Topics in Digital TV: DTV Basics (Advanced)

IEEE BTS DL Seminar
Rich Chernock
CSO
Triveni Digital

IEEE BTS Distinguished Lecturer Program

- Mission: To serve the needs of the members of the BT Society to enhance their professional knowledge and vitality by keeping them informed of the latest research results and their practical applications. The BTS Distinguished Lecturer Program exists for the purpose of providing BTS Chapters with a list of quality lecturers who can potentially give talks at local chapter meetings. The program provides a means for chapters to have access to individuals who are well known educators and authors in the fields of broadcast technology to lecture at Chapter meetings.

Agenda

- **DTV Overview**
- Transport Basics
- Timing
- Coding
- Loudness
- Future

Environment

The Old Days: Television involves wiggling voltages in the right way at the right times so the receiver can recreate the pictures

The New Paradigm: Television involves transmitting database information and parameters to allow the pictures to be calculated.

Observations:

A TV is not expected to behave like a computer

Going black is **NOT** an option

Viewers should not need training to watch DTV

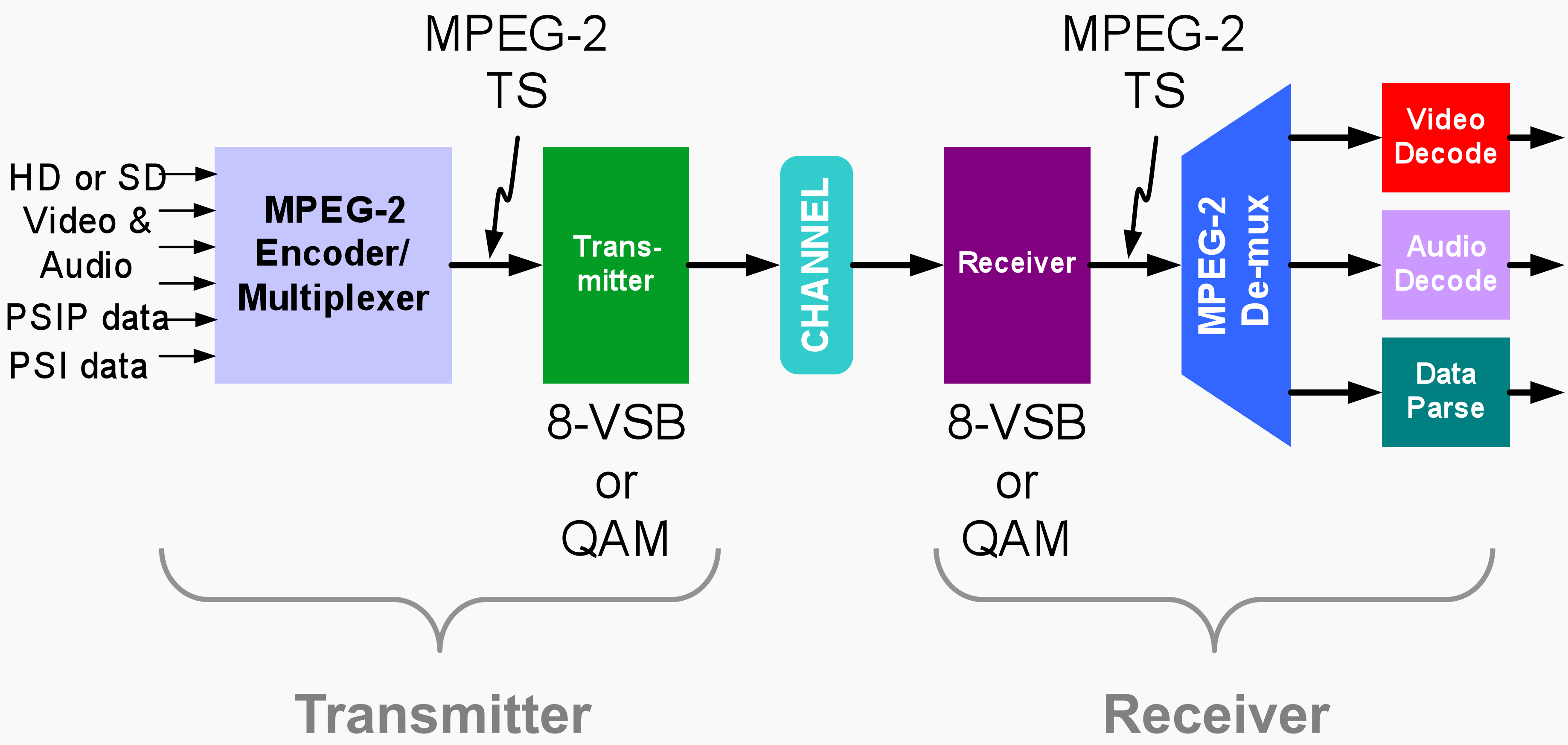
Going digital offers new revenue opportunities to broadcasters

$R_iG_jB_k$ Screen of Death

DTV System Assumptions/Goals

- Interoperability between any sender & any receiver
  - Standards
- One way broadcast
  - Everything needs to work without a backchannel
- Inexpensive receivers
  - Place complexity / cost at transmission side
  - 1 transmitter, potentially millions of receivers
- Minimize assumptions
- Hide the complexity

