

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação

Descrição e Proposição de uma Solução de IPTV.

Autor: Marcus Pereira Magri
Orientador: Prof. Dr. Yuzo Iano

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: **Telecomunicações e Telemática.**

Banca Examinadora

Yuzo IanoDecom/Feec/Unicamp
Phillip Mark Seymour BurtPTC/POLI/USP
Dalton Soares ArantesDecom/Feec/Unicamp
Luís Geraldo Pedroso MeloniDecom/Feec/Unicamp

Campinas, SP

Agosto/2007

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

M276d Magri, Marcus Pereira
Descrição e proposição de uma solução de IPTV. /
Marcus Pereira Magri. --Campinas, SP: [s.n.], 2007.

Orientador: Yuzo Iano
Dissertação (Mestrado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica e de
Computação.

1. Compressão de imagens. 2. MPEG (Padrão de
codificação de vídeo). 3. TCP/IP (Protocolos de rede de
computação). 4. Redes de computação/Protocolos. 5.
Direitos autorais – Cinema. 6. Anúncios em televisão. 7.
Audiências de televisão. 8. Televisão digital. 9. Televisão
por assinatura. I. Iano, Yuzo. II. Universidade Estadual de
Campinas. Faculdade de Engenharia Elétrica e de
Computação. III. Título.

Título em Inglês: Description and proposal of an end-to-end IPTV solution

Palavras-chave em Inglês: Image coding, IPTV, Triple play, STB, VOD, IGMP,
MPEG, H.264, MPEG-2 TS, RTSP, MIDDLEWARE,
DRM, Multicast

Área de concentração: Telecomunicações e Telemática

Titulação: Mestre em Engenharia Elétrica

Banca examinadora: Phillip Mark Seymour Burt, Dalton Soares Arantes, Luís
Geraldo Pedroso Meloni

Data da defesa: 21/08/2007

Programa de Pós-Graduação: Engenharia Elétrica

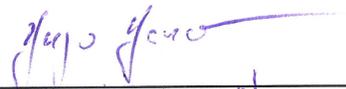
COMISSÃO JULGADORA - TESE DE MESTRADO

Candidato: Marcus Pereira Magri

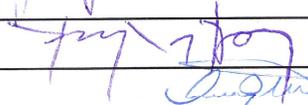
Data da Defesa: 21 de agosto de 2007

Título da Tese: "Descrição e Proposição de uma Solução de IPTV"

Prof. Dr. Yuzo Iano (Presidente):



Prof. Dr. Phillip Mark Seymour Burt:



Prof. Dr. Luís Geraldo Pedroso Meloni:



Prof. Dr. Dalton Soares Arantes:



Página em branco

Resumo

O objetivo deste trabalho é tanto estudar quanto analisar uma solução de IPTV (TV por IP). Tal serviço está emergindo nas operadoras de telecomunicações de todo o mundo. Essa solução está sendo bem aceita e mudando radicalmente a visão de negócios dos provedores de telecomunicações em todo o mundo. Consumidores finais vão demandar serviços personalizados através de múltiplos dispositivos, incluindo canais de TV, conteúdo sob demanda, TV interativa, comunicações de vídeo e de voz, música, compartilhamento de fotos e arquivos e jogos *on-line*. Esses são alguns dos serviços de usuário final, entregues de forma transparente e combinada. Isso tudo requer mudanças inovadoras nas operadoras de telecomunicações, desde uma nova arquitetura de rede, com qualidade de serviço fim-a-fim, confiabilidade, gerência, e segurança. Contribui-se com uma descrição detalhada de como arquitetar essa solução, cobrindo todos os desafios para se implementar um sistema de IPTV, descrevendo-se todos os componentes e serviços da solução, e também qual seria a evolução natural dessa solução. É também parte de nossa contribuição a proposição de algumas melhorias para reduzir o tempo de troca de canal, além de sugestões para se implementar uma solução de inserção de comercial e de inserção de legenda.

Palavras-chave: Codificação de imagem, Imagem, IPTV, *Triple Play*, STB, VOD, IGMP.

Abstract

This work concerns the study and the analysis of an IPTV solution (Internet Protocol Television). This kind of service is emerging in the Telco all over the world and its good acceptance is radically changing the business vision of the telecommunications providers. Final consumers will demand personalized services using multiple devices including TV channels, video-on-demand, interactive TV, video and voice communications, music, photo and file sharing and online games. All these services require novelties and changes at the telecoms field embracing new network architecture using P2P service quality, reliability, customer management and security. The contribution of this work concerns a detailed description of how to design the solution for these challenges, how to implement a new IPTV system describing all the components and services that have to be provided. The natural evolution of this approach is also considered. Additionally, this work proposes new solutions to reduce the channel zapping time and to insert commercials and subtitles.

Keywords: Image coding, Imaging, IPTV, Triple Play, STB, VOD, IGMP, MPEG, H.264, Multicast, MPEG-2 TS, RTSP, Middleware, DRM.

A Deus por tudo aquilo que Ele representa nas nossas vidas.

Agradecimentos

À minha esposa Miriam, meus filhos Marcus Eduardo e Gabriela por toda inspiração, dedicação, suporte e carinho.

Aos meus pais, José Eduardo Magri e Maria Pereira Magri, meus irmão Júnior e Nana por sua lição de vida, amor e apoio durante esta jornada.

Ao meu orientador, Prof. Yuzo Iano pelas suas valiosas contribuições, dedicação, paciência e ajuda.

Aos amigos e colegas da Lucent Technologies, especialmente ao meu gerente Celso Barros que me apoiaram desde o início.

Aos demais colegas de pós-graduação, pelas sugestões e conselhos.

Sumário

Lista de Figuras	xii
Lista de Tabelas	xv
Lista de Símbolos	xv
Abreviaturas	xvi
Publicações	xxi
1. Introdução	1
1.1 História	2
1.2 Motivações e Objetivos do Trabalho	3
1.3 TV a Cabo x IPTV	4
1.4 Desafios para implementar IPTV	6
1.5 Organização da Dissertação	7
1.6 Resumo das Contribuições	8
2. Compressão de Vídeo e Áudio, Fluxo de Transporte MPEG, IP Multicast e RTSP	9
2.1 Compressão de Vídeo	9
2.1.1 MPEG-2	10
2.1.2 H.264/MPEG-4 AVC	22
2.1.3 VC-1	24
2.2 Compressão de Áudio	26
2.2.1 MPEG Áudio	26
2.2.2 Dolby Digital (AC-3)	31
2.2.3 MPEG-2 AAC	32
2.3 Tecnologia de Transporte de Conteúdo	32
2.3.1 MPEG-2 <i>Transport Stream</i>	32
2.3.2 IP Multicast	43
2.3.3 RTSP	48
3. Componentes e Redes do IPTV	51
3.1 Introdução	51
3.2 <i>Head-End</i>	52
3.2.1 Aquisição do Conteúdo	53
3.2.2 Provedor de Conteúdo	56
3.2.3 Inserção de Legenda	57
3.3 Inserção de Comercial	58
3.4 <i>Middleware</i>	60

3.4.1	Integração com OSS/BSS	65
3.5	Servidor de Vídeo	68
3.6	Proteção do Conteúdo e Segurança de Rede	74
3.6.1	Acesso Condicional - CAS	74
3.6.2	Gerência dos Direitos Digitais - DRM	77
3.7	Arquitetura de Rede	79
3.8	Rede Residencial	81
3.8.1	<i>Modem</i>	81
3.8.2	<i>Set Top Box</i>	82
3.9	Rede de Acesso	82
3.9.1	DSL	82
3.9.2	FTTH	85
3.9.3	WiMAX	86
3.10	Rede de Agregação	86
3.11	Núcleo da Rede	87
3.12	Rede de Conteúdo	87
3.13	Qualidade de Serviço (QoS)	88
3.14	Medição de Audiência	91
3.15	Serviços de IPTV	94
3.15.1	Guia de Programação Eletrônica	95
3.15.2	Canais de TV ou de Áudio	95
3.15.3	Controle Paterno	96
3.15.4	<i>Pay Per View</i>	96
3.15.5	<i>Video on Demand</i>	97
3.15.6	<i>Near Video on Demand</i>	98
3.15.7	<i>Subscription Video on Demand</i>	99
3.15.8	<i>Music on Demand</i>	99
3.15.9	<i>Personal Video Recorder (PVR)</i>	99
3.15.10	<i>Pause Live TV</i>	100
3.15.11	<i>Network Personal Video Recorder</i>	101
3.15.12	<i>Games on Demand</i>	103
3.15.13	<i>Walled Garden</i>	104
3.15.14	<i>Web TV</i>	104
3.15.15	Vídeo Conferência	105
3.15.16	Interface de Usuário de Vários Idiomas	106
3.15.17	PiP e Mosaico Dinâmico	106
3.15.18	Múltiplos Ângulos e Múltiplos PiPs	107
3.15.19	Lembrete	107
3.15.20	Canal Pessoal	108
3.15.21	Assistir Programa com Outras Pessoas Remotas	108
3.15.22	Identificação Interativa de BINA na TV	109
3.15.23	<i>T-Commerce</i>	110
3.16	<i>Gerência de Rede</i>	111

4. Sistema de IPTV: Proposta e Implementação.....	113
4.1 <i>Head-End</i>	113
4.1.1 Redundância	113
4.1.2 Inserção de Legenda	113
4.1.3 Inserção de Comercial	115
4.2 <i>Middleware</i>	118
4.2.1 Interface dos Usuários	118
4.2.2 Serviços.....	118
4.2.3 Redundância	119
4.3 Rede	120
4.3.1 Mudança Rápida de Canal	121
5. Conclusões	125
5.1 Sugestões para Trabalhos Futuros	126
Apêndice A: Configuração de <i>Head-end</i>	129
A.1 Codificação de Vídeo	129
A.2 Codificação de Áudio	131
A.3 Codificação de Fluxo de Transporte.....	132
A.4 Criptografia de Canais em Tempo Real.....	135
Apêndice B: Configuração de VOD.....	137
B.1 Criação de Conteúdo VOD.....	137
Referências Bibliográficas	139

Lista de Figuras

Figura 1.1 Sistema de TV a cabo.	5
Figura 1.2 Comparação entre TV a cabo e IPTV.	6
Figura 2.1 - Localização espacial das amostras de luminância e crominância para o formato 4:2:0.	12
Figura 2.2 - Diagrama geral de um codificador.	14
Figura 2.3 - Subamostragens de crominância MPEG-2.	14
Figura 2.4 - Estrutura de um quadro não-entrelaçado de formato 4:2:0.	15
Figura 2.5 - Exemplo de estrutura temporal de seqüência (m = distância entre âncoras; n = distância entre quadros tipo I).	16
Figura 2.6 - Codificação espacial.	18
Figura 2.7 - Codificação temporal.	18
Figura 2.8 - Estimação de movimento.	19
Figura 2.9 - Modos de varredura dos coeficientes do bloco transformado.	20
Figura 2.10 - Diagrama geral de um codificador H.264.	22
Figura 2.11 - Mascaramento de frequências.	27
Figura 2.12 - Mascaramento de frequências nas bandas críticas MPEG.	27
Figura 2.13 - Bloco de compressão de áudio genérico para as três camadas.	29
Figura 2.14 - Áudio multicanal.	31
Figura 2.15 - Diagrama de blocos do MPEG-2 TS.	33
Figura 2.16 - Criação do PES.	34
Figura 2.17 - Cabeçalho do PES.	34
Figura 2.18 - PID no fluxo elementar.	36
Figura 2.19 - Diagrama de blocos do MPEG-2 TS.	38
Figura 2.20 - Tabelas PSI do MPEG-2 TS e PIDs.	39
Figura 2.21 - Sinal PCR.	39
Figura 2.22 - Seqüência de tempo no fluxo de transporte.	40
Figura 2.23 - Fluxo de Transporte MPEG-2 TS com tabelas PSI.	41
Figura 2.24 - Processo de decodificação.	42
Figura 2.25 - Diagrama simplificado de codificação e inserção de legenda.	43
Figura 2.26 - Exemplo de uma transmissão <i>unicast</i>	45
Figura 2.27 - Exemplo de uma transmissão sem e com <i>multicast</i>	46
Figura 2.28 - Protocolos de <i>multicast</i> IGMP x PIM.	47
Figura 2.29 - IGMP <i>join</i>	48
Figura 3.1 - Diagrama de blocos para um sistema de IPTV.	52
Figura 3.2 - Elementos Principais de uma rede de IPTV.	52

Figura 3.3 - Recepção de TV via satélite.	53
Figura 3.4 - Recepção satelital.	54
Figura 3.5 - Recepção de TV aberta do ar.	55
Figura 3.6 - Provedor de conteúdo.	57
Figura 3.7 - Inserção de comercial – <i>splicer</i>	58
Figura 3.8 - Diagrama de blocos da inserção de comercial.	59
Figura 3.9 - Componentes do <i>middleware</i>	61
Figura 3.10 - Exemplo de criação de um canal de TV.	63
Figura 3.11 - Exemplo de criação de um pacote de TV.	63
Figura 3.12 - Exemplo de criação de um usuário.	64
Figura 3.13 - Fluxo para serviços de TV.	65
Figura 3.14 - <i>Application Programing Interfaces</i> (API).	65
Figura 3.15 - Integração entre sistemas.	67
Figura 3.16 - Servidor de vídeo.	69
Figura 3.17 - RAID 5.	70
Figura 3.18 - Codificação de um conteúdo para VOD.	71
Figura 3.19 - Comandos RTSP.	74
Figura 3.20 – Criptografia.	75
Figura 3.21 - CAS/DRM e protocolos usados pelo STB.	76
Figura 3.22 - Situação da rede anterior e situação da rede atual.	80
Figura 3.23 - Arquitetura de Rede.	80
Figura 3.24 - Rede residencial.	81
Figura 3.25 - Rede de acesso DSL.	83
Figura 3.26 - Taxa <i>downstream</i> x Comprimento do <i>loop</i>	85
Figura 3.27 - Rede de agregação.	87
Figura 3.28 - Modelo DiffServ.	89
Figura 3.29 - Nó sem QoS x nó com QoS.	90
Figura 3.30 - Medição de audiência.	92
Figura 3.31 - Medição de audiência instantânea para canais de TV.	93
Figura 3.32 - Medição de audiência VOD.	94
Figura 3.33 - Guia de programação eletrônica.	95
Figura 3.34 - Canais de TV e mini guia.	96
Figura 3.35 - Controle paterno.	96
Figura 3.36 - <i>Pay per view</i>	97
Figura 3.37 - <i>Video on demand</i>	98
Figura 3.38 – <i>Subscription video on demand</i>	99
Figura 3.39 - <i>Personal video recording</i>	100
Figura 3.40 - <i>Pause live TV</i>	101
Figura 3.41 – NPVR.	103
Figura 3.42 - <i>Walled garden</i>	104
Figura 3.43 - <i>Vídeo conferência</i>	105
Figura 3.44 - Interface de usuário de vários idiomas.	106
Figura 3.45 - Mosaico dinâmico.	107
Figura 3.46 - Múltiplos ângulos.	107

Figura 3.47 – Lembrete.	108
Figura 3.48 - Canal pessoal.	108
Figura 3.49 - Assistir programas com pessoas remotas.	109
Figura 3.50 - <i>Interactive TV caller ID</i>	110
Figura 3.51 - <i>T-Commerce</i>	110
Figura 3.52 - Sistema de gerência SNMP.	111
Figura 4.1 - Inserção de legenda - <i>TV broadcast</i>	114
Figura 4.2 - Inserção de legenda – VOD.	115
Figura 4.3 - Diagrama de blocos da inserção de comercial.	116
Figura 4.4 - Agendamento de um comercial.	117
Figura 4.5 - Servidores do <i>middleware</i> em alta disponibilidade.	120
Figura 4.6 - Mudança rápida de canal.	122
Figura 4.7 - <i>Buffer</i> circular.	123
Figura A.1 - Parâmetros básicos de codificação de vídeo (Codificador da Harmonic).	130
Figura A.2 - Parâmetros avançados de codificação de vídeo (Codificador da Harmonic).	130
Figura A.3 - Parâmetros de codificação de áudio (Codificador da Harmonic).	131
Figura A.4 - Parâmetros de codificação de áudio MPEG L2 (Codificador da Harmonic).	132
Figura A.5 - Parâmetros gerais de fluxo de transporte MPEG-2 TS (Codificador da Harmonic).	133
Figura A.6 - Parâmetros IP de fluxo de transporte MPEG-2 TS (Codificador da Harmonic). ...	134
Figura A.7 - Lista de fluxos (Codificador da Harmonic).	135
Figura A.8 - Criptografia de canais (Sistema de CAS/DRM da Widevine).	136
Figura B.1 - Criação de metadados (Middleware da Alcatel-Lucent).	136

Lista de Tabelas

Tabela 2.1 - Taxa de compressão de vídeo	26
Tabela 2.2 - Características centrais das três camadas de compressão de áudio MPEG	28
Tabela 2.3 - Características centrais das três camadas de compressão de áudio MPEG	31
Tabela 2.4 - Cabeçalho do PES	35
Tabela 2.5 - Cabeçalho do PES - campos opcionais	36

Lista de Símbolos

$MSE(i, j)$	Erro quadrático médio.
$Q_a(x, y)$	Quadro atual.
$Q_{ref}(x, y)$	Quadro de referência.
(i, j)	Deslocamento entre os dois blocos.
(a, b)	Posição do bloco atual.

Abreviaturas

- ADSL – *Asymmetrical Digital Subscriber Line*, Linha de Assinante Digital Assimétrica
- ADSL2+ – *Asymmetrical Digital Subscriber Line second generation*, Linha de Assinante Digital Assimétrica de Segunda Geração
- AES – *Advanced Encryption Standard*, Padrão de Criptografia Avançada
- AGC – *Automatic Gain Control*, Controle de Ganho Automático
- API – *Application Programming Interface*, Interface de Programação
- AS – *Application Server*, Servidor de Aplicação
- ASO – *Arbitrary Slice Ordering*, Ordenação Arbitrária de Fatias
- ATM – *Asynchronous Transfer Mode*, Modo de Transferência Assíncrona
- ATSC – *Advanced Television Systems Committee*
- AVC – *Advanced Video Compression*, Compressão de Vídeo Avançada
- BSS – *Billing Support System*, Sistema de Apoio ao Faturamento
- CABAC – *Context Adaptive Binary Arithmetic Coding*, Codificação Aritmética Binária de Contexto Adaptativo
- CAT – *Conditional Access Table*, Tabela de Acesso Condicional
- CAVLC – *Context Adaptive Variable-Length Coding*, Codificação Aritmética Binária Adaptável segundo o Contexto
- CBR – *Constant Bit Rate*, Taxa de Bits Constante
- DB – *Data Base*, Base de Dados
- DCT – *Discrete Cosine Transform*, Transformada Discreta de Cosseno
- DES – *Data Encryption Standard*, Padrão de Criptografia de Dados
- DFT – *Discrete Fourier Transform*, Transformada Discreta de Fourier
- DOS – *Denial of Service*, Ataque Massivo ao Serviço (tentativa de tirá-lo de operação)

DP	- <i>Data Partitioning</i> , Partição de Dados
DPCM	- <i>Differential Pulse Code Modulation</i> , Modulação por Código de Pulso Diferencial
DPI	- <i>Digital Program Insertion</i> , Inserção de Programa Digital
DRM	- <i>Digital Rights Management</i> , Gerência dos Direitos Digitais
DSL	- <i>Digital Subscriber Line</i> , Linha de Assinante Digital
DSLAM	- <i>Digital Subscriber Line Access Multiplexer</i> , Multiplexador de Linhas de Acesso Digital
DTMF	- <i>Dual-Tone Multi-Frequency</i>
DTS	- <i>Decoding Time Stamp</i> , Selo de Tempo da Decodificação
DVB	- <i>Digital Vídeo Broadcast</i>
ECM	- <i>Entitlement Control Message</i>
EMM	- <i>Entitlement Management Message</i>
EPG	- <i>Electronic Programming Guide</i> , Guia de Programação Eletrônica
ES	- <i>Elementary Stream</i> , Fluxo Elementar
FIFO	- <i>First in, First out</i> , Primeiro que Entra, Primeiro que Sai
FMO	- <i>Flexible Macroblock Ordering</i> , Ordenação de Macroblocos Flexível
GOD	- <i>Games on Demand</i> , Jogos sob Demanda
GOP	- <i>Group of Pictures</i> , Grupo de Quadros
HD-SDI	- <i>High Definition Serial Digital Interface</i> , Interface Serial Digital de Alta Definição
HDTV	- <i>High Definition Television</i> , Televisão de Alta Definição
HTTP	- <i>HyperText Transfer Protocol</i> , Protocolo de Transferência Hipertexto
IETF	- <i>Internet Engineering Task Force</i>
IGMP	- <i>Internet Group Management Protocol</i>
IP	- <i>Internet Protocol</i> , Protocolo Internet
IPTV	- <i>Internet Protocol Television</i> , Televisão sobre Protocolo de Internet
IRD	- <i>Integrated Receiver Decoder</i> , Receptor e Decodificador Integrado
ISO	- <i>International Standards Organization</i> , Organização de Padrões Internacionais
ITU	- <i>International Telecommunication Union</i> , União Internacional de Telecomunicações

JPEG	- <i>Joint Photographic Experts Group</i>
LNB	- <i>Low Noise Block</i>
MDCT	- <i>Modified Discrete Cosine Transform</i> , Transformada de Cosseno Discreta Modificada
MOD	- <i>Music on Demand</i> , Música sob Demanda
MPEG	- <i>Motion Picture Experts Group</i>
MPEG-2 TS	- <i>Motion Picture Experts Group Transport Stream</i> , Fluxo de Transporte MPEG-2
MPTS	- <i>Multiple Program Transport Stream</i> , Fluxo de Transporte de Vários Programas
MSE	- <i>Mean Square Error</i> , Erro Quadrático Médio
NBC	- <i>Non Backward Compatible</i>
NIT	- <i>Network Information Table</i> , Tabela de Informação de Rede
NPVR	- <i>Network Personal Video Recorder</i> , Gravação de Vídeo Pessoal na Rede
NTP	- <i>Network Time Protocol</i> , Protocolo de Tempo na Rede
NVOD	- <i>Near Video on Demand</i> , Variante do Serviço PPV
OSS	- <i>Operating Support System</i> , Sistema de Apoio à Operação
PAT	- <i>Program Association Table</i> , Tabela de Associação do Programa
PCM	- <i>Pulse-Code Modulation</i> , Modulação por Código de pulso
PCR	- <i>Program Clock Reference</i> , Referência de Relógio
PES	- <i>Packetized Elementary Streams</i> , Fluxos Elementares Empacotados
PID	- <i>Program Identification</i> , Identificação do Programa
PIM	- <i>Protocol Independent Multicast</i>
PiP	- <i>Picture in Picture</i> , Quadro sobre Quadro
PLL	- <i>Phase Lock Loop</i>
PMT	- <i>Program Map Table</i> , Tabela de Mapeamento do Programa
PON	- <i>Passive Optical Network</i> , Rede Óptica Passiva
POTS	- <i>Plain Old Telephone Service</i> , Serviço de Telefonia Tradicional
PPV	- <i>Pay Per View</i> , Pague para Ver

PSI	- <i>Program Specific Information</i> , Informações Específicas do Programa
PTS	- <i>Presentation Time Stamp</i> , Selo de Tempo da Apresentação
PVR	- <i>Personal Video Recorder</i> , Gravação de Vídeo Pessoal
QoS	- <i>Quality of Service</i> , Qualidade de Serviço
QPSK	- <i>Quadrature Phase Shift Keying modulation</i>
RG	- <i>Residential Gateways</i> , Gateways Residenciais
RLE	- <i>Run Length Encoding</i>
RS	- <i>Redundant Slices</i> , Fatias Redundantes
RTP	- <i>Real Time Protocol</i> , Protocolo de Tempo Real
RTPC	- Rede de Telefonica Pública Comutada
RTSP	- <i>Real Time Stream Protocol</i> , Protocolo de Fluxo em Tempo Real
SCSI	- <i>Small Computer System Interface</i>
SDI	- <i>Serial Digital Interface</i> , Interface Digital Serial
SDTV	- <i>Standard Definition Television</i> , Televisão de Definição Padrão
SGML	- <i>Standard Generalized Markup Language</i>
SMPTE	- <i>Society of Motion Picture and Television Engineers</i>
SMR	- <i>Signal to Mask Ratio</i> , Razão Sinal Máscara
SMS	- <i>Short Message Service</i>
SNMP	- <i>Simple Network Management Protocol</i> , Protocolo de Gerência Simples de Rede
SOAP	- <i>Simple Object Access Protocol</i>
SPTS	- <i>Single Program Transport Stream</i> , Fluxo de Transporte de Programa Único
STB	- <i>Set Top Box</i> , Decodificador
STC	- <i>System Time Clock</i> , Relógio de Tempo do Sistema
SVOD	- <i>Subscription Video on Demand</i> , VOD por Subscrição
TCP	- <i>Transmission Control Protocol</i> , Protocolo de Controle de Transmissão
TSTV	- <i>Time Shift TV</i> , TV Deslocada no Tempo
UDP	- <i>User Datagram Protocol</i>

- UEP - *Unequal Error Protection*, Protecção de Erro Desigual
- URL - *Uniform Resource Locator*, Localização do Recurso
- VBR - *Variable Bit Rate*, Taxa de Bits Variável
- VDSL - *Very High Speed Digital Subscriber Line*, Linha de Assinante Digital de Alta Velocidade
- VDSL2 - *Very High Speed Digital Subscriber Line Second Generation*, Linha de Assinante Digital de Alta Velocidade de Segunda Geração
- VLC - *Variable Length Coding*
- VOD - *Video on Demand*, Vídeo sob Demanda
- VoIP - *Voice over Internet Protocol*, Voz sobre IP
- VTR - *Video Tape Recorder*
- WiMAX - *Worldwide Interoperability for Microwave Access*, Acesso Micro-Ondas de Banda Larga sem Fio
- XML - *Extensible Markup Language*, Linguagens de Marcação

Publicações

1. Marcus Pereira Magri, Yuzo Iano. “How to architect an IPTV system”. Revista IEEE Transactions on Circuits and System I, 2007. TCAS-I 4423 (submetido)
2. Marcus Pereira Magri, Yuzo Iano. “How to architect an IPTV System”, 8th International Workshop on Image Analysis for Multimedia Interactive Services, Grécia, Junho, 2007. (publicado)
3. Marcus Pereira Magri, Yuzo Iano. “Visão Geral da Solução de IPTV”. Rev. Ciência e Tecnologia Unisal. São Paulo, SP, Edição 16, Junho, 2007, ISSN 1677-9649. (publicado)
4. Marcus Pereira Magri, Yuzo Iano. “Como arquitetar uma solução de IPTV”. Telecomunicações, Inatel, 2007. #50 (submetido)
5. Yuzo Iano, Marcus Pereira Magri. “IPTV Solution Description”. Revista SBrT - Sociedade Brasileira de Telecomunicações. 2007. JCIS_Regular_2006_005 (submetido)
6. Marcus Pereira Magri, Yuzo Iano, Miguel Ángel Triguero Romero, “Cómo diseñar una solución de IPTV”. Revista IEEE América Latina, 2007. Paper ID: 402 (submetido)
7. Marcus Pereira Magri, Yuzo Iano. “Consumer IP Video”. Revista IEEE Transactions on Multimedia, 2007. MM001724 (submetido)
8. Marcus Pereira Magri, Yuzo Iano. “Description of IPTV solution”. Revista IEEE Transactions on Consumer Eletronics, 2007. #3435 (submetido)

Página em branco

Capítulo 1

Introdução

IPTV (televisão sobre protocolo de *Internet*), também conhecida com *triple play*, é uma solução capaz de oferecer aos clientes voz, dados e vídeo sobre uma simples conexão de banda larga que pode ser *ethernet*, FTTH (*Fiber To The Home* ou fibra óptica até a casa do usuário), ou mais tipicamente hoje em dia DSL (*Digital Subscriber Line* ou linha de assinante digital). A voz pode ser ou POTS (*Plain Old Telephone Service* ou serviço de telefonia tradicional) ou VoIP (Voz sobre IP ou *Voice over IP*). Serviço de dados é tipicamente uma conexão de alta velocidade via *Internet*. Vídeo inclui todos os serviços tradicionais de TV, bem como serviços avançados que são significativamente melhores que os atuais serviços de TV por cabo, ou por satélite, oferecendo serviços como vídeo sob demanda (VOD ou *Video on Demand*), gravação pessoal de vídeo na rede, bem como serviços convergentes com integração de voz, dados e vídeo. Além disso, a solução oferece uma interface intuitiva com mudanças instantâneas de canal e múltiplos PiPs (mosaico de canais) para aumentar a satisfação do cliente visando torná-lo mais fiel.

O verdadeiro *triple play* é aquele capaz de oferecer os três serviços integrados e transparentes. O termo integrado significa que os três serviços devem ser oferecidos de uma plataforma comum (IP) e gerenciados também através do mesmo sistema de gerência. Entregues de forma transparente, isso significa que um serviço pode tirar proveito do outro. Por exemplo, enquanto um usuário assiste a TV, se ele recebe uma chamada telefônica, a mesma poderia ser mostrada na tela da TV, indicando o número e o nome do chamador. Esse usuário poderia transferir a chamada para caixa postal através de um simples toque no controle remoto.

Devido ao aumento da competição e pressões de preços, muitos provedores de serviço estão agora procurando novas formas de aumentar a receita, através de serviços de valores agregados, usando investimentos já feitos na infra-estrutura de banda larga.

IPTV para os provedores de serviço significa primeiramente uma maneira de reter a receita dos clientes de voz e dados. Os lucros obtidos do serviço de vídeo são a segunda

razão para se investir nesta tecnologia. A capacidade de oferecer todos os serviços em uma única fatura é, sem dúvida, outra grande vantagem para uma operadora. Tal fato vai acirrar ainda mais a batalha de preços com seus concorrentes e favorecer o usuário final. Entretanto, as operadoras devem se diferenciar das companhias de TV a cabo e provedores de *Internet*, fornecendo novos serviços até hoje inexistentes, além de oferecer novos conteúdos.

Operadoras de IPTV devem assegurar um atraente catálogo de filmes sob demanda, conteúdos ao vivo exclusivos, como esportes, múltiplos ângulos, seleção de fotos e vídeos da família e programas em línguas estrangeiras, de forma a competir com os serviços *Pay Per View* (PPV) da TV paga.

É muito mais difícil para uma operadora existente, especialmente TV por satélite, oferecer serviços integrados, como a habilidade de ver a guia de TV no celular quando longe de casa, ou ter a identificação do usuário chamador (*Caller ID*) mostrado na tela da TV, ou o pedido de gravação de um conteúdo via SMS ou *Internet*.

Os fatores que contribuem para alavancar a solução de IPTV incluem os avanços na compressão de vídeo; tecnologias de gerência dos direitos digitais (DRM - *Digital Rights Management*); interfaces e protocolos abertos para integrar com os sistemas legados; padrões abertos; protocolo IP *multicast*; qualidade de serviço para protocolo IP (QoS ou *Quality of Service*); e o crescimento das implementações de redes de banda larga baseada no protocolo IP.

Nesta dissertação, apresenta-se uma descrição detalhada do que é IPTV, todos os aspectos críticos para arquitetar uma nova solução de IPTV, os alavancadores e os desafios para essa solução emergente, bem como a descrição dos componentes, além de se propor algumas sugestões novas para uma solução do tipo abordado.

Áreas de concentração

Este trabalho aborda diversas áreas de conhecimento, incluindo as áreas de compressão de vídeo, compressão de áudio, MPEG-2 *Transport Stream*, redes IP, sistemas multimídia, tecnologias de IP *Multicast*, criptografia, base de dados e servidores *web*.

História

Existem algumas empresas de comunicações começando a oferecer IPTV no mundo. Os três maiores provedores de serviços de IPTV no mundo são a China, com a PCCW's que conta com 450 mil usuários, seguido da Telefônica da Espanha com 350 mil clientes e a Belgacom da Bélgica com 110 mil usuários. Entretanto, o modelo apresentado nesta dissertação difere em alguns aspectos daqueles. A maior diferença é o fato de que eles

ainda trabalham com a compressão de vídeo MPEG-2.

Há dois anos, parecia que a chamada IPTV era a saída para as operadoras de telecomunicações completarem sua oferta de telefonia, vídeo e *Internet*.

A maior barreira, no entanto, é regulatória. Como elas só poderiam oferecer vídeo sob demanda, o modelo de negócios não se sustenta. O consumidor dificilmente assinaria um serviço de programas *à la carte* das operadoras de telefonia e outro com grade de canais de uma segunda empresa. Uma caixa híbrida ajudaria a resolver o nó regulatório, que inclui, além de impedimentos por escrito, uma demora da Agência Nacional de Telecomunicações (ANATEL) em aprovar ou rejeitar a compra de empresas de TV paga por operadoras de telecomunicações. Neste caso, os canais de televisão poderiam vir por satélite e o vídeo sob demanda por IPTV.

Motivações e Objetivos do Trabalho

As constantes inovações nas tecnologias redirecionam de forma rápida e natural o mercado das telecomunicações, destruindo mercados antigos e criando outros novos no seu lugar. Como exemplo, a solução de VoIP tem ajudado a destruir o mercado de telefonia de longa distância. Similarmente, a tecnologia de celular tem reduzido a dependência dos usuários da telefonia fixa. Como VoIP e a telefonia celular têm crescido, o faturamento das operadoras locais e de longa distância tem reduzido. Devido ao crescimento destas duas tecnologias, as companhias de telefonia fixa têm de procurar novas maneiras para recuperar o faturamento.

A boa notícia é que, mesmo com a queda do faturamento do mercado de telefonia, tecnologias emergentes têm criado novas oportunidades para as operadoras.

Avanços na tecnologia DSL permitem um aumento na largura de banda, acoplado com o progresso da compressão de vídeo, como o H.264/MPEG-4 AVC e o Windows Media 9 (VC-1). Tais padrões propiciam às companhias de telefonia transmitir *broadcast* de vídeo sobre suas linhas de par trançado. Essas inovações tecnológicas e a competição para entregar conteúdo ao cliente estão levando ao maior investimento em infra-estrutura de rede desde o final dos anos 90.

A transmissão de vídeo para as residências exige um drástico aumento da largura de banda. Uma vez que as redes DSL atuais foram projetadas essencialmente para *Internet*, faz com que as operadoras de telefonia tenham que atualizar de forma significativa sua rede de acesso. Esse ciclo de atualização na rede de acesso exige também atualizações no núcleo da rede IP, a fim de suportar um potencial maior em número de usuários de TV.

Uma plataforma baseada em IP permite significativas oportunidades para fazer a experiência do assinante de TV mais interativa e personalizada. Outra vantagem de uma

rede IP é a oportunidade para integração e convergência. Serviços convergentes implicam em uma integração dos serviços existentes de forma transparente para criar novos serviços de valor agregado.

Dessa forma, os objetivos desta dissertação estão voltados principalmente para o estudo de uma solução de IPTV fim a fim e de como arquitetar essa solução, cobrindo todos os aspectos críticos e finalmente sugerindo melhorias nessa implementação. De acordo com várias fontes de pesquisas, a solução de IPTV é bastante promissora para as operadoras de telecomunicações.

TV a Cabo x IPTV

Tecnicamente, a televisão por cabo envolve a distribuição de um número limitado de canais de televisão coletados em um local central, para assinantes dentro de uma comunidade através de uma rede de fibra óptica e/ou cabos coaxiais e amplificadores de banda larga como, mostra a Figura 1.1.

A distribuição é feita usando-se diferentes frequências para cada canal. Essa distribuição permite que muitos canais sejam distribuídos simultaneamente através do mesmo cabo. O sintonizador da TV ou o receptor de TV seleciona um canal de seu sinal multiplexado.

Como se pode ver na Figura 1.1 Sistema de TV a cabo 1.1, o mesmo sinal é enviado para todas as residências e o equipamento do cliente é que se encarrega de bloquear e filtrar os canais de acordo com o pacote do assinante.

Essa distribuição utiliza amplificadores de sinal ao longo do caminho, mas o sinal que trafega por ele é o mesmo. Além disso, uma vez que o sinal é modulado em frequência, o cabo permite um número limitado de canais.

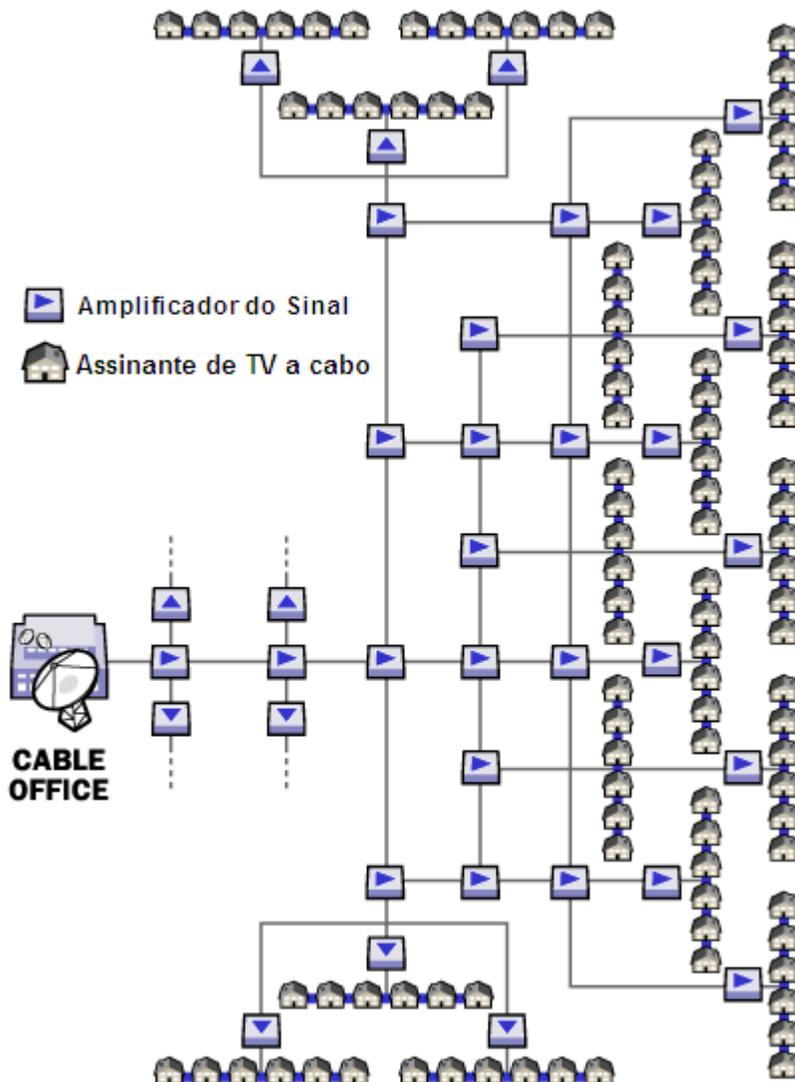


Figura 1.1 Sistema de TV a cabo.

Já a distribuição de TV por IPTV é substancialmente diferente da distribuição de TV por cabo ou satélite. Como mostrado na Figura 1.2 Comparação entre TV a cabo e IPTV, onde cada canal de TV tem uma cor diferente. No caso do DSL, somente o canal que está sendo assistido é transmitido no par trançado para a casa do assinante. Quando o usuário muda de canal, o antigo canal deixa de ser transmitido para que o novo canal possa ser recebido.

Por outro lado, no caso da TV a cabo, todos os canais são distribuídos simultaneamente para todos os usuários.

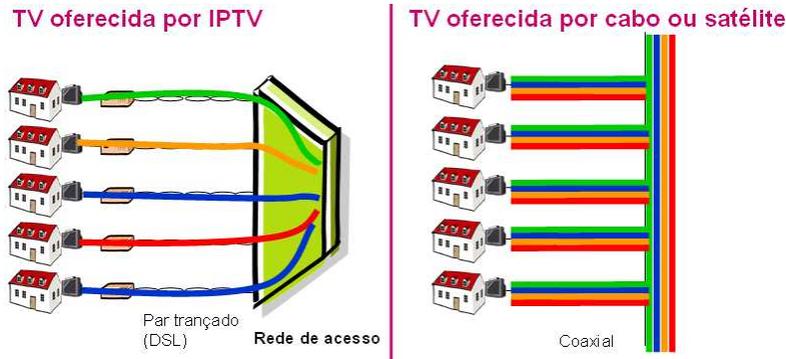


Figura 1.2 Comparação entre TV a cabo e IPTV.

Como se pode perceber, a rede de TV a cabo é orientada a *broadcast* e todos os canais trafegam por todos os pontos ao mesmo tempo, independentemente do canal escolhido pelo usuário. Além disso, a rede de TV a cabo é unidirecional, o que impossibilita uma interatividade. Muitas operadoras estão atualizando suas redes, tornando-a bidirecional e oferecendo outros serviços como o de *Internet* via cabo. Entretanto, como todos os sinais trafegam ao mesmo tempo no cabo, isso impossibilita oferecer um número maior de canais ou mesmo VOD, diferente do sistema de IPTV.

Desafios para implementar IPTV

Empresas de telecomunicações devem oferecer significativos avanços na experiência de como se assiste TV para ter uma vantagem competitiva sustentada. Na medida em que usuários adquiram novas formas de entretenimento de vídeo e serviços mais personalizados, interativos e acessíveis, operadoras de telefonia devem oferecer serviços únicos para preencher uma lacuna no mercado.

Existem alguns desafios sérios na implementação dessa solução, mas o serviço de IPTV tem conseguido vencer todos eles. O primeiro é comprimir um sinal de vídeo de forma mais eficiente do que o atual MPEG-2. Um canal MPEG-2 precisaria de cinco a seis Mbps para SDTV e cerca de 19 Mbps para HDTV.

Existe também o problema do canal único. Ao contrário do sistema de TV a cabo, onde todos os canais fluem através do cabo coaxial, em IPTV apenas um canal flui no fio de telefone por vez devido às restrições de largura de banda. Dessa forma, o tempo de mudança de canal (*zapping time*) é um fator bastante crítico em IPTV. Em geral, esse tempo depende de vários parâmetros: tempo de processamento dos comandos, tempo de atraso de rede, tempo de atraso pelo *Set Top Box* (STB), tempo de armazenamento do *buffer* do STB e tempo de decodificação de vídeo. O tempo de processamento dos comandos é o intervalo de tempo entre a ação do controle remoto e a transmissão da

mensagem IGMP *join* (comando *multicast* que será explicado posteriormente). O tempo de atraso de rede é o intervalo de tempo entre transmissão de mensagens *join* e a recepção do primeiro pacote *multicast* do canal requerido. O tempo de atraso pelo STB é o intervalo de tempo necessário pelo STB para processar os pacotes IPs recebidos e entregar o conteúdo para o processo de decodificação MPEG.

Com uma previsão de mercado mundial para 72 milhões de assinantes em 2010, uma IPTV oferece significativas oportunidades para se captar novas receitas (Fonte IDC).

Provedores de TV a Cabo têm atacado agressivamente o mercado residencial adicionando serviços de voz junto com a sua oferta de vídeo além de *Internet* de banda larga. Dada a força das operadoras de TV a cabo e aos avanços de tecnologias como voz sobre IP (VoIP), provedores de serviços de telecomunicações devem responder de maneira rápida a essa competição, lançando serviços de IPTV e aplicações multimídias.

Outro fator importante é a capacidade de integrar tudo isso com os sistemas existentes de OSS/BSS (*Operating Support System/Billing Support System* ou Sistema de Apoio à Operação/Sistema de Apoio à Tarifação).

Outra característica crucial que vem à tona com o IPTV é a garantia de alta disponibilidade, escalabilidade e segurança, especialmente para se bloquear ataques de DOS (*Denial-Of-Service*), ataque massivo ao serviço na tentativa de tirá-lo de operação.

Organização da Dissertação

Todas as informações envolvidas no desenvolvimento desta dissertação estão descritas em cada um dos capítulos que compõe a presente pesquisa.

No Capítulo 1 apresenta-se uma introdução e a estrutura da dissertação.

No Capítulo 2 apresentam-se conceitos, definições básicas e várias tecnologias envolvidas em uma solução de IPTV. Abordam-se as compressões de vídeo MPEG-2, H.264 e VC-1 e as compressões de áudio MPEG Áudio, AC-3 e AAC. Enfocam-se também as tecnologias de transporte MPEG-2 TS, IP *multicast*, bem como protocolo RTSP (*Real Time Stream Protocol*).

No Capítulo 3 apresenta-se uma solução de IPTV fim a fim, incluindo as áreas do *head-end*, inserção de comercial, *middleware*, servidor de vídeo, proteção de conteúdo, arquitetura de rede de acesso, agregação e transporte, qualidade de serviço, medição de audiência e serviços finais de IPTV.

No Capítulo 4 propõem-se algumas melhorias e uma maneira de se implementar algumas funcionalidades e serviços, como por exemplo, na área do *head-end*, foi proposta uma forma de se implementar uma inserção de legenda e uma inserção de comerciais. Na área do *middleware* e de redes, propõe-se uma maneira de se diminuir o tempo para troca

de canal.

No Capítulo 5 apresentam-se as conclusões finais e sugestões para trabalhos futuros.

Resumo das contribuições

As seguintes contribuições foram feitas nesta tese:

- Detalhamento do funcionamento dos blocos funcionais do sistema de IPTV;
- Detalhamento das tecnologias envolvidas em uma solução de IPTV;
- Proposta de algumas melhorias para uma solução de IPTV;
- Publicações e submissões de artigos pertinentes (relacionados anteriormente).

Capítulo 2

Compressão de Vídeo e Áudio, Fluxo de Transporte MPEG, IP *Multicast* e RTSP

Neste capítulo são apresentados os principais padrões para compressão de áudio e vídeo, o padrão de transporte MPEG, bem como os diferentes métodos de transporte em IP que são o IP *Multicast* e o RTSP (*Real Time Stream Protocol* ou protocolo de fluxo em tempo real). Tais padrões são utilizados pelo sistema de IPTV.

Compressão de vídeo

Para que um sinal de áudio/vídeo possa ser transmitido em uma rede, é necessário que ele seja codificado. Existem muitos esquemas de codificação tanto para áudio quanto para vídeo, com diferentes características de desempenho, complexidade e tempo de processamento. Comparando-se com os fluxos de texto e gráficos, o fluxo de um sinal de vídeo gerado por uma aplicação típica pode ser considerado de grande volume e, por isso, a compressão de dados, especialmente de áudio e vídeo, é necessária para se otimizar a utilização da largura de banda existente nas redes atuais e limitar a demanda por maior capacidade de armazenamento e transferência.

Os padrões mais importantes para compressão de vídeo são a família MPEG (*Moving Picture Experts Group*), estabelecida pela ITU (União Internacional de Telecomunicação). Os codificadores MPEG oferecem três grandes vantagens sobre os demais esquemas existentes: a) compatibilidade universal, b) taxas de compressão grandes e c) perda aceitável de qualidade na imagem final exibida.

O padrão MPEG trata separadamente vídeo e áudio, especificando como esses sinais são associados e sincronizados, possuindo assim três níveis: a camada de sistema, a camada de vídeo e a camada de áudio. A camada do sistema contém as informações sobre sincronização, acesso aleatório, administração de *buffers* e uma marcação de tempo (*time*

code) para cada quadro de vídeo.

O objetivo de um sistema para compressão de vídeo é reduzir a taxa de transmissão, removendo a redundância do sinal antes da transmissão.

MPEG-2

O grupo MPEG (*Moving Picture Experts Group*) iniciou seus trabalhos em 1988 como um grupo de trabalho ISO (*International Standards Organization*) com o objetivo de definir padrões para compressão digital de sinais de áudio e vídeo. O grupo tomou como base o padrão para videoconferência JPEG (*Joint Photographic Experts Group*), hoje conhecido como H.261, o qual foi inicialmente desenvolvido para comprimir imagens estáticas tais como aquelas usadas em fotografia eletrônica [9, 35, 54].

O primeiro objetivo de grupo MPEG foi definir um algoritmo de codificação de vídeo para armazenamento em meio digital, em particular o CD-ROM. O padrão resultante foi publicado em 1993 e compreendia de três partes:

1. Aspectos de sistema (incluindo multiplexação e sincronização);
2. Codificação de vídeo;
3. Codificação de áudio.

O padrão chamado MPEG-1, foi aplicado no sistema *Interactive CD* (CDi) para possibilitar a execução de *playback* em CDs. O sistema CDi foi largamente utilizado em aplicações PC para as quais existe uma gama de codificadores de *hardware* e *software*. O padrão é restrito a formatos de vídeo não entrelaçados e, primeiramente, tinha como objetivo a codificação de vídeo a taxas de bits não maiores que 1,5 Mbps.

Em um sinal de vídeo entrelaçado, a varredura das linhas é tal que as linhas pares são percorridas primeiramente (*even field*) seguidas da varredura das linhas ímpares (*odd field*). O sinal de vídeo entrelaçado visa reduzir o efeito de "cintilação" no visual humano.

Em 1990 iniciaram-se os estudos para um segundo padrão de vídeo que tinha como objetivo ser capaz de codificar imagens entrelaçadas e, originalmente, suportar aplicações de alta qualidade a taxas de transmissão entre 5 e 10 Mbps. Esse segundo padrão, que foi denominado MPEG-2, também suporta formatos de alta definição a taxas de bits no intervalo de 15 a 30 Mbps. Assim como no padrão MPEG-1, o padrão MPEG-2 publicado em 1994, compreende três partes: sistemas, vídeo e áudio.

É importante notar que os padrões MPEG especificam somente a sintaxe, a semântica das seqüências de bits e o processo de decodificação. Não especificam, no entanto, o processo de codificação, ficando livre a proposta de novas técnicas de codificação que visem melhorar o desempenho do sistema.

Um sistema de compressão de vídeo visa reduzir a taxa de transmissão removendo a

redundância e/ou informações de menor importância do sinal antes da transmissão. Tal sistema é implementado pelo codificador de fonte de um transmissor digital. No receptor, o decodificador de fonte reconstrói uma aproximação da imagem a partir da informação remanescente após o processo de compressão.

Uma seqüência de vídeo digital é um conjunto de quadros ordenados temporalmente. Cada quadro é uma matriz cujos elementos contêm informações de cor e representam uma imagem. Os pontos dessa matriz são chamados de *pixels*. No padrão MPEG-2 as cores são representadas através de uma combinação de três componentes, uma de luminância, Y, e duas de cromaticidade, Cr e Cb. A luminância representa a intensidade, ou seja, a imagem em preto-e-branco (brilho da cena). As componentes de cromaticidade contêm a informação de cor (associada a matiz e saturação em sistemas convencionais - nesse caso, fase e módulo do sinal de cromaticidade modulado em quadratura). Em HDTV (*High Definition Television* ou televisão de alta definição), têm-se os dois sinais diferenças de cores Cr e Cb que podem ser associados ao sinal de cromaticidade.

