● Digital headend – series of encoder, MUX, scrambler, modulator for both SD and HD

● Integrated digital headend – one carrier in one RU

● CAS / SMS – Conditional Access System

● EPG – Electronic Program Guide

● Pay-per-view (PPV)

● Remote CAS

● STB- cable, terrestrial, satellite, and MMDS, both SD and HD
Overview

Purpose:

- To gain a basic understanding of MPEG video transmission systems.

Objective:

- To understand the basics of the MPEG protocol
- To learn about the MPEG frame types & elementary streams
- To understand PTS, DTS, and PCR
- To understand the PSI table structure
What does MPEG stand for?
In 1988, the ISO Moving Picture Experts Group (MPEG) began developing a set of coding (audio and video) and transport (multiplexing and synchronization) standards for the efficient storage and delivery of video, audio and data. These standards (with additional extensions by the DVB and ATSC organizations) have pushed MPEG into cable systems everywhere.

MPEG offers:

- Superior video quality at substantial bandwidth savings
- Increased utilization in satellite and cable distribution networks
- Cost savings for storage and retrieval of programming
- Further savings due to MPEG standardization and interoperability
MPEG-1 Audio and Video (first started in July 1989)

500kbps - 1.5Mbps ideally

IS 11172-1: (System) describes synchronization and multiplexing of video and audio.
IS 11172-2: (Video) describes compression of non-interlaced video signals.
IS 11172-3: (Audio) describes compression of audio signals using high performance coding.

- The MPEG-1 standard, established in 1992, is designed to produce reasonable quality images and sound at low bit rates.

- MPEG-1 compression of audio signals, specifies a family of three audio coding schemes, simply called Layer-1, -2, -3, with increasing encoder complexity and performance (sound quality per bit-rate). MPEG-1 Layer-3 is more popularly known as MP3 and has revolutionized the digital music domain.

- MPEG-1 is intended to fit the bandwidth of CD-ROM, Video-CD and CD-i. MPEG-1 usually comes in Standard Interchange Format (SIF), which is established at 352x240 pixels at 1.5 megabits per second, a quality level about on par with VHS. MPEG-1 can be encoded at bit rates as high as 4-5Mbits/sec, but the strength of MPEG-1 is its high compression ratio with relatively high quality. MPEG-1 is also used to transmit video over ADSL, VOD, Video Kiosks, and corporate presentations. MPEG-1 is also used as audio-only (MP3).
MPEG-2 Audio and Video (first started July 1991)
2 - 8 Mbps for SDTV, up to 19 Mbps for HDTV

- Established in 1994, MPEG-2 is designed to produce higher quality images at higher bit rates.

- MPEG-2 streams at lower MPEG-1 bit rates won't look as good as MPEG-1. But at its specified bit rates between 3-10Mbits/sec, MPEG-2 at the full CCIR-601 resolution of 720x480 pixels NTSC delivers true broadcast quality video.

- MPEG-2 was engineered so that any MPEG-2 decoder will play back an MPEG-1 stream, ensuring a side-grade path for users who enter into MPEG with the lower priced MPEG-1 encoding hardware.

- The primary users of MPEG-2 are broadcast and cable companies who demand broadcast quality digital video and utilize satellite transponders and cable networks for delivery of cable television and direct broadcast satellite. It's also the standard specified for DVD encoding.
MPEG-3 was initially intended to cover HDTV, providing larger sampling dimensions and bit rates between 20-40Mbits/sec. It was later discovered that MPEG-2 could be used to cover the requirements of HDTV, so the MPEG-3 standard was dropped.
MPEG-4 [H.264/AVC] (first started in July 1995)
384 Kbps – 8+ Mbps

- The MPEG-4 standard was initiated in 1995 and was finalized at the end of 1998. This standard was initially specified for very low bit rates but now it supports up to 8+ Mbps.

- MPEG-4 is designed for use in broadcast, interactive and conversational environments. The way MPEG-4 is built allows MPEG-4 to be used in television and Web environments simultaneously, not just the one after the other, but also facilitates integration of content coming from both channels in the same multimedia scene.

- MPEG-4 is based on the MPEG-1 and -2 standards and VRML (Virtual Reality Modeling Language).