System Block Diagram

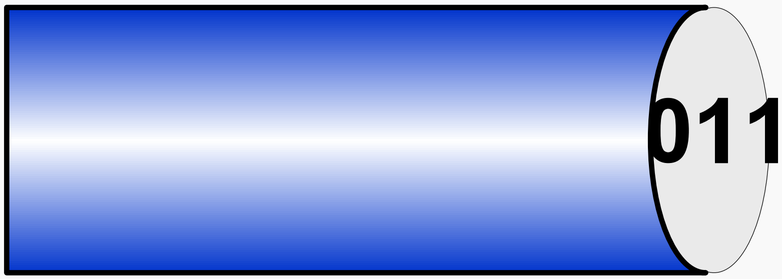


DTV = VIDEO + AUDIO + DATA + Metdata

MPEG-2 Transport Streams carrying multiplexed:

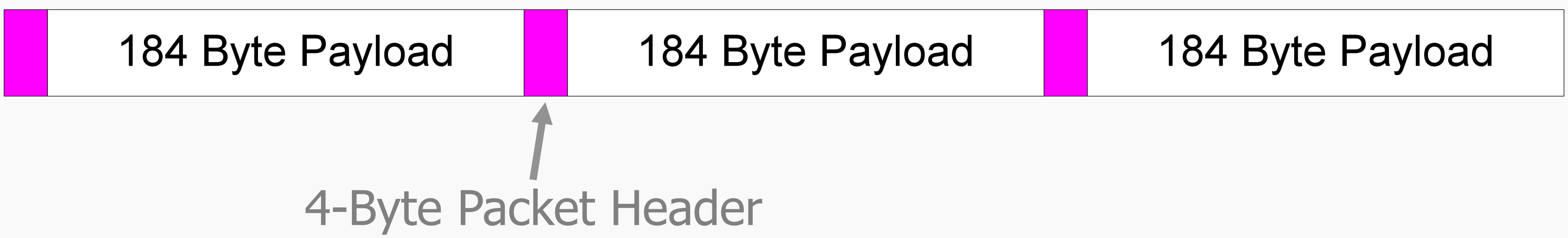
- **Service Information (DVB SI + MPEG-2 PSI)**
- **Audio, video and <u>*data elementary streams*</u>**