Devido às características psico-visuais humanas, é possível subamostrar os sinais de cromaticidade sem perda de qualidade perceptível. Isso ocorre porque o sistema visual humano é menos sensível às altas freqüências dos sinais de cromaticidade do que às altas freqüências do sinal de luminância. Por isso, os padrões HDTV utilizam o formato 4:2:0, em que os sinais de cromaticidade são subamostrados por um fator de 2 nas duas direções, horizontal e vertical. A subamostragem deve ser precedida de uma pré-filtragem, pois caso contrário ocorrerá uma distorção do sinal devido ao *aliasing*, que é uma sobreposição de componentes espectrais. Assim, da mesma forma que nos sistemas convencionais, os detalhes da imagem são reproduzidos pelo sinal de luminância.

A Figura 2.1 apresenta a disposição espacial das amostras de luminância e cromaticidade para o formato 4:2:0. Observa-se que as amostras de cromaticidade estão posicionadas no sentido vertical, entre as amostras de luminância e, por isso, é necessário que as amostras de cromaticidade sejam obtidas através de interpolação.

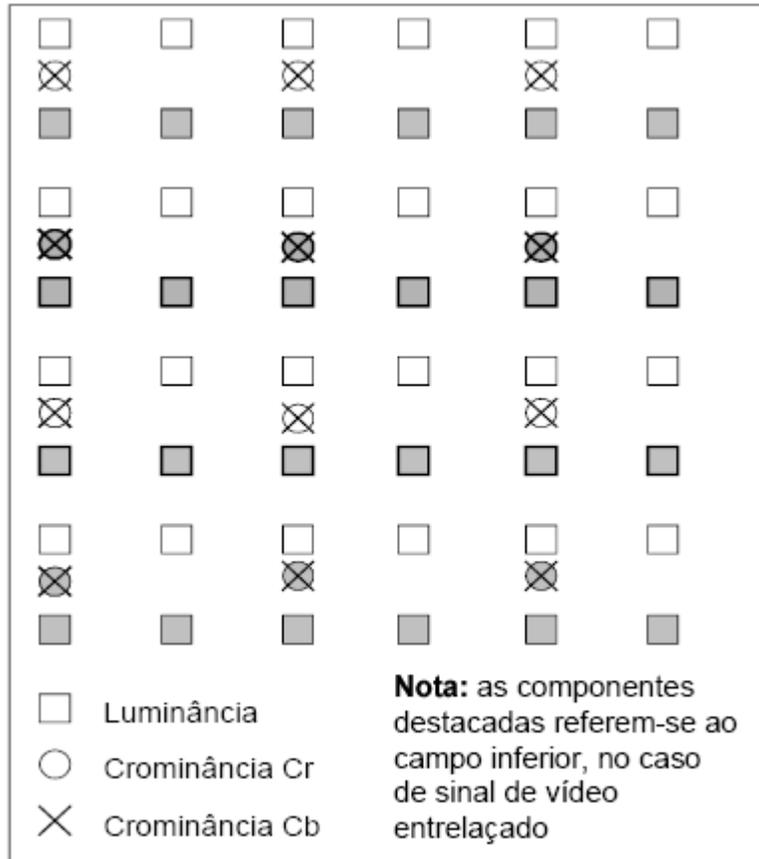


Figura 2.1 - Localização espacial das amostras de luminância e crominância para o formato 4:2:0.

Tipicamente, as seqüências de vídeo apresentam uma alta correlação, tanto temporal quanto espacial. Dessa forma, o padrão MPEG-2 comprime seqüências de vídeo digitalizadas, basicamente explorando os três tipos de redundância presentes nessas seqüências:

1. **Redundância temporal e espacial:** valores de *pixels* não são independentes, mas são correlacionados com seus vizinhos, tanto dentro do mesmo quadro (redundância espacial) através da transformada DCT (*Discrete Cosine Transform* ou transformada discreta de cosseno) quanto entre quadros consecutivos (redundância temporal) através da codificação preditiva entre quadros. Assim, dentro de alguns limites, o valor de um *pixel* pode ser predito a partir dos valores dos *pixels* vizinhos, assim como regiões de um quadro futuro podem ser preditas a partir do quadro atual.
2. **Redundância em entropia:** para qualquer sinal digitalizado não-aleatório alguns valores codificados ocorrem com mais freqüência do que outros. Essa característica pode ser explorada através da codificação dos valores que ocorrem

mais freqüentemente com códigos menores, enquanto que códigos maiores podem ser usados para valores mais raros em ocorrência (codificação por entropia - código de *Huffman*).

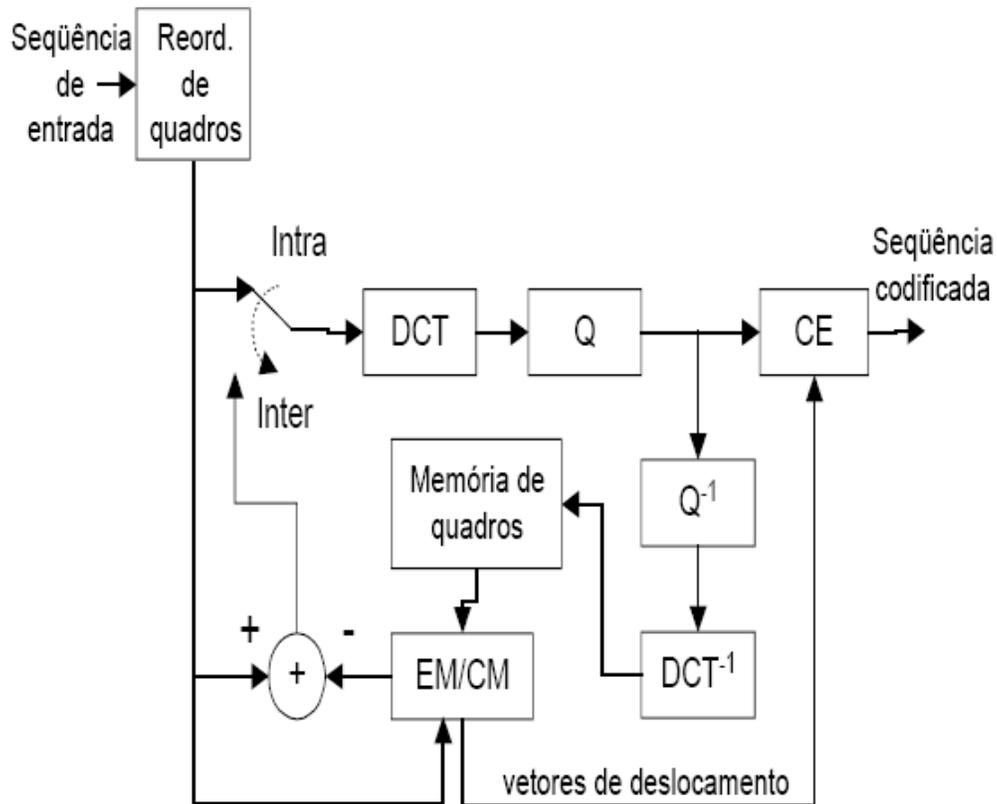
3. **Redundância psico-visual:** Essa forma de remoção de redundância resulta do princípio de funcionamento do olho e do cérebro humano (sistema visual humano) através de subamostragens da cromaticidade. Tanto os limites de definição fina de detalhes que o olho pode resolver (limites de resolução espacial) quanto os limites na habilidade de acompanhar imagens que se movem rapidamente (limites de resolução temporal) são utilizados como limites para que seja descartado aquele subconjunto do fluxo de informação de vídeo que ultrapassa esses limites. Visto que o sistema visual humano não é capaz de perceber esse tipo de informação, não há razão para que ela seja transmitida, resultando, assim, em compressão.

O diagrama geral de um codificador MPEG-2 é apresentado na Figura 2.2 - Diagrama geral de um codificador. Cada etapa desse processo será brevemente discutida; entretanto, primeiro é apresentada a estrutura adotada para a seqüência de vídeo, ou seja, a forma como os dados são divididos, a fim de serem codificados.

Cada um dos quadros da seqüência é dividido em unidades utilizadas no processo de codificação. A menor unidade de codificação no algoritmo MPEG é o bloco. Ele é utilizado na aplicação da transformada DCT, e tem tamanho 8×8 pixels. O bloco pode ser de luminância (Y), ou de cromaticidade (Cr ou Cb).

A DCT é similar à DFT (*Discrete Fourier Transform* ou transformada discreta de Fourier) e tem como propósito usar essa transformação ortogonal para ajudar a remover redundância espacial por meio da concentração de energia do sinal em relativamente poucos coeficientes.

O macro-bloco é a unidade básica de codificação no algoritmo MPEG. Ele consiste de segmentos de 16×16 pixels, e é a unidade utilizada nas etapas de estimação e compensação de movimento. Na Figura 2.3 - Subamostragens de cromaticidade MPEG-2, é apresentado um exemplo de macro bloco de formato 4:2:0, composto de quatro blocos de luminância e dois de cromaticidade. O formato 4:2:0 refere-se à forma como subamostra-se a cromaticidade do sinal, nesse caso, a cromaticidade é subamostrada por um fator de dois nas duas direções, horizontal e vertical. Outros dois formatos são também possíveis: 4:2:2 em que a cromaticidade é subamostrada por um fator de dois somente na horizontal, e o 4:4:4, em que a cromaticidade não é subamostrada, também vistos na Figura 2.3 - Subamostragens de cromaticidade MPEG-2.



Legenda:
 CE - Codificação de entropia (*Huffman* e corrida de zeros).
 DCT - Transformação DCT de blocos de 8x8 *pixels*
 DCT⁻¹ - Transformação DCT inversa
 Q - Quantizador
 Q⁻¹ - Quantizador inverso
 EM/CM - Estimativa e compensação de movimento

Figura 2.2 - Diagrama geral de um codificador.



Figura 2.3 - Subamostragens de crominância MPEG-2.

O *slice* é uma fatia horizontal de macro blocos e serve como unidade de re-sincronismo. Uma fatia horizontal de macro blocos de um quadro pode conter mais de um *slice*. Além

disso, podem existir áreas do quadro onde não há fatias e que, portanto, não são codificadas. Entretanto, todos os perfis (*profiles*) do padrão utilizam uma estrutura de *slice* restrita (*restricted slice structure*), em que todo o quadro é segmentado em *slices*, como no exemplo da Figura 2.4 - Estrutura de um quadro não-entrelaçado de formato 4:2:0.

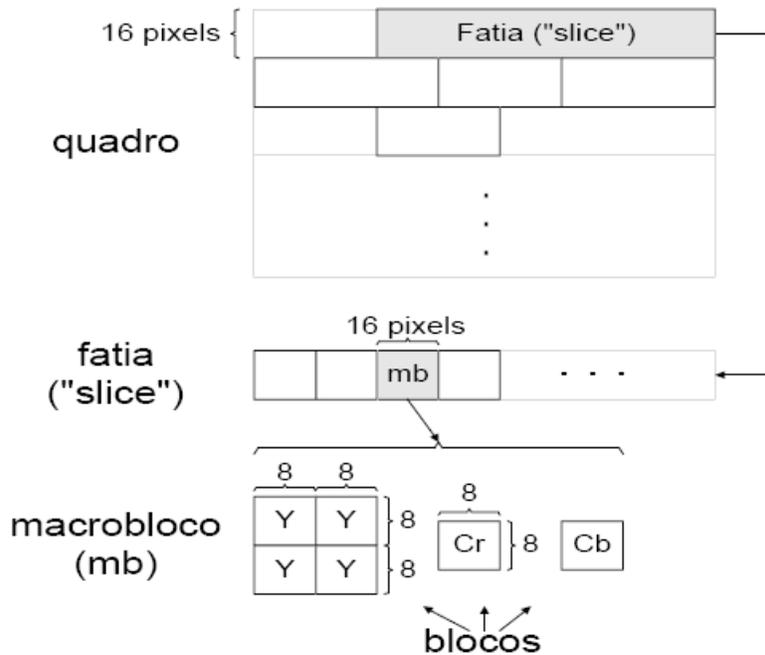


Figura 2.4 - Estrutura de um quadro não-entrelaçado de formato 4:2:0.

Os quadros também são divididos em grupos chamados GOP (*Group of Pictures* ou grupos de quadros). O GOP é um conjunto de quadros que possibilita acesso randômico. No padrão MPEG-2, o conceito de GOP é um pouco diferente do padrão MPEG-1. No primeiro caso sua utilização é opcional, enquanto no segundo, é obrigatória. Tipicamente, 6 a 15 quadros são agrupados em um GOP e é estabelecida uma ordem de tipos de quadros que se repete ao longo da seqüência.

Os quadros podem ser divididos em três tipos, de acordo com a codificação preditiva utilizada:

- Quadro I (intracodificado): esse quadro é codificado de maneira independente dos outros quadros de forma semelhante à utilizada pelo JPEG, isto é, codificado espacialmente e sem nenhuma referência a outro quadro. Dessa forma, o quadro do tipo I pode ser decodificado sem conhecer nenhum outro quadro. Por causa disso, em uma seqüência de quadros, o primeiro quadro é sempre o I, também conhecido com quadro âncora.
- Quadro P (predito): esse quadro é codificado em modo preditivo usando como referência quadros passados (quadros anteriores do tipo I ou P). O residual é transformado usando a DCT. Dessa forma, para decodificar esse quadro, necessita-se de informação de

outros quadros. Os quadros Ps têm cerca de 50-30% do tamanho de um quadro I.

- Quadro B (interpolado): esse quadro é codificado em modo preditivo tanto com referência a quadros passados como a quadros futuros (quadros Is ou Ps). Como contém muito pouca informação, o quadro B nunca é usado como referência para os demais. Ele tem aproximadamente 50% do tamanho do quadro P.

Cada um dos quadros Ps ou Bs é codificado tomando-se a diferença entre os quadros originais e uma estimativa desse quadro. Para o cálculo dessa estimativa, no caso de quadros Ps, o quadro de referência anterior é utilizado em um processo que envolve estimação e compensação de movimento (EM/CM). Para os quadros Bs, o processo é semelhante, só que, nesse caso, são utilizados dois quadros de referência (um anterior e outro posterior) que, após o processo de EM/CM, são interpolados para formarem o quadro estimado. Os quadros de referência são os quadros Ps ou Is temporalmente mais próximos do quadro atual, um anterior e outro posterior. Esses quadros são recuperados no codificador e armazenados na memória de quadros para serem utilizados no processo de EM/CM. Os quadros do tipo B nunca são usados como referência. A Figura 2.5 - Exemplo de estrutura temporal de seqüência, apresenta um exemplo de uma seqüência de quadros ordenados, temporalmente. Nesse exemplo, para cada quadro, é dado o tipo de predição e os quadros de referência relacionados.

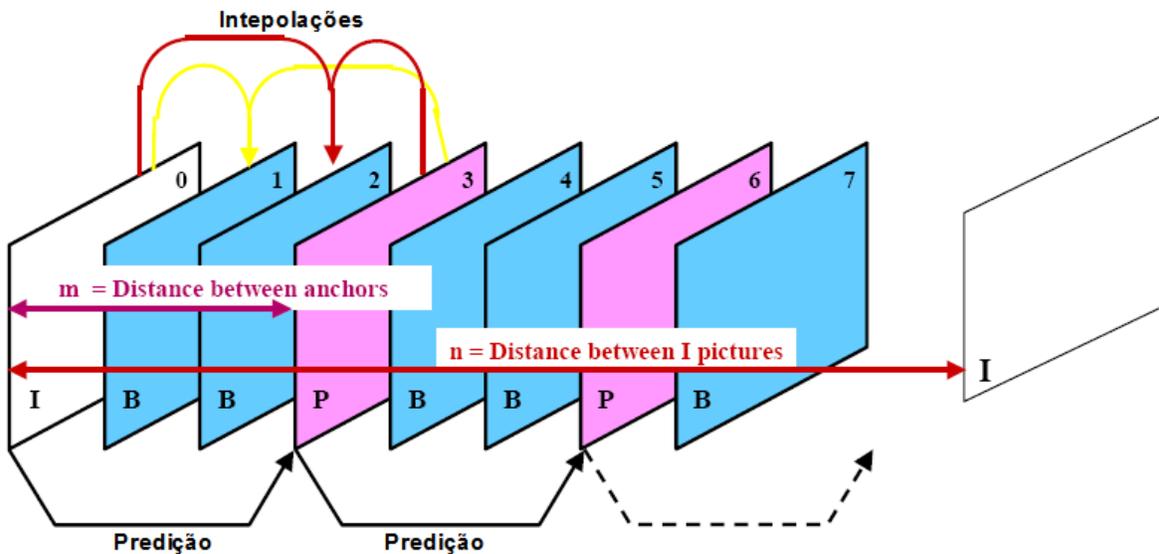


Figura 2.5 - Exemplo de estrutura temporal de seqüência (m = distância entre âncoras; n = distância entre quadros tipo I).

Os quadros a serem codificados geralmente estão dispostos em ordem temporal, ou seja, de acordo com a ordem de apresentação na tela quando a seqüência é reproduzida. Quando os quadros são codificados, eles são incluídos no fluxo de bits de acordo com a ordem em

que são decodificados, sendo esse ordenamento chamado de ordem de codificação. No decodificador é feito um novo re-ordenamento, sendo que os quadros são novamente dispostos em ordem temporal.

Quando a seqüência não contém quadros do tipo B, a ordem de codificação é igual à ordem temporal. Entretanto, se existem quadros B na seqüência, a ordem de codificação é diferente da ordem temporal, pois se devem codificar as duas referências de um quadro B antes de codificá-lo. O seguinte exemplo ilustra esta situação:

Na entrada do codificador tem-se:

I₁ B₂ B₃ P₄ B₅ B₆ P₇ P₈ B₉ I₁₀ P₁₁

Na saída do codificador e na entrada do decodificador tem-se:

I₁ P₄ B₂ B₃ P₇ B₅ B₆ P₈ I₁₀ B₉ P₁₁

Na saída do decodificador tem-se:

I₁ B₂ B₃ P₄ B₅ B₆ P₇ P₈ B₉ I₁₀ P₁₁

Para se explorar a correlação temporal existente na seqüência, utiliza-se uma técnica DPCM (*Differential Pulse Code Modulation* ou modulação por código de pulso diferencial). Basicamente, essa técnica consiste em se codificar a diferença entre o sinal e sua predição ao invés de se codificar o próprio sinal. No caso do padrão MPEG-2 essa diferença é calculada entre o macro bloco atual e sua predição, que pode ser uma região de um quadro anterior ou uma média entre regiões de um quadro anterior e outro posterior.

O modo mais simples de se calcular uma predição para um macro bloco é utilizar macro blocos de um ou dois quadros da mesma região espacial (Figura 2.6 - Codificação espacial). Essa técnica é eficiente em regiões da cena em que não há movimento. Entretanto, em áreas onde há movimento, a correlação entre *pixels* de mesma posição espacial, geralmente é pequena. Por isso, são utilizadas técnicas de estimação e compensação de movimento para tornar mais eficiente esse processo de predição (Figura 2.7 - Codificação temporal). As técnicas de estimação de movimento utilizadas pelo padrão MPEG-2 são baseadas em casamento de blocos. Para cada bloco é realizada uma busca do vetor de deslocamento dentro de uma região do quadro de referência chamada janela de busca. Essa busca consiste em minimizar uma medida de disparidade, por exemplo, o erro quadrático médio (MSE ou *Mean Square Error*), definido pela Equação (2.1).

$$MSE(i, j) = \frac{1}{M \cdot N} \sum_{m=a}^{M+a} \sum_{n=b}^{N+b} [Q_a(m, n) - Q_{ref}(m+1, n+j)]^2 \quad (2.1)$$

onde,

$Q_a(x, y)$ é o quadro atual,

$Q_{ref}(x, y)$ é o quadro de referência,

(i, j) é o deslocamento entre os dois blocos,

(a,b) é a posição do bloco atual.

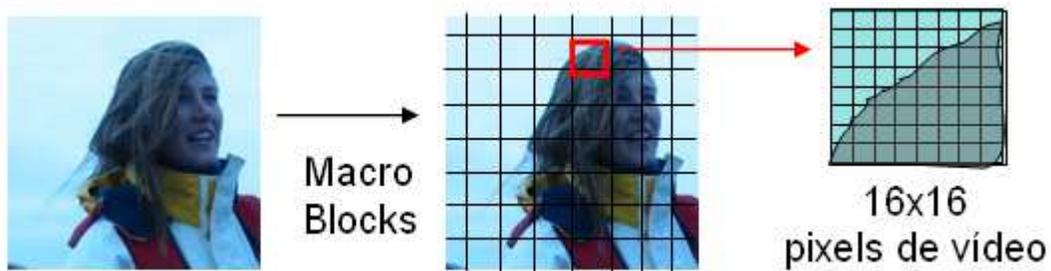


Figura 2.6 - Codificação espacial.

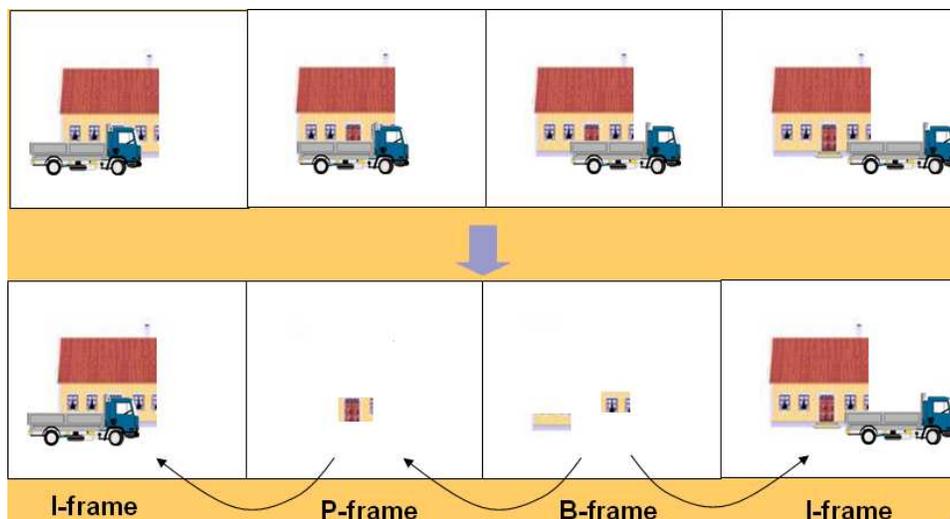


Figura 2.7 - Codificação temporal.

A Figura 2.8 - Estimação de movimento apresenta um exemplo de estimação de movimento em que o deslocamento do bloco atual em relação à mesma posição espacial no quadro de referência é representado pelo vetor \vec{d} (vetor de deslocamento).

Quanto à forma de predição utilizada, basicamente, um macro bloco pode ser codificado de quatro modos:

- Intracodificado;
- Predito utilizando-se o quadro de referência anterior;
- Predito utilizando-se o quadro de referência posterior;
- Ou interpolado (utiliza os dois quadros de referência para formar a predição).

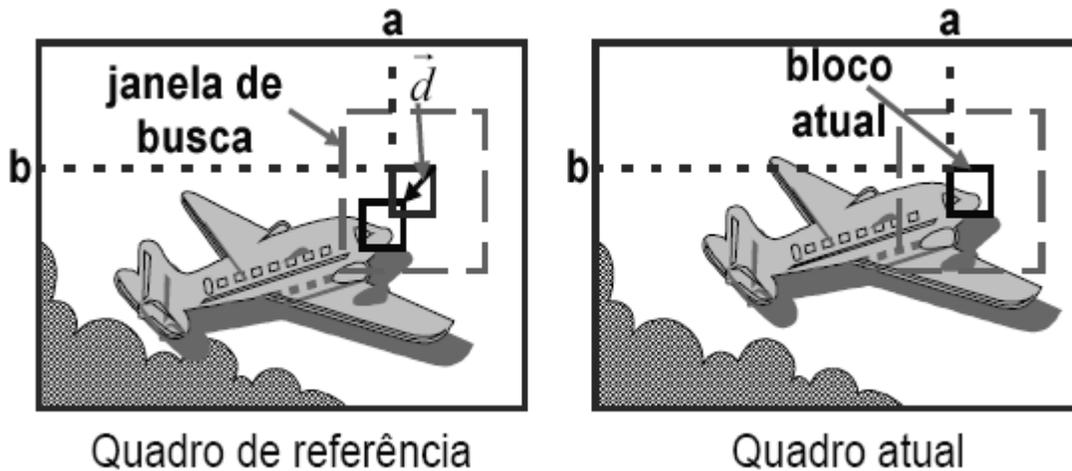


Figura 2.8 - Estimação de movimento.

O padrão MPEG-2 contempla algumas variações desses quatro modos básicos, por exemplo, o *dual-prime* e o 16x8. Nesse último, o macro bloco é dividido em dois e o processo de EM/CM é realizado para cada uma das partes de tamanho 16x8 *pixels*. Esse modo só é utilizado quando os campos do quadro são codificados individualmente.

Os processos de predição e interpolação envolvem estimação e compensação de movimento. Se o macro bloco atual é predito, seja em relação ao quadro de referência anterior ou posterior, o macro bloco do quadro de referência que apresentou melhor casamento é utilizado como predição. No caso de macro blocos interpolados, a predição é formada tomando-se a média entre os macros blocos de melhor casamento de cada um dos quadros de referência.

Os dados a serem codificados e que são referentes a um macro bloco predito ou interpolado, resultam da diferença entre esse macro bloco e a sua predição. Tem-se então um macro bloco-diferença que é codificado de maneira semelhante a um macro bloco intracodificado, utilizando-se transformada DCT e codificação de entropia (código de *Huffman* e *run-length*). Os vetores de deslocamento resultantes do processo de EM/CM são codificados e incluídos no fluxo de bits (*bitstream*).

Todos os macros blocos de um quadro I são intracodificados. Nos quadros Ps pode-se ter macro blocos preditos (em relação ao quadro de referência anterior), ou macro blocos intracodificados, o que equivale a utilizar um macro bloco nulo (todos os elementos iguais a zero) como predição. Finalmente, nos quadros Bs pode-se ter macro blocos interpolados, macro blocos preditos (em relação ao quadro de referência anterior ou posterior), ou macro blocos intracodificados. A escolha de como é formada a predição para um macro bloco fica a cargo do projetista do codificador e, geralmente, depende da correlação entre o macro bloco atual e o macro bloco de melhor casamento de cada quadro de referência. Nos quadros B e P, existe a possibilidade de não se codificar determinado macro bloco. Nesse

caso, o macro bloco é reconstruído no decodificador utilizando-se somente os quadros de referência e considerando nulos os vetores de movimento, se o quadro atual é do tipo P, ou tomando-se os vetores de movimento do último macro bloco codificado, se o quadro atual é do tipo B.

Os dados, resultantes desse processo, sejam eles correspondentes a um macro bloco original ou a um macro bloco-diferença, são divididos em blocos de tamanho 8x8. A cada bloco é aplicada uma transformação DCT bidimensional. A transformação DCT não resulta em compressão, entretanto descorrelaciona os coeficientes tornando mais eficiente o processo de compressão através de outras técnicas [54]. Após a transformação DCT, os coeficientes resultantes são quantizados. O padrão MPEG-2 só define o processo de quantização inversa, permitindo flexibilidade no projeto do quantizador.

Após a quantização, os blocos de 8x8 coeficientes quantizados são varridos em um padrão zig-zag para transformar os 64 coeficientes do bloco em uma seqüência serial de coeficientes quantizados. A varredura alternada, apresentada na Figura 2.9 - Modos de varredura dos coeficientes do bloco transformado, é utilizada quando a seqüência de vídeo é entrelaçada e o bloco codificado contém informação dos campos par e ímpar. A varredura é feita como mostrado na Figura 2.9.

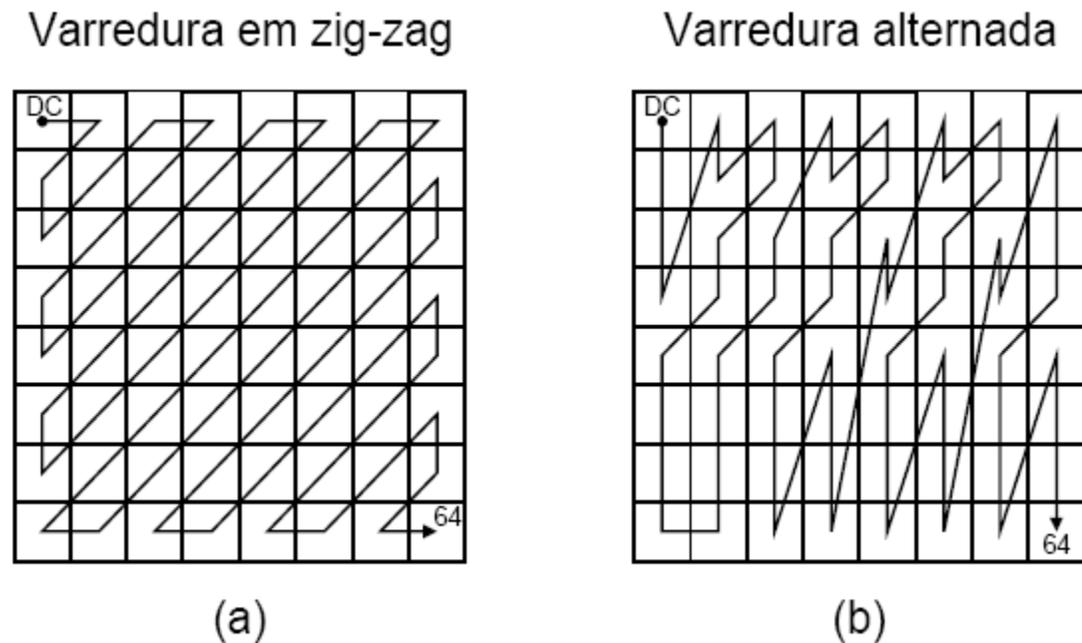


Figura 2.9 - Modos de varredura dos coeficientes do bloco transformado.

Note que, resultante da quantização, ocorrerão longas seqüências de coeficientes nulos. As seqüências de coeficientes produzidas pela varredura zig-zag são codificadas pela contagem do número r_0 de coeficientes zeros que precedem um coeficiente c^* não-zero,

processo denominado codificação RLE (*Run-Length Encoding*). A cada valor r_0 há um respectivo valor c^* associado. Cada par $\{r_0, c^*\}$ é então codificado usando-se um código de comprimento variável VLC (*Variable Length Coding*) basicamente similar ao código de *Huffman*.

Quando se utiliza uma codificação diferencial, um dos problemas que podem ocorrer é a propagação de erros. Por exemplo, se ocorre erro na decodificação de um macro bloco que é utilizado como referência para se codificar outros macros blocos, o erro vai propagar até que esse macro bloco de referência seja substituído por outro decodificado sem erros. Por isso, o padrão MPEG-2 recomenda que cada macro bloco que ocupa uma determinada posição espacial deve ser intracodificado, no máximo a cada 132 quadros.

Para os sistemas digitais recomenda-se que a atualização de quadros, ou seja, a codificação de quadros intracodificados (quadros I), ocorra em intervalos de no máximo 0,5 segundo. Além disso, recomenda-se que seja enviado o cabeçalho de seqüência antes de cada quadro intracodificado. Esses procedimentos são necessários, pois quando um usuário sintoniza um determinado canal, essas informações (cabeçalho de seqüência e quadro I) serão necessárias para o início do processo de decodificação e reprodução do sinal de vídeo. O desempenho da qualidade de mudança (troca) de canal pelo usuário pode ser bastante afetado caso esses procedimentos não sejam observados.

A utilização de códigos de comprimento variável (VLC) e codificação preditiva no processo de compressão fazem com que a taxa de bits seja variável ao longo do tempo em função da estatística do sinal de vídeo original. Como vários sistemas, incluindo o IPTV, utilizam taxas de bits constantes, é necessário um procedimento para controle de taxa. Esse controle é realizado através da utilização de um *buffer* na saída do codificador que simula o buffer da entrada do decodificador. O tamanho desse *buffer* deve ser tal que os atrasos de transmissão decorrentes da inserção desse *buffer* sejam toleráveis pelo sistema. Mecanismos de realimentação são utilizados para alterar parâmetros dos algoritmos de compressão de forma a evitar que o *buffer* do codificador não seja sobrecarregado (*overflow*), nem esvaziado abaixo de um determinado limiar (*underflow*). Os parâmetros mais utilizados para esse controle são os relacionados com os passos de quantização dos coeficientes.

A qualidade da imagem de um codificador/decodificador depende fortemente do conteúdo da imagem. Os conteúdos mais difíceis de codificar/decodificar são os que contêm cenas de ação como um filme ou jogo de basquete, uma imagem muito escura, ou muito clara, com relâmpagos, imagens com movimento de água (mar ou chuva). Entretanto, devido aos avanços dos algoritmos de compressão/descompressão, tem-se obtido excelentes resultados, com uma taxa de bits relativamente baixa, mesmo para esses conteúdos.

H.264/MPEG-4 AVC

H.264, MPEG-4-parte 10, ou MPEG-4 AVC (*Advanced Video Compression*) é o mais recente padrão de compressão e desenvolvido em conjunto pelo ITU-T VCEG (*Video Coding Experts Group*) e pelo ISO/IEC MPEG [10, 25, 54, 55]. A idéia do H.264/AVC foi a de criar um padrão capaz de proporcionar uma boa qualidade com taxas binárias notavelmente inferiores aos padrões anteriores, além de não incrementar a complexidade do seu projeto. O aumento na taxa de compressão pode chegar a 50% se comparado ao padrão MPEG-2, um dos padrões mais utilizados atualmente.

O H.264/MPEG-4 AVC representa alguns avanços na tecnologia de codificação de vídeo em termos de melhoras na eficiência da codificação e flexibilidade para o uso eficaz sobre vários tipos de redes e de aplicações. Sua codificação de vídeo é baseada nos conceitos de codificação de vídeo convencional de compensação de movimento baseado em blocos, mas com algumas importantes diferenças em relação ao padrão anterior.

Os principais blocos de um codificador H.264 são: estimação de movimento, compensação de movimento, predição intra, transformadas diretas (T) e inversas (T^{-1}), quantização direta (Q) e inversa (Q^{-1}), filtro e codificação de entropia, como mostrados na Figura 2.10 - Diagrama geral de um codificador H.264.

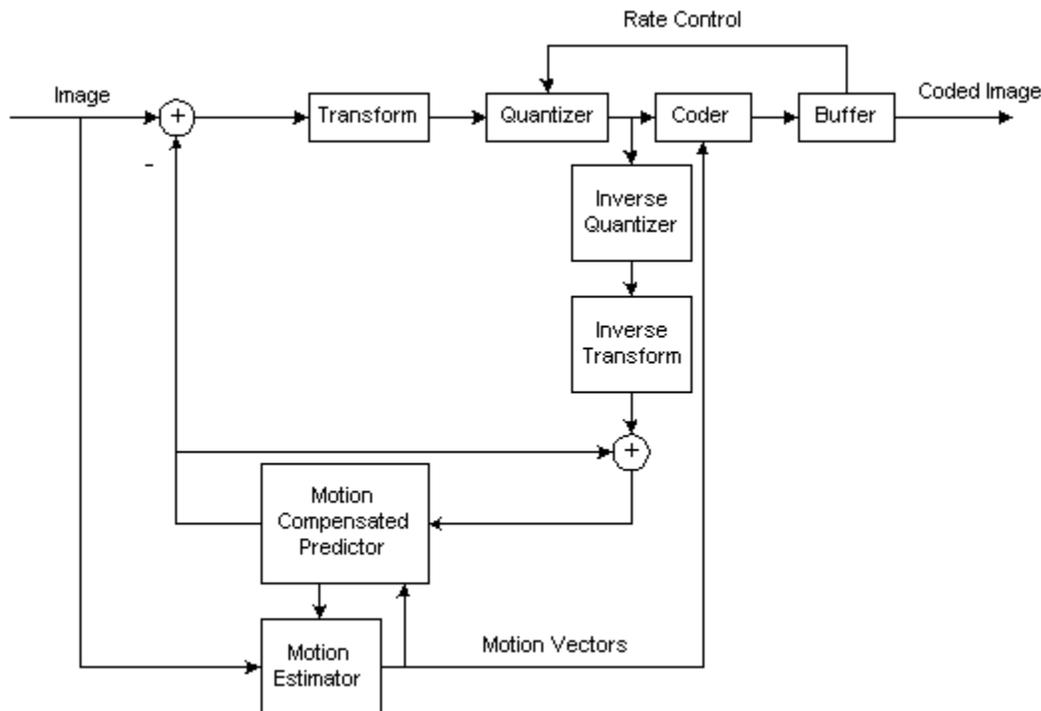


Figura 2.10 - Diagrama geral de um codificador H.264.

O H.264/MPEG-4 AVC contém algumas características que lhe permitem comprimir vídeo de uma maneira muito mais eficiente do que padrões mais antigos, e traz mais flexibilidade para aplicações em diversos ambientes de rede, como citado a seguir.

- Compensação de movimento com amostra precisa. A maioria dos padrões prévios permite calcular com exatidão o vetor do movimento. O novo padrão continua calculando esse vetor com exatidão, mas reduz a complexidade do processo de interpolação quando comparado ao MPEG-2.
- Compensação de movimento multi imagem usando imagens codificadas anteriormente como referência, permitindo usar 32 imagens de referência contra as 1 ou 2 imagens usadas em padrões anteriores. Essa característica em particular permite melhorias consideráveis na taxa binária e na qualidade de imagem.
- Compensação de movimento de bloco de tamanho variável com blocos 16×16 e 4×4 permitindo uma segmentação precisa de regiões de movimento. Esse padrão suporta mais flexibilidade na seleção de tamanhos e de formas de bloco da compensação do movimento quando comparado com o padrão precedente.
- Vetores de movimentos além dos limites do quadro. Enquanto que os vetores de movimento do MPEG-2 e de seus predecessores eram requeridos para apontar somente às áreas dentro do quadro de referência previamente decodificado, a técnica de extrapolação dos limites encontrada primeiramente como uma característica opcional no H.263 é incluída no H.264/AVC.
- Precisão de 1/4 de *pixel* para compensação de movimento, permitindo uma descrição muito precisa de deslocamentos de áreas de movimento. Para a cromaância a resolução é tipicamente reduzida à metade e, portanto a precisão de compensação de movimento se reduz para 1/8 de *pixel*.
- Precisão ponderada que permite ao codificador especificar o escalonamento quando executa compensação de movimento. Isso pode melhorar drasticamente a eficiência da codificação para cenas que contêm *fades* (*fade-to-black*, *fade-in*, e *cross-fade*).
- Codificação aritmética binária adaptável segundo o contexto (CABAC, *Context-Adaptive Binary Arithmetic Coding*), que é uma técnica inteligente para comprimir, sem perdas, elementos de sintaxe no fluxo de vídeo sabendo as probabilidades de elementos de sintaxe num dado contexto.
- Codificação de comprimento variável adaptável segundo o contexto (CAVLC, *Context-Adaptive Variable-Length Coding*), que é uma alternativa de baixa complexidade ao CABAC para a codificação de valores de coeficientes quantificados. Embora tenha uma complexidade mais baixa do que CABAC, CAVLC é mais elaborada e mais eficiente do que os métodos usados tipicamente para codificar coeficientes em outros projetos prévios.

- Ordenação de macro blocos flexível (FMO, *Flexible Macroblock Ordering*) e ordenação arbitrária de fatias (ASO, *Arbitrary Slice Ordering*): são técnicas para reestruturar a maneira como se representa a ordenação das regiões fundamentais (chamadas macro blocos) nas imagens. Tais ordenações FMO e ASO, consideradas tipicamente como fazendo parte de uma característica de robustez de erros/perdas, também podem ser usados para outras finalidades. O ASO pode melhorar o atraso em aplicações de tempo real, particularmente quando usada em rede que tem comportamento de entrega fora de ordem, como redes IP.
- Partição de dados (DP, *Data Partitioning*), uma característica que permite separar elementos mais importantes e menos importantes da sintaxe em diferentes pacotes de dados, possibilitando a aplicação da proteção de erro desigual (UEP, *Unequal Error Protection*) e outros tipos de melhoria da robustez de erros/perdas.
- Fatias redundantes (RS, *Redundant Slices*), uma característica de robustez de erros/perda que permite que um codificador emita uma representação extra de uma região da imagem (tipicamente de baixa fidelidade) para poder ser usada quando a representação preliminar estiver corrompida ou perdida.
- Um processo automático simples para impedir a emulação acidental dos códigos iniciais, que são seqüências especiais de bits nos dados codificados que permitem o acesso aleatório no fluxo de bits e a recuperação do alinhamento de bytes em sistemas que podem perder a sincronização de bytes.
- Contagem da ordem da imagem, uma característica que serve para manter a ordem das imagens e dos valores das amostras das imagens decodificadas e isoladas da informação de sincronismo (permitindo que a informação de sincronismo seja transportada e controlada/modificada separadamente por um sistema sem afetar o conteúdo decodificado da imagem).

Essas técnicas, junto com muitas outras, ajudam o H.264 a ter um desempenho significativamente melhor do que qualquer outro padrão anterior. O H.264 pode ter um desempenho radicalmente melhor do que MPEG-2 vídeo (tipicamente obtém a mesma qualidade, mas com metade, ou menos de metade, do débito binário).

VC-1

O VC-1 é uma especificação de compressão de vídeo padronizada pela SMPTE 421M em 2006 (*Society of Motion Picture and Television Engineers*) e implementada pela Microsoft como Microsoft® Windows Media® (WMV) 9 [36].

O VC-1 foi projetado para alcançar o estado da arte da compressão de vídeo com largura

de banda bem baixa ou bem alta, dependendo da resolução.

A funcionalidade básica do VC-1 envolve a compensação de movimentos baseada em blocos e o esquema da transformada espacial similar ao usado pelos outros padrões de compressão. Entretanto, VC-1 inclui um número de inovações e otimizações que o diferem dos outros esquemas de compressão, resultando em uma excelente qualidade e eficiência.

A seguir apresenta-se uma breve descrição de algumas dessas inovações:

-Transformada de tamanho do bloco adaptativo: Tradicionalmente, as transformadas 8×8 têm sido usadas para a codificação de imagens e vídeos. Entretanto, existe uma evidência que sugere que transformadas 4×4 podem reduzir artefatos nas bordas e descontinuidades. VC-1 é capaz de codificar um bloco 8×8 usando uma transformada 8×8 , duas transformadas 8×4 , duas transformadas 4×8 , ou quatro transformadas 4×4 . Essa característica permite que a codificação tire proveito dos diferentes tamanhos de transformada para uma melhor qualidade de imagem.

-De forma a minimizar a complexidade computacional do decodificador, VC-1 usa transformadas de 16-bits. Isso também tem a vantagem de facilitar a implementação do *hardware* do decodificador fazendo com que ele seja tão eficiente quanto possível.

-A compensação de movimento é o processo de gerar a predição do quadro de vídeo a partir do quadro de referência. Tipicamente, a predição é formada por blocos de dados de 8×8 ou macro blocos de 16×16 . O deslocamento dos dados devido ao movimento é definido por um vetor de movimento, que captura o deslocamento ao longo dos eixos X e Y.

-A eficiência do codificador é afetada pelo tamanho do bloco predito, pela granularidade dos dados que podem ser capturados, e pelo tipo de filtro usado para gerar a predição. VC-1 usa blocos de 16×16 para a predição, com a habilidade de gerar vários quadros de blocos de 16×16 e 8×8 . A granularidade mais fina suportada pelo VC-1 é de $1/4$ de *pixel*.

-VC-1 combina os ajustes do vetor do movimento definidos pelo tamanho de bloco, pela granularidade e pelo tipo do filtro. O resultado é quatro modalidades de compensação de movimento que servem para diversas situações.

-*Loop Filtering*: VC-1 usa filtro de desbloqueio *in-loop* que tenta remover descontinuidades das bordas introduzidas por erros de quantização nos quadros interpolados. Essas descontinuidades podem causar artefatos visíveis nos quadros decodificados sendo que elas podem impactar a qualidade do quadro como um preditor para um quadro interpolado futuro.

-O filtro *in-loop* leva em consideração a transformada de tamanho de bloco de forma adaptativa. O filtro é também otimizado para reduzir o número das operações requeridas.

-Codificação entrelaçada (*Interlace Coding*): o conteúdo de vídeo entrelaçado é usado extensamente na transmissão de TV. Ao codificar o conteúdo entrelaçado, o codificador VC-1 pode tirar vantagem das características de quadros entrelaçados para melhorar a

compressão. Isto é alcançado usando dados de ambos os campos para predizer a compensação de movimento nos quadros interpolados.

-Codificação avançada do quadro B: O VC-1 inclui diversas otimizações que fazem os quadros tipo B mais eficientes.

-Devido à natureza da compressão que usa a compensação do movimento, a codificação dos quadros de vídeo que contêm efeito de desvanecimento (*fade*) para ou do preto é muito ineficiente. O VC-1 inclui a compensação desvanecimento (*fading*) e quando detecta esse efeito, faz uso de métodos alternados para ajustar a luminância. Essa característica melhora a eficiência da compressão para seqüências com desvanecimento e outras mudanças globais de iluminação

-As taxas de compressão vídeo para MPEG-2, H.264, e VC-1 para canais de definição padrão e alta definição, são mostradas na Tabela 2.1 - Taxa de Compressão de Vídeo.

	SDTV	HDTV
MPEG-2	2 - 4 Mbps	16 - 19 Mbps
H.264	1,5 - 2 Mbps	6 - 8 Mbps
WM9 (VC-1)	1,5 - 2 Mbps	6 - 8 Mbps

Tabela 2.1 - Taxa de Compressão de Vídeo.

Compressão de Áudio

O objetivo desta sessão é fornecer uma visão geral dos aspectos mais importantes da compressão de áudio e das suas técnicas de codificação utilizadas.

MPEG Áudio

As normas que descrevem a codificação de áudio MPEG-1 e MPEG-2 são a ISO/IEC 11172-3 e a ISO/IEC 13818-3 respectivamente [8, 27, 28, 29, 30, 31, 35]. Note que assim como a compressão de vídeo MPEG-2, MPEG não define os modelos dos codificadores, ela somente define o formato da seqüência de bits e o modelo de decodificação de referência. A parte de áudio da norma MPEG-2 é descrita a seguir.

MPEG áudio leva em conta o modelo psico-acústico humano para realizar uma compressão “perceptualmente sem perdas”. O modelo divide o domínio de freqüência audível (entre 20 Hz e 20 kHz) em 32 bandas, cada uma delas chamada de banda crítica. O sistema de audição tem uma resolução limitada e dependente da freqüência. A medida perceptualmente uniforme de freqüências pode ser expressa em termos das larguras das

bandas críticas.

O modelo leva em conta o mascaramento de frequências, característica do ouvido humano que quando submetido a um sinal de certa amplitude em uma dada frequência, mascara as outras frequências ao redor, que possuam uma amplitude abaixo de certo limite, como mostrado na Figura 2.11 - Mascaramento de frequências.

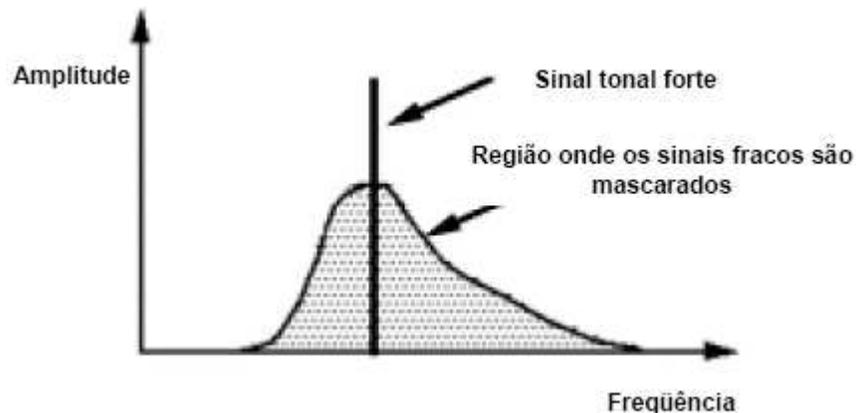


Figura 2.11 - Mascaramento de frequências.

MPEG áudio transforma o sinal para o domínio da frequência e aplica o mascaramento de frequências, codificando apenas aquelas componentes de frequência que não são mascaradas, como mostrado na Figura 2.12 - Mascaramento de frequências nas bandas críticas MPEG.

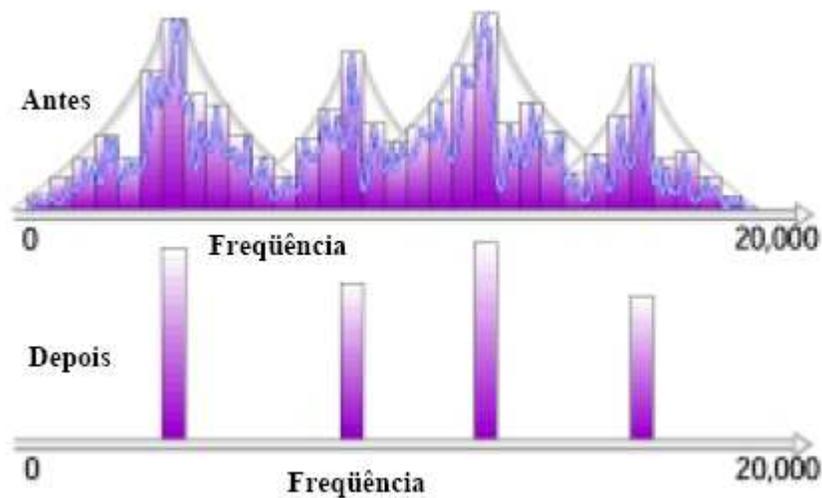


Figura 2.12 - Mascaramento de frequências nas bandas críticas MPEG.

Na codificação usando-se o padrão MPEG, para cada intervalo de tempo de áudio codificado (isto é, para cada conjunto de amostras) existe um número fixo de bits total para todas as 32 sub-bandas. Escolhe-se o número de bits de uma banda de forma a minimizar a percepção auditiva do ruído de quantização, levando-se em conta, como já mencionado, o mascaramento de frequências.

A compressão de áudio MPEG-2 descreve três níveis de compressão denominados camadas 1, 2 e 3 (*Layers* 1, 2 e 3). O nível de compressão, a demanda por mais processamento e a qualidade do som aumentam proporcionalmente com o número da camada. Ao contrário, a largura de banda requerida, diminui com o aumento do número da camada.