MPEG-4 adds the following to MPEG-1, 2 and VRML:
- Support for 2D and 3D content
- Support for several types of interactivity
- Coding at very low rates (2 Kbps for speech, 5 Kbps for video) to very high ones (5 Mbps for transparent quality Video, 64 Kbps per channel for CD quality Audio)
- Native support for natural content and real-time streamed content, using URLs
- Several forms of support over networks with bandwidth unknown at the time of encoding.
- AAC (Advanced Audio Codec) was standardized as an adjunct to MPEG-2 (as Part 7) before MPEG-4 was issued.
MPEG-4 consists of several standards—termed "parts"—including the following:

- **Part 2** (ISO/IEC 14496-2): Visual: A compression codec for visual data (video, still textures, synthetic images, etc.). One of the many "profiles" in Part 2 is the Advanced Simple Profile (ASP).
- **Part 3** (ISO/IEC 14496-3): Audio: A set of compression codecs for perceptual coding of audio signals, including some variations of Advanced Audio Coding (AAC) as well as other audio/speech coding tools.
- **Part 4** (ISO/IEC 14496-4): Conformance: Describes procedures for testing conformance to other parts of the standard.
- **Part 5** (ISO/IEC 14496-5): Reference Software: Provides software for demonstrating and clarifying the other parts of the standard.
- **Part 7** (ISO/IEC 14496-7): Optimized Reference Software: Provides examples of how to make improved implementations (e.g., in relation to Part 5).
- **Part 9** (ISO/IEC 14496-9): Reference Hardware: Provides hardware designs for demonstrating how to implement other parts of the standard.
- **Part 10** (ISO/IEC 14496-10): Advanced Video Coding (AVC): A codec for video signals which is technically identical to the ITU-T H.264 standard.
- **Part 11** (ISO/IEC 14496-11): Scene description and Application engine, also called BIFS; can be used for rich, interactive content with multiple profiles, including 2D and 3D versions.
- **Part 16** (ISO/IEC 14496-16): Animation Framework eXtension (AFX).
- **Part 17** (ISO/IEC 14496-17): Timed Text subtitle format.
- **Part 18** (ISO/IEC 14496-18): Font Compression and Streaming (for OpenType fonts).
- **Part 20** (ISO/IEC 14496-20): Lightweight Scene Representation (LASer).
- **Part 22** (ISO/IEC 14496-22): Open Font Format Specification (OFFS) based on OpenType (not finished - reached "CD" stage in July 2005)
- **Part 23** (ISO/IEC 14496-23): Symbolic Music Representation (SMR) (not yet finished - reached "FCD" stage in October 2006)

Profiles are also defined within the "parts", so an implementation of a part is not always an implement of an entire part.
MPEG offers superior video/audio quality at significant bandwidth savings. Digitizing a standard U.S. movie without means of compression will produce a transport stream of about 168 Mbps. Using 256 QAM modulation, it would take 5 separate 6 MHz channels to transmit one movie.

Using an MPEG-II transport stream, this bit rate can be reduced to a 3-4 Mbps elementary stream (a 50:1 reduction) without significantly affecting perceived video quality.

This allows an MSO to transmit this movie plus 9 or more movies in just one 256-QAM 6 MHz channel.
In the world of analog, video is nothing more than a collection of still pictures, which when presented in quick succession on a Television, appears to be fluid motion. Video is captured at 30 pictures (called frames) per second and presented at the same rate.

The world of digital works much the same way. A stream of 1’s and 0’s (bits) of data are sent to a decoder, which reassembles them to again create 30 frames (29.97 to be exact) frames per second. The difference is in the bits. Each point (pixel) in the picture must be represented by these bits, which happens when the initial pictures are “coded“.

Factoid: Movies viewed at the movie theater are only broadcast at 24 frames per second.
Intra-frame coding

- In this technique, a frame is encoded using only information from the same analog frame (itself). Typical images contain large areas in which adjacent samples are spatially correlated.

- Macroblocks are 16x16 pixel (individual points in a picture) areas of original image. A macroblock usually consists of 4 Y blocks, 1 Cr block, and 1 Cb block.

- Research into the Human Visual System (HVS) has shown that the eye is most sensitive to changes in luminance, and less sensitive to chrominance variations. Since compression is key, MPEG operates on a color space that can effectively take advantage of the eye’s different sensitivity to luminance and chrominance information.

- As such, MPEG uses the YCbCr color space to represent the data values instead of RGB, where Y is the luminance signal, Cb is the blue color difference signal, and Cr is the red color difference signal.
Intra-frame coding, macroblocks

- The macroblocks described previously and shown below are created using a mathematical technique known as a Discrete Cosine Transform (DCT).
- A macroblock can be represented in several different manners when referring to the YCbCr color space. The 3 formats used are known as 4:4:4, 4:2:2, and 4:2:0 video. 4:4:4 is full bandwidth YCbCr video, and each macroblock consists of 4 Y blocks, 4 Cb blocks, and 4 Cr blocks. Being full bandwidth, this format contains as much information as the data would if it were in the RGB color space. 4:2:2 contains half as much chrominance information as 4:4:4, and 4:2:0 contains one quarter of the chrominance information.
- Although MPEG-2 has provisions to handle the higher chrominance formats for professional applications, most cable and satellite feeds use the normal 4:2:0 mode.
**4:4:4**  
No chroma sub sampling, each pixel has Y, Cr and Cb values. 4:2:2 -> Horizontally sub sample Cr, Cb signals by a factor of 2.

