theory and Pratice in Distributed Systems. Springer-Verlag, LNCS 938, 1994.
- [Willrich, 96a] R. Willrich, P. Séanc, M. Diaz, P. de Saqui-Sannes. A Formal Framework for the Specification, Analysis and Generation of Standardized Hypermedia Documents. Third IEEE International Conference on Multimedia Computing and Systems (ICMCS'96), pp. 399-406, Japão, 1996.
- [Willrich, 96b] R. Willrich, P. Séanc, P. de Saqui-Sannes, M. Diaz. Towards Hypermedia Documents Design. 14o Simpósio Brasileiro de Redes de Computadores (SBRC'96), pp. 473-491. Fortaleza, 1996.
- [Willrich, 96c] R. Willrich. Conception Formelle de Documents Hypermédias Portables. Tese apresentada no Laboratoire d'Analyse et d'Architecture des Systèmes du C.N.R.S. em vista de obtenção do título de Docteur pela Universidade Paul Sabatier. Toulouse (França), Setembro de 1996.
- [Willrich, 97] R. Willrich, P. de Saqui-Sannes. *Concepção Formal de Aplicações Multimídia Java*. Anais do XIV Simpósio Brasileiro de Redes de Computadores, pp 135-149, São Carlos (SP), 1997.
- [Witowski] Witowski, Willian E., IP Telephone Design and Implementation Issues, Telogy Networks Inc., <http://www.telogy.com>.
- [Wittig, 94] H. Wittig, J. Winckler, J. Sandvoss. Network Layer Scaling: Congestion Control in Multimedia Communication with Heterogeneous Networks and Receivers. Proc. Multimedia Transport and Teleservices, Springer LNCS 882, pp. 274-293, 1994.
- [Wong, 95] J.W.T Wong, K.K. Pang. Random Access Support for Large Scale VoD. Proc. Australian Telecommunication Networks and Applications Conference, pp. 225-230, 1995.

- [Wynblatt, 95] M. Wynblatt. Position Statement on Multimedia Synchronization. In proc. of the IEEE Workshop on Multimedia Synchronization (Sync'95). In URL: <http://spiderman.bu.edu-sync95-sync95.html>. Tysons Corner, Virginia, 95.
- [Zhang, 94] L. Zhang et al. ReSource ReserVation Protocol (RSVP) Internet Draft, Março, 1994.