A camada 1 tem a menor taxa de compressão, cerca de quatro vezes, demanda a menor quantidade de processamento e tem o mais baixo atraso (*delay*), que é menor que 50 ms. *Layer* 1 exige também mais largura de banda para transmissão e tem uma taxa de transmissão de saída variando de 32 kbps (mono) a 448 kbps (Estéreo). Além disso, a qualidade de som de um sinal *Layer* 1 é inferior ao obtido pelas camadas 2 e 3.

Por outro lado, a camada 3 (MPEG-1 *Layer*-3, ou mais conhecido como simplesmente MP3) apresenta a melhor qualidade de som das três camadas, enquanto alcança um fator de compressão de 1:10. Por outro lado, o tempo de processamento é três vezes maior. As características mais importantes dessas camadas são comparadas na Tabela 2.2 - Características centrais das três camadas de compressão de áudio MPEG.

Camada	Taxa de compressão aproximada	Largura de banda desejada	Faixa de largura de banda permitida	Atraso	Atraso mínimo teórico
1	1:4	192 kbps	32-448 kbps	< 50 ms	19 ms
2	1:6	128 kbps	32-384 kbps	100 ms	35 ms
3	1:10	64 kbps	32-320 kbps	150 ms	58 ms

Tabela 2.2 - Características centrais das três camadas de compressão de áudio MPEG.

As três camadas são compatíveis no sentido de que a camada N pode decodificar a camada N e as camadas inferiores. Por exemplo, um decodificador de camada 3 pode decodificar as camadas 1, 2 e 3. Porém um decodificador de camada 2 só pode decodificar as camadas 1 e 2.

A razão de se ter mais de uma camada para endereçar cada uma das necessidades baseia-se parcialmente nas aplicações e na história. Se, por exemplo, uma reprodução de alta qualidade com uma menor largura de banda é o requisito chave, e o processamento e custo é o segundo requisito, então a camada 3 seria a opção lógica. A camada 2 têm uma

complexidade intermediária, largura de banda de 128 kbps por canal e é mais usada para transmissão de áudio digital.

O desenvolvimento das camadas foi gradual. As especificações da camada 3 foram finalizadas depois das camadas 1 e 2.

As três camadas têm técnicas de compressão e codificação em comum. A seguir pode-se ver uma introdução para a estrutura e função desse processo, conforme mostra a Figura 2.13 - Bloco de compressão de áudio genérico para as três camadas.

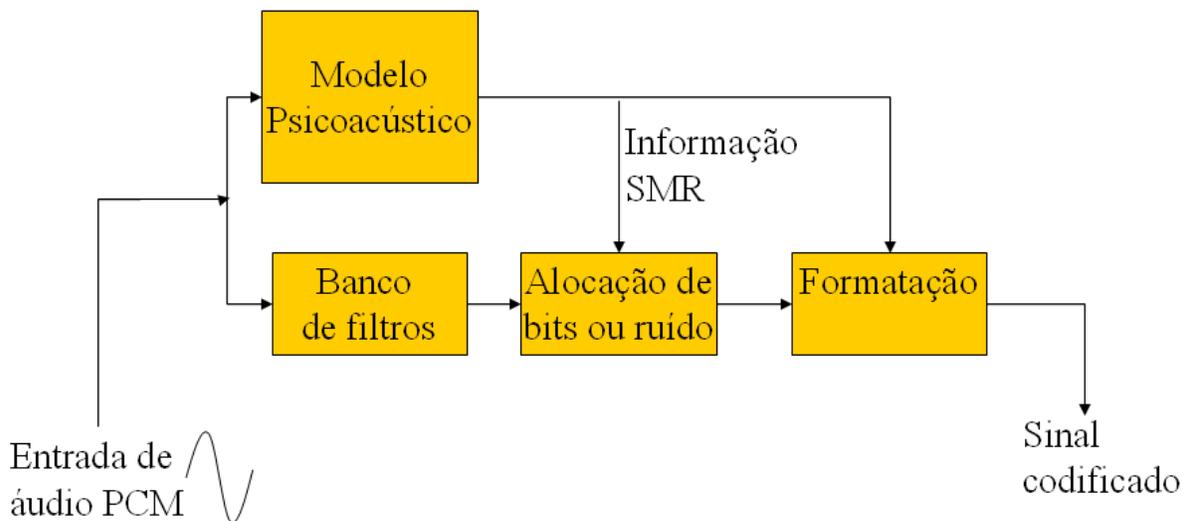


Figura 2.13 - Bloco de compressão de áudio genérico para as três camadas.

A entrada do processo é um sinal de áudio digital (PCM) que entra em um bloco chamado banco de filtros polifásico. No caso de um sinal com mais de um canal, como um som estéreo, cada canal é tratado separadamente. O banco de filtros usado na codificação de áudio pode ser de dois tipos, polifásico ou um híbrido polifásico e MDCT (*Modified Discrete Cosine Transform*). Independentes do tipo, as amostragens do domínio do tempo são convertidas para o domínio da frequência. A saída do banco de filtros é um número de sub-bandas de frequência de igual largura de banda. A codificação de camada 1 agrupa 12 amostras para cada uma das 32 sub-bandas. Já a de camada 2 agrupa 36 amostras para cada grupo de 32 sub-bandas. Cada grupo de 12 recebe então os bits para codificação e, se o número de bits não é zero, um fator de escala. Na camada 3, o número de sub-bandas pode ser de 192 ou 576. Os filtros são relativamente simples e provêm uma boa resolução no tempo com razoável resolução de frequência.

Em paralelo com o banco de filtros, o modelo psico-acústico calcula a razão sinal para máscara (SMR ou *Signal to Mask Ration*) para cada sub-banda. A principal razão do modelo psico-acústico é calcular uma nova alocação de bits para as amostragens de

freqüência nas sub-bandas. O objetivo desse bloco é alocar de forma eficiente a quantidade de bits para cada sub-banda. Portanto, se não existe potência em uma dada sub-banda, nenhum bit é alocado. O princípio central aplicado é o fato que freqüências de alta potência faz com que freqüência de baixa potência próxima às altas seja inaudível, ou seja, um som pode fazer outro inaudível. A alocação de bits é calculada separadamente para cada sub-banda. Em outras palavras, o princípio é a relação entre o que chega ao ouvido humano e o que realmente se escuta. Dessa forma, não se necessita transmitir aquilo que o ouvido não escuta.

Para calcular a SMR, é necessário primeiro fazer a conversão do tempo para o domínio da freqüência do sinal de áudio original. Essa conversão é feita pela transformada rápida de Fourier. Com base no domínio da freqüência, a máxima potência é calculada em cada sub-banda e, dessa forma se pode calcular a parte tonal e a parte não tonal. Com isso, se pode calcular a razão SMR.

O processo de alocação de bits ou ruído usa as amostragens do banco de filtros e a informação SMR do modelo psico-acústico para determinar a quantidade de ruído quantizado que é tolerado para cada sub-banda. Quanto maior a quantidade de ruído quantizado permitido, menor o número de bits necessário para representar cada amostra. Ou seja, o objetivo desse bloco é remover as componentes inaudíveis e quantizar as componentes restantes usando-se um menor número de bits (*Huffman*).

Camadas 2 e 3 ainda levam em conta uma outra característica psico-acústica, ou seja, o mascaramento temporal. Quando é emitido um tom em uma dada freqüência e com certa amplitude, esse tom mascara os tons, na mesma freqüência, abaixo de certa amplitude, que varia no tempo.

O último estágio consiste em se formatar os bits. A saída do banco de filtros quantizada e a alocação de ruído juntamente com outras informações requeridas são coletadas, codificadas e formatadas.

Diferentes modelos de codificação podem ser usados na compressão de áudio MPEG, conforme mostrados na Tabela 2.3 - Características centrais das três camadas de compressão de áudio MPEG.

O sinal *Joint Estéreo* também leva em consideração resultados psico-acústicos que mostram que, para freqüências maiores do que 2 kHz, o ouvido percebe a imagem estereofônica baseado mais na envoltória temporal do áudio do que em sua estrutura mais refinada. Assim, no modo *intensity stereo*, o codificador soma as freqüências mais altas do sinal estereofônico em um único sinal. Os canais de esquerda e direita são reconstruídos com a mesma forma, mas com magnitudes diferentes, baseadas em fatores de escala.

Canais	Descrição
<i>Mono (Single)</i>	Um sinal monofônico é transmitido.
<i>Mono (Dual)</i>	Dois sinais monofônicos independentes são transmitidos.
<i>Estéreo</i>	Um sinal estereofônico é transmitido (Canais de áudio esquerdo/direito são transmitidos separadamente).
<i>Joint Estéreo (Combinado)</i>	Um sinal estereofônico é transmitido, mas a informação do canal esquerdo e direito acima de certa frequência (sub-banda) é combinada de maneira a comprimir o sinal.

Tabela 2.3 - Características centrais das três camadas de compressão de áudio MPEG.

Dolby Digital (AC-3)

O AC-3 é o antigo nome do *Dolby Digital*. Trata-se de um sistema de compressão de áudio mais avançado do que o MPEG desenvolvido em 1992. Essa codificação é proprietária da empresa *Dolby*, mas vem sendo utilizada em DVDs e em vários sistemas de transmissão digital. Elas utilizam seis canais discretos de áudio, sendo cinco com qualidade de CD (20 Hz a 20 kHz) e um apenas para as baixas frequências (20 a 120 Hz). A taxa dessa codificação é de cerca de 384 Kbps. A Figura 2.14 - Áudio multicanal apresenta a distribuição sugerida de alto-falantes para os seis canais de áudio.

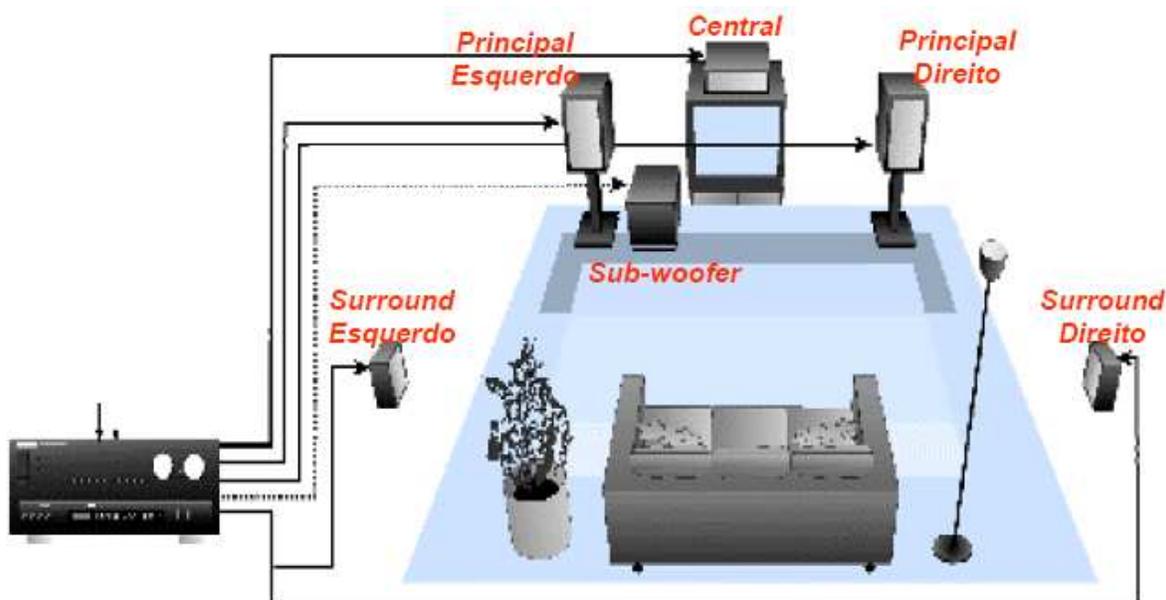


Figura 2.14 - Áudio multicanal.

Dolby Digital divide o espectro de áudio em bandas de frequências estreitas usando modelos matemáticos derivados das características do ouvido humano e analisa cada banda separadamente. Para maximizar a eficiência de dados, os maiores números de bits representam os sinais mais audíveis e menos bits representam os sinais menos audíveis. Para determinar a audibilidade dos sinais, o sistema realiza o mascaramento, fenômeno já explicado na sessão anterior.

MPEG-2 AAC

A codificação AAC é uma novidade do áudio MPEG 2, conhecida como NBC (*Non Backward Compatible*), de não compatível com MPEG 1. Essa codificação é mais eficiente do que o MPEG 1 e permite até 48 canais de áudio e 15 canais de frequências baixas, com taxas de amostragem de 8 a 96 KHz. A AAC necessita de menos processamento e, conseqüentemente, tem retardo menor na (de) codificação.

Tecnologia de transporte de conteúdo

MPEG-2 Transport Stream

Fluxo de Transporte MPEG-2 (MPEG-2 TS ou MPEG-2 *Transport Stream*) é um protocolo de multiplexação e transmissão de áudio e vídeo especificado pela norma ISO/IEC 13818-1, MPEG-2, Parte 1, Sistemas [7, 32, 33, 34]. O MPEG-2 TS oferece características para a correção de erro e é usado em aplicações da transmissão tais como DVB, ATSC e *streaming*.

A tecnologia de *streaming* de vídeo, diferente do *download*, é uma seqüência de imagens, que de forma comprimida e contínua são transmitidas em tempo real. Uma vez que o vídeo é enviado em um *stream* contínuo, ele é reproduzido à medida que chega ao destino. À medida que o vídeo é recebido vai sendo criado um pequeno *buffer*. Quando há informação suficiente no *buffer*, o decodificador começa a reproduzir o vídeo.

O MPEG-2 TS tem como objetivo multiplexar o fluxo de vídeo, de áudio(s), encapsulá-los num formato típico e garantir a sincronização entre áudio e vídeo. Ele tem ainda a função de sincronizar a saída.

Como foi visto nas sessões anteriores, os codificadores de áudio e vídeo são responsáveis pela compressão e codificação dos sinais de áudio e vídeo. Esse fluxo codificado recebe o nome de fluxo elementar (*Elementary Stream* ou ES). Esses são multiplexados a fim de se produzir um programa.

Um programa consiste em um ou mais fluxos elementares, mas tipicamente para IPTV,

ele constituirá de um vídeo, um ou mais ES de áudio, um ES de dados, e zero ou mais ESs de legendas como a seguir:

- *Stream* de dados: informações específicas do programa (*Program Specific Information* ou Tabelas PSI);
- *Stream* de vídeo: MPEG-2, H.264, ou VC-1;
- *Stream* de áudio: MPEG *Layer II*, *Dolby AC-3*, ou AAC (1 ou mais);
- *Stream* de legenda: *DVB bitmap* (0 ou mais).

O MPEG-2 TS pode transmitir muitos programas ou serviços com áudio, vídeo e dados entrelaçados juntos. O decodificador deve ser capaz de classificar o fluxo de transporte, localizar e organizar o vídeo, áudio e dados por programa ou serviço e apresentar o programa para o requisitante. O MPEG-2 TS, então, especifica a estrutura do fluxo de transporte e o mecanismo para transmissão dos dados comprimidos. Dentre outras coisas, essa estrutura provê mecanismos para uma rápida sincronização e correção de erro no decodificador. O MPEG-2 TS também define as tabelas PSI (*Program Specific Information* ou informação específica do programa), que atuam como um índice e permite que o decodificador classifique os dados de forma rápida e acesse a informação desejada no fluxo de transporte.

A seguir é mostrado como se cria um fluxo de MPEG-2 TS. O multiplexador MPEG-2 é alimentado com fluxos elementares de áudio e vídeo gerados pelos codificadores, dados privados, informação de serviço, controle de acesso, legendas e informação de sincronização, como pode ser visto pela Figura 2.15 - Diagrama de blocos do MPEG-2 TS.



Figura 2.15 - Diagrama de blocos do MPEG-2 TS.

Dentro do multiplexador MPEG-2, é feito um agrupamento inicial dos pacotes, onde os fluxos elementares de áudio e vídeo produzem pacotes denominados PES (*Packetized*

Elementary Streams). Esses pacotes, feitos de forma similar ao transporte IP, irão conter dados e cabeçalho, como mostrado na Figura 2.17 - Cabeçalho do PES. Em seguida, são gerados pacotes de transporte a partir dos pacotes PES formando assim um fluxo de transporte que é no fundo a multiplexagem dos pacotes elementares com dados adicionais, como dados de sincronização ou controle de acesso.

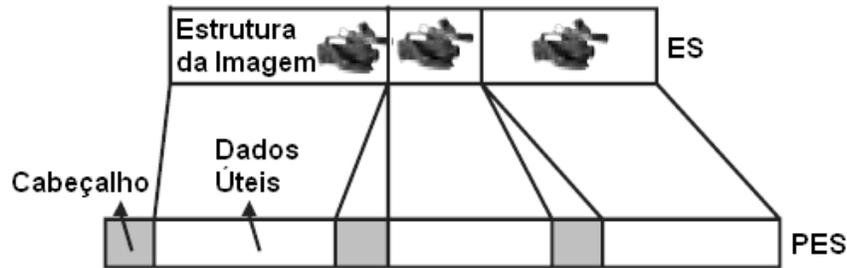


Figura 2.16 - Criação do PES.

Os PESs são divididos em tamanhos fixos de 188 bytes cada, sendo 184 bytes de informação útil e 4 bytes de cabeçalho, conforme mostra a Figura 2.17 - Cabeçalho do PES.

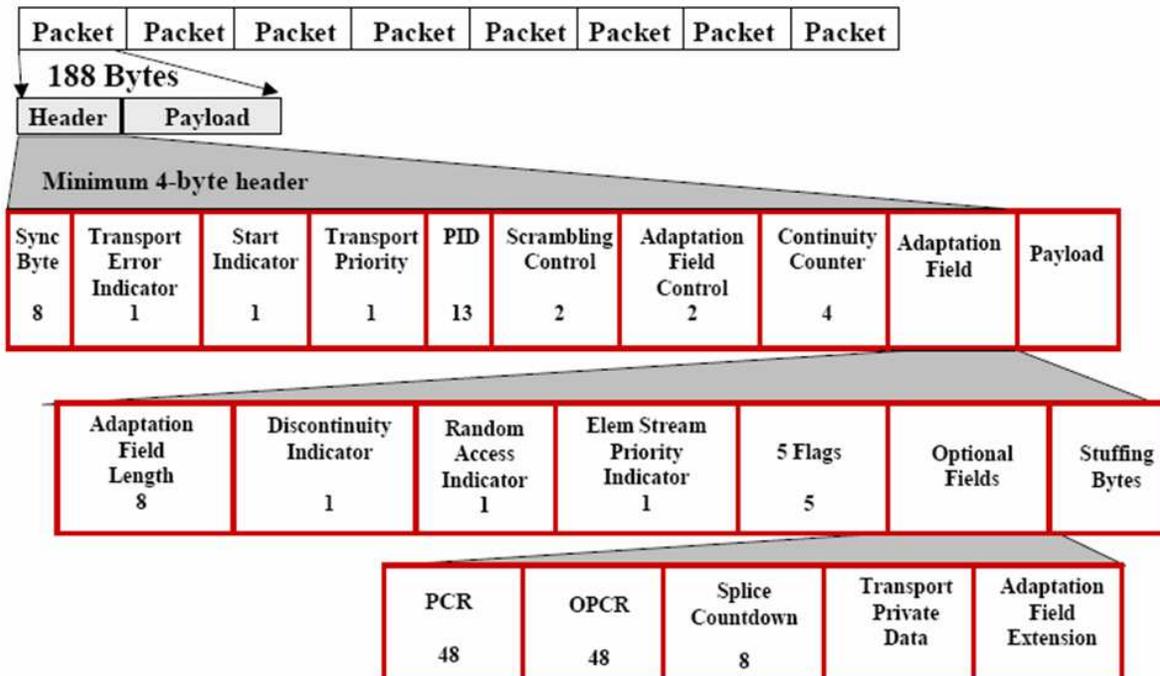


Figura 2.17 - Cabeçalho do PES.

A Tabela 2.4 - Cabeçalho do PES mostra a descrição desses campos.

Nome	N ^o de Bits	Descrição
Byte de Sincronismo	8	Valor 0x47 em hexadecimal
Indicador de erro de transporte	1	Indica presença de um erro não corrigido
Início do <i>payload</i>	1	Indicação do início do <i>payload</i>
Prioridade	1	1 maior prioridade, 0 menor prioridade
PID	13	Identificação do programa
Controle de embaralhamento	2	'00' = não embaralhado
Controle do campo de adaptação	2	Se primeiro bit é setado, campo de adaptação existe. Se segundo bit é setado, existem dados.
Contador de continuidade	4	Campo incrementa para cada pacote enviado no mesmo PID
Campo de adaptação	0 ou mais	Depende do <i>flag</i> acima
Dados de <i>payload</i>	0 ou mais	Depende do <i>flag</i> acima

Tabela 2.4 - Cabeçalho do PES.

A Tabela 2.5 - Cabeçalho do PES - Campos Opcionais mostra a descrição dos campos opcionais.

Cada fluxo elementar é identificado pelo campo de identificação do programa (PID), que tem um papel fundamental na operação do MPEG-2 TS. Um de-multiplexador extrai os fluxos elementares do MPEG-2 TS procurando os pacotes identificados pelo mesmo PID. Na maioria das aplicações, a multiplexação por divisão de tempo será usada para quão freqüente um PID particular aparece no fluxo de transporte. Como mostrado nas Tabela 2.4 - Cabeçalho do PES e Tabela 2.5 - Cabeçalho do PES - Campos Opcionais, o cabeçalho do PES têm vários campos, entretanto como o PID é um valor que identifica o programa, em geral identifica-se o cabeçalho usando-se somente seu PID, como mostrado Figura 2.18 - PID no fluxo elementar.

Nome	Nº de Bits	Descrição
Tamanho do campo adaptação	8	Comprimento do campo de adaptação
Indicador de descontinuidade	1	1 se descontinuidade ocorreu
Indicador de acesso aleatório	1	1 se pacote PES inicia uma seqüência de vídeo/áudio
Indicador de prioridade do ES	1	1 = prioridade mais alta
<i>Flag</i> PCR	1	Se 1, existe campo PCR
<i>Flag</i> OPCR	1	Se 1, existe campo OPCR
<i>Flag</i> do ponto de junção	1	Se 1, existe campo ponto de junção
<i>Flag</i> de dados privados	1	Se 1, existe campo de dados privado
<i>Flag</i> extensão campo adaptação	1	Se 1, existe campo extensão
Campos abaixo são opcionais	Variável	Dependendo dos <i>flags</i>
PCR	48	Referência de relógio
OPCR	48	Referência de PCR original
Contagem regressiva do <i>splice</i> (<i>Splice Countdown</i>)	8	Quantos pacotes TS desde o ponto emendando (pode ser negativo)
Bytes de enchimento	Variável	

Tabela 2.5 - Cabeçalho do PES - Campos Opcionais.



Figura 2.18 - PID no fluxo elementar.

Um programa pode ter vários fluxos de áudio e vários fluxos de legenda, além do seu fluxo elementar de vídeo. Caso estejam disponíveis, esses vários fluxos de áudio e ou legendas (cada um com seu próprio PID), o decodificador poderá selecionar qual fluxo de áudio e qual fluxo de legenda deseja ouvir e apresentar. Entretanto, isso vai depender de algumas outras implementações no *middleware* que serão discutidas no próximo capítulo,

além disso, o decodificador tem que suportar os formatos de compressão de áudio e legenda utilizados na codificação.

Dessa forma, a informação de vídeo e áudio trafega dentro do fluxo de transporte. Para poder coletar os pacotes de um mesmo programa, o receptor tem que conhecer os PIDs correspondentes. Por isso, o MPEG-2 TS inclui a informação necessária para a localização dos distintos pacotes em algumas tabelas que também trafegam multiplexados no fluxo de transporte. Essas tabelas são chamadas de PSI (*Program Specific Information* ou informações específicas do programa) e usam valores reservados de PID.

As tabelas PSI atuam como um índice para o fluxo de transporte, fazendo com que o decodificador possa encontrar o dado que procura de forma rápida.

As tabelas mais usuais são descritas a seguir.

A tabela PAT (*Program Association Table* ou tabela de associação do programa) lista os PIDs de todas as tabelas PMTs e tem sempre o endereço de PID 0x0. Essa é a primeira tabela que o decodificador procura, e por isso todo fluxo de transporte tem sempre que contê-la.

A tabela PMT (*Program Map Table* ou tabela de mapeamento do programa) contém informação dos programas. Para cada programa existe uma tabela PMT correspondente e com um valor de PID único. Dentro de cada PMT vão existir os PIDs correspondentes ao seu programa. Por exemplo, se um programa contém um fluxo elementar de vídeo H.264, a tabela PMT vai descrever isso como um fluxo de vídeo e o tipo de vídeo que ela contém. A tabela PMT terá ainda descritores adicionais, incluindo sincronismo e PID(s) de legenda(s) se existente(s).

O sistema de transmissão por TV a cabo, onde trafegam todos os canais ao mesmo tempo, conterà uma única tabela PAT e várias tabelas PMTs para cada um de seus programas ou canais (MPTS - *Multiple Program Transport Stream*). No caso do IPTV, onde apenas o canal que está sendo transmitido chega à casa do usuário (SPTS - *Single Program Transport Stream*), o fluxo de transporte vai ter uma tabela PAT e uma única tabela PMT.

A tabela PCR (*Program Clock Reference*) define uma referência de relógio, que pode ser ajustada e usualmente é inserida no fluxo elementar de vídeo e é usada para sincronizar o decodificador com o codificador.

Existem outras tabelas opcionais e de menor importância que as tabelas acima que são a CAT (*Conditional Access Table*) e a NIT (*Network Information Table*).

A CAT contém informação a cerca da criptografia do acesso condicional e são usadas junto com as tabelas ECM (*Entitlement Control Message*) e EMM (*Entitlement Management Message*). Isso será discutido no próximo capítulo, onde será abordado o tema de acesso condicional.

A NIT, encontrada com PID igual a 0x10, contém detalhes da rede, incluindo frequência da portadora, é mais usada em sistemas de TV a cabo e não é usada no IPTV.

A Figura 2.19 - Diagrama de blocos do MPEG-2 TS mostra os passos na criação do fluxo de transporte MPEG-2.

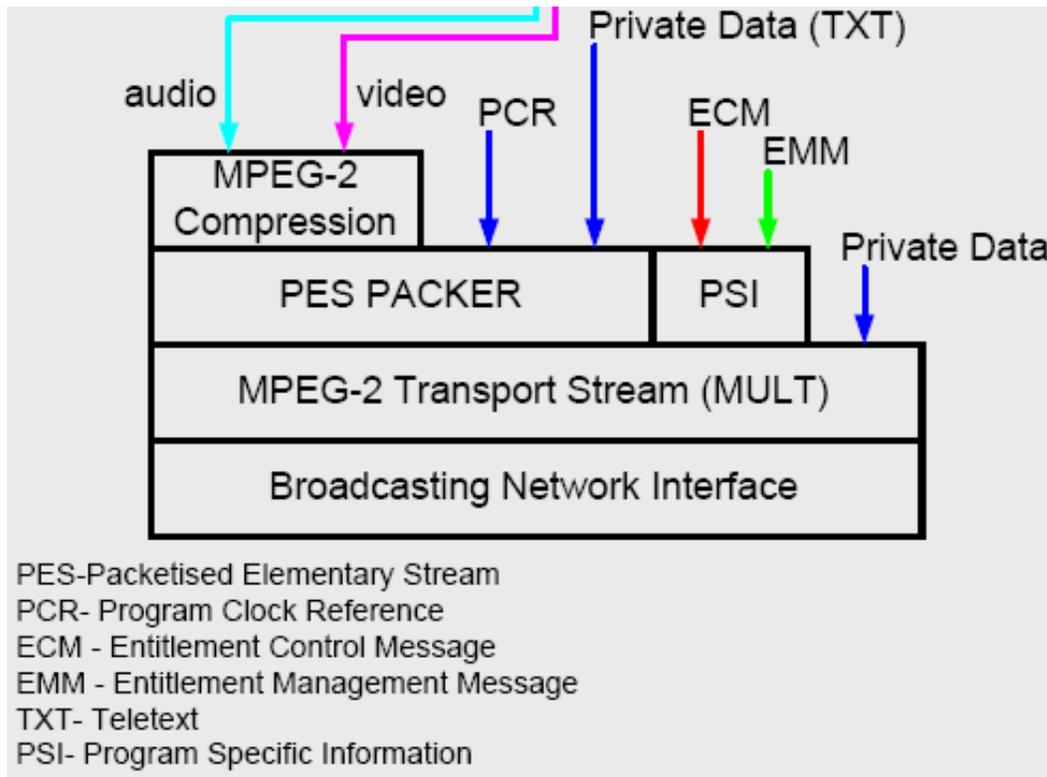


Figura 2.19 - Diagrama de blocos do MPEG-2 TS.

A Figura 2.20 - Tabelas PSI do MPEG-2 TS e PIDs mostra a representação das tabelas PSI inseridas no fluxo de transporte.

O sincronismo no fluxo de transporte é baseado no sinal STC (*System Time Clock* ou relógio de tempo do sistema) de 27 MHz do codificador. Para assegurar a sincronização durante o processo de decodificação o relógio do decodificador precisa ser travado ao STC do codificador. Para alcançar esse travamento, o codificador precisa inserir no fluxo de transporte um selo de tempo (*time stamp*) de 27 MHz para cada programa.

O selo de tempo é conhecido como PCR. Usando o PCR, o decodificador gera um relógio local de 27 MHz que é travado ao STC do codificador. A Figura 2.21 - Sinal PCR mostra como o decodificador trava o relógio como codificador.

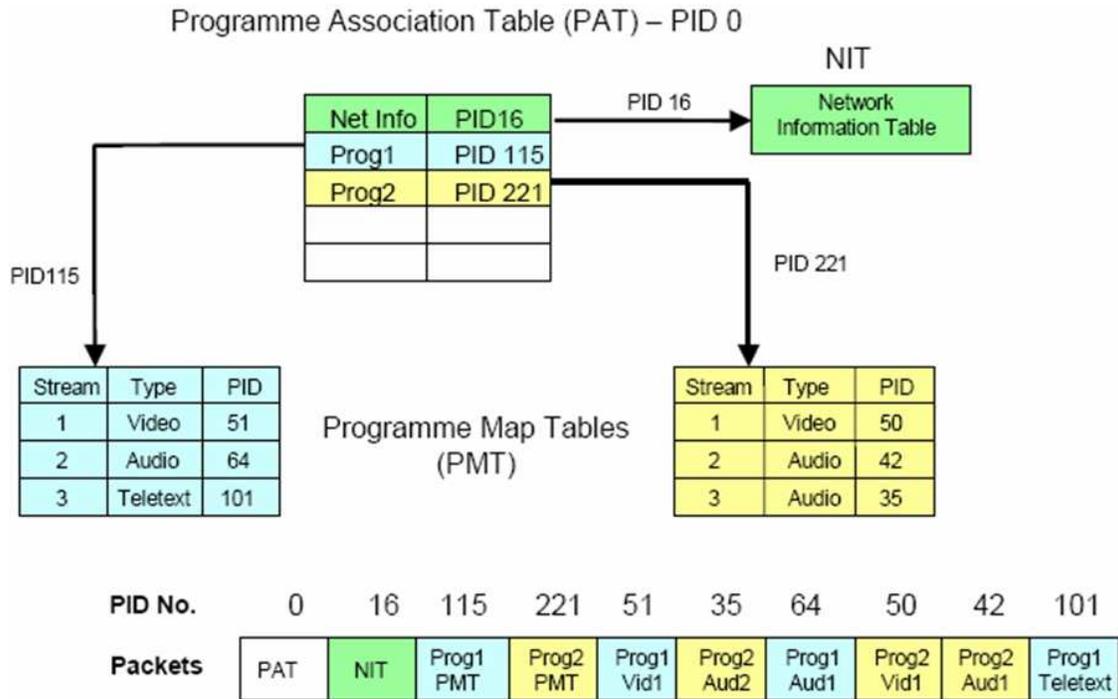


Figura 2.20 - Tabelas PSI do MPEG-2 TS e PIDs.

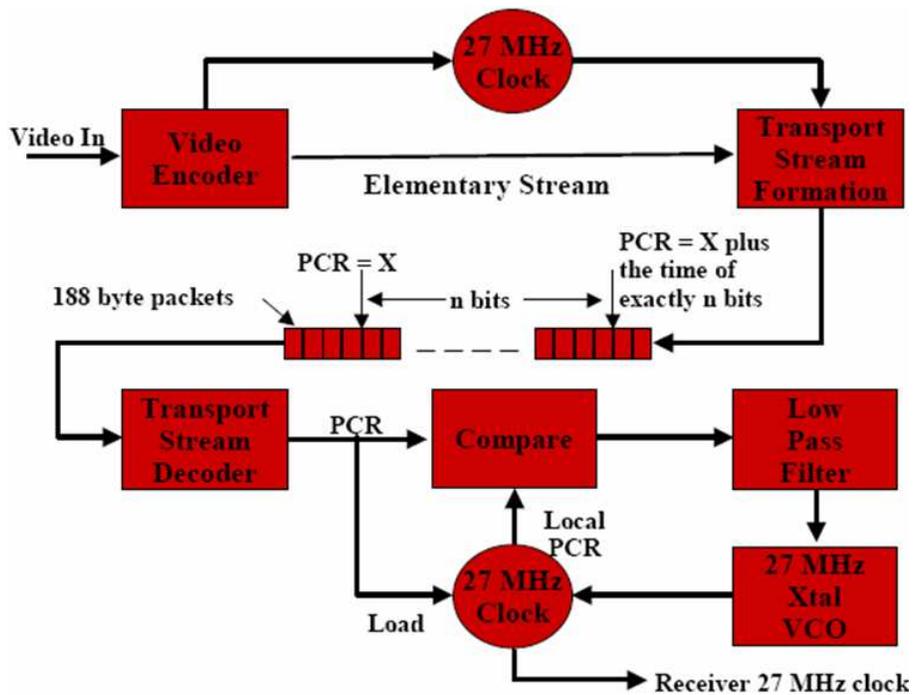


Figura 2.21 - Sinal PCR.

Os quadros de vídeo comprimidos são frequentemente transmitidos fora de ordem. Isso significa que um quadro tipo I usado para regenerar um quadro B precedente deve estar disponível no decodificador bem antes do tempo de apresentação. Para gerenciar esse processo crítico de tempo, existem dois selos de tempo no cabeçalho do pacote PES, o DTS (*Decoding Time Stamp*) e o PTS (*Presentation Time Stamp*). Esses selos informam ao decodificador quando um frame deve ser decodificado (DTS) e mostrado (PTS). Se o DTS para um quadro precede seu PTS consideravelmente, o quadro é decodificado e guardado no *buffer* até que seu selo de apresentação chegue.

A Figura 2.22 - Sequência de tempo no fluxo de transporte mostra a inserção dos tempos no MPEG-2 TS e a decodificação. Antes do fluxo de transporte ser criado, o codificador acrescenta os PTSs e os DTSs para cada quadro no PES. Ele também coloca o PCR para cada programa dentro do MPEG-2 TS. Dentro do decodificador, o PCR passa por um algoritmo PLL (*Phase Lock Loop*), que trava (*lock*) o relógio do decodificador ao STC do codificador. Isso sincroniza o decodificador com o codificador para evitar que o *buffer* do decodificador esvazie (*underflow*) ou encha (*overflow*).

Uma vez que o relógio do decodificador esteja sincronizado, ele começa a decodificar e apresentar os programas como especificados pelo PTS e DTS para cada quadro de vídeo e áudio.

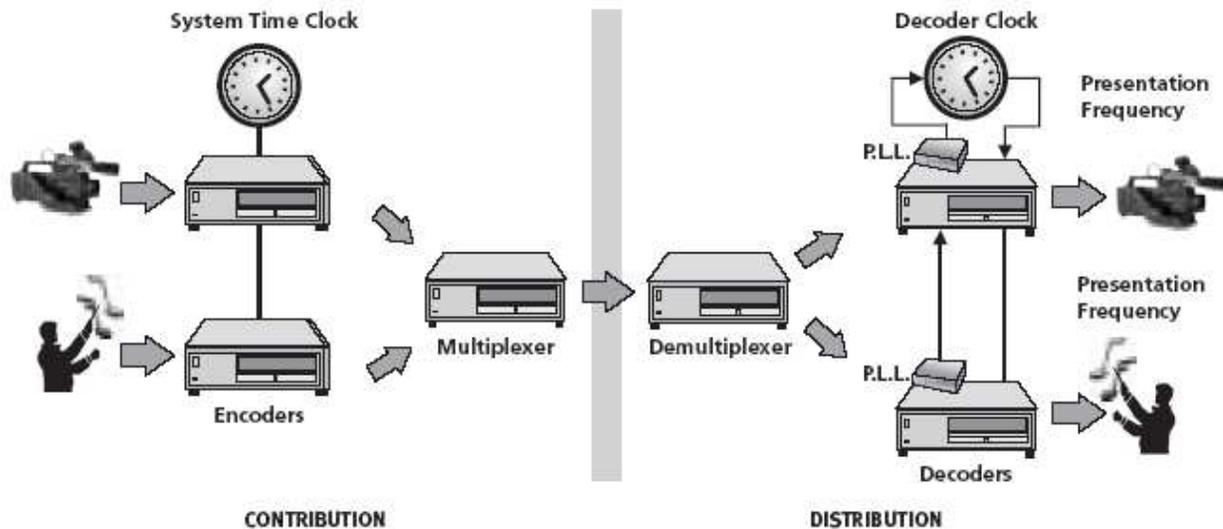


Figura 2.22 - Sequência de tempo no fluxo de transporte.

Um decodificador deve classificar através do MPEG-2 TS, organizando os fluxos do vídeo, do áudio e de dados pelo programa ou pelo serviço. Deve também saber quando apresentar cada parte do programa ou do serviço ao usuário. Esse é o lugar aonde MPEG-2

TS entra. A camada do sistema especifica a estrutura do MPEG-2 TS, ou seja, o mecanismo da transmissão para dados comprimidos. Entre outras coisas, essa estrutura providencia uma rápida sincronização e correção de erro no decodificador. Como explicado anteriormente, as tabelas PSI agem como um índice, permitindo que o decodificador rapidamente classifique e acesse a informação no MPEG-2 TS.

As tabelas PSI ajudam o decodificador a localizar o áudio e o vídeo para cada programa no MPEG-2 TS e verificam permissões ao acesso condicional (CA). As tabelas são repetidas freqüentemente no MPEG-2 TS para suportar o acesso aleatório requerido por um decodificador ao mudar os canais, como mostrado na Figura 2.23 - Fluxo de Transporte MPEG-2 TS com tabelas PSI.

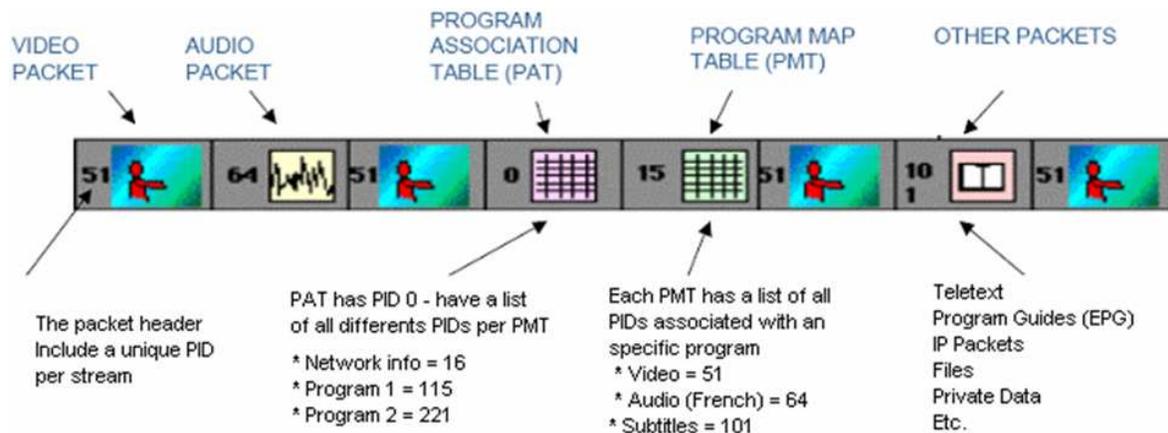


Figura 2.23 - Fluxo de Transporte MPEG-2 TS com tabelas PSI.

Como pode ser visto na Figura 2.24 - Processo de decodificação, o decodificador começa a ler o fluxo de transporte e procura pelo PID 0x0 que corresponde à tabela PAT. A tabela PAT vai conter o endereço PID da tabela PMT.

A seguir o decodificador vai de-multiplexar os pacotes que transportam a PMT desejada e lê-la. A tabela PMT (em IPTV só uma PMT por vez) vai conter os PIDs dos fluxos elementares (vídeo, áudios, legendas), bem como o formato usado para codificá-los. Baseado nos valores dos PIDs definidos na PMT, o decodificador começa a de-multiplexar os fluxos elementares correspondentes.

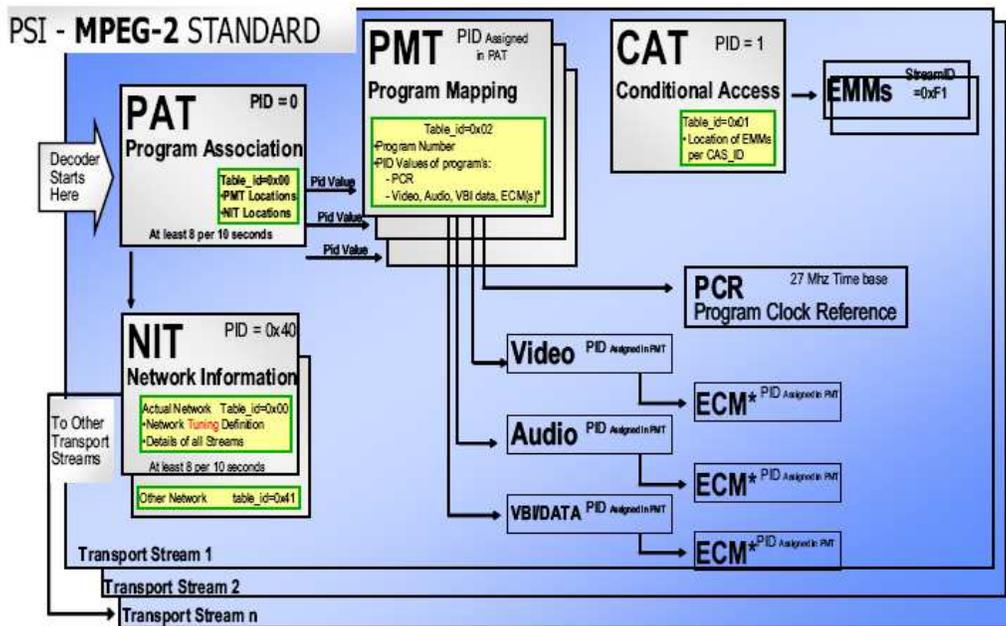


Figura 2.24 - Processo de decodificação.

Legenda

O grupo DVB especificou um meio pelo qual um ou mais fluxos de legenda podem ser parte do fluxo de transporte. A especificação é conhecida como *DVD subtitling* padronizada pela norma ETSI 300 743 [37]. Ambas as extremidades (codificador e decodificador) têm que suportar o padrão *DVD subtitling*.

O método de transmissão de legenda mais usado é o *DVB bitmap*, que opera convertendo cada linha da legenda em uma imagem gráfica e transmitindo isso como um objeto *bitmap*.

Cada caracter em uma legenda é renderizado em imagens *bitmap* seguindo regras de codificação definidas localmente pelo processador de legendas DVB. Isso estabelece a altura e a largura de cada caracter, bem como qualquer outra formatação como *itálico*, *negrito* ou *sublinhado*. O contorno e o foco são também tratados pelo processo de renderização.

A cor de cada *pixel* na imagem é determinada por uma tabela de cores que contém 256 definições de cores.

Uma região tem que ser criada na TV para cada linha da legenda que pode ter uma cor de fundo ou um nível transparente associado a ela.

Como foi visto na sessão anterior, a legenda é um fluxo elementar e referenciado por seu próprio PID. Dessa forma, o fluxo de transporte pode ou não conter um ou mais fluxos

elementares de legenda. Uma vez que a legenda é um ES presente no MPEG-2 TS, ela pode ser mostrada, escondida ou selecionada uma outra (caso tenha mais de uma no fluxo de transporte), tanto para conteúdos ao vivo, quanto para conteúdos gravados como vídeo sob demanda.

O PID do fluxo de legenda deverá estar indicado na tabela PMT e permitirá que o decodificador a selecione.

As legendas são transportadas como objetos com sua própria identificação e, cada objeto inclui seu PTS associado, que indicará ao decodificador em que instante relativo ao PCR o objeto deve aparecer na tela. Esse mecanismo otimiza a largura de banda e simplifica a entrega do serviço.

Um único fluxo de legenda *bitmap* requer tipicamente uma largura de banda de 50 a 100 kbps dependendo do seu tamanho e da densidade da sua cor.

As legendas são codificadas no formato DVB, usando um processador de legenda e, então, o fluxo de legenda é multiplexado com o fluxo de vídeo, áudio e dados PSI, formando um sinal completo para transmissão *broadcast*.

A Figura 2.25 - Diagrama simplificado de codificação e inserção de legenda mostra o processador de legenda, contribuindo para criar o fluxo de legenda, da mesma forma que o codificador MPEG cria os fluxos de vídeo e áudio.

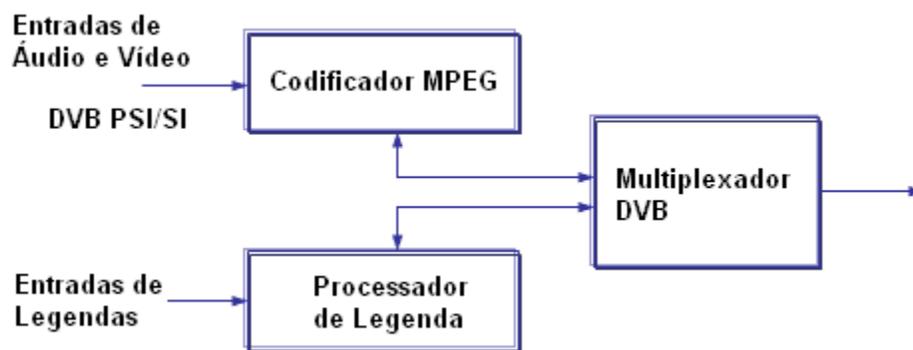


Figura 2.25 - Diagrama simplificado de codificação e inserção de legenda.

Algumas vezes o provedor de conteúdo transmite a legenda em um formato proprietário, que o decodificador não entende. Nesse caso, uma transcodificação do formato da legenda será necessária.

IP Multicast

O protocolo IP usa dois tipos de protocolo para transmissão em camada de rede: o TCP

(*Transmission Control Protocol*) e o UDP (*User Datagram Protocol*). O TCP provê confiabilidade em um canal de conexão baseado numa conexão. Ele quebra grandes cadeias de dados em pacotes menores e assegura que os dados que foram enviados serão recebidos. O UDP, por outro lado, é um protocolo não orientado a conexão e não garante a entrega dos dados enviados, sem que haja um mecanismo próprio de controle de erro.

Em geral o TCP é usado quando uma comunicação confiável é preferida sobre uma de menor latência. A confiabilidade do TCP através de retransmissões pode introduzir um atraso significativo. O UDP, por outro lado, não prove mecanismos de retransmissão de dados perdidos e, por isso, não introduz atrasos. Devido a isso, o UDP é mais conveniente para aplicações em tempo real como um vídeo. O dimensionamento da rede deve ser muito bem feito a fim de se evitar atraso, *jitter* (variação no atraso) e perda de pacotes na rede. As considerações de rede serão discutidas no capítulo 3.

Existem três métodos para transmissão de dados (vídeo) em uma rede IP. O *broadcast*, o *multicast* e o *unicast*.

O *broadcast* é um método de transmissão de um para todos, mesmo se o usuário não quer receber o dado, ainda assim ele irá recebê-lo. Geralmente dados *Broadcast* são restritos a um segmento de rede de comunicação local e não tem uso prático para transmissões de vídeo no IPTV.

O *unicast* é um método de transmissão baseado em uma comunicação ponto a ponto. Nesse caso, os pacotes de dados são enviados somente para um destinatário e nenhum outro computador da rede terá que processar essa informação. A transmissão *unicast* é tipicamente utilizada para enviar um *streaming* de vídeo do servidor para um único cliente. *Unicast* é relativamente simples de se implementar e cada usuário recebe o mesmo endereço para conectar e solicitar o streaming quando ele quiser. O uso da transmissão *unicast* não é eficiente quando muitos usuários estão recebendo a mesma informação ao mesmo tempo porque uma conexão separada precisa ser mantida para cada usuário. Se o conteúdo é acessado por centenas ou milhares de usuários, a largura de banda para tal servidor deve ser centena ou milhares de vezes maiores que a requerida para um único usuário. A Figura 2.26 - Exemplo de uma transmissão *unicast* mostra um caso em que três usuários estão acessando o mesmo conteúdo ao mesmo tempo. Nesse caso, três conexões ponto a ponto são requeridas.

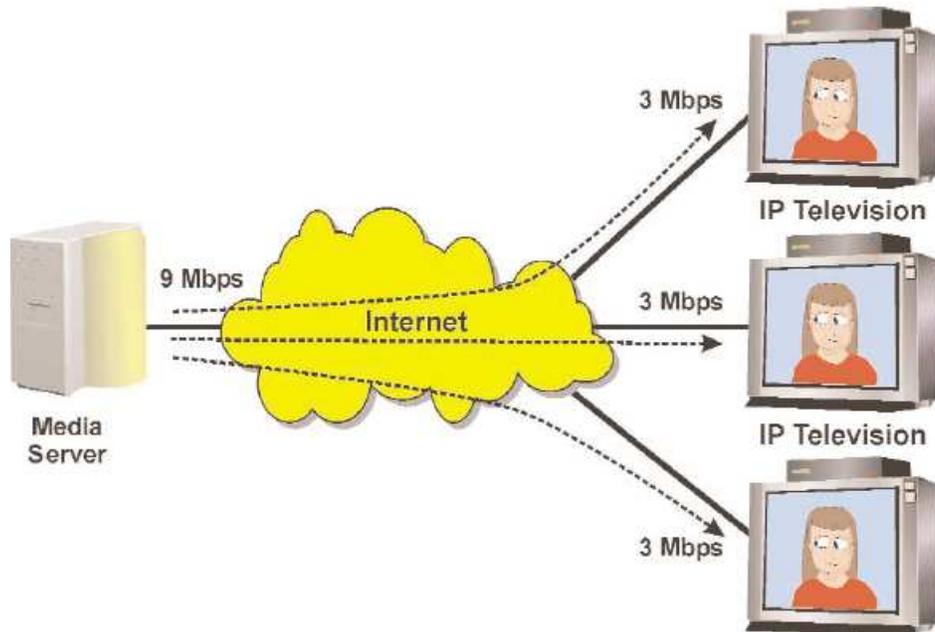


Figura 2.26 - Exemplo de uma transmissão *unicast*.

