

**UNIVERSIDADE FEDERAL DE SANTA CATARINA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA  
COMPUTAÇÃO**

**TRANSMISSÃO DE FLUXOS MPEG-2 COM QOS  
A PARTIR DE SERVIDORES MULTIMÍDIA  
EM REDES ATM**

**Patrícia Lima Seixas Vieira de Melo**

*Florianópolis  
Fevereiro - 2001*

**UNIVERSIDADE FEDERAL DE SANTA CATARINA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA**  
**COMPUTAÇÃO**

*Patrícia Lima Seixas Vieira de Melo*

**TRANSMISSÃO DE FLUXOS MPEG-2 COM QOS**  
**A PARTIR DE SERVIDORES MULTIMÍDIA**  
**EM REDES ATM**

*Dissertação submetida à Universidade Federal de  
Santa Catarina como parte dos requisitos para a  
obtenção do grau de Mestre em Ciência da  
Computação.*

**Orientador: Prof. Dr. Roberto Willrich**

*Florianópolis*  
*Fevereiro - 2001*

**TRANSMISSÃO DE FLUXOS MPEG-2 COM QOS  
A PARTIR DE SERVIDORES MULTIMÍDIA  
EM REDES ATM**

*Patrícia Lima Seixas Vieira de Melo*

Esta Dissertação foi julgada adequada para a obtenção do título de Mestre em Ciência da Computação, Área de Concentração Sistemas de Computação, e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciência da Computação.



---

Prof. Fernando A. O. Gauthier



---

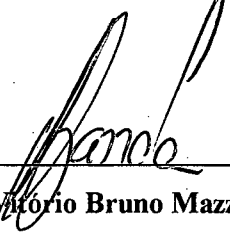
Prof. Dr. Roberto Willrich

**Banca Examinadora:**



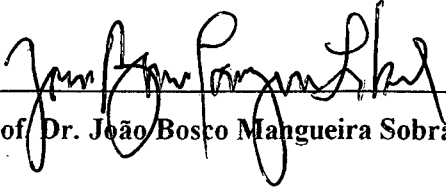
---

Prof. Dr. Roberto Willrich



---

Prof. Dr. Vitorio Bruno Mazzola



---

Prof. Dr. João Bosco Mangueira Sobral

*Ao meu marido Gustavo e ao meu  
anjinho, cujo rostinho ainda não  
conhecemos, mas que já enche a  
família de tanta felicidade.*

## AGRADECIMENTOS

*A realização deste trabalho foi possível não só pela minha vontade e esforço próprios, porém, devido a uma série de incentivos oriundos de pessoas diferentes. Pessoas as quais devo não somente carinho e admiração, mas também meus sinceros agradecimentos.*

*Desta forma, agradeço às pessoas que dirigem o Curso de Mestrado em Ciência da Computação desta instituição, especialmente, ao Professor Roberto Willrich, pela paciência e dedicação com que conduziu junto comigo este projeto.*

*Aos meus amigos, principalmente, à minha sempre amiga-irmã Anamélia, e familiares, pelas palavras de incentivo e conforto, palavras muito importantes nos momentos mais difíceis.*

*Contudo, devo agradecimentos muito especiais o meu marido, pelo amor dedicado, pela presença constante e pela confiança no meu potencial e aos meus pais e irmãos, meus maiores incentivadores, os quais sempre presentes durante o decorrer de minha vida, exercem papel fundamental na realização de todos os projetos do qual faço parte.*

*Agradeço ao meu grandioso Papai do Céu por razões que somente eu e Ele conhecemos.*

## RESUMO

*Os servidores de armazenamento multimídia surgiram da necessidade de manipulação das informações multimídia. Diante das características deste tipo de informação, as quais diferenciam-se amplamente das informações alfanuméricas, o conceito de Qualidade de Serviço (QoS) precisou ser incorporado ao projeto deste tipo de servidor, visando proporcionar a entrega satisfatória dos serviços solicitados. A qualidade originalmente associada ao serviço, a fim de atender as necessidades dos usuários, pode ser modificada e desta forma, passar por um processo de adaptação às novas exigências dos mesmos. O trabalho de Dissertação de Mestrado apresentado tem como objetivo propor um modelo para a transmissão de fluxos MPEG-2 com QoS, a partir de servidores de armazenamento multimídia. Esta contribuição poderá posteriormente fazer parte de uma infra-estrutura maior, haja vista que o modelo proposto trata ainda dos processos de mapeamento e negociação dos parâmetros de Qualidade de Serviço executados durante a fase de configuração destas transmissões. O presente trabalho baseia-se especificamente em servidores de vídeo MPEG-2 tendo como meio de transporte a Rede ATM, uma vez que o mesmo interessa-se somente na definição de QoS ao nível desta tecnologia. A fim de que o embasamento necessário para a apresentação desta Dissertação fosse obtido, a exposição de alguns conceitos iniciais tornou-se essencialmente importante. Portanto, este trabalho introduz também os principais conceitos relativos a Aplicações Multimídia Distribuídas, Padrão de Compressão MPEG-2, Qualidade de Serviço, Tecnologia ATM e Sistema de Gerenciamento de Banco de Dados Multimídia e Servidores de Armazenamento Multimídia.*

**ABSTRACT**

*Multimedia storage servers appeared due to the necessity of manipulating multimedia information. Taking into account the characteristics of this kind of information, which greatly differ from alphanumeric information, Quality of Service (QoS) concepts needed to be incorporated to the designing of multimedia servers, in order to provide suitable delivery of requested services. With the purpose of satisfying clients' requirements, the quality originally associated to the service can be modified and, at that moment, go through an adaptation process, looking forward to meeting the new requests. In this sense, the present Master's Dissertation proposes a model for the transmission of MPEG-2 streams with QoS from multimedia storage servers. Later on, this contribution could be included in a wider infrastructure, since it also deals with mapping and negotiating process of quality of service parameters in multimedia storage servers' transmissions. These processes are executed during the configuration stage of these transmissions. This work will specially deal with MPEG-2 video servers implemented in ATM Networks. So, it is essentially important to define some basic concepts in order to obtain the necessary support to the proposal presentation. Therefore, this work introduces the principal concepts about Distributed Multimedia Applications, MPEG-2 Standard, Quality of Service, ATM Technology and Multimedia Database Management Systems, and Multimedia Storage Servers*

## SUMÁRIO

<b>Capítulo I - Introdução .....</b>	<b>01</b>
1.1. Introdução .....	01
1.2. Revisão Bibliográfica .....	02
1.3. Objetivos do Trabalho .....	06
1.4. Organização do Trabalho .....	06
<b>Capítulo II – Aplicações Multimídia Distribuídas .....</b>	<b>08</b>
2.1. Introdução .....	08
2.2. Características das Aplicações Multimídia Distribuídas .....	09
2.2.1. Tipo de Informação .....	09
2.2.2. Modo de Entrega de Informação .....	10
2.2.3. Simetria .....	13
2.2.4. Número de Participantes .....	14
2.3. Requisitos de Comunicação das Aplicações Multimídia Distribuídas .....	15
2.3.1. Largura de Banda e Espaço de Armazenamento .....	15
2.3.2. Suporte à Compressão de Dados .....	16
2.3.3. Atraso Fim-a-Fim e Variação de Atraso .....	17
2.3.4. Relação Temporal e Espacial entre as Mídias .....	19
2.3.5. Tolerância a Erros e Perdas de Dados Multimídia .....	22
2.3.6. Suporte a Comunicações <i>Multicast</i> e <i>Broadcast</i> .....	23
2.4. Qualidade de Serviço em Sistemas Multimídia Distribuídos .....	23
2.5. Considerações Finais .....	25
<b>Capítulo III – Padrão de Compressão MPEG-2 .....</b>	<b>26</b>
3.1. Introdução .....	26
3.2. Família de Padrões MPEG-2 .....	27
3.3. Padrão MPEG-2 <i>Video</i> .....	27
3.3.1. Objetos Elementares .....	28
3.3.2. Tipos de Imagem MPEG .....	30
3.3.3. Princípios do Processo de Compressão de Vídeo MPEG-2 .....	31
3.3.4. Processamento de Imagens Entrelaçadas .....	34



3.3.5. Níveis e Perfis .....	35
3.3.6. Escalabilidade do Padrão MPEG-2 <i>Video</i> .....	37
3.4. Padrão MPEG-2 <i>Audio</i> .....	39
3.4.1. Camadas de Compressão .....	39
3.4.2. Princípios do Processo de Compressão de Áudio .....	40
3.4.3. Formato do Quadro de Áudio MPEG .....	40
3.4.4. Modos de Codificação .....	41
3.5. Padrão MPEG-2 <i>Systems</i> .....	42
3.5.1. Processo de Conversão dos Fluxos de Informação em Fluxos de Pacotes .....	43
3.5.2. Mapeamento dos Pacotes PES em Pacotes de Transporte .....	45
3.5.3. Estrutura de um Pacote PES .....	46
3.5.4. Estrutura de um Pacote de Transporte .....	47
3.5.5. Detecção de Erros .....	51
3.5.6. Informações Específicas de Programa .....	52
3.6. Considerações Finais .....	54
<b>Capítulo IV - Qualidade de Serviço .....</b>	<b>55</b>
4.1. Introdução .....	55
4.2. Definição .....	56
4.3. Parâmetros de QoS .....	57
4.4. Modelo de QoS .....	59
4.4.1. Modelo Genérico de QoS .....	60
4.5. Considerações Gerais .....	64
<b>Capítulo V - Tecnologia ATM .....</b>	<b>65</b>
5.1. Introdução .....	65
5.2. Tecnologia ATM .....	66
5.3. Modelo de Arquitetura de uma Rede ATM .....	70
5.3.1. Sistemas Fins .....	70
5.3.2. Comutadores .....	71
5.3.3. <i>Gateways</i> .....	71
5.3.4. Interligação entre Dispositivos ATM .....	72
5.4. Modelo de Referência de Protocolos .....	73

5.5. Modelo de Camadas ATM .....	74
5.5.1. Camada Física .....	74
5.5.2. Camada ATM .....	76
5.5.3. Camada de Adaptação ATM (AAL) .....	78
5.6. Troca de Unidades de Informação entre Camadas .....	82
5.7. Categorias de Serviço ATM .....	85
5.8. Contrato de Tráfego e Parâmetros de Tráfego ATM .....	88
5.8.1. Parâmetros de QoS ATM .....	89
5.9. Processo de Sinalização ATM .....	92
5.9.1. Troca de Mensagens de Sinalização .....	93
5.10. Mapeamento de Pacotes de Transporte MPEG-2 em Células ATM .....	96
5.11. Considerações Finais .....	97
<b>Capítulo VI – Sistema de Gerenciamento de Banco de Dados Multimídia e</b>	
<b>    Servidores de Armazenamento Multimídia .....</b>	<b>98</b>
6.1. Introdução .....	98
6.2. Sistema de Gerenciamento de Banco de Dados Multimídia .....	99
6.2.1. Informações Multimídia e Metadados .....	99
6.3. Arquitetura de Servidor de Armazenamento Multimídia .....	100
6.3.1. Gerenciamento dos Requisitos de Espaço de Armazenamento .....	102
6.3.2. Mecanismos de Configuração de Discos de Armazenamento .....	103
6.3.3. Mecanismos de Posicionamento de Blocos de Dados no Disco .....	107
6.4. Algoritmos de Controle de Admissão e Escalonamento de Discos .....	112
6.4.1. Algoritmos de Controle de Admissão .....	112
6.4.2. Algoritmos de Escalonamento de Disco .....	114
6.5. Configuração do Servidor e Conexões de Rede .....	117
6.6. Considerações Finais .....	119
<b>Capítulo VII – Transmissão de Fluxos MPEG-2 com QoS a partir de Servidores</b>	
<b>    Multimídia em Redes ATM .....</b>	<b>121</b>
7.1. Introdução .....	121
7.2. Visão Geral do Modelo Proposto .....	122
7.3. Descrição do Funcionamento do Modelo Proposto .....	123

7.3.1. Seleção do Vídeo MPEG-2 e Especificação do Nível de QoS .....	123
7.3.2. Mapeamento e Negociação dos Parâmetros de QoS .....	124
7.3.3. Recuperação do Vídeo MPEG-2 Transmitido .....	125
7.4. Gerenciamento de Vídeos MPEG-2 .....	126
7.4.1. Esquema Básico para o Gerenciamento de Vídeos MPEG-2 .....	127
7.5. Camada de Mapeamento de QoS .....	131
7.5.1. Descrição do Serviço de Mapeamento de QoS .....	132
7.5.2. Descrição do Protocolo de Mapeamento de QoS .....	132
7.6. Exemplificação do Funcionamento do Modelo Proposto .....	134
7.6.1. Seleção do Vídeo MPEG-2 .....	135
7.6.2. Especificação do Nível de QoS .....	136
7.6.3. Apresentação do Vídeo MPEG-2 .....	137
7.7. Considerações Finais .....	138
<b>Capítulo VIII – Conclusão .....</b>	<b>139</b>
<b>Referências Bibliográficas .....</b>	<b>141</b>
<b>Bibliografia Consultada .....</b>	<b>144</b>

## LISTA DE FIGURAS

Fig. 2.1 – Aplicação Simétrica .....	13
Fig. 2.2 – Aplicação Assimétrica .....	14
Fig. 2.3 – Visão Simplificada do Processo de Compressão de Dados .....	16
Fig. 2.4 – Sincronização Espacial .....	19
Fig. 2.5 – Sincronização Intrafluxo .....	20
Fig. 2.6 – Sincronização Interfluxo .....	21
Fig. 3.1 – Hierarquia de Camadas de Vídeo MPEG .....	28
Fig. 3.2 – Concatenação das Imagens MPEG .....	30
Fig. 3.3 – Matriz de Quantização .....	32
Fig. 3.4 – Leitura em Zig-Zag .....	33
Fig. 3.5 – Processo de Compensação de Movimento .....	34
Fig. 3.6 – Imagens Progressivas .....	35
Fig. 3.7 – Imagens Entrelaçadas .....	35
Fig. 3.8 – Formato Simplificado do Quadro de Áudio MPEG .....	40
Fig. 3.9 – Estrutura Simplificada do Padrão MPEG-2 <i>Systems</i> .....	43
Fig. 3.10 - Mapeamento dos Pacotes PES em Pacotes de Transporte .....	45
Fig. 3.11 – Estrutura Simplificada de um Pacote PES .....	46
Fig. 3.12 – Estrutura de um Pacote de Transporte MPEG-2 .....	48
Fig. 3.13 – Relação entre as Tabelas de Informações Específicas de Programa .....	52
Fig. 3.14 - Mapeamento de uma Seção em um Pacote de Transporte MPEG-2 .....	54
Fig. 4.1 – Espectro de QoS .....	57
Fig. 4.2 – Início do Processamento da QoS .....	63
Fig. 4.3 – Gerenciamento da QoS .....	63
Fig. 4.4 – Finalização da QoS .....	64
Fig. 5.1 – Formato Genérico da Célula ATM .....	66
Fig. 5.2 – Modelo de Arquitetura de uma Rede ATM .....	70
Fig. 5.3 – Direcionamento das Informações nos Comutadores .....	71
Fig. 5.4 – Modelo de Referência de Protocolos para RDSI-FL .....	73
Fig. 5.5 – Modelo de Camadas ATM .....	74
Fig. 5.6 – Formatos da Célula ATM na UNI e na NNI .....	76

Fig. 5.7 – Caminhos e Canais Virtuais .....	78
Fig. 5.8 – Camada AAL .....	79
Fig. 5.9 – Troca de Unidades de Informação entre Camadas .....	82
Fig. 5.10 – Troca de Unidades de Informações entre Camada AAL e Camada ATM .....	83
Fig. 5.11 – Campos da CPCS–PDU .....	84
Fig. 5.12 – Troca de Mensagens de Sinalização .....	94
Fig. 5.13 – Mapeamento de Pacotes de Fluxo de Transporte MPEG–2 em Células ATM .....	96
Fig. 6.1 – RAID Nível 1 .....	104
Fig. 6.2 – Hierarquia de Armazenamento Distribuída .....	105
Fig. 6.3 – Vetor de Servidores sem Particionamento de Dados .....	106
Fig. 6.4 – Vetor de Servidores com Particionamento de Dados .....	107
Fig. 6.5 – Posicionamento Restrito de Dados .....	110
Fig. 6.6 – Posicionamento Estruturado de Registro de Dados .....	110
Fig. 6.7 – Escalamento Round–Robin com Posicionamento de Dados Restrito .....	116
Fig. 6.8 – Arquitetura de um Servidor Multimídia .....	118
Fig. 6.9 – Arquitetura SCAM .....	119
Fig. 7.1 – Visão Geral do Modelo Proposto .....	122
Fig. 7.2 – Transmissão do Fluxo MPEG-2 em Redes ATM .....	125
Fig. 7.3 – Esquema Básico para o Gerenciamento de Vídeos MPEG-2 .....	127
Fig. 7.4 - Camada de Mapeamento de QoS .....	131
Fig. 7.5. – Conexões ATM Estabelecidas .....	134
Fig. 7.6 – Processo de Seleção do Vídeo MPEG-2 .....	135
Fig. 7.7 – Especificação do Nível de QoS .....	136
Fig. 7.8 – Apresentação do Vídeo MPEG-2 Solicitado .....	137

## LISTA DE TABELAS

Tab. 2.1 – Simetria e Número de Participantes .....	14
Tab. 3.1 – Família de Padrões MPEG-2 .....	27
Tab. 3.2 – Níveis e Perfis .....	36
Tab. 3.3 – Principais Campos do Campo <i>Indicators</i> .....	46
Tab. 3.4 – Campos do Cabeçalho de um Pacote de Transporte .....	48
Tab. 3.5 – Campos do Campo de Adaptação .....	49
Tab. 5.1 – Classes de Serviços e Protocolos da Camada de Adaptação ATM .....	80
Tab. 5.2 – Parâmetros de QoS ATM .....	90
Tab. 6.1 – Requisitos de Espaço de Armazenamento para Dados Multimídia Digitais não Compactados .....	101

# CAPÍTULO I

## INTRODUÇÃO

*Este capítulo tem como finalidade apresentar uma introdução acerca dos temas tratados neste trabalho, bem como seus objetivos e a forma como o texto encontra-se organizado.*

### 1.1. INTRODUÇÃO

Os Sistemas Multimídia surgiram como uma evolução dos sistemas convencionais, isto é, os sistemas alfanuméricos, e dos avanços no desenvolvimento da tecnologia digital, possibilitando o processamento de forma integrada dos diversos tipos de informação, tais como textos, imagens, vídeos e áudios.

Visando a transmissão das informações multimídia através de redes de comunicação, estes sistemas, agora chamados Sistemas Multimídia Distribuídos, introduziram ao ambiente computacional e de comunicação rigorosos requisitos de desempenho fim-a-fim. Estes requisitos devem ser satisfeitos em busca do adequado funcionamento do sistema, bem como de que o resultado final da apresentação obtenha a qualidade esperada pelo usuário da aplicação multimídia.

Em vista disso, o conceito de Qualidade de Serviço (QoS) foi incorporado ao ambiente multimídia. Este conceito permite que os usuários/aplicações especifiquem os requisitos de desempenho necessários à execução do serviço e o sistema, por sua vez, forneça garantias de que os requisitos especificados serão satisfeitos.

No final dos anos 80, a Tecnologia ATM (*Asynchronous Transfer Mode*) surgiu como uma proposta de servir de meio de transporte comum para todos os tipos de informação existentes, inclusive utilizando-se dos conceitos de QoS e, conseqüentemente, satisfazendo os requisitos de cada tipo de informação.

Como as informações multimídia têm características e requisitos bastante diferentes, também os dispositivos de armazenamento e processamento voltados para informações convencionais precisaram ser adaptados, dando início a um grande interesse pelo desenvolvimento de servidores de armazenamento multimídia, os quais suportam tanto informações multimídia quanto alfanuméricas.

## 1.2. REVISÃO BIBLIOGRÁFICA

O servidor de armazenamento multimídia, ou somente servidor multimídia, assim como um servidor convencional, deve proporcionar o armazenamento, o gerenciamento e a recuperação de informações.

Contudo, como as informações manipuladas por este servidor possuem estrutura semântica mais complexa e ocupam espaço de armazenamento extraordinariamente maior, para que o funcionamento adequado de todo o sistema seja alcançado, o servidor multimídia deve proporcionar um bom desempenho fim-a-fim durante a execução de suas operações.

E para tanto, uma série de requisitos impostos ao sistema pelas informações multimídia precisa ser satisfeita, tal como a utilização de maneira otimizada dos recursos computacionais disponíveis no sistema. Desta forma, é fundamental que o servidor multimídia baseie suas operações nos conceitos de Qualidade de Serviço (QoS).

[Vogel, 1995] define QoS principalmente sob o ponto de vista da aplicação e do usuário: *“Qualidade de Serviço representa o conjunto das características quantitativas e qualitativas de um sistema multimídia distribuído, necessário para se conseguir a funcionalidade desejada de uma aplicação, a qual inclui a apresentação dos dados multimídia para o usuário e sua referente satisfação”*.

Neste sentido, muitos projetos de pesquisa têm sido desencadeados, a fim de tornar a recuperação, a transmissão e a apresentação de informações multimídia tarefas adequadamente executáveis do ponto de vista da rede de comunicação, da aplicação multimídia e, principalmente, do usuário.

De uma forma geral, em se tratando de Qualidade de Serviço (QoS), várias arquiteturas de gerenciamento de QoS têm sido apresentadas com o intuito de permitir que redes e sistemas-fins (usuários e servidores), usufruam do suporte à QoS e que, conseqüentemente, aplicações multimídia distribuídas baseadas em servidores multimídia possam ser colocadas em prática mais eficientemente.

Partindo do princípio que a probabilidade de existirem variações na quantidade de recursos inicialmente alocados às aplicações multimídia é muito grande, atualmente, dá-se muita ênfase ao que se denominou de *Arquitetura de QoS Adaptativa*. Nestas



arquiteturas, a alocação de recursos e a potencial renegociação dos mesmos ao longo da sessão multimídia são operações executadas dinamicamente pela aplicação conforme sua necessidade.

A QoS adaptativa foi abordada por [Witana, 1999] através de um modelo de software voltado ao desenvolvimento de aplicações adaptativas. Este modelo permite a especificação de regras dependentes da aplicação que controlam o comportamento da mesma. Estas regras ditam como as preferências dos usuários influenciam a QoS e quais ações serão tomadas frente à mudanças na disponibilidade de recursos da rede.

A fim de que os conceitos atribuídos a este modelo pudessem ser observados e, por conseguinte, validados, uma aplicação de Vídeo sob Demanda (VoD), chamada SuperNOVA, foi desenvolvida e utilizada como ferramenta à execução dos testes necessários. Esta aplicação oferece facilidades que permitem ao usuário especificar o grau de satisfação e o custo do serviço requisitado, e controlar as preferências de QoS que lhe são cabíveis. E é com base nestas especificações que os recursos da rede são alocados.

A partir da observação de que o usuário da aplicação multimídia é o principal medidor da qualidade do serviço que a aplicação em questão se dispõe a oferecer, estabeleceu-se o interesse pelo desenvolvimento de aplicações multimídia distribuídas que permitissem, de acordo com a qualidade percebida pelo usuário, a especificação e o controle do material multimídia a ser apresentado e do comportamento do serviço.

Contudo, para a implementação de aplicações desta natureza, os conceitos de QoS precisaram passar por uma reformulação que acrescentasse o que se convencionou chamar de Qualidade Perceptiva (*Perceptual Quality*), onde os parâmetros de QoS seriam especificados de acordo com a qualidade na qual o usuário vê e ouve o material multimídia.

Neste contexto, [Richards, 1998] propôs um sistema de adaptação que permite ao usuário a especificação, o monitoramento e a adaptação da QoS perceptível, através do controle de um único parâmetro, o custo. [Richards, 1998] introduziu, então, o conceito de “satisfação”, de maneira que a QoS fornecida pelo sistema pudesse ser quantificada. Por conseguinte, foram estudados o mapeamento do custo em termos de satisfação, e a configuração do sistema.

Outros projetos focalizam-se no melhoramento e até mesmo no desenvolvimento de novas arquiteturas de servidor multimídia, favorecendo o uso satisfatório dos recursos disponíveis na rede e a otimização do número de acessos simultâneos.

Dentre estes projetos pode-se citar o projeto *Elvira Video Server* [Sandstã, 1997], o qual constitui-se de uma arquitetura de servidor multimídia experimental, implementada em uma Rede ATM, utilizada como parte do modelo *VideoSTAR*. Este modelo visa possibilitar o compartilhamento e a reutilização dos dados multimídia e de suas respectivas informações de controle, isto é, metadados, os quais descrevem o conteúdo, a estrutura e a semântica dos dados multimídia aos quais estão associados.

A arquitetura apresentada no projeto *Elvira Video Server* foi desenvolvida para satisfazer uma série de novos requisitos atribuídas às aplicações de vídeo. Além dos requisitos de armazenamento e entrega de material multimídia de maneira eficiente, esta arquitetura também foi projetada para suportar vários padrões de codificação de vídeo e aplicações interativas, ou seja, aplicações que permitem interatividade entre o usuário e o material apresentado.

[Silveira, 2000] propôs um Servidor de Qualidade (SoQ) para a diminuição, quando necessário, da taxa de bits associada aos arquivos de vídeo transmitidos em tempo real através de uma Rede ATM, a fim de que a largura de banda disponível na rede fosse adequadamente utilizada.

Para tanto, uma rede neural “backpropagation” foi utilizada durante a fase de negociação dos parâmetros de QoS, de forma que o SoQ pudesse reajustar os coeficientes de quantização MPEG de maneira não linear, possibilitando, em vista disso, o acesso por parte dos usuários do material multimídia desejado, mesmo que a rede não seja capaz de permitir a transmissão na qualidade máxima associada ao mesmo. Os vídeos encontram-se previamente codificados em MPEG-1 ou MPEG-2 e armazenados em um servidor de vídeo ou em memória terciária.

Na fase de estabelecimento da conexão ATM e, por conseguinte, antes da transmissão do vídeo, o usuário deve ser informado sobre a possibilidade de ocorrerem variações na qualidade originalmente associada ao vídeo selecionado. Neste caso, é apresentada ao usuário uma lista contendo cinco amostras de qualidades de vídeo diferentes, através da

qual o mesmo pode selecionar a qualidade mínima aceitável para a apresentação do vídeo escolhido.

Desta forma, o usuário pode acessar o material multimídia com uma qualidade inferior àquela em que ele foi digitalizado, contudo, o fator determinante para esta diminuição de qualidade é a largura de banda disponível da rede e não as especificações dos usuários quanto a qualidade em que a transmissão e, por consequência, a apresentação do material devem ser executadas.

O trabalho descrito nesta dissertação apresenta um modelo voltado ao projeto de aplicações multimídia distribuídas implementadas em servidores multimídia, onde os conceitos de QoS Adaptativa e Perceptiva foram aplicados. A principal característica deste trabalho é a implementação de uma camada, a qual denominou-se Camada de Mapeamento de QoS, que permite que a qualidade do vídeo especificada pelo usuário seja dinamicamente fornecida pelo sistema, sempre que possível.

Para tanto, o sistema oferece mecanismos para que os parâmetros de qualidade das mídias contínuas, no caso vídeos MPEG-2, sejam especificados de maneira mais realística, isto é, com base na percepção humana quanto à qualidade do material multimídia selecionado, e para que haja resposta imediata às mudanças efetuadas pelos usuários da aplicação, no que se refere a qualidade percebida da mídia durante sua apresentação.

O que diferencia o modelo apresentado neste artigo dos demais modelos citados nesta seção, é o fato dos usuários, através de uma interface gráfica, terem acesso aos parâmetros de QoS mais diretamente ligados à qualidade do vídeo, tais como resolução da imagem e taxa de quadros de vídeo, que irão influenciar o modo como a conexão ATM será estabelecida e, conseqüentemente, como a recuperação, a transmissão e a apresentação do fluxo multimídia serão executadas.

O modelo proposto por [Richards, 1998], por exemplo, permite que o usuário especifique somente o custo do serviço e o nível de satisfação. Da mesma forma, o modelo de [Silveira, 2000] não prevê a interferência direta do usuário na fase de especificação dos parâmetros de QoS, ou seja, durante a escolha da qualidade de apresentação do vídeo. O usuário pode somente selecionar qual a qualidade mínima aceitável para a apresentação do vídeo escolhido.

Igualmente, não há referência nos outros modelos citados sobre os procedimentos executados pelo Sistema de Banco de Dados Multimídia durante o armazenamento e a recuperação do material multimídia e seus componentes, ou seja, suas correspondentes informações de controle, e sobre a execução do mapeamento dos parâmetros de QoS nas camadas do sistema.

O modelo proposto define um esquema para o gerenciamento do vídeos MPEG-2 armazenados no servidor, para que com base no nível de QoS especificado pelo usuário, a qualidade na qual o vídeo será recuperado, transmitido e apresentado seja calculada. Posteriormente, a Camada de Mapeamento de QoS definida pode dar início ao mapeamento dos parâmetros de QoS em todas as camadas do sistema.

### **1.3. OBJETIVOS DO TRABALHO**

Esta Dissertação de Mestrado apresenta um modelo voltado para aplicações multimídia distribuídas, fundamentadas em servidores de armazenamento multimídia.

O modelo proposto baseia-se em transmissões de fluxos multimídia, mais especificamente vídeos MPEG-2, com Qualidade de Serviço e visa prover facilidades que permitam a interação entre usuário e o conteúdo multimídia, no que diz respeito à especificação do nível de QoS no qual espera-se que o vídeo seja apresentado. Desta forma, também faz parte do escopo do modelo o mapeamento e a negociação dos parâmetros de QoS envolvidos na transmissão.

O modelo utiliza a Tecnologia ATM como meio de transporte, a fim de se beneficiar dos mecanismos oferecidos pela mesma quanto ao processamento dos parâmetros de QoS e, portanto, trata somente da QoS ao nível desta tecnologia. Antes, porém, da apresentação do modelo, a realização de um estudo inicial sobre os principais aspectos atribuídos à mesma torna-se necessária.

### **1.4. ORGANIZAÇÃO DO TRABALHO**

Além deste capítulo introdutório, o trabalho apresenta-se constituído de sete outros capítulos. O Capítulo II trata dos conceitos relacionados às Aplicações Multimídia Distribuídas, dentre os quais podemos citar suas características e seus

requisitos de transmissão. O Capítulo III destina-se a apresentar os conceitos acerca do Padrão de Compressão MPEG-2, tais como principais padrões e processo de compressão de vídeo. O Capítulo IV apresenta os conceitos relativos à Qualidade de Serviço (QoS), assim como o Modelo Genérico de QoS adotado por este trabalho. O Capítulo V descreve a Tecnologia ATM e seus principais elementos. No Capítulo VI são vistos os principais conceitos acerca de Sistema de Gerenciamento de Banco de Dados Multimídia e Servidores de Armazenamento Multimídia. O Capítulo VII destina-se a apresentar os elementos que compõem o modelo proposto neste trabalho de Dissertação de Mestrado. E finalmente, o Capítulo VIII trata das considerações finais que envolvem todo o desenvolvimento do trabalho, tais como os benefícios e desvantagens encontrados na implementação de um modelo da natureza do modelo proposto.

## CAPÍTULO II

### APLICAÇÕES MULTIMÍDIA DISTRIBUÍDAS

*Nos últimos anos muitos esforços têm sido aplicados no que se refere à implementação e ao aperfeiçoamento de aplicações multimídia distribuídas. O aparecimento deste tipo de aplicação ocorreu logo que o conceito de multimídia passou a fazer parte do cenário de processamento de informações. Neste capítulo as principais questões e características referentes a aplicações multimídia distribuídas serão abordadas.*

#### 2.1. INTRODUÇÃO

Atualmente, um dos conceitos mais utilizados no contexto da indústria de processamento de informações é o de **Multimídia**. Seu surgimento aconteceu a partir do desenvolvimento tecnológico dos sistemas de informação e de comunicação à alta velocidade.

Diferentes interpretações podem ser encontradas para o que multimídia realmente significa, mas de uma maneira geral pode-se conceituá-la como o modo de representação das informações através da combinação de vários tipos de mídia, tais como vídeo, áudio, textos e gráficos.

Desta forma, o aparecimento deste conceito apresenta-se como uma evolução natural dos sistemas de informação, já que o mundo real é constituído não só de textos e gráficos, como são baseados os sistemas de informações convencionais, ou seja, os sistemas alfanuméricos, mas também de imagens em movimento e sons.

Existem duas linhas de pesquisa no que se refere aos sistemas multimídia. A primeira trabalha com multimídia implementada em sistemas *standalone*, isto é, o processamento da informação é todo executado no ambiente local. A segunda concentra seus esforços na combinação dos conceitos relacionados à computação multimídia e aos sistemas distribuídos, surgindo assim os sistemas multimídia distribuídos.

Esta última apresenta-se como principal estimuladora para o desenvolvimento de novas aplicações baseadas em sistemas multimídia distribuídos, as quais são de grande importância para a indústria, o meio acadêmico e instituições de padronização, uma vez que possibilita a criação de aplicações do tipo videoconferência, vídeo sob demanda,

ensino a distância entre outras baseadas em tecnologias de comunicação de banda larga (RDSI-FL).

Este capítulo visa a descrição dos aspectos relativos a aplicações multimídia distribuídas, focalizando-se principalmente na descrição dos requisitos de comunicação impostos por estas aplicações aos componentes do sistema, tais como redes e sistemas-fins, envolvidos na transmissão das informações, visto que para que tais aplicações sejam adequadamente executadas determinadas condições devem ser atendidas, como alta largura de banda de transmissão, devido à natureza volumosa dos dados multimídia.

## **2.2. CARACTERÍSTICAS DAS APLICAÇÕES MULTIMÍDIA DISTRIBUÍDAS**

As características de uma aplicação multimídia distribuída determinam qual o tipo desta aplicação e, conseqüentemente, influenciam os requisitos de comunicação que a mesma impõe aos componentes do sistema envolvidos na transmissão das informações [Kwok, 1997].

Desta forma, as principais características destas aplicações serão apresentadas nesta seção, como requisito fundamental para a posterior descrição dos requisitos de comunicação de tais aplicações.

### **2.2.1. TIPO DE INFORMAÇÃO**

Como dito na seção anterior, um sistema multimídia é caracterizado por manipular informações representadas através da combinação de diversos tipos de mídia. Porém, para que um sistema seja efetivamente considerado um sistema multimídia, o mesmo deve ser capaz de manipular digitalmente pelo menos um tipo de mídia estática e um tipo de mídia dinâmica, onde:

- **Mídias Estáticas ou Discretas:** são aquelas baseadas somente na dimensão espaço (textos, imagens, gráficos, exceto gráficos com animações), seus conteúdos e significados não dependem do tempo.
- **Mídias Dinâmicas ou Contínuas ou Isócronas:** são mídias dependentes do tempo (vídeos, áudios e animações), ou seja, seus conteúdos e significados

dependem da taxa com que são apresentadas. Também são chamadas de *mídias contínuas*, pois as amostras<sup>1</sup> destas mídias devem ser recuperadas e apresentadas em intervalos de tempo regulares e devido à relação fixa entre cada unidade de mídia e o tempo, também são chamadas de *mídias isócronas*.

Normalmente, uma aplicação distribuída pode manipular os dois tipos de informação ao mesmo tempo. E desta maneira, as questões de sincronização entre os vários fluxos de informação gerados, na maioria das vezes de fontes distintas, tornam-se de essencial importância.

### 2.2.2. MODO DE ENTREGA DE INFORMAÇÃO

Baseado na forma como a informação é entregue ao destino, a aplicação distribuída pode ser baseada em tempo real ou baseada em tempo não real.

- ***Aplicação em Tempo Real:*** a informação (baseada no tempo ou não) deve ser apresentada imediatamente após a sua chegada ao destino. Por conseguinte, tais aplicações, por exemplo, videofonia, requerem garantias de largura de banda, atraso, variação de atraso e perda de dados para que a entrega seja executada em tempo hábil.
- ***Aplicação em Tempo Não Real:*** a informação é transmitida e, após a sua chegada ao destino, armazenada para posterior apresentação. Este tipo de aplicação requer que o receptor possua suficiente espaço de armazenamento.

Cada tipo de mídia possui seus próprios requisitos de comunicação e características, algumas das quais podem ser consideradas de acordo com a classificação da natureza do tráfego gerado, que pode ser [Soares, 1997]:

- ***Tráfego Contínuo com Taxa Constante (Constant Bit Rate - CBR):*** o tráfego é sempre constante e, conseqüentemente, a taxa média é igual à taxa máxima.

---

<sup>1</sup> Segmentos de informação digitalizada que descrevem a amplitude da informação em função do tempo



- **Tráfego Contínuo com Taxa Variável (Variable Bit Rate - VBR):** o tráfego é contínuo, porém, a taxa de transmissão sofre variações ao longo do tempo.
- **Tráfego em Rajadas (bursty):** o tráfego possui períodos ativos intercalados por períodos inativos. Durante os períodos ativos, há alta geração de informação pela fonte que opera à taxa máxima, contudo, durante os períodos de inatividade, nenhum tráfego é produzido pela fonte.

A seguir, serão classificadas as diversas mídias quanto aos seus requisitos de comunicação, considerando que características, tais como atraso máximo de transferência, variação estatística do atraso, vazão<sup>2</sup> média, entre outras, variam muito de uma mídia para outra.

#### TEXTO

- Normalmente o tráfego é gerado em rajadas.
- Suporta atraso na transferência dos dados e variação estatística do atraso, exceto para aplicações em tempo real.
- Não tolera erro de bits na transferência de dados.

#### IMAGEM GRÁFICA

- Tráfego gerado em rajadas, exceto no caso de imagens gráficas animadas, cujo tráfego é contínuo com taxa constante.
- Suporta atraso na transferência dos dados e variação estatística do atraso.
- Pode ser representada no formato vetorial, ou seja, pelas coordenadas dos segmentos de reta que a compõem e no formato matricial ou bitmap (categoria de imagens gráficas representadas ponto a ponto, isto é, formadas por *pixels*), neste caso representadas por uma matriz de pontos, onde cada componente da

---

<sup>2</sup> Neste trabalho o termo vazão refere-se à taxa de transferência de dados, ou seja, a taxa pela qual os dados são recuperados ou armazenados em um dispositivo de armazenamento.

matriz carrega uma informação de cor do ponto.

- A taxa de erro de bits para imagens no formato matricial e sem compressão pode ser maior que a taxa de erro de pacotes, visto que será praticamente imperceptível se um *pixel* sair com cor diferente da cor original. Porém, o mesmo não ocorre no caso de um bloco inteiro da imagem se perder.
- Não tolera erros de bits em imagens vetoriais e naquelas (vetoriais ou matriciais) onde forem utilizadas técnicas de compressão ou compactação<sup>3</sup>.

## ÁUDIO

- O tráfego gerado para comunicação é do tipo CBR quando não houver compactação ou compressão, caso contrário, o tráfego é do tipo VBR e ainda, na ocorrência de voz com detecção de silêncio, o tráfego é caracterizado como rajadas.
- Independentemente do tipo de tráfego gerado e mesmo quando no sinal de voz for realizada compactação por detecção do silêncio, o sinal deve ser reproduzido no destino a uma taxa constante.
- Suporta valores de atraso em intervalos pré-definidos. A variação estatística do atraso deve ser compensada a fim de que a continuidade dos pacotes (os surtos de voz são divididos em pacotes) seja mantida e, conseqüentemente, também a inteligibilidade da informação no destino. Basicamente, os algoritmos de compensação trabalham de forma a garantir a reserva de pacotes antes do início da reprodução dos mesmos, inserindo, a cada surto de voz, um atraso inicial, limitado pelo atraso de transferência máximo permitido ao surto de voz.
- Os problemas de atraso de transferência máximo são críticos somente em aplicações que exigem comunicação interativa em tempo real. Pois, se o atraso de transferência for grande demais, um efeito de ruptura começa a ocorrer na comunicação, podendo torná-la inviável.