The Digital Pipe

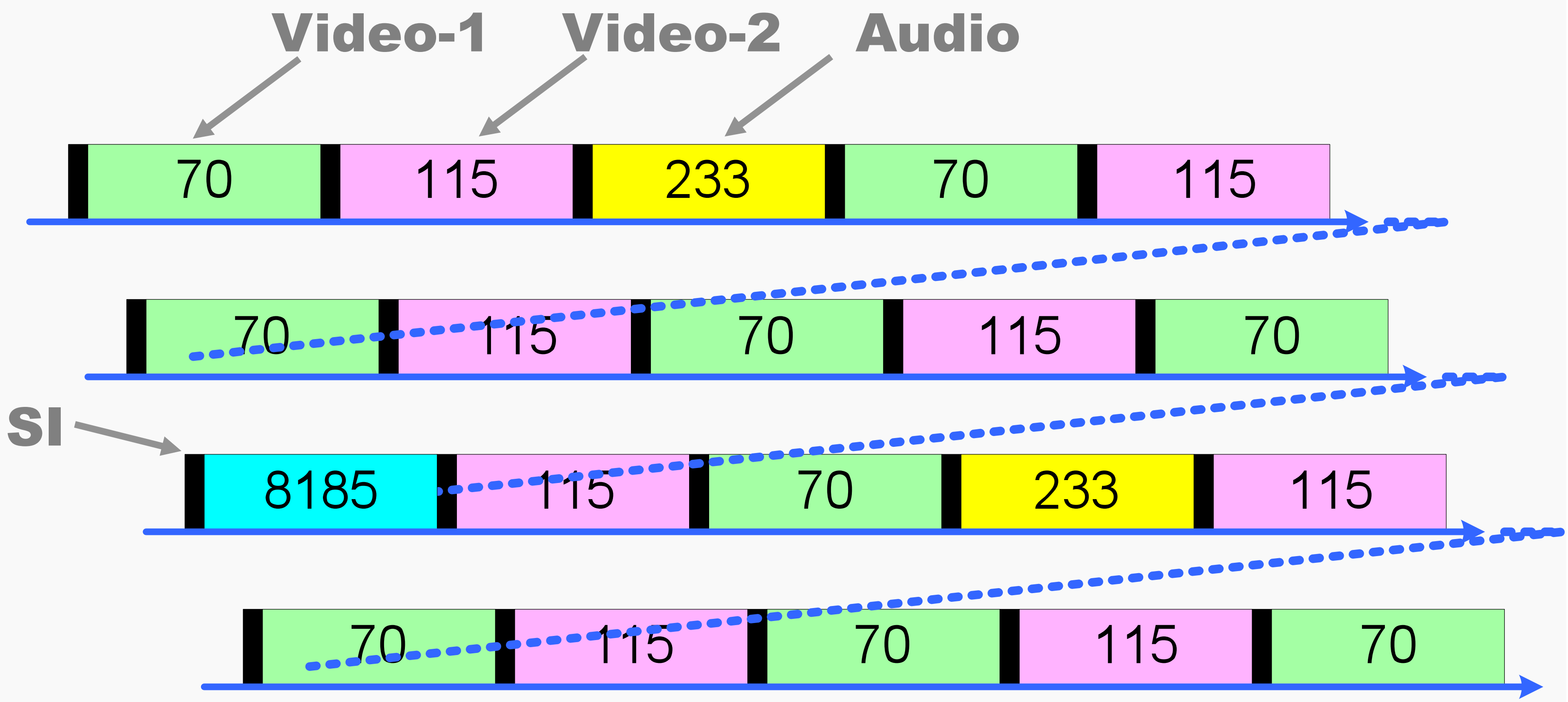


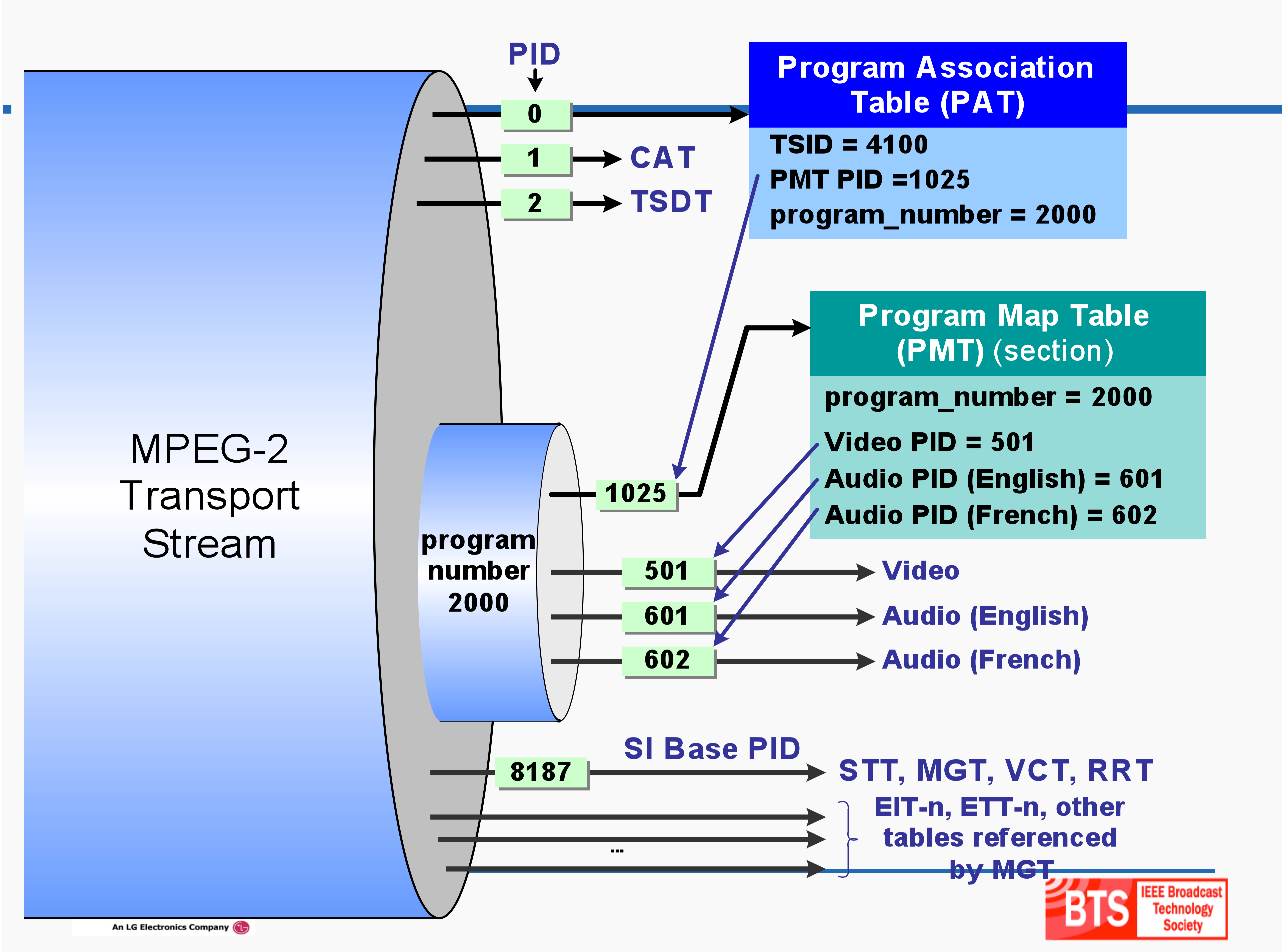
011001000101010 1...