O *multicast* é a comunicação entre um único emissor e tendo-se múltiplos receptores na rede. A tecnologia *multicast* é utilizada para reduzir tráfego na rede quando muitos receptores querem ver o mesmo conteúdo fonte simultaneamente. Dessa forma, a fonte transmite uma única cópia do fluxo para centenas de recipientes. A maior diferença comparada com o *unicast* é que o *stream* de vídeo é enviado somente uma vez. Já a diferença do *multicast* para o *broadcast* é que no primeiro caso, somente os usuários que desejarem é que vão receber o *streaming*. *Multicasting* (também conhecido como *IP-Multicasting*) é tipicamente usado em conjunto com transmissões em tempo real [6, 26, 43].

A Figura 2.27 - Exemplo de uma transmissão sem e com *multicast* mostra a comparação de uma situação sem e com *multicast*.

Um serviço de TV *broadcast* entrega um único *stream* de vídeo para muitos usuários. Por exemplo, se 1000 usuários estão assistindo um canal de notícias da CNN, todos eles receberão a cópia do mesmo *stream*. No caso de um serviço *unicast*, onde cada usuário recebe seu próprio conteúdo fim-a-fim, isso requer uma enorme quantidade de largura de banda. Nesse caso, a largura de banda não é usada de forma eficiente.

Um uso mais eficiente seria a replicação do conteúdo na rede. Do satélite somente um *stream* será enviado para a rede, e o canal de vídeo é replicado somente no nó de acesso. Se não existe nenhum usuário assistindo esse canal em um determinado nó de acesso, então esse não recebe uma cópia. Dessa forma, pode-se economizar largura de banda, entregando um único canal de informação para milhares de assinantes.

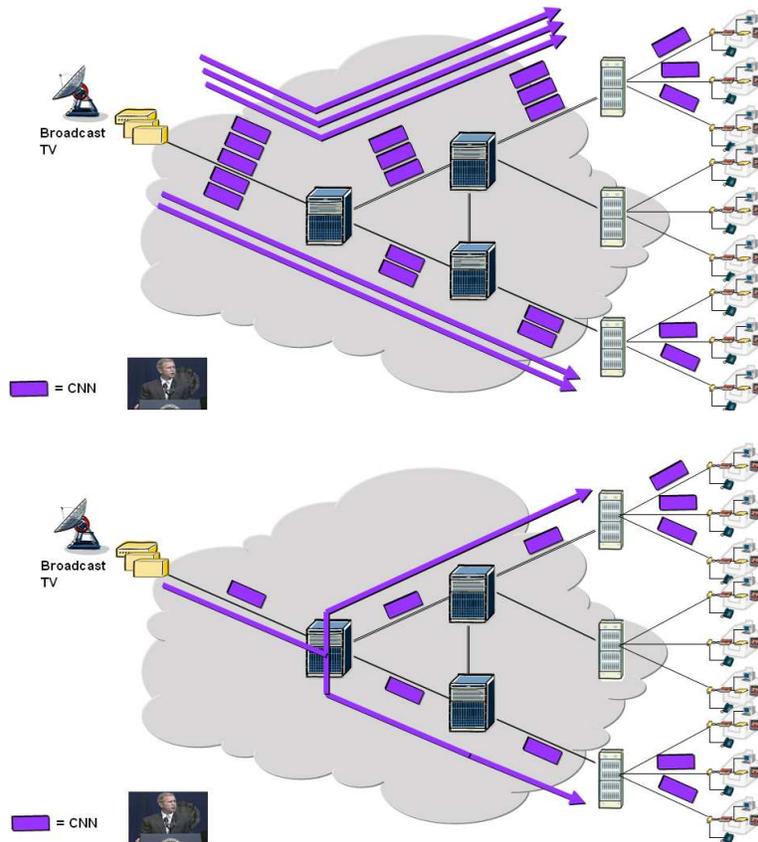


Figura 2.27 - Exemplo de uma transmissão sem e com *multicast*.

Quando os dados são transmitidos por *broadcast*, uma única cópia dos dados é emitida a todos os clientes na rede. *Unicast* desperdiça larguras de banda emitindo múltiplas cópias dos dados a vários usuários. A transmissão *broadcast* pode também afetar o desempenho dos equipamentos dos clientes, uma vez que cada um tem que processar o *broadcast*, independente do interesse ou não do usuário.

O *multicast* utiliza o que há de melhor nas duas situações acima, transmitindo uma única cópia dos dados àqueles clientes que os solicitam. Dessa forma, múltiplas cópias dos dados não são emitidas através da rede, nem são enviadas para clientes que não as querem. Com isso, a transmissão em *multicast* pode reduzir drasticamente a largura de banda que as aplicações multimídia necessitam na rede.

O usuário precisa indicar se ele quer receber uma cópia de um determinado canal. Isso é feito usando o protocolo chamado IGMP (*Internet Group Multicast Protocol*). Ou seja, o IGMP especifica comunicação somente entre o usuário e o primeiro roteador da rede. Na rede IP, os roteadores também têm de indicar se querem ou não receber o *multicast* para outros roteadores. Assim, nessa parte da rede, o protocolo mais usado é o PIM (*Protocol Independent Multicast*). Isso é mostrado pela Figura 2.28 - Protocolos de *multicast* IGMP x PIM.

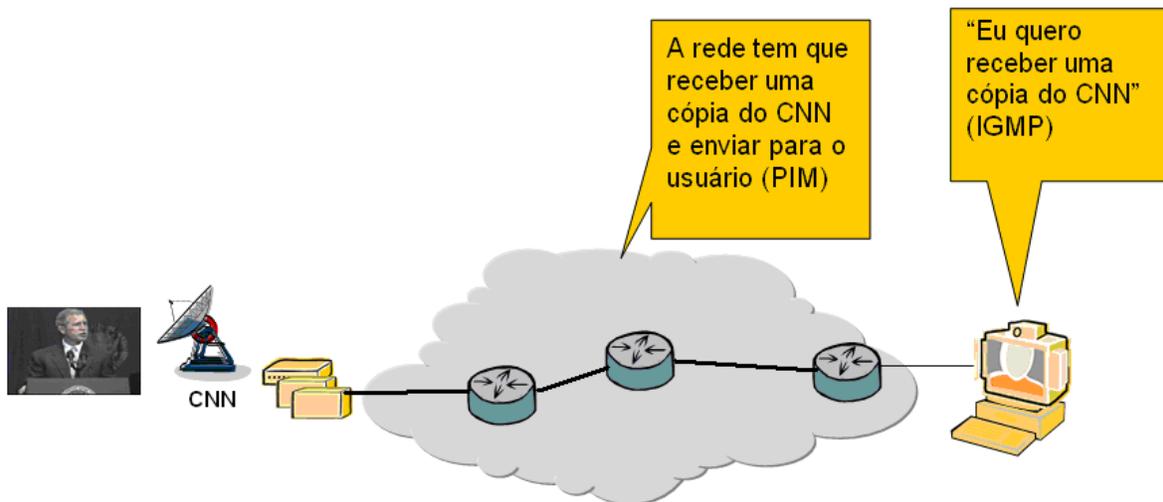


Figura 2.28 - Protocolos de *multicast* IGMP x PIM.

Em IPTV o MPEG-2 TS é encapsulado pelo IP *multicast* no nível *ethernet*, o que é outra grande vantagem, devido ao baixo custo das redes *ethernet* comparado com outras tecnologias.

Quando um cliente se "junta" a um grupo, dois processos são iniciados: primeiro, uma mensagem de IGMP (IGMP *join*) é enviada ao roteador local do cliente para informar que o cliente quer receber o MPEG-2 TS daquele novo grupo. Em seguida, o cliente ajusta sua interface de rede IP para receber o *multicast* no novo endereço e porta de rede (ver Figura 2.29 - IGMP *join*).

Quando um cliente deixa um grupo, uma mensagem (IGMP *leave*) é enviada ao roteador local. Se esse era o último cliente desse grupo, o grupo deixa de existir após um tempo, liberando assim recursos na rede.

A faixa reservada para endereços IP *multicast*, também chamado de endereços classe D, vai de 224.0.0.0 até 239.255.255.255, mas

- 224.0.0.0 é reservado e não pode ser utilizado por nenhum grupo;
- 224.0.0.1 até 224.0.0.255 reservado para o uso de protocolos de roteamento dinâmico e outros protocolos;
- 224.0.1.0 até 239.255.255.255 são designados para varias aplicações *multicast* ou permanecem disponíveis;
- 239.0.0.0 até 239.255.255.255 *muticast* privados.

Receptores *multicast* geralmente submetem seus pedidos para o equipamento mais próximo para se "juntar" a um grupo (ou sessão) ativo. Caso não exista esse grupo, ou seja, caso ninguém tenha solicitado o conteúdo de um determinado grupo e um usuário o deseja, ele será criado dinamicamente.

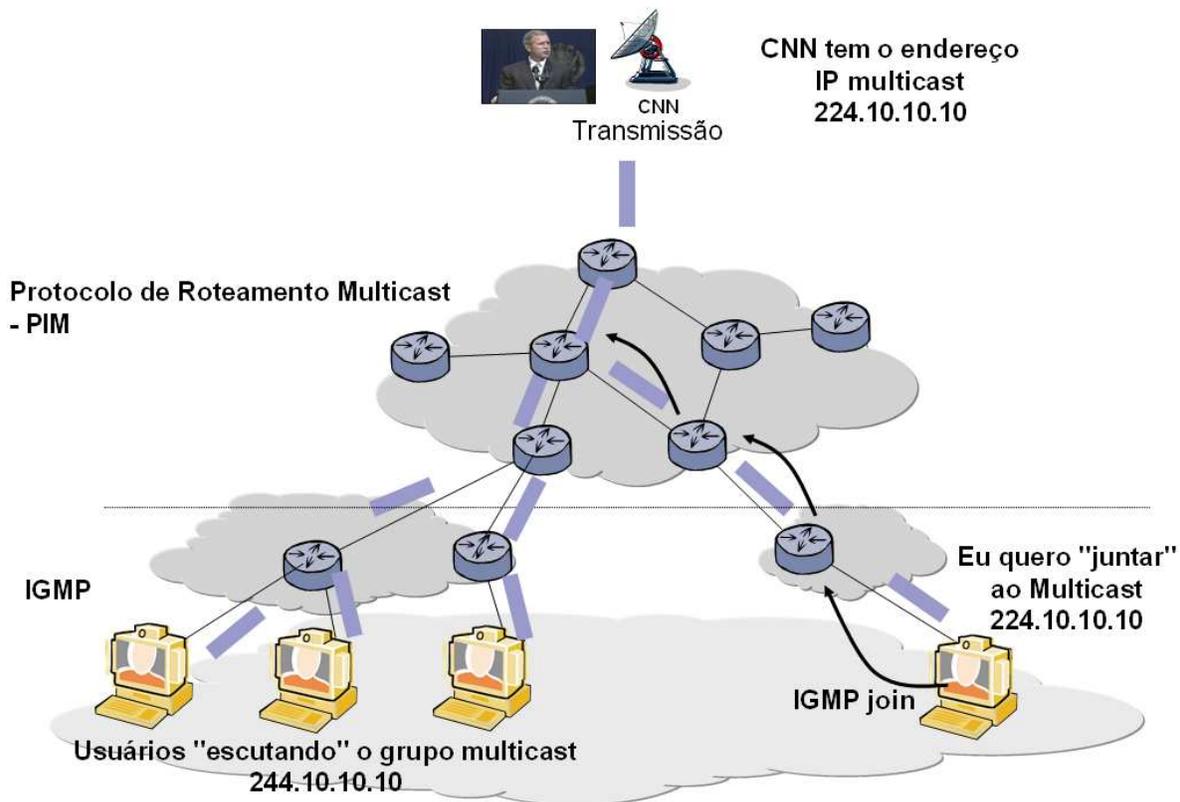


Figura 2.29 - IGMP join.

Para que um sistema *multicast* possa funcionar, todos os nodos (roteadores) da rede têm que ser capazes de tratar as sessões, seja pelo protocolo IGMP (usuário para nodo de acesso) ou pelo PIM (nodo para nodo no núcleo da rede).

Transmissões de TV ao vivo e PPV (*Pay-Per-View*) para IPTV usam o *multicast*. Já o VOD (*Video on Demand*) e o MOD (*Music on Demand*) utilizam o *unicast*. Isso se deve à característica dos serviços sob demanda, que são individuais. Esses serviços serão discutidos com mais detalhes no capítulo 3.

RTSP

O RTSP (*Real Time Streaming Protocol* ou Protocolo de *Streaming* em Tempo Real) é um protocolo no nível de aplicação desenvolvido pela IETF (*Internet Engineering Task Force*) em 1998 com a RFC 2326 para controle na transferência de dados com propriedades de tempo real. RTSP torna possível a transferência, sob demanda, de dados em tempo real como áudio e vídeo além de permitir que o cliente controle remotamente o fluxo, executando comandos como se fosse um DVD, como pausar, avançar, retroceder, parar e reproduzir [6, 48, 49].

O RTSP é um protocolo de aplicação e não é encarregado de transportar o conteúdo.

Esse transporte pode ser o TCP ou o UDP. Como foi apresentado na sessão anterior, o UDP é mais adequado para *streaming* de vídeo e será sempre utilizado para esses casos.

No caso do IPTV, um servidor de vídeo irá armazenar seus conteúdos no formato MPEG-2 TS. O formato é exatamente igual ao de um fluxo de tempo real e a diferença é que esse fluxo será capturado e guardado em um arquivo digital. Dessa forma, esse conteúdo poderá ter mais de um áudio e legendas, como mostrados na sessão do MPEG-2 TS. Nesse caso, o usuário poderá mudar o áudio ou selecionar legendas, dependendo da implementação.

A transmissão desse conteúdo pode ser feita em *multicast* ou *unicast*, dependendo da implementação. Pelo que foi discutido na sessão do *IP Multicast*, o *unicast* não é muito otimizado quando vários usuários estão solicitando o mesmo conteúdo ao mesmo tempo. Dessa forma, se esse for o caso, melhor seria o uso do *multicast*.

Entretanto, caso o usuário esteja disposto a pagar um valor mais caro para ter um conteúdo somente para ele e ter controle total sobre o fluxo, o *unicast* deve ser usado.

As diferentes formas de implementação e serviços, bem como as características do servidor de vídeo serão discutidas no capítulo 3.

O protocolo RTSP é similar em sintaxe e operação ao protocolo HTTP (*HyperText Transfer Protocol*), mas o RTSP adiciona novas solicitações. Enquanto o HTTP não é orientado a conexão (*stateless*), o RTSP necessita estabelecer uma conexão antes (*statefull*) e uma identificação da sessão é usada para rastreá-la sempre que necessário e dessa forma, não necessita de uma conexão TCP.

As mensagens RTSP são enviadas do cliente para um servidor sendo que um número típico de pedidos HTTP é frequentemente usado. Essas mensagens podem ser vistas na Figura 3.19 - Comandos RTSP.

Página em branco

Capítulo 3

Componentes e Redes do IPTV

Introdução

A solução *triple play* oferece uma única rede IP capaz de entregar voz, vídeo e *Internet* usando-se uma conexão de banda larga. Isso tudo devido à arquitetura flexível da rede IP que, além de ser a infra-estrutura emergente dos últimos dez anos para inovação de serviços, suporta os diferentes requisitos para cada serviço, como Qualidade de Serviço (QoS), largura de banda e segurança. Mais importante do que isso, uma vez que todos os três serviços usam uma rede IP em comum, tem sido possível desenvolver novos serviços combinando-se elementos de voz, vídeo e dados. Existem algumas soluções que se dizem *triple play*, mas que na verdade simplesmente entregam todos os três serviços (voz, dados e TV) juntos. Já o IPTV, que é conhecido como *triple play* real, não somente entrega os serviços juntos desde uma mesma conexão de banda larga, como também os integra de forma transparente [1, 13, 14, 17]. Dessa forma, um usuário enquanto assiste TV, pode receber notificações de correio eletrônico na tela da TV, a identificação de uma chamada também na TV, a visualização da programação da TV na tela do computador e uma série de outros serviços já existentes ou que estão para serem lançados. Uma descrição dos principais serviços de IPTV será mostrada no final deste capítulo [23].

As partes funcionais da arquitetura de IPTV são o *head-end*, onde se tem a aquisição e tratamento do conteúdo, integrada com codificadores e ferramentas de encapsulamento; *middleware*, responsável pela gerência, entrega e consumo dos serviços; proteção do conteúdo, realizando a criptografia em tempo real e a gerência dos direitos digitais (DRM) dos conteúdos; e as premissas dos clientes, com os equipamentos de terminação e decodificação do lado do usuário (ver Figura 3.1 - Diagrama de blocos para um sistema de IPTV).

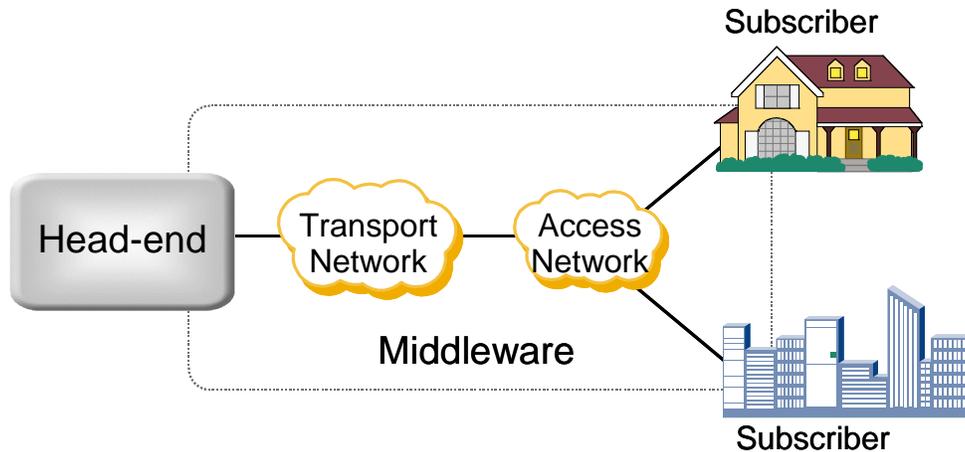


Figura 3.1 - Diagrama de blocos para um sistema de IPTV.

O *head-end*, o *middleware*, os servidores de VOD, a rede de transporte e de acesso, o acesso condicional, os servidores de proteção do conteúdo (DRM), e os decodificadores do cliente (STB - *Set Top Box*) são os elementos principais de uma solução de IPTV e são mostrados na Figura 3.2 - Elementos Principais de uma rede de IPTV.

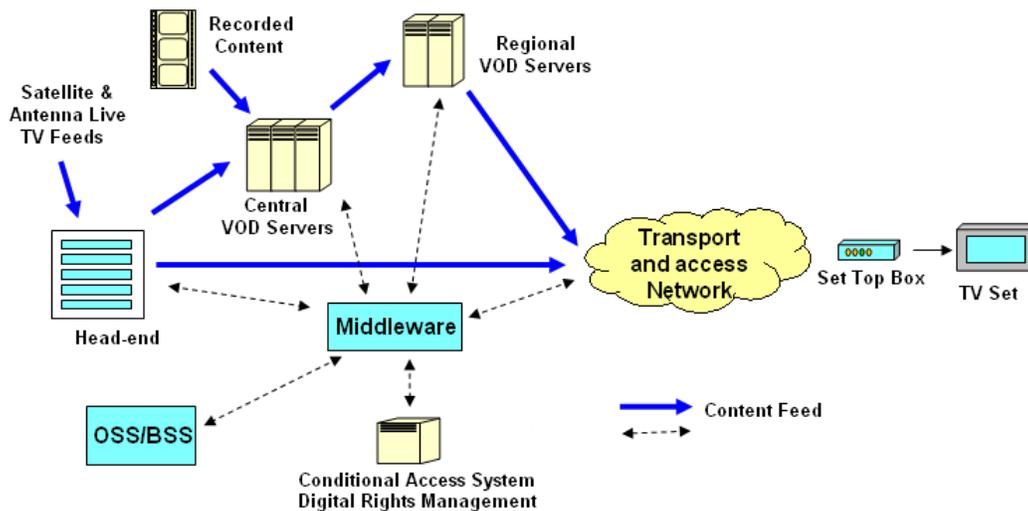


Figura 3.2 - Elementos Principais de uma rede de IPTV.

Head-End

O *head-end*, ou cabeceira de vídeo é a área primária onde o conteúdo é adquirido e tratado para alimentar a rede do IPTV, e assim executa várias funções.

Assim como nas soluções de TV a cabo ou por satélite, um serviço de IPTV requer um

head-end. Esse é o ponto da rede onde os canais de TV e os conteúdos sob demanda são recebidos, comprimidos, tratados e formatados para distribuição na rede IP.

Aquisição do Conteúdo

O *head-end* contempla também o parque de antenas e receptores de satélite (IRD - *Integrated Receiver Decoder*) para a aquisição dos sinais dos diversos satélites.

A antena de satélite recebe o sinal do ar proveniente do provedor de conteúdo. Esse sinal, comumente codificado e cifrado, é enviado ao IRD, como mostrado na Figura 3.3 - Recepção de TV via satélite.

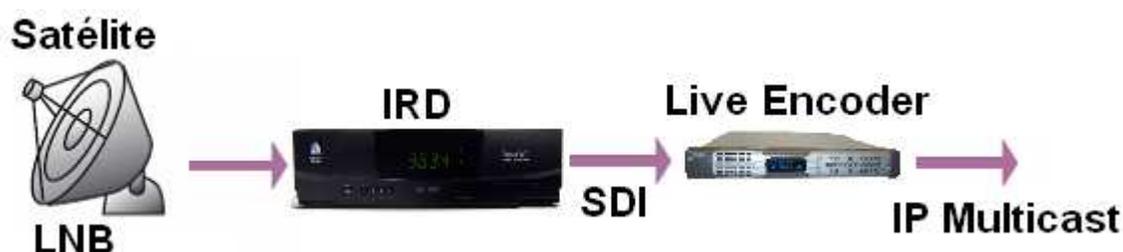


Figura 3.3 - Recepção de TV via satélite.

Antes da transmissão os provedores de conteúdo cifram o sinal com chave secreta, para evitar que outras pessoas possam captar o sinal do ar. Além disso, a taxa de compressão utilizada não é apropriada para envio para os usuários finais, porque o codificador utilizado pelo provedor de conteúdo, tentando otimizar o seu uso, combina diferentes canais. Exemplo: canal de esporte com canal de notícias. Quando um canal é codificado de forma separada, é possível tirar o máximo de proveito da tecnologia de compressão e reduzir a largura de banda.

Então das antenas são recebidos sinais MPEG-2 TS modulados em QPSK das bandas C, L e Ku. Um conversor LNB (*Low Noise Block*) é usado na recepção por satélite e é usualmente fixado no prato da antena. O papel do LNB é transformar uma banda de frequências relativamente altas e convertê-las para frequências mais baixas. Essas baixas frequências trafegam pelo cabo até o IRD com uma menor atenuação, além de baratear o uso dos circuitos eletrônicos envolvidos.

O LNB é também usado para amplificação do sinal antes da atenuação do cabo e o bloco é livre de equipamentos adicionais como fonte de alimentação ou receptor digital. Isso tudo contribui para que o ruído na saída seja o mais baixo possível.

Alguns satélites aceitam o uso de um LNB universal (Banda Ku) com polarização linear (Vertical ou Horizontal). Entretanto, alguns outros exigem uma polarização circular

(Sentido horário ou anti-horário). Isso tudo depende do provedor de conteúdo. Na Figura 3.4 - Recepção satelital é mostrada uma recepção satelital típica de vários satélites adjacentes, onde um prato de antena fixo é necessário para cada satélite.

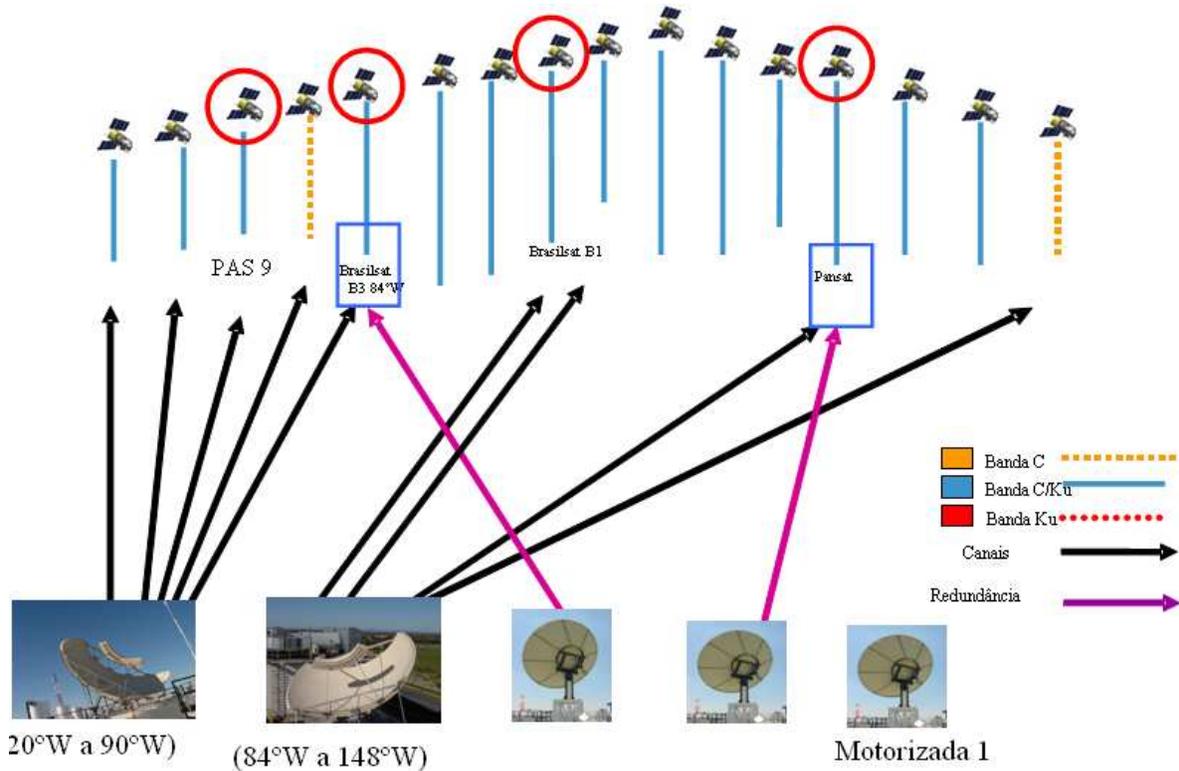


Figura 3.4 - Recepção satelital.

Do LNB, o sinal é enviado ao IRD via cabo coaxial. O IRD é utilizado para demodular, decifrar e decodificar o conteúdo recebido por satélite. A saída do IRD será um sinal sem compressão, usualmente um sinal SDI (*Serial Digital Interface* ou interface digital serial). O fluxo SDI é transmitido do IRD para o codificador.

SDI é um padrão ITU-R BT.656 e SMPTE-259M que é uma interface de vídeo digital, com definição padrão (SD), sem compressão usada em transmissão *broadcast*, com largura de banda típica de 270 Mbps. Uma tecnologia relacionada é a chamada HD-SDI (*High Definition Serial Digital Interface* ou SDI de alta definição), padronizada pela norma SMPTE-292M, com velocidade 1.485 Gbps [44, 45, 46].

Uma tecnologia emergente é a SMPTE-372M, uma interface nominal de 3 Gbps usada em algumas aplicações, como cinema digital. Ela requer maior fidelidade e resolução do que um HDTV padrão pode fornecer.

Esses padrões são usados para transmissão sem compressão, sem criptografia e

apropriadas para distâncias curtas.

Um *head-end* típico também recebe sinais de TV aberta pelo ar. Uma vez que cada canal terrestre representa uma frequência definida, é necessária uma antena receptora para cada canal que o provedor queira receber. Essas antenas são frequentemente montadas em uma estrutura de torre.

Alternativamente, outras fontes de programação podem ser distribuídas via fibra óptica, microondas ou outro canal dedicado.

Para o recebimento de canais de TV aberta, faz-se uso de antenas e de-moduladores, como mostrados na Figura 3.5 - Recepção de TV aberta do ar.

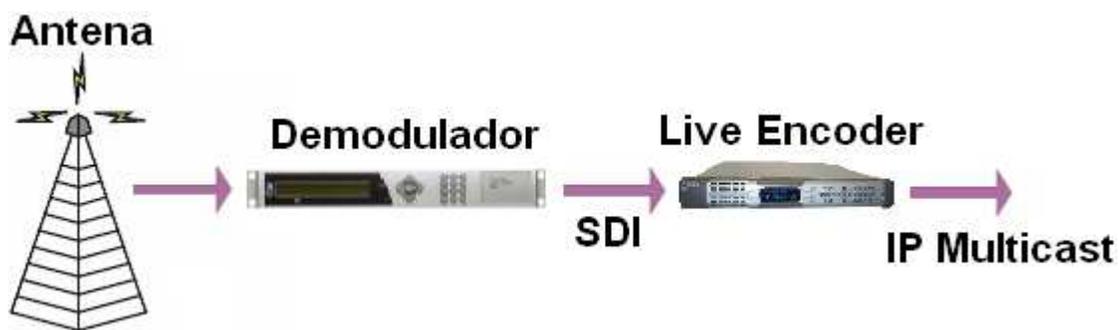


Figura 3.5 - Recepção de TV aberta do ar.

Para ambos os casos, via satélite ou via antena, os sinais SDI são enviados para o codificador (*encoder*) que vai comprimir os fluxos elementares de vídeo e áudio(s), gerar o MPEG-2 TS e fazer o encapsulamento *IP multicast*.

Os codificadores são usados para comprimir os *streams* na menor taxa possível e com taxa de bits constante (CBR ou *Constant Bit Rate*). O sinal de vídeo pode ser comprimido usando MPEG-2, H.264 ou VC-1. Os codificadores podem também comprimir o sinal com uma resolução e uma velocidade bem mais baixa, que poderá ser usado para criação de serviços PiP (*Picture in Picture*), mosaico dinâmico ou conteúdo para dispositivos móveis. A velocidade desse canal é da ordem de 200 kbps.

Os codificadores podem ser baseados em *hardware* ou em *software*. Entretanto, os codificadores em *hardware* fornecem uma melhor confiabilidade e desempenho, comparados com os baseados em *software*.

Na configuração dos codificadores, podem se configurar diversos parâmetros incluindo o padrão de compressão para vídeo e áudio, a resolução, a largura de banda para cada fluxo elementar, a largura de banda do fluxo de transporte e quais tabelas PSI devem ser inseridas no MPEG-2 TS. Além desses valores típicos, é possível também a configuração de parâmetros avançados como tamanho do GOP, se o GOP vai ser fixo ou variável, ou até

mesmo do número de quadros Bs entre os quadros Ps.

Uma vez configurados esses valores, e além de informar o número de canais de áudio existentes, dependendo do que está chegando do provedor do conteúdo, o codificador vai criar o MPEG-2 TS e fazer o encapsulamento de tudo isso em IP *multicast*. Após a atuação dos codificadores se fará uso de um *switch ethernet* para fazer a agregação de todos os canais. Isso será visto na sessão de rede.

Alguns codificadores podem incluir no MPEG-2 TS uma entrada de áudio pré-codificada, se disponibilizada pelo provedor de conteúdo, como por exemplo, um *Dolby AC-3* com 5.1 canais. Nesse caso, o STB do usuário final poderá extrair o conteúdo e usufruir o uso desses cinco canais.

Os codificadores estão constantemente enviando os dados em direção a rede IP, em tempo real (*live streaming*), ou seja, os codificadores fazem a compressão e o *streaming*. O usuário final vai receber esse *streaming* e tocá-lo tão logo o receba, também em tempo real, diferente de baixar todo o conteúdo e depois tocar, como acontece em alguns casos na *Internet*.

A saída dos codificadores também pode ser desviada e capturada para um armazenamento posterior nos servidores de VOD. Note que a compressão será usada para TV e para vídeo sob demanda.

A compressão é necessária para reduzir a largura de banda requerida para a transmissão para os assinantes. Por exemplo, um sinal sem compressão consome 270 Mbps. Os codificadores atuais têm a capacidade de fornecer imagens de boa qualidade com taxas da ordem de 2 Mbps, usando-se codificadores de nova geração (H.264 ou VC-1).

Os *head-ends* podem ser ainda nacionais ou regionais. É muito comum que o provedor de serviço tenha um *head-end* nacional ou principal para uma grande parte dos canais e pequenos *head-ends* locais para a aquisição dos canais de uma determinada cidade ou região.

Provedor de Conteúdo

Os provedores de conteúdo de TV comercializam seus conteúdos para os diversos provedores de serviços. O conteúdo é um dos fatores mais importantes para a boa qualidade dos canais de IPTV já que o assinante busca a diversidade e interatividade. Esse último aspecto diferencia uma IPTV da TV por assinatura.

O grande desafio dos provedores de conteúdo é criar uma proteção contra pirataria, já que o assinante tem a possibilidade de programar os conteúdos, horários de exibição e até mesmo gravar utilizando o *Set Top Box* (STB). Hoje os provedores de conteúdo buscam parcerias com os grandes produtores de cinema e TV, principalmente os norte-americanos.

Dessa forma, os provedores de conteúdo impõem várias restrições a seus provedores de serviços e isso varia de programadora para programadora. A lista de possíveis restrições inclui: os canais devem ser cifrados ou ter algum tipo de proteção; alguns canais não podem ser gravados; a qualidade final da imagem deve ser excelente; inserção de anúncios de comerciais para alguns canais não é liberada; dentre outras.

Além disso, alguns provedores de conteúdo podem ainda cobrar uma taxa adicional de acordo com o número de usuários finais do provedor de serviço.

Caso alguma restrição não seja respeitada, o provedor de conteúdo pode cobrar multas ou até mesmo retirar o canal da lista do ar para aquele provedor de serviço.

A Figura 3.6 - Provedor de conteúdo mostra o envio de conteúdo via satélite do provedor de conteúdo para provedor de serviço.

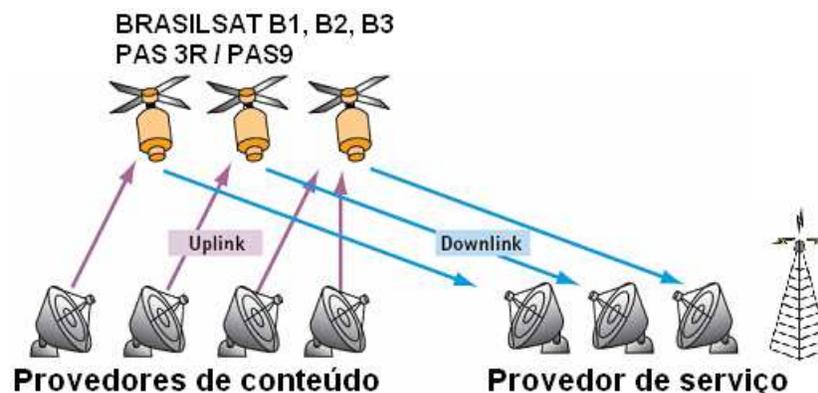


Figura 3.6 - Provedor de conteúdo.

Inserção de Legenda

Como discutido no capítulo 2, os fluxos de legenda são considerados fluxos elementares. Dessa forma, a primeira premissa para que o sistema seja implementado no sistema de IPTV é fazer uma inserção do(s) fluxo(s) elementar de legenda no MPEG-2 TS. Além disso, a aplicação de TV tem que suportar essa função e com isso a aplicação de TV habilitará um botão para que o usuário possa selecionar uma legenda. Se vários fluxos de legendas, com diferentes PIDs, são inseridos no fluxo de transporte (MPEG-2 TS), então o usuário terá a flexibilidade de escolher qual legenda ele quer ver, ou mesmo, escondê-la. Finalmente, o *Set Top Box* deve ser capaz de decodificar o formato da legenda.

O formato padrão de legenda é o DVD *bitmap*, ETS 300 743 (*Digital Video Broadcasting Subtitling System*), que geralmente difere do padrão proprietário enviado pelo provedor de conteúdo. Para resolver esse problema, são usados transcodificadores de

legenda para um formato compreensível pelo STB [7, 37].

As legendas que constituem uma parte da solução de IPTV, é um serviço interativo e os usuários podem selecionar um idioma usando um controle remoto. Esse serviço começa a se tornar uma importante característica para as necessidades do mercado.

Os fluxos de legenda são (trans) codificados e inseridos no MPEG-2 TS e transmitidos para os STBs.

Inserção de comercial

A solução de IPTV oferece a habilidade para o provedor de serviço inserir, criar, substituir e programar conteúdos de publicidade na sua grade de programação. Essa funcionalidade é conhecida como a inserção de comercial (*Ad Insertion*) e é uma característica muito atrativa, sendo assim uma poderosa fonte de faturamento.

A inserção de comercial é o processo de inserir conteúdo de publicidade nos canais de TV. Para sistemas *broadcast*, os comerciais são tipicamente inseridos numa base nacional ou geográfica. No sistema de IPTV esses anúncios podem ser direcionados para um público específico com base no seu perfil.

Para arquitetar essa solução, diversos equipamentos adicionais são necessários. Na maioria dos casos, a programadora manda os tons (*cue tones* ou DTMF) embebidos no canal esquerdo de áudio, indicando início ou fim de comercial. Pode-se ainda fazer uso de uma tecnologia digital, conhecida como SCTE-35 que é usado como gatilho (*trigger*) digital [7, 26, 47].

Por último, é necessário um servidor de comerciais bem como um equipamento conhecido como *splicer*, que serão responsáveis por substituir a programação original pelo comercial apropriado. A Figura 3.7 - Inserção de comercial – *splicer* mostra o *splicer* que seleciona a saída de duas ou mais entradas para produzir uma saída.

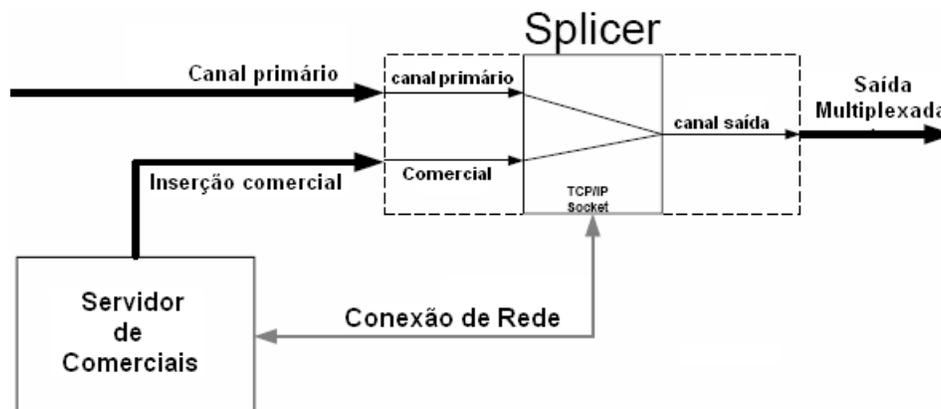


Figura 3.7 - Inserção de comercial – *splicer*.

O *splicer* recebe as mensagens DTMF (*cue tones*) de início e fim de comercial que informam quando ele deve comutar as entradas.

Os comerciais são, em geral, conteúdos de curta duração de 10 ou 15 segundos e depois disso, o *splicer* volta a comutar para a entrada original. O mesmo princípio pode ser usado para inserção de programação local de maior duração como, por exemplo, notícias locais.

Uma vez que a mídia digital é composta de quadros Is, Ps e Bs que compõem uma estrutura de GOP, a inserção de comercial digital é mais complexa do que a inserção analógica. Por isso, o elemento chave da inserção, o *splicer*, deve ter a capacidade de comutar exatamente em um quadro I, para evitar sobreposição ou distorção de imagens.

Na Figura 3.8 - Diagrama de blocos da inserção de comercial são mostrados os principais equipamentos da solução.

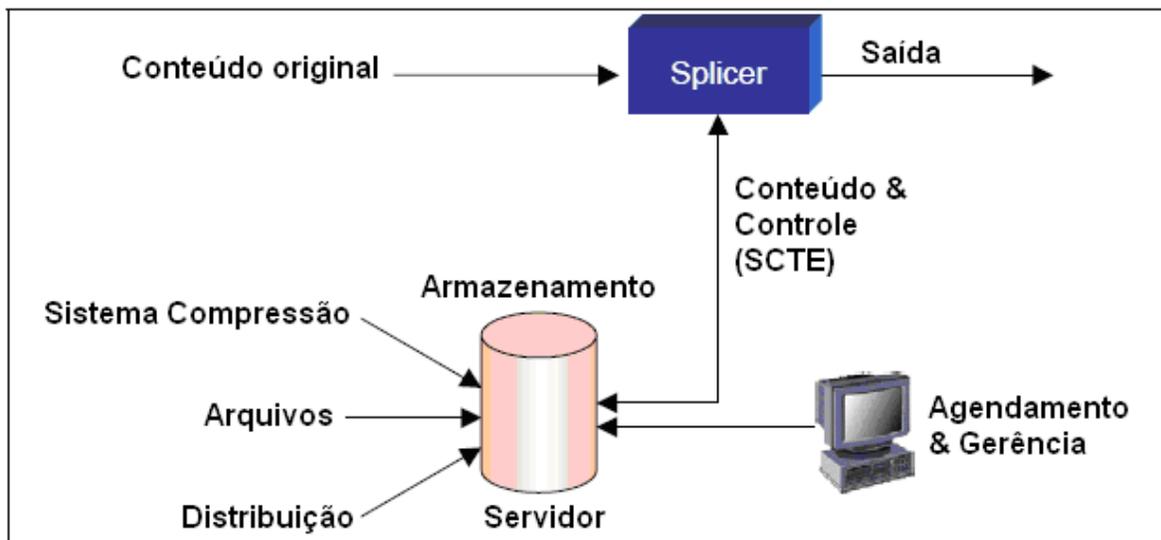


Figura 3.8 - Diagrama de blocos da inserção de comercial.

Um servidor de comercial é usado para guardar todos os comerciais que devem ser tocados. Uma estação de agendamento e controle é usada para programar os comerciais por canal e por horário. Por exemplo, se chegar um *cue-tone* no canal 1, entre as 12h55min e 13h05min, então deve-se tocar o comercial da “coca-cola”.

Vale ainda ressaltar que a solução de inserção de comercial junto com uma medição de audiência real é uma ferramenta poderosa para o pessoal de *marketing*. É possível saber ao certo quantas pessoas assistiram a um determinado comercial. Isso será visto na sessão de medição de audiência.

Além disso, os operadores de IPTV podem ganhar mais dinheiro com publicidade, medindo a reação dos telespectadores e direcionando melhor os comerciais. É possível

saber se o usuário mudou de canal na hora do comercial ou se o assistiu por completo.

Middleware

Tipicamente uma arquitetura cliente/servidor, onde o cliente reside no equipamento terminal do usuário (STB), o *middleware* foi criado para designar camadas de *software* que não constituem diretamente aplicações, mas que facilitam o uso de ambientes ricos em tecnologia da informação. A camada de *middleware* concentra serviços como identificação, autenticação, autorização, tarifação, diretórios, certificados digitais e outras ferramentas para segurança.

O *middleware* é responsável por oferecer uma gerência central dos componentes de IPTV, como os servidores de vídeo, os servidores de acesso condicional (CAS), a gerência dos direitos digitais (DRM), os terminais de usuários (STB) e a garantia da entrega do conteúdo. É também função do *middleware* a criação do modelo de negócio: definição de preços dos serviços, pagamentos, descontos, limites de usuários, controle paterno, administração de usuários, proteção de conteúdo, além de estar integrado com os sistemas de faturamento e demais sistemas legados das operadoras.

Além disso, ele é responsável por prover a navegação e a visualização das telas pelo cliente, isto é ele hospeda e oferece a interface de usuário, comumente por meio de páginas *web*.

Apesar de já existir um IPTV fórum, ainda não foi recomendado nenhuma padronização para os servidores do *middleware*.

O *middleware* inclui servidores de *web cache*, servidores de aplicação ou servidores *web*, servidores de base de dados, sistemas de gerência de conteúdos e ferramentas que suportam o desenvolvimento e a entrega de serviços.

Servidores *web cache*, são servidores temporários que armazenam os últimos dados solicitados pelo cliente e, se usado de forma adequada, eliminam a necessidade de solicitar repetidos dados aos servidores de aplicação (AS ou *Application Servers*) e aos servidores de base de dados (DB ou *Data Base*). Além de acelerar a entrega de dados comuns, os servidores de *web cache* também otimizam o desempenho, reduzindo a carga nos AS e DB. A Figura 3.9 - Componentes do *middleware* mostra os servidores *web cache*, os AS e o servidores DB que são os principais componentes do *middleware*.

Além da função de armazenar dados estáticos e dinâmicos, os servidores *web cache* servem também como uma barreira de segurança para as máquinas principais do *middleware*. Dessa forma, os clientes (STBs) não conseguem alcançar as máquinas dos ASs e muito menos as dos servidores DBs. Assim, um *hacker* jamais conseguirá conectar-se no servidor da base de dados.

Uma vez que os servidores de *web cache* podem guardar dados dinâmicos, como dados de programação eletrônica, eles devem conter também algumas políticas para invalidar os dados periodicamente.

O fluxo do STB para o *web cache* funciona da seguinte forma: o cliente STB solicita dados ao servidor *web cache* que responde imediatamente caso o objeto exista. Se o objeto não está armazenado, o *web cache* o solicita para os servidores de aplicação que podem ou não consultar a base de dados. O AS retorna o valor para o *web cache* e se a resposta for um valor armazenável, o *web cache* vai armazenar uma cópia para a próxima consulta e responder ao STB.

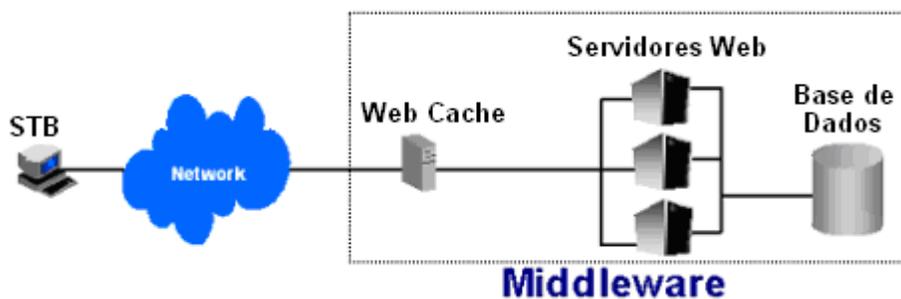


Figura 3.9 - Componentes do *middleware*.

Parte também do pacote do *middleware*, os servidores de aplicação (AS) tratam a maioria, se não toda a lógica da aplicação, como uma parte complementar nas premissas do cliente. Eles executam a gerência do conteúdo, gerência dos usuários, relatórios e gerência do negócio, além de naturalmente prover a interface final do usuário.

Os ASs são também servidores *web* que são sistemas de computadores usados para prover acesso aos dados armazenados e recuperados por comandos HTTP. O protocolo HTTP é usado para solicitar e coordenar a transferência de documentos entre o servidor *web* e o cliente. Tipicamente os servidores *web* permitem que navegadores de *Internet* (*Browsers*) solicitem e processem as informações da *Internet*.

No caso do IPTV, os servidores *web* vão hospedar as diferentes páginas ou os *frameworks* que serão mostrados na TV do cliente.

Os servidores de base de dados representam o núcleo da solução. É a parte mais importante da solução, porque vai conter os dados de todos os clientes, os diferentes produtos disponíveis para compra, os preços de cada serviço, os consumos de cada cliente e os limites de compra. Assim, é função dos servidores DB a autenticação dos usuários (STBs), a autorização a serviços e produtos, dependendo do pacote de cada cliente e a tarifação de todos eles.

Os clientes (STBs) têm um navegador de *Internet* no seu sistema operacional, mas o cliente não percebe, uma vez que os botões e barras de ferramentas estão escondidas. Entretanto, quando se liga um STB pela primeira vez, ele vai executar um comando HTTP para o servidor *web cache*, tentando alcançar sua página principal (*Home Page*) que estará previamente configurada em cada STB. Como esse é o primeiro acesso, o usuário tem que ser autenticado. Dessa forma, os servidores *web cache*, solicitam o acesso a um servidor AS. Esse servidor AS por sua vez vai consultar a base de dados, para saber se aquele usuário existe e se ele está ativo. Essa autenticação pode ser feita de diversas maneiras, mediante o uso de um par usuário/senha, que aparecerá na tela do usuário na primeira vez, ou através do seu endereço de rede (*MAC address*), ou através do seu IP, ou mesmo através de certificados Microsoft.

Uma vez que esse usuário é corretamente autenticado, entra agora uma parte da autorização, que irá autorizar tudo o que o cliente pode fazer e tudo o que ele pode comprar. Aqui entram os pacotes de TV aos quais ele vai ter acesso, quanto ele pode gastar por mês, se ele pode ter acesso a conteúdo adulto, se ele pode comprar filmes, etc.

Como foi visto no capítulo de *head-end*, cada canal de TV, terá um endereço *multicast* saindo do codificador, isto é, um MPEG-2 TS encapsulado por um endereço IP *multicast*, como por exemplo, o endereço 224.10.10.5 com a porta 1003. Entretanto, os usuários finais não sabem o que é endereço IP. Eles sabem, por exemplo, que o canal *Fox* é o número 53. Dessa forma, é função do *middleware* também a criação dos canais, que nada mais é que a associação de um endereço IP *multicast*, porta, a um nome e número de canal. A associação de um ou mais canais de TV constituem um pacote de TV. Esse pacote de TV terá um preço que será debitado da fatura do assinante. A Figura 3.10 - Exemplo de criação de um canal de TV mostra uma janela típica para criação de um canal de TV.