---

<sup>3</sup> *Técnicas de compressão são aquelas cuja redução da taxa gerada de dados causa perda de informação, normalmente imperceptível para ser o humano. Diferentemente, as técnicas de compactação não causam perda de informação.*

- Devido o alto grau de redundância dos sinais de áudio, as taxas de erros de bits ou de pacotes podem ser relativamente altas, desde que os pacotes gerados não sejam muito grandes.

## VÍDEO

- O tráfego pode ser do tipo CBR ou VBR, dependendo da técnica de codificação utilizada.
- Independentemente do tipo de tráfego gerado, o sinal deve ser reproduzido no destino a uma taxa constante.
- Suporta valores de atraso em intervalos pré-definidos. No caso de ocorrer variação estatística do atraso, a mesma deve ser compensada. Como normalmente o vídeo vem acompanhado de áudio, satisfeitos os requisitos de atraso deste, os requisitos de vídeo também serão obedecidos.
- A taxa de erros de bits pode ser maior que a taxa de erro de pacote, no entanto, como a imagem é dinâmica e vários quadros são gerados por segundo, a taxa de erro de pacotes não é tão nociva.

### 2.2.3. SIMETRIA

Uma aplicação distribuída pode ser unidirecional, ou seja, a transmissão ocorre em um único sentido, ou bidirecional, onde a comunicação ocorre nos dois sentidos. As aplicações bidirecionais são mais comuns e podem ser classificadas em *simétricas* e *assimétricas*.

Uma aplicação é simétrica quando nas duas direções as informações transmitidas têm características similares. Caso contrário, a aplicação é classificada como assimétrica.

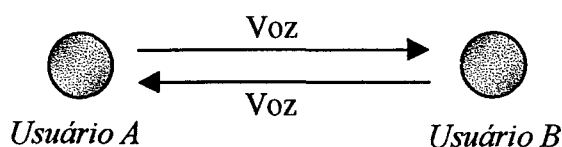
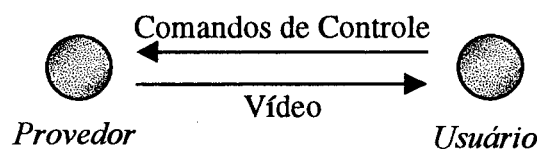


Fig. 2.1 - Aplicação Simétrica

A figura 2.1 ilustra uma aplicação simétrica de chamada de voz, enquanto que a figura 2.2 apresenta uma aplicação assimétrica de vídeo conferência, onde através de uma direção são enviados os comandos de controle e através da outra ocorre a transmissão do vídeo solicitado pelo usuário.



*Fig. 2.2 - Aplicação Assimétrica*

#### 2.2.4. NÚMERO DE PARTICIPANTES

Dependendo do número de participantes envolvidos na comunicação, uma aplicação distribuída pode ser ponto-a-ponto ou multiponto. Uma aplicação ponto-a-ponto envolve somente dois participantes, um em cada extremidade da comunicação, enquanto que em uma aplicação multiponto mais de dois participantes compartilham a mesma comunicação.

A simetria e o número de participantes envolvidos em uma comunicação são características que descrevem a natureza da conexão estabelecida por uma aplicação distribuída, logo influenciam fortemente os requisitos de comunicação desta aplicação. A tabela abaixo apresenta algumas aplicações com diferentes combinações para a simetria e o número de participantes.

NÚMERO DE PARTICIPANTES	SIMETRIA DA CONEXÃO	
	SIMÉTRICA	ASSIMÉTRICA
PONTO-A-PONTO	<i>Chamada de voz</i>	<i>Vídeo sob demanda</i>
MULTIPONTO	<i>Videoconferência</i>	<i>BroadcastTV</i>

*Tab. 2.1 – Simetria e Número de Participantes*

### **2.3. REQUISITOS DE COMUNICAÇÃO DAS APLICAÇÕES MULTIMÍDIA DISTRIBUÍDAS**

Para que o adequado funcionamento de uma aplicação multimídia distribuída seja alcançado, as condições impostas por estas aplicações ao sistema devem ser atendidas. Porém, não só as aplicações impõem condições ao sistema, mas também os usuários destas aplicações, principalmente no que se refere à qualidade da apresentação, tolerância à falhas e tempo de resposta.

Desta forma, pode-se dizer que os requisitos de comunicação de uma aplicação multimídia distribuída são especificados em termos dos requisitos associados ao tráfego gerado pela aplicação e dos requisitos associados às exigências dos usuários em relação à qualidade do serviço fornecido. Tais requisitos especificam a forma como os dados serão enviados através da rede, quais os recursos que devem ser alocados para que o tráfego multimídia seja adequadamente suportado, quais as condições de atraso e erro dos pacotes enviados através da rede devem ser atendidas, etc.

Todos estes requisitos devem ser satisfeitos para que o adequado funcionamento do sistema multimídia e o resultado esperado do processamento das mídias sejam alcançados, resultando na inteira satisfação do usuário quanto à qualidade do serviço solicitado.

Os principais requisitos de comunicação estabelecidos pelas aplicações multimídia distribuídas serão apresentadas nesta seção. Visto que o objetivo geral deste trabalho é o estudo de aplicações multimídia distribuídas, a partir desta seção, tais aplicações serão referenciadas somente como aplicações multimídia.

#### **2.3.1. LARGURA DE BANDA E ESPAÇO DE ARMAZENAMENTO**

Devido à grande quantidade de dados associada ao tráfego multimídia, principalmente no que se refere às mídias contínuas, a rede de comunicação deve prover dispositivos de armazenamento de alta capacidade e alta largura de banda de transmissão (ou alta taxa de transmissão), uma vez que o tráfego multimídia, especialmente o tráfego de vídeo, consome a maioria da largura de banda da rede, quando informações multimídia são transmitidas [Raghavan, 1998; Zheng, 1999].

Contudo, somente aplicações em tempo real têm rigorosas restrições quanto à alocação de largura de banda, visto que para aplicações em tempo não real a largura de banda necessária para a transmissão das informações não é uma questão tão relevante, pois a mesma é bastante baixa se comparada à largura de banda total da rede.

### 2.3.2. SUPORTE À COMPRESSÃO DE DADOS

Para que as informações multimídia sejam devidamente transmitidas do emissor ao receptor, a rede deve ter implementado mecanismos que controlem a alocação de largura de banda da rede para cada sessão estabelecida [Fairhurst, 1999].

Para que estes requisitos sejam minimizados, o sistema multimídia pode implementar técnicas de compressão de dados, cujo princípio básico é reduzir ao máximo, sem comprometer o conteúdo da informação, o número de bits utilizados na representação da mesma, de maneira a diminuir o volume de armazenamento e limitar a taxa de bits necessária para transmissão através da rede [Fluckiger, 1995].

As técnicas de compressão implementam algoritmos de compressão de dados genéricos aplicáveis a qualquer tipo de dado. As técnicas de compressão de vídeo exploram a redundância espacial (a correlação de pontos vizinhos dentro de uma imagem) e as particularidades da percepção visual humana (carência de sensibilidade do olho a determinados detalhes), tanto quanto as imagens estáticas e a alta redundância temporal existente entre imagens sucessivas, como a imagens em movimento. Similarmente, as técnicas de compressão de áudio exploram as particularidades da percepção auditiva humana, a fim de reduzir as taxas de bits através da eliminação de informações audíveis [Benoit, 1997].

Em geral, os termos compressão e codificação são utilizados com o mesmo sentido, sendo codificação o termo mais genérico. Basicamente, o processo de compressão de dados consiste na compressão da informação na fonte (emissor) e descompressão da mesma, após transmissão na rede, no destino (receptor), ver figura. 2.3.



Fig. 2.3 – Visão Simplificada do Processo de Compressão de Dados

## **COMPRESSÃO À TAXA DE BITS CONSTANTE OU À TAXA DE BITS VARIÁVEL**

Durante a conversão de sinais de vídeo e áudio analógicos para a forma digital, as amostras destas mídias são feitas em intervalos de tempo regulares, cada qual representada através do mesmo número de bits. Em vista disso, vídeo e áudio não codificados são fluxos gerados a taxa de bits constante, de maneira que, em um intervalo de tempo fixo, a mesma quantidade de dados para um fluxo em particular será encontrada.

Algumas técnicas de compressão produzem fluxos à taxa de bits constante e outras não. As primeiras são chamadas *técnicas de compressão à taxa de bits constante (CBR)*, as demais são chamadas *técnicas de compressão à taxa de bits variável (VBR)*.

No momento da implementação de uma ou outra técnica de compressão, esta classificação deve ser considerada por dois motivos. Primeiro, o conteúdo das mídias varia com o tempo. Quanto mais complexo for o conteúdo, mais dados serão necessários para representá-lo e vice-versa. Para tanto, técnicas de compressão VBR são mais eficientes e produzem vídeo e áudio de maior qualidade a uma mesma taxa de compressão. Segundo, a especificação e a modelagem de fluxos VBR são mais trabalhosas e em virtude disso, os sistemas de comunicação multimídia têm dificuldade em suportá-los [Lu, 1996].

No capítulo que será apresentado a seguir, o Padrão de Compressão MPEG-2 será analisado mais detalhadamente devido a sua grande utilização para codificação de vídeo armazenado e, principalmente, devido a sua escalabilidade, característica muito importante para o desenvolvimento do trabalho, desde que permite o tratamento das modificações da qualidade do serviço solicitado pelo cliente e, por conseguinte, a execução de adaptações nas mídias já armazenadas.

### **2.3.3. ATRASO FIM-A-FIM E VARIAÇÃO DE ATRASO**

Como dito na seção 2.2.1, as mídias contínuas são dependentes do tempo e, por conseguinte, para se obter apresentação de boa qualidade as amostras destas mídias devem ser recuperadas e apresentadas continuamente na mesma taxa em que foram

amostradas. Desta forma, a qualidade da apresentação depende do valor das amostras e do intervalo em que as mesmas são apresentadas.

Normalmente, em sistemas multimídia distribuídos a informação é capturada em um ambiente do sistema e apresentada em outro. A soma dos atrasos durante o período em que a informação está em transição, inclusive dos atrasos proporcionados por todos os componentes do sistema pelo qual a informação trafega, tais como decodificadores e roteadores, refere-se ao *atraso fim-a-fim* ou *latência*.

Baseando-se em redes de comutação de pacotes, observa-se que cada pacote transmitido chega ao destino em intervalos de tempo variáveis, o que é chamado de *variação de atraso* ou *delay jitter*.

Os requisitos de atraso e variação de atraso só se aplicam às aplicações baseadas em tempo-real, principalmente para aquelas cuja interatividade é implementada, uma vez que para tais aplicações a continuidade da apresentação deve ser preservada [Kwok, 1997; Hafid, 1996].

Para que os requisitos necessários para a transmissão de mídias contínuas sejam satisfeitos, a rede deve manter o atraso e a variação de atraso em intervalos de tempo pequenos e pré-definidos, favorecendo a devida sincronização das mídias no destino.

Se o intervalo especificado para o atraso for ultrapassado, a qualidade da apresentação é prejudicada, pois os pacotes atrasados são considerados pacotes perdidos e, portanto, simplesmente descartados (as questões a cerca da retransmissão dos dados perdidos serão tratadas na seção 2.3.5). Da mesma forma, se a variação de atraso for muito grande, a informação multimídia não pode ser apresentada imediatamente após a sua chegada ao destino e, conseqüentemente, algum mecanismo, tal como o armazenamento, deve ser executado, a fim de adequar o fluxo da informação à apresentação.

Para o usuário estes erros podem ser percebidos como distorções na imagem e no som, falta de sincronismo entre as mídias envolvidas, redução da frequência de exibição e/ou até mesmo paralisação da apresentação. Por exemplo, para determinadas aplicações de telefonia, o atraso de até aproximadamente 250 milisegundos pode ser tolerado, além deste valor, a voz começará a ficar distorcida [Liu, 1998].



O tempo de atraso aceitável depende da aplicação em questão. Para aplicações de conversação ao vivo, os requisitos para que a interatividade natural da conversação seja mantida são muito mais rígidos se comparados aos requisitos impostos por uma aplicação de recuperação de informações. Contudo, para a maioria das aplicações, um tempo de resposta aceitável não pode passar de poucos segundos.

#### 2.3.4. RELAÇÃO TEMPORAL E ESPACIAL ENTRE AS MÍDIAS

Por definição, o tráfego gerado por uma aplicação multimídia envolve várias mídias. Estas devem ser transmitidas e recuperadas de maneira coordenada a fim de que a relação temporal e espacial existente entre elas seja mantida e que, conseqüentemente, a qualidade esperada da apresentação seja obtida. Para tanto, além dos requisitos de atraso e variação de atraso, acima descritos, também determinados requisitos de sincronização devem ser satisfeitos.

O objetivo da sincronização é eliminar todos os atrasos e variações de atraso introduzidos nos vários fluxos de mídia durante a transmissão e manter estes fluxos sincronizados [Furht, 1994]. Existem dois tipos de sincronização: *sincronização espacial* e *sincronização temporal*. A sincronização espacial, ilustrada na figura 2.4, refere-se à correta apresentação dos dados multimídia nos dispositivos de saída (por exemplo, monitor), em um determinado momento, ou seja, de acordo com as especificações do autor da aplicação sobre as coordenadas em que as mídias devem ser apresentadas, tratando do tamanho, da rotação e do posicionamento destes dados no dispositivo.

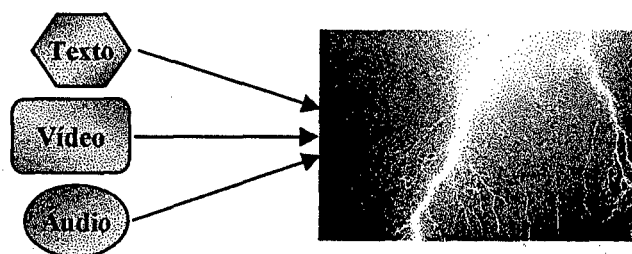


Fig. 2.4 – Sincronização Espacial

A sincronização temporal trata da sincronização dos elementos de dados multimídia, por exemplo, um fluxo de vídeo, um quadro de um fluxo de vídeo ou um pixel de um

quadro de vídeo, nos dispositivos de saída, mantendo a ordem temporal de apresentação destes elementos, além de regras para a exibição, inibição e edição dos elementos de acordo com a especificação do autor. Segundo [Hafid, 1996], a sincronização temporal pode ser classificada em: sincronização de fluxo, sincronização baseada em evento e sincronização de grupo.

- **Sincronização de Fluxo:** trata das questões de atraso e variação de atraso no que se refere a um único fluxo (sincronização intrafluxo), bem como da manutenção da relação temporal existente entre dois ou mais fluxos durante a fase de apresentação (sincronização interfluxo). Como visto a sincronização de fluxo pode ser classificada em:

- **Sincronização Intrafluxo (Intramídia):** independentemente de o fluxo ser isócrono ou não, a sincronização intramídia equivale à relação temporal existente entre dados sucessivos dentro de um mesmo fluxo. Se o fluxo for isócrono, o intervalo de apresentação de dados consecutivos deve ser fixo. A figura abaixo ilustra a transmissão de quadros de vídeo, onde conforme a sincronização intramídia, os mesmos devem ser apresentados a cada  $T$  ms.

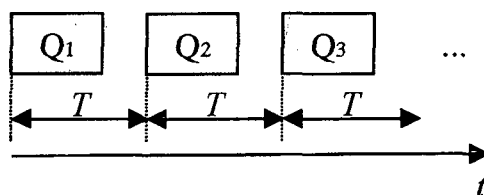
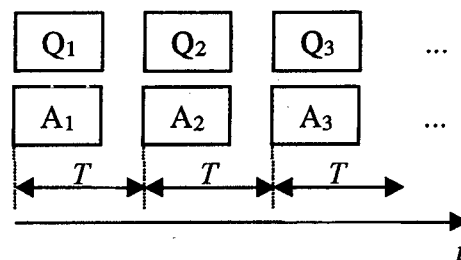


Fig. 2.5. Sincronização Intrafluxo

- **Sincronização Interfluxo (Intermídia):** a sincronização intermídia também independe do fluxo ser isócrono ou não; a mesma diz respeito à relação temporal entre diferentes fluxos, uma vez que cada fluxo, quando transmitido através da rede, está sujeito aos atrasos introduzidos por cada componente pelo qual trafega, além de poder seguir caminhos diferentes. A mesma também é aplicada nos casos onde existem duas ou mais instâncias de um mesmo tipo de mídia, tal como sincronização de dois fluxos de áudio. A figura 2.6 apresenta a sincronização interfluxo, onde cada quadro de vídeo

transmitido deve ser sincronizado simultaneamente a sua respectiva amostra de áudio.



*Fig. 2.6. - Sincronização Interfluxo*

- **Sincronização Baseada em Eventos:** permite que a aplicação multimídia, em resposta às notificações de eventos ocorridos durante a apresentação das mídias, execute determinadas ações. Contudo, a execução destas ações introduzirá ao sistema rigorosas restrições temporais, uma vez que as mesmas podem afetar a construção de valores de dados (por exemplo, modificação do documento multimídia) ou o comportamento temporal da apresentação de mídias contínuas (por exemplo, operações de pausa, retrocesso e avanço).
- **Sincronização de Grupo:** atinge aplicações cujo conteúdo multimídia deve ser apresentado a diversos usuários, tais como aplicações de ensino a distância, onde cada participante deve ter simultaneamente a mesma visão da janela que está sendo compartilhada. A sincronização de grupo é especialmente importante quando tais aplicações visam a execução de operações de interação entre determinado número de usuários, os quais devem receber a mesma informação exatamente ao mesmo tempo, como é o caso das aplicações de teleconferência.

Nestes termos, no que se refere à sincronização multimídia, existem duas áreas de trabalho. A primeira propõe o desenvolvimento de mecanismos e ferramentas que ofereçam aos autores de aplicações uma maneira fácil de especificar a relação temporal entre as mídias envolvidas na aplicação. A segunda visa garantir a relação temporal especificada de forma a acabar com a natureza indeterminística dos sistemas de comunicação.

### 2.3.5. TOLERÂNCIA A ERROS E PERDAS DE DADOS MULTIMÍDIA

O propósito de uma aplicação multimídia distribuída é transmitir, através da rede, informações multimídia entre um ou mais usuários, qualquer perda ou erro que venha a acontecer nos dados transmitidos irá afetar a apresentação da informação, pois a entrega dos dados não será completamente executada. Desta forma, a aplicação multimídia está exposta a um conjunto de requisitos de erros e perdas de dados.

Ao contrário do que acontece com as aplicações alfanuméricas, cuja perda ou erro de dados não é permitido, para alguns tipos de aplicações multimídia isto pode ser, até certo ponto, tolerado, já que para o ser humano pequenos erros de bits e perdas de dados de vídeo, áudio e imagens digitais são imperceptíveis.

Comumente, as aplicações baseadas na transmissão de fluxo em tempo real podem permitir uma quantidade limitada de perda de dados (dependendo da capacidade de recuperação do decodificador e do tipo de aplicação) e, contrariamente, para aplicações baseadas na transmissão de blocos em tempo real e aplicações em tempo não real a perda de dados não é tolerada, o que significa que se isto ocorrer, os dados devem ser, na medida do possível, recuperados.

No caso de aplicações em tempo real, principalmente aplicações baseadas na transmissão de blocos em tempo real, a recuperação de erros e perda de dados através da retransmissão não é a solução mais adequada, devido o atraso acrescentado. Como a percepção humana é falha em relação a certos erros de dados multimídia, tais aplicações podem ser projetadas de maneira a tirar proveito deste fato e, por conseguinte, evitar a retransmissão e o aumento da largura de banda de transmissão proporcionado pelo *overhead* adicional. Todavia, para aplicações tempo não real os efeitos causados pela retransmissão (atrasos e *overhead*) não são uma questão importante, visto que tais aplicações não têm restrições temporais e consomem, em geral, o mínimo de largura de banda.

Independentemente do tipo de aplicação implementada, os requisitos para a taxa de perda e erros de pacotes de dados são mais rigorosos do que os impostos pela taxa de bits, pois a perda de pacotes pode afetar profundamente a apresentação da mídia.

Contudo, quando técnicas de compressão de dados são utilizadas, a taxa de erros de bits permitida é muito baixa. Normalmente, quanto maior a taxa de compressão, mais baixa

é a tolerância a erros, uma vez que mais informações são transportadas por bit comprimido de informação.

### **2.3.6. SUPORTE A COMUNICAÇÕES MULTICAST E BROADCAST**

Na maioria das vezes as aplicações multimídia são multiponto, ou seja, a transmissão de uma informação é compartilhada por vários usuários. Para tais aplicações, é importante que a entrega das informações a todos os destinos seja feita em um intervalo de tempo aceitável. Para tanto, é essencial que a rede suporte comunicações *multicast* ou *broadcast* a taxa de dados e atraso fim-a-fim específicos.

A implementação de multicast e broadcast resulta na redução do tempo e dos recursos necessários, tais como largura de banda e memória, a entrega de um mesmo fluxo de informação a vários usuários, e na diminuição do processamento envolvido com a transmissão e do *overhead* gerado pelos protocolos.

### **2.4. QUALIDADE DE SERVIÇO EM SISTEMAS MULTIMÍDIA DISTRIBUÍDOS**

De acordo com o que foi dito na seção 2.3, os requisitos de comunicação de uma aplicação multimídia distribuída estão intimamente ligados às condições, de transmissão e de apresentação das informações multimídia, impostas ao sistema multimídia.

O conceito de *Qualidade de Serviço (QoS)* foi apresentado como uma proposta para estruturar e garantir o cumprimento destes requisitos, denominados a partir de então como *Requisitos de Qualidade de Serviço* ou somente *Requisitos de QoS*.

A QoS pode ser especificada como um conjunto de parâmetros de requisitos ou simplesmente *Parâmetros de QoS*. Além dos parâmetros de QoS comuns a qualquer sistema, em sistemas multimídia distribuídos existem aqueles ligados direta (por exemplo, resolução e frequência de quadros em uma animação) ou indiretamente (por exemplo, atraso fim-a-fim, variação de atraso e taxa de erros de bits) a qualidade da apresentação dos fluxos multimídia no receptor.

Para que os requisitos fim-a-fim das aplicações multimídia sejam suportados, todos os componentes do sistema envolvidos, entre eles, sistemas-fins e comutadores, devem implementar determinado nível de QoS.

Os parâmetros de QoS referentes a estes componentes podem ser caracterizados por parâmetros como, atraso de transmissão, variação de atraso de transmissão, taxa de perda de células, entre outros, uma vez que também executam funções específicas de processamento de fluxo, as quais geram pelo menos um fluxo de entrada e outro de saída, cada qual com propriedades de tempo real diferentes.

Quando mais de um fluxo, por exemplo, um fluxo de vídeo e outro de áudio, é gerado por somente uma sessão, existe a possibilidade destes fluxos serem transmitidos através de um único canal ou através de vários canais, um para cada fluxo [Hafid, 1996]:

- *Único Canal:* se forem transmitidos através de um único canal, a relação temporal entre os fluxos de mídias deve ser mantida durante toda a transmissão, sendo necessário apenas um pequeno número de pacotes de sinalização. Contudo, existem algumas desvantagens agregadas a esta abordagem:
  - Apesar de cada mídia impor seus próprios requisitos de QoS, somente os requisitos associados a uma única mídia podem ser estabelecidos por canal. Por conseguinte, para que a QoS de todas as mídias seja satisfeita, a QoS reservada ao canal deve é basear-se na qualidade da mídia com requisitos mais altos, o que acarreta em certa ineficiência no uso dos recursos da rede.
  - No caso de interrupção do canal, a transmissão da informação é paralisada, desfavorecendo determinados tipos de aplicação, tal como videofonia, onde é melhor receber parte da informação do que nenhuma.
  - Quando técnicas de multiplexação forem utilizadas, é provável que algumas dificuldades apareçam, como por exemplo, necessidade de utilização, no destino, de mecanismos complexos de reconstrução de informações.
  - Existe a impossibilidade da execução de processamento paralelo das mídias envolvidas.
- *Vários Canais:* quando cada fluxo de mídia é transmitido através de seu próprio canal, a QoS associada aos canais refere-se aos requisitos de QoS necessários à transmissão da mídia em questão. Desta maneira, mecanismos de sincronização devem ser utilizados no destino, de forma que a relação temporal e espacial entre

os diferentes fluxos de mídia sejam mantidas, uma vez que cada fluxo pode ser conduzido por caminhos diferentes, ocasionando diferentes atrasos e variações de atrasos de transmissão.

De acordo com [Hafid, 1996], existem casos onde a melhor solução é aquela que utiliza a integração de ambas as abordagens. No capítulo III, o tema Qualidade de Serviço será abordado em maiores detalhes.

## **2.5. CONSIDERAÇÕES FINAIS**

Este capítulo abordou as principais características acerca das aplicações multimídia distribuídas. Dentre estas características destacam-se os requisitos impostos por tais aplicações no que se refere ao suporte a comunicação, a necessidade da utilização de técnicas de compressão de dados, como forma de minimizar os requisitos de espaço de armazenamento e largura de banda da rede e, finalmente, a necessidade da aplicação dos conceitos de QoS. Como faz parte do escopo do trabalho abordar a transmissão com QoS de fluxos multimídia MPEG-2, as próximas seções destinam-se ao estudo dos principais conceitos e características do Padrão MPEG-2 e do suporte à QoS.

## CAPÍTULO III

### PADRÃO DE COMPRESSÃO MPEG-2

*Este capítulo tem como finalidade descrever alguns aspectos relativos ao Padrão de Compressão MPEG-2 e aos três dos principais padrões que o compõe – Video, Audio e Systems. Atualmente, o mesmo está sendo amplamente estudado como forma de compressão de tráfego de informações multimídia.*

#### 3.1. INTRODUÇÃO

O Padrão MPEG-2 “*Generic Coding of Moving Pictures and Associated Audio*” foi o segundo padrão de compressão desenvolvido pelo Grupo MPEG “*Moving Picture Experts Group*”. Iniciados em 1990, os estudos em relação ao padrão tiveram seus primeiros resultados publicados quatro anos mais tarde sob a referência ISO/IEC 13818. Na verdade, o padrão MPEG-2 é composto por uma família de padrões, a qual abrange todas as diferentes questões relativas à representação, transmissão e entrega de vídeo e áudio digitais.

O esquema de compressão de dados implementado pelo padrão foi projetado para ser genérico e flexível, a fim de atender os requisitos de uma gama de aplicações multimídia e tornar a troca de fluxos de bits entre diferentes aplicações, transmissões e meios de armazenamento mais simples de ser executada. Além disso, funcionalidades importantes, tais como codificação eficiente de imagens entrelaçadas e escalabilidade, são consideradas, uma vez que o padrão lida também com transmissões de televisão digital.

A não padronização do processo de compressão e o conseqüente baixo custo associado aos decodificadores MPEG-2, abriram um leque de possibilidades para o desenvolvimento de aplicações baseadas na alta qualidade de vídeo proporcionada pelo padrão. Atualmente, os Padrões MPEG-2 estão sendo empregados em serviços de vídeo digital e na maioria dos sistemas de transmissão digital via satélite. Especificamente, o Padrão MPEG-2 Video está sendo adotado como o padrão de compressão de vídeo para aplicações de televisão de alta definição (HDTV).



As seções seguintes destinam-se à descrição dos principais aspectos relacionados à compressão, multiplexação e reconstrução dos fluxos de vídeo e áudio no destino.

### 3.2. FAMÍLIA DE PADRÕES MPEG-2

Apesar de ser, na maioria das vezes, associado somente à compressão de vídeo, o Padrão MPEG-2 também considera a representação e transmissão de vídeo e áudio digitais. Suas funcionalidades são executadas por oito padrões (partes) distintos, os quais encontram-se listados a seguir. Os mais importantes aspectos relacionados aos três principais padrões – *Video*, *Audio* e *Systems* – serão descritos nas próximas seções.

PADRÃO	DESCRIÇÃO
MPEG-2 Systems (ISO/IEC 13818-1)	Define a multiplexação de um ou mais fluxos de vídeo, áudio e outros dados em um único fluxo para transmissão e armazenamento
MPEG-2 Video (ISO/IEC 13818-2)	Define o formato e a codificação dos fluxos de vídeo
MPEG-2 Audio (ISO/IEC 13818-3)	Define o formato e a codificação dos fluxos de áudio
MPEG-2 Compliance Testing (ISO/IEC 13818-4)	Determina como os testes, que verificam se os fluxos de bits e decodificadores estão satisfazendo os requisitos especificados pelos três padrões anteriores, devem ser projetados
MPEG-2 Software Simulation (Future TR) (ISO/IEC 13818-5)	É considerado um relatório técnico, que descreve a implementação de software completa dos três primeiros padrões
MPEG-2 Digital Storage Media Command and Control Protocol – DSM-CC (ISO/IEC 13818-6)	Especifica um conjunto de protocolos, que fornece operações e funções de controle específicas para o gerenciamento dos fluxos de bits MPEG
MPEG-2 Real Time Interface for Systems Decoders (ISO/IEC 13818-9)	É a especificação da interface de tempo-real (RTI – <i>Real Time Interface</i> ) para decodificadores de Fluxo de Transporte.
MPEG-2 DSM Reference Script Format (ISO/IEC 13818-10)	É a parte referente ao teste de configuração do DSM-CC. Este padrão ainda está sob desenvolvimento

Tab. 3.1 – Família de Padrões MPEG-2

### 3.3. PADRÃO MPEG-2 VIDEO

O Padrão MPEG-2 *Video* não define o processo de compressão, o mesmo é responsável somente pela definição de um formato (sintaxe) que possa descrever o fluxo de bits de vídeo gerado por este processo e pela definição das informações necessárias à

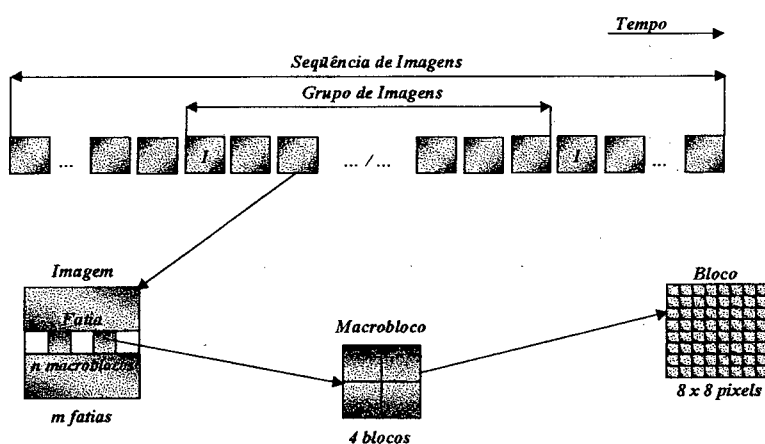
interpretação destes fluxos pelo decodificador (semântica de decodificação). Desta forma, cada aplicação pode utilizar o processo de compressão que mais se adequar aos seus requisitos e limitações.

A primeira vista, a não especificação do processo de compressão pode parecer problemática, todavia, os Padrões MPEG, de um modo geral, seguem o princípio de especificar o mínimo possível, a fim de manter a interoperabilidade e tornar o processo aberto a melhorias, tais como redução do tempo de codificação e aumento da qualidade da imagem [Chiariglione, 1996; Orzessek, 1998].

Dentre todas as funcionalidades atribuídas ao padrão, o processamento de imagens entrelaçadas, a escalabilidade e a introdução do conceito de níveis e perfis são as que mais se destacam. Estas e outras funcionalidades serão apresentadas nas seções seguintes.

### 3.3.1. OBJETOS ELEMENTARES

O Padrão MPEG-2 *Video* lida com um número de objetos elementares, usados na estruturação da informação de vídeo. Uma seqüência de vídeo MPEG é constituída por uma hierarquia de camadas, onde cada camada é responsável por uma ou mais funções específicas no processo MPEG de compressão de vídeo, ver figura 3.1.



**Fig. 3.1 – Hierarquia de Camadas de Vídeo MPEG**

- **Seqüência de Imagens:** é a camada mais alta da hierarquia; representa um número de imagens de vídeo ou um grupo de imagens de vídeo (e não o vídeo inteiro) e

define o contexto válido para toda a seqüência, tal como parâmetros básicos de vídeo.

- **Grupo de Imagens (GOP – Group of Pictures):** um GOP é formado por uma seqüência de vídeo e sempre começa com uma imagem do tipo I, também permite o acesso randômico a uma seqüência de imagens.
- **Imagem:** é a unidade de apresentação fundamental, a qual pode ser de três tipos (I, P ou B), como será visto na próxima seção.
- **Fatia:** uma fatia permite acesso randômico dentro de uma imagem; é composta por uma série de macroblocos e contém informações sobre o posicionamento dos mesmos na tela. Na presença de erros de transmissão e perda de dados da imagem, as fatias podem ser utilizadas para dar continuidade ao processo de apresentação, pois o decodificador, ao invés de ignorar toda a imagem, é capaz de dar segmento a apresentação a partir do início da próxima fatia. Na verdade, para a compressão de dados as fatias não são necessárias, as mesmas são somente codificadas para que na imagem existam pontos de re-sincronização e endereçamento intraquadro.
- **Macrobloco:** é utilizado para compensação e avaliação de movimento. Cada macrobloco tem tamanho igual a 16 x 16 pixels e é composto por quatro blocos com informações de luminância<sup>1</sup> e, dependendo do formato das amostras utilizadas na digitalização do vídeo, mais um certo número de blocos com informações de crominância<sup>2</sup>.
- **Bloco:** uma imagem é dividida em um número de blocos, onde cada bloco contém oito linhas e cada linha contém oito amostras de valores de pixel de luminância e crominância de um quadro. Desta forma, um bloco é definido por 64 valores de pixel (8 x 8 pixels) de luminância e crominância .

---

<sup>1</sup> Refere-se à condição de luminosidade da imagem

<sup>2</sup> Subportadora de cor; tom e saturação de uma imagem representada pelos sinais U (tonalidade), V (saturação) e Y (luminescência)

### 3.3.2. TIPOS DE IMAGEM MPEG

O Padrão MPEG define três tipos diferentes de imagem, ver figura 3.2.

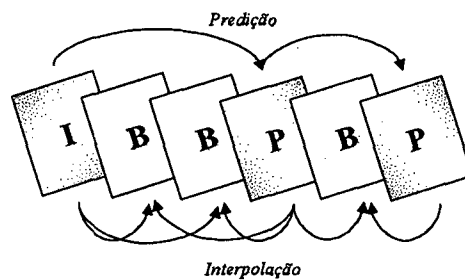


Fig. 3.2 – Concatenação das Imagens MPEG

- **Imagens I (Intracoded):** são codificadas sem qualquer referência às outras imagens (P e B), visto que contém todas as informações necessárias à sua própria reconstrução e à reconstrução de imagens P e B pelo decodificador, conseqüentemente, a taxa de compressão para estas imagens é relativamente baixa. As imagens do tipo I são ponto de entrada essencial para o acesso a uma seqüência de imagens, pois o processo de decodificação não pode ser iniciado até que uma imagem deste tipo seja recebida e, portanto, as mesmas devem aparecer regularmente na seqüência.
- **Imagens P (Predictive):** através de técnicas de compensação de movimento com predição, estas imagens são codificadas através de informações obtidas por imagens I ou P anteriormente apresentadas, as quais recebem o nome de *Imagens de Referência*. O número de imagens P entre duas imagens I não pode ser muito extenso. A taxa de compressão destas imagens é significativamente maior do que a taxa proporcionada por imagens I.
- **Imagens B (Bidirectional):** são codificadas através da interpolação entre imagens I ou P, as quais são apresentadas anteriormente ou posteriormente às imagens B. Estas não difundem erros de codificação, pois não são utilizadas na codificação de outras imagens. Logo, a taxa de compressão obtida é a mais alta dentre os três tipos de imagens MPEG.

### 3.3.3. PRINCÍPIOS DO PROCESSO DE COMPRESSÃO DE VÍDEO MPEG-2

O processo de compressão de vídeo MPEG-2 baseia-se em duas características. Primeiro, em como o olho humano e os centros visuais no cérebro reconhecem as imagens. E segundo, na identificação de regiões similares em imagens adjacentes dentro de uma seqüência de imagens.

Através da **codificação intraquadro** e da **codificação interquadro**, respectivamente, o mesmo também explora as redundâncias espacial (correlação entre pontos vizinhos dentro de uma imagem) e temporal (correlação entre imagens sucessivas, no caso de imagens em movimento) freqüentemente encontradas entre os quadros transmitidos em um fluxo de vídeo.

A codificação intraquadro codifica cada quadro separadamente, enquanto que a codificação interquadro, tira vantagem do fato de que quadros sucessivos são freqüentemente similares, a fim de obter maior nível de redução da quantidade de informação necessária para a transmissão ou o armazenamento.

Na presença de imagens em movimento, a correlação temporal entre os quadros é muito baixa e muitas vezes inexistente. Nestas situações, o fluxo de vídeo assemelha-se a um conjunto de imagens estáticas não relacionadas e, desta maneira, para que compressão de dados eficiente seja obtida, a correlação espacial e, portanto, a codificação intraquadro apresenta-se como a técnica mais adequada.

#### **DCT (DISCRETE COSINE TRANSFORM)**

Uma vez que o olho humano, devido a sua estrutura interna, apresenta-se insensível a determinadas particularidades de cor da imagem, principalmente, a mudanças de cores, o Padrão MPEG-2 utiliza um método baseado no DCT, na identificação e remoção destas particularidades, aproximando as informações de luminância e crominância em cada bloco.

Um conjunto de coeficientes de freqüência, os quais descrevem as mudanças de cor em cada bloco, é calculado e utilizado no lugar dos valores de cor originais. O processo DCT é reversível, o que significa que não há perda da informação.

## PROCESSO DE QUANTIZAÇÃO

Porém, sem que a qualidade da imagem seja prejudicada, é possível que os coeficientes abaixo de um determinado valor sejam eliminados e substituídos por zero. Os coeficientes restantes, isto é, aqueles não substituídos por zero, são quantizados com uma precisão que decresce de acordo com o aumento das frequências espaciais, pois na presença de altas frequências espaciais, a visão humana tem sensibilidade reduzida.

Este processo é chamado **Processo de Quantização** e os coeficientes resultantes são obtidos através de uma matriz denominada **Matriz de Quantização**, ver figura 3.3. Como, o processo de quantização é irreversível, visto que as informações da imagem que foram removidas não podem mais ser recuperadas, o processo de compressão MPEG-2 é considerado um processo de compressão de dados com perdas.

1033	7	20	12	37	31	-2	0	⇒	129	1	1	1	2	1	0	0
-4	12	10	-26	-12	8	7	13		0	1	1	-1	-1	0	0	0
-8	-4	-1	-2	6	-18	0	9		0	0	0	0	0	-1	0	0
0	-16	-11	14	-1	-3	15	2		0	-1	0	1	0	0	0	0
-17	11	26	-3	-1	4	-3	7		-1	0	1	0	0	0	0	0
20	11	0	-2	-7	-8	3	4		1	0	0	0	0	0	0	0
24	-9	20	8	-6	24	6	11		-1	0	1	0	0	1	0	0
11	11	2	8	-3	-10	-7	0		0	0	0	0	0	0	0	0

Fig. 3.3 – Matriz de Quantização

Após a execução deste processo, os coeficientes são lidos em zig-zag (figura 3.4), transformando a matriz de coeficientes em um fluxo de bits serial. Como o processo de quantização produz um grande número de zeros no espectro de altas frequências (por exemplo, 2, 0, 0, 1, 0, 0,1), a leitura em zig-zag permite que estas longas séries de coeficientes nulos sejam obtidos o quanto antes, a fim de contribuir para a eficiência da próxima técnica a ser aplicada, no caso a técnica VCL (*Variable Length Coding*).