- MPEG-2 Transport Stream



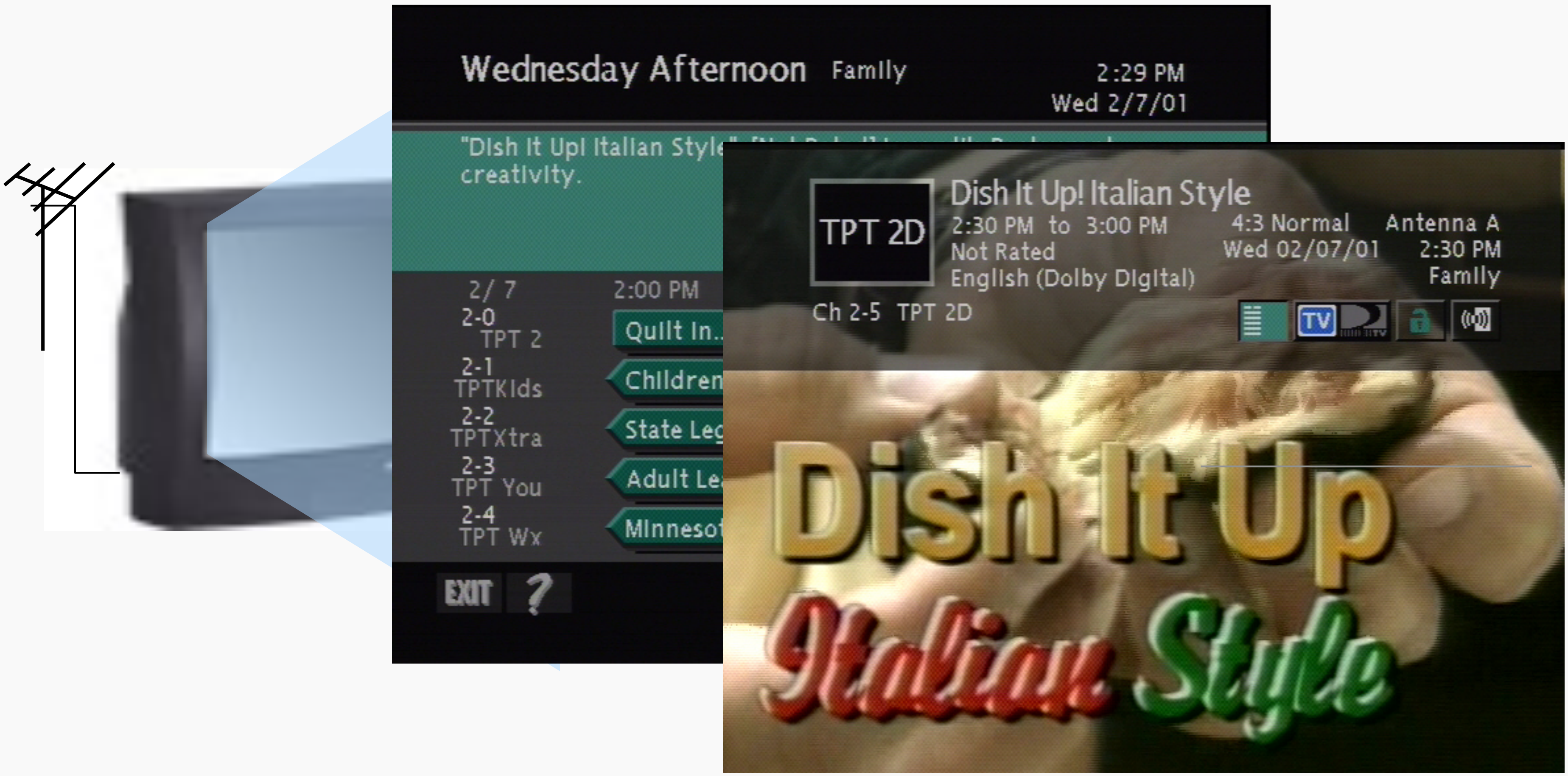
MPEG-2 Transport Stream





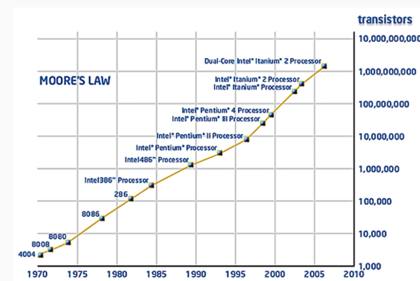
PID
0
1
2
CAT
TSDT
Program Association Table (PAT)
TSID = 4100
PMT PID =1025
program_number = 2000
Program Map Table (PMT) (section)
program_number = 2000
Video PID = 501
Audio PID (English) = 601
Audio PID (French) = 602
MPEG-2 Transport Stream
program number 2000
1025
501
601
602
Video
Audio (English)
Audio (French)
SI Base PID
8187
STT, MGT, VCT, RRT
...
EIT-n, ETT-n, other tables referenced by MGT
An LG Electronics Company
BTS IEEE Broadcast Technology Society

End Result - Television



Standards: Toolboxes & Tools

- A standard is equivalent to defining a toolbox
  - What tools are available
  - How are they to be used (constraints)
  - By defining the tools & usage, it is possible to independently build systems that interoperate (without requiring conversations)
- Standards evolution
  - Capitalize on technology advances:
    - Define new tools?
    - Extensions on existing tools?
    - Relaxation of constraints?
    - Combination of above…



Agenda

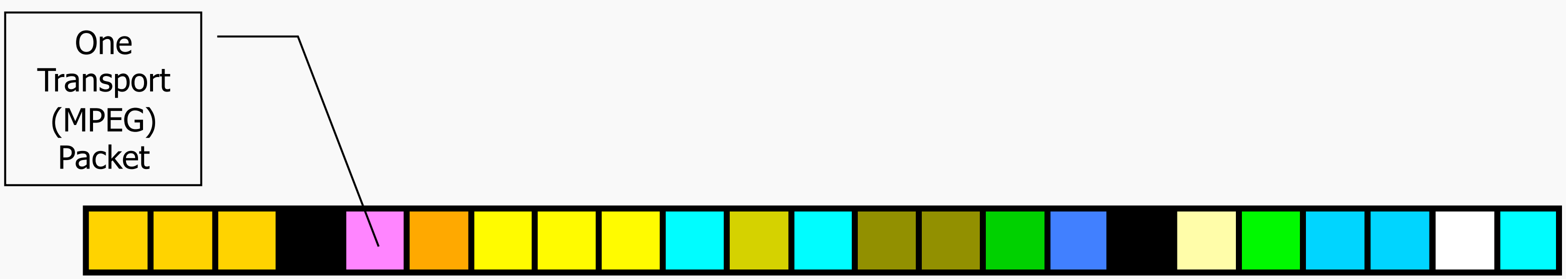
- ✓ *DTV Overview*
- ➢ **Transport Basics**
- Timing
- Coding
- Loudness
- Future