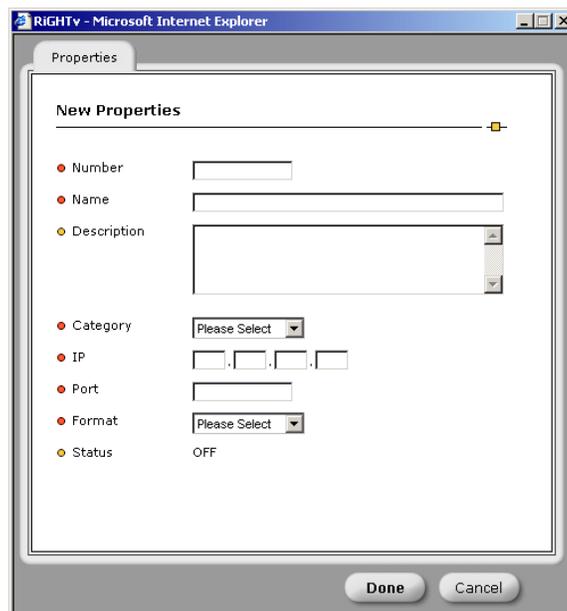


Figura 3.10 - Exemplo de criação de um canal de TV.

A Figura 3.11 - Exemplo de criação de um pacote de TV mostra uma associação de um ou mais canais de TV formando um pacote. Esse pacote terá um preço definido pelo provedor de serviço.

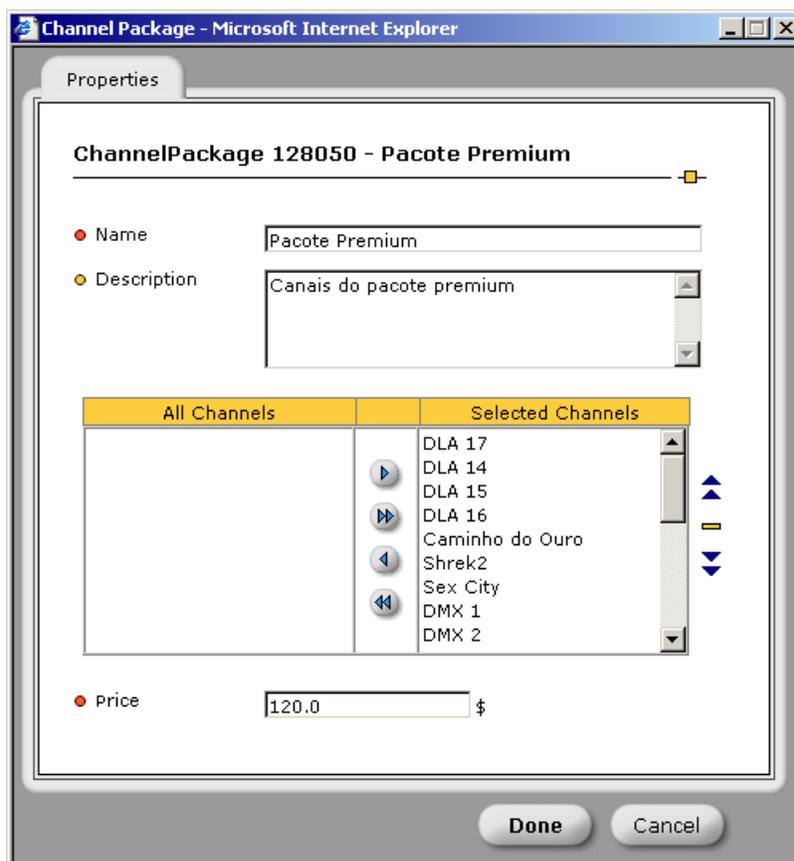


Figura 3.11 - Exemplo de criação de um pacote de TV.

A Figura 3.12 - Exemplo de criação de um usuário mostra uma janela típica para criação de um usuário. Nesse exemplo aqui, o usuário tem um nome de usuário e senha para acesso, além de senhas para compra de serviços *on-line*. O usuário aqui é ainda identificado por seu endereço MAC e tem um ou mais pacotes de TV associados a sua conta.

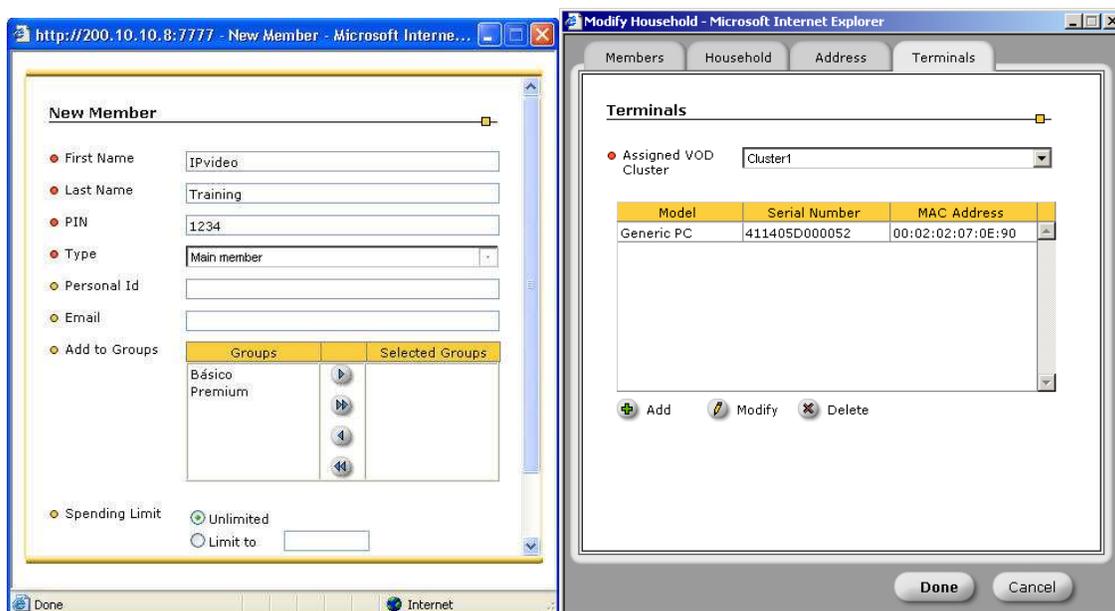


Figura 3.12 - Exemplo de criação de um usuário.

Vale ressaltar, que os fluxos de canais não passam pelo *middleware*. O *middleware* faz apenas um controle e um gerenciamento externo. A Figura 3.13 - Fluxo para serviços de TV mostra os elementos envolvidos na parte de serviços de TV.

Dessa forma, quando um usuário aperta o número 53 no seu controle remoto, o *middleware* vai verificar na base de dados se esse usuário tem acesso a esse pacote/canal e traduzir esse número em um comando “igmp://224.10.10.5:1003” que será enviado de volta ao STB. Dessa forma, o STB vai fazer uma solicitação para a rede do comando acima, que é um comando *multicast*.

Entretanto, usuários mais avançados (um *hacker*) poderiam tentar executar o comando acima para um canal que ele não tenha acesso. Para que um *hacker* possa fazer isso, ele deve configurar um computador com os mesmos dados de endereço IP do STB e se fazer passar por um STB. Nesse caso, esse usuário mal intencionado, não se autenticaria no *middleware*, mas tentaria apenas assistir aos canais de TV aos quais ele não tem acesso. Caso isso ocorra, ele não irá passar pelo *middleware*, que por sua vez não fará a devida verificação. Devido a isso, uma proteção adicional deve ser inserida. Essa proteção pode ser através da criptografia dos canais ou mesmo através de controle de acesso dos endereços IPs no equipamento de acesso ao cliente. Esse tópico será abordado em mais detalhes no item de proteção de conteúdo.

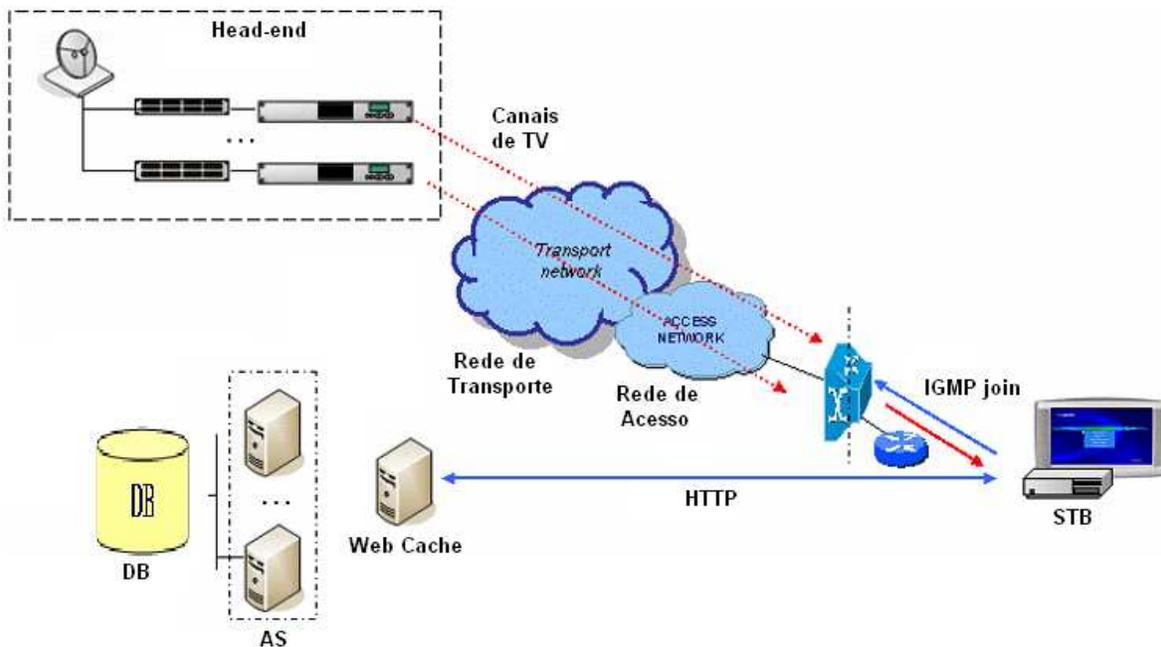


Figura 3.13 - Fluxo para serviços de TV.

Integração com OSS/BSS

A solução de IPTV, por meio de *middleware*, é integrada com os sistemas de apoio a operações (OSS ou *Operating Support System*), e sistemas de apoio ao faturamento (BSS ou *Billing Support System*) da operadora através de suas interfaces de programação (API ou *Application Programming Interface*) conforme mostra a Figura 3.14 - *Application Programming Interfaces*.

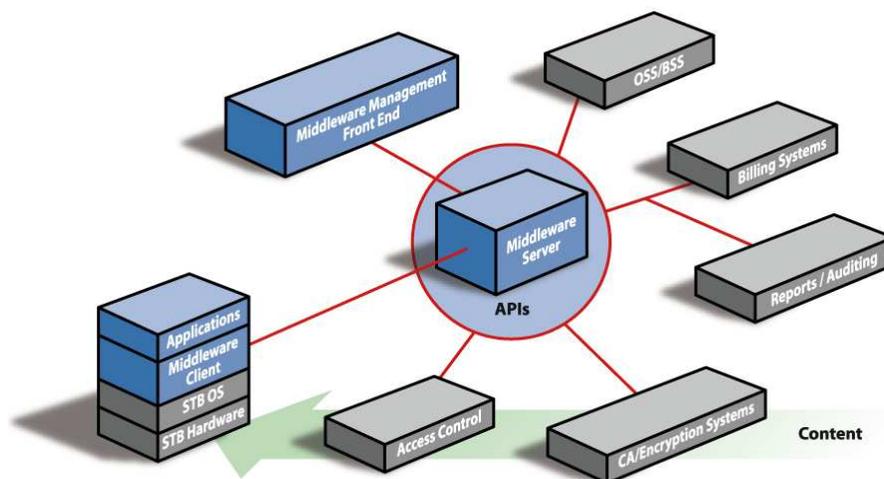


Figura 3.14 - *Application Programming Interfaces* (API).

Em geral, essas APIs fazem uso de interfaces modernas como XML (*Extensible Markup Language* ou linguagens de marcação para necessidades especiais), SOAP (*Simple Object Access Protocol*) e serviços *web* que são usados para a integração entre os mais diversos servidores.

XML é uma linguagem padronizada de marcação capaz de descrever diversos tipos de dados. Seu propósito principal é a facilidade de compartilhamento de informações por meio de *Internet*.

Estimulado pela insatisfação com os formatos existentes, um grupo de empresas e organizações que se autodenominou *World Wide Web Consortium* (W3C) começou a trabalhar em meados da década de 1990 em uma linguagem de marcação que combinasse a flexibilidade da SGML (*Standard Generalized Markup Language*) com a simplicidade da HTML. O princípio do projeto era criar uma linguagem que pudesse ser lida por *software*, e integrar-se com as demais linguagens. Sua filosofia seria incorporada por vários princípios importantes incluindo a separação entre conteúdo e formatação; a simplicidade e a legibilidade, tanto para humanos quanto para computadores; a possibilidade de criação de *tags* sem limitação; a criação de arquivos para validação de estrutura; a interligação de bancos de dados distintos; e a concentração na estrutura da informação, mas não na sua aparência.

O XML é considerado um bom formato para a criação de documentos com dados organizados de forma hierárquica, como se vê frequentemente em documentos de texto formatados, imagens vetoriais ou bancos de dados. Pela sua portabilidade, um banco de dados pode, por meio de uma aplicação, escrever em um arquivo XML, e outro banco distinto pode ler então esses mesmos dados.

O SOAP é um protocolo usado para troca de informações, sendo essas estruturadas em uma plataforma descentralizada e distribuída, utilizando-se tecnologias baseadas em XML. Sua especificação define um *framework* que provê maneiras para se construir mensagens que podem trafegar por meio de diversos protocolos e que foi especificado de forma a ser independente de qualquer modelo de programação ou outra implementação específica.

Geralmente servidores SOAP são implementados utilizando-se servidores HTTP, embora isto não seja uma restrição para funcionamento do protocolo. As mensagens SOAP são documentos XML que aderem a uma especificação fornecida pelo órgão W3C.

Os servidores do *middleware* devem ser integrados não somente aos sistemas OSS/BSS como também aos diversos outros servidores da solução de IPTV. A Figura 3.15 - Integração entre sistemas mostra as integrações típicas de uma solução, que são interfaces com um sistema de provisionamento, com sistemas de tarifação, com servidores de vídeo, com servidores de EPG e com os servidores de proteção de conteúdo CAS/DRM todas elas usando APIs SOAP.

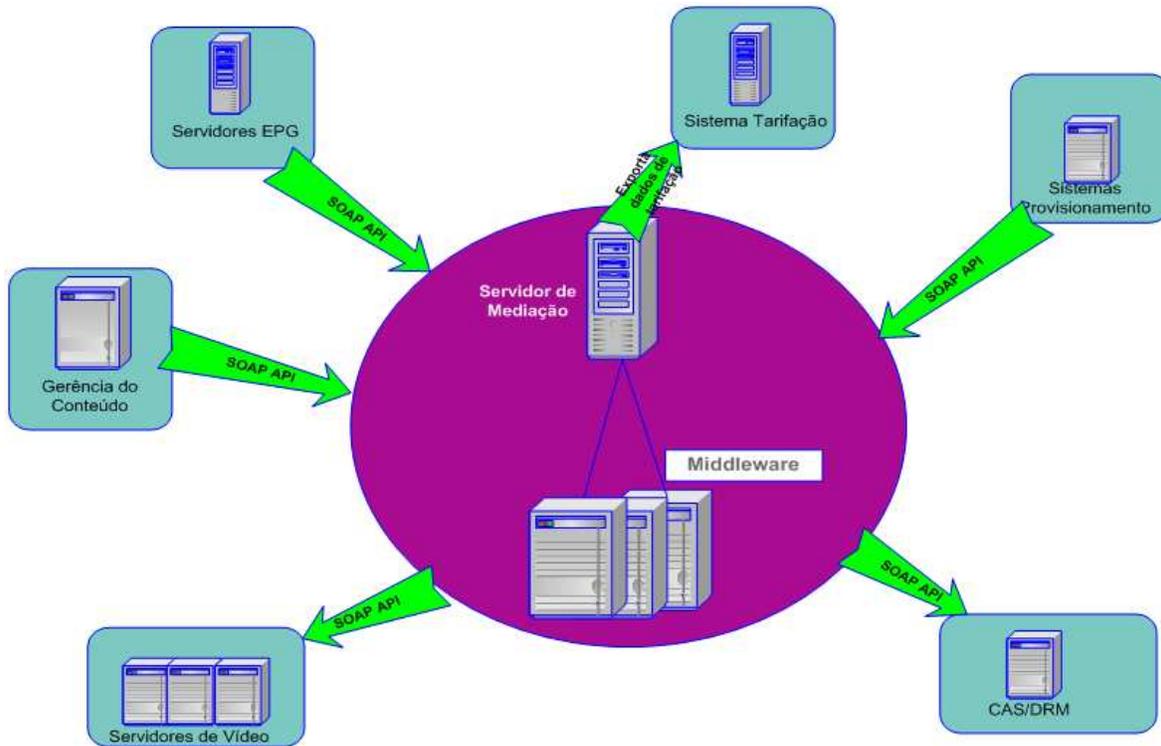


Figura 3.15 - Integração entre sistemas.

Os provedores de serviços já têm um sistema de provisionamento para configurar os mais diversos sistemas e serviços. Dessa forma, esse sistema já existente e conhecido dos operadores pode ser integrado para ser usado também para configurar as novas tarefas de IPTV. Sendo assim, para se criar um usuário de IPTV, por exemplo, o operador pode usar o próprio sistema já conhecido por ele. Essa integração, usando uma API SOAP, é bem simples. Nesse caso, o sistema de provisionamento já existente envia um arquivo XML de configuração que será recebido na máquina correspondente do *middleware*. Essa integração é bastante útil para o provedor de serviço, uma vez que seus operadores já estão acostumados com a interface e com as telas do seu sistema de provisionamento.

Uma integração tão ou mais importante é a integração do *middleware* com um sistema de tarifação. Todas as compras e consumos, *on-line* ou mensais, realizados pela plataforma de IPTV serão armazenados nos servidores de base de dados do *middleware*. Esses dados devem ser faturados de forma automática e em tempo real. Dessa forma, essas informações devem ser exportadas de alguma forma para o sistema de tarifação e faturamento da operadora. Mais uma vez, as API SOAP podem ser utilizadas.

O sistema de tarifação é uma combinação de *software* e *hardware* que recebem informações detalhadas de uso dos serviços, agrupam essas informações por cliente ou por conta, cria fatura de cobrança mensal, além de criar relatórios para gerência e registros.

Para que essa tarefa seja integrada de forma simples e rápida, um sistema de mediação pode ser usado.

Um dispositivo de mediação é um *software* ou equipamento que filtra e isola as informações dos equipamentos de rede. O dispositivo de mediação recebe, processa e formata as informações de uso para um formato adequado para processamento pelo sistema de tarifação. Essa informação processada é periodicamente ou continuamente usada pelo sistema de tarifação do provedor de serviço.

O dispositivo de mediação é comumente usado pelo sistema de bilhetagem, uma vez que ele pode receber dados proprietários ou não padronizados dos servidores de IPTV e formatá-los em mensagens em um formato que o sistema de bilhetagem possa compreender.

Outra informação que deve ser inserida no *middleware* é a informação de programação eletrônica dos canais (EPG ou *Electronic Programming Guide*) que contém a lista dos programas, horários, sinopses, atores etc. para todos os canais para todo o mês, por exemplo. Isso poderia ser feito de forma manual pelo operador, o que representaria um árduo trabalho de digitação.

Entretanto, existem empresas especializadas em consolidar a programação em um arquivo XML, por exemplo. Esse arquivo poderia ser facilmente importado para a base de dados. Para isso, o provedor de serviço deve fazer um acordo com o provedor de EPG para que esses dados consolidados possam ser enviados periodicamente.

Servidor de Vídeo

Controlado pelo *middleware* e usando tipicamente API SOAP, o servidor de vídeo sob demanda (servidor VOD), ou servidor de *streaming* de vídeo, é um elemento composto por um conjunto de *hardware* e *software* dedicados com grande capacidade de armazenamento e com alta capacidade de *streaming*. Esse servidor irá armazenar todos os conteúdos que serão usados sob demanda, tais como filmes, músicas e *shows*.

O servidor de VOD tem a capacidade de transmitir (fazer *streaming*) de centenas ou milhares de títulos ao mesmo tempo. Ele transmite um conteúdo armazenado no formato MPEG-2 TS para os clientes usando o protocolo RTSP ou *multicast* [5, 6, 23, 26]. Quando transmitido em RTSP (*unicast*) esse servidor pode fazer a transmissão de um mesmo conteúdo para vários usuários ao mesmo tempo. Nesse caso, como é uma conexão ponto a ponto do servidor para cada usuário, cada usuário terá controle total sobre o seu fluxo. Assim, enquanto um usuário está avançando um filme X, outro pode estar começando e outro terminando o mesmo filme, e assim, se a transmissão é *unicast*, cada usuário pode interagir com o seu próprio conteúdo, avançando, retrocedendo, ou até mesmo parando o

fluxo.

Alguns servidores de vídeo sob demanda têm ainda a capacidade de transmitir usando *multicast*. Nesse caso, como visto no capítulo 2, todos os usuários sintonizados no mesmo endereço *multicast*, irão assistir ao mesmo conteúdo ao mesmo tempo e obviamente não podem interagir com a transmissão. Na Capítulo 3 será mostrado um interessante uso da transmissão de VOD via *multicast*.

Existem várias empresas especializadas em desenvolver servidores de vídeo e, em geral, usam um *hardware* proprietário com uma altíssima capacidade de armazenamento, bem como, uma grande capacidade de *streaming*.

A Figura 3.16 - Servidor de vídeo mostra um servidor de vídeo típico com várias posições para discos, em geral são discos SCSI (*Small Computer System Interface*) de 400 GB, além de várias placas de saída *Ethernet* para *streaming* de conteúdos. Em geral, eles têm ainda uma grande quantidade de memória *cache* que pode conter dezenas de filmes. Essa memória é usada para otimizar os *streams* em progresso.

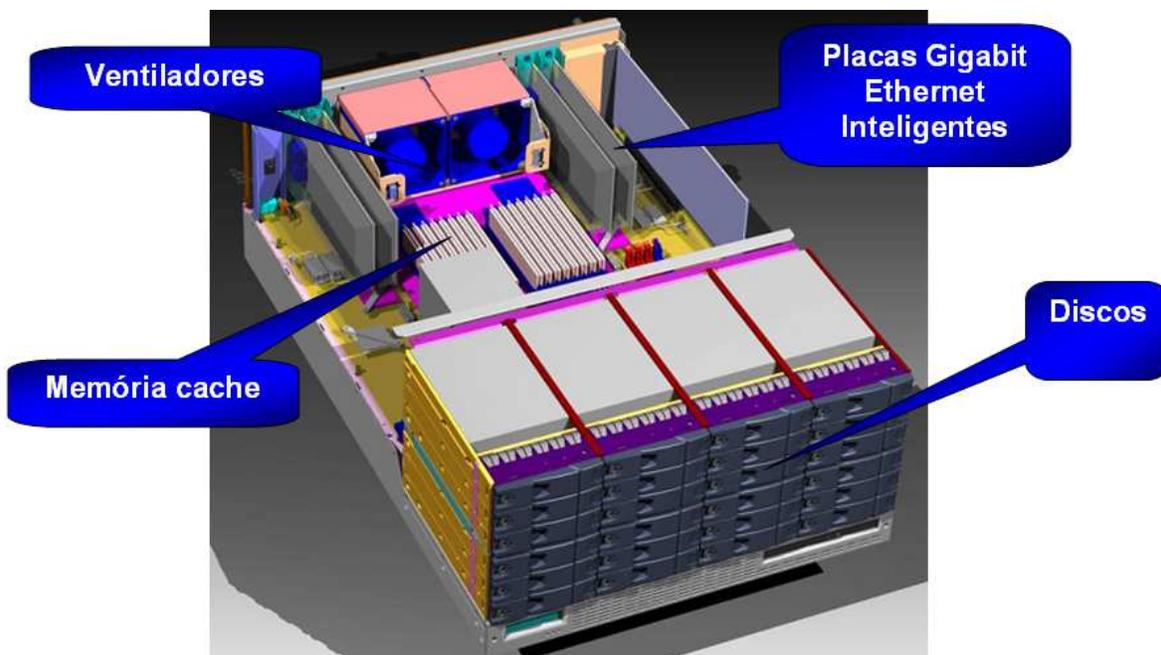


Figura 3.16 - Servidor de vídeo.

Os servidores de vídeo, em geral, têm também mecanismos de proteção de redundância tanto para armazenamento, quanto para *streaming*.

Os discos usam o RAID (*Redundant Array of Independent Disks*). Quando um conteúdo é copiado para o servidor, esse conteúdo é repartido em blocos e cada bloco é copiado em um disco, sendo que um dos discos irá conter um bloco de paridade. O sistema mais usado

nesses casos é o RAID 5, que apresenta excelentes resultados e uma boa tolerância a falhas. O RAID 5 rotaciona a paridade, ou seja, copia o bloco de paridade a cada momento em um disco, como visto na Figura 3.17 - RAID 5.

Dessa forma, um conteúdo não é armazenado em um único disco e sim no servidor por completo. Com isso, pode-se até mesmo perder um disco, que o sistema consegue recuperar os dados desse disco pelo bloco de paridade correspondente.

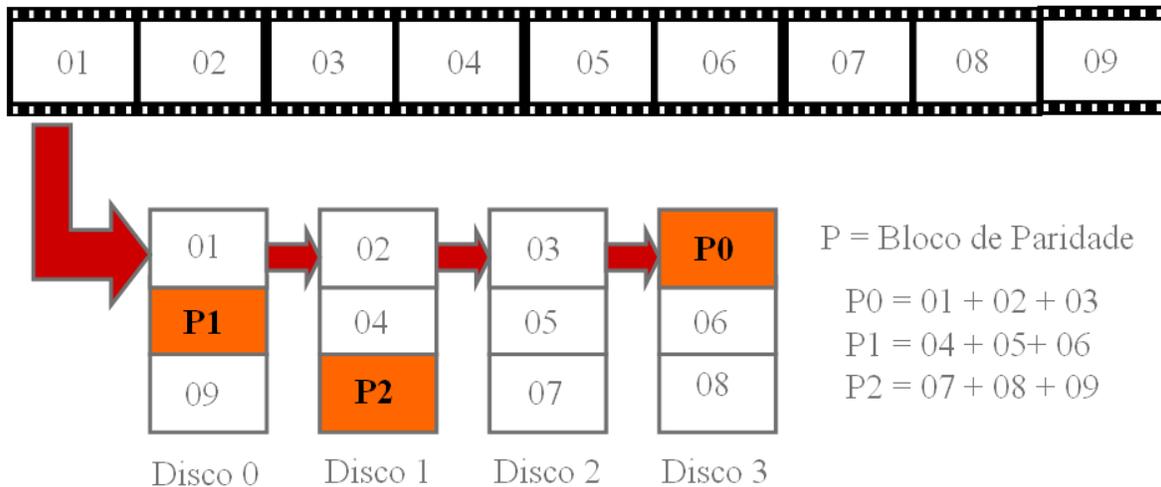


Figura 3.17 - RAID 5.

Os conteúdos VOD devem ser codificados usando o padrão MPEG-2 TS. Isso pode ser feito de diversas formas. Os provedores de conteúdo podem enviar arquivos de filmes já codificados no formato certo, ou isso pode ficar a cargo do provedor de serviço.

Um caso típico é o envio dos conteúdos usando fitas *betacam* (formato profissional desenvolvido pela Sony) e VTR (*Video Tape Recorder*) ou DVD profissional como material fonte. Nesse caso, o provedor de serviço pode codificar o conteúdo com os parâmetros que ele quer, como a escolha da resolução ou largura de banda que seja mais adequada a sua rede e ao conteúdo.

Tanto o VTR quanto o DVD profissional apresentam uma excelente qualidade e inclusive têm saída SDI. A Figura 3.18 - Codificação de um conteúdo para VOD mostra os equipamentos usados na codificação e criação de um arquivo MPEG-2 TS. O VTR é a fonte do codificador. O codificador é o mesmo usado para aplicações de canais *multicast*, só que, nesse caso ficará *off-line* e, será usado, somente quando for necessária a codificação de títulos VOD. A saída do codificador será exatamente igual à saída de um codificador para canais ao vivo, ou seja, MPEG-2 TS encapsulado por IP *Multicast*. A solução necessita ainda de uma estação para gravação e controle dos equipamentos. Essa estação

controla o VTR, configura o codificador com os parâmetros necessários e grava o conteúdo *multicast* em um arquivo com extensão MPG. O operador deve rebobinar a fita para o início e começar a gravar exatamente segundos antes de começar o filme e parar assim que o filme acabe. Apesar da extensão ser MPG, o formato nada mais é que o MPEG-2 TS, com a codificação de vídeo em H.264, por exemplo. Esse arquivo será utilizado para inserção no servidor de vídeo. Como visto, no tópico de MPEG-2 TS, esse conteúdo pode conter legendas e mais de um áudio. Dessa forma, o conteúdo VOD poderá ser visto com todos esses recursos.

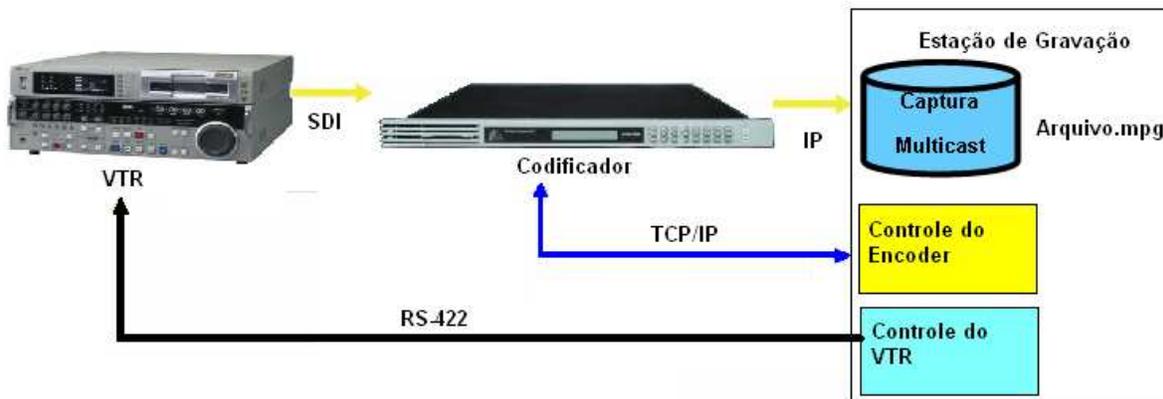


Figura 3.18 - Codificação de um conteúdo para VOD.

Para o processo de ingestão do conteúdo, necessita-se ainda de informações adicionais do filme, incluindo nome, duração, atores, diretores, sinopse, classificação, preço, idiomas e por quantas horas o conteúdo ficará disponível após a compra. Esses dados são conhecidos como metadados e são comumente usados para prover informações técnicas e descritivas sobre os dados. Essas informações serão visualizadas posteriormente pelos usuários na hora da compra do conteúdo. A associação dos metadados com o arquivo digital é chamada de *asset*. Esse *asset* deve ser carregado no servidor de vídeo num processo chamado de ingestão.

Esse processo é mais uma vez feito pelos servidores de *middleware* que, por meio das APIs, fazem as interfaces com servidor de vídeo como visto na Figura 3.15 - Integração entre sistemas.

Os conteúdos devem respeitar uma janela de lançamento que deve ser rigorosamente respeitada. Primeiro, os conteúdos são lançados nos cinemas americanos e nacionais. Algumas vezes em paralelo. Em seguida, eles são distribuídos para as redes de hotéis e linhas aéreas. Após certo tempo, eles são lançados nas locadoras e pouco depois para os sistemas de VOD e PPV. Após isso, vão para a TV paga e finalmente chega à TV aberta

após uns três ou quatro anos.

Para respeitar essa cadeia, o conteúdo deve ser protegido em todo o seu ciclo, desde o recebimento das fitas *betacam*, passando pela ingestão e armazenamento. Esses conteúdos são geralmente cifrados antes do armazenamento, ou seja, os títulos são armazenados cifrados para evitar que os próprios funcionários do provedor de serviço os copiem. Esse tópico será discutido com mais detalhes no capítulo de proteção de conteúdo.

O provedor de conteúdo, usando essa janela de lançamento vai determinar também quando um conteúdo pode ser comercializado e quando deve ser retirado de comercialização, com diversas penalidades previstas caso o provedor de serviço não respeite algumas dessas condições.

Entretanto, o provedor de serviço vai receber o conteúdo com certa antecedência, para que ele tenha tempo de codificá-lo e prepará-lo para o lançamento na data devida.

Uma vez que o conteúdo foi carregado com sucesso e a data de lançamento foi atingida, o conteúdo estará disponível para a compra pelos usuários. Isso também será feito de forma automática. Os metadados irão conter, além dos dados comentados acima, a data e a hora de lançamento. Esses valores chegarão à base de dados e só serão mostrados pelos servidores de aplicação para os usuários, quando essa data for alcançada. Quando isso ocorrer, o usuário vai poder navegar pelo título, verificando informações adicionais e comprá-lo *on-line*. Tudo isso utilizando HTTP entre o STB e o AS. Quando o usuário realiza a compra, ele deve colocar sua senha de compra que será verificada na base de dados. Uma vez que a senha está correta, um novo consumo é gerado para aquele usuário para posterior cobrança. O AS retorna então uma URL (*Uniform Resource Locator*) que conterá o endereço IP do servidor de vídeo mais próximo do usuário e o conteúdo comprado. Um exemplo de uma URL típica para esse caso seria:

```
rtsp://200.10.10.9:554/SeaChange?NodeGroupId=1&ProviderId=hbo&ProviderAssetId=batman
```

O STB pode também receber juntamente com essa URL, uma chave para decifrar um conteúdo, caso o mesmo tenha sido armazenado cifrado no servidor de vídeo, o que vai ocorrer comumente para conteúdos VOD,

Quando o usuário apertar o botão de *Play*, na verdade o que será feito é uma solicitação RSTP do STB diretamente ao servidor de vídeo. O servidor fará então a transmissão daquele conteúdo cifrado para o assinante que fez a solicitação. Esse STB, de posse da chave de criptografia, irá remover a criptografia, decodificar o conteúdo e transformá-lo para o formato de TV.

Existe outra tarefa que é executada durante o processo de ingestão. Quando o servidor de vídeo recebe um novo conteúdo, cifrado ou não, ele irá criar alguns outros arquivos, que

são chamados de *trick files*. Esses arquivos serão usados para fazer o avanço e o retrocesso e são baseados nos quadros Is do MPEG. Por conter apenas quadros Is eles ocupam menor espaço em disco.

Dessa forma, quando um usuário está assistindo um conteúdo e aperta a tecla de avanço, o servidor de vídeo, usando um arquivo de índice vai determinar o ponto que ele está assistindo e o ponto correspondente no arquivo de avanço e começar o *streaming* para ele daquele ponto. Daí vem o nome de *trick file*, porque o servidor de vídeo faz um truque alterando o arquivo. Um arquivo é criado para cada sentido de avanço e retrocesso. Se o provedor de serviços quiser oferecer múltiplas velocidades, outros arquivos deverão ser criados em cada velocidade e ao mesmo tempo excluindo quadros Is intermediários. Isso é uma definição feita pelo provedor de serviço e valerá para todos os títulos. Os arquivos de avanço e retrocesso serão baseados nos quadros Is e, dependendo da velocidade, poderá ainda ser a cada 2, 3 ou mais quadros Is. Quando o usuário está assistindo um conteúdo no modo avançado, ele, na verdade, está vendo outro conteúdo. Finalmente quando o usuário aperta a tecla de *play* de novo, um novo cálculo de posição é feito para determinar o ponto atual no arquivo original correspondente ao ponto no arquivo avançado.

A Figura 3.19 - Comandos RTSP mostra o recebimento da URL via HTTP e os comandos RTSP para estabelecimento e controle da sessão. Essa sessão continua estabelecida até que o usuário aperte o botão de *stop*.

O servidor de VOD é instalado próximo ao usuário na rede, assim, quando um usuário pedir um conteúdo, o *stream* não precisa trafegar por toda a rede. Isso se deve ao fato de que esse *stream unicast* em escala poderia impactar toda a rede. Dessa forma, servidores de vídeo podem ser instalados por região geográfica para atender um grupo de usuários.

Os servidores de vídeo podem ter uma topologia centralizada ou distribuída. A topologia centralizada irá armazenar todos os títulos numa única localidade e requer uma grande capacidade de armazenamento e uma alta escalabilidade. Entretanto, sua gerência é mais simples.

Por outro lado, uma topologia distribuída pode ter uma menor capacidade de armazenamento, uma vez que existirão vários servidores de vídeo geograficamente distribuídos. Além disso, como os servidores são distribuídos, o tráfego *unicast* do servidor mais próximo até o STB que o solicitou ocupará menor trecho da rede.

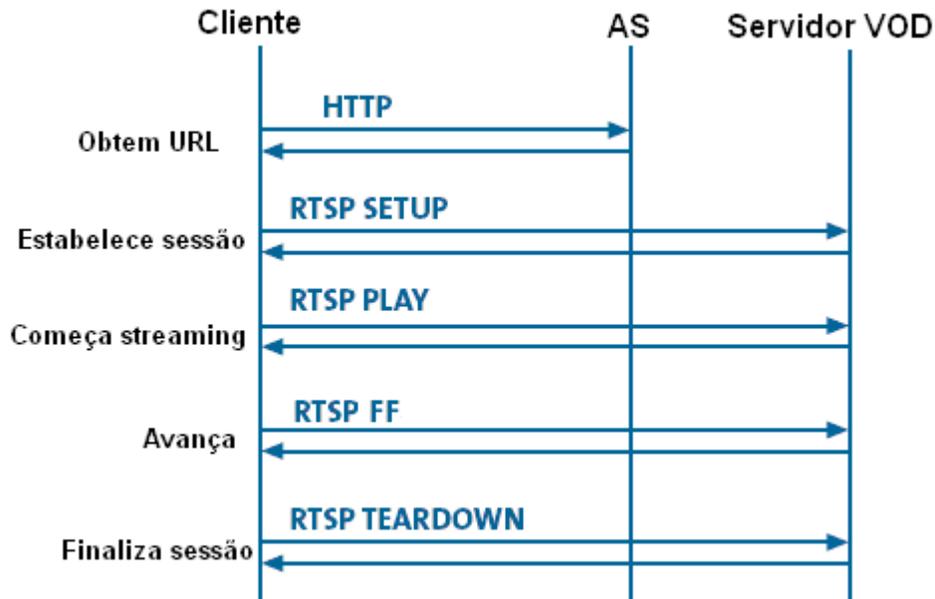


Figura 3.19 - Comandos RTSP.

Proteção do Conteúdo e Segurança de Rede

Os provedores de conteúdo consideram a proteção do conteúdo como a principal preocupação, a questão mais importante e mandatória em IPTV. Essa função é tratada pelo *middleware* e tem como característica a habilidade de proteger o negócio impedindo cópias ilegais pela gerência dos direitos digitais (DRM ou *Digital Rights Management*) e assegurar o pagamento pelo uso através do sistema de acesso condicional (CAS ou *Conditional Access System*).

O CAS impede e limita o direito de visualização somente para os clientes que estão pagando pelo serviço.

Acesso Condicional - CAS

Uma das etapas do processo do acesso condicional é a autenticação do STB na rede. Autenticação é o processo de troca de informação entre o STB e a rede de comunicação que permite ao operador da rede confirmar a verdadeira identidade do usuário ou dispositivo.

Essa validação da autenticidade do STB permite que o provedor de serviço possa negar serviços para os usuários que não foram identificados, inibindo assim o uso fraudulento do STB.

A autenticação pode ser feita baseada no endereço MAC do STB, no seu endereço IP, no uso de um par usuário/senha, por meio de porta de terminal de acesso, ou certificados ou

mesmo uma combinação de alguns desses valores.

Se o STB não está autenticado, ele não conseguirá acessar nenhum serviço da rede, uma vez que o STB faz interface diretamente com o *middleware*.

O CAS controla também o sinal de TV, limitando o acesso dos canais de TV somente para os usuários autorizados, ou seja, somente para os usuários que estão pagando pelo serviço.

Existem várias formas de proteger o conteúdo *broadcast*. Ela pode ser feita usando a criptografia dos canais *broadcast* em tempo real, ou fazendo-se filtros de endereços *multicast* na conexão do usuário. Nesse último caso, somente o equipamento de acesso, próximo ao usuário é que irá filtrar determinado canal. É adicionado na configuração do equipamento de acesso um filtro IGMP para todos os canais que ele não pode assistir.

A criptografia é um processo de proteção da informação para que esta não seja usada ou não seja interpretada por usuários não autorizados.

Um sistema de criptografia tipicamente usa uma combinação de chave(s) e um algoritmo de criptografia para modificar os dados. Uma chave de criptografia é um código único que é usado para cifrar dados para protegê-los de usuários não autorizados. Uma chave de criptografia é geralmente mantida privada ou secreta dos outros usuários.

Sistemas de criptografia podem usar a mesma chave para cifrar e decifrar o conteúdo (criptografia simétrica) ou o sistema pode usar diferentes chaves para cifrar e decifrar (criptografia assimétrica).

A informação que não é cifrada é chamada de texto limpo (*clear text*) e a informação cifrada é chamada de texto cifrado (*cypher text*), conforme mostra a Figura 3.20 – Criptografia.

O comprimento da chave de criptografia é a quantidade de dígitos ou bits que são usados para cifrar o processo. Geralmente, quanto maior o comprimento da chave, mais forte é a proteção.

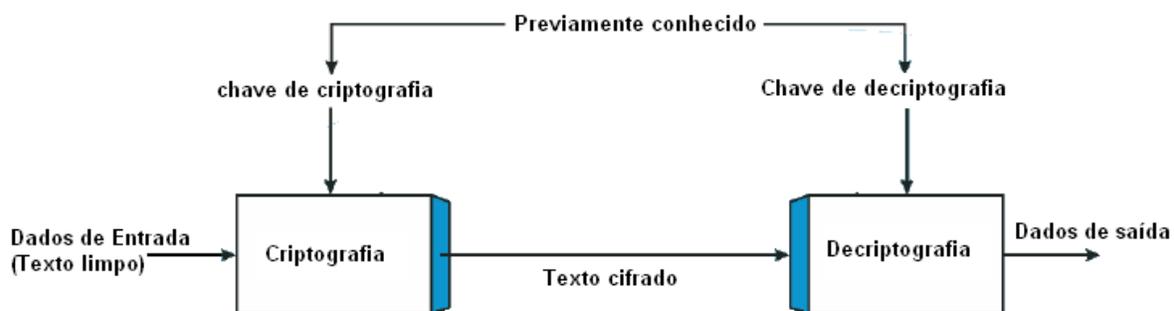


Figura 3.20 – Criptografia.

Existem vários algoritmos de criptografia. O mais novo e também considerado o mais seguro, usado inclusive pelo governo norte americano é o AES (*Advanced Encryption Standard* ou padrão de criptografia avançada). O AES é baseado no algoritmo de criptografia de Rijndael e usa blocos de 128 bits de tamanho e pode usar chaves de 128, 192 ou 256 bits de comprimento [38, 39, 40, 41, 42]. O padrão AES deve supostamente substituir o padrão DES (*Data Encryption Standard* ou padrão de criptografia de dados).

Modernos sistemas de criptografia para IPTV usam o algoritmo AES de 128 bits e, quando usada na criptografia em tempo real, pode alterar (rodar) a chave de criptografia a cada poucos segundos ou minutos. Assim, mesmo que um hacker consiga descobrir uma chave, o que é praticamente impossível, dado que a chave só vale poucos minutos, essa chave valeria apenas para aquele canal e para aquele período de poucos minutos.

Esses sistemas são extremamente eficientes e capturam automaticamente todos os MPEG-2 TS de todos os codificadores e os cifram em tempo real. Cada canal é cifrado com uma chave diferente. Além disso, os servidores de CAS fazem interface com o *middleware*. Sendo assim, quando o *middleware* associa um pacote de TV a um usuário, o *middleware* irá comunicar isso ao servidor de CAS para que esse último envie as chaves de decriptografia de todos os canais daquele pacote diretamente para aquele assinante. A Figura 3.21 - CAS/DRM e protocolos usados pelo STB ilustra bem esse processo.

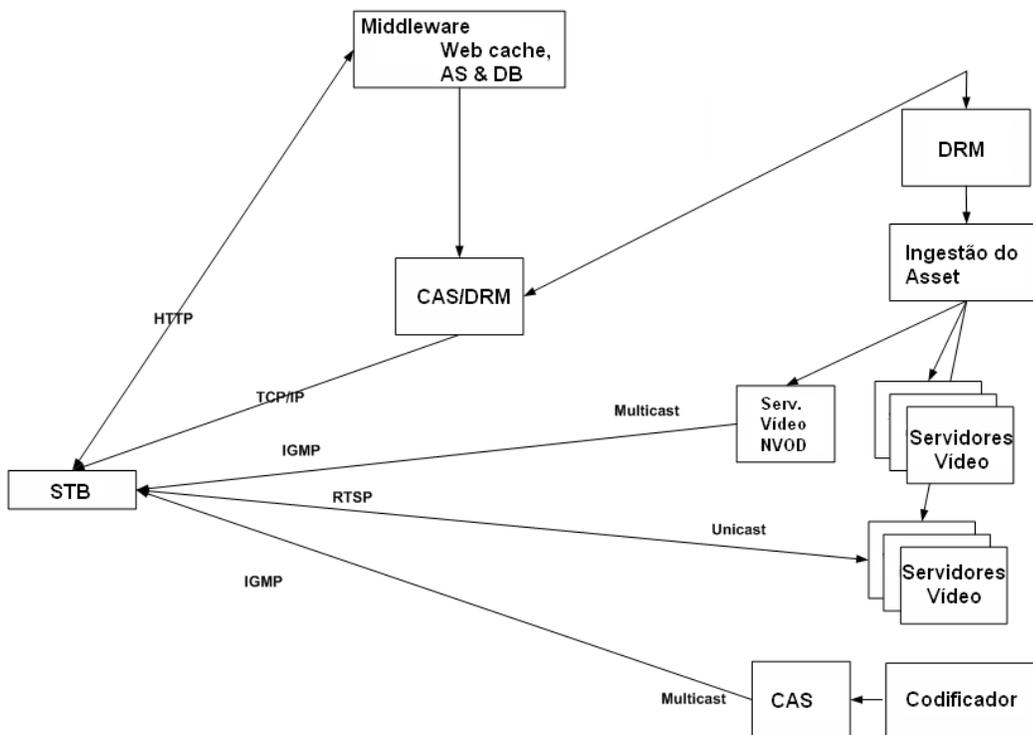


Figura 3.21 - CAS/DRM e protocolos usados pelo STB.

Todos os canais *multicast* que saem dos codificadores serão capturados automaticamente pelo servidor CAS. Isso se deve ao fato do servidor CAS ter uma interface no modo promíscuo. Esse servidor vai então cifrar todo MPEG-2 TS e inserir de novo o encapsulamento com o mesmo IP *multicast*. Esses servidores são bastante rápidos e geram um atraso de apenas poucos mili-segundos nesse processo.

Foi mostrado que o STB deve ser autenticado e, além disso, ele terá um pacote de TV associado. Sendo assim, caso um usuário tente apertar um número do controle remoto que não faz parte da sua programação, o *middleware* vai apresentar uma mensagem avisando-o de que aquele canal não está habilitado.

Entretanto, uma pessoa com um pouco mais de conhecimento, pode tentar substituir o STB por um PC e utilizar o mesmo endereço IP do STB. Muito provavelmente esse PC não conseguirá autenticar no *middleware*, mas ainda assim ele terá a conectividade IP com as outras interfaces mostradas na Figura 3.21 - CAS/DRM e protocolos usados pelo STB. Então ele poderia utilizar um tocador de *multicast*/MPEG-2 TS e executar o comando “igmp://224.10.10.5:1008” para um endereço *multicast* conhecido. Nesse caso, então ele estará transpondo a proteção do *middleware*. Entretanto, como mostrado, o canal está cifrado e ele não terá a chave para decifrar e receberá uma imagem completamente ilegível.

A proteção apresentada acima é extremamente segura e eficiente. Cabe ressaltar que ela pode ser muito cara, uma vez que em geral, as empresas especializadas nesse ramo, cobram um adicional pelo número de assinantes do provedor de serviço.

Outro tipo de solução que poderia ser utilizada é o uso de filtro IGMP na porta do equipamento de acesso. Por exemplo, suponha que um usuário tenha acesso a um pacote básico. O equipamento de acesso desse usuário terá uma configuração de filtro IGMP para todos os canais que não pertencem ao pacote básico. Sendo assim, se um *hacker* tentar executar o comando IGMP, por exemplo, o “igmp://224.10.10.5:1008”, ele não receberá nada.

Essa solução é suficientemente barata, mas protege somente após o equipamento de acesso. Assim, caso uma pessoa tenha acesso ao núcleo da rede, como por exemplo, um funcionário da própria operadora, ela conseguirá assistir a todos os canais sem criptografia.

Gerência dos Direitos Digitais - DRM

Agora que um vídeo é apenas outra forma de transmissão de dados digital, os usuários podem gravar seus programas e assisti-los onde e quando quiserem, além de poderem avançar os comerciais. Isso pode ameaçar a indústria de entretenimento de várias formas, primeiro porque os usuários podem gravar o conteúdo e compartilhar com parentes e amigos, aumentando a pirataria. Caso um usuário consiga copiar esse conteúdo sem

criptografia, ele pode inclusive editar o arquivo e excluir os comerciais. O prejuízo é real porque os desenvolvedores de comerciais pagam uma fortuna para os provedores de conteúdo a fim de anunciar os seus mais diversos produtos. Dessa forma, várias empresas, como Microsoft, Disney, BBC, entre outras, estão buscando novas formas de desenvolver sofisticadas técnicas de novos sistemas de DRM.

Cuidados e proteções ainda maiores são necessários para os conteúdos VOD, para evitar que se quebre a janela de tempo de lançamento. Dessa forma, uma exigência dos provedores de conteúdo é o uso de criptografia para VOD e eles só irão comercializar seus conteúdos de VOD para os provedores de serviços caso essa exigência seja cumprida. Além disso, o sistema de DRM usado deve passar por uma série de auditorias feitas pelos grandes estúdios de cinema para provar que é realmente segura e forte.

Os sistemas DRM podem usar uma combinação de marca d'água digital, certificados digitais, assinaturas digitais, sistema de acesso condicional, códigos de ativação, autenticação e criptografia para prover a segurança necessária para o conteúdo e sua distribuição. Além disso, têm-se usado uma combinação de processos, programas e *hardware* para assegurar essas tarefas.