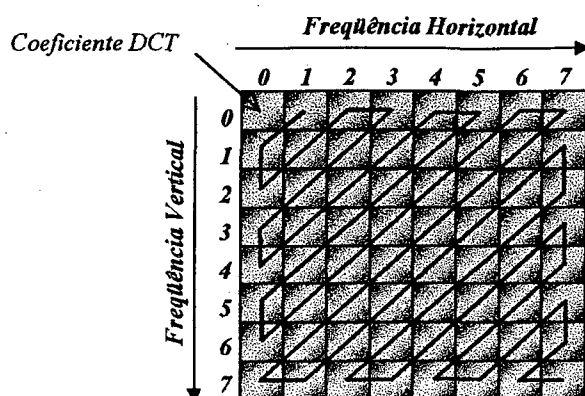


Fig. 3.4 – Leitura em Zig-Zag

### VLC (VARIABLE LENGTH CODING)

A técnica VCL utiliza uma tabela de conversão para a codificação com códigos mais curtos (somente poucos bits) dos coeficientes que aparecem com maior frequência na sequência de coeficientes e com códigos mais longos dos que aparecem com menor frequência. O Padrão MPEG-2 define um número de tabelas com códigos, onde cada código destina-se a codificação de um padrão de coeficientes específico.

### AVALIAÇÃO DE MOVIMENTO

Partindo do princípio de que imagens sucessivas, muito provavelmente, contém regiões similares, o Padrão MPEG-2 *Video* utiliza-se de técnicas de avaliação de movimento no processo de compressão de vídeo, as quais, através da definição de um **Vetor de Movimento**, baseiam-se na identificação de regiões similares em imagens adjacentes.

A avaliação de movimento é feita ao nível de macroblocos. Para cada macrobloco, o codificador pesquisa nas imagens anteriores, no caso de imagens P e anteriores e posteriores, para imagens B, por um macrobloco similar ao atual. Se encontrado, a diferença na posição entre os macroblocos é calculada, codificada através do processo DCT e, posteriormente, pela técnica VCL juntamente com o vetor de movimento do macrobloco em questão, num processo chamado **Compensação de Movimento**, ver figura 3.5.

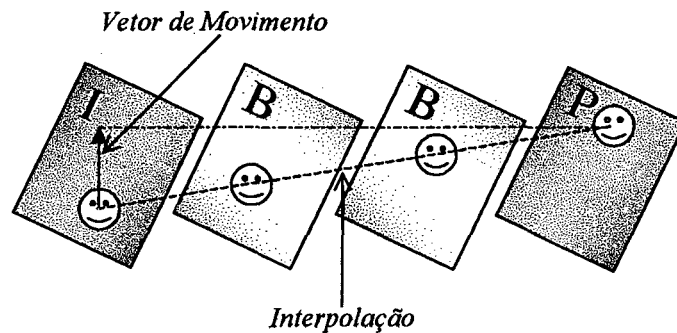


Fig. 3.5 – Processo de Compensação de Movimento

Quando a predição feita pelo processo de avaliação de movimento não retorna um resultado satisfatório, como na presença de cenas com muitos objetos em movimento, as partes correspondentes da imagem são intracodificadas (codificação intraquadro).

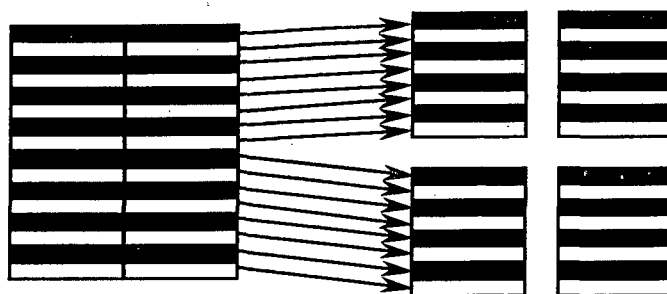
O algoritmo de compressão de vídeo utilizado pelo Padrão MPEG-2 foi projetado de forma a colocar a maioria do trabalho de processamento no codificador, conseqüentemente, o decodificador pode ser implementado em hardware ou software com menos poder de processamento.

#### 3.3.4. PROCESSAMENTO DE IMAGENS ENTRELAÇADAS

A fim de lidar com vídeo digital, o Padrão MPEG-2 *Video* considera o conceito de entrelaçamento de vídeo (vídeo entrelaçado), que consiste na divisão da imagem (um quadro) que deve ser exibida, em duas outras imagens (dois campos), as quais contêm metade das informações da imagem original.

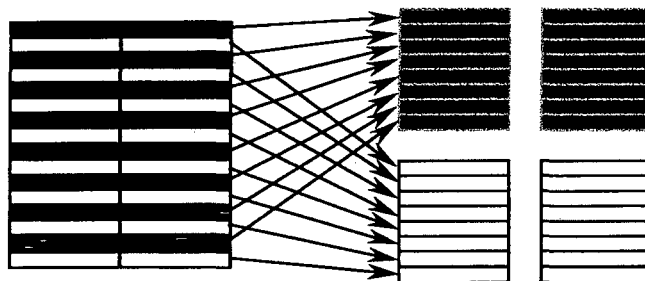
No caso da codificação convencional de uma seqüência de vídeo progressivo, isto é, da estrutura de quadro, o fluxo de entrada consiste da codificação de um campo ímpar, seguido de um campo par e assim por diante, ver figura 3.6. Esta estrutura é mais adequada quando, entre dois campos sucessivos, existir pouco movimento e muitos detalhes.





*Fig. 3.6 – Imagens Progressivas*

No caso de imagens entrelaçadas, ou seja, da estrutura de campos, o fluxo de entrada do codificador consiste de uma série de campos ímpares (parte superior) e pares (parte inferior), separados periodicamente por um campo relacionado ao tempo, figura 3.7. Esta estrutura é melhor aplicada na presença de imagens em movimentos, porém, sem muitos detalhes. Todavia, há situações em que campos e quadros são utilizados na codificação de uma única seqüência de vídeo.



*Fig. 3.7 – Imagens Entrelaçadas*

### 3.3.5. NÍVEIS E PERFIS

As especificações do Padrão MPEG-2 tornaram-se bastante complexas, principalmente, devido a grande variedade de aplicações que o mesmo destina-se a atender. Como para a maioria das aplicações a implementação de toda a sintaxe especificada pelo padrão não é viável, os conceitos de **níveis** e **perfis** foram desenvolvidos e implementados, de modo que subgrupos das funcionalidades do Padrão MPEG-2 *Video* fossem definidos e que somente as características necessárias a cada aplicação fossem especificadas e, por conseguinte, o custo dos equipamentos MPEG-2 fossem reduzidos.

O Padrão MPEG-2 define quatro níveis e cinco perfis, ver tabela abaixo. Cada nível especifica o conjunto de parâmetros que será utilizado pela aplicação, tais como resolução da imagem, que pode estender-se de SIF a HDTV, e taxa máxima de bits por perfil. Cada perfil determina o conjunto de ferramentas de vídeo que será utilizado.

NÍVEIS	DESCRIÇÃO
<i>Low Level (LL)</i>	Corresponde à resolução SIF utilizada no MPEG-1, acima de 360 x 288
<i>Main Level (ML)</i>	Corresponde à resolução padrão 4:2:0, acima de 720 x 576
<i>High 1440 Level (H14)</i>	Foi projetado para HDTV, resolução acima de 1440 x 1152
<i>High Level (HL)</i>	É otimizado para HDTV tela larga, resolução acima de 1920 x 1152
PERFIS	DESCRIÇÃO
<i>Simple Profile (SP)</i>	Não utiliza predição bidirecional (imagens B), logo gera taxas de bits mais altas e simplifica a estrutura do codificador e decodificador
<i>Main Profile (MP)</i>	Utiliza todos os três tipos de imagem (I, P e B). Apesar de acrescentar complexidade ao codificador e decodificador, é o melhor compromisso entre taxa de compressão e custo
<i>Scalables Profiles</i>	Tencionado para uso futuro. Permitirá a transmissão de uma imagem de qualidade básica, em relação à resolução espacial ( <i>Spatially Scalable Profile</i> ) ou precisão de quantização ( <i>SNR Scalable Profile</i> ), e de informações suplementares, permitindo que as características da mesma sejam realçadas
<i>High Profile (HP)</i>	Projetado para aplicações de transmissão de HDTV em formato 4:2:0 ou 4:2:2

*Tab. 3.2 – Níveis e Perfis*

O que determina qual subgrupo de funcionalidades será implementado em determinada aplicação e a habilidade do decodificador em manipular um fluxo de bits específico é a combinação de um nível e com um perfil. Dentre todas as combinações existentes, a mais importante é ML com MP (ML@MP - *Main Level at Main Profile*), a qual define um subgrupo de funcionalidades suficiente para a implementação de aplicações de transmissão de TV, com qualidade PAL<sup>3</sup> ou NTSC<sup>4</sup>.

<sup>3</sup> *Phase Alternate Line* - padrão de TV europeu que transmite sinais analógicos à taxa de 625 linhas de resolução transmitidas em 25 quadros entrelaçados por segundo

<sup>4</sup> *National Television Standards Committee* - padrão de sinal de televisão cuja transmissão ocorre à taxa de 625 linhas de resolução transmitidas em 30 quadros entrelaçados por segundo, incluindo frequência modulada para áudio e sinal para estéreo

Como os níveis e os perfis são organizados hierarquicamente, existe uma compatibilidade crescente entre diferentes níveis e perfis, sendo assim o decodificador de um determinado perfil é capaz de decodificar os perfis mais baixos.

### 3.3.6. ESCALABILIDADE DO PADRÃO MPEG-2 VIDEO

Uma das características mais importantes do Padrão MPEG-2 é que o mesmo utiliza uma abordagem escalável, o que o faz suportar uma variedade de aplicações de vídeo, tais como bibliotecas digitais e HDTV, sem que para isso diferentes padrões e, conseqüentemente, diferentes formatos de fluxos de bits, sejam definidos, a fim de que cada tipo de aplicação seja atendida [Orzessek, 1998].

A codificação escalável (ou em camadas) pode ser aplicada a diferentes aspectos da apresentação de vídeo e possibilita a obtenção simultânea de vídeos com mais de uma resolução ou qualidade. A mesma também proporciona transmissão segura de tráfego multimídia em ambientes de transmissão propensos a erros e é condescendente com as restrições de largura de banda da rede. Através da implementação dos conceitos de fluxos de vídeo com prioridade e recuperação ou correção de erros, a codificação escalável contribui para a melhora da segurança da transmissão [Gringeri, 1998].

A escalabilidade obtida pelo padrão é fruto da estruturação em duas ou mais camadas de todo o fluxo de bits, a começar por uma camada base, seguida por um número de camadas de reforço.

Durante a transmissão, os parâmetros de QoS associados a cada camada podem diferir. Para que a qualidade do vídeo seja controlada pelo usuário ou se adeque a disponibilidade da largura de banda da rede, a camada base deve ser transmitida com prioridade mais alta do que as camadas de reforço.

O Padrão MPEG-2 *Video* especifica várias técnicas de escalabilidade, a saber:

- **Particionamento de Dados:** consiste na fragmentação, no emissor, do fluxo de bits de vídeo em camada base e camada de reforço. A camada base contém os campos do cabeçalho mais importantes, vetores de movimento e coeficientes DCT de baixa frequência, enquanto que a camada de reforço contém informações menos críticas, tal como coeficientes DCT de alta frequência.

O particionamento do fluxo de vídeo é executado com base em um elemento chamado PBP (*Priority Break Point*), o qual especifica o conteúdo de cada camada ou partição. A escolha do PBP pode ser baseada no tipo de quadro de vídeo MPEG-2, nas características do enlace de transmissão, na largura de banda disponível ou em qualquer outro requisito específico da aplicação.

Durante a transmissão, a camada base e a camada de reforço são enviadas através de canais de alto e baixo desempenho, respectivamente. Desta forma, quando transmitidos em Redes ATM, um canal com requisitos de QoS garantidos é reservado à camada base e outro menos confiável à camada de reforço.

No receptor, o fluxo de vídeo original é obtido através da combinação das duas camadas. Sendo que na presença de erros na camada de reforço, somente a camada base é utilizada na recuperação do fluxo de vídeo.

- ***SNR (Signal to Noise Ratio)***: consiste na transmissão de duas camadas do fluxo de vídeo com resolução temporal e espacial similares, porém, com níveis de qualidade de vídeo diferentes. A qualidade da informação de vídeo fornecida pela camada base, a qual é transmitida através de um canal de alta prioridade, pode ser melhorada pela junção desta com a informação fornecida pela camada de reforço.
- ***Escalabilidade Espacial***: possibilita a manipulação de imagens com resoluções espaciais diversas, uma vez que, a partir de um mesmo fluxo de vídeo, duas camadas de vídeo com resoluções espaciais diferentes são geradas. A camada mais baixa pode, por exemplo, ser composta de um formato de vídeo diferente da camada mais alta, além de possibilitar a interoperabilidade entre as camadas. Esta funcionalidade pode ser utilizada em sistemas HDTV/TV, de maneira a permitir a migração de um serviço de TV digital para um serviço de resolução espacial mais alta HDTV.
- ***Escalabilidade Temporal***: possibilita a manipulação de taxas de imagens diferentes em um único fluxo de vídeo, através do particionamento dos quadros do fluxo de entrada em duas ou mais camadas, as quais podem compartilhar ou não a mesma resolução temporal. Quando combinadas, as camadas proporcionam a mesma resolução temporal fornecida pelo fluxo de vídeo de entrada original.

- **Escalabilidade Híbrida:** possibilita, através da combinação de mais de uma técnica de escalabilidade, a produção de três ou mais camadas de vídeo. Todavia, quanto maior o número de camadas, maior é a complexidade atribuída à técnica aplicada.

### 3.4. PADRÃO MPEG-2 *AUDIO*

Da mesma forma que o Padrão MPEG-2 *Video* explora as limitações da visão humana, o Padrão MPEG-2 *Audio* explora as limitações da audição humana, de maneira que a quantidade de informação necessária para a codificação dos sinais de áudio seja reduzida e a conseqüente deterioração da qualidade do som a ser reproduzido não seja percebida.

O Padrão MPEG-2 *Audio* inclui todas as funcionalidades do Padrão MPEG-1 *Audio* e mais uma série de novas funcionalidades e características, por exemplo, o mesmo gera metade da taxa de amostragem produzida pelo MPEG-1, e ainda assim obtém um som de boa qualidade, e possibilita o processamento de sinais de áudio multicanais. Por conseguinte, os equipamentos MPEG-1, através da decodificação da parte que lhe confere, são capazes de decodificar parcialmente os fluxos MPEG-2 e os equipamentos MPEG-2, por sua vez, podem decodificar totalmente os fluxos MPEG-1.

#### 3.4.1. CAMADAS DE COMPRESSÃO

O Padrão MPEG *Audio* especifica três camadas de compressão – camadas 1, 2 e 3 – as quais oferecem diferentes taxas de compressão e níveis de qualidade para um dado sinal de áudio. As mesmas são compatíveis entre si, o que significa que a camada mais alta pode decodificar as camadas mais baixas.

Quanto maior o número de camadas, maior é a taxa de compressão, o poder de processamento requerido para a manipulação do sinal de áudio e melhor será a qualidade do som, porém, menor será a largura de banda necessária para a transmissão do sinal [Orzessek, 1998]. A complexidade do codificador e decodificador também aumenta proporcionalmente de acordo com a quantidade de camadas.

Por exemplo, a camada 3 gera taxa de compressão mais alta (1:10) e som de melhor qualidade, a custo de alto poder de processamento. Enquanto que a camada 1 gera taxa de compressão mais baixa (1:4) e, conseqüentemente, demanda poder de processamento mais baixo e os mais altos requisitos de largura de banda de transmissão, além disso, a mesma proporciona som de qualidade inferior ao som produzido pelas camadas 2 e 3.

### 3.4.2. PRINCÍPIOS DO PROCESSO DE COMPRESSÃO DE ÁUDIO

O processo de compressão de áudio MPEG, aplicável às três camadas de compressão, é iniciado quando o sinal de áudio digital, através de um processo chamado *filter bank*, é convertido em amostras de áudio no domínio da frequência sem perda de informação. Este processo gera um número de sub-bandas de mesmo tamanho, logo, que requerem mesma largura de banda de transmissão. No caso de sinais com mais de um canal de áudio, como por exemplo, em som estéreo, cada canal é tratado separadamente.

Paralelamente ao processo anterior, outro processo, chamado *psychoacoustic model*, elimina todos os sinais cuja audição humana não seria capaz de identificar e define a precisão de quantização necessária a cada sub-banda, ou seja, a quantidade de bits que será utilizada na representação de cada amostra de áudio, a fim de que as faixas de frequência onde a audição humana apresenta-se mais sensível possam ser quantizadas com maior precisão.

### 3.4.3. FORMATO DO QUADRO DE ÁUDIO MPEG

Um quadro de áudio é a unidade de acesso elementar dentro de uma seqüência de amostras de áudio MPEG. Cada quadro transporta todas as informações necessárias à decodificação do sinal de áudio no destino. Ele é dividido em quatro partes principais, ver figura 3.8.

Cabeçalho	CRC	Dados de Áudio			DA
<i>Sistema</i>	<i>Paridade</i>	<i>Bits de Alocação</i>	<i>Informação de Seleção</i>	<i>Fator de Escalamento</i>	<i>Amostras de Áudio</i>
32 bits	16 bits	Tamanho Variável			Tamanho Variável

Fig. 3.8 – Formato Simplificado do Quadro de Áudio MPEG

- **Cabeçalho:** contém informações de sincronização e de sistema. Sua estrutura é a mesma para todas as camadas de compressão;
- **Campo de Paridade (CRC):** campo opcional, utilizado para a detecção de erros no fluxo de áudio;
- **Dados de Áudio:** contém as informações de áudio, ou seja, as amostras de áudio transmitidas e algumas informações adicionais necessárias. Este campo tem estruturas diferentes para as camadas 1, 2 e 3. É composto pelos seguintes campos:
  - **Bits de Alocação:** definem a resolução da codificação das amostras de uma das 32 sub-bandas. No caso das camadas 2 e 3, este campo também especifica se estas sub-bandas são ou não agrupadas em árvores;
  - **Informação de Seleção de Fator de Escalamento:** este campo só aparece na estrutura dos quadros da camada 2 e 3. Indica se o fator de escalamento é válido para toda a duração do quadro ou se existem dois ou mais fatores diferentes;
  - **Fator de Escalamento:** especifica o fator de multiplicação das amostras de uma sub-banda. No caso das camadas 2 e 3, o valor deste fator deve estar de acordo com o que foi especificado no campo acima;
  - **Amostras de áudio:** são as informações de áudio propriamente ditas, como apresentado anteriormente;
- **Dados Auxiliares (DA):** este campo pode armazenar informações adicionais de naturezas diferentes quando necessário, tal como informações referentes ao som multicanal.

#### 3.4.4. MODOS DE CODIFICAÇÃO

As capacidades de transmissão dos canais de áudio MPEG podem ser exploradas de várias maneiras. Para o Padrão MPEG-2 *Audio* quatro modos de codificação diferentes para estes canais são definidos:

- **Mono (Single):** consiste na transmissão de um único canal de áudio;

- **Mono (Dual):** consiste na transmissão de dois canais de áudio, cujos sons não se encontram correlacionados;
- **Stereo:** consiste na transmissão independente de dois canais, um canal da esquerda e outro da direita;
- **Joint Stereo:** explora a redundância existente entre os canais da esquerda e direita, a fim de que a taxa de bits de áudio seja reduzida e que, por consequência, uma melhor taxa de compressão seja obtida através da combinação deste dois canais. Existem duas maneiras de se obter a codificação para este modo: **Intensity Stereo** e **MS Stereo**. As camadas 1 e 2 utilizam somente o **Intensity Stereo**, enquanto que a camada 3 utiliza tanto o **Intensity Stereo**, quanto o **MS Stereo**, além da combinação de ambas.

### 3.5. PADRÃO MPEG-2 SYSTEMS

Os fluxos de bits gerados pelos codificadores de vídeo e áudio, bem como os fluxos de bits definidos pelos usuários, nomeados **Fluxos de Bits de Dados Privados**, são denominados **Fluxos Elementares (ES – Elementary Stream)**.

Em se tratando especificamente de fluxos de vídeo e áudio, cada fluxo elementar pode ser organizado em unidades chamadas **unidades de acesso**, que dependendo da natureza do fluxo elementar, podem ser constituídas de uma imagem ou um quadro de áudio decodificado.

Os fluxos elementares devem ser combinados de maneira organizada e complementados com informações adicionais que permitam sua segmentação pelo decodificador, bem como a sincronização da imagem e do som durante a apresentação e a seleção, por parte do usuário, dos componentes específicos de seu interesse.

Estas funções são de responsabilidade do Padrão MPEG-2 *Systems*. O mesmo trata da multiplexação dos fluxos de vídeo, áudio e dados privados em um único fluxo de bits, adequado ao armazenamento e a transmissão, e da definição das informações de controle e gerenciamento destes diferentes fluxos, fornecendo a sintaxe necessária à sincronização dos mesmos durante a decodificação e apresentação das informações de vídeo e áudio no receptor. A figura 3.9 ilustra a estrutura do Padrão MPEG-2 *Systems*.



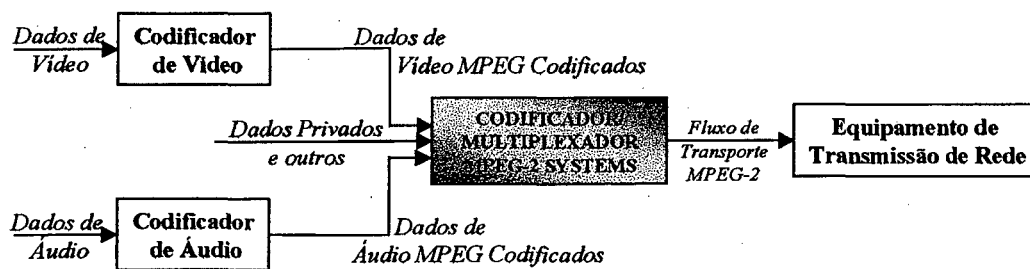


Fig. 3.9 – Estrutura Simplificada do Padrão MPEG-2 Systems

Basicamente, o que o Padrão MPEG-2 *Systems* faz é definir como a **Camada de Sistema** executará suas funções. Esta camada é responsável por um conjunto de funções básicas que envolvem outra camada, denominada **Camada de Compressão**, a qual é composta pelos fluxos elementares. Dentre estas funções pode-se citar:

- Conversão do fluxo contínuo de informação em pacotes e multiplexação dos vários fluxos de pacotes gerados em um único fluxo;
- Adição de *timestamps*<sup>5</sup> nos fluxos elementares, a fim de que os mesmos possam ser sincronizados durante a apresentação, e
- Inicialização e gerenciamento das áreas de armazenamento necessárias à decodificação dos fluxos elementares no destino.

### 3.5.1. PROCESSO DE CONVERSÃO DOS FLUXOS DE INFORMAÇÃO EM FLUXOS DE PACOTES

Para que as funções citadas acima sejam executadas, cada fluxo elementar de informação contínua deve ser convertido em um fluxo de pacotes, chamado **Fluxo Elementar de Pacotes (PES - Packetized Elementary Stream)**.

Cada pacote PES, ou seja, cada pacote que compõe o fluxo elementar de pacotes, é constituído de cabeçalho e de carga útil, ambos de tamanhos e formatos variados. A estrutura do pacote de PES encontra-se melhor apresentada na seção 3.5.3.

Dependendo do tipo de aplicação, os PESs podem ser multiplexados de duas maneiras diferentes - **Fluxo de Programa (PS - Program Stream)** e **Fluxo de Transporte (TS -**

<sup>5</sup> Refere-se às informações de tempo utilizadas para a distribuição e sincronização dos relógios

**Transport Stream**) - a fim de formar dois tipos diferentes de fluxos de bits e, conseqüentemente, atender aos diferentes requisitos impostos por tais aplicações.

O Fluxo de Programa MPEG-2 utiliza pacotes longos de tamanhos variados, os quais, quando multiplexados, devem compartilhar necessariamente a mesma base de tempo. Estes fluxos são mais adequados às aplicações de armazenamento e transmissão baseadas em ambientes supostamente não susceptíveis a erros, pois, devido ao tamanho longo dos pacotes, a presença de pacotes corrompidos pode resultar em desgaste na qualidade da informação de vídeo e áudio.

Contrariamente, os Fluxos de Transporte MPEG-2 utilizam pacotes de transporte de tamanho pequeno e fixo (188 bytes), adequando-os a ambientes propensos a erros, uma vez que torna possível a implementação eficiente de técnicas que possam minimizar os efeitos causados por pacotes corrompidos ou perdidos. Os fluxos de transporte podem ser de dois tipos:

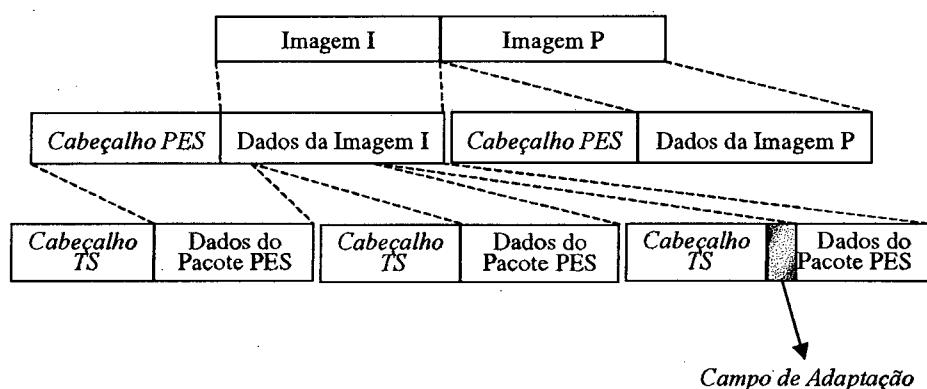
- **Simple Program Transport Stream (SPTS)**: consiste da multiplexação de vários fluxos PES (vídeo, áudio ou dados privados) pertencentes ao mesmo programa, os quais devem necessariamente utilizar a mesma base de tempo, a fim de simplificar a sincronização durante a decodificação. Este tipo de fluxo foi originalmente projetado para aplicações baseadas na transmissão de programas de TV via satélite, cabo ou terrestre;
- **Multi Program Transport Stream (MPTS)**: consiste da multiplexação de vários programas (SPTS), os quais não necessariamente precisam compartilhar a mesma base de tempo.

Uma vez que o trabalho trata de transmissões multimídia em Redes ATM e que, como será visto nos capítulos futuros, o ATM Forum padronizou o mapeamento de pacotes de transporte MPEG-2 em células AAL5, neste capítulo somente os aspectos relacionados aos pacotes de transporte serão abordados.

### 3.5.2. MAPEAMENTO DOS PACOTES PES EM PACOTES DE TRANSPORTE

Antes de transmitidos, os pacotes PES devem ser mapeados em **Pacotes de Transporte (PT)**, os quais também se constituem de cabeçalho (4 bytes) e carga útil (até 184 bytes), precedidos por um campo de adaptação opcional, só utilizado quando o tamanho da informação contida na carga útil for menor do que 184 bytes.

O Padrão MPEG-2 Systems determina que um pacote de transporte deve conter somente os dados derivados de um único pacote PES e que este deve sempre começar e terminar no início e no final, respectivamente, da carga útil do pacote de transporte, ver figura 3.10.



*Fig. 3.10 - Mapeamento dos Pacotes PES em Pacotes de Transporte*

Como os pacotes PES são mais longos que os pacotes de transporte, os mesmos devem ser fragmentados em blocos de dados de 184 bytes. Quando o tamanho do pacote PES não for múltiplo de 184, o que frequentemente acontecerá, o último pacote de transporte deverá ser iniciado com o campo de adaptação, cujo tamanho pode ser calculado através subtração do tamanho total da carga útil (184 bytes) pelo tamanho do último pacote PES.

É importante ressaltar que o cabeçalho dos pacotes PES e o cabeçalho dos pacotes de transporte transmitem informações diferentes. O primeiro contém informações ligadas diretamente ao fluxo elementar, tal como tipo de fluxo, enquanto que o segundo, contém informações utilizadas durante a multiplexação, a transmissão e a entrega do fluxo de informação [Orzessek, 1998].

### 3.5.3. ESTRUTURA DE UM PACOTE PES

Como visto na seção 3.5.1, um pacote PES é composto do cabeçalho, que contém informações sobre qual fluxo elementar o pacote faz parte e qual o tipo de informação (vídeo, áudio ou dados privados) este fluxo transporta, e da carga útil, ou seja, dos dados propriamente ditos.

A figura 3.11 ilustra de maneira simplificada a estrutura de um pacote PES. Dentre todos os campos contidos no cabeçalho de um pacote PES, os mais importantes são *stream\_id*, *indicators* e *flags*. O *stream\_id* especifica o formato do pacote PES, ou seja, qual o tipo de informação que o mesmo carrega. Os *indicators* apresentam informações adicionais sobre o conteúdo do pacote, a tabela 3.3 descreve os campos contidos neste campo.

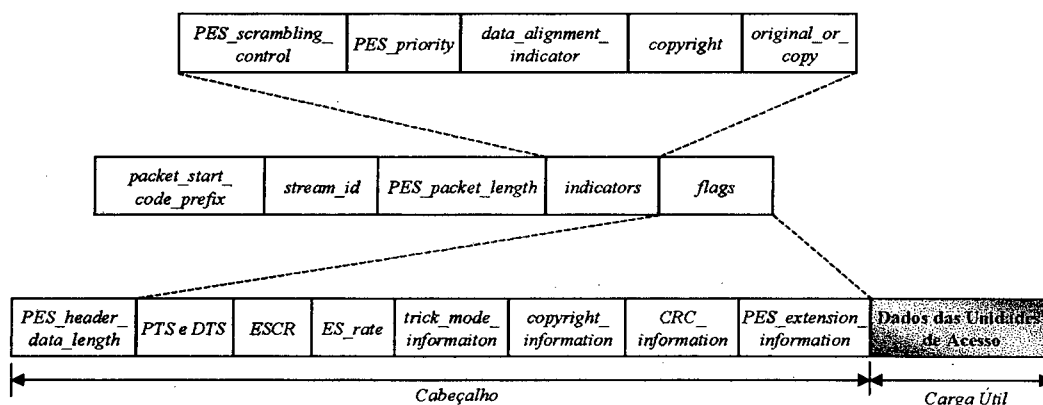


Fig. 3.11 – Estrutura Simplificada de um Pacote PES

CAMPOS DO CAMPO INDICATORS	DESCRIÇÃO
<i>PES_scrambling_control</i>	Especifica se a carga útil encontra-se embaralhada. Caso positivo, o tipo de embaralhamento definido pelo usuário é utilizado
<i>PES_priority</i>	É utilizado para associar maior prioridade ao pacote PES
<i>data_alignment_indicator</i>	Indica que a carga útil começa com um código de início de vídeo ou áudio
<i>copyright</i>	Indica que a carga útil é protegida com direitos autorais
<i>original_or_copy</i>	Indica se a carga útil é original ou se é uma cópia

Tab. 3.3 – Principais Campos do Campo Indicators

Os *flags*, também são usados na definição do formato do pacote PES, os mesmos são utilizados para indicar a presença dos seguintes campos:

- *presentation\_time\_stamp (PTS) e decode\_time\_stamp (DTS)*: no caso de fluxos de vídeo ou áudio, fornecem informações de *timestamp*, indicando quando os dados devem ser decodificados e apresentados. O campo DTS sempre aparece associado a um campo PTS, porém, este é um campo opcional só utilizado quando o tempo de decodificação da unidade de acesso for diferente do tempo de apresentação atual;
- *elementary\_stream\_clock\_reference (ESCR) e elementary\_stream\_rate (ES\_rate information)*: fornecem informações de tempo adicionais ao decodificador;
- *trick\_mode\_information*: informa quando a carga útil está representando um fluxo especial. Por exemplo, quando a função avanço rápido for executada em uma aplicação de vídeo sob demanda interativa, este campo pode ser associado a um valor que represente “avanço rápido”;
- *copyright\_information*: os dados do pacote PES estarão protegidos com direitos autorais se este campo for igual a 1. Esta informação pode ser utilizada na implementação de funções que evitem que o pacote seja copiado;
- *CRC\_information*: este campo é opcional. Ele permite que o pacote contenha um valor CRC<sup>6</sup>, calculado a partir da carga útil dos pacotes PES anteriores;
- *PES\_extension\_information*: contém vários outros campos que possibilitam a manipulação de espaços de armazenamento e de fluxos de sistema MPEG-1.

#### 3.5.4. ESTRUTURA DE UM PACOTE DE TRANSPORTE

Como dito anteriormente, um pacote de transporte é constituído de 188 bytes, dos quais 4 são reservados ao cabeçalho e 184 à carga útil. A estrutura do pacote de transporte e os campos que compõem o cabeçalho estão ilustrados na figura 3.12.

---

<sup>6</sup> *Cyclical Redundancy Check* - método de detecção de erros em transmissão de dados.

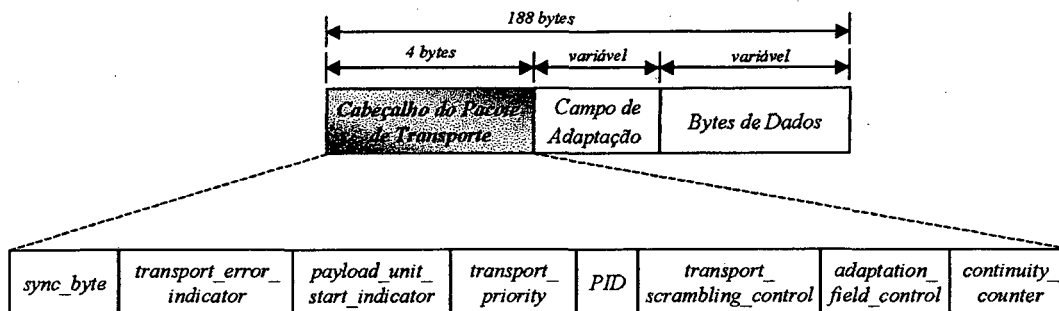


Fig. 3.12 – Estrutura de um Pacote de Transporte MPEG-2

A tabela 3.4 lista os campos que constituem o cabeçalho de um pacote de transporte, assim com suas funcionalidades. Estes campos são chamados campos de controle, uma vez que fornecem informações importantes sobre a carga útil do pacote e indicam o aspecto de outros campos no cabeçalho.

CAMPOS DO CABEÇALHO DE UM PACOTE DE TRANSPORTE	DESCRIÇÃO
<i>sync_byte</i>	É utilizado na identificação do início do pacote de transporte
<i>transport_error_indicator</i>	Indica um erro de bit no pacote de transporte
<i>payload_unit_start_indicator</i>	Indica que o primeiro byte da carga útil do pacote de transporte é o início de uma unidade de carga útil, por exemplo, um pacote PES
<i>transport_priority</i>	É utilizado para indicar prioridade relativa dos pacotes de transporte
<i>PID (Packet Identifier)</i>	Define qual o tipo dos dados transportados na carga útil do pacote de transporte. Também é utilizado na identificação de pacotes de transporte que contém pacotes PES procedentes do mesmo fluxo elementar.
<i>transport_scrambling_control</i>	Indica o tipo de embaralhamento utilizado na carga útil do pacote de transporte
<i>adaptation_field_control</i>	Indica se o pacote de transporte é seguido por um campo de adaptação e/ou carga útil
<i>continuity_counter</i>	É um contador incrementado por cada pacote de transporte que compartilha o mesmo PID e que contém carga útil. O mesmo recebe valor igual a zero, antes que seu valor máximo seja alcançado. Este campo pode ser utilizado, por exemplo, para verificação de perda de pacotes.

Tab. 3.4 – Campos do Cabeçalho de um Pacote de Transporte

## CAMPO DE ADAPTAÇÃO

Apesar do campo de adaptação transportar informações importantes ao processamento dos fluxos de transporte MPEG-2, tais como informações para a recuperação do relógio<sup>7</sup> e para a execução de funções de concatenação, o mesmo não precisa necessariamente estar associado a todos os pacotes de transporte, tornando a sua utilização opcional.

Como dito na seção 3.5.2, este campo é normalmente utilizado quando o tamanho do pacote PES não se ajustar perfeitamente ao tamanho da carga útil do pacote de transporte. Os campos que compõem o campo de adaptação encontram-se descritos na tabela 3.5.

CAMPOS DO CAMPO DE ADAPTAÇÃO	DESCRIÇÃO
<i>adaptation_field_lenght</i>	Indica o tamanho do campo de adaptação
<i>flags and indicators</i>	Os <i>flags</i> determinam qual a estrutura dos demais campos, enquanto que os <i>indicators</i> fornecem informações sobre a carga útil
<i>PCR fields</i>	Contém o <i>Program Clock Reference</i> , o qual transporta as informações de <i>timestamps</i> utilizadas pelo decodificador durante a sincronização de seu relógio com o relógio do codificador
<i>OPCR fields</i>	Contém o <i>Program Clock Reference</i> original
<i>splice_countdown</i>	É utilizado para que as funções de concatenação sejam suportadas
<i>private_data fields</i>	É composto por um campo de tamanho e pelos bytes de dados privados, os quais têm formato definido pelo usuário
<i>adaptation field extention</i>	Fornece informações adicionais ao suporte de fluxos de transporte concatenados e multiplexados

Tab. 3.5 – Campos do Campo de Adaptação

## CONCATENAÇÃO DE FLUXOS DE TRANSPORTE

Como dito anteriormente, o fluxo de transporte suporta a multiplexação e demultiplexação de diferentes programas, gerados por diferentes fluxos de bits de vídeo e áudio. O campo de adaptação, através do campo *adaptation field extention*, define os elementos de sintaxe necessária para que dois fluxos de pacotes PES diferentes sejam concatenados em um único fluxo de transporte [Orzessek, 1998].

<sup>7</sup> Circuito eletrônico que, através da produção de pulsos uniformemente espaçados, determina a velocidade de processamento e sincroniza as atividades do computador.

Para a execução das funções de concatenação, é necessário que o multiplexador de fluxo de transporte tenha conhecimento sobre onde começa e onde termina uma unidade de acesso de vídeo ou áudio, a fim de que partes destas diferentes unidades não sejam embaralhadas e, conseqüentemente, a apresentação do vídeo ou áudio não seja afetada.

Como este processo é bastante complexo, visto que o multiplexador teria que decodificar não somente os pacotes PES, mas também as unidades de acesso de vídeo ou áudio, pontos de concatenação são utilizados com o propósito de indicar onde, no fluxo de transporte em questão, um novo programa pode ser introduzido.

### ***PROGRAM CLOCK REFERENCE (PCR)***

O Padrão MPEG-2 especifica, através de alguns dos campos contidos no cabeçalho do pacote PES e do pacote de transporte MPEG-2, a distribuição e a sincronização de relógios entre codificador e decodificador.