**4:1:1**  
Horizontally sub sampled by a factor of 4.

**4:2:0**  
Sub sampled in both the horizontal and vertical dimensions by a factor of 2. Theoretically, the chroma pixel is positioned between the rows and columns as shown in the figure.

4:1:1 and 4:2:0 are mostly used in JPEG and MPEG.
Inter-frame coding

Intra-frame coding techniques were limited to processing the video signal relative only to information within the current video frame. Considerably more compression efficiency can be obtained however, if time-based redundancies, are exploited as well. With televisions displaying pictures at 30 frames per second, very little information will change from frame to frame in the span of a second or less, provided no scene changes occur. Temporal processing to exploit this redundancy uses a technique known as block-based motion compensated prediction, using motion estimation.

Starting with an I frame, the encoder can forward predict a future frame. This is commonly referred to as a P frame, and it may also be predicted from other P frames, although only in a forward time manner. As an example, consider a group of pictures that lasts for 6 frames. In this case, the frame ordering is given as I,P,P,P,P,P,I,P,P,P,P,P. Each P frame in this sequence is predicted from the frame immediately preceding it, whether it is an I frame or a P frame. As a reminder, I frames are coded spatially with no reference to any other frame in the sequence.
Statistical redundancy (Entropy or Huffman coding)

Statistical methods are used to code the most frequently occurring sequences with fewer bits (pulled from a lookup table) compared to those which are less likely.

Factoid- In 1951, David Huffman and his MIT information theory classmates were given the choice of a term paper or a final exam. The professor, Robert M. Fano, assigned a term paper on the problem of finding the most efficient binary code. Huffman, unable to prove any codes were the most efficient, was about to give up and start studying for the final when he hit upon the idea of using a frequency-sorted binary tree, and quickly proved this method the most efficient.
MPEG Video Frames in review

I Frames
Intra-coded only – This is a reference frame for future predictions. Moderate compression (on order of 10:1), limits the propagation of transmission of errors, supports random access and fast forward/fast reverse.

P Frames
Forward prediction from either previous I frames or previous P frames. Reference for future P or B frames. Good compression savings (20:1).

B Frames
Bi-directional interpolated prediction from two sources. Previous reference I or P frames (forward prediction). Future reference I or P frames (backwards prediction). Highest compression (50:1).
Group of pictures (GOP)
The first frame of a group and all additional frames dependant on each other for prediction. The number of frames in a group is user selectable (Typically 15)

Video sequence
A series of frames
Digital Video is all about timing...with 3 sources of timing!

PCR: Program Clock References
- Inserted into transport stream every 100 ms max.
- Based upon a 27MHz clock at the encoder
- Synchronizes program timing

DTS: Decode Time Stamp
- Instructs the decoder (ie STB) when to decode specific GOP’s

PTS: Presentation Time Stamp
- Instructs the decoder when to present specific GOP’s
Audio can be synchronized with video or be standalone

MPEG-1 Audio: A 2-channel (mono, dual mono or stereo) w/sampling rates of 44.1 kHz (CD), 48 kHz (DAT), 32 kHz (digital voice), compressed bit rates of 32-192 kbps/channel

MPEG-2 BC: (Backward Compatible) is multi-channel (up to 5.1) w/additional lower 16, 22.05, and 24 kHz sampling rates

MPEG-2 AAC: (Advanced Audio Coding) is a very high-quality audio coding standard for 1 to 48 channels at sampling rates of 8 to 96 kHz, with multi-channel, multi-lingual, and multi-program capabilities.
Dolby AC-3 is not part of the MPEG standard (although it can be carried in an MPEG-2 Transport Stream as a private data stream)

- It has been adopted by the ATSC (for HDTV), North American Digital Cable (most stb’s now support both MPEG-2 Audio and Dolby AC-3) and is part of the DVD audio spec (along with MPEG-1 & 2 Audio and Linear PCM)

- The compression algorithm is proprietary in nature and must be licensed from Dolby Labs, Inc.

- Compressed bit rates of 32-640 kbps

- Mono to full 5.1 channels

Factoid: Most feeds found in a Cable Headend or off satellite are encoded using AC-3 with a 48 KHz sampling rate at a bit rate of 192 or 384 KHz.
Sources are made up of elementary streams. Each program elementary stream (PES) is encoded as a separate stream, then synchronized through a common program PCR.
A Transport stream is a multiplex of one or more programs (each of which may have their own independent time base)

Provides a synchronization mechanism for dealing with delay variations in packet delivery

Provides for the more sophisticated handling of data (private data, CA)

Fixed length 188 or 204 byte packets. Most North American Digital Cable Systems are designed for 188 byte packets.