Os conteúdos VOD, principalmente, devem ser protegidos em todas as fases. Por exemplo, se o provedor de serviços recebe seus conteúdos por meio de fitas *betacam* essa proteção tem que começar na forma do envio das fitas. Tem que utilizar um transporte seguro e somente pessoal autorizado pode receber essas fitas. Após isso, o conteúdo tem que ser digitalizado e, isso deve ser feito novamente em uma sala protegida, por uma equipe selecionada e especializada. Esse ambiente em geral deve ter câmeras de segurança e não pode entrar com nenhum tipo de mídia ou dispositivo de armazenamento, como *laptops*, discos externos, CD-ROM etc. Após a codificação do conteúdo, as fitas devem ser armazenadas em local protegido, o arquivo digital deve ser cifrado pelo sistema de DRM e o arquivo limpo deve ser removido.

Novamente aqui a criptografia que deve ser usada é o AES e com o tamanho de chave típico de 128 bits. A chave usada para cifrar cada filme é diferente e é armazenada na base de dados de forma cifrada, também.

Esse conteúdo cifrado é inserido no servidor de vídeo que, mesmo com o conteúdo cifrado, consegue gerar os arquivos de *trick files* e o arquivo de índice.

Quando o usuário compra um determinado filme, o AS consulta a chave na base de dados e a envia para o usuário. Esse envio é feito de forma cifrada e usando uma chave de sessão conhecida apenas pelo STB e pelo AS.

Quando o STB solicita o conteúdo ao servidor de vídeo, esse último faz o *streaming* do conteúdo cifrado, mas como o STB já tem a chave, ele conseguirá decifrá-lo. O *middleware* deve ainda garantir outra proteção, ou seja, impedir que o usuário final grave o conteúdo.

Isso é feito de duas formas. A primeira feita pelo *middleware* impede que o usuário grave o conteúdo, mesmo que o STB tenha um disco interno. Durante o recebimento do fluxo, o usuário não tem nenhuma opção para gravar.

Finalmente, deve-se ainda ter uma proteção na saída analógica dos STBs, para evitar que os usuários usem vídeo cassete ou gravadores de DVD. Essa proteção inclui o *Macrovision*.

Macrovision é uma empresa que criou um esquema de prevenção contra cópia, estabelecida em 1983. O nome é também usado para fazer referência ao sistema de proteção contra cópias.

A maneira como a proteção funciona é bem interessante. O sinal de vídeo tem um tipo especial de ruído que os aparelhos de TV não notam, mas que os gravadores não podem tratar. Esse sinal confunde um componente conhecido como circuito AGC (*Automatic Gain Control* ou Controle de Ganho Automático) nos gravadores e grava o sinal incorretamente.

O AGC normalmente é usado para amplificar os sinais fracos e atenuar os sinais fortes. O *Macrovision* insere alguns sinais não padronizados na área não visível da TV. Essa informação de vídeo adicional faz com o que o AGC pense que é uma imagem muito clara ou muito escura e tenta ajustar a saída para o que ele acha que deveria ser, fazendo com o que a imagem fique muito clara ou muito escura.

Esse truque é repetido várias vezes e de forma aleatória. Assim, o AGC provoca o aparecimento de fortes contrastes nas imagens, tornando-as muito escuras ou muito claras, várias vezes e de forma aleatória. Isso torna o conteúdo irritante de se ver. O áudio permanece inalterado, porque o circuito que grava o áudio é separado do vídeo.

Entretanto, esse efeito só aparece quando o sinal é redirecionado para um gravador de DVD ou VCR, uma vez que a maioria dos aparelhos de TV não tem AGC.

Arquitetura de Rede

Na Figura 3.22 - Situação da rede anterior e situação da rede atual observam-se a situação da rede anterior, onde existem três redes separadas, a rede de telefonia pública comutada (RTPC), a *Internet* via rede IP e a rede de TV a cabo ou satélite. Na nova situação, o operador faz uso de uma única rede IP. Entretanto, o sistema de IPTV requer uma maior largura de banda, principalmente para o novo componente de vídeo, qualidade de serviço (QoS) diferenciado e priorizando cada um dos serviços, suporte a *multicast* e *unicast*, e uma securização de toda a rede. Tudo isso demanda uma transformação total da rede.

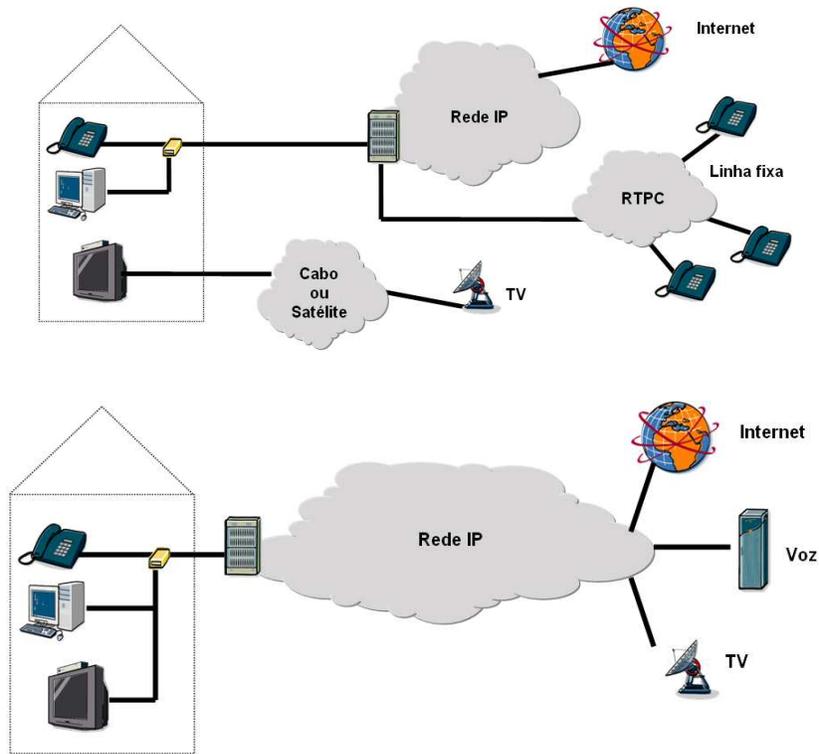


Figura 3.22 - Situação da rede anterior e situação da rede atual.

A arquitetura de rede é composta da rede residencial, rede de acesso, rede de agregação, rede de transporte (*core*) e rede de conteúdos como mostrado na Figura 3.23 - Arquitetura de Rede.

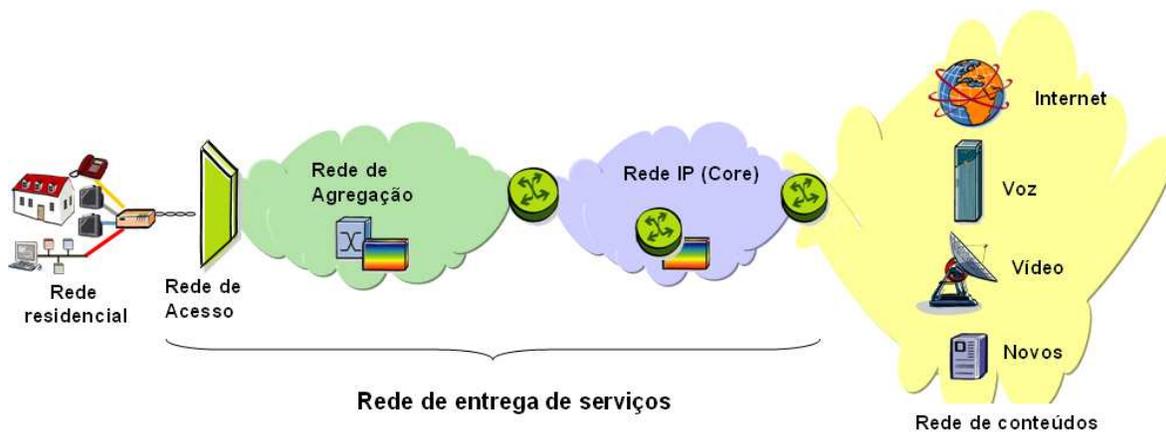


Figura 3.23 - Arquitetura de Rede.

Rede residencial

A rede residencial é a posição onde é feita a terminação da conexão de banda larga e é composta de vários dispositivos como PC, telefone IP ou telefone tradicional, TV com *Set Top Box* (STB), todos eles compartilhando essa mesma conexão, como mostrado na Figura 3.24 - Rede residencial [16, 17, 20, 21].

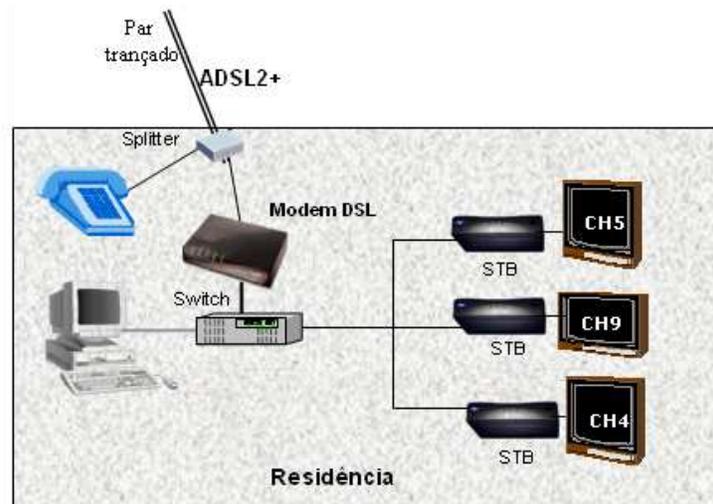


Figura 3.24 - Rede residencial.

Modem

O modem provê conexão de banda larga ao provedor de serviços. Como o próprio nome diz, a funcionalidade do modem é a modulação e demodulação da interface física (ex. linha ADSL - *Asymmetric Digital Subscriber Line*). É muito comum o uso de *gateways* residenciais (RG ou *Residential Gateways*) integrados ao *modem* ADSL, equipamento esse que oferece uma maior inteligência para os serviços, encaminhando os pacotes para a interface IP correspondente e tratando os pacotes com diferentes políticas de qualidade de serviço.

No caso do uso da tecnologia ADSL2+, é necessário um *modem* ADSL2+ que termina a conexão ADSL e a apresenta a camada *ethernet*, ou seja, o *modem* tem que ser compatível com o equipamento de acesso. A tecnologia DSL será explicada no tópico de rede de acesso.

Quando não se faz uso dos *gateways* inteligentes (RG), pode-se ainda trabalhar com *modem* ADSL com um *switch IP ethernet* integrado. Uma das portas *ethernet* é conectada ao computador para acesso de banda larga, outra porta é conectada ao receptor de IPTV,

STB (*Set Top Box*). Sobram ainda outras portas, que podem ser conectadas a um telefone de VoIP (Voz sobre IP) ou a outros STBs. A Figura 3.24 - Rede residencial mostra essas conexões, com a diferença de que a figura mostra um telefone tradicional e um *splitter* (filtro) para separar as frequências, uma vez que o DSL trabalha nas altas frequências e o telefone nas baixas frequências.

Set Top Box

O *Set Top Box* (STB) é um dispositivo final requerido para transformar o serviço de vídeo para TV e se conecta ao *modem* através de uma conexão *ethernet*. Além disso, é necessário um STB para cada aparelho de TV.

A função básica do STB é decodificar e decifrar os *streams* de vídeo e de áudio e transformar esse conteúdo para o formato requerido pela TV.

Além disso, os STBs habilitam a navegação do cliente aos diferentes serviços e telas por meio de uma conexão com o servidor de aplicação do *middleware*, tipicamente utilizando HTML. Esse STB pode ainda acessar serviços *multicast* usando o protocolo IGMP, ou conteúdos sob demanda através do protocolo RTSP. Para o caso desse último, que é um conteúdo *unicast*, o RTSP precisa ser iniciado pelo próprio STB.

Na Figura 3.21 - CAS/DRM e protocolos usados pelo STB são mostrados os diversos protocolos que devem ser suportados pelo STB. Além disso, o STB deve ter o cliente para o CAS/DRM e o cliente para o *middleware*.

Rede de Acesso

A rede de acesso, também conhecida como "última milha", é a parte da rede que conecta os usuários finais ao provedor de serviço (ver a Figura 3.23 - Arquitetura de Rede).

Ela se refere ao conjunto de equipamentos, cabos, fios que estão entre os consumidores finais e o provedor de serviço.

As redes de acesso podem também ser de vários tipos, como DSL (*Digital Subscriber Line* ou linha de assinante digital), FTTH (*Fiber to the Home* ou fibra até a casa), ou até mesmo uma rede sem fio como a emergente tecnologia do WiMAX (*Worldwide Interoperability for Microwave Access*).

DSL

A rede de acesso mais utilizada no momento para IPTV é a DSL, devido à grande quantidade de linhas já instaladas e, nesse caso, multiplexa linhas DSL de várias residências como pode ser visto na Figura 3.25 - Rede de acesso DSL. O equipamento da

rede de acesso é conhecido como DSLAM (*Digital Subscriber Line Access Multiplexer* ou multiplexador de acesso de linhas de assinante digital), que tem a função de agregar múltiplas linhas DSL e conectar para a rede de agregação num sentido e no outro sentido (sentido do usuário), ele combina múltiplos serviços em um único par trançado.

O link que vai do DSLAM em direção a rede de agregação pode ser ATM (*Asynchronous Transfer Mode* ou modo de transferência assíncrona) ou *ethernet*. Nota-se uma grande evolução para *ethernet*, uma vez que ela é mais barata. A Figura 3.25 - Rede de acesso DSL mostra uma conexão DSL saindo da casa do usuário até chegar ao DSLAM.

O DSLAM também vai ter funções de camada 3, funcionando também como um roteador.

As taxas alcançadas por conexões DSL são diretamente afetadas pela quantidade de ruídos e atenuação da linha, o que faz com que exista uma variação de velocidade grande entre uma linha e outra. O caminho do fio telefônico entre a casa do usuário até a central telefônica é chamado de segmento local, sendo seu comprimento crítico para as conexões DSL. Quanto maior o seu comprimento, menor será a taxa alcançada pela conexão.

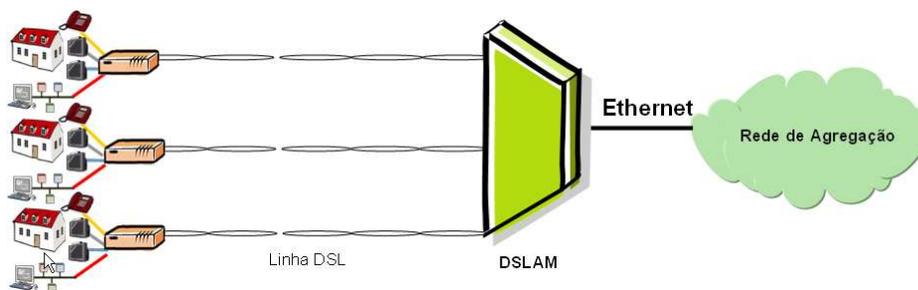


Figura 3.25 - Rede de acesso DSL.

Existem várias tecnologias DSL com diferentes aplicações e diferentes larguras de banda para o usuário, como ADSL (*Asymmetric Digital Subscriber Line* ou linha de assinante digital assimétrica), ADSL2 (*Asymmetric Digital Subscriber Line second generation* ou linha de assinante digital assimétrica de segunda geração), ADSL2+, VDSL (*Very high speed Digital Subscriber Line* ou linha de assinante digital de altíssima velocidade) [2, 16, 20, 21, 50, 51, 52].

A tecnologia ADSL (padrão ITU G.992.1) foi desenvolvida usando-se técnicas avançadas de processamento de sinal por meio de *chips* especiais denominados DSP, além de enormes avanços nas áreas de transformadores de linha, filtros analógicos e conversores A/D. Técnicas avançadas foram aplicadas para que se possam tratar os altos níveis de atenuação e ruído, a fim de se manter a separação entre canais e a recuperação dos padrões

de modulação.

A assimetria do ADSL se deve ao fato de que a velocidade no sentido DSLAM ao usuário (*downstream*) é diferente da velocidade no sentido usuário ao DSLAM (*upstream*). As taxas teóricas (nas condições ideais e com distâncias muito curtas) desse padrão são de 8,0 Mbps para *downstream* e 1,0 Mbps para *upstream*. Entretanto, na prática esse valor é bem menor, visto que os usuários estão a poucos quilômetros de distância do DSLAM.

O ADSL2, nova geração da tecnologia ADSL, padronizada pelo ITU-T G.992.3 tem uma melhor eficiência de modulação e menor cabeçalho (*overhead*). Devido a isso, aumenta em 50% a taxa de *downstream*, chegando a 12 Mbps e mantendo a taxa de *upstream* em 1 Mbps. O aumento da velocidade não é a única vantagem do ADSL2, que ainda traz benefícios da taxa adaptativa e a capacidade de monitoração contínua das condições de ruído, desempenho e qualidade da linha utilizada. A taxa adaptativa permite que a velocidade de uma conexão ADSL2 seja mudada, sem que haja interrupção de serviço ou qualquer tipo de perda para o usuário.

Em 2003 foi lançado pelo ITU-T o novo padrão G.992.5, também chamado de ADSL2+, que dobra a largura de banda usada pelos seus antecessores ADSL e ADSL2. O novo padrão utiliza uma largura de banda de 2,2 MHz e permite taxas de até 26 Mbps a 0 metro, que é reduzida para 20 Mbps em linhas de até 1,5 km. No sentido *upstream*, as taxas chegam até a 1 Mbps.

Além das altas velocidades e da taxa adaptativa, o ADSL2+ também permite a redução do consumo de energia durante períodos de baixa ou não-utilização, tendo a capacidade de adequar seu consumo elétrico sob demanda e conforme a necessidade do usuário.

Para os casos mais típicos, o ADSL2+ suporta uma largura de banda de uns 10 a 12 Mbps para assinantes que estão a 3 km do DSLAM. Tudo isso, é claro, de acordo com as condições do par metálico. Essa banda é suficiente para suportar 2 ou 3 pontos de STB na casa do cliente.

Na Figura 3.26 - Taxa *downstream* x Comprimento do *loop* pode ser visto uma comparação da taxa para as tecnologias ADSL, ADSL2, ADSL2+ e FTTH (*Fiber To The Home* ou Fibra até a residência do usuário) com o comprimento do cabo telefônico do *modem* ao DSLAM. Como se pode perceber consegue-se maior largura de banda para os usuários que estão mais próximos do DSLAM. Exceção apenas para FTTH que oferece uma alta velocidade sem limitações de distância.

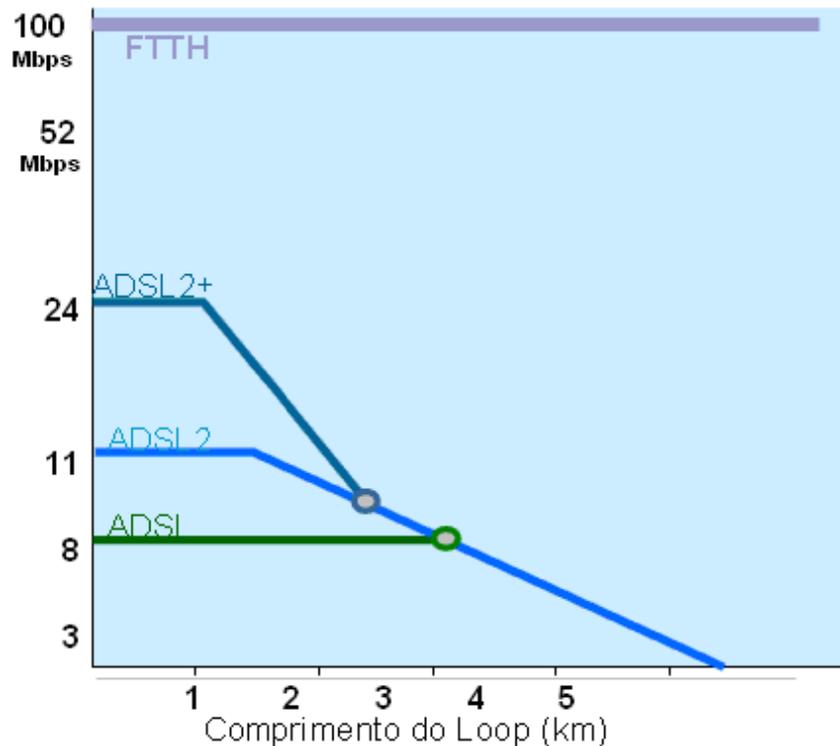


Figura 3.26 - Taxa *downstream* x Comprimento do loop.

FTTH

O FTTH usa fibra óptica entre a casa do assinante até o nó de acesso [3, 22]. Sua implementação é usualmente mais cara do que a do DSL. Entretanto, uma vez instalado, não são necessários *upgrades* adicionais. Como os avanços são feitos na velocidade sobre a fibra, somente os equipamentos de cada ponta da fibra precisam ser substituídos e não a fibra. Algumas redes FTTH usam PON (*Passive Optical Network* ou rede óptica passiva) e, dessa forma nenhuma alimentação é necessária na rede. A rede de fibra requer menos manutenção. Esses fatores reduzem o custo operacional de uma rede FTTH comparado ao custo operacional de uma rede de cobre.

O custo de instalação de um sistema FTTH é variável, dependendo dos custos de construção envolvidos no lançamento da fibra. Essa é a razão primária do porquê o FTTH é mais caro do que usar os cabos de cobre existentes.

A grande vantagem da rede FTTH é a largura de banda. As tecnologias disponíveis hoje conseguem oferecer 100 Mbps para a residência do usuário, o que seria suficiente para múltiplos canais HDTV.

O modelo de negócio do IPTV está intimamente ligado ao acesso e a tecnologia de compressão usada para se implementar o serviço. Por isso, o FTTH continua ainda sendo

uma promessa na medida em que os custos estejam sendo reduzidos.

WiMAX

WiMAX é o nome dado ao padrão IEEE 802.16 que é uma tecnologia de rede de acesso banda larga, sem fio, com altas taxas de transferência (1 Mbps até 70 Mbps) [53]. Ela é, atualmente, praticamente a única tecnologia capaz de oferecer altas velocidades de banda larga sobre uma rede sem fio e funciona na frequência de 2-11 GHz ou de 10-66 GHz.

Atualmente, existem duas aplicações principais para o WiMAX: aplicação WiMAX fixa que é uma tecnologia de acesso de banda larga para residências e empresas ou aplicação para celular.

Essa é uma tecnologia emergente e pode vir a ser usada para cobrir áreas rurais, áreas mais distantes aonde o DSL não chega, ou mesmo áreas móveis como por exemplo, serviços de transportes de trens.

Rede de Agregação

A rede de agregação, como o próprio nome diz, agrega vários nós de acesso (DSLAM) e os conecta ao núcleo da rede. Como comentado na sessão anterior, percebe-se uma evolução muito grande do *uplink ethernet* vindo do DSLAM, por ser mais barato. Pela mesma razão, na rede de agregação, emprega-se também o uso de *switches ethernet*. Outra vantagem da *ethernet* é que ela é uma tecnologia mais fácil de implementar do que a ATM.

A rede de agregação é opcional e se poderiam conectar os DSLAMs diretamente à rede IP, entretanto é mais econômico conectar um grupo de DSLAMs aos *switches ethernet* e os *switches ethernet* à rede IP [12, 18, 21, 26].

Além disso, os *switches* permitem a construção de topologias em anel que provê redundância na conexão entre o nó de acesso ao roteador IP.

Os *switches* podem assumir algumas funções dos roteadores, permitindo que os roteadores consigam tratar um número maior de usuários.

A Figura 3.27 - Rede de agregação mostra a localização dos DSLAMs e dos *switches ethernet*.

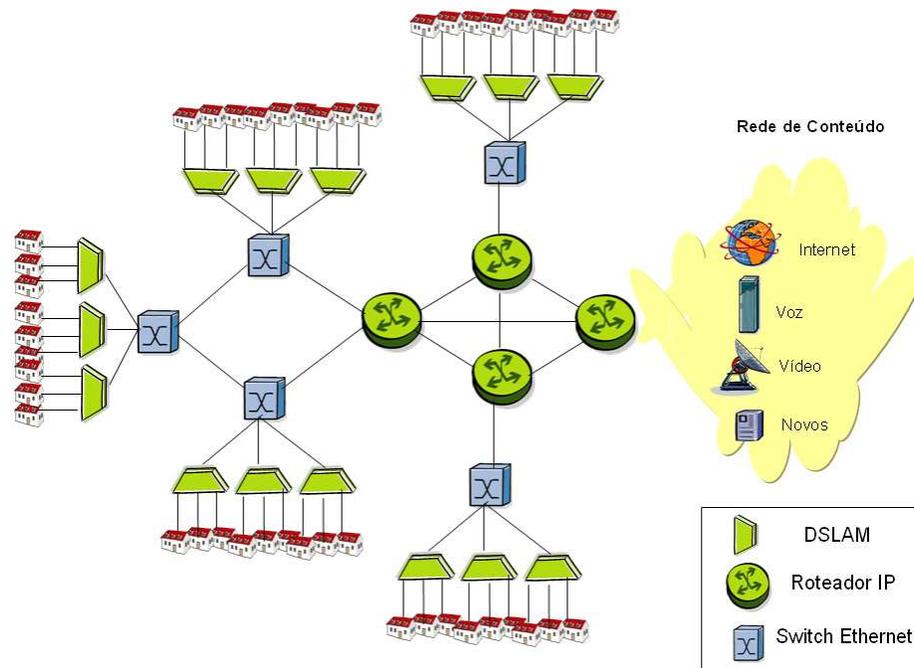


Figura 3.27 - Rede de agregação.

Núcleo da Rede

O núcleo da rede ou a rede *backbone* é composto por um conjunto de roteadores IP que conectam a rede de agregação à rede de conteúdo [12, 21, 26].

O equipamento de conexão entre a rede de agregação à rede IP é conhecido como IP *edge*. Esse roteador faz a terminação da agregação *ethernet* e provê roteamento em direção ao *backbone* IP.

O núcleo da rede IP será composto de roteadores que são puramente roteadores de camada 3 (roteador *core* ao invés de roteador *edge*).

O objetivo do *backbone* IP é rotear o tráfego, da maneira mais eficiente possível, para a aplicação correta na rede de conteúdo.

Resumindo, a rede de transporte pode ser de vários tipos, sendo que a única premissa é que essa rede deve ter uma excelente qualidade de serviço, suportar o *multicast*, ter uma baixa latência e, obviamente suportar uma largura de banda adicional para o vídeo, tanto *unicast* (VOD) quanto *multicast* (canais de TV e PPV).

Rede de Conteúdo

Como mostrado na seção do *head-end*, os codificadores alimentam o núcleo da rede com

os diversos MPEG-2 TS e fazem parte da rede de conteúdo juntamente com uma parte de *Internet*, voz e novos serviços, realizando assim o uso de serviços combinados.

Qualidade de Serviço (QoS)

De maneira intuitiva, qualidade de serviço (*Quality of Service* ou QoS) representa quão rápido os dados podem ser transferidos, quanto tempo o receptor precisa esperar, quão correto os dados são recebidos e quantos dados podem ser perdidos [11, 12, 26].

Tecnicamente, QoS é a medição de três parâmetros:

- *Atraso (Delay)*: consiste no atraso que é introduzido quando um roteador ou nó na rede não é capaz de processar e encaminhar pacotes no tempo esperado. Isso ocorre quando o nó está sobrecarregado ou quando o nó usa mecanismos de prioridade.
- *Jitter*: ele é introduzido sobre uma conexão quando o atraso não é constante sobre uma série de pacotes pertencentes à mesma conexão. Isso pode ser causado pela variação de carga em um roteador.
- *Razão de perda de pacote (ou Packet Loss Ratio)*: esse fator é dependente da carga nos nós.

A maioria dos assinantes está acostumada a um alto nível de qualidade de voz da rede de telefonia pública comutada (RTPC) tradicional bem como ao som e imagem perfeitos provenientes de um aparelho de DVD, e também à transmissão de TV ininterrupta do serviço de TV a cabo ou satélite. Dessa forma, os provedores de serviço devem ser capazes de atender às expectativas de seus assinantes.

O serviço de *Internet* utiliza o melhor esforço (*best effort*). Isso significa que a rede pode descartar ou atrasar pacotes quando ocorrer um congestionamento como resultado de muito tráfego na rede. Melhor esforço representa o tipo de serviço mais simples que uma rede pode oferecer e não provê nenhuma garantia. Quando um *link* está congestionado, os pacotes são simplesmente descartados.

Algumas aplicações, tais como correio eletrônico, navegação via *Internet* ou transferência de arquivos funcionam bem com o melhor esforço, porque, como não são em tempo real, e caso haja alguma perda ou atraso de pacote, a própria aplicação pede a retransmissão do mesmo e o usuário final muitas vezes nem sequer percebe.

Entretanto, o serviço de melhor esforço não é adequado para aplicações em tempo real como conversa usando voz ou o *streaming* de vídeo e, dessa forma, essa opção não é adequada para um IPTV. Caso haja uma perda ou atraso de pacote nesses serviços, isso irá impactar diretamente na qualidade.

Para o caso da aplicação de voz, a largura de banda requerida é muito baixa e os problemas podem ser evitados reservando-se recursos de rede para tal aplicação. Para o

caso do vídeo, como a largura de banda necessária é alta, outro tipo de QoS deve ser aplicado.

QoS leva em consideração a disponibilidade, a qualidade, a prioridade e o atraso. QoS avalia a disponibilidade da largura de banda. Se um cliente quiser menos atraso no vídeo e na voz, é necessário selecionar uma qualidade de serviço apropriada para tal, com mais largura de banda e menor latência. Diversos padrões vêm sendo usados para esses casos como os modos de taxa constante (CBR ou *Constant Bit Rate*) e os de taxa variável (VBR - *Variable Bit rate*).

Os tráfegos de voz, vídeo e dados requerem tratamento diferenciados de QoS. Na maioria dos casos, essa diferenciação se faz por meio de prioridades. Por exemplo, um canal de IPTV prêmio deve ter mais prioridade que o conteúdo multimídia grátis pela *Internet*.

Existem algumas maneiras de se implementar QoS em redes IP. A seguir, apresenta-se o modelo DiffServ (serviço diferenciado ou *Differentiated Services*) que é bastante utilizado.

O modelo DiffServ permite priorizar pacotes de um serviço sobre pacotes de outro serviço. Na Figura 3.28 - Modelo DiffServ vê-se que o roteador da direita recebe pacotes de *Internet* e pacotes de TV. Os pacotes de *Internet* podem tolerar um certo atraso, e com isso, a página *web* será apresentada um pouco depois na tela do computador, mas pacotes de TV não podem tolerar atrasos, caso se deseje que o usuário receba um *stream* de vídeo de boa qualidade. Dessa forma, o roteador deve dar prioridade para os pacotes de vídeo e tratá-los primeiro, e somente depois disso, tratar os pacotes de *Internet*.

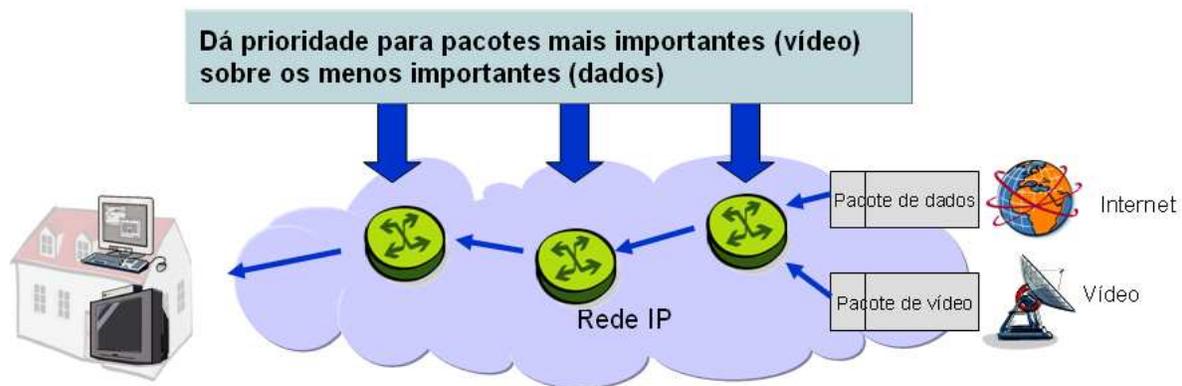


Figura 3.28 - Modelo DiffServ.

Os mecanismos de QoS são somente relevantes em um momento de congestionamento. Caso a rede esteja sobredimensionada, todos os pacotes (*Internet*, vídeo e voz) podem ser facilmente tratados pela rede e todos chegarão ao mesmo tempo.

Na Figura 3.28 - Modelo DiffServ é mostrada uma rede simplificada com apenas roteadores, mas na realidade todos os elementos de rede devem ser capazes de dar prioridade para que os pacotes de vídeo possam ser entregues com uma boa qualidade de serviço. Qualidade do serviço precisa ser feita fim-a-fim: modem, nós de acesso (DSLAM), nós *ethernet* de agregação, roteadores IP, ou seja, todos os elementos envolvidos devem ser capazes de distinguir entre pacotes de diferentes classes de serviços.

Para tornar clara a implementação do QoS, compara-se um nó sem QoS com um com QoS (ver a Figura 3.29 - Nó sem QoS x nó com QoS). Um nó pode ser qualquer elemento de rede que é envolvido no transporte de pacotes, como por exemplo, um roteador, um *switch ethernet*, um nó de acesso ou um *modem*.

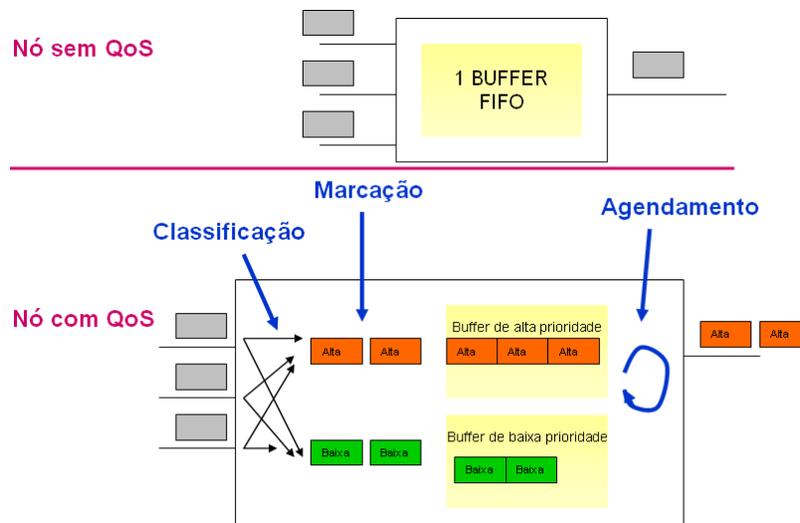


Figura 3.29 - Nó sem QoS x nó com QoS.

O nó sem QoS não é capaz de fazer distinção entre pacotes. Ele recebe os pacotes de vários níveis de importância em seu *link*, mas não pode distingui-los. Para todos os nós os pacotes são iguais, no caso aqui pacotes da cor cinza, usando a regra do FIFO (*First in, First out* ou primeiro que entra, primeiro que sai).

Por outro lado, o nó com QoS é capaz de distinguir entre pacotes de alta ou baixa prioridade. Por exemplo, um nó pode dar prioridade baseado na porta física, por exemplo, todo pacote que chega à porta 1 recebe uma maior prioridade do que os que chegam à porta 2. Existem várias outras possibilidades, incluindo dar prioridade baseado no endereço IP de origem ou destino, porta TCP, etc. A distinção entre prioridade de pacotes é chamada de classificação.

Uma vez que a classificação é feita, o nó pode colocar uma etiqueta nos pacotes para indicar a prioridade. Esse processo é chamado de marcação e os nós posteriores podem

fazer uso dessa marcação já feita por aquele nó. Nesse caso, a classificação poderia ser baseada nos bits QoS e pular o processo de marcação.

O nó coloca os pacotes de alta prioridade em um *buffer* separado dos de menor prioridade. Os nós sempre processam os pacotes do *buffer* e, quando um *buffer* está cheio e chega um novo pacote para aquele *buffer*, o pacote será descartado. A Figura 3.29 - Nó sem QoS x nó com QoS mostra dois *buffers*, mas poderia ser mais.

O agendamento irá decidir como esvaziar o *buffer*. Uma maneira de se implementar poderia ser a de tratar os pacotes de baixa prioridade somente depois que o *buffer* de alta prioridade se esvaziar, mas esse método pode por em risco os pacotes de baixa prioridade que eventualmente nunca serão tratados. Outra possibilidade é, por exemplo, que o nó trate uma grande parte dos pacotes de alta prioridade, então tratar parte do *buffer* de menor prioridade e depois voltar para o *buffer* de maior prioridade.

Note que esse mecanismo dá um controle sobre todos os três parâmetros do QoS:

- Perda de pacote: pacotes de alta prioridade serão tratados primeiro pelo agendamento, fazendo com que a chance de que o *buffer* fique sobrecarregado e os pacotes sejam descartados se torne pequena.
- Atraso: *buffer* de alta prioridade será tipicamente pequeno. Um *buffer* pequeno significa uma menor fila, fazendo com que os pacotes nunca esperem muito tempo para serem processados.
- *Jitter*, variação no atraso: um atraso será sempre pequeno, e então a variação no atraso também será pequena. Um pacote será tratado muito rápido.

Medição de Audiência

Outra característica muito atrativa oferecida pelo IPTV é a medição de audiência em tempo real, com relatórios do uso dos canais de TV. A plataforma de IPTV permite um conhecimento completo e real sobre as preferências e o comportamento de maneira não intrusiva.

Isso é feito através das mensagens de log de IGMP recebidas pelo DSLAM e enviadas para um servidor de audiência. Sempre que um usuário muda um canal, um comando de IGMP *leave* (desconexão) será enviado a um DSLAM seguido por um IGMP *join* (conexão). Assim o DSLAM coletará todas as mudanças de canal (*zapping*) de todos os usuários.

Dessa forma, o DSLAM pode ser configurado para enviar as informações de IGMP para um servidor de audiência local, sempre que algum usuário trocar de canal.

Sendo assim, quando um usuário muda de canal, algumas informações serão enviadas para o servidor de audiência local: dados do usuário, endereço *multicast* do canal de

origem, endereço *multicast* do canal de destino e dados do DSLAM.

Os servidores de audiência local vão filtrar as informações, torná-las anônimas e enviá-las para o servidor de audiência central, que pode também ter a função de servidor *web*. A Figura 3.30 - Medição de audiência mostra uma representação típica para a coleta de audiência de canais de TV.

Para evitar uma avalanche de dados em um servidor, deve-se ainda limitar o número de DSLAMs que enviarão os dados para um servidor de audiência local.

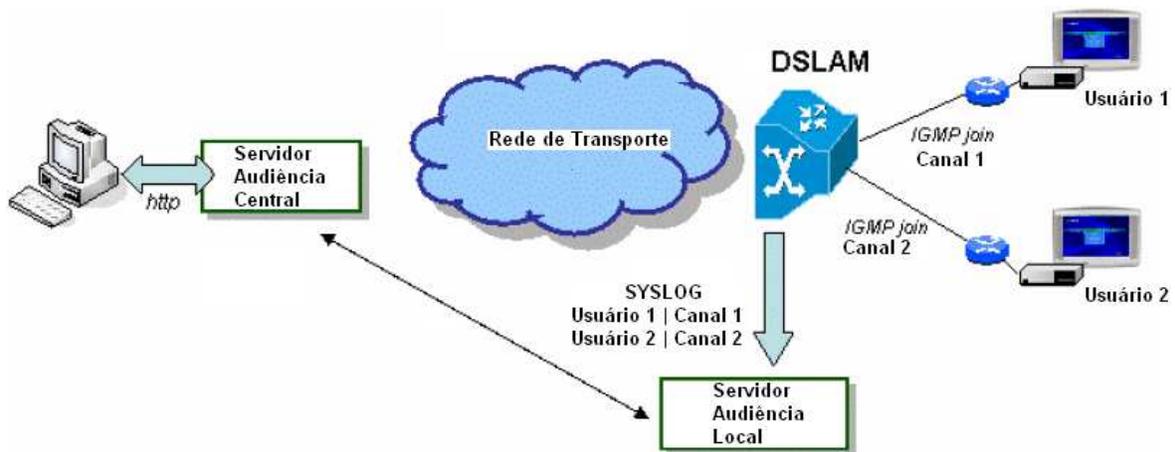


Figura 3.30 - Medição de audiência.

O servidor de audiência central pode ser implementado de forma tal que se possa ver em tempo real quantos usuários assistiram a um determinado programa, quantos usuários mudaram de canal durante um comercial, ou quais os programas mais assistidos pelos usuários de uma determinada região. Pode-se ainda gerar relatórios contendo dados de audiência máxima, mínima, média ou real para todos os canais, ou mesmo para um conjunto de canais. Isso se deve ao fato de que essas informações podem ser combinadas com as informações da base de dados e as informações de programação eletrônica dos canais.

Um cuidado que deve ser tomado na medição de audiência é o de tornar os dados anônimos por questões legais em diversos países. Não se pode associar o nome de uma pessoa a um canal. Isso pode ser feito também pelo servidor de audiência local.

A Figura 3.31 - Medição de audiência instantânea para canais de TV mostra um exemplo de uma tela de medição de audiência.

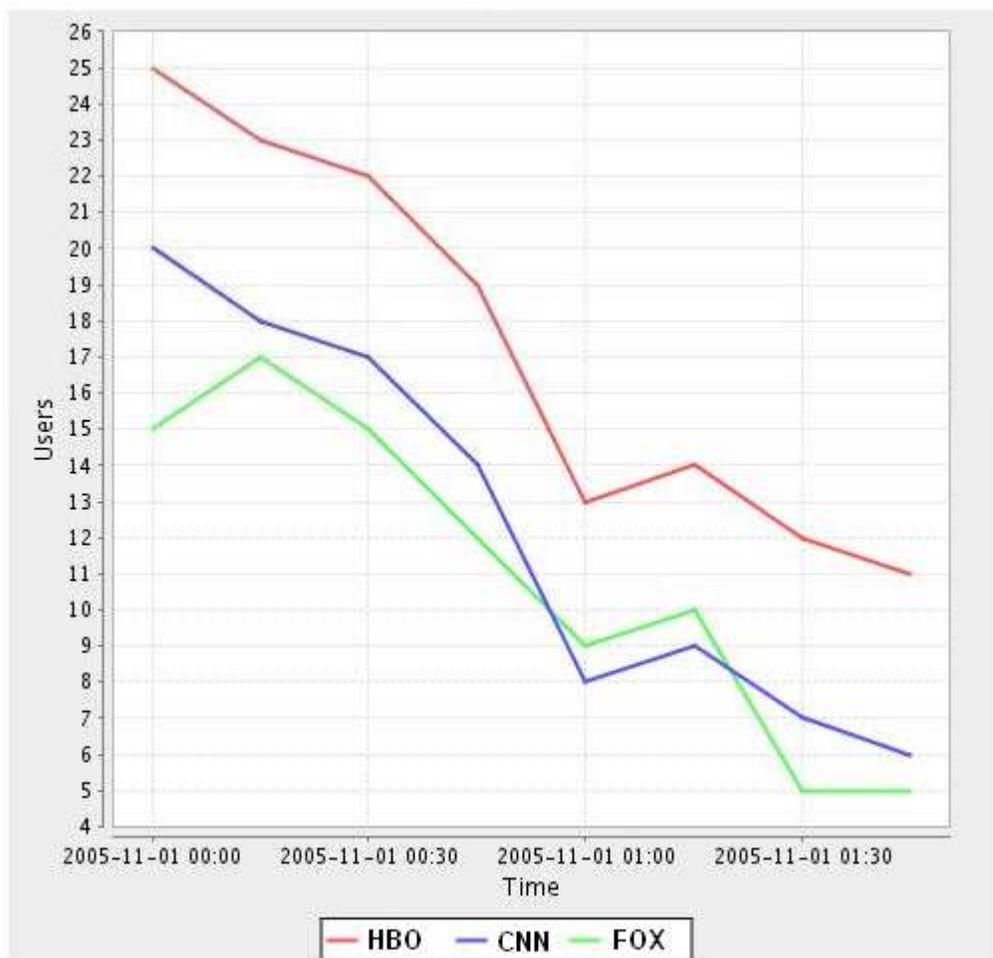


Figura 3.31 - Medição de audiência instantânea para canais de TV.

A medição de audiência pode ainda ser usada, coletando-se dados dos servidores de aplicação do *middleware* para conteúdos sobre demanda. Dessa forma, poderia ser possível gerar também relatório dos filmes mais assistidos, ou menos assistidos num determinado período. Na Figura 3.32 - Medição de audiência VOD mostra-se uma tela de exemplo para medição de dados de VOD, nesse caso, os 10 filmes mais assistidos no período de 01/01/2007 à 01/07/2007.

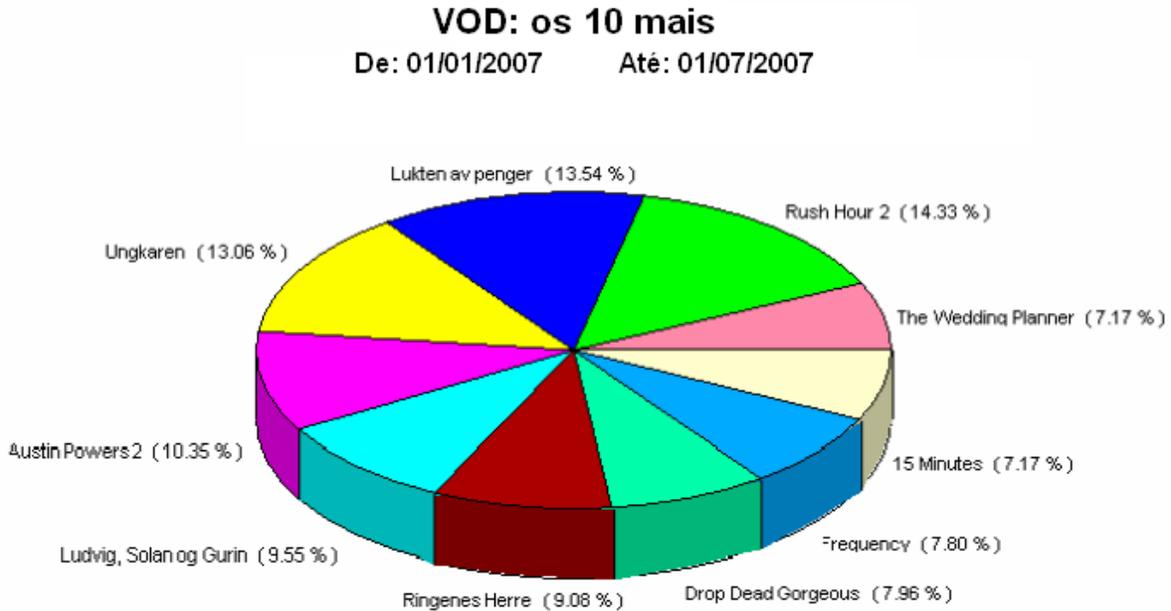


Figura 3.32 - Medição de audiência VOD.

Essas informações podem ser guardadas e pode-se com isso gerar um perfil para cada usuário, indicando quais os programas que cada usuário gosta de assistir. Isso poderia propiciar o uso de comerciais customizados e direcionados para cada público.

A medição de audiência real é um diferencial que a solução de IPTV apresenta em relação a outras soluções de TV e pode ser explorada de uma forma ainda mais lucrativa do que a já adotada pela solução de TV por cabo, por exemplo.

Serviços de IPTV

Tipicamente uma solução de IPTV suporta uma grande gama de serviços, fáceis de usar, que conduzem a uma experiência inovadora na TV através de uma interface de usuário bastante intuitiva, incluindo serviços de TV e sob demanda. A maioria dos serviços de usuário final já está pronta para uso e alguns desses serviços são descritos a seguir.

Outros serviços estão sendo criados, podendo inclusive ser uma combinação de vários outros, todos eles integrados e transparentes para um assinante.

O preço dos serviços vai depender de vários fatores, mas um fator predominante será se o serviço é *multicast* ou *unicast*. Como mostrado anteriormente, conteúdos *unicast* são ponto a ponto e, se usados em escala, pode comprometer toda a rede. Já os conteúdos *multicast* podem ser compartilhados entre vários usuários. Devido a isso, os conteúdos *unicast* devem ser mais caros, uma vez que o usuário vai usar um trecho da rede somente para ele.

A seguir são descritos diversos serviços finais para os usuários de IPTV.

Guia de Programação Eletrônica

O guia de programação eletrônica (*Electronic Programming Guide* ou EPG) é um portal interativo que lista todos os conteúdos disponíveis, além de conter mecanismos avançados de busca por programa, por horário ou por gênero (ver a Figura 3.33 - Guia de programação eletrônica).



Figura 3.33 - Guia de programação eletrônica.

Canais de TV ou de Áudio

Conforme foi citado no tópico abordado sobre *middleware*, os canais devem ser criados e associados a um pacote de TV. Esses canais devem ainda ser agrupados em pacotes, onde se define um preço mensal. Os usuários então devem se inscrever a um determinado pacote para ter acesso aos canais de TV. Enquanto o usuário assiste ao canal ele pode ainda fazer uso do mini guia ou mini EPG, enquanto se assiste ao canal. Para se ter acesso a esse serviço, o assinante deve pagar uma quantidade fixa por mês (ver a Figura 3.34 - Canais de TV e mini guia).

De forma análoga, podem-se criar canais de áudio apenas no *middleware* e também associá-los a pacotes.

Um canal de áudio tem um formato exatamente igual ao de um canal de vídeo, que é um MPEG-2 TS, encapsulado por IP *multicast*. A diferença é que ele não vai ter o fluxo elementar de vídeo.

Do ponto de vista do *head-end* esse canal pode ser recebido já codificado no formato certo, ou ainda ser recebido no formato analógico e codificado pelo próprio provedor de serviço. Nesse caso, faz-se necessário o uso de codificadores de áudio apenas.



Figura 3.34 - Canais de TV e mini guia.

Controle Paterno

Controle paterno (*Parental Control*) é um serviço bastante conhecido, mas que no IPTV permite que o bloqueio seja feito não somente por canal, mas também pela classificação do programa. Dessa forma, o campo de classificação do programa, existente na EPG pode ainda ser usado para impedir que seus filhos tenham acesso a determinados programas (ver a Figura 3.35 - Controle paterno).