Durante o processo de decodificação, o decodificador MPEG-2 seleciona todos os pacotes de transporte pertencentes ao mesmo fluxo elementar, ou seja, com mesmo PID, e reconstrói as unidades de acesso correspondentes a este fluxo. Posteriormente, com base nos *timestamps* contidos nos campos DTS e PTS, o decodificador executa a decodificação e a apresentação destas unidades.

Para que estas operações sejam executadas no momento correto, o relógio do decodificador deve estar sincronizado com o relógio do codificador, isto é, com o relógio que foi utilizado quando os *timestamps* de apresentação e decodificação foram criados.

Para tanto, o pacote de transporte MPEG-2 contém um campo chamado PCR (*Program Clock Reference*), o qual é gerado durante a codificação do fluxo de transporte e, juntamente com o fluxo, é enviado periodicamente ao decodificador, a fim de preservar a sincronização.

O relógio do decodificador é inicializado pelo primeiro PCR transmitido e deve se manter exatamente igual ao valor de *timestamp* marcado pelos PCRs posteriormente recebidos. Caso contrário, o mesmo deve ser ajustado com base na diferença entre os valores marcados pelo seu relógio e pelo PCR atual.



Quem define em quais pacotes de transporte o campo PCR para o programa será encontrado, é a tabela de mapeamento de programa (apresentada nas próximas seções), isto é feito através da especificação dos valores de PID destes pacotes.

### 3.5.5. DETECÇÃO DE ERROS

A detecção de erros nos fluxos de transporte executada pelo Padrão MPEG-2 *Systems*, baseia-se principalmente em determinados campos do cabeçalho do pacote de transporte, são eles:

- ***transport\_error\_indicator***: o equipamento MPEG-2, tal como um multiplexador de fluxo de transporte MPEG-2, utiliza este campo para indicar a presença de erros no cabeçalho ou carga útil do pacote de transporte. De acordo com esta informação, o decodificador pode, por exemplo, executar mecanismos adicionais de cancelamento de erros durante a decodificação e assim, contribuir para o aumento do nível de qualidade dos serviços de rede.
- ***transport\_priority***: indica que um pacote de transporte tem alta prioridade. Desta forma, na presença de congestionamento, onde determinados pacotes precisam ser descartados, o mesmo deve receber tratamento de prioridade mais alta. Como alguns protocolos de rede, por exemplo, ATM, também utilizam bits que determinam prioridade, a correlação entre prioridade de pacotes de transporte MPEG-2 e prioridade de rede pode ser executada.
- ***continuity\_counter***: é incrementado por cada pacote de transporte com mesmo PID, para a detecção de perda de pacotes.

Não só os pacotes de transporte possuem campos para a verificação de erros e perda de pacotes. O MPEG-2 também utiliza o método CRC na proteção da informações transmitidas. Por exemplo, os CRCs são utilizados por todas as tabelas PSI e também estão presentes no cabeçalho do pacote PES, neste caso seus valores são calculados a partir dos bytes de dados dos pacotes PES anteriores. Os pacotes PES ainda utilizam o campo *elementary\_stream\_priority*, o qual funciona de maneira similar ao *transport\_priority*.

### 3.5.6. INFORMAÇÕES ESPECÍFICAS DE PROGRAMA

Como visto anteriormente, o fluxo de transporte suporta a transmissão de mais de um programa simultaneamente, cada qual composto por um ou mais fluxos de pacotes elementares (pacotes PES).

O Padrão MPEG-2 *Systems* define um programa como um conjunto de fluxos elementares que necessariamente compartilham a mesma base de tempo. Para que estes diferentes fluxos elementares sejam agrupados em programas e corretamente identificados pelo receptor, é necessário que o fluxo de transporte transmita, além das informações de vídeo e áudio, informações de controle e gerenciamento, tal como informações periódicas sobre a base de tempo.

Estas informações são agrupadas em um grupo de tabelas interligadas, chamado **Informações Específicas de Programa (IEP)**. Dependendo de sua importância, cada tabela pode ser composta por uma ou mais seções, ver figura 3.13. Existem quatro tabelas IEP especificadas pelo Padrão MPEG-2 *Systems*:

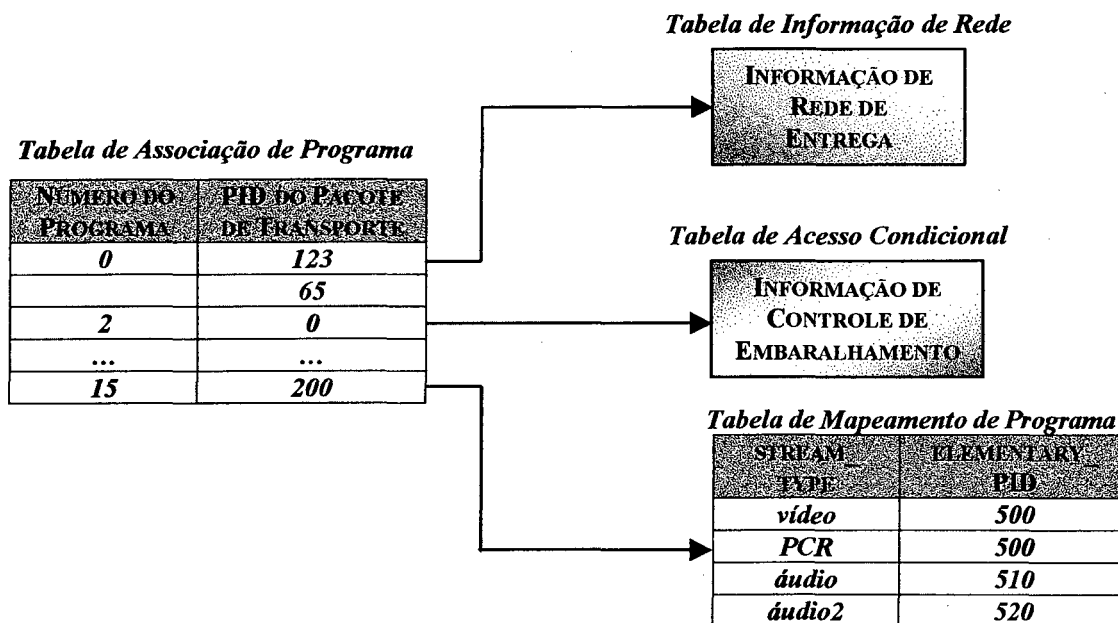


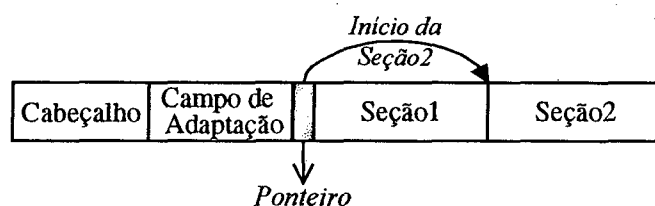
Fig. 3.13 – Relação entre as Tabelas de Informações Específicas de Programa

- **Tabela de Alocação de Programa (TAP):** a presença desta tabela é obrigatória para os pacotes de transporte cujo campo PID é igual a zero. Sua função é fornecer ao decodificador as primeiras informações sobre quais programas estão sendo transmitidos pelo fluxo de transporte. Estes programas possuem duas

entradas armazenadas na tabela, uma com o número do programa e outra com seu respectivo PID, onde o valor do PID indica quais os pacotes de transporte que carregam a *tabela de mapeamento de programa*. Por exemplo, na figura acima, a tabela TAP indica que para o programa de número 15, a *tabela de mapeamento de programa* pode ser encontrada nos pacotes de transporte com PID igual a 200. Somente o programa de número zero possui um significado especial, uma vez que os PIDs associados a ele identificam os pacotes cujo conteúdo é a *tabela de informações de rede*. Mesmo que todos os programas estejam embaralhados, esta tabela é sempre claramente transmitida.

- ***Tabela de Mapeamento de Programa (TMP)***: cada programa transmitido pelo fluxo de transporte possui uma tabela TMP, cujo conteúdo consiste dos campos *elementary\_PID* e *stream-type*, e de um campo opcional chamado *stream descriptors*, o qual fornece informações adicionais sobre os fluxos pertencentes ao programa, como combinação de nível e perfis e taxa de quadro. O campo *elementary\_stream* contém todos os PIDs dos pacotes de transporte que carregam os pacotes PES de um programa específico, enquanto que o campo *stream\_type* define o tipo de fluxo PES presente nos pacotes de transporte identificados pelo campo anterior. Na figura 3.13, os pacotes PES do programa de número 15, cujo conteúdo é vídeo, podem ser encontrados nos pacotes de transporte com PID 500, assim como os pacotes de áudio, podem ser encontrados nos pacotes de transporte de PID 510 e 520.
- ***Tabela de Informação de Rede (TIR)***: o conteúdo desta tabela não é definido pelo Padrão MPEG-2, o mesmo deixa esta função sob a responsabilidade do provedor de rede, contudo, seu conteúdo refere-se às informações sobre a rede de entrega utilizada na transmissão dos fluxos de transporte.
- ***Tabela de Acesso Condicional (TAC)***: é transmitida por pacotes de transporte com PID igual a zero. A mesma contém informações sobre os métodos de criptografia utilizados para os dados de vídeo e áudio, assim como entradas para os PIDs que identificam os pacotes que transportam informações de controle para sistemas de embaralhamento.

Como abordado previamente, as informações das tabelas que compõem o IEP podem ser segmentadas em seções, que por sua vez devem ser mapeadas em pacotes de transporte MPEG-2. Diferentemente dos pacote PES, as seções podem ser armazenadas em qualquer posição da carga útil do pacote de transporte, ou seja, o pacote pode começar com o final de uma outra seção, independente da mesma ser ou não precedida por um campo de adaptação. Porém, a fim de que o início de cada seção seja identificada, o primeiro byte da carga útil tem a finalidade de apontar para o início de uma nova seção, ver figura 3.14.



**Fig. 3.14 - Mapeamento de uma Seção em um Pacote de Transporte MPEG-2**

Os campos que compõem o cabeçalho de uma seção fornecem informações sobre o tipo, o tamanho, o número da seção e o número total de seções que fazem parte de uma mesma tabela IEP. Somente para as seções de dados privados, os campos *número da seção* e *número total de seções* são opcionais.

### 3.6. CONSIDERAÇÕES FINAIS

Este capítulo destinou-se a apresentação do Padrão de Compressão MPEG-2, seus componentes e suas características, sendo que somente as três principais partes que compõem o padrão – *Systems*, *Audio* e *Video* – foram descritas com mais detalhamento, visto serem estas as partes relevantes ao trabalho. No próximo capítulo, os principais conceitos sobre Qualidade de Serviço (QoS) serão abordados, uma vez que o trabalho baseia-se na transmissão de fluxos MPEG-2 com QoS.

## CAPÍTULO IV

### QUALIDADE DE SERVIÇO

*O conceito de Qualidade de Serviço (QoS) tornou-se nos últimos anos muito discutido no âmbito de Redes de Computadores e Sistemas Multimídia. Várias são as definições dadas à QoS, cada uma das quais, muitas das vezes, tenta se adequar a ambientes e aplicações particulares. Neste capítulo serão apresentados os principais conceitos relacionados à QoS, bem como o Modelo de QoS adotado por este trabalho.*

#### 4.1. INTRODUÇÃO

O conceito de Qualidade de Serviço (QoS) foi parcialmente formalizado nos anos 80 pela ISO (*International Organization for Standardization*). Inicialmente, as pesquisas em torno de QoS eram feitas dentro do contexto de camadas de arquiteturas individuais, tais como plataforma de sistemas distribuídos, sistema operacional e subsistema de rede e de transporte.

Contudo, diante do aparecimento dos sistemas multimídia distribuídos, as arquiteturas de QoS definidas até então precisaram estender-se de maneira a abranger também as questões de suporte fim-a-fim para comunicações multimídia, uma vez que não satisfaziam as exigências destes sistemas e não definiam nenhum mecanismo que garantisse que os requisitos de QoS especificados seriam realmente satisfeitos.

Na verdade, QoS é um conceito complexo cujo significado depende do ponto de vista sobre o qual está sendo observado. Para [Fluckiger, 1995], o conceito de QoS baseia-se na declaração de que nem todas as aplicações necessitam do mesmo desempenho sobre a rede nas quais estão sendo executadas. Sendo assim, o ideal seria que as aplicações indicassem à rede seus próprios requisitos, antes de efetivamente iniciarem a transmissão dos dados.

Principalmente, após o surgimento dos serviços RDSI-FL (Redes Digitais de Serviços Integrados de Faixa Larga) e suas tecnologias de suporte, tal como ATM (*Asynchronous Transfer Mode*), o interesse pela implementação deste conceito se firmou ainda mais, levando a novas definições, com o propósito de abranger os requisitos de comunicação multimídia. Na seção a seguir serão apresentados alguns dos conceitos encontrados na literatura referente ao assunto.

## 4.2. DEFINIÇÃO

Vários foram os autores que definiram Qualidade de Serviço. Alguns o fizeram sob a perspectiva de sistemas, outros de aplicações e de usuários. Existem ainda aquelas definições que abrangem mais de uma perspectiva.

[Fluckiger, 1995] define QoS sob a óptica do sistema de comunicação, uma vez que limita os requisitos de QoS a este sistema. Segundo ele, *QoS é formada pelos requisitos que as aplicações impõem sobre a rede nas quais estão sendo executadas, onde a especificação de tais requisitos deve acontecer antes da transmissão dos dados.*

Como ponto de partida para uma definição mais abrangente pode-se considerar a definição do ITU-T (*International Telecommunications Union*): *“QoS é o efeito coletivo do desempenho de serviço, o qual determina o grau de satisfação de um usuário deste serviço”*. E a definição dada pelo Projeto RACE, da Comissão Européia: *“QoS é a medida de quanto um serviço apresentado ao usuário é bom, medido através de uma linguagem compreensível e que se manifesta em um número de parâmetros, que têm tanto valores subjetivos quanto objetivos”*.

Igualmente, como nas duas últimas definições, [Vogel, 1995] define QoS principalmente sob o ponto de vista da aplicação e do usuário: *“Qualidade de Serviço representa o conjunto das características quantitativas e qualitativas de um sistema multimídia distribuído, necessário para se conseguir a funcionalidade desejada de uma aplicação, a qual inclui a apresentação dos dados multimídia para o usuário e sua referente satisfação”*.

Baseando-se nas definições citadas, há dois aspectos de QoS a serem considerados: *requisitos de QoS especificados pelos usuários/aplicações* e *garantias de QoS fornecidas pelos sistemas*. Os requisitos de QoS são especificados através de um grupo de parâmetros, o qual encontra-se exposto à negociação entre os diversos componentes do sistema multimídia envolvidos.

Como exposto anteriormente, não há ainda uma definição única para QoS. Todavia, de maneira geral, pode-se dizer que QoS trata das características que, de uma forma ou de outra, afetam a qualidade de uma aplicação. Neste trabalho, a definição abordada por [Vogel, 1995] será considerada, uma vez que a mesma aplica-se a sistemas multimídia

distribuídos e logo, concentra-se não somente nos requisitos do sistema, mas, principalmente, nos requisitos da aplicação e do usuário.

### 4.3. PARÂMETROS DE QoS

Os parâmetros de QoS descrevem atributos específicos de um recurso do sistema, tais como taxa de bits, limite e variação de atraso. Na prática, os mesmos não recebem um valor fixo, um ou mais valores lhes são associados. O valor de um parâmetro pode variar de aplicação para aplicação e influenciar direta ou indiretamente os valores de outros parâmetros. Por exemplo, a diminuição da taxa de erro através da retransmissão implica em aumento no atraso de transmissão.

Além disso, diferentes parâmetros são utilizados em diferentes camadas ou subsistemas de um sistema multimídia distribuído [Lu, 1996], pois estes tratam objetos de dados diferentes. Por exemplo, a escolha da resolução e da taxa de quadro de um vídeo implica em requisitos de largura de banda.

As aplicações têm os parâmetros de QoS demarcados por dois pontos - *QoS Desejada* e *QoS Aceitável* - os quais definem o espectro da QoS [Raghavan, 1998], como mostra a figura 4.1.

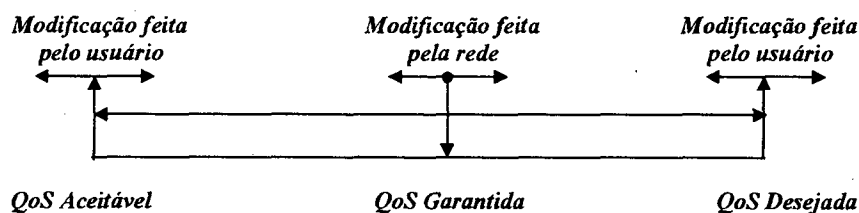


Fig. 4.1 – Espectro de QoS

A QoS desejada corresponde às condições ideais sob as quais a aplicação deve ser executada. A QoS aceitável corresponde à situação mínima aceitável, o que significa que abaixo deste ponto a execução da aplicação torna-se inviável.

O conjunto de parâmetros de QoS constitui-se de parâmetros que expressam o desempenho do comportamento do sistema e de outros que expressam outras características do mesmo. Desta forma, os parâmetros de QoS podem ser classificados da seguinte maneira [Hafid, 1996]:

- **Parâmetros Objetivos:** podem ser diretamente observados e medidos no ponto em que o serviço está sendo acessado pelo usuário. Por exemplo, atraso e vazão.
- **Parâmetros Subjetivos:** não podem ser diretamente medidos, visto que dependem do equipamento e da percepção do usuário. Por exemplo, qualidade do alto-falante.

Além de descrever os atributos de determinados recursos do sistema, a especificação dos parâmetros de QoS também determina o tipo de serviço que está sendo implementado, uma vez que, na prática, o usuário pode especificar o nível de garantia que deseja para tal serviço. Existem pelo menos três tipos de serviços ou níveis de garantias diferentes [Steinmetz, 1995; Lu, 1996]:

- **Garantia Determinística ou Serviço Garantido:** fornece garantia de que a QoS especificada pelo usuário será totalmente satisfeita. Baseia-se tanto na representação estatística, que trata dos limites estatísticos dos parâmetros de QoS, tal como limite estatístico de taxa de erro, como na representação determinística, onde os parâmetros de QoS podem ser obtidos através de valores únicos, de um par de valores ou de um intervalo de valores;
- **Garantia Estatística ou Serviço Previsível:** os parâmetros de QoS são estimados com base no comportamento anterior da rede, ao qual o serviço tenta corresponder;
- **Serviço de Melhor Esforço:** não fornece garantia de que a QoS requerida será satisfeita ou fornece garantia parcial. No primeiro caso, os parâmetros de QoS necessários não são especificados. No segundo, alguns limites para estes parâmetros, nas formas estatística ou determinística são considerados.

De maneira simplificada, o processamento dos requisitos de QoS especificados pelo usuário/aplicação ao sistema, pode ser descrito da seguinte forma:

- Quando uma aplicação precisa iniciar uma sessão, ela submete ao sistema uma solicitação contendo os requisitos de QoS a serem satisfeitos. Tais requisitos



foram especificados com base nas exigências do usuário quanto à qualidade do serviço que foi solicitado, como por exemplo, em relação ao desempenho, custo e sincronização;

- Os requisitos de QoS são então mapeados em parâmetros de QoS para as várias camadas e subsistemas. Estes, por sua vez, são negociados entre os componentes do sistema multimídia envolvidos, a fim de que os mesmos avaliem se podem ou não oferecer garantias de que os requisitos solicitados serão consistentemente satisfeitos;
- Terminada a negociação, um contrato entre o sistema e a aplicação é elaborado, os recursos são reservados para que a QoS possa ser proporcionada e a sessão pode finalmente ser iniciada. Porém, se por ventura a QoS requerida não puder ser oferecida, o sistema pode negar a inicialização de uma nova sessão para a aplicação.

#### 4.4. MODELO DE QoS

Segundo [Aurrecoechea, 1998], os modelos arquiteturais de suporte à QoS propostos até então precisam ser melhor estudados, em virtude de ainda oferecerem certas deficiências, principalmente, quanto a:

- **Interfaces:** as interfaces atuais geralmente não são configuráveis quanto a QoS, além de oferecerem somente um pequeno grupo de facilidades necessárias para o controle e gerenciamento de fluxos multimídia;
- **Mecanismos de Suporte a Garantias de QoS:** ainda é necessário o desenvolvimento de mais pesquisas no que se refere ao controle distribuído, mecanismos de monitoramento e manutenção de QoS, a fim de que os níveis de serviços contratados possam ser previsíveis e garantidos;
- **Modelo Geral:** há a necessidade de que um modelo geral de arquitetura de QoS seja desenvolvido, em favor de que toda a base conceitual de QoS, em diferentes níveis de sistemas e entre arquiteturas de redes diferentes, seja implementada.

Diante destas carências, muitas pesquisas foram desencadeadas com o intuito de definir um conjunto de interfaces de QoS configuráveis que pudessem formalizar a QoS em redes e sistemas fins, fornecendo um modelo para a integração de mecanismos de gerenciamento e controle de QoS. No decorrer destas pesquisas, algumas arquiteturas foram propostas com a finalidade de proporcionar abordagens que fornecessem garantias de QoS fim-a-fim. A estas arquiteturas geralmente, dá-se o nome de *Modelo de QoS*.

O Modelo Genérico de QoS apresentado a seguir baseia-se nos modelos abordados por [Aurrecochea, 1998], [Hafid, 1996] e [Lu, 1996], e tem como objetivo tratar a QoS em sistemas multimídia distribuídos.

#### 4.4.1. MODELO GENÉRICO DE QoS

Apesar de parecer simples, o processamento dos requisitos de QoS está sujeito a várias complicações, tais como mudanças nos requisitos originalmente especificados no decorrer da sessão e impossibilidade, por parte do sistema, de conservação dos parâmetros negociados em virtude de problemas, como congestionamento da rede.

Nestes casos, os requisitos de QoS inicialmente negociados podem ser renegociados, a fim de que os mesmos se adequem às novas exigências do usuário ou a nova realidade do sistema. Por conseguinte, é evidente a necessidade da implementação de mecanismos que possibilitem a alocação dinâmica de recursos, o monitoramento constante da QoS fornecida, bem como a adequação desta QoS quando necessário.

Este modelo consiste de um conjunto de mecanismos utilizados para a provisão de QoS, a saber:

- **Especificação da QoS:** este mecanismo permite que os usuários/aplicações especifiquem os requisitos de QoS, ou seja, os valores para os parâmetros de QoS que devem ser satisfeitos para que a qualidade desejada seja obtida;
- **Mapeamento da QoS:** faz o mapeamento automático dos parâmetros de QoS entre as várias camadas e subsistemas do sistema multimídia, visto que as mesmas tratam objetos de dados diferentes;

- **Negociação da QoS:** a fim de que o maior número possível de aplicações seja atendido, este mecanismo visa estabelecer um acordo entre o usuário/aplicação e o sistema em relação aos parâmetros de QoS especificados. Sua tarefa baseia-se nos resultados obtidos pelos mecanismos de controle de admissão e reserva de recursos, descritos abaixo;
- **Controle de Admissão:** com base nos parâmetros de QoS especificados, este mecanismo, através da execução de testes de avaliação, determina se novas sessões podem ser estabelecidas sem que com isso a QoS das sessões já em andamento seja afetada;
- **Reserva de Recursos:** se o mecanismo de controle de admissão determinar o estabelecimento de uma nova sessão, a função deste mecanismo é, através da reserva de determinada quantidade de recursos da rede e dos sistemas fins, manter certo nível de QoS para que os requisitos de QoS especificados sejam satisfeitos;
- **Monitoramento da QoS:** a QoS das sessões já em andamento deve ser monitorada, para que ações adequadas possam ser tomadas, em virtude de eventuais problemas no fornecimento da QoS especificada, e para que as informações relevantes ao funcionamento do sistema possam ser armazenadas, com o objetivo de produzir estatísticas, tais como tempo gasto para a recuperação de um problema e causa mais freqüente dos problemas;
- **Policimento da QoS:** é composto por um conjunto de ações executadas pelo sistema, a fim de que o tráfego gerado pelas aplicações seja monitorado e controlado, objetivando mantê-lo dentro da QoS combinada e preservar as garantias de QoS existentes. Quando o tráfego gerado é violado, o mecanismo de policiamento pode, por exemplo, ignorar a violação do usuário e somente notificá-lo ou encerrar a sessão, ou seja, a transmissão do fluxo;
- **Adaptação da QoS:** este mecanismo deve reagir adaptativamente no caso de ocorrerem *mudanças no ambiente do sistema*, ou seja, o mesmo é utilizado para que a QoS combinada durante a fase de negociação seja mantida o maior tempo possível. Quando a QoS não puder ser mantida, este mecanismo deve degradar a qualidade da sessão;

- **Renegociação da QoS:** este mecanismo é necessário se acaso ocorrerem *mudanças nas especificações de QoS iniciais*. O mecanismo de renegociação pode ser iniciado pelo usuário, caso ele queira aumentar ou diminuir os requisitos de QoS originalmente negociados, ou pelo próprio sistema, devido à falta de recursos, objetivando reduzir a qualidade atualmente fornecida de forma a evitar interrupção do serviço. Neste último caso, este mecanismo só é utilizado quando o mecanismo de adaptação não puder ser executado;
- **Contabilização da QoS:** este mecanismo é muito importante para o gerenciamento da QoS uma vez que é responsável pela contabilização do custo relativo ao serviço solicitado pelo usuário. Não havendo a noção de custo do serviço, o usuário terá razões suficientes para solicitar sempre a melhor qualidade, o que poderá implicar em sobrecarga do sistema e, por conseguinte, na diminuição do número de usuários aceitos, visto que a quantidade de recursos será limitada;
- **Finalização da QoS:** terminado o serviço, o mecanismo de finalização deve enviar a todos os componentes do sistema envolvidos uma notificação indicando que os mesmos já podem liberar todos os recursos reservados para a sessão e, da mesma forma, que todos os processos de programas, tais como codificação/decodificação e envio/recebimento de dados, devem ser destruídos.

Todos estes mecanismos compõem e executam as funções do que se denomina **Gerenciamento da QoS**, cuja finalidade é cumprir adequadamente os requisitos de QoS dos usuários. Segundo o modelo exposto, o processamento dos parâmetros da QoS é constituído de três etapas:

- **Início do Processamento da QoS:** leva à inicialização da sessão multimídia de um determinado usuário (Fig. 4.2);

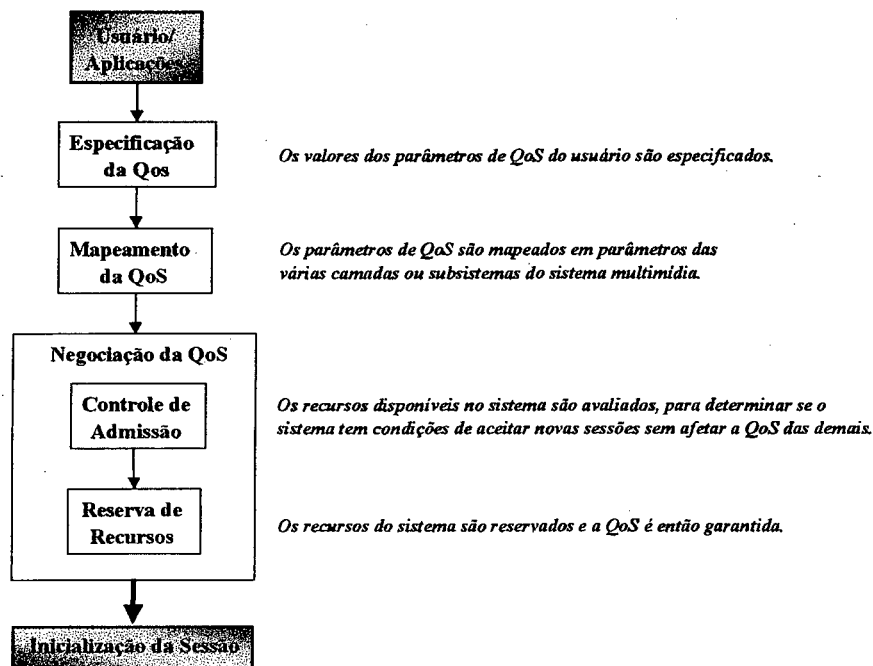


Fig. 4.2 - Início do Processamento da QoS

- **Gerenciamento da QoS:** sua finalidade é administrar a QoS que está sendo fornecida (Fig. 4.3);

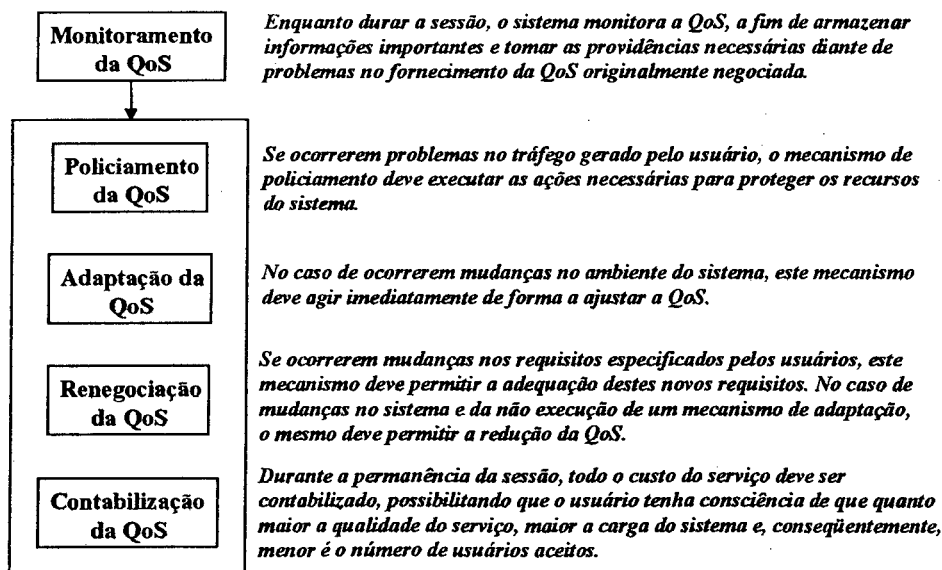
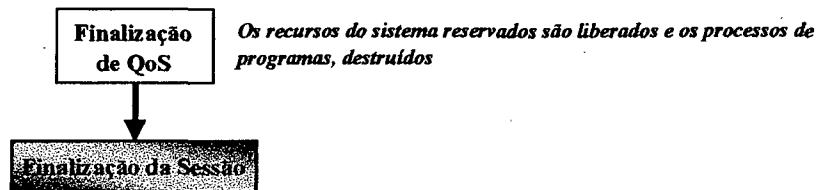


Fig. 4.3 - Gerenciamento da QoS

- **Finalização da QoS:** leva ao fim da sessão (Fig. 4.4).



*Fig. 4.4 - Finalização da QoS*

De acordo com o modelo exposto, existem algumas questões que merecem maior atenção. Primeiro, deve-se dar muita importância à contabilização da QoS, uma vez que a mesma tem grande influência na execução das funções de Gerenciamento da QoS, principalmente, nas funções de negociação e renegociação. Segundo, a execução de mecanismos de adaptação, diante da violação da QoS, deve ser feita automaticamente sempre que possível, a fim de evitar a intervenção do usuário/aplicação.

Finalmente, a interface do usuário deve ser a mais apropriada possível, permitindo que o mesmo especifique facilmente seus requisitos, bem como gerencie e controle a execução do serviço solicitado, por exemplo, através de operações de parada, pausa, avanço, etc.

#### 4.5. CONSIDERAÇÕES FINAIS

Como o trabalho descrito nesta dissertação concentra-se no estudo de transmissões de fluxos multimídia, mais especificamente, transmissões de fluxos MPEG-2 com QoS, a apresentação, nesta seção, dos conceitos referentes à QoS foi fundamental. Como o trabalho também considera que todo o processamento dos fluxos é executado dentro de uma Rede ATM, o próximo capítulo dedicar-se-á ao estudo desta tecnologia.

## CAPÍTULO V

### TECNOLOGIA ATM

*Atualmente, ATM é vista como a tecnologia mais promissora no que diz respeito à implementação de redes de suporte à integração de serviços. Além disso, às especificações desta tecnologia está associado o conceito de Qualidade de Serviço (QoS). Neste capítulo, serão tratados os principais aspectos relacionados à Tecnologia ATM, dentre os quais pode-se citar Modelo de Arquitetura, Modelo de Camadas e Qualidade de Serviço em Redes ATM.*

#### 5.1. INTRODUÇÃO

Inicialmente, os sistemas de comunicação eram voltados ao suporte de tipos de serviços específicos. Contudo, o desenvolvimento da tecnologia digital e a evolução das redes de telefonia proporcionaram o aparecimento das *Redes Digitais de Serviços Integrados (RDSI)*, que são um padrão ITU-T (*International Telecommunications Union*).

O objetivo destas redes é fornecer uma única tecnologia capaz de integrar serviços diversos e, conseqüentemente, tráfegos de naturezas diferentes (como vídeo, áudio e texto), a altas velocidades de transmissão e com economia no compartilhamento de recursos.

Dentre as duas variações existentes para as RDSI, as RDSI-FL (Redes Digitais de Serviços Integrados de Faixa Larga) foram concebidas a partir do aparecimento da fibra óptica, que possibilitou a integração destes serviços por um mesmo meio físico e a altas velocidades, proporcionando um maior aproveitamento da largura de banda.

Desta maneira, as RDSI-FL puderam operar a velocidades mais altas que as RDSI-FE (Redes Digitais de Serviços Integrados de Faixa Estreita), denominação dada às primeiras RDSI, as quais eram tidas como uma evolução natural das redes de telefonia.

De acordo com o ITU-T, a Tecnologia ATM (*Asynchronous Transfer Mode*) é o padrão para a implementação e provisão de RDSI-FL, visto que suporta diferentes serviços, de maneira a satisfazer os requisitos exigidos pelos diferentes tráfegos, a altas velocidades de transmissão.

A criação do *ATM Forum* - consórcio industrial autônomo - deu origem a um conjunto de acordos para a implementação da Tecnologia ATM. Todavia, o mesmo não tem autoridade para a produção de padrões, cabendo esta tarefa a órgãos de domínio internacional. Apesar disto, ainda assim o *ATM Forum* é uma organização importante no desenvolvimento desta tecnologia, uma vez que objetiva favorecer e acelerar sua entrada no mercado, tendo como ponto principal a interoperabilidade de produtos e serviços.

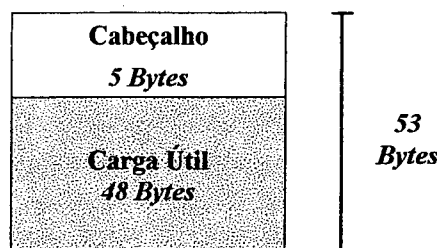
As seções a seguir concentram-se na apresentação dos principais aspectos e características referentes à Tecnologia ATM.

## 5.2. TECNOLOGIA ATM

O Modo de Transferência Assíncrono (ATM) é uma tecnologia de comunicação de dados que se propõe a servir de transporte comum a tráfegos de naturezas diferentes, possibilitando uma forma de padronização de todos os meios de transmissão. Originalmente, foi projetada para Redes de Longa Distância (WANs), porém, devido suas características, tornou-se uma expectativa como solução também para a interconexão de Redes Locais (LANs) a altas velocidades .

### UNIDADES DE TRANSFERÊNCIA DE INFORMAÇÃO

A Tecnologia ATM baseia-se no tráfego de unidades de transferência de informação de tamanho fixo e formato padronizado, chamadas *células*. Cada célula ATM é formada por 53 bytes, dos quais 5 bytes são reservados ao cabeçalho e 48 à carga útil, ou seja, a informação propriamente dita, ver figura 5.1.



*Fig. 5.1 - Formato Genérico da Célula ATM*



Para a rede, a única informação relevante que as células contém é o cabeçalho, uma vez que o mesmo contém as informações necessárias à rede entregar a célula ao receptor correto. Desta forma, para a rede, a natureza e o conteúdo das informações contidas na carga útil são desconsiderados.

Utilizar unidades de tamanho fixo e pequeno implica em algumas vantagens [Soares, 1997; Cereda, 1997]:

- Unidades de tamanho fixo facilitam a previsão de retardos na rede, favorecendo os projetos da mesma e dos serviços que ela pode suportar, visto que determinados serviços, principalmente serviços multimídia, impõem rigorosas restrições aos valores admissíveis dos retardos entre os extremos envolvidos na comunicação;
- O tempo de empacotamento em relação a unidades maiores é menor, favorecendo principalmente sinais de vídeo e áudio, uma vez que o tempo de atraso de transferência é também menor;
- Previsibilidade acrescida no projeto do hardware de suporte dos protocolos de comunicação e simplificação do projeto modular de elementos comutadores da rede.

A única desvantagem encontrada na utilização de unidades pequenas de informação é a diminuição da capacidade efetiva de transmissão da rede, ocasionada pelo grande *overhead* introduzido pelo cabeçalho obrigatório agregado a cada célula. No caso da Tecnologia ATM, este *overhead* aproxima-se a 9,5%.

## **MODO DE TRANSFERÊNCIA**

O termo *modo de transferência* refere-se aos mecanismos de multiplexação e comutação implementados por uma determinada tecnologia [Lu, 1996]. As redes baseadas no Modo de Transferência Assíncrono utilizam-se de mecanismos de Multiplexação Estatística, também denominada Multiplexação por Divisão do Tempo Assíncrono (TDM Assíncrono) e Comutação de Pacotes.

ATM é uma tecnologia de comutação rápida de pacotes, mais especificamente baseia-se em *Cell Relay*, ou seja, no tráfego de células. O equipamento do usuário é responsável por transformar todos os fluxos ou blocos de dados gerados pelas aplicações, para o formato de células, as quais são entregues à rede como uma seqüência de células de tamanho fixo. De acordo com as informações contidas no cabeçalho das células, as mesmas são entregues ao receptor apropriado e reconstituídas para seu formato original.

A Multiplexação Estatística consiste na divisão do tempo sobre o meio de transmissão em segmentos fixos de informação, chamados células. Neste mecanismo não há alocação de canal, nem estabelecimento de conexão. As parcelas de tempo são alocadas dinamicamente conforme a demanda das estações, de forma que quando não houverem usuários dispostos a transmitir, o meio de transmissão fica inativo. Conseqüentemente, não há desperdício da capacidade de transmissão, pois o tempo não utilizado encontra-se sempre disponível caso algum usuário decida transmitir.

Vale enfatizar que ATM é assíncrona no que se refere a células, visto que estas são estatisticamente multiplexadas, ou seja, podem ser transmitidas a qualquer instante, sem que para isso tenham que esperar por uma parcela de tempo periodicamente alocada a cada uma delas. E é síncrona em relação à transmissão de bits, já que o início de uma transmissão só pode ocorrer em intervalos de tempo bem definidos, isto é, quando os relógios do transmissor e receptor estiverem sincronizados, sem a necessidade de bits de início e fim.

## CONEXÕES ATM

Como ocorre com algumas tecnologias de comutação de pacotes, ATM é orientada à conexão, ou seja, um circuito virtual, chamado conexão de canal virtual (VCC – *Virtual Channel Connection*), deve ser estabelecido entre emissor e receptor antes que a transmissão se inicie, fornecendo conexões fim-a-fim compostas da concatenação de enlaces de canais virtuais (VCL – *Virtual Channel Links*). Desta maneira, os recursos da rede necessários à comunicação podem ser alocados antecipadamente, simplificando e agilizando o processo de comutação.