DTV Broadcast Stream

- Special case of ***MPEG-2 transport stream***
- May contain multiple ***services***
  - Video channels
    - A video stream
    - One or more audio streams
    - Possibly one or more data streams
  - Audio channels
    - One or more audio streams
    - Possibly one or more data streams
  - Data-only channels
    - One or more data streams

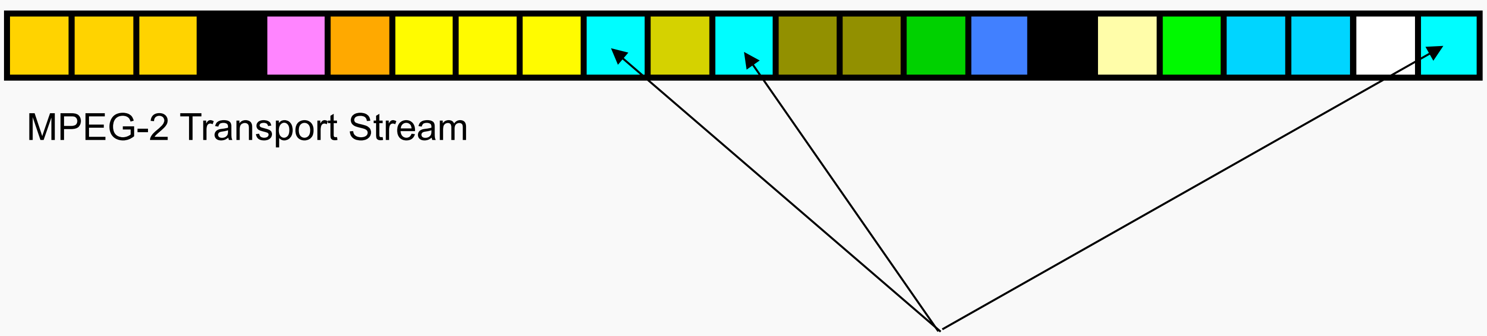
MPEG-2 Transport Stream

- Made up of 188-byte ***transport packets***, each with 4 byte header & 184 byte payload
- Each packet contains any ONE kind of information– audio, video, data, PSI, …



MPEG-2 Transport Stream (Contd.)

- We say transport packets have multiple interleaved ***elementary streams*** -- audio, video, data, PSI, ...
- Packets belonging to the same elementary stream are identified by ***packet id*** (PID) in packet header (same color in our illustrations).

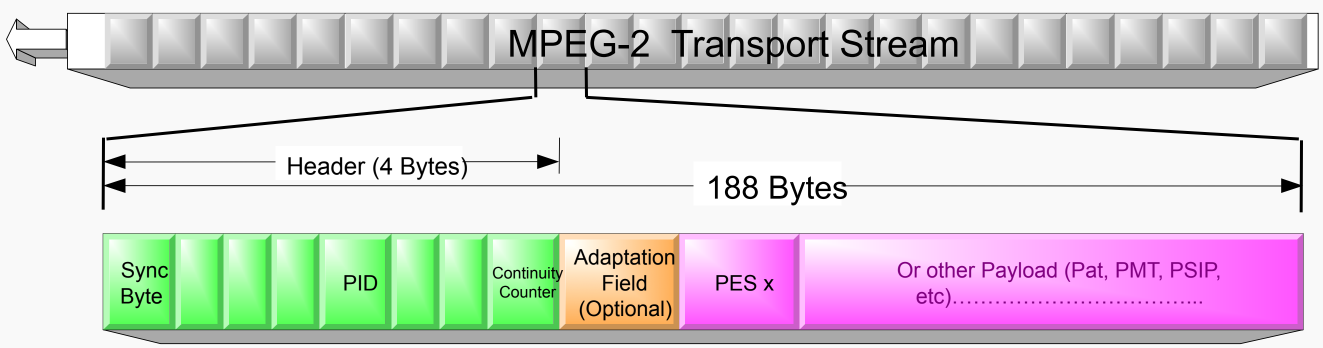


These three packets are the same color. They have the same PID and belong to the same Elementary stream.