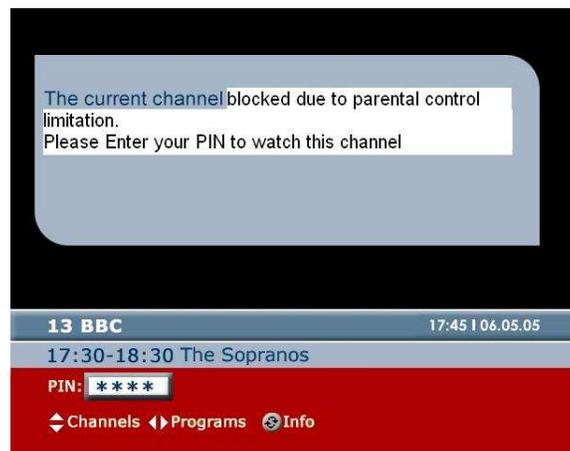


Figura 3.35 - Controle paterno.

Pay Per View

Pay Per View (PPV ou pague para ver) permite aos provedores de serviços oferecerem

conteúdos ao vivo e pré-gravados para os assinantes em horários agendados e acessíveis somente através de uma compra adicional. Do ponto de vista do *head-end*, um canal PPV é exatamente igual a qualquer outro canal transmitido por *multicast*. Entretanto, para o *middleware* esse canal será criado de maneira diferente, como sendo PPV. Isso significa que quando um usuário apertar o número de um canal PPV, ele não terá uma imagem na tela, mas sim uma interface mostrando a ele todos os programas PPVs disponíveis para compra. Caso o assinante, resolva comprar aquele programa, o canal (*stream*) será liberado para ele minutos antes de começar o respectivo programa (ver a Figura 3.36 - *Pay per view*).

Do ponto de vista do provedor de serviço, qualquer canal pode ser comercializado como sendo PPV. Como por exemplo, canais com jogos de futebol, ou lutas de boxe, canais de filmes, ou canais de músicas.



Figura 3.36 - *Pay per view*.

Video on Demand

Vídeo sob demanda (*Video on Demand* ou VOD) habilita aos clientes comprar um vídeo só para ele (*unicast*) por um período (por exemplo, por 24 horas), usando-se uma listagem de títulos. Também conhecido como locadora virtual, esse serviço permite que o usuário assista o conteúdo diversas vezes, avance, pause ou retrocesse. O serviço de VOD permite ao usuário uma experiência semelhante ao do DVD, incluindo avanços e retrocessos em várias velocidades, avanço ou retrocesso de capítulos e pausa (ver a Figura 3.37 - *Video on demand*).

Esse é outro atrativo que o IPTV oferece e como comentado no tópico de servidor de vídeo, pode-se usar uma estrutura de servidores de vídeo descentralizada, de forma a não ocupar toda a rede com tráfegos *unicast*.



Figura 3.37 - Video on demand.

Near Video on Demand

Near Video On Demand (NVOD) é uma variação do PPV para conteúdos pré-gravados, permitindo ao usuário comprar um conteúdo que está sendo transmitido periodicamente em um intervalo curto de tempo, como por exemplo, 5 ou 15 minutos.

Esse é um serviço extremamente interessante e pode ser bastante útil para um provedor de serviço. Imagine que um filme muito popular é apresentado a um provedor de serviço. Caso esse provedor lance esse filme como VOD e, como é um filme atípico e muito esperado, no dia ou na semana do lançamento, vários usuários podem comprá-lo. Isso geraria uma avalanche de *unicasts* na rede, que poderia não ser suportado pela rede. Entretanto, se esse provedor lançar esse filme, na primeira semana, como sendo NVOD, ele consegue controlar a quantidade de *streams* na rede nessa semana.

Se esse filme tem a duração de 2 horas e o operador cria um NVOD espaçado de 15 minutos, significa que a rede terá um total de 8 *streams multicast* para aquele conteúdo. E sempre que um deles acaba ele começa em seguida, em *loop*.

Na semana seguinte ele pode retirar esse filme NVOD de circulação e lançá-lo como VOD porque provavelmente várias pessoas já assistiram e não vão assistir de novo.

Esse serviço por ser *multicast* será mais barato para o usuário final, mas obviamente, não poderá ser pausado, avançado, ou retrocedido. Entretanto o usuário pode alterar entre os outros *streams*, como se fosse um canal de TV. Assim, ele pode navegar para o *stream* que está a alguns minutos defasados.

Finalmente, o servidor de vídeo tem que ser capaz de fazer *stream* em *multicast* para que esse serviço funcione.

Subscription Video on Demand

Subscription Video On Demand (SVOD ou VOD por subscrição) permite que provedores de serviço ofereçam conteúdos *unicast* a uma taxa mensal por subscrição e não por título. Como por exemplo, um usuário pode assistir qualquer conteúdo que esteja na categoria de "filmes antigos" por uma taxa mensal fixa, não importando quantas vezes ele assista. Pode também ser usado para documentários ou séries de TV.

Esse serviço é exatamente igual ao VOD do ponto de vista técnico, sendo a única diferença a forma de cobrança (ver a Figura 3.38 – *Subscription video on demand*).



Figura 3.38 – *Subscription video on demand*.

Music on Demand

Music on Demand (MOD, música sob demanda ou áudio sob demanda) permite ao usuário comprar uma única música, várias músicas de um álbum ou de um grupo musical só para ele (*unicast*) por um período, por meio de uma listagem de títulos.

Como os títulos de música são de curta duração e tem uma taxa bastante baixa, o provedor de serviço pode, alternativamente, comercializar esse serviço como sendo por subscrição mensal.

Personal Video Recorder (PVR)

Personal Video Recording (PVR ou gravação de conteúdo pessoal local) permite ao usuário gravar conteúdos ao vivo, utilizando o seu próprio STB. Nesse caso, o STB do usuário necessita ter um HD (*Hard Disk*) embutido. O PVR tem o mesmo comportamento do gravador de vídeo pessoal clássico, sendo a diferença o fato de que o usuário simplesmente escolhe qual programa quer gravar pelo EPG e não precisa se preocupar com horário de início e de fim. Mesmo que o programa mude de horário, o EPG será atualizado

e a gravação será realizada corretamente (ver a Figura 3.39 - *Personal video recording*).

Além disso, o usuário não precisa esperar o programa terminar para começar a assistir, como nos gravadores clássicos. Isso pode ser feito a qualquer momento após o início da gravação. Essa funcionalidade está relacionada ao próximo serviço que será discutido a seguir, que é o de pausa na programação ao vivo.

Programas sob demanda não podem ser gravados por restrição do provedor de conteúdo. Assim, nem sequer irá aparecer uma opção para se gravar.

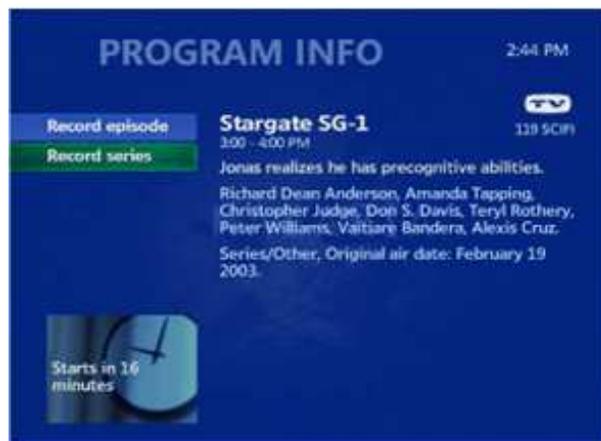


Figura 3.39 - *Personal video recording*.

Pause Live TV

Pause Live TV ou pausa na TV ao vivo permite que o usuário aperte o botão de pausa em uma programação ao vivo. Tendo um funcionamento bastante similar ao PVR, quando o usuário aperta o botão de *pause*, seu STB começa imediatamente a gravar a programação ao vivo. O usuário pode a qualquer momento voltar a assistir ao programa e inclusive avançar essa programação até chegar de novo ao ponto ao vivo que pode ser visto na barra de progresso da Figura 3.40 - *Pause live TV*.



Figura 3.40 - *Pause live TV*

Network Personal Video Recorder

Network Personal Video Recording (NPVR ou PVR na rede) permite a habilidade de gravar conteúdos ao vivo, utilizando-se os recursos da rede (servidores VOD), e uma posterior transmissão *unicast* sob demanda com as mesmas funcionalidades do VOD. Esse serviço é similar ao PVR, com a diferença de que o STB não precisa ter HD.

Entretanto, para que esse serviço funcione, algumas considerações devem ser feitas.

O servidor de vídeo deve ter interfaces para os canais IGMP que vem do *head-end* e devem ser capazes de gravar programas quando solicitado.

Além disso, os provedores de conteúdo podem impor algumas restrições para que determinado programa não possa ser gravado.

A Figura 3.41 – NPVR mostra um diagrama típico de uma solução de NPVR e dependendo do número de assinantes, o servidor de NPVR central e local podem estar na mesma máquina.

O operador deve definir, por meio dos dados de programação eletrônica, quais os programas que podem ser gravados, para todos os canais. Como foi dito anteriormente, os dados de EPG serão mantidos na base de dados. Por sua vez, o servidor de aplicação que está constantemente lendo essa base de dados, irá mostrar na tela do usuário um botão de gravar, caso o programa esteja marcado como gravável e caso o usuário tenha assinado o serviço de NPVR.

Em geral, os usuários têm uma cota em horas para gravação NPVR no servidor de vídeo, durante certo número de dias.

Caso esse usuário aperte o botão de gravar, seu STB irá enviar uma mensagem para o AS, que por sua vez irá solicitar acesso ao servidor de NPVR Central. O servidor de NPVR central vai verificar se o usuário tem horas disponíveis para gravação, qual o servidor de vídeo mais próximo dele e se ele é o primeiro a solicitar a gravação daquele programa naquele servidor.

O servidor NPVR central tem uma inteligência tal que se mais de um usuário pedir para gravar o mesmo programa no mesmo servidor, somente uma cópia será feita.

Os usuários podem apagar seu conteúdo após assistirem um programa, mas da mesma forma, o mesmo só será excluído do servidor de vídeo quando todos os usuários que solicitaram a gravação pedirem para apagar. Dessa forma, o usuário tem a sensação de que existe uma área do servidor de vídeo só para ele, mas o sistema vai usar isso de forma inteligente.

Quando estiver próximo do horário de início do programa, o servidor de NPVR central irá solicitar ao servidor de NPVR local mais próximo do assinante que uma gravação seja feita no canal "X", com a hora de início e fim do programa (por exemplo, "igmp://224.10.1.3:2222 das 13:30 às 14:15").

O servidor de NPVR local por sua vez vai solicitar ao servidor de vídeo que faça a gravação. Esse último irá então fazer um comando de IGMP *join* para aquele canal e capturar o fluxo no período especificado.

Após a gravação o conteúdo estará disponível para os usuários que o solicitaram e será assistido como *unicast*. Dessa forma, os usuários poderão avançar, retroceder, parar o conteúdo ou mesmo, assisti-lo várias vezes.

Uma variante desse serviço é conhecida popularmente como "TV de ontem" ou NPVR de operador ou *Time Shifted TV* (TSTV). Nesse caso, o operador irá selecionar todos os programas de alguns canais que sempre serão gravados e com isso o sistema irá automaticamente efetuar a sua gravação, sem que qualquer usuário a solicite.

Em geral, a operadora define alguns canais, que são mais assistidos, e gravam todos os seus conteúdos e os mantêm por 2 ou 3 dias, por exemplo. Assim, os usuários que assinaram esse serviço podem assistir à programação defasada no tempo, daí o nome TSTV.

O TSTV e o NPVR são serviços independentes e, em geral, cobrados mensalmente do usuário, independente de se ele está ou não fazendo uso. Um usuário pode assinar somente o serviço de TSTV, somente o serviço de NPVR, os dois serviços ou nenhum deles. Entretanto, para o servidor de NPVR central, os serviços não diferem. Por exemplo, se um programa já vai ser gravado por definição da operadora e se alguém solicitar para gravá-lo, pelo NPVR, novamente, somente uma cópia será feita.

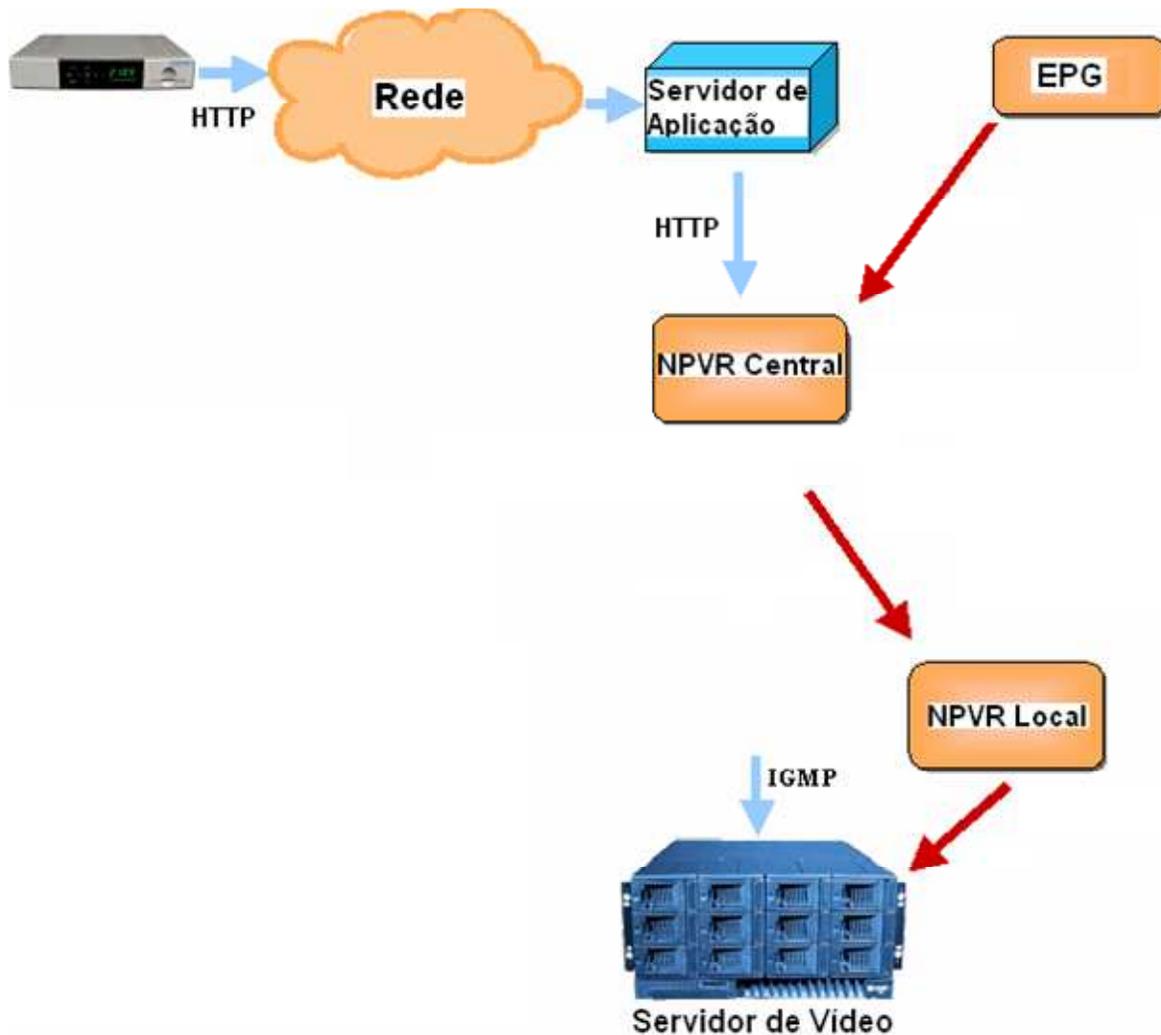


Figura 3.41 – NPVR.

Games on Demand

Games on Demand (GOD ou jogos sob demanda) permite que usuários selecionem jogos de uma listagem e os joguem contra o sistema ou mesmo com outros jogadores. Esse serviço exige um servidor de jogos bastante poderoso e um controle (*joystick*) conectado na porta USB do STB. Além disso, o jogo deve ser quebrado em várias cenas e inserido no servidor. O usuário recebe o jogo como se fosse um VOD, via *unicast* e quando ele aperta um botão para dar um tiro, por exemplo, um comando, com suas coordenadas, é enviado para o servidor de jogos que muda a cena. Tudo isso é realizado de forma bastante rápida e transparente, sendo que o usuário dificilmente percebe um atraso.

Esse serviço oferece várias vantagens e é um forte aliado no combate à pirataria, uma vez que os jogos são armazenados cifrados no servidor. Além disso, o usuário não consegue

copiá-lo, uma vez que receberá somente o *stream* em tempo real no momento em que estiver jogando. Para o usuário existe a vantagem de não necessitar de um computador potente, com muita memória e com boa placa de vídeo, uma vez que o STB que vai decodificar e decifrar seu conteúdo.

Walled Garden

O serviço *Walled Garden* permite aos usuários navegar num portal usando seu aparelho de TV. Nesse serviço o usuário pode navegar somente em um ambiente controlado e, terão acesso somente aos conteúdos que são oferecidos pelo provedor. Mais uma vez, o provedor pode usar essas páginas amarelas para fazer anúncios de publicidade. A Figura 3.42 - *Walled garden* mostra algumas telas desse serviço, onde podem ser disponibilizadas diversas informações úteis, incluindo listas de cinema, clima e tempo, ações e muitas outras.

Esse serviço exige outro servidor *web* para hospedar toda a informação em um formato já adequado para a tela da TV.



Figura 3.42 - *Walled garden*.

Web TV

Trata-se de um serviço de acesso à *Internet* que permite aos usuários navegar na rede usando seu aparelho de TV. Entretanto, como se sabe, o formato das páginas de *Internet* para computador é diferente do tamanho da tela da TV. Dessa forma, se o provedor de *Internet* não está preparado para isso, a experiência que o usuário vai ter não será muito boa, pois ele terá que usar as teclas de flechas de controle remoto para ver todo o conteúdo.

Para que esse serviço funcione de forma adequada, duas alternativas podem ser usadas. A primeira seria a adequação por parte dos provedores de *web*, como por exemplo, *Google*

e CNN que verificam de qual navegador se está requisitando a página e automaticamente ajustar os dados para o tamanho correto. Outra alternativa seria o uso de uma máquina intermediária para fazer essa conversão, o que não é muito simples.

Devido a isso, é mais comum encontrar o serviço de *Walled Garden*, que é mais preferido pelos provedores de serviço. Alguns ainda fazem uso do serviço mesmo com a formatação inadequada.

Vídeo Conferência

O serviço de vídeo conferência permite uma conversa com vídeo através da rede. Nesse caso, as câmeras de vídeo são conectadas à porta USB do STB, mas elas não são câmeras tipo padrão. São câmeras que fazem compressão de imagem e que têm microfone embutido.

Na Figura 3.43 - *Vídeo conferência* são mostradas algumas telas e os usuários devem ser cadastrados como se fossem para um serviço MSN ou *Yahoo*. Caso o usuário esteja *on-line*, ele poderá ser convidado para uma conversa com vídeo.

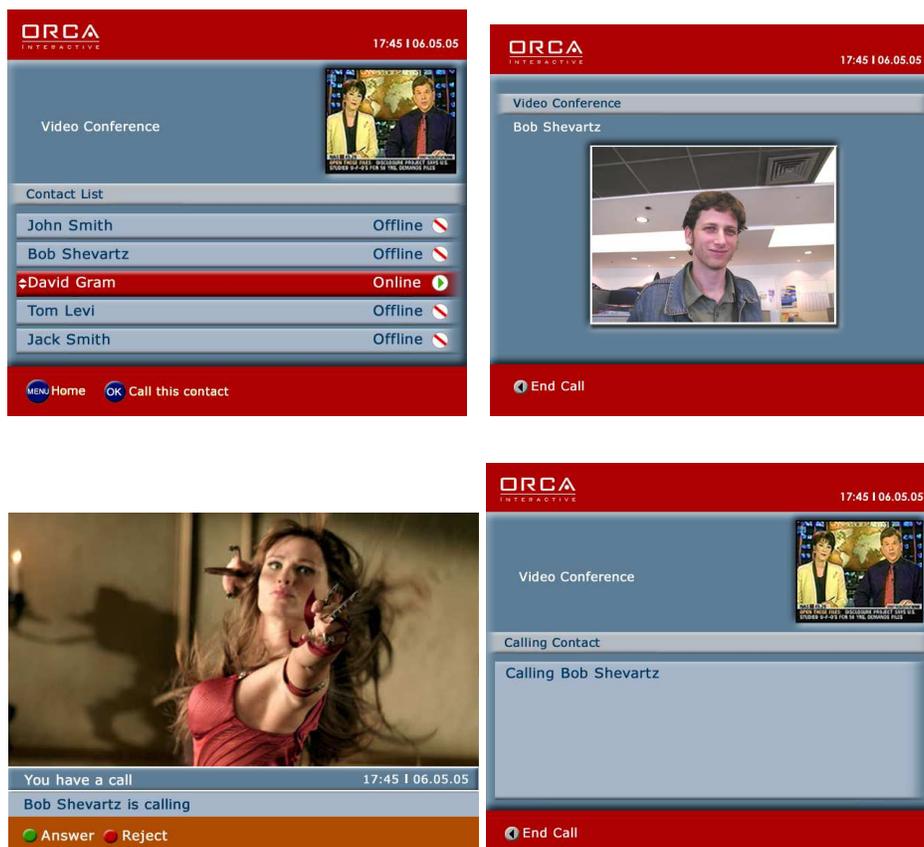


Figura 3.43 - *Vídeo conferência*.

Interface de Usuário de Vários Idiomas

Por meio das interfaces de usuário, o assinante pode selecionar o idioma em que quer ver as mensagens, a sinopse dos filmes e a programação de EPG. Para que isso funcione bem o operador deve carregar os dados em todos os idiomas já pré-selecionados. Uma janela típica para selecionar o idioma é mostrada na Figura 3.44 - Interface de usuário de vários idiomas.



Figura 3.44 - Interface de usuário de vários idiomas.

PiP e Mosaico Dinâmico

O serviço de PiP (*Picture in Picture*) permite que o usuário possa assistir a um canal de TV em tela cheia e outro em um quadrado menor em um dos cantos da tela. O canal de PiP tem uma menor resolução e deve ser disponibilizado pelo *head-end* com um endereço *multicast* diferente. Além disso, o STB deve ser capaz de decodificar dois canais ao mesmo tempo.

Já o serviço de mosaico dinâmico, como mostrado na Figura 3.45 - Mosaico dinâmico disponibiliza vários canais de uma só vez para o usuário. Diferente do PiP que são canais distintos, o mosaico é apenas um canal e é gerado por equipamentos especializados no *head-end*. Esse equipamento recebe todos os canais de PiP e monta alguns canais de acordo com a especificação do provedor de serviço. Como por exemplo, canais de esporte, ou filmes ou notícias são agrupados em um único canal por esse equipamento. Além disso, esse equipamento permite que o usuário selecione cada uma das janelas pequenas e, caso ele deseje, ele pode mudar para o canal em tela cheia. Caso isso ocorra, o STB do assinante vai executar um comando IGMP *leave* de canal mosaico e IGMP *join* para o canal em tela cheia.



Figura 3.45 - Mosaico dinâmico.

Múltiplos Ângulos e Múltiplos PiPs

Essa opção é uma variação do serviço anterior e permite que o usuário veja um programa de vários ângulos diferentes, como por exemplo, em um jogo de futebol, com vários PiPs como mostra a Figura 3.46 - Múltiplos ângulos.



Figura 3.46 - Múltiplos ângulos.

Lembrete

É possível criar lembretes para programas favoritos. Assim, um minuto antes de se começar um programa, o sistema apresentará uma mensagem perguntando se ele deseja mover para o canal favorito ou se quer ignorar a mensagem. Os usuários usando seu controle remoto podem adicionar quantos lembretes eles desejarem e essa informação será armazenada na base de dados. A Figura 3.47 – Lembrete mostra um exemplo desse serviço.

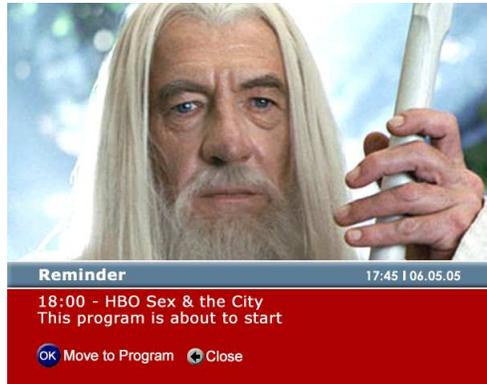


Figura 3.47 – Lembrete.

Canal Pessoal

É possível que o usuário crie o seu próprio canal de TV, injetando vídeo ou fotos no sistema e publicando para familiares e amigos que ele deseja. Para que isso funcione, o operador deve disponibilizar um acesso remoto aos seus usuários que irão inserir esses dados no servidor de vídeo. A Figura 3.48 - Canal pessoal mostra um exemplo desse serviço.



Figura 3.48 - Canal pessoal.

Assistir Programa com Outras Pessoas Remotas

Esse serviço permite que um usuário assista a um programa juntamente com seus amigos, cada um na sua própria casa e envie mensagens de texto ou "carinhas" na tela, expressando seu humor como mostra a Figura 3.49 - Assistir programas com pessoas

remotas.

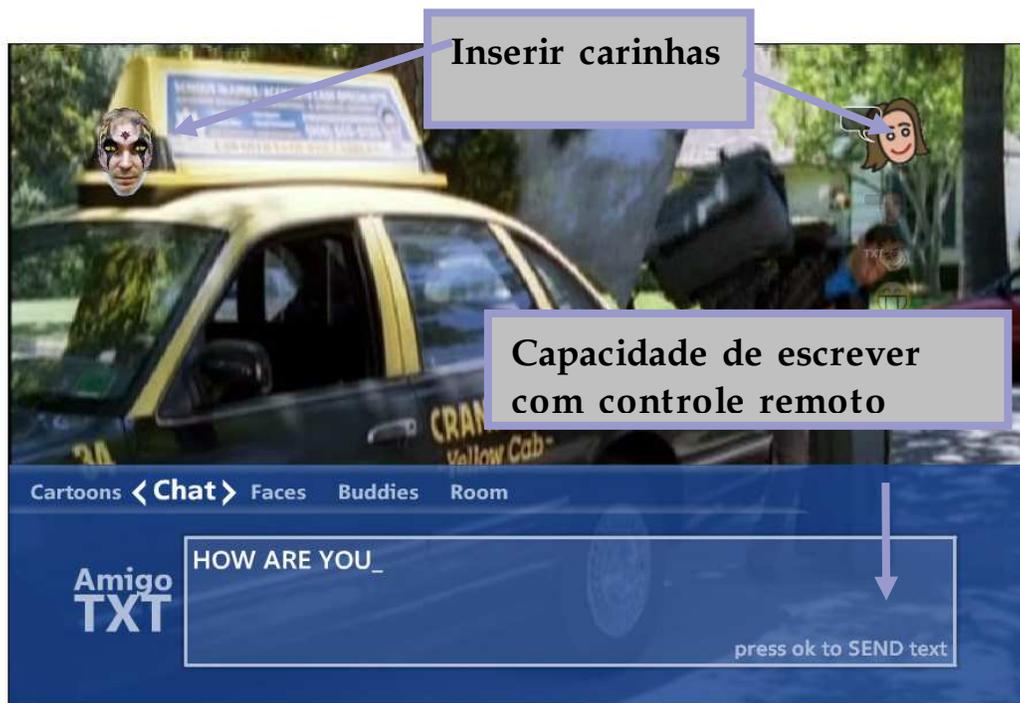


Figura 3.49 - Assistir programas com pessoas remotas.

Identificação Interativa de BINA na TV

Esse serviço, também conhecido como *Interactive TV Caller ID*, é um exemplo de um serviço convergente e integrado que tira vantagens dos serviços de telefonia e TV. Nesse caso o provedor de serviço pode notificar o assinante na sua TV quando ele receber uma chamada telefônica. Esse serviço permite ainda que o usuário, usando seu controle remoto, tome uma ação, como transferir a chamada para a caixa postal ou rejeitar a chamada, como é mostrado na Figura 3.50 - *Interactive TV caller ID*.

Para que essa funcionalidade opere de modo satisfatório, o *middleware* deve estar integrado com um componente do IMS conhecido como *Service Broker* (SB). O SB tem várias interfaces e protocolos diferentes e é usado para integração de vários equipamentos. Assim, quando alguém liga para esse usuário, a chamada será encaminhada para o SB, que por sua vez contacta o *middleware* que envia os comandos para o STB.



Figura 3.50 - *Interactive TV caller ID.*

T-Commerce

T-commerce ou comércio eletrônico pela TV é um poderoso serviço que possibilita aos assinantes de IPTV pedir informações adicionais ou realizar uma compra diretamente de sua tela de TV. Novamente aqui, o *middleware* tem que estar integrado com o site de compra e dessa forma, o usuário não necessitaria de entrar com endereço ou nenhuma outra informação, mas somente com sua senha (ver a Figura 3.51 - *T-Commerce*). Assim, quando um usuário pedir uma pizza, por exemplo, ele só terá que escolher qual sabor e o produto chegará diretamente na casa dele. A forma de pagamento, nesse caso, seria feita diretamente na sua fatura. Esse serviço tem algumas vantagens para o provedor de serviço, primeiro porque o usuário não vai sair da frente da TV e segundo porque a operadora pode cobrar uma pequena comissão dos vendedores por esse serviço.



Figura 3.51 - *T-Commerce.*

Gerência de Rede

Como a solução de IPTV envolve vários servidores e equipamentos de diferentes fabricantes, recomenda-se o uso de um sistema de gerência superior que vai coletar alarmes de todos os sistemas. Em rede IP utiliza-se muito o SNMP (*Simple Network Management Protocol* ou protocolo de gerência simples de rede).

O SNMP é um protocolo usado para comunicar uma informação de gerência entre a estação de gerência de rede e os agentes, como por exemplo, roteadores, *switches*, DSLAMs e servidores. Em conformidade com esse protocolo, os equipamentos desenvolvidos por diferentes fabricantes podem ser gerenciados por um único programa. O protocolo SNMP é amplamente usado, via protocolo IP e opera sobre UDP com as portas 161 e 162. Seus alarmes são chamados de *trap* SNMP e seguem um padrão específico.

Um diagrama de blocos é mostrado na Figura 3.52 - Sistema de gerência SNMP.

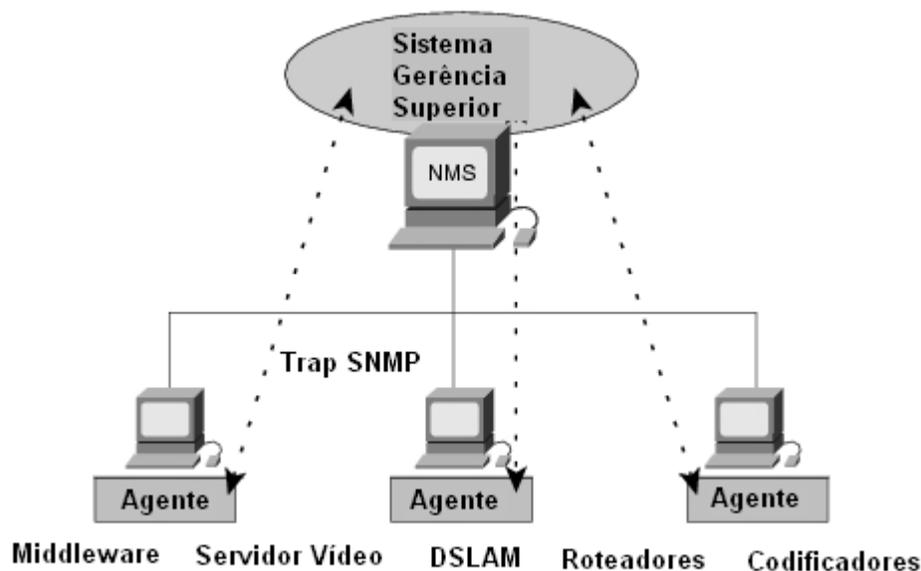


Figura 3.52 - Sistema de gerência SNMP.

Página em branco

Capítulo 4

Sistema de IPTV: Proposta e Implementação

Nesse capítulo, propõem-se algumas melhorias e sugestões que poderiam vir a ser implementadas e aplicadas em uma solução de IPTV.

Head-End

Redundância

No *head-end*, além do que já foi exposto no capítulo 3, sugere-se o uso da proteção de redundância para todos os equipamentos críticos, visto que se ocorrer alguma falha em qualquer desses dispositivos, um ou mais canais ficarão indisponíveis para todos os assinantes.

- Redundância para as antenas de satélites, com o uso de uma antena motorizada de *backup*;
- Redundância para os receptores de satélite (IRDs);
- Redundância para os codificadores;
- Redundância para os sistemas de gerência do *head-end*;
- Redundância dos *switches ethernet* que irão receber o tráfego *multicast* de todos os codificadores.

Inserção de legenda

Ainda na parte do *head-end*, propõe-se a implementação de uma solução de inserção de legenda. Essa solução foi implementada e testada no laboratório da Alcatel-Lucent.

Como comentado anteriormente, é muito comum o envio, por parte dos provedores de conteúdo, da legenda juntamente com o fluxo de vídeo e de áudio. Entretanto, cada provedor, usa um formato diferente e proprietário para se codificar uma legenda.

A Figura 4.1 - Inserção de legenda - TV *broadcast* mostra um modelo que poderia ser usado para transcodificar a legenda para um formato conhecido e fazer a sua inserção dentro do MPEG-2 TS.

Nesse caso, o receptor de satélite extrai a legenda em formato proprietário e envia para o transcodificador de legenda, que irá transformá-la em DVD *bitmap*. Ao mesmo tempo, o receptor de satélite envia o sinal de áudio/vídeo SDI para o codificador. O codificador irá comprimir os sinais, criar o MPEG-2 TS e fazer o encapsulamento IP *multicast*. Um multiplexador IP/ASI, vai receber esse fluxo IP *multicast* vindo do codificador, remover o encapsulamento e abrir o MPEG-2 TS. O multiplexador vai então criar um novo MPEG-2 TS, adicionar os fluxos elementares do MPEG-2 TS vindo do encoder, adicionar a legenda recebida do transcodificador e encapsulá-lo novamente para outro IP *multicast*, agora com uma legenda.

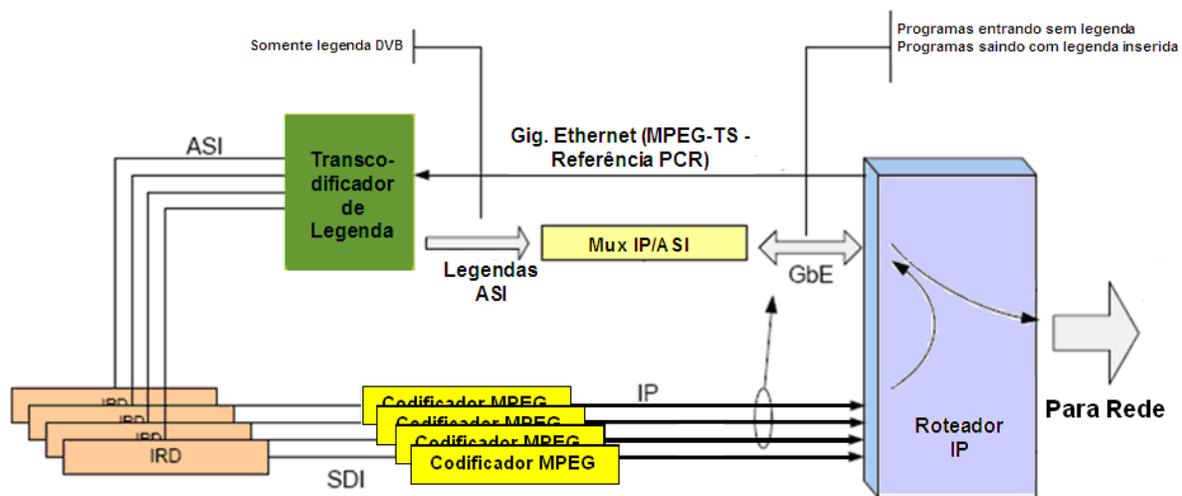


Figura 4.1 - Inserção de legenda - TV *broadcast*.

Note que o sinal do PCR deve ser usado para sincronizar a legenda com o áudio e vídeo. Os usuários quando solicitam esse canal irão receber um fluxo MPEG-2 TS com um ou mais fluxos elementares de legenda inseridos e seu STB, por meio do controle do *middleware*, irá poder selecionar entre as legendas ou mesmo escondê-la.

Essa solução é uma sugestão de implementação para canal de TV ao vivo, mas poderia ser adaptada para o caso de VOD e nesse caso o conteúdo VOD seria criado já com a(s) legenda(s), como mostra a Figura 4.2 - Inserção de legenda – VOD. Nesse caso, a legenda

de entrada poderia ser um arquivo texto, já com os tempos de espaçamento definidos, bastando apenas sincronizá-la com o fluxo.

Ainda nesse caso, a estação de gravação irá continuar gerenciando o codificador para modificar alguns parâmetros de compressão de áudio e vídeo, mas a gravação será feita, não do IP *multicast* que sai do codificador, mas do endereço *multicast* que vem do MUX IP/ASI já com a legenda inserida. Essa solução foi implementada e testada no laboratório da Alcatel-Lucent.

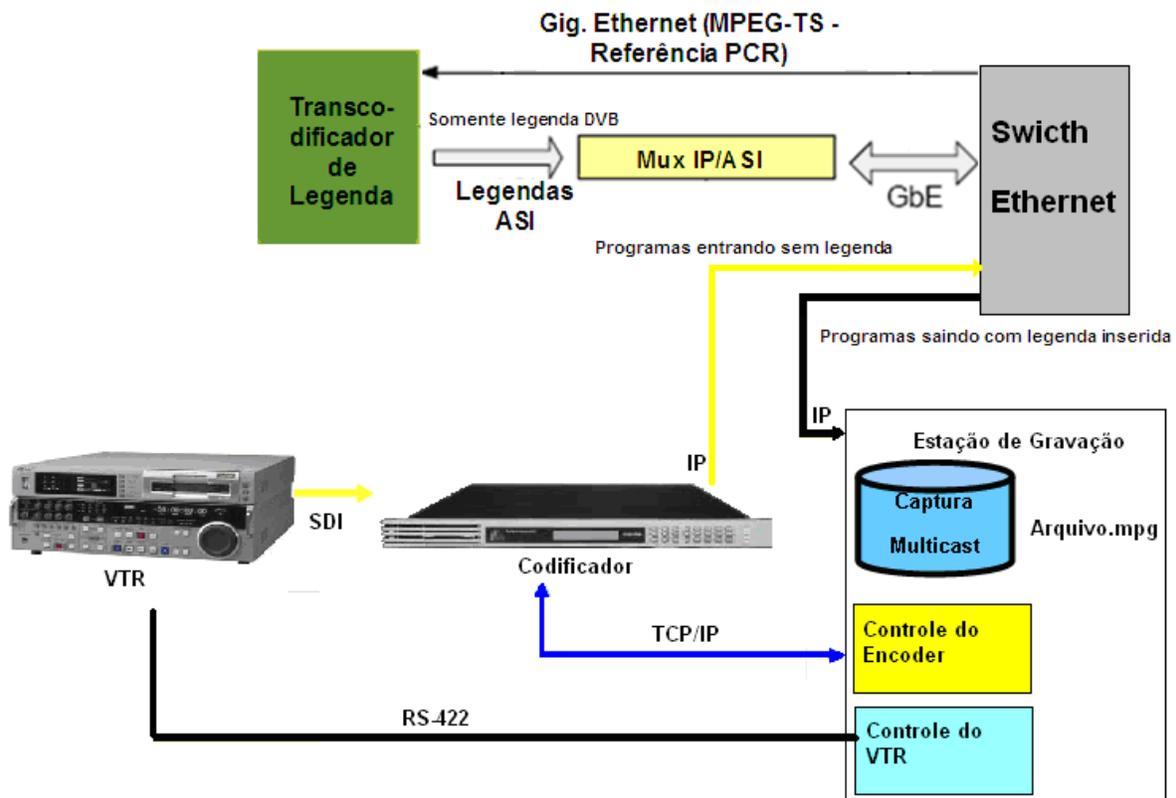


Figura 4.2 - Inserção de legenda – VOD.

Inserção de comercial

Finalmente a última proposta na área do *head-end*, seria para a parte de inserção de comercial. Em geral, as programadoras já mandam os chamados *cue-tones* indicando início de comercial no canal esquerdo de áudio. É um sinal audível que pode ser escutado até mesmo pelo usuário final.

A Figura 4.3 - Diagrama de blocos da inserção de comercial mostra uma proposta para uma possível implementação. Essa solução foi implementada e testada no laboratório da

Alcatel-Lucent. A sinalização enviada é composta, geralmente, de três dígitos DTMF pré-estabelecidos para cada canal. Dessa forma, poder-se-ia utilizar um equipamento de contato aberto/fechado para detectar essa sinalização. Esse equipamento deve fazer uso de chaves seletoras para indicar qual o DTMF esperado.

Suponha então que o DTMF esperado é o 901 para o canal *Cartoon*. Se chegar esse valor pré-ajustado no decodificador de *cue-tones*, ele irá fechar um contato elétrico. Quando ele fechar um contato elétrico, esse equipamento vai gerar um alarme no equipamento seguinte que vai enviar um *trap* SNMP para o sistema de gerência dos *encoders*. Esse sistema, já vai estar previamente configurado e interpretar esse alarme como um DPI (*Digital Program Insertion* ou inserção de programa digital) e inserir uma sinalização no MPEG-2 TS correspondente ao canal *Cartoon*, campo *Splice Countdown* (ver a Figura 2.17 - Cabeçalho do e Tabela 2.5 - Cabeçalho do PES - Campos Opcionais). Esse campo é padronizado pela norma SCTE-35 (*Digital Program Insertion Cueing Message for Cable*).

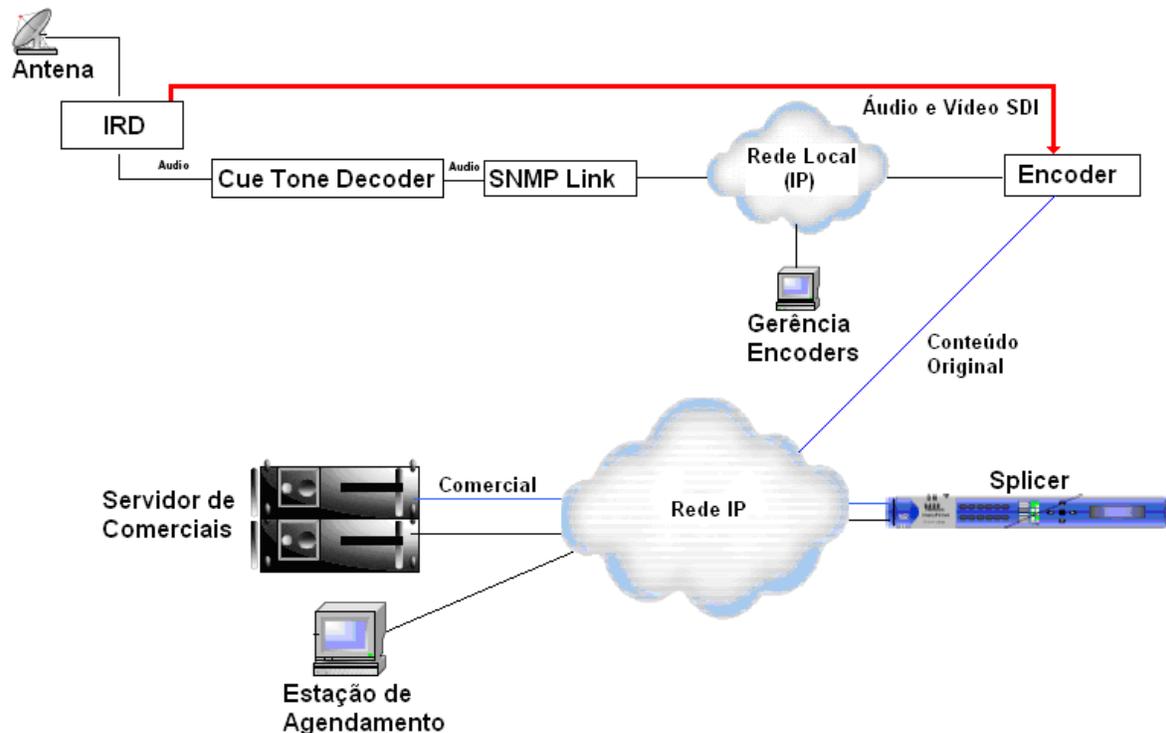


Figura 4.3 - Diagrama de blocos da inserção de comercial.

Isso tudo deve ser feito porque os provedores de conteúdo enviam sinalização analógica. Agora quando o MPEG-2 TS com a mensagem de DPI chegar ao *Splicer*, ele irá substituir o conteúdo ao vivo pelo conteúdo armazenado no servidor de comercial.

Como se pode ver na Figura 4.3 - Diagrama de blocos da inserção de comercial, a rede

IP conterà dois canais, um canal *multicast* depois do codificador e um canal com outro endereço *multicast* após o *splicer*. O canal que será usado pelo *middleware* é o segundo.

Essa programação deve ser feita previamente, como mostra uma tela de exemplo na Figura 4.4 - Agendamento de um comercial, que diz que se uma sinalização chegar pelo canal *Cartoon* das 17h55min às 18h05min, ele deve tocar o comercial LUC_SOCCER2. Essa programação é feita no servidor de comercial que já vai estar preparada para isso, mas será comandada pelo *Splicer*, que irá solicitar um *unicast* do servidor de comercial para ele e inseri-lo no seu *multicast* de saída.

Essa sinalização chega com alguns poucos segundos de antecedência para que a imagem seja cortada no momento devido.

O *splicer* tem uma inteligência tal que vai substituir uma programação em um quadro I e, além disso, se algo ocorrer e se ele não receber o comercial, ele segue enviando para frente o conteúdo que chega da programadora.

Video ID	Status
LUC_SOCCER2	Unchecked

Figura 4.4 - Agendamento de um comercial.

Outro cuidado que deve ser tomado é para que o comercial gravado tenha mais ou menos as mesmas características do canal ao vivo, tais como resolução, formatos de compressão e taxas. Isso tudo para evitar que o STB sofra mudanças abruptas e dessa forma, não gere nenhum tipo de imperfeição na tela do usuário.

Quando o comercial termina ou quando receber uma mensagem de fim de comercial, o *splicer* comuta novamente e envia para sua saída a programação do canal ao vivo.

É importante ressaltar que todas as máquinas devem estar precisamente sincronizadas e, como se está usando uma solução sobre uma rede IP, o protocolo NTP (*Network Time Protocol* ou protocolo de tempo na rede) é mais requisitado.

O NTP é um protocolo para sincronização de relógios de computadores usando a rede e trabalha com o protocolo UDP, porta 123. Aproveita-se para dizer, que não somente na inserção de comercial, mas toda a rede deve estar sincronizada. Em geral, usa-se uma máquina como sendo servidor de NTP para atualizar todas as máquinas. O servidor de NTP por sua vez se atualiza com uma hierarquia superior e mais precisa.

Voltando ao item de comercial e publicidade novamente, esse é um tema tão interessante e lucrativo que os desenvolvedores de comerciais estão também mudando a maneira de como criar sua publicidade com novas formas de prender a atenção dos usuários. Primeiro, por meio de uso de comerciais mais criativos atraindo os mais diversos espectadores. Segundo, fazendo comerciais de menor duração, por exemplo, de 5 segundos. Terceiro, eles estão gerando anúncios de tal maneira que mesmo que o usuário avance a programação ele consegue ver um logotipo ou uma imagem fixa na tela. Alguns estudos americanos revelaram que esses usuários, ao verem uma imagem interessante na tela, retrocedem e assistem ao comercial inteiro.

Middleware

Interface dos Usuários

Outra característica muito importante que uma solução de IPTV deve ter refere-se às interfaces dos usuários. Essas interfaces podem ser customizadas de várias maneiras, mas deve-se tomar um cuidado de desenvolver interfaces ao mesmo tempo intuitivas, fáceis de se usar, com uma excelente visualização gráfica e que não consumam muito processamento do STB.

Serviços

Outra sugestão seria a criação de um serviço de usuário final. Como se sabe, durante a transmissão de uma corrida de fórmula-1, cada carro de corrida tem sua própria câmera, além de várias outras espalhadas pelo circuito. Como o IPTV suporta um número ilimitado de canais, visto que somente um canal é enviado por vez para a casa do assinante, poderia ser criado um pacote de TV para F1 e o usuário final escolher qual câmera ele quer assistir. Isso exigiria apenas alguns contratos comerciais e a disponibilização das imagens de todas as câmeras para o provedor de serviço. Esse serviço pode ser extrapolado para vários tipos

de esportes ou programação.

Redundância

O *middleware*, como é o centro da solução de IPTV, também deve ser redundante para evitar que uma parada em qualquer serviço ou servidor, comprometa toda a solução. Os servidores de *web caches*, os servidores de aplicação e os servidores de base de dados, como mostrados na Figura 4.5 - Servidores do *middleware* em alta disponibilidade.

Os servidores de *web cache* poderiam ainda ser usados como balanceadores de carga e isso traria várias vantagens. Primeiro, porque aumentaria a disponibilidade, uma vez que os dados estariam balanceados e mesmo que ocorresse uma falha ou manutenção de um servidor de aplicação, os outros ASs conseguiriam atender às solicitações. Segundo, esses servidores "escondem" os servidores internos dos usuários, minimizando assim os ataques de segurança.

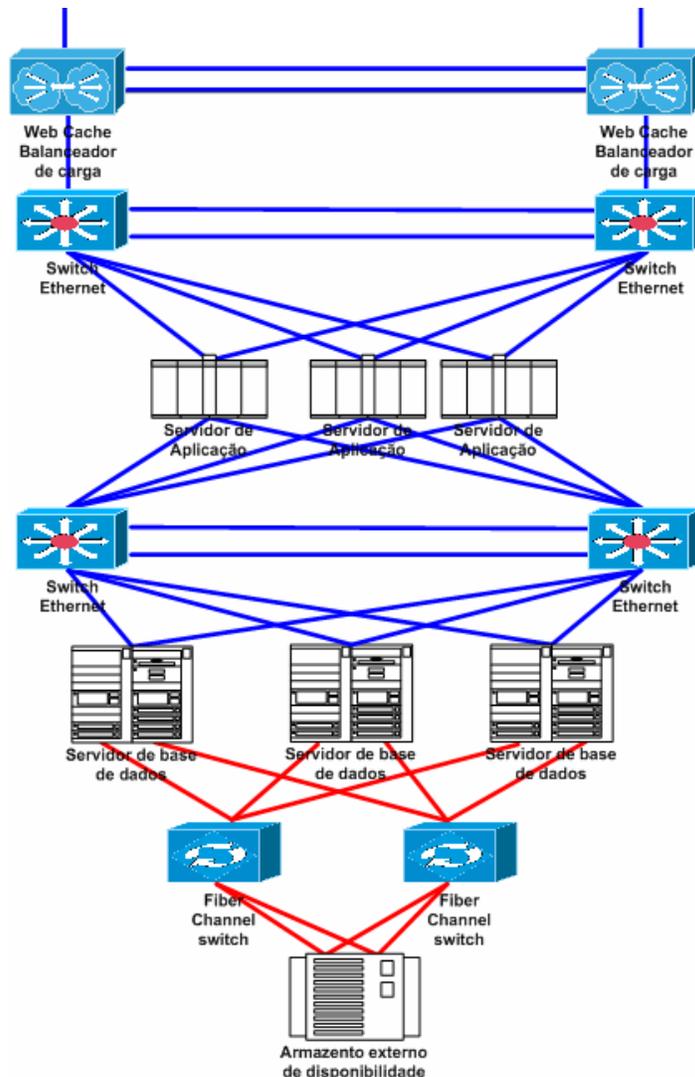


Figura 4.5 - Servidores do *middleware* em alta disponibilidade.