Existem dois tipos de conexões ATM:

- *Circuitos Virtuais Permanentes (PVC - Permanent Virtual Circuits)*: normalmente são configurados estatisticamente por algum mecanismo externo, geralmente uma plataforma de gerenciamento de rede, através da intervenção manual do administrador da rede. Estas conexões são fixas e mantidas, freqüentemente por um longo período de tempo, enquanto não houver outra intervenção;
- *Circuitos Virtuais Comutados (SVC - Switched Virtual Circuits)*: são estabelecidos através de um protocolo de sinalização (ver seção 5.9) ou através da integração de vários destes protocolos, logo, não necessitam de intervenção manual. Desta forma, a conexão deve ser estabelecida antes que a comunicação se inicie e liberada após o término da mesma.

O ITU-T definiu quatro tipos de configurações para as conexões ATM: ponto-a-ponto, ponto-a-multiponto, multiponto-a-ponto e multiponto-a-multiponto. Porém, somente conexões ponto-a-ponto e ponto-a-multiponto foram até o momento tratadas nas Especificações UNI 3.1 e 4.0.

Como o trabalho trata da transmissão de fluxos multimídia entre usuário e servidor de armazenamento multimídia e considera que a troca de informações entre estes é executada somente em conexões comutadas ponto-a-ponto, somente este tipo de conexão é abordado no trabalho.

As conexões ATM ponto-a-ponto são caracterizadas como conexões bidirecionais, onde o emissor e o receptor tem liberdade para especificar seus requisitos independentemente. Por conseguinte, largura de banda assimétrica pode ser suportada. O caminho físico estabelecido entre emissor e receptor e vice-versa deve ser idêntico em ambas as direções. Da mesma forma, a fim de simplificar o gerenciamento da tabela VCC, os valores escolhidos para o endereçamento (VPI/VCI) da conexão (ver seção 5.2.2) devem ser os mesmos para cada enlace no decorrer do caminho.

### 5.3. MODELO DE ARQUITETURA DE UMA REDE ATM

Uma Rede ATM é composta por diversos equipamentos interligados de acordo com o Padrão ATM. Desta forma, o Modelo de Arquitetura ATM é constituído pelos tipos de interconexão, pelos equipamentos necessários para permitir o estabelecimento de conexões e pelas regras necessárias ao bom funcionamento de uma Rede ATM (Fig. 5.2), sendo estes elementos responsáveis pela execução das principais funções desempenhadas nesta arquitetura.

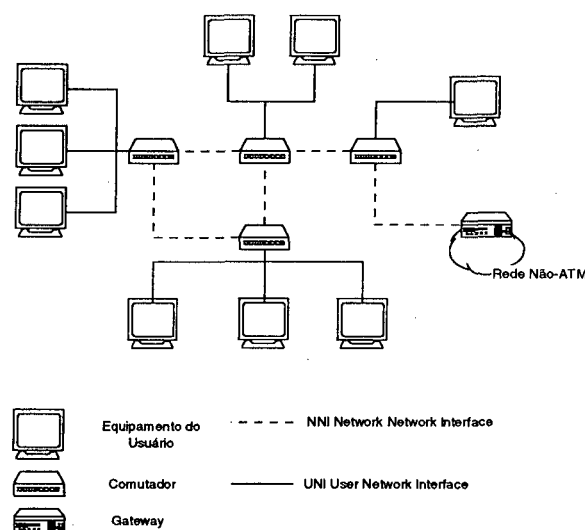


Fig. 5.2 – Modelo de Arquitetura de uma Rede ATM

#### 5.3.1. SISTEMAS FINIS

Os sistemas fins, também chamados equipamentos terminais, ou seja, os equipamentos dos usuário (por exemplo, PCs ou estações de trabalho) e os servidores, necessitam de adaptação para serem incorporados à Rede ATM. Normalmente, isto ocorre através de dispositivos e programas adicionais.

Um exemplo destes dispositivos são as placas adaptadoras, também chamadas NIC (*Network Interface Card*). Estas placas são compatíveis com as características internas do equipamento com que irão interagir. Porém, além destas, a utilização de um software, cuja função é fazer com que o equipamento interaja de maneira correta com a rede, é absolutamente necessária.

### 5.3.2. COMUTADORES

O comutador é o equipamento que faz a interconexão entre os vários sistemas fins instalados na rede, alterando estas conexões quando necessário, para o adequado direcionamento das informações.

Os comutadores possuem duas formas distintas de direcionar as informações (Fig. 5.3): *comutação espacial* e *comutação temporal*. Na comutação espacial há a necessidade de se fazer uma conexão prévia entre a porta de origem e a porta de destino. Na comutação temporal isto não ocorre e logo, a informação precisa trazer explicitamente todos os dados de endereçamento necessários para que a mesma seja enviada ao receptor correto.

O comutador também é responsável por verificar e atualizar os dados contidos no cabeçalho de cada célula ATM, a fim de adaptá-las ao caminho que devem seguir até o próximo comutador.

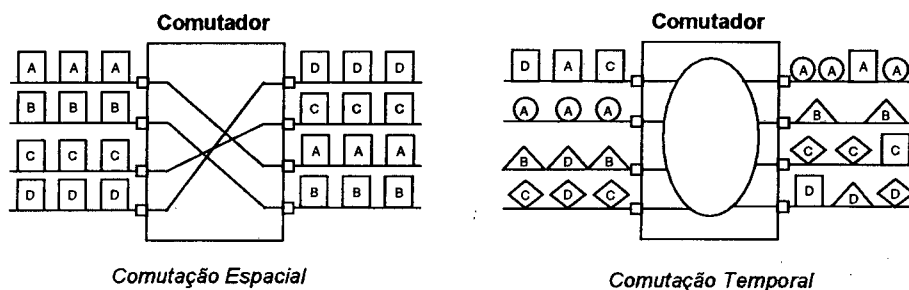


Fig. 5.3 – Direcionamento das Informações nos Comutadores

Outra função do comutador é o armazenamento temporário de células, processo chamado de enfileiramento (*buffering* ou *queuing*). Através do atraso de algumas células, o enfileiramento evita a colisão de células vindas de origens diferentes, porém, com receptor comum.

### 5.3.3. GATEWAYS

Os *gateways* interligam redes distintas logo, são usados para interligar uma Rede ATM a uma Rede não-ATM. Também são responsáveis pela conversão de protocolos de comunicação entre uma rede e outra e pela execução de funções características de um

comutador, ou seja, comutação espacial e temporal, tratamento de cabeçalhos e armazenamento temporário de células.

#### **5.3.4. INTERLIGAÇÃO ENTRE DISPOSITIVOS ATM**

Dois tipos de interconexão entre os dispositivos que fazem parte do Modelo de Arquitetura ATM são alvo de estudos, principalmente, do *ATM Forum* e do ITU-T. São eles: UNI (*User Network Interface*) e NNI (*Network Network Interface*). Ambos têm o objetivo de estabelecer regras para a viabilização de interconexão e interoperabilidade de sistemas em rede.

##### **UNI (*USER NETWORK INTERFACE*)**

A UNI define todas as características da interface entre o sistema do usuário e a rede, incluindo níveis elétricos e físicos, a maneira como as células serão delimitadas e formatadas, os mecanismos para identificação das várias conexões estabelecidas sobre a interface e os mecanismos para os sistemas de usuários especificarem os parâmetros de qualidade de serviço [Fluckiger, 1996].

A UNI tolera facilmente qualquer composição de aplicações no tráfego que passa por ela, a menos que a capacidade máxima do tráfego seja excedida. Isto contribui para que a Tecnologia ATM suporte o tráfego de informações compostas por mídias diferentes [Cereda, 1997].

##### **NNI (*NETWORK NETWORK INTERFACE*)**

A NNI descreve interconexão de redes fazendo esta ligação entre elementos comutadores. A demora de uma especificação por parte do *ATM Forum* para esta interface se deve em grande parte a sua complexidade, principalmente, no que se refere às questões que abrangem meios físicos de comunicação e protocolos envolvidos na operação e gerência da rede, uma vez que entre os elementos comutadores pode existir uma gama de ligações NNI, por onde as informações devem ser roteadas até seu receptor.

#### 5.4. MODELO DE REFERÊNCIA DE PROTOCOLOS

O Modelo de Camadas utilizado no estudo da Tecnologia ATM baseia-se no Modelo de Referência de Protocolos (PRM - *Protocol Reference Model*), definido pela recomendação I.321 do ITU-T para RDSI-FL (Fig. 5.4), a qual suporta serviços e aplicações de vários tipos sobre uma mesma rede [Lu, 1996].

No modelo, as funções executadas na Rede ATM são representadas de forma tridimensional, onde as camadas referentes às funções de controle encontram-se separadas daquelas destinadas às funções que tratam das informações dos usuários.

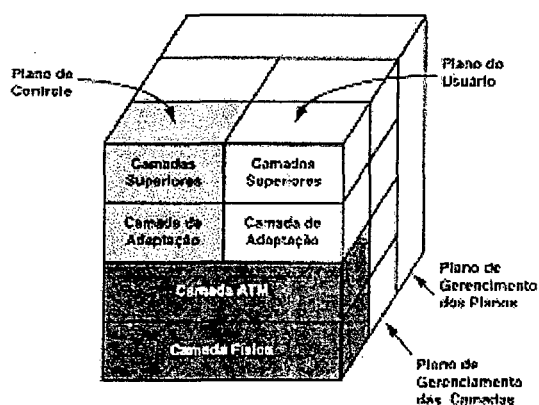


Fig. 5.4 - Modelo de Referência de Protocolos para RDSI-FL

O Modelo de Referência de Protocolos é constituído de três planos:

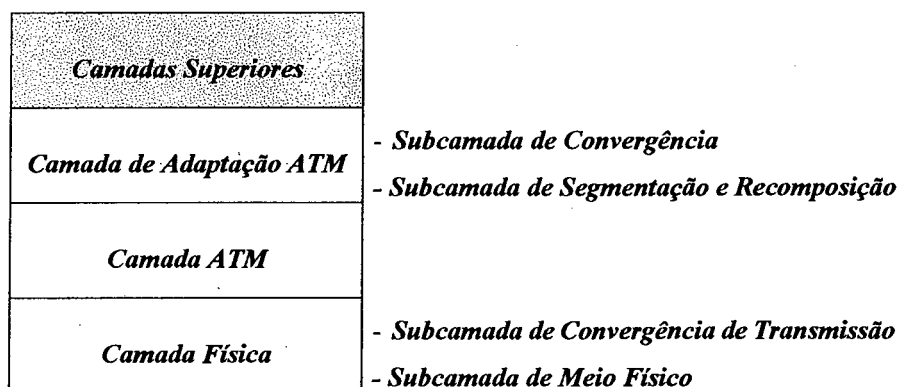
- **Plano do Usuário:** responsável pelo transporte das informações dos usuários, controle de fluxo e recuperação de erros. Divide-se em três camadas inferiores e camadas superiores;
- **Plano de Controle:** responsável pelos aspectos de sinalização, tais como sinalização necessária para estabelecimento, manutenção e finalização de chamadas (conjunto de conexões) e conexões;
- **Plano de Gerenciamento:** responsável por manter a rede e executar funções operacionais. É dividido em dois outros planos:
  - **Plano de Gerenciamento dos Planos:** define funções de gerenciamento relativas ao sistema. Gerencia, de forma integrada, os planos do usuário, de

controle e o próprio plano de gerenciamento. Não se limita a uma camada específica;

- **Plano de Gerenciamento de Camadas:** responsável pela operação e manutenção relativas a cada camada, compreendendo o gerenciamento de recursos e parâmetros associados às entidades de protocolo.

## 5.5. MODELO DE CAMADAS ATM

Como dito anteriormente, o Plano de Usuário é dividido em três camadas inferiores - *Camada Física*, *Camada ATM* e *Camada de Adaptação ATM* - e camadas superiores, as quais não fazem parte das recomendações da RDSI-FL baseadas no modelo acima. A figura 5.5 apresenta a estrutura em camadas do Plano de Usuário e as subdivisões que as Camadas Física e de Adaptação ATM ainda possuem.



*Fig. 5.5 - Modelo de Camadas ATM*

### 5.5.1. CAMADA FÍSICA

Em termos gerais, a camada física é responsável pela transmissão de bits através de um meio de transmissão. Ela define as características elétricas e as interfaces de rede, além da taxa de bits de transmissão. Para transmissão em meios ópticos, as taxas de bits inicialmente propostas para a Tecnologia ATM foram 155 Mbps e 622 Mbps. Esta tecnologia não se prende a um determinado tipo de meio de transmissão, porém, para redes de longa distância o meio mais adequado é a fibra óptica.



A camada física, assim como a camada ATM, está presente em todos os equipamentos da rede e é comum a todos os tipos de tráfego, serviços ou planos, fornecendo facilidades de transmissão e comutação de células [Soares, 1997]. A camada física é dividida em duas subcamadas:

- **Subcamada de Convergência de Transmissão:** é responsável por um conjunto de serviços que devem ser oferecidos à camada ATM, através da execução das seguintes operações:
  - **Geração e Verificação do HEC:** o HEC (*Header Error Control*) corresponde a um campo de controle de erro encontrado no cabeçalho da célula, através do qual erros no cabeçalho podem ser detectados e eventualmente corrigidos. Cabe a esta subcamada, na transmissão, o cálculo e a inserção deste campo no cabeçalho e na recepção, a verificação do mesmo;
  - **Delineamento de células:** delimita a fronteira das células de modo que o receptor saiba onde começa e termina uma célula dentro de uma seqüência de bits ou bytes;
  - **Desacoplamento da Taxa de Células:** uma vez que a Tecnologia ATM utiliza-se de transmissão assíncrona de células, em determinados períodos podem não haver células a transmitir. Desta forma, durante a transmissão, é necessário que células ociosas sejam enviadas com o intuito de manter constante a taxa de geração de informação para a subcamada de meio físico, que normalmente é baseada na transmissão contínua de bits ou bytes pelo meio físico. Na recepção as células ociosas devem ser descartadas.
- **Subcamada de Meio Físico:** é responsável pelas funções dependentes do meio físico, ou seja, características mecânicas, elétricas e ópticas dos meios de transmissão adotados, bem como pela transmissão e recepção adequada do conjunto de bits ou bytes enviados pela subcamada de convergência de transmissão.

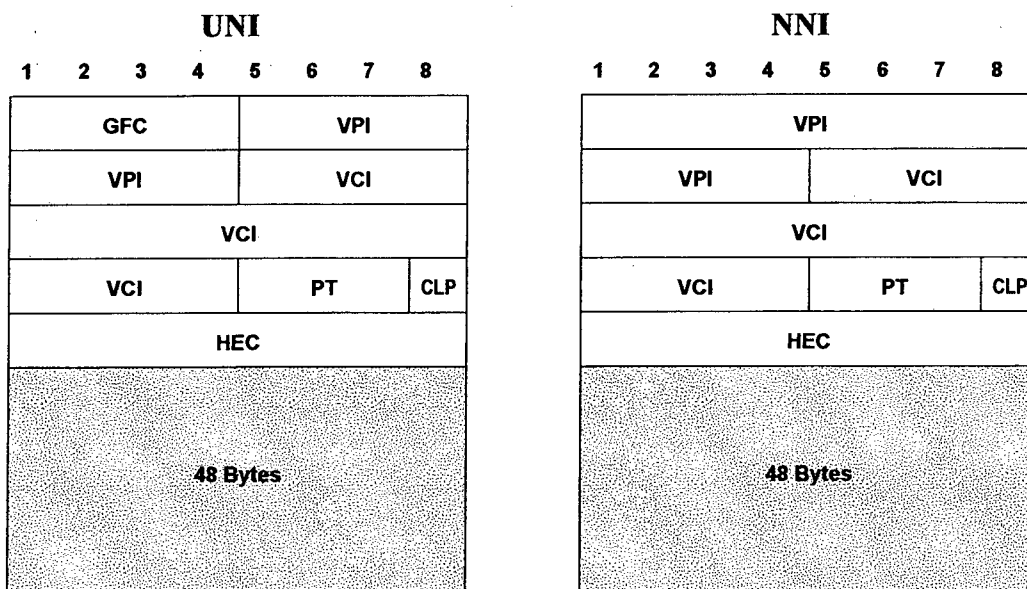
As operações executadas por estas duas subcamadas podem ser resumidas da seguinte maneira: a camada ATM envia à subcamada de convergência de transmissão um fluxo

de células, sobre o qual esta subcamada executa as funções citadas acima, além de transformar o fluxo de células recebido em um fluxo de bits ou bytes. Este fluxo é então enviado à subcamada de meio físico, para que a transmissão ao receptor correto seja finalmente iniciada. No receptor o processo é inverso.

### 5.5.2. CAMADA ATM

A camada ATM é uma camada de multiplexação e comutação completamente independente da camada física. Sua função é direcionar as informações recebidas. Se tiverem chegado ao seu receptor final, as mesmas são enviadas às camadas superiores, caso contrário, de acordo com a conexão previamente estabelecida, são enviadas ao próximo ponto da rede.

Todo o processamento executado pela camada ATM baseia-se na geração e verificação dos campos do cabeçalho da célula, cuja estrutura difere na UNI e na NNI (Fig. 5.6).



*Fig. 5.6 - Formatos da Célula ATM na UNI e na NNI*

Como dito anteriormente, dos 53 bytes da célula ATM 5 destinam-se ao cabeçalho. Somente os campos VPI, VCI e HEC são modificados cada vez que a célula passa por um comutador. A seguir, encontram-se descritos todos os campos do cabeçalho de uma célula:

- **GFC (*Generic Flow Control* - Controle de Fluxo Genérico):** este campo não é ainda utilizado, pois não foi completamente definido. No entanto, segundo o ITU-T o mesmo será utilizado para acesso controlado a determinado fluxo, visando o compartilhamento do meio;
- **VPI (*Virtual Path Identifier* - Identificador de Caminho Virtual):** sua função é identificar uma conexão, ou seja, o VCC. Tem comprimento igual a 8 bits na UNI e igual a 12 bits na NNI, o que possibilita a identificação de um maior número de conexões entre elementos comutadores.
- **VCI (*Virtual Channel Identifier* - Identificador de Canal Virtual):** juntamente com o VPI, possibilita o chaveamento de células por parte dos comutadores. Também identifica a conexão, isto é, o VCL, sendo a parte menos significativa do código formado pelo par VPI + VCI. É composto de 16 bits.
- **PT (*Payload Type* - Tipo de Informação):** identifica o tipo de informação contida na célula e, em alguns casos, pode armazenar informações adicionais, tais como se a célula passou por algum nó congestionado. É formado por 3 bits.
- **CLP (*Cell Lost Priority* - Prioridade de Perda de Célula):** indica prioridade no descarte de células, podendo ser modificado pelo usuário ou por qualquer outro elemento de rede. No caso de eventuais problemas na rede, as células que tiverem o campo CLP igual a 1 são as primeiras a serem descartadas. Campo de 1 bit.
- **HEC (*Header Error Control* - Controle de Erro de Cabeçalho):** como visto anteriormente, a camada física é responsável pela sua inserção e verificação. É utilizado para o delineamento das células e para a identificação de erros no cabeçalho das mesmas, os quais dependendo da complexidade, podem ou não ser corrigidos. Campo composto de 8 bits.

Uma conexão ATM é identificada pelos campos VPI e VCI. Comparando o modo como uma conexão ATM é identificada com o sistema de números telefônicos, pode-se dizer que o VPI corresponde ao código regional, enquanto que o VCI corresponde ao número local de um número telefônico.

Embora cada sistema fim ATM tenha um identificador fixo, a Tecnologia ATM baseia-se nos campos VPI e VCI para fazer o roteamento das células, os quais são alocados dinamicamente para cada conexão.

Quando um comutador recebe uma célula, primeiramente ele examina o campo VPI, a fim de determinar se o receptor da mesma é local. Se a resposta for positiva, o comutador verifica o campo VCI com a finalidade de determinar para qual sistema fim a célula é destinada. Caso contrário, a mesma é repassada a outro comutador.

Sobre um mesmo meio físico, vários caminhos virtuais (VPIs) podem ser estabelecidos. Da mesma forma, vários canais virtuais (VCIs) podem ser encontrados sobre um único caminho virtual, ver figura 5.7. Diante disto, entre pares de sistemas fins podem ser estabelecidas várias conexões. A vantagem deste tipo de endereçamento é que tanto a comutação de células quanto o estabelecimento de chamadas e conexões tornam-se mais rápidos.

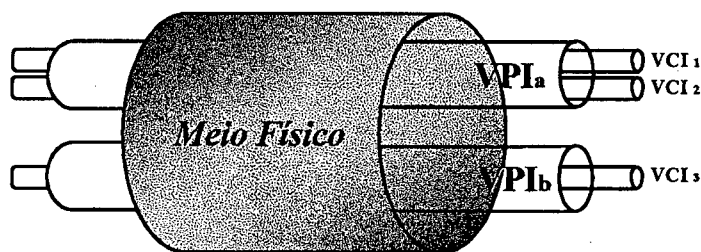


Fig. 5.7 – Caminhos e Canais Virtuais

### 5.5.3. CAMADA DE ADAPTAÇÃO ATM (AAL)

A função da camada de adaptação ATM é fazer o mapeamento entre a camada ATM e a camada superior mais próxima, ou seja, fazer uma adaptação melhorada dos serviços fornecidos pela camada ATM aos requisitos da camada mais alta. [Raghavan, 1998].

O termo *camada de adaptação* refere-se ao processo de adaptação dos dados na forma adequada para a Tecnologia ATM. Mais especificamente, refere-se a todos os protocolos utilizados para a segmentação de dados em células e o reagrupamento das mesmas em dados. Esta é a primeira camada de protocolo fim-a-fim do Modelo de Referência da RDSI-FL (Fig. 5.8).

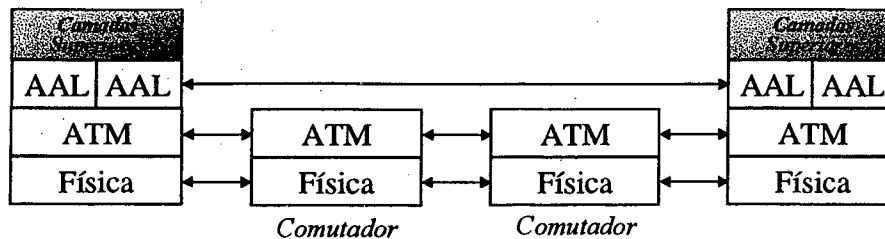


Fig. 5.8 - Camada AAL

Desta maneira, o objetivo da AAL é empacotar eficientemente os vários tipos de dados (arquivos alfanuméricos, amostras de áudio, quadros de vídeo, etc.), enviados das camadas superiores à camada ATM, em uma série de células, as quais serão enviadas através de conexões ATM e reconstituídas no formato apropriado de dados no receptor [Lu, 1996].

As funções da camada de adaptação ATM são divididas entre duas subcamadas:

- **Subcamada de Convergência:** é responsável pela multiplexação de serviços, detecção de erros na carga útil e perdas de células, e recuperação da relação temporal da informação original no receptor. Porém, a execução de uma ou outra função dependerá do tipo de serviço que está sendo oferecido.
- **Subcamada de Segmentação e Recomposição:** na transmissão, sua função é segmentar os dados, cujo tamanho varia de acordo com a aplicação, em células de 48 bytes, aos quais posteriormente a camada ATM acrescentará os 5 bytes de cabeçalho. Na recepção, sua função é fazer a recomposição dos dados vindos da camada ATM às camadas superiores;

Como exposto anteriormente, a AAL deve oferecer serviços às camadas mais altas. Como cada serviço possui seus próprios requisitos, o ITU-T definiu para a AAL quatro classes de serviços, rotuladas de A a D, a fim de fazer com que a mesma suporte os diferentes requisitos das aplicações. Estas classes são classificadas segundo:

- **Relação temporal entre emissor e receptor (requerido ou não-requerido):** dependerá do tipo de tráfego ser ou não sensível a atrasos e variações de atrasos durante a transmissão e a entrega dos dados ao receptor;

- **Taxa de bits (constante ou variável):** considera-se as taxas com que as informações foram geradas pela aplicação, para transmissão pela Rede ATM, e não a taxa de transmissão pelo meio físico.
- **Modo de conexão (orientado a conexão ou não-orientado a conexão):** considera-se a natureza da aplicação nas camadas superiores à AAL.

Após a definição do modelo de classes de serviços baseado na camada AAL, quatro tipos de protocolos foram definidos. Para cada protocolo, algumas características foram especificadas, tal como o formato da unidade de informação que será apresentado à subcamada de segmentação e recomposição .

Estes protocolos devem dar suporte aos diferentes tipos de tráfegos gerados pelos diferentes tipos de serviços. A tabela abaixo caracteriza as quatro classes de serviços, relacionando-as aos quatro protocolos AAL.

	CLASSE A	CLASSE B	CLASSE C	CLASSE D
<b>Relação Temporal entre Emissor e Receptor</b>	<i>Requerido</i>		<i>Não-Requerido</i>	
<b>Taxa de Bits</b>	<i>Constante</i>	<i>Variável</i>		
<b>Modo de Conexão</b>	<i>Orientado a Conexão</i>			<i>Não-Orientado a Conexão</i>
<b>Protocolo</b>	<i>AAL1</i>	<i>AAL2</i>	<i>AAL3/4</i>	<i>AAL3/4</i>

**Tab. 5.1 – Classes de Serviços e Protocolos da Camada de Adaptação ATM**

Existem ainda o protocolo AAL0, também conhecido como AAL Nulo, o mesmo determina ausência de funções da camada AAL, porém, representa o fornecimento de serviços da camada ATM e a AAL de Sinalização (como será apresentado mais adiante), a qual é utilizada no suporte às conexões de sinalização entre sistemas fins e comutadores.

Como se pode perceber, os protocolos AAL3 e AAL4 foram convertidos a um único protocolo (AAL3/4). Isto se deve ao fato de ambos executarem procedimentos comuns.

O *ATM Forum*, com o objetivo de simplificar as tarefas executadas por este protocolo, propôs a criação de um quinto protocolo, chamado AAL5.

Mesmo não oferecendo todas as funções do AAL3/4 e apesar de mais simples, o AAL5 opera de maneira mais eficiente, com melhor capacidade de detecção e correção de erros, menor *overhead* e requisitos de processamento mais simples. Assim como o AAL3/4, o AAL5 fornece serviços de classe C e D.

Atualmente, o *ATM Forum* está trabalhando em um sexto protocolo, o AAL6, que visa o empacotamento de fluxos multimídia, em especial fluxos de vídeo MPEG e MPEG-2. A idéia é utilizar técnicas de correção de erros (FEC) para aumentar a confiabilidade da comunicação de maneira que a recuperação de erros extras não seja necessária e que os requisitos de sincronização MPEG sejam suportados [Lu, 1996].

#### CAMADA DE ADAPTAÇÃO AAL5

Devido as suas características e o fato de ter implementação garantida em quase todos os equipamentos ATM [Cereda, 1997], o AAL5 tornou-se o protocolo da camada AAL mais importante.

Denominado originalmente de SEAL (*Simple and Efficient Adaptation Layer* – Camada de Adaptação Eficiente e Simples), o AAL5 foi projetado para operar de maneira mais eficiente que o AAL3/4, pois a indústria de computação acreditava que este protocolo implicava em muito processamento e overhead de transmissão. O AAL5 apresenta funções comuns ao AAL3/4, com exceção à multiplexação:

- **Detecção de Erros na Carga Útil:** ao invés de detectar erros ao nível de células, o protocolo AAL5 fornece proteção CRC (*Cyclic Redundancy Check*) ao nível de unidades de informação de protocolo (PDU – *Protocol Data Unit*), partindo do princípio que a fibra óptica será largamente utilizada nas Redes ATM e que, por conseguinte, os enlaces de transmissão de dados serão razoavelmente não susceptíveis a erros. O esquema de proteção de erros AAL5 baseia-se em CRC de 32 bits e pode ser utilizado na detecção de erros de bits e perda de células. Para a detecção de perda de células, o AAL5 utiliza ainda checagem de tamanho.

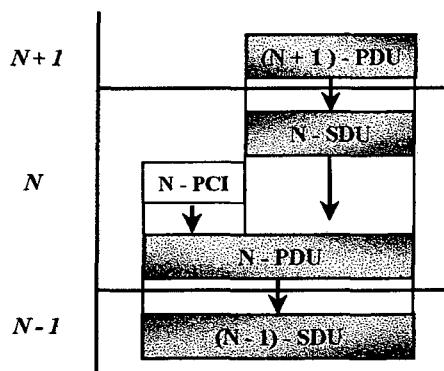
- **Delineamento de Dados:** ao invés de indicar a posição de cada célula na AAL5-PDU, o protocolo AAL5 indica, através de um campo chamado PTI (*Payload Indicator*), somente o final da AAL5-PDU, ou seja, a última célula pertencente ao mesmo grupo.

Como visto anteriormente, a camada AAL é subdividida em subcamada de convergência e subcamada de segmentação e recomposição. Para a AAL5 a subcamada de convergência dividi-se ainda em duas outras subcamadas – **Subcamada de Convergência Específica do Serviço (SSCS - Service Specific Convergence Sublayer)** e **Parte Comum da Subcamada de Convergência (CPCS - Common Part Convergence Sublayer)**.

A SSCS é responsável por procedimentos específicos executados no suporte de serviços também específicos. Esta subcamada pode receber valor zero, caso a execução de funções específicas não será necessária. A CPCS é responsável pela execução dos procedimentos comuns a todas as classes de serviço.

## 5.6. TROCA DE UNIDADES DE INFORMAÇÃO ENTRE CAMADAS

O processo de transmissão do fluxo de informações entre as camadas AAL é executado através de um processo de troca de unidades de informações, ver figura abaixo, onde:



*Fig. 5.9 – Troca de Unidades de Informação entre Camadas*



- **SDU (Service Data Unit):** é a unidade de informação de serviço recebida por uma camada através da interface com a camada superior;
- **PDU (Protocol Data Unit):** é a unidade de informação de protocolo entregue a camada inferior. Esta unidade é gerada a partir da adição de informações de controle da camada (PCI) à SDU recebida da camada superior;
- **PCI (Protocol Control Information):** é a informação de controle de protocolo, que adicionada à SDU, transforma-se em uma PDU da camada.

O fluxo de informações e o processo de troca de unidades de informações encontram-se ilustrados na figura abaixo:

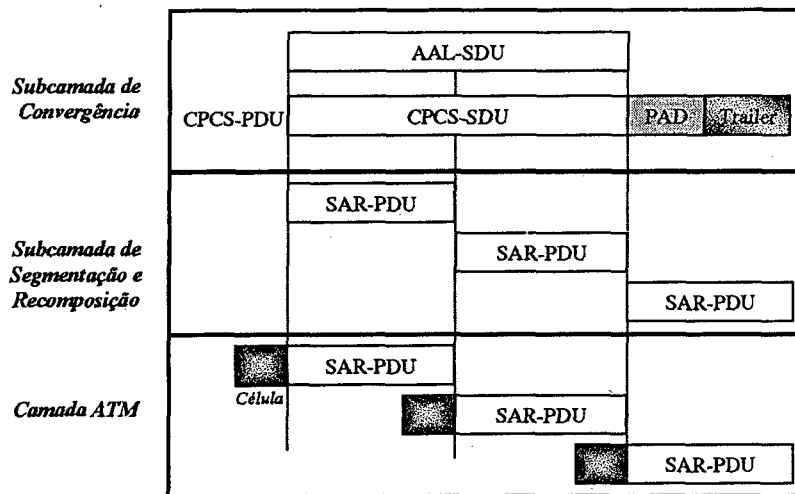


Fig. 5.10 – Troca de Unidades de Informações entre Camada AAL e Camada ATM

- 1- O processo se inicia quando a **AAL-SAP** (ponto de acesso a serviços) recebe uma **AAL-SDU**, a qual pode ou não ser processada por uma **SSCS**;
- 2- A **AAL-SDU** é então repassada a camada **CPCS**, transformando-se em uma **CPCS-SDU**;
- 3- A esta **CPCS-SDU** é adicionado um *trailer*, cujo tamanho é 8 bytes. Se necessário, alguns bits de adaptação também são adicionados como ajuste, caso o tamanho da **CPCS-SDU** não seja múltiplo de 48. Desta forma, uma **CPCS-PDU** é obtida.

- 4- A **CPCS-PDU** é então entregue à subcamada de segmentação e recomposição (SAR) tornando-se uma **SAR-SDU**;
- 5- A camada de segmentação e recomposição segmenta a **SAR-SDU** transformando-a em **SAR-PDUs** de 48 bytes, sem adicionar campos de overhead e indica, através do campo PTI, qual a última célula do grupo de células originado pela mesma **SAR-SDU**;
- 6- Estes segmentos são entregues à camada ATM, transformando-se, conseqüentemente, em **ATM-SDUs**;
- 7- Finalmente, a camada ATM adiciona a cada célula o cabeçalho de 5 bytes e inclui no campo PT a indicação da última célula do grupo.

No destino o processo inverso é executado. Os campos presentes na CPCS-PDU estão ilustrados na figura abaixo e encontram-se descritos logo a seguir.



Fig. 5.11 – Campos da CPCS-PDU

- **CPCS-SDU**: no caso de ausência da subcamada SSCS, este campo compõe-se da própria AAL-SDU;
- **PAD**: este campo só é utilizado caso seja necessário ajustar o tamanho da PDU a um número múltiplo de 48;
- **CPCS-UU (CPCS User to User Indicator)**: este campo é utilizado para o transporte de informações entre CPCSs;
- **CPI (Common Part Indicator)**: a utilização deste campo ainda esta em fase de estudos, sua provável função será o alinhamento do *trailer*;
- **Data Length**: informa o tamanho da informação. Desta forma é possível saber o quanto da CPCS-PDU é informação e o quanto é PAD;

- **Data Length:** informa o tamanho da informação. Desta forma é possível saber o quanto da CPCS-PDU é informação e o quanto é PAD;
- **CRC (Cyclic Redundancy Check):** esquema utilizado na detecção de erros na CPCS-PDU.

### 5.7. CATEGORIAS DE SERVIÇO ATM

Uma das maiores características da Tecnologia ATM é a capacidade de suportar pedidos de serviços com características diferentes, durante o estabelecimento de uma conexão. Diante disto, como visto na última seção, baseado nas diferenças qualitativas dos diversos tipos de aplicação, o ITU-T definiu quatro classes de serviços para a camada AAL.

Porém, os atributos qualitativos utilizados nesta classificação, tal como modo de conexão, são irrelevantes à camada ATM e descrevem de maneira insuficiente o serviço a ser oferecido, não proporcionando um modelo operacional que usuário e rede possam utilizar efetivamente durante o estabelecimento de chamadas e conexões e alocação de recursos.

Sendo assim, um novo modelo de classe de serviço foi proposto, como uma abordagem baseada em parâmetros, que relacionam características de tráfego e requisitos de QoS para o comportamento (desempenho) da rede, sem dividir os serviços em classes de aplicações.

Como somente parâmetros quantitativos são utilizados na descrição de uma aplicação. Segundo este novo modelo, todas as aplicações são representadas como parte de uma série contínua de possíveis combinações de valores de todos estes parâmetros [Kwok, 1997].

Com o objetivo de diferenciar as classes de serviços baseadas na camada AAL daquelas baseadas na camada ATM, as primeiras serão tratadas como *classes de serviço ATM*, enquanto que as posteriores serão referenciadas como *categorias de serviço ATM*.

As funções da Rede ATM, tais como estabelecimento de conexões e alocação de largura de banda, são estruturadas de maneira diferente em cada categoria de serviço. Da mesma forma são estruturados os métodos utilizados para obtenção de garantias de

serviço fornecidas pela rede e os mecanismos implementados nos sistemas fins e protocolos de camadas mais altas para a realização de cada categoria de serviço.

Dentre as cinco categorias de serviços definidas para a camada ATM, duas caracterizam-se como serviços de tempo real (CBR e rt-VBR) e três como serviços de tempo não real (nrt-VBR, ABR e UBR). As categorias de serviço ATM são:

- **CBR (Constant Bit Rate):** destina-se às aplicações tempo real, cujo tráfego é transmitido a uma taxa constante de bits e o emissor e o receptor encontram-se sincronizados no tempo. O serviço CBR suporta qualquer tipo de dados, desde que a transmissão ocorra a um tempo de resposta previsível e que a quantidade de largura de banda, durante toda a conexão, permaneça estática e continuamente disponível. O serviço CBR é altamente sensível a atrasos e intolerante no que se refere à perda de células;
- **rt-VBR (Real Time Variable Bit Rate):** este serviço é utilizado por conexões que transmitem à taxa variável de bits com determinadas restrições de tempo. Este tipo de serviço é utilizado por aplicações sensíveis ao tempo, tais como aplicações de videoconferência, cujo tráfego gerado é codificado e aplicações multimídia, cuja perda de células, desde que em pequena quantidade, é tolerada sem que a degradação da qualidade da apresentação seja notada. Para este tipo de serviço o atraso fim-a-fim é crítico;
- **nrt-VBR (Non Real Time Variable Bit Rate):** é utilizado para transmissão a taxa de bits variável. Não há necessidade de que os sistemas fins estejam sincronizados no tempo, porém, a largura de banda e a latência devem ser garantidas. Tanto o serviço rt-VBR quanto o nrt-VBR podem ser utilizados por qualquer tipo de aplicação que possa se beneficiar do envio de dados a taxas variáveis, a fim de utilizar mais eficientemente os recursos da rede. Este serviço é utilizado por aplicações orientadas a transações, tais como sistemas de reserva interativos, onde o tráfego é esporádico e em rajadas. Não tem restrições quanto a atraso e variação de atraso;
- **ABR (Available Bit Rate):** é similar ao serviço nrt-VBR, pois também é utilizado por conexões que transmitem a taxas variáveis de bits, para o qual não há garantias de sincronização no tempo entre emissor e receptor. Contudo, o serviço

ABR não garante largura de banda e latência. O ABR fornece um serviço de transmissão de melhor esforço, no qual mecanismos de controle de fluxos são utilizados no ajuste dinâmico da quantidade de largura de banda disponível para o emissor. A princípio, o ABR destina-se a qualquer tipo de tráfego insensível ao tempo e que não necessite de garantias de serviço. O mesmo visa atingir fontes de tráfego que sejam capazes de reduzir ou aumentar a taxa de informação gerada quando necessário, permitindo explorar as mudanças nas características de tráfego da camada ATM (por exemplo, largura de banda disponível) logo após o estabelecimento da conexão. O serviço ABR garante somente uma quantidade mínima de largura de banda e pode limitar-se à taxa de transmissão máxima especificada;

- **UBR (*Unspecified Bit Rate*)**: também é similar ao nrt-VBR, contudo, diferentemente do ABR, não implementa mecanismos de controle de fluxo. Normalmente, é usado por aplicações que não impõem restrições quanto ao atraso, variação de atraso, perda de células e especificação de qualidade de serviço. O serviço UBR também tem sido muito utilizado em Redes Locais Internet e em ambientes de Redes de Longa Distância (WAN - *Wide Area Network*) para armazenamento e transmissão de tráfego, tais como transferência de arquivos e e-mail. É um serviço de melhor esforço, que tem o objetivo de atender aplicações não-críticas, uma vez que compartilha a largura de banda restante na rede sem qualquer mecanismo de realimentação específico, não incluindo noções de negociação de largura de banda por conexão e não especificando garantias de serviço relacionadas ao tráfego.

Descritas as categorias de serviço ATM, pode-se fazer uma associação entre estas e as classes de serviço ATM:

- **Classe de Serviço A**: destina-se a atender aos tráfegos gerados a taxa de bits constante (CBR).
- **Classe de Serviço B**: destina-se a atender aos tráfegos gerados a taxa de bits variável em tempo real (rt-VBR).