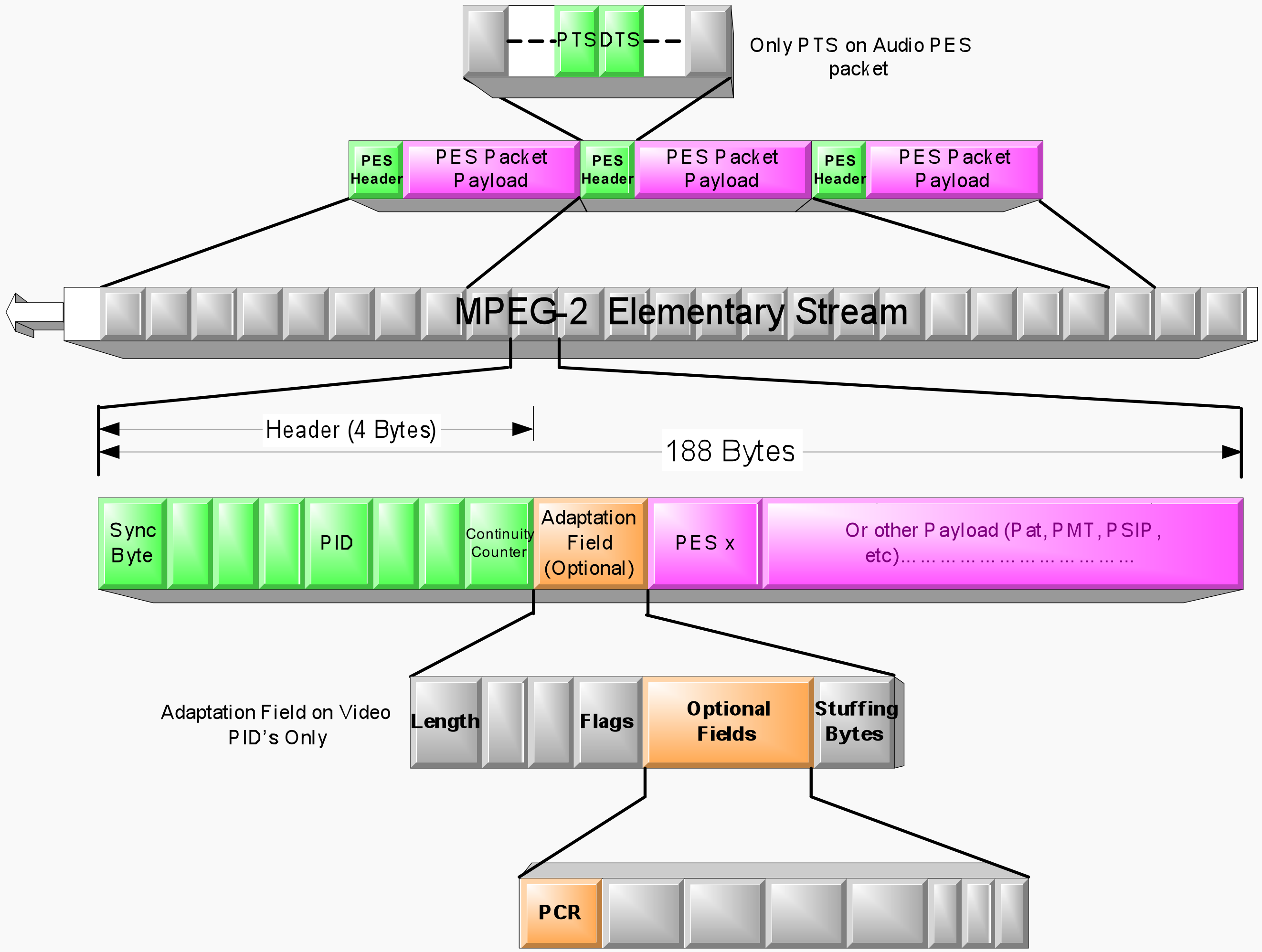
MPEG-2 Transport Stream – Header Fields

Noteworthy fields:

1) Sync Byte – Find packet boundary
2) PID – Used while demultiplexing stream
3) Continuity Counter – Identify packet loss
4) PCR stamp in adaptation field – Clock sync

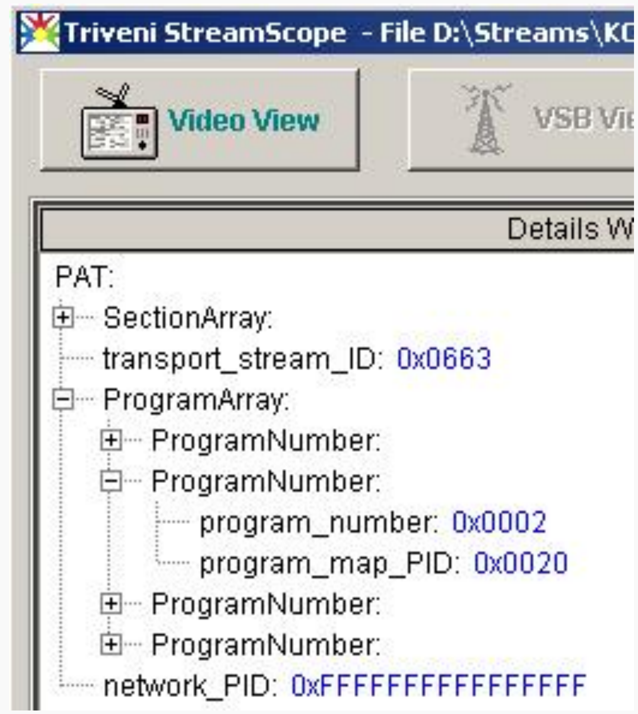


MPEG-2 Elementary Stream



MPEG-2 Sections

- MPEG-2 Sections carry "table" data
- Most often Metadata – PSI, DVB-SI
- Header describes what's inside
  - Length, type, version…
- Payload carries information
- Most often, via loops



MPEG Header Fields: Sync Byte

- When a decoder first tunes, all it sees are a stream of 0's and 1's
- The decoder must first identify the beginning of packets before it can interpret the stream
- The decoder uses the Sync Byte field to do this

01010001111010010101101010001100011110010111000

MPEG Data Prior to Packet Sync

MPEG Header Fields: Sync Byte (Contd.)