Rede

A rede de transporte pode ser de vários tipos, sendo as únicas premissas que essa rede deva ter uma excelente qualidade de serviço, suportar o *multicast*, ter uma baixa latência e, obviamente suportar uma largura de banda adicional para o *unicast*.

Quanto à rede de acesso, a operadora deve fazer um estudo para tentar viabilizar o uso de FTTH, o que possibilitaria canais e conteúdos de alta-definição.

Mais uma vez e, tão importante quanto, a rede também deve ter proteção de redundância. Os roteadores e *switches ethernet* devem ter, não somente, equipamentos redundantes, como também caminhos redundantes.

Mudança Rápida de Canal

Como foi comentado o tempo mudança de canal (*zapping time*) é um fator bastante crítico no sistema de IPTV e que esse tempo depende de vários parâmetros: tempo de processamento dos comandos, tempo de atraso de rede, tempo de atraso pelo STB, tempo de armazenamento do *buffer* do STB e tempo de decodificação de vídeo. É possível otimizar vários desses parâmetros com uma dificuldade maior para o tempo de decodificação MPEG [4, 24, 43]. Além disso, o codificador também possui um *buffer* que simula o *buffer* do STB e é usado para que o envio de informação seja feito com uma taxa de bits constante (CBR). O tempo total pode ser da ordem de 2 a 3 segundos, similar a TV por satélite.

Sabe-se que o decodificador começa lendo um fluxo de transporte e procura pelo PID 0x0 que corresponde à tabela PAT. A tabela PAT irá conter o endereço PID da tabela PMT.

A seguir, o STB vai de-multiplexar os pacotes que transportam a PMT desejada e lê-la. A tabela PMT irá conter os PIDs dos fluxos elementares (vídeo, áudios, legendas), bem como o formato usado para codificá-los. Baseado nos valores dos PIDs definidos na PMT, o decodificador começa a de-multiplexar os fluxos elementares correspondentes. Após isso, o STB deve encher o seu *buffer* para começar a decodificação. Esse processo todo pode levar 2, 3 ou até 4 segundos, o que é um tempo muito alto. Primeiro, porque todo STB necessita de um *buffer* para ser capaz de reconstruir o vídeo dos quadros Is, Ps e Bs que ele recebe. O decodificador não pode reconstruir um quadro B até que ele receba o próximo quadro. Segundo, quando um usuário muda de canal, é muito grande a chance de ele receber um fluxo no meio de um GOP, significando que o quadro I já passou. Uma vez que ele não tem o quadro I do GOP, ele não pode reconstruir os quadros Bs e Ps e o STB não irá mostrar nada na tela. Isso implica que o usuário terá que esperar o próximo quadro I. Esse tempo depende do tamanho do GOP e para um GOP de 60 quadros a uma taxa de 30 quadros por segundo, isso pode levar 2 segundos. O operador pode diminuir esse tempo, diminuindo o tamanho do GOP, mas isso pode aumentar a taxa de transmissão.

Propõe-se um método que poderia reduzir esse tempo. Propõe-se a inclusão de um novo servidor que teria acesso a todos os canais *multicast* e que vai fazer um encapsulamento RTP (*Real Time Protocol* ou protocolo de tempo real) em todos os pacotes. Chama-se aqui esse servidor de servidor Y, com mostrado na Figura 4.6 - Mudança rápida de canal. Estima-se que essa redução poderia ser de 30 a 40%.

O RTP vai acrescentar funções de transporte necessárias para prover sincronização e número de seqüência dos pacotes. O RTP utiliza o transporte UDP/IP e é adequado para aplicações em tempo real.

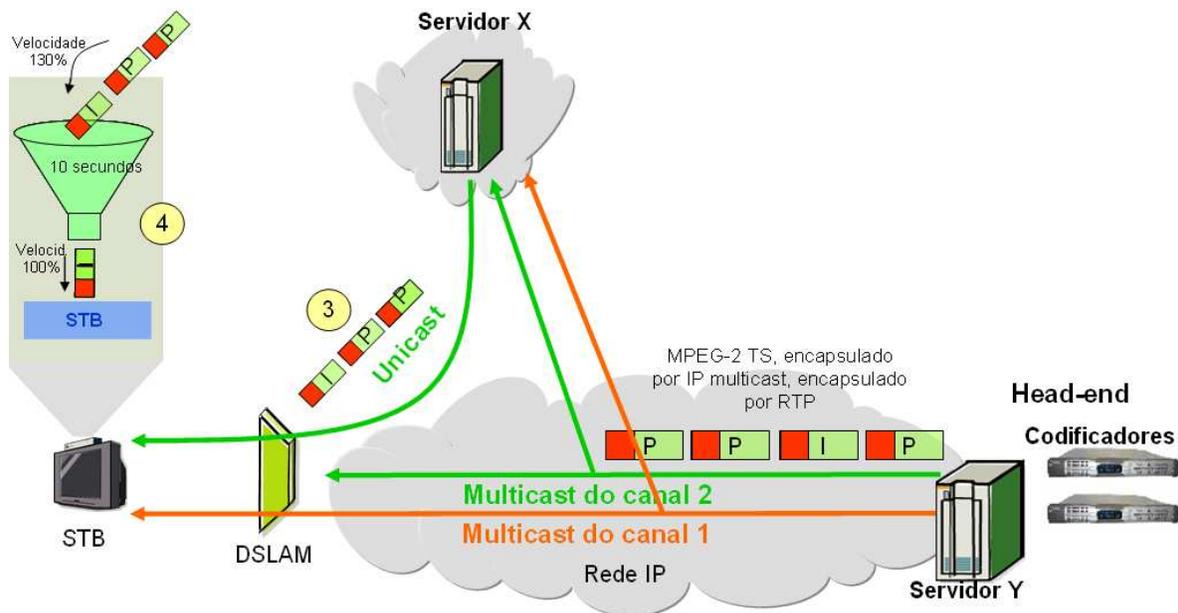


Figura 4.6 - Mudança rápida de canal.

Um novo servidor (servidor X), também irá receber todos os fluxos *multicast* do servidor Y e gravá-los em um *buffer* circular para cada canal de, por exemplo, 10 segundos, como mostra a Figura 4.7 - *Buffer* circular, que significa que uma vez que o *buffer* se encher, os quadros antigos são descartados. Chama-se aqui esse servidor de servidor X.

Dessa forma, quando o STB muda de canal, essa máquina poderia ser envolvida no processo transitório e reduzir esse tempo.

O processo seria feito da seguinte forma, como mostrado na Figura 4.6 - Mudança rápida de canal.

O STB está assistindo o canal 1 via *multicast*, diretamente dos codificadores. Depois disso, ele muda para o canal 2.

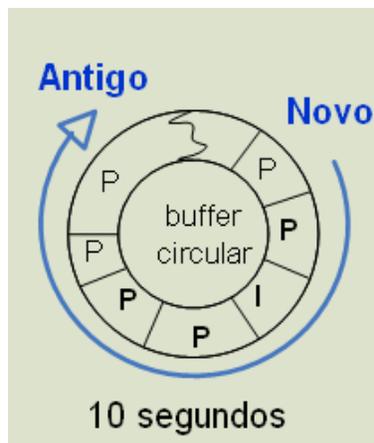


Figura 4.7 - Buffer circular.

Ao mesmo tempo o servidor X está recebendo o canal 2 via *multicast* do *head-end* e os armazena no seu *buffer* circular. Vamos supor um *buffer* de 10 segundos de tamanho e os quadros Bs não são mostrados apenas para simplificar a figura.

- 1- Quando o usuário muda para canal 2, ele solicita ao DSLAM para sair do canal 1 usando o *IGMP fast leave*. O *IGMP fast leave* é uma modificação do protocolo IGMP que corta a transmissão do canal mais rápido, eliminando uma mensagem de reconhecimento que é normalmente usada;
- 2- O STB solicita ao servidor X via HTTP que ele quer receber o canal 2;
- 3- O servidor X inicia a transmissão dos quadros do canal 2 que estão armazenados no *buffer* circular para o STB via *unicast* a uma velocidade mais rápida (por exemplo, 30%) do que o tempo real. O servidor X começa a transmissão desde um quadro I para começar a mostrar o vídeo imediatamente;
- 4- Essa transmissão mais rápida chega ao buffer do STB, representado por um funil de 10 segundos. Esse *buffer*, por sua vez, envia os pacotes numa velocidade normal de 100% e o cliente não vai ter uma imagem acelerada, mas com uma velocidade normal. Uma vez que o primeiro quadro recebido no decodificador é um quadro I, o usuário pode começar a ver o canal 2 imediatamente;
- 5- Enquanto o STB está assistindo ao canal 2 do servidor X, via *unicast*, ele envia um comando *IGMP join* do canal 2 ao DSLAM. Esse por sua vez inicia o envio do canal 2 *multicast* para o usuário. O STB continua a encher o buffer, mas agora com quadros do canal 2, via *multicast*;
- 6- O STB comuta de forma transparente do *unicast* para o *multicast*. Pode existir uma sobreposição de quadros no STB, mas o cliente do STB vai verificar e descartar os quadros desnecessários. Assim, o *unicast* não é mais necessário e o servidor X pára de mandar quadros do canal 2.

Página em branco

Capítulo 5

Conclusões

Este trabalho baseou-se no estudo de um sistema de IPTV, cobrindo todos os aspectos da solução, desde o *head-end*, o *middleware*, o CAS/DRM, a rede de transporte e acesso, qualidade de serviço, incluindo as tecnologias envolvidas, como *multicast*, RTSP, tecnologias de compressão de áudio e vídeo, MPEG-2 TS e até o ADSL2+.

Além disso, foi proposta uma forma de se reduzir o tempo de mudança de canal. Propôs-se também alguns serviços finais que poderiam ser criados e a implementação de um modelo para inserção de legenda e um modelo para inserção de publicidade.

Um sistema de IPTV bem arquitetado oferecerá serviços de TV digital que as pessoas estão demandando hoje, como HDTV, serviços de gravação, serviços sob demanda e aplicações interativas. Além disso, uma vez que a conexão de banda larga utiliza os dois sentidos de transmissão simultaneamente, o IPTV irá habilitar novas experiências que vão mudar drasticamente a maneira de como as pessoas assistem a uma TV.

Outros desafios para um sistema de IPTV resultam da capacidade de se oferecer três ou mais pontos de conexão na casa do usuário. Isso requer muito mais largura de banda e, atualmente a rede de acesso DSL não pode oferecer essa facilidade. Esse aspecto é diferente da TV a cabo em que todos os canais chegam por meio de cabo coaxial na casa do usuário e com o uso de um divisor de frequência pode-se ter quantos pontos quanto se queira na residência.

A indústria da TV está migrando para uma nova geração, que é o uso da TV em alta definição (HDTV). O sistema HDTV com sua alta resolução requer muito mais largura de banda que o DSL atual é capaz de oferecer. Felizmente, a tecnologia do DSL também está evoluindo e já se pode encontrar no ramo da pesquisa e desenvolvimento o quase pronto padrão VDSL2, que possibilitará velocidades reais na ordem de 20 a 50 Mbps. Na atualização das redes de acesso legalizadas, o VDSL2 pode vir a ser uma tecnologia muito usada com custos relativamente baixos, além de grande largura de banda e confiabilidade.

Nem todas as companhias de telefone ficarão limitadas à velocidade dos fios de cobre.

Muitas delas já começaram a atualizar a rede de cobre, substituindo-a pela fibra óptica até a casa do assinante (FTTH).

Devido ao fato do IPTV ser bi-direcional, é possível inclusive criar comerciais interativos. Por exemplo, durante o comercial de um carro, se um usuário apertar, por exemplo, um botão vermelho, ele poderia ver quantos quilômetros por litro faz o carro, ou seu desempenho, ou seu conforto.

Essa interatividade tende a crescer na medida em que o serviço é lançado e isso vai propiciar serviços muito imaginativos.

O sistema de IPTV ainda não foi regulamentado no Brasil e as operadoras poderiam comercializar apenas o VOD, usando a norma SCM (Sistema de Comunicação Multimídia). Entretanto, enquanto os órgãos brasileiros não derem uma concessão de TV para as operadoras de telefonia, dificilmente elas conseguiriam emplacar seus serviços massivamente. Já existe uma pressão muito grande junto a ANATEL, mas enquanto isso o que poderia ser feito é o uso de STB híbridos, que funcionariam com o sistema de TV a cabo e com o sistema de IPTV ao mesmo tempo. Devido a isso, já se pode perceber algumas aquisições de empresas de TV a cabo ou satélite por operadoras de telefonia no mercado.

Finalmente, IPTV tem a sustentação para inovadoras aplicações como a arquitetura IMS (*IP Multimedia Subsystem* ou subsistemas multimídia IP). O IMS é um elemento chave na arquitetura 3G e torna possível o uso das aplicações IPTV em qualquer outro aparelho, como um celular ou PDA. O IMS é uma solução bem mais ampla e complexa que permitirá oferecer qualquer serviço para qualquer dispositivo. Dessa forma, o IPTV passa a ser uma das diversas aplicações do IMS, como por exemplo, o uso de vídeo no celular.

Sugestões para Trabalhos Futuros

O estudo aponta questões importantes a serem investigadas no prosseguimento deste trabalho. Apresentam-se a seguir possíveis extensões que podem resultar em novas linhas de pesquisa, como por exemplo:

- Desenvolver um *software* para simular o tempo de mudança de canal;
- Estudar o uso da tecnologia VDSL2 em ambientes com interferência ou longas distâncias;
- Verificar a viabilidade do uso da tecnologia WiMAX como acesso ao IPTV, o que propiciaria o uso dos serviços em localidades remotas ou rurais;
- Estudar a proposição e implementação de solução de IMS e verificar como que o IPTV poderia ser integrado a ela, propiciando todos os seus serviços no celular;
- Criar um simulador para o DRM e verificar como torná-lo cada vez mais seguro.

Como se sabe, o DRM vai ser sempre uma preocupação para os provedores de conteúdo e a idéia aqui seria uma maneira de estar sempre utilizando um DRM mais forte possível;

- Estudar a escalabilidade de uma solução de IPTV e verificar como se comportaria o sistema para o caso de um milhão de assinantes;
- Elaborar métodos para criar conteúdos codificados em H.264 que mesmo em modo de avanço rápido se possa perceber uma imagem estática na tela.

Página em branco

Apêndices

Apêndice A: Configuração do *Head-end*

A.1 Codificação de Vídeo

Para melhorar a qualidade da compressão, pode-se alterar parâmetros avançados de codificação de vídeo. Essa configuração é feita no software de gerência dos codificadores e é suportada por quase todos os fabricantes. Exemplos dessas configurações podem ser vistos nas Figura A.1 - Parâmetros básicos de codificação de vídeo e Figura A.2 - Parâmetros avançados de codificação de vídeo, que são telas capturadas do sistema de gerência dos codificadores da Harmonic.

Na Figura A.1 - Parâmetros básicos de codificação de vídeo podem ser alterados o modo e o padrão de compressão de vídeo. A maioria dos codificadores suporta simultaneamente o MPEG-2, o H.264 e o VC-1. Nesse caso, o provedor de serviço precisa definir um padrão para configurar o codificador corretamente. Pode-se ainda alterar o número de linhas de resolução horizontal. A configuração da resolução vertical, definição padrão convencional ou alta definição, requer, em geral, mudança no *hardware* do codificador. Finalmente nessa tela, deve ser definida a largura de banda para o fluxo elementar de vídeo.

Já na Figura A.2 - Parâmetros avançados de codificação de vídeo podem ser alterados o tamanho do GOP, o formato da tela 4:3 ou 16:9, bem como o modo de varredura.

Foram efetuados vários testes no laboratório da Alcatel-Lucent para se encontrar a melhor resolução e taxa. Foram usadas as seguintes combinações de resolução 720x480, 704x480, 544x480 e 480x480. Os resultados variam muito com a versão de software dos codificadores e dos decodificadores e por essas razões os resultados não foram publicados.



Figura A.1 - Parâmetros básicos de codificação de vídeo (Codificador da Harmonic).

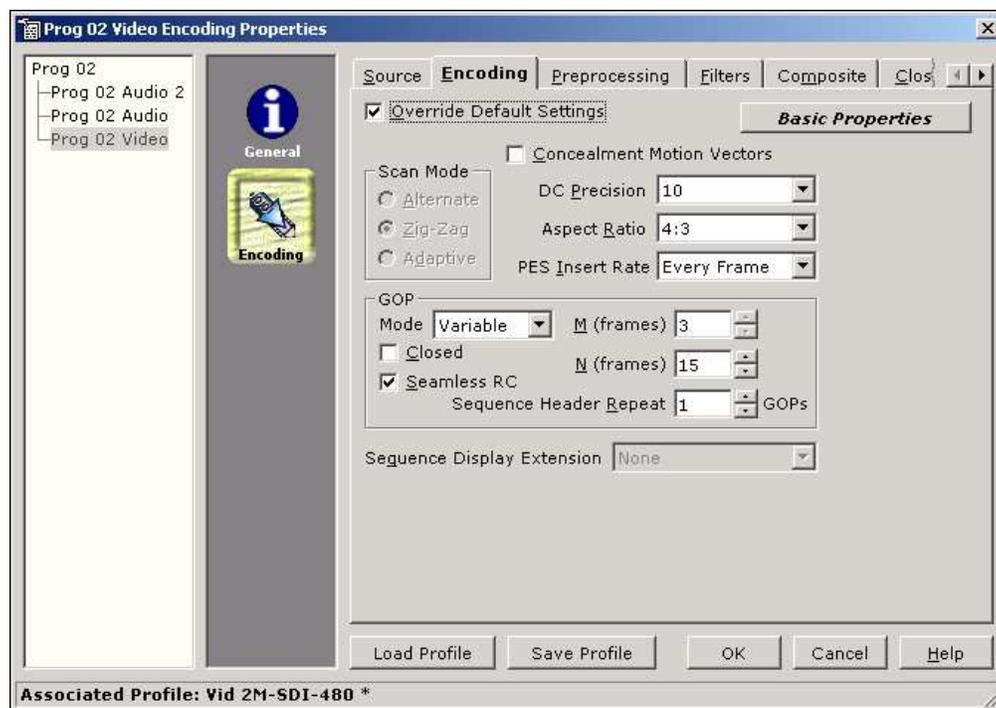


Figura A.2 - Parâmetros avançados de codificação de vídeo (Codificador da Harmonic).

A.2 Codificação de Áudio

Da mesma forma, os parâmetros de compressão do áudio, bem como o formato adotado devem ser configurados no mesmo software de gerência, como mostra Figura A.3 - Parâmetros de codificação de áudio e Figura A.4 - Parâmetros de codificação de áudio MPEG L2 (Codificador da Harmonic).. De acordo com o padrão de compressão escolhido na Figura A.3 - Parâmetros de codificação de áudio (Codificador da Harmonic).., o nome e as opções da última aba serão alterados como mostrado na Figura A.4 - Parâmetros de codificação de áudio MPEG L2 (Codificador da Harmonic)..

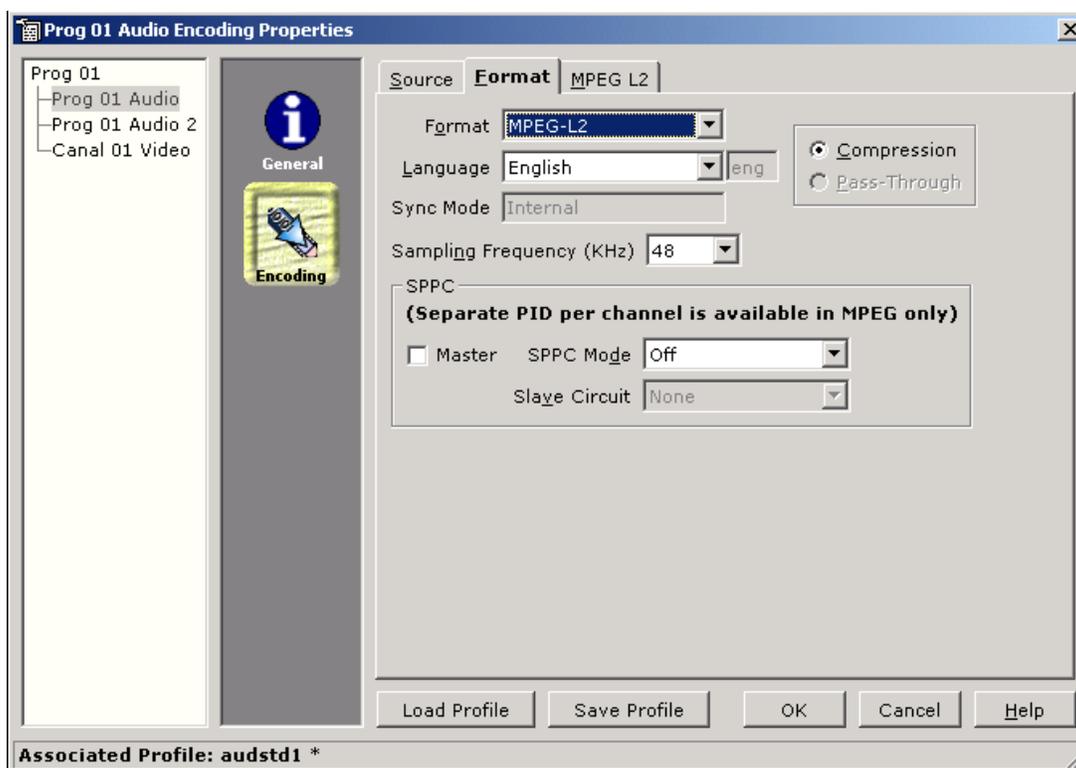


Figura A.3 - Parâmetros de codificação de áudio (Codificador da Harmonic).

Na Figura A.4 - Parâmetros de codificação de áudio MPEG L2 (Codificador da Harmonic). são mostrados os parâmetros de largura de banda para o fluxo elementar de áudio, o modo de codificação adotado, bem como, a possibilidade de se inserir a referência PCR no fluxo de áudio. Como explicado no capítulo 2, o PCR é, em geral, inserido no fluxo elementar de vídeo. Entretanto, para canais de áudio (rádio), o PCR deve ser inserido,

uma vez que o fluxo de transporte não conterá fluxo de vídeo.

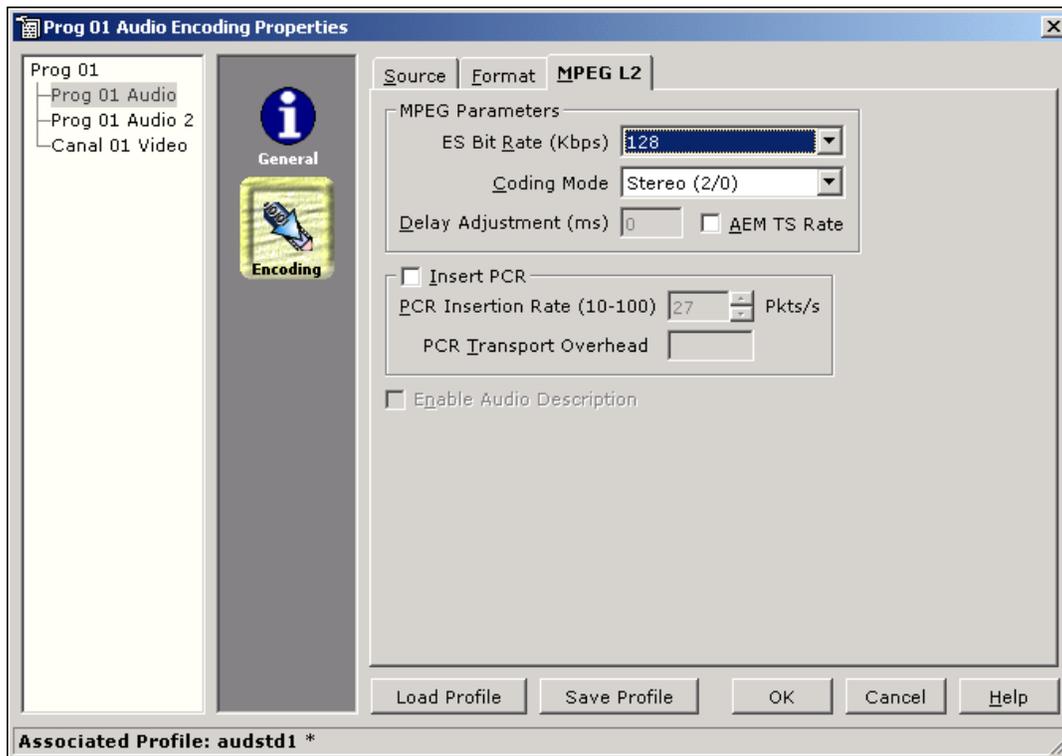


Figura A.4 - Parâmetros de codificação de áudio MPEG L2 (Codificador da Harmonic).

A.3 Codificação do Fluxo de Transporte

Finalmente, devem-se configurar os parâmetros do fluxo de transporte, como mostrado nas Figura A.5 - Parâmetros gerais de fluxo de transporte MPEG-2 TS (Codificador da Harmonic). e Figura A.6 - Parâmetros IP de fluxo de transporte MPEG-2 TS (Codificador da Harmonic)..

Um parâmetro importante, mostrado na Figura A.5 - Parâmetros gerais de fluxo de transporte MPEG-2 TS (Codificador da Harmonic)., é a largura de banda para todo o fluxo de transporte, que deve ser maior do que a soma de todos os fluxos elementares, incluindo as tabelas PSI.

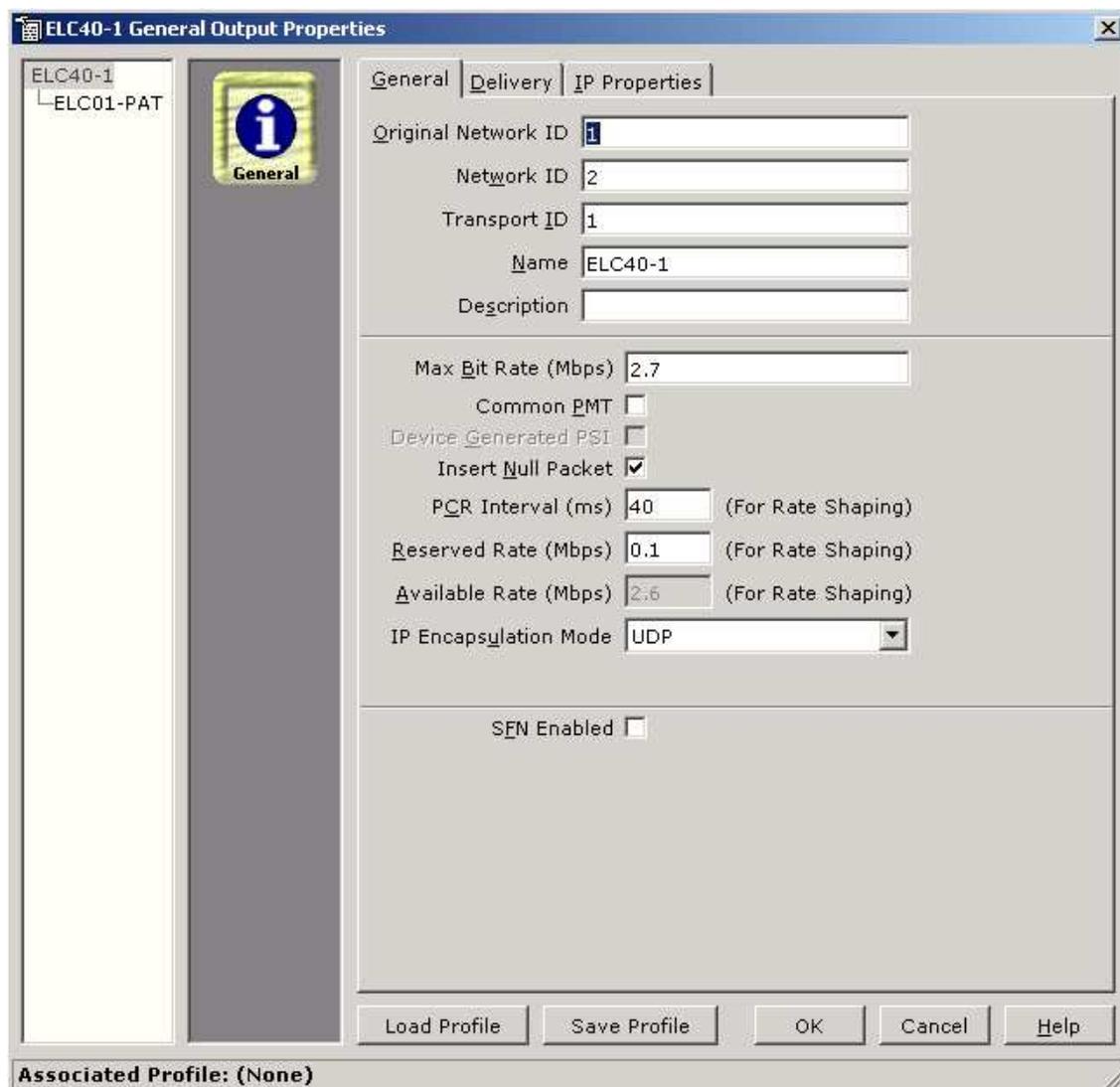


Figura A.5 - Parâmetros gerais de fluxo de transporte MPEG-2 TS (Codificador da Harmonic).

Finalmente na Figura A.6 - Parâmetros IP de fluxo de transporte MPEG-2 TS (Codificador da Harmonic), configuram-se os valores de endereço IP de destino ou endereço IP *multicast* (*Destination IP Address*) e a porta usada para transmissão. Esses dois valores representam um canal e serão usados posteriormente pelo *middleware*.

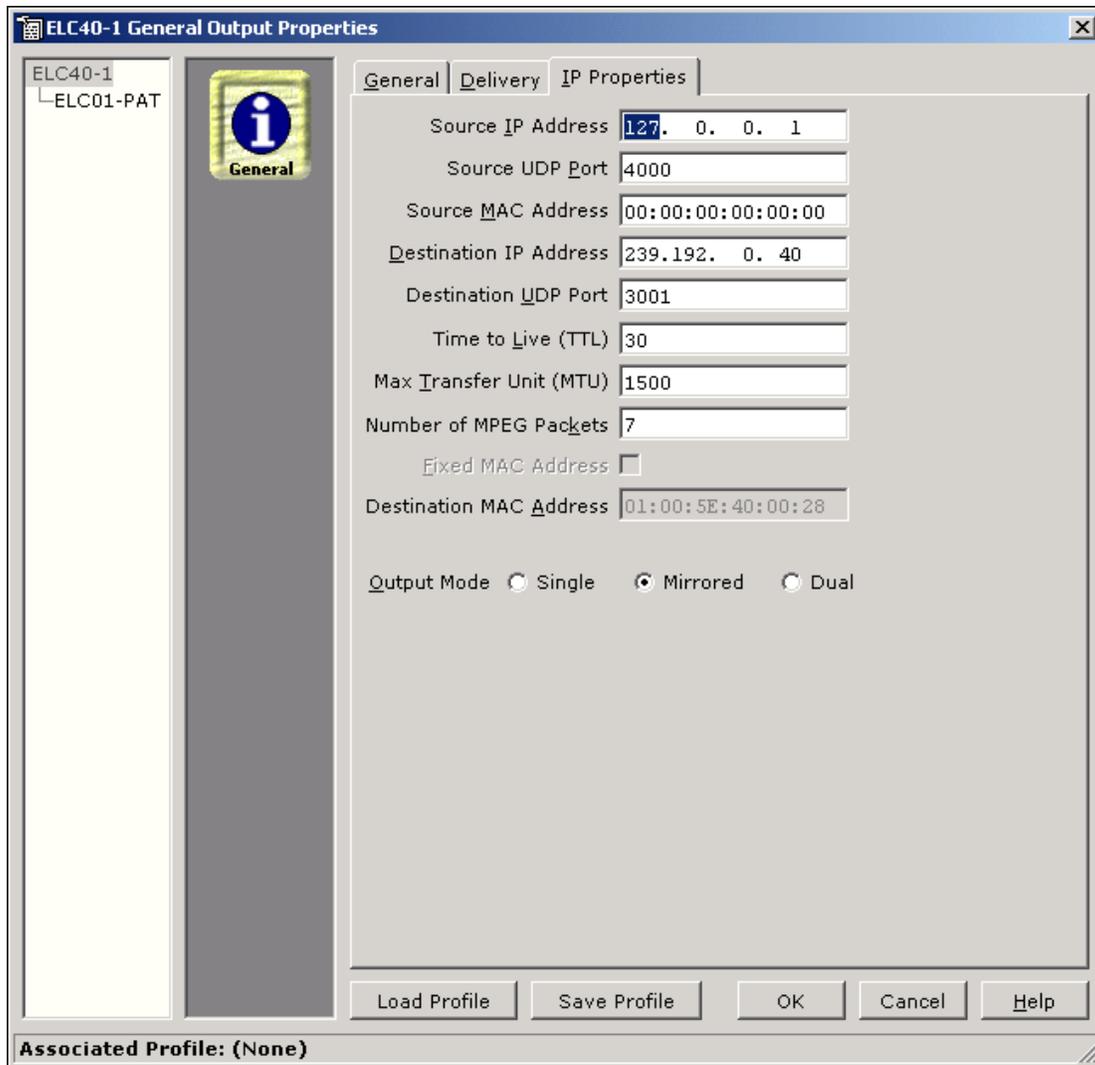


Figura A.6 - Parâmetros IP de fluxo de transporte MPEG-2 TS (Codificador da Harmonic).

A Figura A.7 - Lista de fluxos (Codificador da Harmonic). mostra todo o conteúdo encapsulado pelo IP *multicast*. Nesse caso, tem-se um fluxo elementar de vídeo, além de dois fluxos elementares de áudio e das tabelas PAT e PMT, que compõem o MPEG-2 TS.

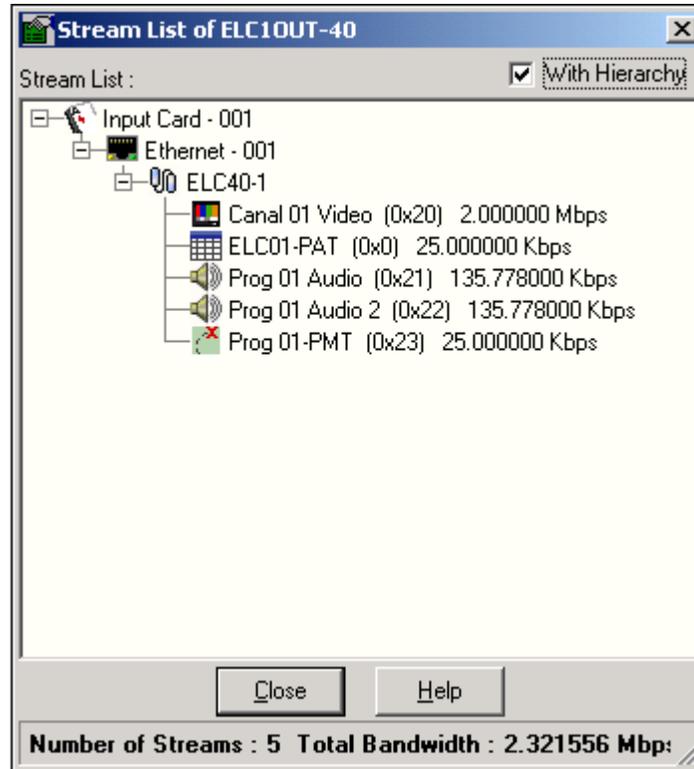


Figura A.7 - Lista de fluxos (Codificador da Harmonic).

A.4 Criptografia de Canais em Tempo Real

Ainda no *head-end* e logo após os codificadores, podem residir os equipamentos de criptografia. Em geral, esses equipamentos têm duas interfaces de rede, uma interface para a rede sem criptografia e outra para a rede protegida. Essa primeira interface fica no modo promíscuo (*promiscuous mode*). Esse modo refere-se à configuração da placa de rede que é configurada para receber todo o tráfego da rede e não somente aquele endereçado a ela. Esse modo é muito conhecido pelos *hackers* e é bastante usado por *softwares* que fazem captura de dados (*sniffer*) da rede.

Sendo assim, se a primeira interface está no modo promíscuo, ela vai receber todos os canais *multicast* daquela rede, como mostrado na Figura A.8 - Criptografia de canais (Sistema de CAS/DRM da Widevine).. O operador pode ainda definir quais canais serão criptografados. Em um caso real, em geral, todos os canais são criptografados. Após a criptografia, o equipamento envia os dados, agora criptografados para uma outra interface de rede, que vai "alimentar" a rede de transporte. Nesse caso, nem mesmo os funcionários do provedor de serviço terão acesso aos canais sem criptografia.

EIM IP: 200.10.10.12

Server ID: 1 Bandwidth: 41.384 [Mbps]

Service Manager

- List Services
- Logout
- Home

<input type="checkbox"/>	Network Address	Disposition	Bandwidth [Mbps]	Current Service	Network Resource	Server ID
<input type="checkbox"/>	224.10.10.16:3016	Clear	2.379	128069	128069	1
<input type="checkbox"/>	224.10.10.14:3014	Encrypted	2.309	128065	128065	1
<input type="checkbox"/>	224.10.10.15:3015	Clear	2.948	128067	128067	1
<input type="checkbox"/>	224.10.10.17:3017	Clear	2.914	128060	128060	1
<input type="checkbox"/>	224.10.10.13:3013	Clear	1.693	128244	128244	1
<input type="checkbox"/>	239.192.1.50:2050	Encrypted_Fixed_Key	2.321	Unset	Unset	1
<input type="checkbox"/>	224.10.10.18:3018	Clear	2.835	128246	128246	1
<input type="checkbox"/>	239.192.1.51:2051	Encrypted Fixed Key	2.321	Unset	Unset	1

Local intranet

Figura A.8 - Criptografia de canais (Sistema de CAS/DRM da Widevine).

Apêndice B: Configuração de VOD

B.1 Criação de Conteúdo VOD

Como explicado no capítulo 3, um conteúdo VOD contém um arquivo digital, que é o filme, um arquivo de trailer, os arquivos para avanço e retrocesso e os metadados.

Na Figura B.1, os campos dos metadados são mostrados. Esses parâmetros incluem o título, a descrição, a categoria, a duração, o gênero, os atores, os idiomas e as datas de ativação e desativação.

The screenshot shows a web browser window titled 'Administración de Cbd - Microsoft Internet Explorer' with the address 'http://172.26.46.10:81/principal.aspx'. The main content area is a form for updating a product. The form is titled 'ACTUALIZAR PRODUCTO' and has a dropdown menu set to 'Película' and an 'ID. Prod: 1000067'. The form is divided into several sections:

- DATOS DEL PRODUCTO:** Includes fields for 'Título' (Tesouro), 'Descripción' (National Treasure), 'Categoría' (Estrenos), 'Estado de publicación' (Publicado), 'Estado STI' (Validado), 'Validación Automática' (checked), 'Estado' (Activa, Inactiva, Borrable), and 'Precio' (2,39).
- Periodo de comercialización:** Includes 'Inicio' (10/23/2006), 'Fin' (11/30/2007), 'Borrado' (11/30/2008), and 'Periodo validez (horas)' (72).
- FICHA DEL PRODUCTO:** Includes 'Duración (minutos)' (45), 'Año Producción' (2005), 'Información DVD' (checked), 'Género' (Aventuras), 'Clasificación por edad' (No recomendada para menores de 13 años), 'Macrovisión' (unchecked), 'Productora', 'Directores', 'Actores principales', 'Actores secundarios', 'Nacionalidad', 'Premios', 'Idiomas' (Inglés (Estéreo), Portugués (Estéreo)), and 'Subtítulos' (Portugués).
- URLs:** A table with columns 'U.R.L.' and 'Descripción'.

On the left side, there is a tree view showing a directory structure under 'Videoclub', including folders like 'Estrenos', 'Adultos', and 'Ocio y Cultura', and files like 'tesouro.jpg' and 'tesouro.mpg'.

Figura B.1 - Criação de metadados (Middleware da Alcatel-Lucent).

Página em branco

Referências Bibliográficas

- [1] Kim, J. Y.; Hahm, J. H.; Kim, Y. S.; Choi, J. K.; “NGN Architecture for IPTV Service without Effect on Conversational Services”, Advanced Communication Technology, 2006. ICACT 2006. The 8th International Conference Volume 1, Page(s): 465 - 469, 20-22, Feb. 2006.
- [2] ITU-T Recommendation H.610, "Full service VDSL – System architecture and customer premise equipment", July 2003.
- [3] Wan-Ki Park; Chang-Sik Choi; Dae-Young Kim; Younbg-Kwac Jeong; Kwang-Ro Park “IPTV-aware Multi-Service Home Gateway based on FTTH Access Network” Consumer Electronics, 2005. (ISCE 2005). Proceedings of the Ninth International Symposium, Page(s): 285 – 290, 14-16, June 2005.
- [4] Chunglae Cho, Intak Han, Yongil Jun and Hyeongho Lee, “Improvement of Channel Zapping Time in IPTV Services Using the Adjacent Group Join-Leave Method”, Proc. of Advanced Communication Technology 2004, pp97 1-975, August 2004.
- [5] Wu D., Hou Y., Zhu W., Zhang U-Q., and' Peha J., "Streaming Video over the Internet: Approaches and Directions", IEEE Transactions on Circuits and Systems for Video Technology, Vol. II, No. 3, March 2001.
- [6] Alesso, Peter H., "Producing Internet Video as Broadband Technologies Converge", Addison Wesley, 2000.
- [7] ISO/IEC 13818-1, Generic coding of moving pictures and associated audio information: Systems, 2000.
- [8] ISO/IEC 13818-2, Generic coding of moving pictures and associated audio information: Video, 2000.
- [9] ISO/IEC 13818-3, Generic coding of moving pictures and associated audio information: systems, 2000.
- [10] ISO/IEC 14496-10, Advanced Video Coding, 2004.
- [11] Sangjae Lee; Chunglae Cho; Intark Han; “QoS-adjustable home gateway for IPTV service”, Consumer Electronics, 2005. ICCE. 2005 Digest of Technical Papers. International Conference, Page(s):395 – 396, 8-12, Jan. 2005.
- [12] Ulf Jennehag; Tingting Zhang; “Increasing bandwidth utilization in next generation

- IPTV networks” Image Processing, 2004. ICIP '04. 2004 International Conference on Image processing, Page(s):2075 - 2078 Vol. 3, 24-27, Oct. 2004.
- [13] Alfonsi, B.; “I want my IPTV: Internet Protocol television predicted a winner” Distributed Systems Online, IEEE Volume 6, Issue 2, Feb. 2005.
- [14] Cherry, S.; “The battle for broadband [Internet protocol television]” Spectrum, IEEE, Volume 42, Issue 1, Page(s):24 – 29, Jan. 2005.
- [15] Stienstra, A.J.; “Technologies for DVB services on the Internet”. Proceedings of the IEEE Volume 94, Issue 1, Page(s):228 – 236, Jan. 2006.
- [16] Intark Han; Hong-Shik Park; Youn-Kwae Jeong; Kwang-Roh Park; “An integrated home-server for communication, broadcast reception, and home automation”, Consumer Electronics, IEEE Transactions on Volume 52, Issue 1, Page(s):104 – 109, Feb. 2006.
- [17] Jain, R.; “I Want My IPTV”, Multimedia, IEEE, Volume 12, Issue 3, Page(s): 96 – 96, July-Sept. 2005.
- [18] Cruz-Neira, C.; Lindahl, G.; ”A voyage into virtual reality. Networking our VR lab to Iowa middle schools and high schools”, Computer Graphics and Applications, IEEE, Volume 20, Issue 2, Page(s):16 – 19, March-April 2000.
- [19] Walko, J.; “I love my IPTV”, Communications Engineer, Volume 3, Issue 6, Page(s): 16 – 19, Dec./Jan. 2005/2006.
- [20] Sangjae Lee; Jieun Kim; Youngae Jeon; Sangshung Choi; “Residential HDTV distribution system using UWB and IEEE 1394”, Consumer Electronics, 2006. ICCE '06. 2006 Digest of Technical Papers. International Conference, Page(s): 41 – 42, 7-11 Jan. 2006.
- [21] Faure, J.-P.; “The IEEE P1901 project: broadband over power lines”, Consumer Electronics, 2006. ICCE '06. 2006 Digest of Technical Papers. International Conference, Page(s): 159 – 160, 7-11 Jan. 2006.
- [22] Sangjae Lee; Chunglae Cho; Intark Han; “FTTH residential gateway And IP tuner for IPTV service”, Consumer Communications and Networking Conference, 2006. CCNC 2006. 2006 3rd IEEE, Volume 1, Page(s):497 – 501, 8-10 Jan. 2006.
- [23] Placci, F.M.; “Service delivery architecture for premium IPTV and triple play”, Design of Reliable Communication Networks, 2005. (DRCN 2005). Proceedings. 5th International Workshop, Page(s):331 – 331, 16-19 Oct. 2005.
- [24] Chunglae Cho; Intak Han; Yongil Jun; Hyeongho Lee; “Improvement of channel zapping time in IPTV services using the adjacent groups join-leave method”, Advanced Communication Technology, 2004. The 6th International Conference on Volume 2, Page(s):971 – 975, August 2004.

- [25] Siebert, P.; "Implementation of H.264/MPEG-4 AVC in low cost set top boxes", Consumer Electronics, 2005. (ISCE 2005). Proceedings of the Ninth International Symposium, Page(s):310 – 314, 14-16, June 2005.
- [26] <http://www.lucent.com>, 2005-2007
- [27] K. Brandenburg and H. Popp, An introduction to mpeg layer-3, tech. report, Fraunhofer Institut fr Integrierte Schaltungen (IIS), Erlangen, Germany, October 2001.
- [28] Fraunhofer Institute for Integrated Circuits FhG-IIS A, MP3 And AAC Explained, AES 17th International Conference on High Quality Audio Coding, Erlangen, Germany, August 1999.
- [29] ISO/IEC International Standard IS 11172-3 "Information Technology - Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mb/s - Part 3: Audio", 1993.
- [30] K.Brandenburg, G.Stoll, "The ISO/MPEG-Audio Codec: A Generic Standard for Coding of High Quality Digital Audio", 92nd AES-Convention, preprint 3336, Vienna, 1992.
- [31] ISO/IEC, "Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s - Part 3: Audio", ISO/IEC 11172-3, 1993.
- [32] Advanced Television Standard Committee, "Website oficial." <http://www.atsc.org/standards.html>, 2006.
- [33] Digital Video Broadcast. "Website oficial", <http://www.dvb.org>, 2006
- [34] MPEG, "Website oficial." <http://www.mpeg.org>, 2006.
- [35] Mendes, Luiz R., Iano, Yuzo, Sablón, Vicente I. B., Pietro, René T; "Subsistema de Compressão de e Codificação dos Sinais de Vídeo dos Padrões HDTV", Revista Científica Periódica - Telecomunicações, Volume 3, Número 01, Páginas 19-27, 2000.
- [36] SMPTE 421M, "VC-1 Compressed Video Bitstream Format and Decoding Process", 2006.
- [37] ETS 300 743, Digital Video Broadcasting Subtitling Systems, September 1997.
- [38] William Sattlings, "Cryptography and Network Security", Prentice-Hall, 2nd edition, 1998.
- [39] NIST. Advanced Encryption Standard (AES), 2006.
- [40] <http://csrc.nist.gov/encryption/aes>, 2006.
- [41] NIST. Overview of the AES Development Effort. <http://csrc.nist.gov/encryption/aes/index2.html#overview>, February 2001.
- [42] DAEMEN, Joan e RIJMEN, Pawel. The block cipher Rijndael. Página oficial do Rijndael, disponível em <http://www.esat.kuleuven.ac.be/~rijmen/rijndael/>, September

1999.

- [43] <http://www.ietf.org/rfc.html>, RFC 3376, Internet Group Management Protocol, Version 3. B. Cain, S. Deering, I. Kouvelas, B. Fenner, A. Thyagarajan, October 2002.
- [44] ITU-R BT.656, Interfaces for Digital Component Video Signals in 525-Line and 625-Line Television Systems, January 1986.
- [45] SMPTE-259M, Serial Digital Interface, SDI, March 1999.
- [46] SMPTE-292M, High Definition Serial Digital Interface, HD-SDI, 1998.
- [47] ANSI/SCTE-35, Digital Program Insertion Cueing Message for Cable, 2001.
- [48] <http://www.ietf.org/rfc.html>, RFC 1889, RTP: A Transport Protocol for Real-Time Applications. H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson. January, 1996.
- [49] <http://www.ietf.org/rfc.html>, RFC 2326, RTSP: Real Time Streaming Protocol (RTSP). H. Schulzrinne, A. Rao, R. Lanphier, April 1998.
- [50] ITU-T Recommendation G.992.1, "Asymmetric digital subscriber line (ADSL) transceivers", July 1999.
- [51] ITU-T Recommendation G.992.3, "Asymmetric digital subscriber line transceivers 2 (ADSL2)", December 2003.
- [52] ITU-T Recommendation G.992.5, "Asymmetric Digital Subscriber Line (ADSL) transceivers - Extended bandwidth ADSL2 (ADSL2plus), January 2005.
- [53] IEEE 802.16 Working Group on Broadband Wireless Access Standards, WiMAX, December 2005.
- [54] Vasudev Bhaskaran, Konstantinos Konstantinides, "Image and Video Compression Standards : Algorithms and Architectures", 1997.
- [55] Teraoka, Kleber, "Implementação do codificador de video H.264 com transformada FLICT", Dissertação de mestrado, UNICAMP, 2003.