- **Classe de Serviço C:** destina-se a atender aos tráfegos gerados a taxa de bits variável em tempo não real (nrt-VBR).
- **Classe de Serviço D:** destina-se a atender aos tráfegos gerados tanto a taxa de bits disponível (ABR), quanto à taxa de bits não especificada (UBR).

## 5.8. CONTRATO DE TRÁFEGO E PARÂMETROS DE TRÁFEGO ATM

O *contrato de tráfego* é um acordo firmado entre o sistema fim e a rede antes do estabelecimento da conexão, ou seja, durante a fase de sinalização, que especifica as características de tráfego e os requisitos de QoS da aplicação na camada ATM. Estas características e requisitos são definidos, principalmente, por dois elementos: *descriptor de tráfego da conexão* e *parâmetros de QoS de cada sentido de transmissão*.

O descriptor de tráfego da conexão especifica as características de tráfego da conexão ATM como um todo; o mesmo é constituído dos seguintes elementos:

- **Descriptor de Tráfego do Emissor:** define as características de tráfego da conexão ATM a ser estabelecida pelo emissor, as quais são especificadas pelos *parâmetros de tráfego ATM*;
- **CDVT (Cell Delay Variation Tolerance):** especifica o quanto a Rede ATM é tolerante em receber tráfego cujo valor encontra-se acima do valor estipulado pelo parâmetro de tráfego ATM PCR (descrito logo abaixo). Dependendo do valor do CDVT, do PCR e da taxa de transmissão do meio, células consecutivas podem ser transmitidas e ainda assim serem consideradas de acordo com o que foi especificado no contrato de tráfego;
- **Parâmetro de Definição de Conformidade:** a conformidade é determinada pelo policiamento de tráfego. Com base neste parâmetro pode-se avaliar se o tráfego gerado está ou não de acordo com o que foi especificado no contrato de tráfego. O mesmo especifica a conformidade para conexões CBR, rt-VBR, nrt-VBR e UBR, uma vez que a conformidade para conexões ABR são modificadas com base na carga da rede e nos procedimentos de controle de fluxo.

Os parâmetros de tráfego ATM especificam as características de tráfego do emissor na camada ATM, os mesmos estão relacionados abaixo:

- **Peak Cell Rate (PCR):** refere-se à taxa máxima permissível em que células podem ser transmitidas ao longo de uma conexão na Rede ATM. Este parâmetro é expresso em células por segundo;
- **Sustainable Cell Rate (SCR):** refere-se à taxa média (em células por segundo) permissível em que as células podem ser transmitidas em longo prazo sobre uma conexão específica;
- **Maximum Burst Size (MBS):** refere-se ao número máximo de células que pode ser transmitido continuamente através de uma conexão específica à taxa especificada pelo parâmetro PCR;
- **Minimum Cell Rate (MCR):** refere-se à taxa mínima permissível na qual as células podem ser transmitidas ao longo de uma dada conexão ATM, ou seja, é a mínima largura de banda alocada para uma conexão. Este parâmetro não descreve o comportamento do tráfego e é utilizado por serviços sob demanda (como ABR), a fim de garantir que uma conexão não sofrerá de falta de alimentação quando não houver mais largura de banda disponível.

No momento do estabelecimento da conexão, a rede determina se pode cumprir o contrato de tráfego, o que implica em aceitar ou não a conexão e satisfazer os requisitos de tráfego e de QoS associados à mesma. A partir do momento em que a conexão é aceita, cabe a rede garantir que os requisitos de tráfego e de QoS especificados e negociados no contrato de tráfego serão satisfeitos, desde que a parte do contrato cabível ao usuário seja cumprida.

### 5.8.1. PARÂMETROS DE QOS ATM

A qualidade de serviço em Redes ATM é tratada como o efeito que as características de desempenho de um serviço provocam sobre o sistema, de maneira que o grau de satisfação do usuário possa ser determinado [Soares, 1997]. Sendo assim, pode-se dizer

que os parâmetros de QoS são definidos sob a perspectiva do usuário de um dado serviço.

Os requisitos de QoS do usuário são especificados pelos parâmetros de QoS ATM na Camada ATM. É de responsabilidade das camadas superiores, inclusive da camada de adaptação ATM, o mapeamento destes parâmetros de forma adequada à camada ATM.

Segundo o ITU-T, somente os parâmetros cuja observação e mensuração podem ser feitas diretamente no ponto de acesso do usuário, tais como retardo e taxa de perda de células, têm importância para a definição da QoS na Camada ATM. A tabela abaixo apresenta os parâmetros de QoS ATM:

TIPO	PARÂMETRO	SIGLA	SIGNIFICADO
<b>Características de rede medidas diretamente no receptor</b>	<i>Cell Transfer Delay</i>	CTD	Tempo máximo gasto na entrega de células
	<i>Cell Delay Variation</i>	CDV	Varição de atraso na entrega de células
	<i>Cell Lost Ratio</i>	CLR	Fração de células perdidas ou entregues com atraso
<b>Características de rede não negociáveis</b>	<i>Cell Error Ratio</i>	CER	Fração de células entregue com erro
	<i>Severly Errored Cell Block Ratio</i>	SECBR	Fração de blocos distorcidos
	<i>Cell Misinsertion Rate</i>	CMR	Fração de células entregue ao receptor errado

*Tab. 5.2 - Parâmetros de QoS ATM*

Como visto na seção 5.5.3, o ITU-T definiu para a camada de adaptação ATM quatro classes de serviços ATM. Cada classe de serviço relaciona-se a um tipo de serviço ATM e a uma classe de QoS ATM e seus parâmetros associados. Segundo [Ferguson, 1998], existem dois tipos de classes de QoS ATM:

- **Classe de QoS Especificada:** nesta classe os parâmetros de desempenho são explicitamente especificados;



- **Classe de QoS Não Especificada:** nesta classe os parâmetros de desempenho não são diretamente especificados. Os parâmetros são selecionados pela rede ou escolhidos dentre valores padrões<sup>5</sup>, resultando em uma classe de QoS implícita.

De acordo com estes dois tipos de classes de QoS ATM e da relação entre estas e as classes de serviços ATM, as seguintes classes de QoS especificadas foram definidas:

- **Classe 1:** a QoS suportada por esta classe satisfaz os requisitos de desempenho da classe de serviço ATM A. Esta classe deve prover desempenho similar àquele fornecido pelas linhas privadas digitais;
- **Classe 2:** suporta QoS que atenda aos requisitos de desempenho da classe de serviço ATM B. Visa prover desempenho aceitável durante o empacotamento de vídeo e áudio em aplicações multimídia e de teleconferência;
- **Classe 3:** a QoS suportada por esta classe satisfaz os requisitos de desempenho da classe de serviço ATM C. Objetiva fornecer desempenho aceitável para a interoperabilidade de protocolos orientados a conexão, tal como *Frame Relay*;
- **Classe 4:** suporta QoS que atenda aos requisitos de desempenho da classe de serviço ATM D. Deve fornecer desempenho aceitável para a interoperabilidade de protocolos não orientados a conexão, tais como IP.

Cada conexão ATM é associada a uma dada classe de QoS, a qual especifica um conjunto de parâmetros de desempenho, ou seja, a especificação da classe de QoS, para uma dada conexão ATM, indica qual a QoS solicitada.

Contudo, não é necessário que todas as conexões ATM sejam associadas a uma QoS especificada, uma vez que se uma ou mais conexões não utilizarem completamente a capacidade total da QoS reservada no contrato de tráfego, grande quantidade dos recursos alocados para estas conexões será desperdiçada e, neste caso, a melhor solução é associar ao contrato de tráfego uma classe de QoS não especificada. Esta é a única

---

<sup>5</sup> Estes valores encontram-se indicados na Especificação "Gerenciamento de Tráfego" do ATM Forum.

classe que deve ser suportada por todas as redes e a única permitida quando um serviço de melhor esforço for solicitado.

Na verdade, os parâmetros de QoS que fazem parte do contrato de tráfego e que são negociados antes do estabelecimento da conexão, se enquadram em uma ou outra classe de QoS especificada.

A manutenção da QoS negociada em uma Rede ATM é feita através de técnicas de controle de tráfego e congestionamento, as quais devem ter conhecimento de determinados parâmetros de QoS, a fim de agirem de forma eficiente.

Muito embora os conceitos de QoS em Redes ATM impliquem em grandes possibilidades, existem vários fatores que contribuem para que o processamento da QoS nestas redes ainda seja problemático em certos aspectos.

O maior desafio diz respeito à negociação da QoS de novas conexões, sem que a garantia de QoS das conexões já em andamento seja afetada, principalmente, quando estas atravessarem diferentes ambientes operacionais ou até mesmo diferentes domínios administrativos.

### **5.9. PROCESSO DE SINALIZAÇÃO ATM**

Cabe ao protocolo de sinalização, através da definição de um conjunto de procedimentos e informações de sinalização, tornar a aplicação capaz de estabelecer, monitorar e finalizar uma conexão ATM (circuito virtual) sempre que necessário.

As principais funções das informações de sinalização são indicar quais são os requisitos de rede (requisitos de tráfego e qualidade de serviço) que precisam ser alocados para o estabelecimento de uma conexão e transportar parâmetros para a checagem de compatibilidade entre os sistemas fins.

As informações de sinalização são transmitidas em conexões especiais, chamadas VCC de Sinalização (SVCC – *Signaling VCC*), que se caracterizam como conexões ponto-a-ponto e, normalmente, podem ser identificadas pelo par VPI=0 e VCI=5.

Os procedimentos de sinalização, por sua vez, especificam como e quais são as informações de sinalização que devem ser trocadas para o estabelecimento, a monitoração e a finalização de uma conexão ATM.

Todos os procedimentos de sinalização são de responsabilidade do Plano de Controle, que utiliza a camada ATM como meio de transmissão de células com informações de sinalização e uma camada de adaptação específica, denominada AAL de Sinalização.

### **5.9.1. TROCA DE MENSAGENS DE SINALIZAÇÃO**

A execução de funções de controle e troca de informações para o estabelecimento, monitoração e finalização de uma conexão, durante a fase de sinalização, ocorre através da troca de mensagens entre o emissor, os elementos comutadores e, finalmente, o receptor. As mensagens vindas da UNI para a NNI precisam ser mapeadas em mensagens de sinalização NNI e vice-versa.

Como o trabalho considera somente conexões ponto-a-ponto, apenas as mensagens de sinalização definidas para este tipo de conexão serão descritas:

#### ***1 - Estabelecimento da Conexão***

- **SETUP:** enviada pelo emissor como pedido de estabelecimento de uma nova conexão. Esta mensagem deve conter o endereço do emissor e do receptor, bem como os requisitos especificados no contrato de tráfego e o endereço do circuito virtual que potencialmente será estabelecido;
- **CALL PROCEEDING:** esta mensagem é opcional e é utilizada caso o tempo de espera do emissor por uma resposta precise se estendido por mais alguns segundos, indicando que a conexão ainda permanece em fase de estabelecimento e a mensagem **CONNECT** ainda não pôde ser enviada;
- **CONNECT:** indica que o receptor e o restante dos elementos de rede envolvidos aceitaram o estabelecimento da conexão. Esta mensagem deve conter os identificadores utilizados para a conexão no sentido receptor-emissor;
- **CONNECT ACK:** esta mensagem é opcional e é enviada pelo emissor, a fim de que o estabelecimento da conexão seja confirmado.

## 2 – Monitoração da Conexão

- STATUS ENQUIRY: pedido de informação sobre o comportamento da conexão;
- STATUS: mensagem enviada como resposta à mensagem anterior.

## 3 – Finalização da Conexão

- DISCONNECT: enviada pelo emissor como pedido de finalização da conexão estabelecida;
- RELEASE: indica que os recursos alocados à conexão estão sendo liberados;
- RELEASE COMPLETE: confirmação da liberação dos recursos alocados.

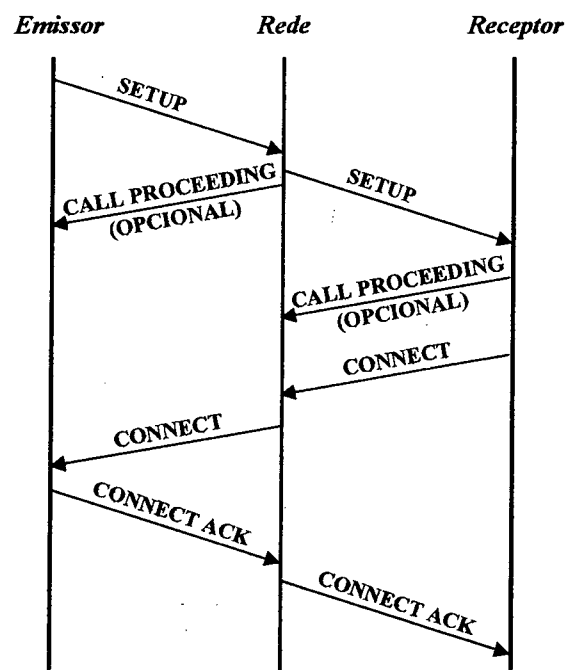


Fig. 5.12 – Troca de Mensagens de Sinalização

A figura 5.12 ilustra a troca de mensagens de sinalização durante a fase de estabelecimento de uma conexão ATM. De acordo com a figura, o emissor inicia o processo após o envio ao receptor, por meio de uma conexão chamada VCC de sinalização, de uma mensagem SETUP.

Após um determinado período de tempo, normalmente 4 segundos, o emissor deve receber uma resposta sobre o possível estabelecimento da conexão. Se a conexão, e conseqüentemente seus requisitos, puderem ser atendidos por todos os componentes de rede envolvidos, inclusive o receptor, a resposta enviada por este último é uma mensagem CONNECT.

A partir deste momento a conexão pode ser estabelecida, uma vez que todos os componentes de rede têm conhecimento sobre quais recursos devem ser alocados e quais os exatos endereços (VPI/VCI) devem ser utilizados.

Se pelo menos um elemento de rede não puder alocar os recursos solicitados, a conexão é rejeitada e, neste caso, cabe ao usuário tentar imediatamente o estabelecimento de uma outra conexão, agora com nível de QoS mais baixo, ou esperar até que a rede tenha recursos suficientes para atender aos requisitos originalmente solicitados.

Como dito anteriormente, durante o estabelecimento de uma conexão, as características de tráfego relevantes à mesma devem ser especificadas. Desta forma, os recursos da rede podem ser controlados e, por conseguinte, as características associadas às conexões em andamento podem ser preservadas. Este fato torna o processo de sinalização o ponto de partida para a implementação de uma gama de mecanismos, tais como:

- **Controle de Admissão de Conexões:** é responsável pelo controle dos recursos da rede. Sua finalidade é determinar se o estabelecimento de uma nova conexão pode ser executado, sem que para isso a QoS atribuída às conexões já em andamento seja prejudicada. Desta maneira, a QoS associada às conexões é garantida e a probabilidade de congestionamentos é minimizada;
- **Policimento da Conexão:** é responsável pelo monitoramento e pelo controle do tráfego gerado pelo emissor no momento do estabelecimento da conexão, a fim de mantê-lo de acordo com o que foi negociado e, conseqüentemente, proteger os recursos da rede contra má utilização e também evitar congestionamentos;
- **Controle de Congestionamentos:** caso a rede encontre-se congestionada, este mecanismo permite que a mesma retorne a um estado aceitável.

### 5.10. MAPEAMENTO DE PACOTES DE TRANSPORTE MPEG-2 EM CÉLULAS ATM

A especificação “*Audiovisual Multimedia Services: Video on Demand, Specification 1.1*” do ATM Forum, define uma série de aspectos relacionados à transmissão de dados audiovisuais, através de redes de banda larga baseadas na Tecnologia ATM, inclusive como os pacotes de fluxo de transporte MPEG-2 devem ser transmitidos através da Rede ATM.

Segundo esta especificação, os pacotes de fluxo de transporte MPEG-2 (SPTS - *Single Program Transport Stream*) devem ser mapeados na Camada de Adaptação ATM do tipo 5 (AAL5), com a Subcamada de Convergência Específica de Serviço de valor nulo. Portanto, um ou N Pacotes de Fluxos de Transporte MPEG-2 devem ser mapeados em uma AAL5-SDU [ATMForum, 1997].

Para Circuitos Virtuais Comutados (SVCs), o valor de N é estabelecido via sinalização ATM, utilizando-se  $N = \text{AAL5 CPCS-SDU}/188$ . Porém, para Circuitos Virtuais Permanentes N recebe 2 como valor padrão e logo, o tamanho máximo do CPCS-SDU é igual a 376 bytes. Com a finalidade de manter um determinado nível de interoperabilidade, todos os equipamentos baseados nesta especificação devem suportar os valores de  $N = 2$  e tamanho máximo do CPCS-SDU = 376, ver figura abaixo.

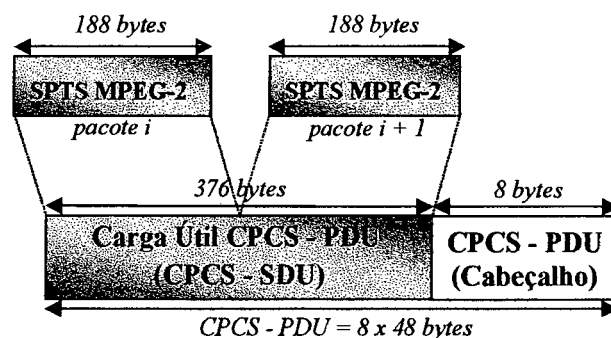


Fig.5.13 – Mapeamento de Pacotes de Fluxo de Transporte MPEG-2 em Células ATM

Considerando que uma CPCS-PDU é constituída de 48 bytes de carga útil e que a última CPCS-PDU deve comportar 8 bytes referentes ao cabeçalho, pode-se concluir que uma AAL5 CPCS-SDU de 376 bytes (dois pacotes de transporte), juntamente com o cabeçalho, requerem 384 bytes e podem ser mapeados exatamente em 8 células ATM com campo PAD de valor zero. Logo, uma AAL5 PDU deve conter dois pacotes de

transporte, a menos que ela contenha o último pacote de transporte do fluxo de transporte (SPTS).

### **5.11. CONSIDERAÇÕES FINAIS**

Uma vez que a finalidade desta dissertação é apresentar um modelo para a transmissão de fluxos MPEG-2 com QoS, a Tecnologia ATM foi utilizada como tecnologia de rede para a implementação deste trabalho, visto que a mesma oferece funcionalidades para o gerenciamento de QoS. Desta forma, este capítulo apresentou os principais conceitos que envolvem esta Tecnologia, principalmente, aqueles que envolvem QoS e mapeamento de pacotes MPEG-2 em células ATM. O próximo capítulo destina-se a descrição do sistema de gerenciamento de bando de dados multimídia e da arquitetura de servidor de armazenamento multimídia.

## CAPÍTULO VI

### SISTEMA DE GERENCIAMENTO DE BANCO DE DADOS MULTIMÍDIA E SERVIDORES DE ARMAZENAMENTO MULTIMÍDIA

*Visando suprir a carência de gerenciamento, processamento e armazenamento de informações multimídia por parte dos Sistemas de Gerenciamento de Banco de Dados e Armazenamento de arquivos tradicionais, existe hoje um amplo interesse no projeto e desenvolvimento de Sistemas de Gerenciamento de Banco de Dados Multimídia e Servidores de Armazenamento Multimídia. Neste capítulo, as principais características e requisitos destes elementos serão apresentados.*

#### 6.1. INTRODUÇÃO

Com o surgimento e, conseqüentemente, com o aumento da utilização dos dados multimídia, tais como áudios e vídeos, os sistemas convencionais, ou seja, aqueles em que o sistema de gerenciamento de banco de dados e a arquitetura de servidor de armazenamento baseiam-se na manipulação de dados alfanuméricos, precisaram atualizar-se, a fim de suportar também a manipulação de dados multimídia e, então, satisfazer os novos requisitos impostos por este tipo de dado.

Os conceitos aplicados por estes novos sistemas, os quais podem ser denominados Sistemas Multimídia, passaram a ser estudados e, posteriormente, utilizados no desenvolvimento e na implementação de uma variedade de aplicações, tais como aplicações de vídeo sob demanda e bibliotecas virtuais.

O projeto dos sistemas multimídia difere dos sistemas convencionais, principalmente no que se refere ao grande espaço de armazenamento e à alta taxa de transmissão de dados impostos pelas mídias digitais. Além disso, diferentemente dos dados alfanuméricos, os dados multimídia requerem que as operações efetuadas pelo servidor multimídia sejam realizadas com base em determinadas garantias de Qualidade de Serviço (QoS), para que o resultado apropriado destas operações seja obtido.

As principais questões acerca de um sistema de gerenciamento de banco de dados multimídia e uma arquitetura de servidor de armazenamento multimídia serão discutidas nas seções a seguir. É importante observar que, dependendo da aplicação que esta sendo implementada, somente algumas destas características deverão ser consideradas. Desta



forma, o tipo de aplicação dita quais as características do sistema de gerenciamento de banco de dados e da arquitetura de servidor.

## 6.2. SISTEMA DE GERENCIAMENTO DE BANCO DE DADOS MULTIMÍDIA

O sistema de gerenciamento de banco de dados de um sistema multimídia distribuído fornece facilidades para a criação e o armazenamento apropriado de objetos multimídia no banco de dados, bem como acesso concorrente a esses objetos e seus atributos (metadados), os quais deverão ser manipulados em um ambiente distribuído transparente ao usuário e à aplicação.

### 6.2.1. INFORMAÇÕES MULTIMÍDIA E METADADOS

São dois os tipos de informações manipuladas por um sistema de gerenciamento de banco de dados multimídia: *informações multimídia*, ou seja, os objetos multimídia (vídeo, áudio, textos, etc.) armazenados e acessados pelo usuário/aplicação, e *metadados*, também chamados *informações de controle*, os quais descrevem o conteúdo, a estrutura e a semântica dos objetos multimídia aos quais estão associados.

Cabe a pessoa responsável pela manutenção do banco de dados relacionar (indexar), de forma automática ou manual, cada objeto multimídia aos seus respectivos metadados, a fim de tornar o acesso a estes objetos o mais eficiente possível.

Segundo [Prabhakaran, 1997], os metadados necessários à manipulação dos objetos multimídia podem ser classificados em três categorias, a saber:

- *Metadados dependente do conteúdo*: a criação destas informações depende do conteúdo dos objetos multimídia. Por exemplo, no caso de uma imagem de uma pessoa, esta informação pode representar uma característica facial (tipo do nariz e cor do cabelo) e no caso de um vídeo, a mesma pode descrever a movimentação da câmera (panorama, zoom ou inclinação) e histograma de cor;
- *Metadados independente do conteúdo*: apesar de relacionadas ao conteúdo dos objetos multimídia, a criação destas informações não depende do conteúdo dos mesmos. Estas informações podem descrever, por exemplo, data de produção do

objeto, nome do fotógrafo, no caso de uma imagem e nome do diretor, no caso de um vídeo;

- ***Metadados que descrevem o conteúdo:*** apesar de estarem diretamente ligadas ao conteúdo dos objetos multimídia, estas informações, bem como algumas informações do tipo anteriormente descrito, não podem ser geradas automaticamente a partir do conteúdo dos mesmos, uma vez que descrevem características dos objetos através do ponto de vista dos usuários ou aplicações e, portanto, devem ser criadas com o suporte destes usuários ou de ferramentas que ofereçam processos cognitivos. Para imagens, são exemplos deste tipo de informações, as expressões faciais (raiva, felicidade) e, para vídeos, a descrição dos objetos (cores, textura) e das ações executadas em cena.

A descrição do conteúdo de um objeto, ou seja, a criação de metadados para um determinado objeto multimídia é subjetiva e depende do tipo de mídia e de quais informações são relevantes à aplicação em questão. Conseqüentemente, dependendo do tipo de aplicação, nem todas as categorias descritas acima precisarão ser criadas.

Para efeito de manutenção do banco de dados, é importante que a criação de metadados seja feita automaticamente, porém, como visto anteriormente, muito freqüentemente a interferência humana será necessária.

A criação de metadados é executada através da implementação de funções que extraem as características relevantes dos objetos multimídia. As técnicas aplicadas por estas funções dependem do tipo do metadado e baseiam-se nas informações intramídia e intermídia associadas aos objetos multimídia [Prabhakaran, 1997]. As informações intramídia tratam da interpretação das informações dentro da própria mídia, enquanto que as informações intermídia tratam da interpretação da relação existente entre várias mídias.

### **6.3. ARQUITETURA DE SERVIDOR DE ARMAZENAMENTO MULTIMÍDIA**

O servidor de armazenamento multimídia, ou somente servidor multimídia, assim como um servidor convencional, deve proporcionar o armazenamento, o gerenciamento e a

recuperação de informações. Contudo, as informações manipuladas por este servidor possuem estrutura semântica mais complexa e ocupam espaço de armazenamento extraordinariamente maior, conseqüentemente, as operações executadas pelo mesmo são mais difíceis de serem colocadas em prática.

Para que o funcionamento adequado de todo o sistema seja alcançado, o servidor multimídia deve proporcionar um bom desempenho fim-a-fim na recuperação das informações multimídia e para tanto o mesmo deve satisfazer uma série de requisitos [Lu, 1996; Sandstã, 1997]:

- O servidor deve utilizar de maneira otimizada os recursos computacionais disponíveis no sistema;
- O servidor deve permitir alta largura de banda de transmissão e prover grande espaço de armazenamento, uma vez que as informações multimídia ocupam espaço muito grande (ver tabela 6.1);

ESPECIFICAÇÕES	REQUISITOS DE LARGURA DE BANDA
Áudio com qualidade de telefone	64 Kbits/s
Áudio com qualidade de CD (2 canais, amostras de 16 bits a 44,1 KHz)	1,4 Mbits/s
Vídeo com qualidade de NTSC <sup>6</sup> (604 x 480 pixels, 8 bits/pixel)	8,7 Mbytes/s
Vídeo com qualidade de HDTV <sup>7</sup> (1024 x 2000 pixels/quadro, 24 bits/pixel)	351 Mbytes/s

*Tab. 6.1 – Requisitos de Espaço de Armazenamento para Dados Multimídia Digitais não Compactados*

- A fim de que o sincronismo das mídias que compõem as informações multimídia seja mantido, o servidor deve possuir capacidade de armazenamento e recuperação em tempo real, pois as mídias contínuas são ditas dependentes do tempo;
- O servidor deve suportar recuperação, entrega e apresentação das informações multimídia de maneira contínua, a fim de que os requisitos de tempo real impostos pelas mídias contínuas sejam satisfeitos;

<sup>6</sup> National Television Standards Committee – padrão de sinal de televisão desenvolvido pela FCC, que trabalha a uma taxa de 625 linhas de resolução, transmitidas em 30 quadros entrelaçados por segundo. É um composto dos sinais RGB, incluindo frequência modulada para áudio e sinal para estéreo.

<sup>7</sup> High Definition TV - refere-se à tecnologia de televisão com melhor qualidade do que a TV de transmissão convencional.

- Para tornar o sistema econômico, o servidor deve permitir acessos concorrentes, ou seja, o mesmo deve atender a vários usuários ao mesmo tempo. Logo, a arquitetura, os mecanismos e os algoritmos implementados no servidor devem ser escaláveis e capazes de suportar diferentes características de tráfego;
- O servidor deve permitir que o usuário execute não só operações convencionais, tais como abrir, ler e escrever, mas também operações de interação, como parada, avanço e retrocesso rápidos, respondendo às mesmas em um intervalo de tempo aceitável;
- Se o servidor não for implementado especificamente para um único tipo de aplicação, o mesmo deve ser capaz de armazenar e transmitir o mesmo conteúdo codificado por diferentes padrões de compressão e, portanto, com diferentes resoluções e a diferentes taxas de transmissão;
- Como as informações multimídia são compostas por mídias diferentes, é necessário o servidor possibilite o armazenamento de metadados e, conseqüentemente, o acesso eficiente destas informações.

### **6.3.1. GERENCIAMENTO DOS REQUISITOS DE ESPAÇO DE ARMAZENAMENTO**

Como visto na seção 6.3, os dispositivos de armazenamento ou a combinação destes dispositivos devem prover alta capacidade de armazenamento e largura de banda de transmissão. É importante observar que a largura de banda de transmissão mencionada é aquela utilizada pelos usuários e não a máxima largura de banda de transmissão dos dispositivos de armazenamento.

Para que estes requisitos sejam satisfeitos e para que o objetivo dos servidores multimídia, de atender aos requisitos temporais e espaciais dos fluxos multimídia e ao mesmo tempo servir simultaneamente ao maior número de usuários possível, seja também obedecido, os servidores devem implementar, com base nas características da aplicação, adequados algoritmos de escalamento de disco, controle de admissão e métodos de posicionamento de blocos de dados no disco.

Esta seção destina-se à descrição destes mecanismos e algoritmos e suas variações, além da descrição de alguns mecanismos de configuração de discos de armazenamento.

### 6.3.2. MECANISMOS DE CONFIGURAÇÃO DE DISCOS DE ARMAZENAMENTO

Os mecanismos de configuração de discos determinam o grau de escalabilidade do sistema, em termos de adição de novos dispositivos e distribuição da carga da rede. Quando necessária, esta adição deve ser executada em tempo real, ou seja, enquanto os dispositivos que já fazem parte do sistema estão em atividade, e a carga da rede imposta pelas várias solicitações enviadas pelos usuários deve ser eventualmente distribuída entre o maior número possível de dispositivos de armazenamento.

Existem vários mecanismos de configuração de discos, os quais serão apresentados a seguir. Cada mecanismo apresenta vantagens e desvantagens em relação à implementação de determinadas aplicações. Cabe a pessoa responsável pelo gerenciamento da rede, avaliar qual mecanismo responderá melhor aos objetivos da aplicação e apresentará os melhores resultados em termos de escalabilidade.

- **VETORES DE DISCOS**

Utilizar um único disco para o armazenamento de informações multimídia limita o número de potenciais acessos concorrentes a estas informações, devido à largura de banda do disco. Em contra partida, aumentar a largura de banda através da replicação destas informações em vários discos implica em aumento do *overhead*.

Uma outra alternativa é distribuir as informações multimídia em vários discos, de forma que a taxa de transferência dos dados cresça efetivamente de acordo com o número de discos utilizados. Esta técnica tornou-se popular com o aparecimento da arquitetura RAID (RAID - *Redundant Arrays of Inexpensive Disks*). Esta arquitetura é constituída de um vetor de discos, que consiste de uma estrutura de discos rígidos, cuja função é permitir que o computador opere em altas velocidades, no que se refere ao uso dos discos, e ao mesmo tempo garanta rápida recuperação de perda destes.

Esta estrutura melhora principalmente a capacidade de armazenamento e a taxa de transmissão, pois os dados de um arquivo são distribuídos em diversos discos, fazendo com que a leitura seja feita paralelamente, ou seja, o setor físico 1 de cada disco é acessado em paralelo como um grande setor lógico 1 e assim por diante [Gemmell, 1994].

Os RAIDs proporcionam, se comparados a discos maiores e mais caros, melhor desempenho, confiabilidade e escalabilidade a um custo menor, em consequência podem ser utilizados em muitos sistemas de computadores e como dispositivos de armazenamento de dados multimídia.

Existem vários níveis (organizações) diferentes de RAID. A escolha de um ou outro dependerá das características da aplicação. Para aplicações de vídeo, os níveis mais adequados são RAID 1 e RAID 3 [Cecilio, 1996]. A figura 6.1 ilustra o RAID nível 1, cuja função é copiar os dados em discos diferentes, a fim de possibilitar a recuperação dos mesmos diante da falha de um dos discos.

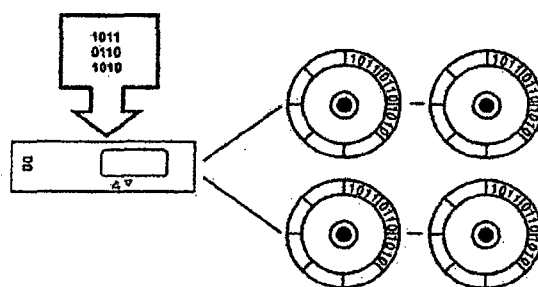


Fig. 6.1 - RAID Nível 1

#### • HIERARQUIAS DE ARMAZENAMENTO

Sabendo que os discos magnéticos, mesmo sendo a melhor opção para o armazenamento de informações multimídia, devido ao baixo tempo de acesso, alta taxa de transferência e capacidade de acesso aleatório, são também a alternativa mais cara e visando simular a alta capacidade dos discos óticos e das fitas magnéticas, uma hierarquia de armazenamento composta de vários tipos de dispositivos de armazenamento pode ser implementada, para que servidores de baixo custo e de grande escala possam ser desenvolvidos.

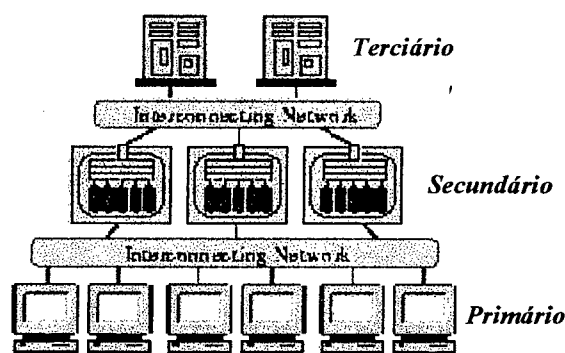
Uma das possibilidades de construção de uma hierarquia de armazenamento é utilizar fitas magnéticas (o que implica em acesso seqüencial) e discos óticos, para o armazenamento de arquivos, e discos magnéticos, somente para o armazenamento do início dos segmentos de arquivos multimídia. Desta maneira, a latência<sup>8</sup> de inicialização é reduzida e transições uniformes na apresentação podem ser obtidas. A desvantagem

<sup>8</sup> Tempo que a cabeça de leitura de um disco terá que esperar, depois de posicionada na trilha do disco, até que o setor que está sendo procurado passe sob ela e seja lido.

desta abordagem é que, devido à baixa taxa de transmissão dos discos óticos e fitas magnéticas, o acesso simultâneo a múltiplos fluxos não é suportado.

Os conceitos acima podem ser estendidos para uma hierarquia de armazenamento distribuída, a qual constitui-se de um número de níveis de armazenamento (Fig. 6.2), onde cada nível é composto de dispositivos de armazenamento diferentes.

O nível de armazenamento primário (discos magnéticos e memória), normalmente, representa o espaço de armazenamento disponível no equipamento do usuário e é responsável pela interação em tempo real com o mesmo. O nível secundário (vetores de discos) é responsável pelo armazenamento dos dados mais frequentemente acessados e, finalmente, o nível terciário (fita magnética) é responsável pelo armazenamento de grande quantidade de dados, o mesmo não é escalado para a apresentação de dados em tempo real, os dados raramente solicitados também são armazenados neste nível.



*Fig. 6.2 – Hierarquia de Armazenamento Distribuída*

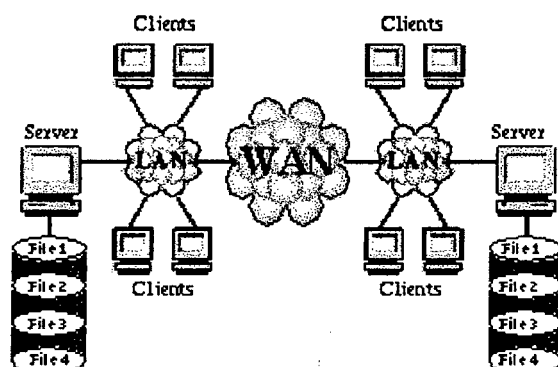
A vantagem de uma hierarquia de armazenamento distribuída é a escalabilidade oferecida em termos de compartilhamento da carga do sistema e a facilidade atribuída à replicação dos dados mais frequentemente acessados. Além disso, na presença de sobrecarga do sistema, novos componentes podem ser adicionados à hierarquia.

- **VETORES DE SERVIDORES**

Uma questão importante em sistemas multimídia é a QoS, a qual pode ser degradada se o desempenho do sistema como um todo for insuficiente. Quando um único servidor é

responsável por atender simultaneamente a fluxos diferentes de mídias de vários usuários, o sistema fica sujeito às limitações deste servidor, que pode suportar ou não esta carga. Com o intuito de manter a QoS, ao invés de um único servidor, vários servidores podem ser utilizados formando o que se conhece por *Vetor de Servidores*.

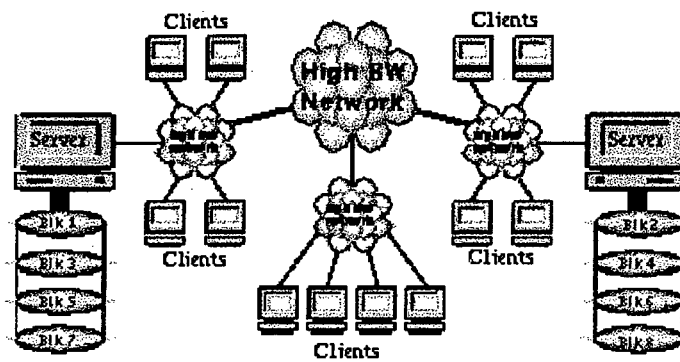
Para a implementação de um vetor de servidores, os arquivos multimídia podem ou não ser particionados entre os servidores que compõem o vetor. Os arquivos não-particionados são distribuídos igualmente entre todos os servidores, os quais, conseqüentemente, possuem exatamente o mesmo conteúdo. Desta forma, o usuário somente necessita conectar-se ao servidor mais próximo ou menos sobrecarregado, considerando também a carga da rede, ver figura 6.3.



*Fig. 6.3 – Vetor de Servidores sem Particionamento de Dados*

No outro caso, ver figura 6.4, os arquivos são particionados em unidades de tamanhos iguais e distribuídos entre alguns servidores do vetor, os quais tornam-se dependentes, a fim de que durante a apresentação todas as unidades estejam disponíveis. Para que o usuário possa recuperar uma informação multimídia, o mesmo deve interagir com todos os servidores envolvidos, dividindo assim a carga ocasionada pela solicitação.





*Fig. 6.4 – Vetor de Servidores com Particionamento de Dados*

Os dois últimos casos citados são eficientes no que se refere ao compartilhamento da carga da rede entre todos os dispositivos de armazenamento, porém, deixam a desejar no que se refere à adição de novos dispositivos.

Na primeira situação, como todos os servidores alcançam sua capacidade mínima de armazenamento ao mesmo tempo, a adição de um novo componente somente aliviará a sobrecarga do sistema, pois este herdará o mesmo conteúdo existente nos outros servidores. Na Segunda situação, sempre que um novo dispositivo precisar ser adicionado ao sistema, todos os arquivos precisarão ser novamente particionados, para que os mesmos fiquem igualmente distribuídos entre os novos e os antigos dispositivos de armazenamento.

### **6.3.3. MECANISMOS DE POSICIONAMENTO DE BLOCOS DE DADOS NO DISCO**

Antes de armazenar qualquer arquivo, o servidor de armazenamento deve dividi-lo em blocos. Estes por sua vez podem ocupar vários blocos dentro do disco físico em que foram armazenados.

Para que estes blocos sejam posicionados adequadamente no disco, basicamente, dois métodos podem ser implementados: posicionamento contínuo e posicionamento distribuído. Algumas variações destes métodos, como é apresentado a seguir, foram propostas para que houvesse melhora no desempenho do servidor.

### • POSICIONAMENTO CONTÍNUO DE DADOS

Segundo este método, não há intervalos dentro de um arquivo, pois os blocos de dados que compõem o mesmo são continuamente armazenados no disco. Desta maneira, para a leitura de um determinado arquivo, somente um tempo de busca é gasto para o posicionamento do cabeçote<sup>10</sup> do disco no primeiro bloco de dados deste arquivo.