- The Sync Byte of a packet is always 0x47 (Hexadecimal) or '01000111' (binary)
- The decoder looks for strings of zeros and ones which match the pattern of the sync byte (see red below)

0101000111101001010110101000110001111001011110000

MPEG Header Fields: Sync Byte (Contd.)

- Once the decoder finds a 0x47 in the stream, it looks 187 bytes down the stream, and looks for another 0x47
- If it finds three Sync Bytes in a row, then the Decoder has Found Sync and assumes packet boundaries from then on
- Each packet is tested for 0x47 as soon as it arrives. If a packet arrives with an incorrect sync byte, the decoder starts over. This is called SYNC LOSS

Found Sync

0x47	187 bytes	0x47	187 bytes	0x47	187 bytes	Packet 188 bytes

Sync Lost

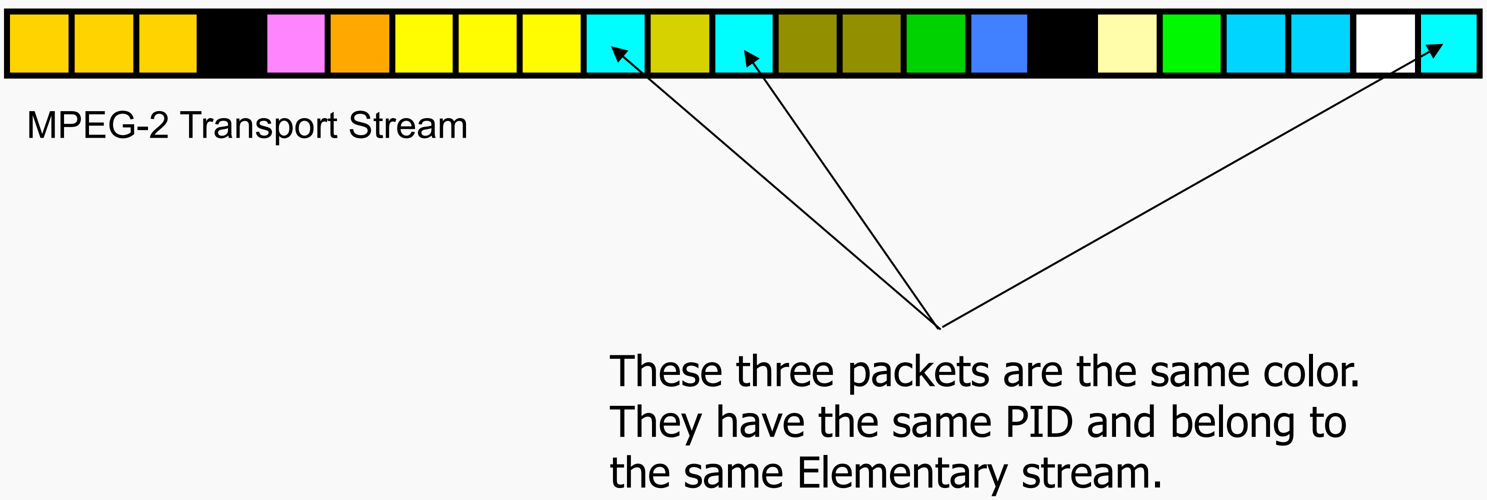
Packet 188 bytes	Packet 188 bytes	0x32	11010101010111011010101010001

MPEG Header Fields: Sync Byte (Contd.)

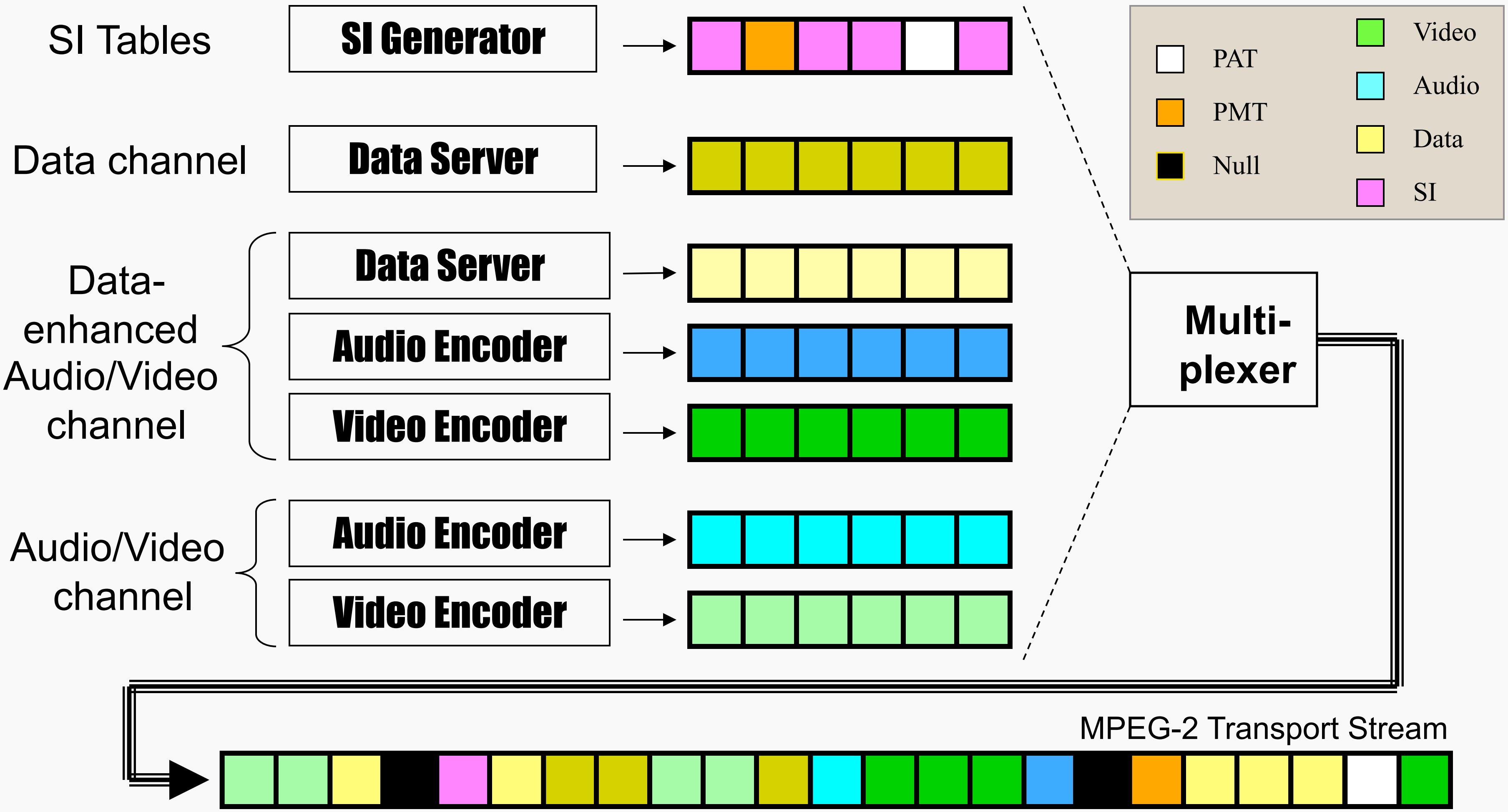
- If you don’t have Packet Sync, the decoder cannot find packet boundaries. You will not be able to decode at all
- Packet Sync problems typically occur in hardware at packet boundaries during format converters, edge devices, demodulators etc:
  - ASI to Gig-E
  - ASI to Microwave or QAM
  - Satellite to ASI