Factoid- The byte size of MPEG transport streams was chosen because they aligned with DES encryption and ATM payloads. The small size of the packets (188 or 204) are also less prone to errors during transport.
Sync is a unique code that indicates the start of every transport packet. Note that the PES packets can be split over transport packet boundaries.
The MPEG-2 standard defines a number of tables which help in the reassembly of individual elementary streams after packet transport across a network:

**Program Association Table (PAT)**
This table has a PID value of zero and contains a list of all programs in the transport and their associated PID values which are used to index into the Program Map Table.

**Program Map Table (PMT)**
This table has a user assigned PID value. Each program carried in the transport stream has a program map table entry associated with it. This entry contains the PID values of each of the elementary streams (video, audio, data) which make up a program.

**Conditional Access Table (CAT)**
This table has a unique PID value of one. If any of the elementary streams are scrambled / encrypted then a CAT entry is usually detailed. This provides the PID values for the transport packets that contain the conditional access management and entitlement information.
The Digital Video Broadcast (DVB) consortium has extended the basic MPEG PSI tables (in documents EN 300 468 and others) in order to support a services model and additional data types such as teletext, subtitles and EPGs (Electronic Program Guides). DVB streams are typically found in Europe and other countries outside North America. These extensions include:

**Network Information Table (NIT)**
This table has a PID value of 0X10 and contains private data which details items such as channel frequencies, satellite transponder details, modulation characteristics, service origin, service name and details of any alternative networks. It also indicates the PID for the NIT.

**Service Description Table (SDT)**
This table has a PID value of 0x11 and contains data describing the services in the system.

**Event Information Table (EIT)**
This table has a PID value of 0x12 and contains data concerning events or programmes, e.g names, start times, duration etc which are used to construct the electronic program guide.
An excellent example of an MPTS (Multiple Program Transport Stream).

- **SPTS**
  - Video 1 PID
  - Audio 1 PID
  - Data 1 PID

- **MPTS**
  - Video 1 PID
  - Audio 1 PID
  - Data 1 PID
  - Video 2 PID
  - Audio 2 PID
  - Audio 3 PID
  - Data 2 PID

- **Other Data** (Teletext, CA, etc)

- **MPEG PSI**
  - PAT, CAT, PMT

- **ATSC PSIP**
  - STT, MGT, VCT, EIT

- **DVB SI**
  - TDT, SDT, SAT

**MPEG-2 MPTS**
Using a 6 Mhz channel for transmission, the data rate depends on modulation.

- Using 8-VSB (Off-Air), the maximum Data Rate is 19.39 Mbps.
- Using 64 QAM (Cable), the maximum Data Rate is 26.97 Mbps.
- Using 256 QAM (Cable), the maximum Data Rate is 38.8 Mbps.

**IMPORTANT**
If the accumulative data rate of all programs exceed the total max allowed, all pictures break up!!!
A program can be encoded as a CBR or a VBR program.

CBR - Constant Bit Rate. The bit rate is constant regardless of scene complexity, resulting in either wasted bandwidth or a low quality program.

VBR - Variable Bit Rate. The bit rate constantly changes based on the complexity of each scene. Quality remains constant.

Typical bit rates (everything except locally encoded is VBR and there are often exceptions):
- Locally Encoded programs - 3-6 Mbps CBR (depending on settings)
- Most satellite delivered DC-II programs - Approx 3-4 Mbps average
- Most satellite delivered Powervue programs - Approx 4-5 Mbps average
- Sports programming - 4-6 Mbps average
- Non Sports HD programs over Satellite - 11-15 Mbps
- Sports HD programs over Satellite (and HDNet) - 17-19 Mbps
- HD programs broadcast Off-Air - 11-19 Mbps

This means that a 256 QAM 6Mhz channel (38.8 Mbps) can comfortably carry about 9-12 programs without lose of quality.
Statistical Remultiplexing involves the use of two functions:

1. Statistical Remultiplexing allows a Cable Operator to pick and choose only the services he desires from multiple MPTSs (Multiple Program Transport Streams) and create a brand new MPTS with the new services chosen.
2. Since most services are VBR (variable bit rate), the bit rate is constantly changing. However, the transport stream each service rides in is always CBR (often 27 or 38.8 Mbps), so the statistical remultiplexer must “manage the bit rate of each service to avoid bit rates which exceed the transport stream bandwidth.

Managing service bit Rates (transrating) is done through several methods:
- Stripping out null padding from encoder generated Constant Bit Rate (CBR) streams. This method does not affect picture quality.
- Time shifting. Because the transport streams are being buffered, recreating a mux allows the opportunity to manage the buffers to ensure no two services peak at the same moment in bit rate. This method does not affect picture quality.
- Re-quantizing the video as required. This is usually achieved through proprietary algorithms. These methods above can often achieve a 25-40% reduction of bit rate without significant loss in quality, depending on the original content.
Thank You