Embora seja um método aparentemente de fácil implementação, existem duas limitações no que se refere à fragmentação do disco durante o armazenamento e a remoção de blocos de dados. Primeiramente, para que a continuidade dos blocos de dados em um arquivo seja mantida, grandes cópias devem ser executadas durante a execução destas operações, uma vez que os blocos encontrados após o ponto de inserção e remoção devem caminhar para trás e para frente, respectivamente, a fim de que o espaço deixado pelo bloco seja preenchido.

A segunda limitação refere-se ao espaço deixado no disco durante a remoção de um arquivo. Se por ventura o arquivo que for armazenado no lugar do arquivo removido for de tamanho menor, o que pode acontecer com frequência, visto a baixa probabilidade de se encontrar um arquivo de mesmo tamanho, o disco ficará fragmentado e, logo, isto ocasionará um desperdício de largura de banda.

Este método torna-se então, mais apropriado para aplicações onde a recuperação de dados ocorre mais frequentemente do que o armazenamento, não havendo edição de dados. Neste contexto enquadram-se os aplicativos de mídia sob demanda.

### • POSICIONAMENTO DISTRIBUÍDO DE DADOS

Este método baseia-se no posicionamento distribuído dos blocos de dados de um arquivo em torno do dispositivo em que estão armazenados.

As limitações quanto a este método referem-se à implementação de mecanismos necessária ao rastreamento dos blocos de um arquivo e ao tempo de busca gasto, para que cada bloco seja lido, durante a recuperação de vários blocos de um mesmo arquivo. Até mesmo quando a quantidade de dados a ser recuperada for pequena, existe a probabilidade de que a metade seja armazenada em um bloco e a outra metade no bloco

---

<sup>10</sup> Dispositivo que lê, grava e apaga informações em um dispositivo de armazenamento

seguinte, neste caso, o tempo de busca refere-se ao tempo de busca intra-arquivo, ou seja, o tempo gasto para a passagem do cabeçote de leitura de um bloco a outro.

No que diz respeito à primeira limitação, um arquivo de mapeamento deve ser mantido, a fim de possibilitar a localização dos próximos blocos de um arquivo, a partir do bloco que está sendo acessado. Existem várias maneiras de se implementar este tipo de mecanismo, algumas delas encontram-se relacionadas abaixo:

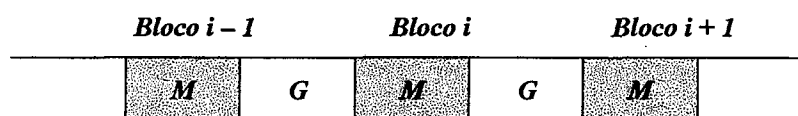
- ***Lista Vinculada:*** cada bloco contém um ponteiro para o bloco seguinte do arquivo. Um arquivo descritor, contendo a lista de blocos de um arquivo, necessita somente apontar para o primeiro bloco. A desvantagem deste método é que acessar aleatoriamente um bloco implica em acessar todos os blocos anteriores a este;
- ***Tabela de Alocação de Arquivos (FAT):*** esta tabela mapea e grava a localização de cada bloco armazenado no disco. Se mantida na memória principal, possibilita que o acesso aleatório seja feito rapidamente. Todavia, isto não ocorre com servidores de multimídia com volume de dados muito grande;
- ***Índice:*** cada arquivo é associado a um índice, o qual contém ponteiros para cada bloco do arquivo. Não há necessidade de que todos os índices de todos os arquivos sejam carregados na memória principal. Porém, a recuperação simultânea de vários arquivos pode fazer com que os índices ocupem um espaço de armazenamento muito grande, tornando o armazenamento na memória principal inviável;
- ***Solução Híbrida:*** baseia-se no armazenamento tanto de uma lista vinculada quanto de índices, onde a primeira é utilizada para apresentação em tempo real e a segunda para acesso aleatório. A desvantagem deste método é o maior espaço de armazenamento necessário.

Com relação à segunda limitação, uma maneira de reduzir a ocorrência de tempo intra-arquivo é fazer com que a quantidade de dados recuperada corresponda a um bloco de dados. Esta abordagem seria de fácil implementação em fluxos codificados a taxa de

bits constante, porém, muito difícil de ser implementada em fluxos codificados a taxa de bits variável.

- **POSICIONAMENTO RESTRITO DE DADOS**

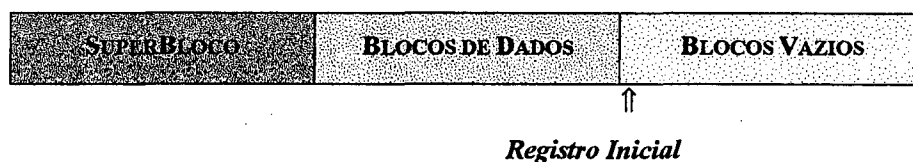
O posicionamento restrito é uma variação do método acima. Partindo do princípio que a taxa de leitura de um cabeçote de disco é mais alta do que a taxa de apresentação de um fluxo, este mecanismo utiliza-se do armazenamento de blocos de tamanhos limitados entre blocos de dados consecutivos, a fim de reduzir o tempo de busca intra-arquivo a um limite razoável. Na figura 6.5, *M* representa a quantidade de dados dentro de cada bloco e *G*, a quantidade de dados correspondente ao intervalo.



*Figura 6.5 – Posicionamento Restrito de Dados*

- **POSICIONAMENTO ESTRUTURADO DE REGISTRO DE DADOS**

Este método baseia-se na composição de vários blocos de dados, onde são armazenados os arquivos de dados, e um superbloco, que armazena informações de gerenciamento, tais como endereços de blocos e ponteiro para o registro inicial, ou seja, o endereço do primeiro bloco vazio disponível (Fig. 6.6).



*Registro Inicial*

*Fig. 6.6 – Posicionamento Estruturado de Registro de Dados*

Os blocos são armazenados continuamente, começando do *registro inicial*. Os intervalos deixados por blocos que acabaram de ser removidos não são preenchidos por novos blocos de dados, o que torna este tipo de posicionamento adequado para recuperação e edição de dados.

Se por acaso não houver edições de dados, as localidades temporais são mantidas, isto é, os dados que estiverem temporariamente próximos uns dos outros são armazenados em blocos contínuos no disco, possibilitando a redução do tempo de busca e que os requisitos de continuidade sejam satisfeitos. Para facilitar a sincronização intermídia, os dados multimídia, tais como vídeo e áudio associados, podem ser intercalados.

- **POSICIONAMENTO DE DADOS EM VETORES DE DISCOS**

No caso de vetores de discos serem utilizados como dispositivos de armazenamento, várias opções podem ser utilizadas para o armazenamento de arquivos. Um método simples consiste em armazenar todo um arquivo multimídia em um único disco e, conseqüentemente, limitar o número de acessos concorrentes, devido à taxa de transferência deste disco.

Por exemplo, no máximo 4 acessos concorrentes serão permitidos se a taxa de transferência de um disco for de 4 Mbytes/s e a taxa de apresentação de um fluxo for de 1 Mbyte/s, o que para arquivos muito solicitados é insuficiente. Uma solução é armazenar uma cópia do arquivo em mais de um disco, porém, com o agravante de que a implementação de discos adicionais implica em gastos extras.

Outro método, ao invés de distribuir uma cópia de um mesmo arquivo entre vários discos, distribui os blocos deste arquivo, favorecendo o acesso de vários usuários ao mesmo arquivo simultaneamente.

Existem duas outras variações para este mesmo método. A primeira organiza os blocos em grupos de acesso, isto é, todos os discos são sincronizados e os mesmos setores sobre todos os discos são acessados no mesmo instante.

A segunda armazena os blocos de um arquivo multimídia de maneira intercalada em um conjunto de discos, ou seja, blocos sucessivos de um mesmo arquivo são armazenados em discos diferentes não sincronizados e, desta maneira, os blocos são acessados independentemente. Contudo, para que o escalamento do disco seja facilitado, os blocos são armazenados de forma cíclica, de maneira a formar um anel lógico.

A combinação destes dois métodos pode ser utilizada na distribuição de arquivos entre um grande número de discos incorporados a um *cluster*<sup>11</sup> de rede de máquinas servidoras. Esta técnica possibilita a construção de servidores multimídia escaláveis que possam viabilizar a entrega de uma única cópia de um arquivo multimídia a um grande número de solicitações. Além disso, técnicas de redundância podem ser aplicadas nos arquivos multimídia para que a disponibilidade e a vazão sejam aumentadas.

#### **6.4. ALGORITMOS DE CONTROLE DE ADMISSÃO E ESCALAMENTO DE DISCOS**

Na maioria das ocasiões, em um dado momento, o servidor multimídia deve processar a recuperação de fluxos multimídia para vários usuários e, ainda, deve responder, em um intervalo de tempo aceitável, às operações de interação executadas pelos mesmos.

Para que a apresentação de um fluxo multimídia (por exemplo, uma seqüência de vídeo ou de áudio) tenha significado, é necessário que a taxa em que o mesmo está sendo apresentado seja contínua. Para tanto, o servidor multimídia deve garantir que o armazenamento e a apresentação de cada fluxo de mídia irá ocorrer em tempo real.

A apresentação contínua de um fluxo multimídia consiste em uma seqüência de tarefas executadas periodicamente, em determinado intervalo de tempo, isto é, consiste na recuperação de blocos de mídia do disco em tempos de apresentação predeterminados.

Na seção seguinte, serão apresentados os algoritmos de controle de admissão e de escalamento de disco, os quais, respectivamente, são utilizados para o controle do número de apresentações simultâneas e para que a continuidade da apresentação dos fluxos seja mantida.

##### **6.4.1. ALGORITMOS DE CONTROLE DE ADMISSÃO**

Os algoritmos de controle de admissão previnem a sobrecarga do sistema, tendo por finalidade a garantia de desempenho do mesmo, isto é, oferecer garantias de qualidade de serviço durante a entrega dos dados. Segundo [Lu, 1996], o controle de admissão em servidores de armazenamento deve basear-se no seguinte critério: “O total de largura de

---

<sup>11</sup> Sistema em servidor de dados, pelo qual dois ou mais equipamentos podem ter acesso ao mesmo banco de dados e compartilhar as mesmas unidades de disco.

banda proporcionado por todos os fluxos solicitados deve ser menor do que a taxa de transferência do disco”.

Os algoritmos de controle de admissão devem, dados alguns requisitos de desempenho em tempo real de cada usuário, ser utilizados pelos servidores multimídia de modo a determinar se um novo usuário pode ou não ser admitido, sem que os requisitos de desempenho dos usuários já atendidos sejam violados.

Um conceito muito importante no âmbito da computação é o de *deadline*, que consiste no tempo limite em que uma tarefa deve ser executada. Visto isto, pode-se supor que dentre os requisitos de desempenho de um usuário, está o de satisfazer todos os *deadlines* de tempo real das mídias contínuas. Desta forma, é possível analisar duas implicações impostas por este requisito:

- Para algumas aplicações multimídia a perda de alguns *deadlines* em alguns casos pode ser tolerada e, conseqüentemente, recompensada pela redução do custo de serviço;
- De modo a garantir que todos os *deadlines* serão satisfeitos, considerações devem ser feitas em relação à latência de busca e latência rotacional dos discos. Por conseguinte, um servidor multimídia deve ser capaz de acomodar usuários adicionais através da implementação de algoritmos de controle de admissão que explorem a variação estatística do tempo de acesso de blocos de mídia do disco.

No que diz respeito a garantias de que os requisitos de *deadlines* serão satisfeitos, alguns autores propuseram níveis de qualidade de serviço (QoS), dentre os quais [Gemmell, 1994] cita:

- **Determinístico:** há garantia de que todos os *deadlines* serão satisfeitos. Para este nível de serviço, o algoritmo de controle de admissão considera os piores ambientes, isto é, ambientes susceptíveis a falhas, na aceitação de novos usuários;
- **Estatístico:** há garantia de que os *deadlines* serão satisfeitos, porém, baseada em uma certa probabilidade. Por exemplo, o usuário pode solicitar um serviço que garanta a satisfação de 90% dos *deadlines* durante certo intervalo. Para fornecer

tais garantias, o algoritmo de controle de admissão deve considerar o comportamento estatístico do sistema enquanto novos usuários são admitidos;

- **Melhor Esforço:** não há garantias de que os *deadlines* serão satisfeitos. O servidor escala os acessos somente quando os usuários enquadrados no nível determinístico e estatístico deixarem tempo disponível.

Fornecer o nível estatístico de garantia de serviço é essencial devido à variação no tempo de busca e latência rotacional e, também, devido à variação nos requisitos de transmissão de dados dos fluxos de mídia compactados.

#### 6.4.2. ALGORITMOS DE ESCALAMENTO DE DISCO

Como já foi abordado, é função do servidor entregar os dados à rede na mesma taxa em que os mesmos serão apresentados ao usuário, além de prevenir a falta de alimentação dos dados e minimizar os requisitos de armazenamento, de maneira a reduzir o custo e longos atrasos de transmissão.

Como a taxa de transferência de um cabeçote de disco é maior do que a taxa de apresentação da maioria dos fluxos multimídia individuais, para que os recursos do servidor sejam utilizados eficientemente, o cabeçote de disco deve servir ao maior número possível de solicitações, o que pode ser conseguido através da implementação de algoritmos de escalamento de disco adequados.

As técnicas de escalamento de disco são necessárias para que a continuidade da apresentação dos dados multimídia seja mantida, apesar das operações de disco serem indeterminísticas, devido ao tempo de acesso randômico e ao compartilhamento com outras aplicações. A seguir serão apresentados os algoritmos de escalamento de disco mais comuns:

- **First-Come-First-Server (FCFS):** as solicitações são atendidas de acordo com a sua ordem de chegada. Desta maneira, a média de tempo de busca é alta, pois a localização e o movimento do cabeçote do disco são ignorados;



- **Shortest Seek Time First (SSTF):** tenta reduzir o tempo de busca atendendo às solicitações cuja localização encontra-se mais próxima ao cabeçote do disco, favorecendo os dados que estiverem localizadas no centro do mesmo. Quando o servidor estiver muito sobrecarregado, possivelmente, as solicitações de transferência dos dados localizados nas trilhas interior e exterior não poderão de maneira nenhuma ser atendidas;
- **Scan:** tenta minimizar o tempo de busca atendendo às solicitações de acordo com a movimentação do cabeçote do disco, ou seja, primeiro todas as solicitações em uma direção são atendidas e, então, o cabeçote do disco é revertido para que as solicitações localizadas na direção oposta sejam atendidas. Assim, uma solicitação pode ser atendida primeiro em uma direção e por último na direção oposta. A apresentação do fluxo deve aguardar até o final do primeiro ciclo<sup>1</sup>.

Estes algoritmos não podem ser diretamente implementados para o escalamento de servidores multimídia, pois não satisfazem os critérios de tempo real dos dados multimídia. Contudo, se o número de solicitações simultâneas for limitado e um tempo para cada solicitação for reservado, o algoritmo *Scan* é capaz de fornecer garantias de serviço para cada solicitação. A seguir serão apresentados os algoritmos de escalamento de disco em tempo real mais utilizados, alguns se apresentam como variações dos algoritmos acima:

- **EARLIEST DEADLINE FIRST - EDF**

Este é o algoritmo mais utilizado para o escalamento de disco em tempo real de tarefas com *deadline*. Os blocos de mídia com *deadline* mais próximos têm prioridade na recuperação. A desvantagem deste algoritmo é que o posicionamento dos blocos de dados é ignorado, implicando em um longo tempo de busca e latência rotacional e, portanto, em uma baixa utilização dos recursos do servidor.

---

<sup>1</sup> Devido à natureza periódica da apresentação, o servidor multimídia pode atender aos vários fluxos simultaneamente executando o processamento em ciclos, onde durante cada ciclo, o servidor recupera uma seqüência de blocos de mídia para cada fluxo.

Este algoritmo não utiliza toda a taxa de transferência do disco, pois muito tempo é desperdiçado para a movimentação do cabeçote do disco em torno do mesmo, sem que dados realmente úteis sejam lidos.

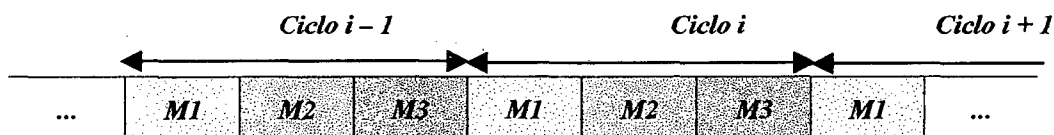
Devido ao EDF não impor um posicionamento de dados específico e o tempo de busca poder variar de 0 até o valor máximo durante o momento em que este estiver atendendo a duas solicitações, é muito difícil desenvolver um algoritmo de controle de admissão que possa simultaneamente utilizar de maneira eficiente os recursos do servidor e satisfazer todos os requisitos de continuidade dos fluxos.

- **ROUND-ROBIN**

Este é o mais simples de todos os algoritmos de escalamento de discos em tempo real. Segundo este algoritmo, a ordem em que os usuários são atendidos não varia de um ciclo para outro. Desta maneira, o intervalo de tempo entre dois serviços sucessivos para um mesmo fluxo é limitado pela duração do ciclo. Se a duração do ciclo for constante, então, o intervalo também será. Já que a ordem da recuperação dos dados é fixa, o posicionamento dos mesmos é importante na determinação do tempo de busca.

A desvantagem deste algoritmo é que ele não explora a posição relativa dos blocos de mídia recuperados durante um ciclo e, por esta razão é, algumas vezes, utilizado com mecanismos de posicionamento de dados restrito, a fim de que o tempo de busca seja mínimo e a continuidade de cada fluxo seja mantida.

Por exemplo, na figura 6.7, três fluxos podem ser recuperados continuamente. M1, M2 e M3 referem-se à quantidade de dados encontrada dentro de cada bloco correspondente aos três fluxos representados.



*Fig. 6.7 – Escalamento Round-Robin com Posicionamento de Dados Restrito*

- **SCAN-EARLIEST DEADLINE FIRST – SCAN-EDF**

O objetivo deste algoritmo é reduzir o tempo de busca médio do algoritmo EDF, combinando este ao algoritmo Scan. Segundo seu funcionamento, as solicitações com *deadline* mais próximos têm prioridade no atendimento.

Quando todos as solicitações tiverem o mesmo *deadline*, o algoritmo Scan-EDF reduz-se somente ao algoritmo Scan, ou seja, tais solicitações terão seus blocos acessados unicamente pelo algoritmo Scan. Contudo, quando as solicitações tiverem *deadline* diferentes, somente o algoritmo EDF é utilizado. Com o objetivo de tornar o algoritmo Scan-EDF mais eficaz, muitas técnicas têm sido propostas com o intuito de aumentar o número de pedidos com mesmo *deadline* [Lu, 1996].

- **GROUP SWEEPING SCHEDULING – GSS**

No GSS cada ciclo é dividido em grupos, para os quais cada usuário é designado. Os grupos são atendidos em uma ordem fixa, assim como no Round-Robin. Dentro de cada grupo, o algoritmo Scan é utilizado para atender aos fluxos referentes ao grupo, reduzindo assim a duração do ciclo, e o algoritmo Round-Robin é utilizado para a redução do intervalo entre serviços sucessivos. Desta maneira, o algoritmo GSS mantém um compromisso entre a duração do ciclo e o intervalo de serviços sucessivos.

## 6.5. CONFIGURAÇÃO DO SERVIDOR E CONEXÕES DE REDE

A arquitetura de um servidor de armazenamento multimídia deve ser escalável e prover alta largura de banda de acesso à rede. Esta seção apresenta, brevemente, como os dispositivos de armazenamento podem ser interconectados a um servidor multimídia e como este pode ser conectado à rede, sabendo que o mesmo deve suportar a entrega do maior número possível de fluxos a vários usuários simultaneamente.

Existem várias propostas para execução das duas operações acima, dentre estas as mais discutidas hoje em dia são as que utilizam a Tecnologia ATM como rede de comunicação (Fig. 6.8), onde entre cada dispositivo de armazenamento e o comutador existe um canal dedicado.

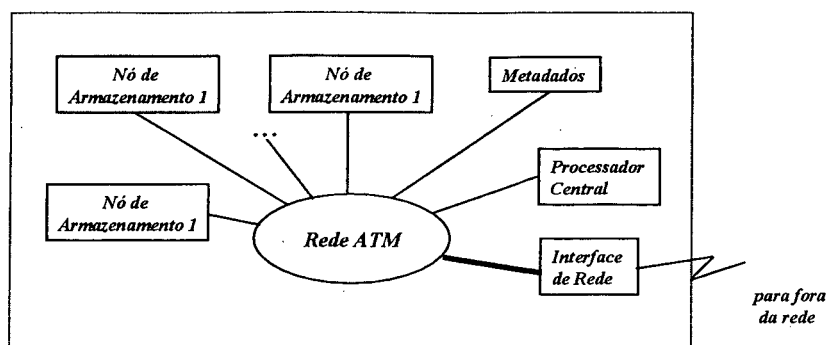
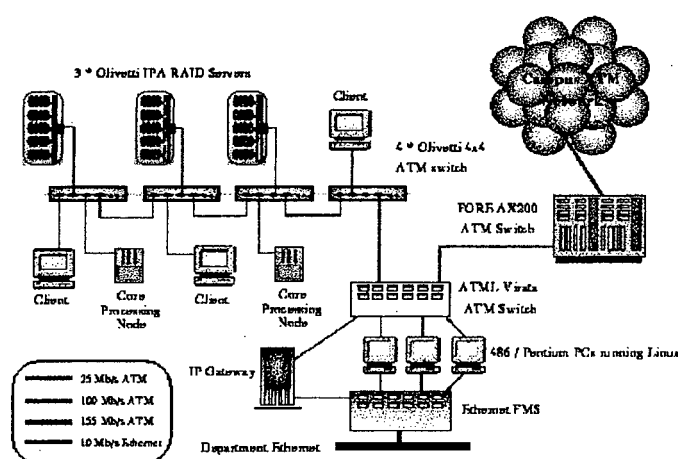


Fig. 6.8 – Arquitetura de um Servidor Multimídia

A arquitetura exposta na figura acima é escalável em algumas situações, visto que, quando necessário, mais dispositivos de armazenamento podem ser adicionados à rede. Porém, a interface de rede pode tornar-se o gargalo de desempenho, sendo preciso a utilização de várias destas interfaces e enlaces de acesso à rede. Os componentes que fazem parte desta arquitetura são:

- **Metadados:** contém informações (título, localização, relação entre as mídias, índices) em documentos multimídia armazenados nos nós de armazenamento;
- **Processador Central:** tem acesso aos metadados dos fluxos armazenados em outros dispositivos de armazenamento. É responsável por receber as solicitações dos usuários, executar testes de controle de admissão e estabelecer conexões para as solicitações aceitas, em favor dos nós de armazenamento;
- **Nós de Armazenamento:** onde são agrupados os dispositivos de armazenamento. Cada nó contém discos magnéticos, fitas magnéticas e/ou fibras ópticas e podem ser independentes uns dos outros ou ter alguma relação devido à divisão de dados. Cada nó também deve ser capaz de enviar dados diretamente à rede. Porém, o número de enlaces de acesso à rede necessários por um servidor dependerá da quantidade de largura de banda solicitada.

Outra proposta para arquitetura de servidor é a arquitetura de hierarquia chamada *SCAlable Multimedia Server (SCAM)*, ver figura 6.9. Esta arquitetura consiste de vários servidores distribuídos em uma WAN e logicamente organizados em vários níveis. Os arquivos multimídia são replicados ou distribuídos entre estes servidores, os quais são intimamente ou independentemente ligados, dependendo da utilização e disponibilidade de largura de banda dos enlaces que os conectam. Durante a operação, os servidores cooperam entre si para atender simultaneamente a vários usuários.



*Fig. 6.9 – Arquitetura SCAM*

Como a carga do tráfego é dividida entre vários servidores, não há possibilidade de um servidor específico tornar-se o gargalo de desempenho da rede, além disso, esta arquitetura é escalável quanto o número de usuário e arquivos multimídia.

## 6.6. CONSIDERAÇÕES FINAIS

Neste capítulo foram vistos os principais aspectos referentes aos Sistemas de Gerenciamento de Banco de Dados Multimídia e Servidores de Armazenamento Multimídia, uma vez faz parte do escopo do trabalho a manipulação de informações multimídia nestes componentes.

O próximo capítulo apresentará o modelo proposto neste trabalho de Dissertação cujo princípio é possibilitar, de maneira mais eficiente, que a transmissão dos dados

multimídia armazenados no Servidor Multimídia e gerenciados pelo Sistema de Banco de Dados Multimídia usufrua mais eficientemente dos conceitos de Qualidade de Serviço apresentados no Capítulo IV.

## CAPÍTULO VII

### TRANSMISSÃO DE FLUXOS MPEG-2 COM QoS A PARTIR DE SERVIDORES MULTIMÍDIA EM REDES ATM

*O propósito deste capítulo é apresentar o modelo proposto neste trabalho de Dissertação. O objetivo deste modelo é tornar a transmissão de fluxos MPEG-2, a partir de servidores de armazenamento multimídia, a mais adequada possível, tendo como base os requisitos de QoS especificados pelos usuários da aplicação multimídia e as funcionalidades atribuídas à Tecnologia ATM, no que se refere ao gerenciamento da QoS.*

#### 7.1. INTRODUÇÃO

Existe uma gama de aplicações multimídia distribuídas baseadas em servidor multimídia, tais como aplicações de vídeo sob demanda e bibliotecas digitais, cujo objetivo é explorar as funcionalidades e a alta qualidade proporcionadas pelo Padrão MPEG-2.

Para tais aplicações, a especificação dos requisitos de qualidade de serviço, ou seja, do nível de qualidade no qual o serviço deve ser recuperado, transmitido e apresentado é um ponto fundamental. Desta maneira, estas aplicações podem se beneficiar dos serviços que o sistema tem para oferecer, como por exemplo, facilidades para a negociação da QoS.

Como basicamente a qualidade de serviço, ao nível de usuário é avaliada através da qualidade perceptível do som e da imagem, existe, a princípio, a necessidade de que a especificação dos requisitos de QoS, que devem ser satisfeitos durante a apresentação do serviço, parta do usuário da aplicação.

Este capítulo apresenta um novo modelo para a transmissão de fluxos MPEG-2 com QoS. O modelo é voltado a aplicações multimídia distribuídas baseadas em servidor de armazenamento multimídia, que permitam interação entre o usuário e o conteúdo apresentado no que se refere à especificação do nível de QoS.

O modelo proposto pressupõe o uso da Tecnologia ATM, visando a utilização por parte do servidor multimídia das funcionalidades oferecidas por esta tecnologia no que diz respeito ao gerenciamento da QoS. Por conseguinte, é relevante para o desenvolvimento do modelo somente à definição da QoS ao nível da Tecnologia ATM.

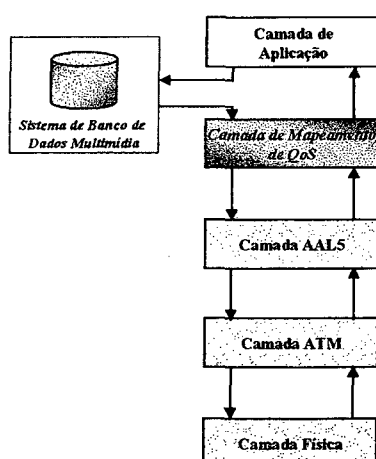
O desenvolvimento de um trabalho desta natureza foi impulsionado principalmente pela necessidade existente de um modelo capaz de satisfazer não somente os requisitos impostos pelas informações multimídia, neste caso, vídeos MPEG-2, mas também aqueles impostos pelo nível de qualidade de serviço especificado pelo usuário da aplicação multimídia.

Sendo assim, este modelo se dispõe a estudar ainda o modo como os processos de mapeamento e negociação dos parâmetros de QoS serão realizados dentro do sistema multimídia distribuído, considerando-se que os requisitos de QoS originalmente agregados a cada vídeo MPEG-2 armazenado no servidor podem ser modificados pelo usuário da aplicação, caso seja conveniente.

## 7.2. VISÃO GERAL DO MODELO PROPOSTO

O modelo proposto neste trabalho permite que o usuário selecione o vídeo MPEG-2 desejado e que, com base na qualidade máxima em que o mesmo encontra-se armazenado no servidor multimídia, especifique o nível de qualidade de serviço no qual o vídeo deve ser recuperado, transmitido e apresentado.

A visão geral dos principais elementos que compõem o modelo pode ser obtida a partir da figura abaixo. Como pode ser observado, o modelo baseia-se na definição da QoS apenas ao nível da Tecnologia ATM.



*Fig. 7.1 – Visão Geral do Modelo Proposto*



Na figura pode-se identificar o *Sistema de Banco de Dados Multimídia*, cujo objetivo é gerenciar o armazenamento, a recuperação e a apresentação dos vídeos MPEG-2 e suas correspondentes informações de controle, chamadas metadados, conceito futuramente apresentado.

E também, a *Camada de Mapeamento de QoS*, cuja finalidade é apresentar mecanismos para a execução do mapeamento dos parâmetros de QoS da aplicação em parâmetros de QoS de rede, para que posteriormente a negociação destes parâmetros em todos os componentes do sistema e, por fim, a transmissão do vídeo MPEG-2 sejam realizados adequadamente através da Tecnologia ATM.

Com o intuito de evitar interferências entre conexões de usuários diferentes e tornar o modelo, em seu estágio atual, o mais simples possível, considera-se que todas as conexões estabelecidas entre usuário e servidor caracterizam-se como conexões unicast ponto-a-ponto e que todo o processamento ocorre dentro de uma Rede Local ATM.

### **7.3. DESCRIÇÃO DO FUNCIONAMENTO DO MODELO PROPOSTO**

Para que as funções exercidas pelo modelo proposto nesta Dissertação fossem melhor expostas, as mesmas foram divididas em três etapas principais, as quais apresentam-se listadas a seguir:

#### **7.3.1. SELEÇÃO DO VÍDEO MPEG-2 E ESPECIFICAÇÃO DO NÍVEL DE QoS**

A interface gráfica aplicada deve prover de funcionalidades que permitam ao usuário o acesso à lista de vídeos MPEG-2 disponíveis no servidor de armazenamento multimídia e a modificação, caso seja conveniente, do nível de qualidade no qual o vídeo deve ser recuperado e apresentado.

Devido à impossibilidade de que melhoramentos sejam feitos na qualidade em que o vídeo foi originalmente digitalizado, o usuário é capaz somente de selecionar um nível de qualidade igual ou inferior àquele em que o vídeo encontra-se armazenado.

A especificação do nível de QoS é realizada na Camada de Aplicação utilizando-se os termos *Excelente*, *Bom*, *Regular* e *Ruim*, isto é, termos que qualificam, e não

quantificam, o nível de QoS no qual o usuário está disposto a visualizar o vídeo selecionado. A utilização de termos desta natureza pode ser justificada pela facilidade atribuída aos mesmos no que se refere ao uso de aplicações de vídeo por parte de usuários que se identifiquem com qualquer tipo de perfil.

A proporção de qualidade representada por cada um destes termos, por exemplo, *Excelente* igual a 100% da qualidade do vídeo e *Bom* igual a 80%, é fortemente dependente do tipo de aplicação de vídeo que está sendo implementada. Sendo assim, fica sob responsabilidade do projetista da aplicação a definição destes valores.

Os parâmetros de QoS correspondentes ao nível de QoS especificado pelo usuário são então calculados a partir da qualidade máxima no qual o vídeo foi digitalizado e armazenado no servidor multimídia. Desta forma, o Sistema de Banco de Dados Multimídia precisa ser acessado, a fim de que a qualidade máxima seja capturada e a qualidade solicitada seja neste caso avaliada.

### **7.3.2. MAPEAMENTO E NEGOCIAÇÃO DOS PARÂMETROS DE QoS**

Terminada a etapa anterior, o nível de QoS especificado pelo usuário na Camada de Aplicação deve ser transparentemente mapeado em todas as camadas do sistema. Para tanto, a Camada de Aplicação repassa à camada de sistema mais próxima, no caso a Camada de Mapeamento de QoS definida, os resultados obtidos durante o cálculo da QoS.

Como dito anteriormente, o nível de QoS especificado na Camada de Aplicação é expresso através de termos que qualificam, de maneira simples e acessível, o serviço que será oferecido. Contudo, a partir do momento que as informações são repassadas à Camada de Mapeamento de QoS a manipulação dos parâmetros de QoS relevantes passa a ser executada em termos de pacotes de dados.

Sendo assim, o nível de QoS especificado pelo usuário é mapeado em parâmetros de QoS compreensíveis pela rede, para posteriormente serem transformados em parâmetros de QoS expressos ao nível de Tecnologia ATM, ou seja, parâmetros de QoS ATM, visto que a entrega do vídeo MPEG-2 solicitado será executado via Rede ATM.

Desta maneira, estes parâmetros encontrar-se-ão na forma apropriada para serem negociados, através de uma conexão ATM de sinalização, entre os vários componentes do sistema envolvidos na possível transmissão do vídeo, inclusive usuário e servidor.

Segundo os princípios fixados pela Tecnologia ATM, se todos os componentes concordarem em alocar os recursos necessários à transmissão do vídeo MPEG-2 solicitado, uma nova conexão ATM é estabelecida, os parâmetros de QoS passam então a fazer parte do contrato de tráfego ATM e a transmissão pode ser iniciada, seguindo o esquema exibido na figura abaixo.



*Fig. 7.2 – Transmissão do Fluxo MPEG-2 em Redes ATM*

Vale ressaltar que os parâmetros contidos no contrato de tráfego foram obtidos com base no cálculo executado sobre o nível de qualidade em que o vídeo MPEG-2 encontra-se armazenado no Banco de Dados Multimídia. O mesmo deve ser respeitado por todos os componentes do sistema durante todo o período em que a conexão permanecer ativa.

Se por ventura um único componente do sistema não puder alocar os recursos necessários, o usuário pode imediatamente tentar o estabelecimento de uma nova conexão, porém, especificando um nível de QoS mais baixo, ou pode esperar até que o sistema tenha recursos suficientes para aceitar o nível de QoS originalmente especificado.

### **7.3.3. RECUPERAÇÃO DO VÍDEO MPEG-2 TRANSMITIDO**

Devido à alta escalabilidade proporcionada pelo vídeo MPEG-2, ver seção 3.3.6, é possível que o usuário modifique a qualidade máxima associada ao vídeo já armazenado no servidor, a uma qualidade inferior.

Desta maneira, de acordo com os parâmetros de QoS especificados no contrato de tráfego ATM, somente uma ou mais camadas do vídeo MPEG-2 precisarão ser mapeadas em células ATM e transmitidas continuamente em tempo-real através da conexão ATM que acabou de ser estabelecida.

Logo, é fundamental que o servidor possua também um elemento que possibilite a seleção do número de camadas do vídeo MPEG-2 que precisam ser transmitidas. Este elemento foi denominado *Escalador de Mídia*, ver figura 7.2.

À medida que o fluxo MPEG-2 for chegando ao usuário, o mesmo vai sendo armazenado temporariamente, a fim de que a relação temporal e espacial entre as unidades de dados enviadas seja mantida. Posteriormente, o fluxo é decodificado e apresentado ao usuário através de seu dispositivo de saída.

#### 7.4. GERENCIAMENTO DE VÍDEOS MPEG-2

Para aplicações de vídeo, a maneira como a recuperação e a entrega de determinado vídeo são executadas é ditada pelo tipo de aplicação. Por exemplo, algumas aplicações baseiam-se na recuperação de uma só vez de todos os segmentos de um vídeo específico, outras se baseiam somente na identificação e recuperação de determinados segmentos deste vídeo.

O esquema definido e apresentado nesta seção visa o acesso de dados de vídeo MPEG-2 e suas características, isto é, seus correspondentes metadados, a fim de que, fundamentada no nível de QoS especificado pelo usuário na Camada de Aplicação, a qualidade na qual o vídeo será recuperado, transmitido e apresentado seja calculada.

Este esquema tem como objetivo prover facilidades para que os usuários possam, de maneira efetiva, acessar o conteúdo armazenado no servidor multimídia simplesmente especificando o vídeo que desejam assistir e a qualidade na qual a apresentação do mesmo deve ser realizada.

Desta forma, a tarefa do Sistema de Banco de Dados Multimídia é gerenciar a recuperação do servidor multimídia de todos os blocos de dados associados ao vídeo selecionado e transmitir os mesmos adequadamente, isto é, satisfazendo todos os

requisitos temporais e espaciais necessários, para que a apresentação no dispositivo de saída do usuário seja efetuada de acordo com a qualidade indicada.

Não cabe ao trabalho especificar as técnicas utilizadas para o armazenamento e a recuperação destes vídeos e suas informações de controle, uma vez que a execução destas operações varia de acordo com a configuração do servidor multimídia e própria localização do mesmo na rede. Portanto, parte-se do princípio de que estes dados foram previamente e devidamente armazenados no servidor multimídia.

#### 7.4.1. ESQUEMA DE GERENCIAMENTO DE VÍDEOS MPEG-2

O esquema de gerenciamento de vídeos MPEG-2 definido neste trabalho baseia-se nos modelos abordados por [Sheth, 1998] e [Elmasri, 2000], o mesmo encontra-se ilustrado na figura 7.3.

De acordo com o esquema, o *Sistema de Gerenciamento de Banco de Dados Multimídia* e a *Arquitetura de Servidor de Armazenamento Multimídia* fazem parte de um sistema maior denominado *Sistema de Banco de Dados Multimídia*, cuja função é gerenciar, de uma maneira geral, o funcionamento dos elementos que lhe compõe. Cada unidade que faz parte do esquema possui uma função específica dentro do Sistema de Banco de Dados Multimídia, as quais serão listadas a seguir.

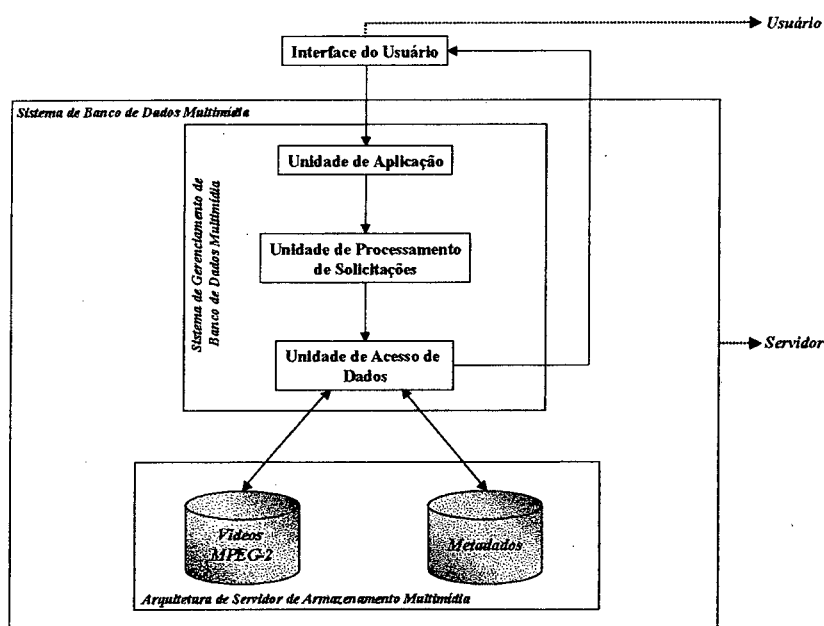


Fig. 7.3 – Esquema de Gerenciamento de Vídeos MPEG-2

- **Interface do Usuário:** a função básica desta interface é servir de ponte entre o usuário e o conteúdo do servidor multimídia, possibilitando que as solicitações sejam formuladas adequadamente através das facilidades fornecidas. A interface deve, por exemplo, permitir que o usuário selecione o vídeo MPEG-2, a partir de uma série de vídeos armazenados no servidor multimídia, associe o nível de qualidade no qual o vídeo deve ser apresentado e receba-o de acordo com o que foi especificado ou, se for o caso, de acordo com o que foi negociado com o sistema.
- **Sistema de Gerenciamento de Banco de Dados Multimídia:** a função deste sistema dentro do esquema apresentado é fornecer facilidades para que as operações de armazenamento e acesso dos vídeos MPEG-2 e seus correspondentes metadados sejam executados transparentemente e de maneira adequada. As unidades que compõe este sistema encontram-se descritas a seguir:
  - **Unidade de Aplicação:** a manipulação do conteúdo do servidor multimídia ao nível de usuário é executada através da elaboração de solicitações, as quais por sua vez são normalmente especificadas utilizando-se uma variedade de linguagens diferentes. A função básica desta unidade é mapear estes diferentes formatos de solicitações em uma linguagem única de representação de dados que possa ser utilizada na localização dos dados de vídeo ou metadados relevantes;
  - **Unidade de Processamento de Solicitações:** o objetivo desta unidade é analisar as solicitações, a fim de identificar quais os dados (vídeos e/ou metadados) devem ser acessados;
  - **Unidade de Acesso de Dados:** com base nas informações recebidas da unidade de processamento de solicitações, esta unidade tem como função acessar do espaço de armazenamento físico e repassar ao dispositivo de saída do usuário os dados identificados anteriormente;
- **Arquitetura de Servidor de Armazenamento Multimídia:** é função desta arquitetura proporcionar o armazenamento, o gerenciamento e a recuperação dos vídeos MPEG-2, bem como de seus metadados, de maneira que os requisitos impostos por estes elementos sejam respeitados. Fazem parte desta arquitetura as unidades descritas abaixo:

- **Vídeos MPEG-2:** representa o espaço de armazenamento físico dentro do servidor multimídia no qual os dados de vídeo MPEG-2 encontram-se armazenados. É função do sistema de Banco de Dados Multimídia gerenciar o armazenamento e a recuperação destes dados, bem como garantir que a entrega dos mesmos aos usuários será feita adequadamente, obedecendo todos os requisitos impostos no que se refere, por exemplo, a transmissão dos elementos que compõem o vídeo e sincronização dos mesmos no destino;
- **Metadados:** representa o espaço de armazenamento físico dentro do servidor multimídia no qual todos os metadados, ou seja, os parâmetros de QoS correspondentes à qualidade na qual cada vídeo foi digitalizado, encontram-se armazenados. É função do sistema de banco de dados multimídia gerenciar o armazenamento e a recuperação de todos os metadados.

É importante observar que, dependendo do tipo de informação representada, o processo de criação de metadados pode ser manual ou automática, ver seção 6.2.1. Da mesma maneira, a natureza destes dados depende do tipo de aplicação que esta sendo implementada e de quais informações são relevantes à mesma [ATMForum, 1997]. Tais metadados podem, por exemplo, fornecer informações do tipo taxa de bits ou taxa de pacote, entre outras informações de QoS.

Para a implementação deste trabalho, os metadados mais interessantes seriam aqueles que descrevem a qualidade na qual o vídeo MPEG-2 foi digitalizado e armazenado, sendo assim pode-se citar: resolução da imagem, taxa de quadro de vídeo e taxa de quadro de áudio.

O que se tem estudado é a importância dos metadados no suporte a facilidades de busca de informações e, principalmente, no suporte eficiente a funções de gerenciamento de QoS, tal como durante a negociação de QoS.

As informações multimídia (vídeos MPEG-2) e os metadados devem ser manipulados pelo Sistema de Banco de Dados Multimídia através de uma linguagem de definição de dados. Esta linguagem permite a especificação, por parte do projetista do banco de dados, dos três principais componentes de uma informação multimídia - estrutura,

conteúdo e apresentação – e dos processos de inserção, recuperação, modificação e eliminação destas informações.

A linguagem de definição de dados deve permitir não somente a definição e a manipulação de informações multimídia, mas também, através da especificação dos requisitos de banco de dados necessários, a definição e o processamento dos metadados, principalmente daqueles ligados à QoS e as regras de localização destas informações, visando o protocolo de negociação de QoS [Vogel, 1995].

A fim de atender aplicações multimídia voltadas aos usuários, o acesso às informações multimídia armazenadas no banco de dados, deve ser feito a partir de mecanismos de consultas que descrevem os metadados associados às informações multimídia desejadas, inclusive as características temporal e espacial destas informações. Para tanto, o usuário deve dispor de uma interface gráfica apropriada que se apresente de acordo com o descrito a seguir [Vogel, 1995]:

- A terminologia utilizada na especificação dos parâmetros deve ser a mais simples e acessível possível, uma vez que para a maioria dos usuários a utilização de termos técnicos seria inviável. Pelo mesmo motivo, a interface deve ocultar sempre que possível os parâmetros de QoS relativos à parte do sistema no qual o usuário não tem acesso;
- A interface deve apresentar ao usuário exemplos, como imagens de tamanhos e resoluções variadas, que lhe permita relacionar sua escolha à qualidade da imagem e som que será oferecido;
- A interface deve informar, de maneira clara, a relação entre custo e desempenho associada à escolha do usuário, a fim de que o mesmo tenha consciência de que quanto melhor a qualidade escolhida, maior o custo do serviço;
- Os requisitos de QoS especificados pelo usuário devem ser mapeados em parâmetros de sistema de maneira automática e transparente;
- A interface deve permitir que o usuário acompanhe, no decorrer da sessão, o comportamento da QoS fornecida, bem como que o mesmo altere os requisitos de QoS originalmente especificados quando necessário;



- A interface deve ser capaz de armazenar as escolhas feitas pelo usuário em perfis, evitando que o mesmo repita o processo de seleção dos requisitos sempre que uma nova sessão tenha que ser inicializada.

## 7.5. CAMADA DE MAPEAMENTO DE QoS

Como dito no capítulo V, é de responsabilidade das Camadas Superiores e da Camada de Adaptação ATM, o mapeamento dos parâmetros de QoS especificados pelo usuário na forma adequada à Camada ATM, ou seja, em parâmetros de QoS ATM, a fim de que a mesma execute as funções de gerenciamento de QoS necessárias.

A camada definida neste trabalho foi denominada *Camada de Mapeamento de QoS*, uma vez que oferece condições para a execução do mapeamento dos parâmetros de QoS especificados pelo usuário e, portanto, parâmetros específicos da aplicação, em parâmetros de QoS específicos de rede.

Posteriormente, este mapeamento possibilitará que a Camada AAL5 dê o tratamento apropriado a estes parâmetros e que a Camada ATM, finalmente, inicie o processo de negociação da QoS em todos os componentes do sistema envolvidos na possível transmissão.

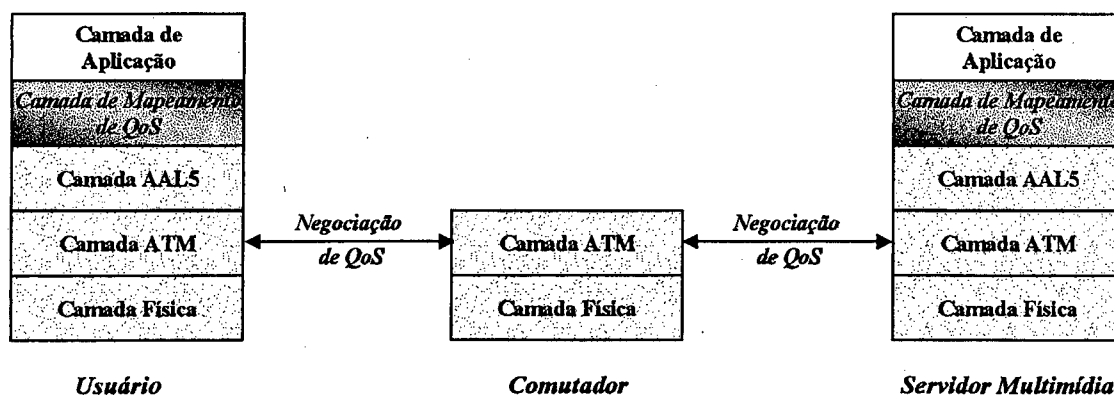


Fig. 7.4 - Camada de Mapeamento de QoS

A localização da Camada de Mapeamento de QoS no Modelo de Camadas ATM é apresentada na figura 7.4. Para que a definição desta camada fosse realizada, o serviço e o protocolo referentes à camada foram discriminados.

O serviço representa o conjunto de funções oferecidas por uma camada à camada superior, através da troca de primitivas de serviços, enquanto que o protocolo refere-se ao conjunto de regras estabelecidas que governam a comunicação entre estas camadas [Soares, 1997].

### 7.5.1. DESCRIÇÃO DO SERVIÇO DE MAPEAMENTO DE QoS

As primitivas de serviço são unidades de dados trocadas somente entre a Camada de Aplicação e a Camada de Mapeamento de QoS. Como a finalidade desta última camada é fazer o mapeamento dos parâmetros de QoS especificados pelo usuário na Camada de Aplicação em parâmetros de QoS inteligíveis pela rede, ou seja, manipular as informações obtidas da camada superior em termos de pacotes e não mais em termos qualitativos, somente as primitivas de serviço descritas abaixo foram definidas:

- *mapear\_parâmetros (conexão\_ID, parâmetros)*: a finalidade desta primitiva de serviço é mapear o nível de QoS especificado pelo usuário na Camada Superior, ou seja, na Camada de Aplicação, em parâmetros de QoS de rede, para que posteriormente a Camada AAL5 transforme-os em parâmetros de QoS na forma adequada à Camada ATM, isto é, em parâmetros de QoS ATM. Somente o endereço ATM e os parâmetros QoS relevantes à aplicação implementada são especificados;
- *confirmar\_mapear\_parâmetros (conexão\_ID, confirmação)*: o objetivo desta primitiva de serviço é indicar se o mapeamento executado pela primitiva anterior ocorreu com sucesso ou não. Neste caso, precisam ser especificados o endereço ATM da conexão e o valor adequado do parâmetro *confirmação*, isto é, positivo caso o mapeamento tenha sido executado e negativo, caso contrário.

### 7.5.2. DESCRIÇÃO DO PROTOCOLO DE MAPEAMENTO DE QoS

O modelo parte da premissa de que a especificação dos parâmetros de QoS é executada somente pelo usuário da aplicação, durante a fase de configuração da conexão que potencialmente será estabelecida. Cabendo ao usuário somente a especificação do nível

de QoS referente à qualidade do som e da imagem. Posteriormente, este nível será mapeado em termos de parâmetros de QoS.

A fase de especificação da QoS é iniciada pelo usuário no momento em que o mesmo seleciona o nível de QoS, o que transparentemente atribui valores a determinados parâmetros de QoS. Este processo ocorre através da interface gráfica disponível.

Os parâmetros especificados pelo usuário na Camada de Aplicação são então utilizados pela Camada de Mapeamento de QoS definida, para que a fase de mapeamento da QoS seja iniciada.

Especificados o vídeo desejado e o nível de QoS sob o qual o mesmo deve ser apresentado, a Camada de Aplicação solicita, através do envio da primitiva *mapear\_parâmetros (conexão\_ID, parâmetros)*, à Camada de Mapeamento de QoS o mapeamento dos parâmetros de QoS especificados pelo usuário, em parâmetros de QoS de rede.

A Camada de Mapeamento de QoS analisa a primitiva e seus atributos, mapea os parâmetros nela contidos em parâmetros de QoS de rede e envia a primitiva *confirmar\_mapear\_parâmetros (conexão\_ID, confirmação)* à Camada de Aplicação, a fim de que o mapeamento seja confirmado ou não.

Caso o mapeamento tenha sido, por um motivo qualquer, cancelado, o usuário deve tentar o envio de uma outra solicitação. Caso contrário, os parâmetros de QoS de rede são repassados à Camada AAL5, para que a mesma transforme-os em parâmetros de QoS ATM e a negociação destes parâmetros possa então ser realizada pela Camada ATM.

Desta forma, as características associadas ao serviço solicitado são mapeadas em termos de células ATM, ou seja, em termos de especificações de tráfego ATM. Por exemplo, a taxa de quadro e a taxa de perda de quadro especificadas serão mapeadas como taxa de célula e taxa de perda de célula, respectivamente.

Além disso, os pacotes MPEG-2 oriundos da Camada de Mapeamento de QoS são transformados em células ATM, de acordo com a especificação “*Audiovisual Multimedia Services: Video on Demand, Specification 1.1*” do ATM Forum descrita na seção 5.10.

A partir deste momento, uma conexão ATM, chamada VCC de Sinalização (figura 7.5) é estabelecida via protocolo de sinalização ATM (ver seção 5.9) e a fase de negociação da QoS é iniciada, em todos os componentes do sistema envolvidos na potencial transmissão do serviço desejado. Durante esta fase, o sistema avalia, com base na quantidade de recursos disponíveis, se pode ou não satisfazer os requisitos de QoS especificados pelo usuário.

Se todos estes componentes puderem e concordarem em satisfazer os requisitos de QoS especificados, os recursos são alocados, uma nova conexão ATM, denominada VCC de Dados, é estabelecida (figura 7.5) e a transmissão do serviço é finalmente iniciada, obedecendo aos requisitos de QoS especificados pelo usuário na fase anterior.

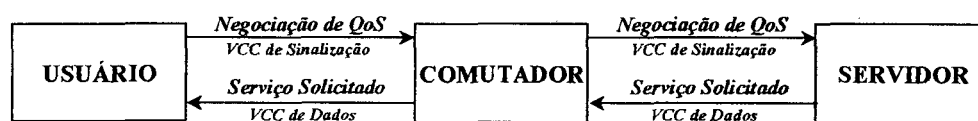


Fig. 7.5. – Conexões ATM Estabelecidas

Caso contrário, o estabelecimento da conexão é negada, cabendo ao usuário tentar imediatamente o estabelecimento de uma nova conexão, porém, com nível de QoS inferior, ou esperar até que o sistema tenha os recursos necessários à satisfação dos requisitos inicialmente solicitados.

## 7.6. EXEMPLIFICAÇÃO DO FUNCIONAMENTO DO MODELO PROPOSTO

Nesta seção o funcionamento do *Modelo para Transmissão de Fluxos MPEG-2 com QoS a partir de Servidores Multimídia em Redes ATM* proposto neste trabalho será detalhadamente exemplificado, tendo como base os conceitos apresentados nas seções anteriores acerca do *Esquema de Gerenciamento de Vídeos MPEG-2* e da *Camada de Mapeamento de QoS* que compõem o modelo citado.

Muito freqüentemente, em uma aplicação multimídia distribuída, o ponto de partida para o processamento dos requisitos de QoS é o usuário. Desta forma, a interface utilizada para a especificação destes requisitos deve ser gráfica e amigável, a fim de que

a escolha dos parâmetros de QoS associados ao serviço que será oferecido seja facilitada.

Idealmente, a forma como os possíveis valores, que podem ser atribuídos aos parâmetros de QoS, são apresentados ao usuário da aplicação deveria obedecer aos critérios qualitativos do material multimídia, neste caso, vídeos MPEG-2, de forma que usuários com qualquer nível de conhecimento pudessem usufruir adequadamente da interface. Desta maneira, os parâmetros de QoS poderiam ser especificados utilizando-se expressões do tipo excelente, bom, razoável, ruim, alto, médio, baixo, etc.

Portanto, para que a exemplificação do funcionamento do modelo fosse melhor visualizada, interfaces gráficas foram elaboradas e utilizadas como parte de uma aplicação virtual, na tentativa de que algumas questões a cerca de uma interface adequada fossem exibidas.

### 7.6.1. SELEÇÃO DO VÍDEO MPEG-2

A partir da interface apresentada na figura 7.6, o usuário tem condições de visualizar quais vídeos MPEG-2 encontram-se armazenados no servidor multimídia e ainda algumas informações, tais como, diretor, atores principais e ano de lançamento, que lhe ajudaram a decidir qual vídeo deve ser apresentado.

**Aplicação Multimídia Distribuída - Protótipo Demonstrativo**

Selecionar Vídeo

<b>Vídeos Disponíveis:</b>	<p><b>101 Dalmátas</b></p> <p>60 Segundos</p> <p>Adoro Problemas</p> <p>Eu, Tu, Eles</p>	<b>Síntese:</b>	<p>Em Londres, um projetista de videogames e uma designer de modas se conhecem e rapidamente se casam. Mas na verdade foi um duplo casamento, pois ambos tinham dalmatas e logo 15 filhotes nascem. Logo depois, uma excêntrica e malévola estilista de moda para quem a dona dos cachorrinhos trabalha, oferece sete mil e quinhentas libras por toda a ninhada.</p>
<b>Ficha Técnica:</b>			
<b>Atores Principais:</b>	Glenn Close		
<b>Direção:</b>	Jeff Daniels		
<b>Ano de Produção:</b>	1996		
<b>Duração:</b>	103 minutos		
<b>Gênero:</b>	Infantil		

**Especificar Qualidade**

Fig. 7.6 – Processo de Seleção do Vídeo MPEG-2

Na verdade, estas informações fazem parte do conjunto de metadados que foram extraídos durante a digitalização do vídeo. Contudo, as mesmas não são consideradas no momento em que o cálculo do nível de QoS está sendo executado, visto que não influenciam na qualidade da apresentação do vídeo.

### 7.6.2. ESPECIFICAÇÃO DO NÍVEL DE QoS

Selecionado o vídeo MPEG-2 desejado, o usuário tem acesso a uma segunda interface (ver figura 7,7) cuja finalidade é disponibilizar informações sobre a qualidade da imagem e do som em que o vídeo selecionado encontra-se armazenado no servidor multimídia, isto é, a qualidade em que o vídeo foi digitalizado.

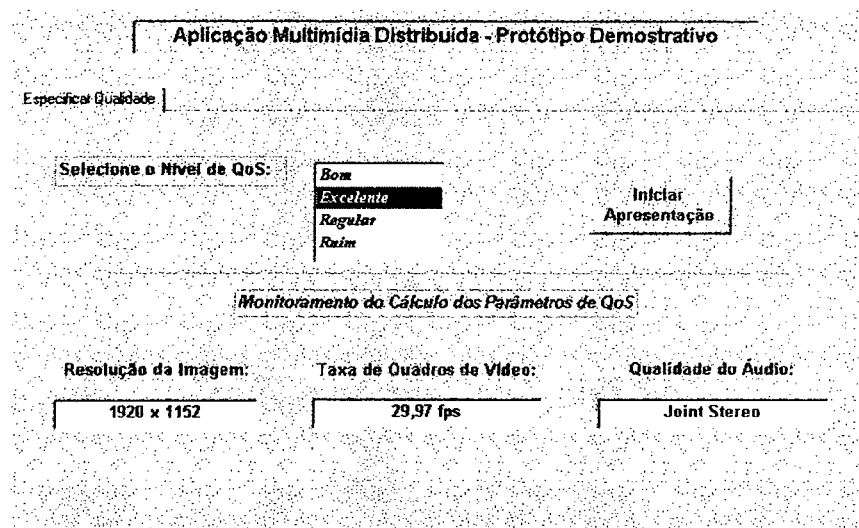


Fig. 7.7 – Especificação do Nível de QoS

Esta interface possibilita, através da opção *Selecione o Nível de QoS*, que o usuário selecione o nível de QoS no qual o vídeo escolhido deve ser apresentado. A fim de obedecer aos critérios qualitativos citados anteriormente e, portanto, facilitar a utilização da aplicação implementada por parte do usuário, esta escolha é feita com base nos termos *Excelente*, *Bom*, *Regular* e *Ruim*.

Escolhido o nível de QoS, o Sistema de Banco de Dados Multimídia é transparentemente acessado, de modo que os parâmetros de QoS correspondentes ao

nível determinado pelo usuário sejam calculados. Vale ressaltar que este cálculo é efetuado tendo como base à qualidade máxima (*Excelente*) no qual o vídeo foi digitalizado e armazenado no servidor multimídia.

Com o intuito de melhor exemplificar o cálculo dos parâmetros de QoS, este processo pode ser acompanhado através dos campos exibidos por esta interface. Para efeito de simplificação, somente a resolução da imagem, taxa de quadro de vídeo e qualidade do áudio foram utilizados como parâmetros de QoS.

Selecionado o vídeo desejado e o nível de qualidade no qual espera-se que o mesmo seja apresentado, o usuário, através de um clique no botão *Iniciar Apresentação*, finaliza a fase de especificação da QoS para, finalmente, iniciar as fases de mapeamento e negociação da QoS, descritas nas seções anteriores.

### 7.6.3. APRESENTAÇÃO DO VÍDEO MPEG-2

Terminadas com sucesso as fases citadas, o servidor está apto a transmitir o vídeo MPEG-2 solicitado, através da conexão ATM estabelecida, e o usuário a assistir o vídeo selecionado, ver figura 7.8. Os processos de transmissão e recepção seguem o esquema ilustrado anteriormente na figura 7.2.

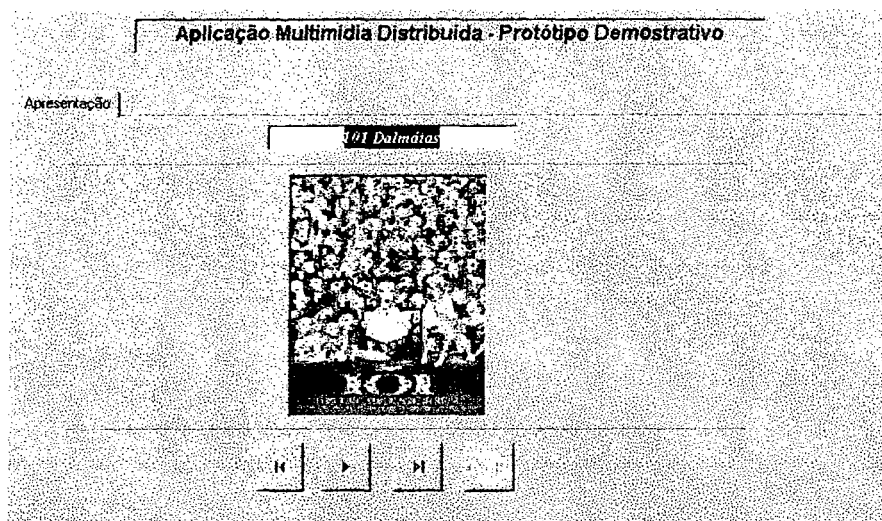


Fig. 7.8 – Apresentação do Vídeo MPEG-2 Solicitado

É importante observar que as interfaces apresentadas nesta seção, não cobrem o mínimo de questões que devem ser levadas em consideração durante o projeto de uma interface gráfica ideal à implementação de aplicações multimídia distribuídas, tal como, descrição da relação entre custo do serviço solicitado e desempenho da rede.

Contudo, as mesmas cumprem o seu propósito que é servir como meio de exemplificação simplificada do processo de seleção e apresentação do vídeo MPEG-2 desejado e especificação dos requisitos de QoS por parte do usuário.

### **7.7. CONSIDERAÇÕES FINAIS**

Este capítulo compôs-se da apresentação do Modelo de Transmissão de Fluxos MPEG-2 com QoS a partir de Servidores Multimídia em Redes ATM proposto nesta dissertação. Segundo a estrutura do modelo, foram descritos o modo como o Gerenciamento de Vídeos MPEG-2 é realizado, a Camada de Mapeamento de QoS definida, que é responsável pelo mapeamento dos parâmetros de QoS especificados pelo usuário na forma inteligível pela rede e, finalmente, a descrição e exemplificação do funcionamento do modelo, tendo como base os elementos apresentados anteriormente.



## CAPÍTULO VIII

### CONCLUSÃO

*A finalidade deste capítulo é apresentar algumas considerações importantes obtidas durante a fase de desenvolvimento deste trabalho de Dissertação de Mestrado, tais como, assuntos estudados, resultados obtidos e dificuldades encontradas no decorrer deste projeto.*

O propósito fundamental do trabalho de Dissertação de Mestrado apresentado é de estudar a transmissão de fluxos multimídia, mais especificamente fluxos MPEG-2, a partir de servidores de armazenamento multimídia, considerando os conceitos de Qualidade de Serviço (QoS).

Durante o desenvolvimento deste trabalho optou-se pela utilização da Tecnologia ATM como meio de transporte comum aos fluxos gerados, uma vez que a mesma dispõe de características essenciais ao tratamento de informações multimídia, tais como, células de tamanho fixo e padronizado e gerenciamento de QoS.

A fim de que o gerenciamento de QoS oferecido por esta tecnologia fosse complementado com facilidades adicionais, ou seja, com facilidades que traduzissem os requisitos de qualidade exigidos pelos usuários das aplicações em uma linguagem comum a esta tecnologia, o trabalho definiu um modelo composto de dois elementos principais: *Sistema de Banco de Dados Multimídia e Camada de Mapeamento de QoS*.

O primeiro elemento é responsável pelo gerenciamento das operações de armazenamento, recuperação e apresentação dos vídeos MPEG-2 e suas correspondentes características, isto é, metadados.

Enquanto que o segundo, oferece mecanismos para a execução do mapeamento dos parâmetros de QoS da aplicação, isto é, parâmetros definidos pelos usuários, em parâmetros de QoS de rede, para que, posteriormente, a negociação destes parâmetros em todos os componentes do sistema e, por fim, a transmissão do vídeo MPEG-2 sejam realizados adequadamente através da Tecnologia ATM.

Durante a fase de definição do Sistema de Banco de Dados Multimídia, a maior dificuldade encontrada foi a decisão de que tipo de dados armazenar no Banco de

Dados. Primeiramente, pensou-se no armazenamento de parâmetros de QoS ao nível da Tecnologia ATM, ou seja, no armazenamento de parâmetros de QoS ATM expressos em células.

Contudo, como a digitalização dos vídeos MPEG-2 é feita ao nível de pacotes, optou-se pelo armazenamento de parâmetros de QoS de rede, tais como, taxa de quadro e taxa de erro de quadro, ao invés de taxa de células e taxa de erro de células.

No que se refere à definição da Camada de Mapeamento de QoS, não foi possível, até a finalização deste projeto, estabelecer a forma exata de execução do mapeamento dos requisitos de QoS, especificados pelos usuários na Camada de Aplicação, em termos de parâmetros de QoS na Camada de Mapeamento de QoS.

Este ainda é um problema inclusive de âmbito internacional no qual vários órgãos e instituições focalizam seus esforços. Desta forma, a resolução deste problema, bem como a validação do modelo proposto através de análise prática, continuam merecendo atenções especiais e ficam como trabalhos futuros.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [ATMFORUM, 1997] ATM FORUM, Technical Committee. **Audiovisual Multimedia Services: Video on Demand Specification 1.1**. 1997.
- [Aurrecoechea, 1998] AURRECOECHEA, Cristina; CAMPBELL, Andrew T; Hauw, Linda. **A Survey of QoS Architectures**. ACM/Springer Verlag Multimedia Systems Journal: Special Issue on QoS Architecture, v.6, n.3, p.138-151, maio, 1998.
- [Benoit, 1997] BENOIT, Herve. **Digital Television: MPEG-1, MPEG-2 and Principles of the DVB Systems**. John Wiley & Sons, 1997.
- [Cecilio, 1996] CECILIO, Edmundo Lopes; RODRIGUES, Rogério Ferreira. **Vídeo sob Demanda**. 1996.
- [Cereda, 1997] CEREDA, Ronaldo Luiz Dias et. al. **ATM: O Futuro das Redes**. São Paulo: Makron Books, 1997.
- [Chiariglione, 1996] CHIARIGLIONE, Leonardo. **MPEG and Multimedia Communications**. 1996. [http://www.csel.stet.it/ufv/leonardo/paper/isce96.htm#about\\_MPEG](http://www.csel.stet.it/ufv/leonardo/paper/isce96.htm#about_MPEG).
- [Elmasri, 2000] ELMASRI, Ramez; NAVATHE, Shamkant B. **Fundamentals of Database Systems**. 3.ed. Addison-Wesley, 2000.
- [Fairhurst, 1999] FAIRHURST, Godred. **Network Delivery of High Quality MPEG-2 Digital Video**. JISC Technology Applications Programme. University of Aberden, 1999.
- [Fluckiger, 1995] FLUCKIGER, François. **Understanding Networked Multimedia: Applications and Technology**. Prentice Hall, 1995.
- [Furht, 1994] FURHT, Borko. **Multimedia Systems: An Overview**. IEEE Multimedia, v.1, n.1, p.47-59, 1994.
- [Gemmel, 1994] GEMMEL, James et. Al. **Multimedia Storage Servers: A Tutorial and Survey**. 1995
- [Gringeri, 1998] GRINGERI, Steven et.al. **Transmission of MPEG-2 Video Streams over ATM**. IEEE Multimedia, v.5, n. 1, p.58-71, jan., 1998.
- [Hafid, 1996] HAFID, Abdelhakim; VON BOCHMANN, Gregor; DSSOULI, Rachida. **Distributed Multimedia Applications and Quality of Service**. 1996. <http://www.iro.umontreal.ca/labs/teleinfo/PubListIndex.html>.
- [Kwok, 1997] KWOK, Timothy. **ATM: The New Paradigm for Internet, Intranet and Residential Broadband Services and Applications**. Prentice Hall PTR, 1997.

- Broadband Services and Applications. Prentice Hall PTR, 1997.**
- [Liu, 1998] LIU, Chunlei. **Multimedia Over IP: RSVP, RTP, RTCP, RTSP.** 1998.  
[http://www.cis.ohio-state.edu/~jain/cis788-97/ip\\_multimedia/index.htm](http://www.cis.ohio-state.edu/~jain/cis788-97/ip_multimedia/index.htm).
- [Lu, 1996] LU, Guojun. **Communication and Computing for Distributed Multimedia System.** Artech House, 1996.
- [Orzessek, 1998] ORZESSEK, Michael; SOMMER, Peter. **ATM & MPEG-2: Integrating Digital Video into Broadband Networks.** Prentice Hall PTR, 1998.
- [Prabhakaran, 1997] PRABHAKARAN, B. **Multimedia Database Management Systems.** Kluwer Academic Publishers, 1997.
- [Raghavan, 1998] RAGHAVAN, S. V.; TRIPATHI, Satish K. **Networked Multimedia Systems: Concepts, Architecture, and Design.** Prentice Hall, 1998.
- [Richards, 1998] RICHARDS, Antony et. al. **Mapping User Level QoS from a Single Parameter.** Second IFIP/IEEE International Conference on Management of Multimedia Networks and Services, Versailles, France, 1998
- [Sandstå, 1997] SANDSTÅ, Olav; LANGØRGEN, Stein; MIDTSTRAUM, Roger. **Video Server on na ATM Connected Cluster of Workstations.** 1997.  
[http://www.idi.ntnu.no/IDT/grupper/DB-grp/tech\\_papers/SCCC97\\_elvira/html/sccc97.html](http://www.idi.ntnu.no/IDT/grupper/DB-grp/tech_papers/SCCC97_elvira/html/sccc97.html)
- [Sheth, 1998] SHETH, Amit; KLAS, Wolfgang. **Multimedia Data Management: Using Metadata to Integrate and Apply Digital Media.** McGraw-Hill, 1998.
- [Silveira, 2000] SILVEIRA, Regina M.; RUGGIERO, Wilson V. **Servidor de Qualidade para Sistema de Video sob Demanda.** VI Simpósio Brasileiro de Multimídia e Hiperemídia, p.325-340. Natal, 2000.
- [Soares, 1997] SOARES, Luiz Fernando Gomes; LEMOS, Guido; COLCHER, Sérgio. **Redes de Computadores: das LANs, MANs e WANs às Redes ATM.** 2.ed. Campus, 1997.
- [Steinmetz, 1995] STEINMETZ, Ralf; NAHRSTEDT, Klara. **Multimedia: Computing, Communications & Applications.** Prentice Hall, 1995.
- [Volgel, 1995] VOGEL, Andreas et. al. **Distributed Multimedia and QoS: A Survey.** IEEE Multimedia. v.2, n.2, p.10-18., 1995.
- [Witana, 1999] WITANA, Varuni; FRY, Michael; ANTONIADES, Mark. **A Software Framework for Application Level QoS Management.** IEEE / IEE Quality of Service, 1999.

IWQoS '99. P.52-61, 1999

[Zheng, 1999]

ZHENG, Bing; ANTIQUZZAMAN, Mohammed. **Traffic Management of Multimedia over ATM Networks**. IEEE Communication Magazine, v.37, n.1, p.33-38, jan., 1999.

### BIBLIOGRAFIAS CONSULTADAS

- [Adams, 1995] ADAMS, Michael. **Real Time MPEG Asset Delivery over ATM**. 1995. <http://www.pahfinder.com/corp/tech/adams/mpegoveratm/mpegoveratm.html>. Acessado em 09/1999
- [Amor, 1999] AMOR, Daniel. **ATM Networks**. <Http://www.cli.di.unipi.it/~amor/atm2.html>. Acesado em agosto de 1999.
- [Cisco, 1999] CISCO. **Provisioning ATM Services**. [http://cisco.com/univercd/cc/td/doc/product/rtrmgmt/cpc/ug\\_1\\_1/pratn.htm](http://cisco.com/univercd/cc/td/doc/product/rtrmgmt/cpc/ug_1_1/pratn.htm). Acessado em Agosto de 1999
- [Coughlan, 1999] COUGHLAN, Ivan. **Video on Demand Technologies**. <http://skynet.ul.ie/~dracula/telecom.html>. Acessado em maio de 1999.
- [Ferguson, 1998] FERGUSON, Paul; HUSTON Geoff. **Quality of Service: Delivering QoS on the Internet and in Corporate Networks**. John Wiley & Sons, 1998.
- [Fry, 1997] FRY, Michael et. al. **Managing QoS in Multimedia Services**. 1997. Journal of Network and Systems Management, v.5, n.3, p.283-300, 1997.
- [Gringeri, 1998] GRINGERI, Steven et.al. **Transmission of MPEG-2 Video Streams over ATM**. IEEE Multimedia, v.5 , n. 1, p.58-71 , jan., 1998.
- [Huoponen, 1996] HUOPONEN, Jani; WAGNER, Thorsten. **Video on Demand: A Survey**. 1996. [http://fiddle.ee.vt.edu/courses/ee4984/Projects1996/huoponen\\_wagner/huoponen\\_wagner.html#ref1](http://fiddle.ee.vt.edu/courses/ee4984/Projects1996/huoponen_wagner/huoponen_wagner.html#ref1). Acessado em maio de 1999.
- [Kansas, 1999] UNIVERSITY OF KANSAS. **ATM Classes**. [http://www.tisl.ukans.edu/EECS/EECS\\_700\\_Nets/qos.html](http://www.tisl.ukans.edu/EECS/EECS_700_Nets/qos.html). Acessado em agosto de 1999.
- [Käppner, 94] KÄPPNER, Thomas; WOLF, Lars C. **Media Scaling in Distributed Multimedia Objects Services**. <http://www.kom.e-technik.tu-darmstadt.de/~lars/pub-lars.html>. Acessado em setembro de 1999.
- [Kerhervé, 1996] KERHERVÉ, Brigitte et. al. **Metadata Modeling for Quality of Service Management in Distributed Multimedia Systems**. IEEE Metadata Conference, 1996
- [Kunert, 2000] KUNERT, Del. **Today's Video Servers: Key Technology Issues**. <http://www.ccur.com/ivod/articles.html>. Acessado em 06/2000.
- [Kuo, 1998] KUO, Franklin; EFFLESBERG, Wolfgang; GARCIA-LUNA-ACEVES, J. J. **Multimedia Communications: Protocols and Applications**. Prentice Hall, 1998.
- [Metropoa, 1999] PROJETO METROPOA/PROCERGS. **Video sob Demanda - Compressão de Vídeo**. <http://www.via-rs.com.br/metropoa/compIntro.htm>. Acessado em junho de 1999.

- [MPEG, 1999] THE MPEG HOME PAGE. <http://www.cselt.stet.it/mpeg/index.htm>. Acessado em junho de 1999.
- [Mussauer, 1997] MUSSAUER, Ronaldo de Lima; RUBINSTEIN, Marcelo Gonçalves; DUARTE, Otto Carlos Muniz B. **Sincronização Multimídia**. Rio de Janeiro: 1997.
- [Nahrstedt, 1996] NAHRSTEDT, Klara; QIAO, Lintian. **Tuning System for Distributed Multimedia Applications**. University of Illinois, 1996. <http://cairo.cs.uiuc.edu/papers.html>
- [Neuffer, 1999] NEUFFER. **What is RAID?** [http://www.uni-mainz.de/~neuffer/scsi/what\\_is RAID.html](http://www.uni-mainz.de/~neuffer/scsi/what_is RAID.html). Acessado em maio de 1999.
- [Niemelä, 1996] NIEMELÄ, Marko. **Prospects for Interactive Video-On-Demand**. [http://stekt.oulu.fi/multimedia/multimedia\\_96/interact/](http://stekt.oulu.fi/multimedia/multimedia_96/interact/). Acessado em maio de 1999.
- [Ott, 1997] OTT, M.; MICHELITSCH, D. Reininger; WELLING, G. **An Architecture for Adaptive QoS and its Application to Multimedia Systems Design**. 1997. <http://ccrl.nj.nec.com/paper/97-R-005/paper.html>.
- [Paek, 1995] PAEK, Seungyup; BOCHECK, Paul; CHANG, Shih-Fu. **Scalable MPEG-2 Video Servers with Heterogeneous QoS on Parallel Disk Arrays**. 5<sup>th</sup> IEEE Workshop on Network & Operating System Support for Digital Audio & Video, New Hampshire, USA, 1995.
- [Pegler, 2000] PEGLER, D.W. **SCALable Multimedia Storage**. <http://tina.lancs.ac.uk/computing/staff/dwp.html>. Acessado em 05/2000.
- [Rajapakshe, 1995] RAJAPAKSHE, Harindra; QUEK, Derek Paul. **Video on Demand**. 1995. [http://www-dse.doc.ic.ac.uk/~nd/surprise\\_95/journal/vol4/shr/report.html](http://www-dse.doc.ic.ac.uk/~nd/surprise_95/journal/vol4/shr/report.html). Acessado em abril de 1999.
- [Recitronic, 1999] **O que é RAID?** <http://www.recitronic.com.br/multi08.html>. Acessado em fevereiro de 1999.
- [Shan, 1997] SHAN, Jinhua; DONG, Wei. **ATM Networks for Multimedia: Quality of Service**. 1997. [http://fiddle.visc.ece.vt.edu/courses/ee4984/Projects1997/dong\\_shan.html](http://fiddle.visc.ece.vt.edu/courses/ee4984/Projects1997/dong_shan.html). Acessado em 02/2000.
- [Sikora, 2000] SIKORA, Thomas. **MPEG-1 and MPEG-2 Digital Video Coding Standards**. [http://wwwam.HHI.DE/mpeg-video/papers/sikora/mpeg1\\_2/mpeg1\\_2.htm](http://wwwam.HHI.DE/mpeg-video/papers/sikora/mpeg1_2/mpeg1_2.htm). Acessado em 03/2000.
- [Tanenbaum, 1997] TANENBAUM, Andrew S. **Redes de Computadores**. Rio de Janeiro: Campus, 1997.
- [Wiltfang, 1997] WILTFANG, Hajo R.; SCHMIDT, C. **QoS Monitoring for ATM-based Networks**. International Conference on Management of Multimedia Networks and Services, Montreal, Canada, 1997.

[Wiseman, 1999] WISEMAN, John. **An Introduction to MPEG Video Compression.**  
<http://members.aol.com/symbandgrl/>. Acessado em junho de 1999.