PIDs Defined

- PID stands for Packet ID
- Each Packet has a PID (indicated by color). Packets belonging to the same source of information have the same PID (same color).

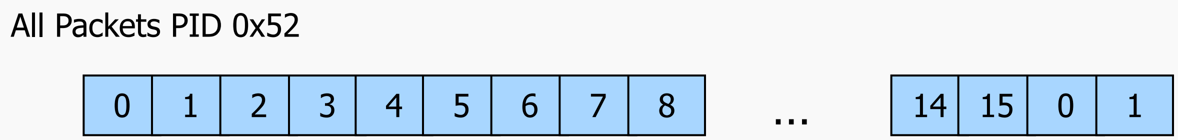


PIDs Defined – DVB Bitstream

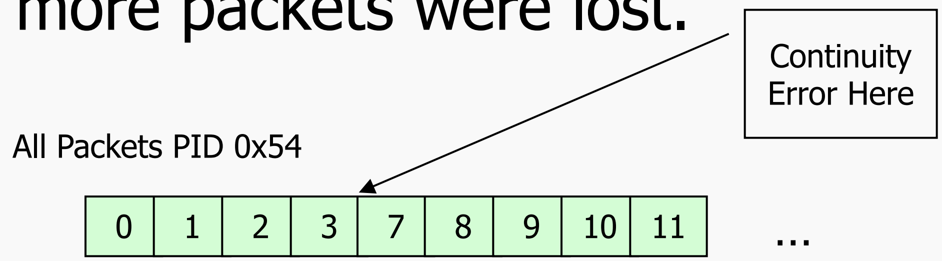


MPEG Header: Continuity Counter

- The continuity counter is a 4 bit field in the header which increments by 1 each time a packet comes out on a specific PID:

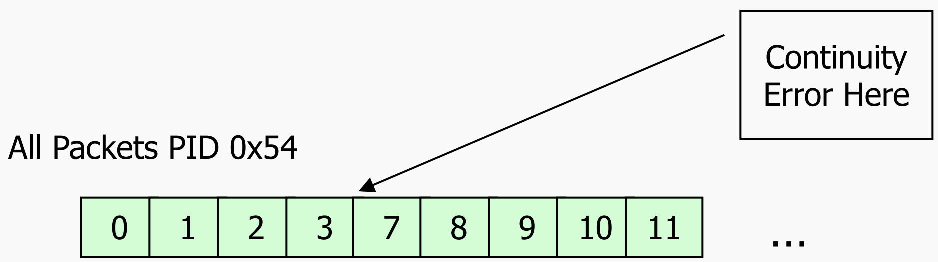


- When a PID ‘skips’ at least one value of the continuity Counter, we call it a ‘Continuity Error.’ This means one or more packets were lost.



How many Packets lost per Continuity error?

- Continuity errors indicate packet loss, but they cannot indicate how MANY packets were lost…
  - Take this example. How many packets were lost?
    - One possible answer is 3 packets. The packets with Continuity Counter values of 4, 5, and 6 were lost.
      - However, that's not the whole story.



Cycle of Continuity Counter

- The Continuity counter works just like an analog clock.
- Just as it is not possible to know if Jack slept for one, 13, 25 hours or longer, one continuity error can mean the loss of any number of packets
- Since even one Continuity Count error can mean the loss of MANY packets, these errors can cause all kinds of havoc in MPEG2 transport networks

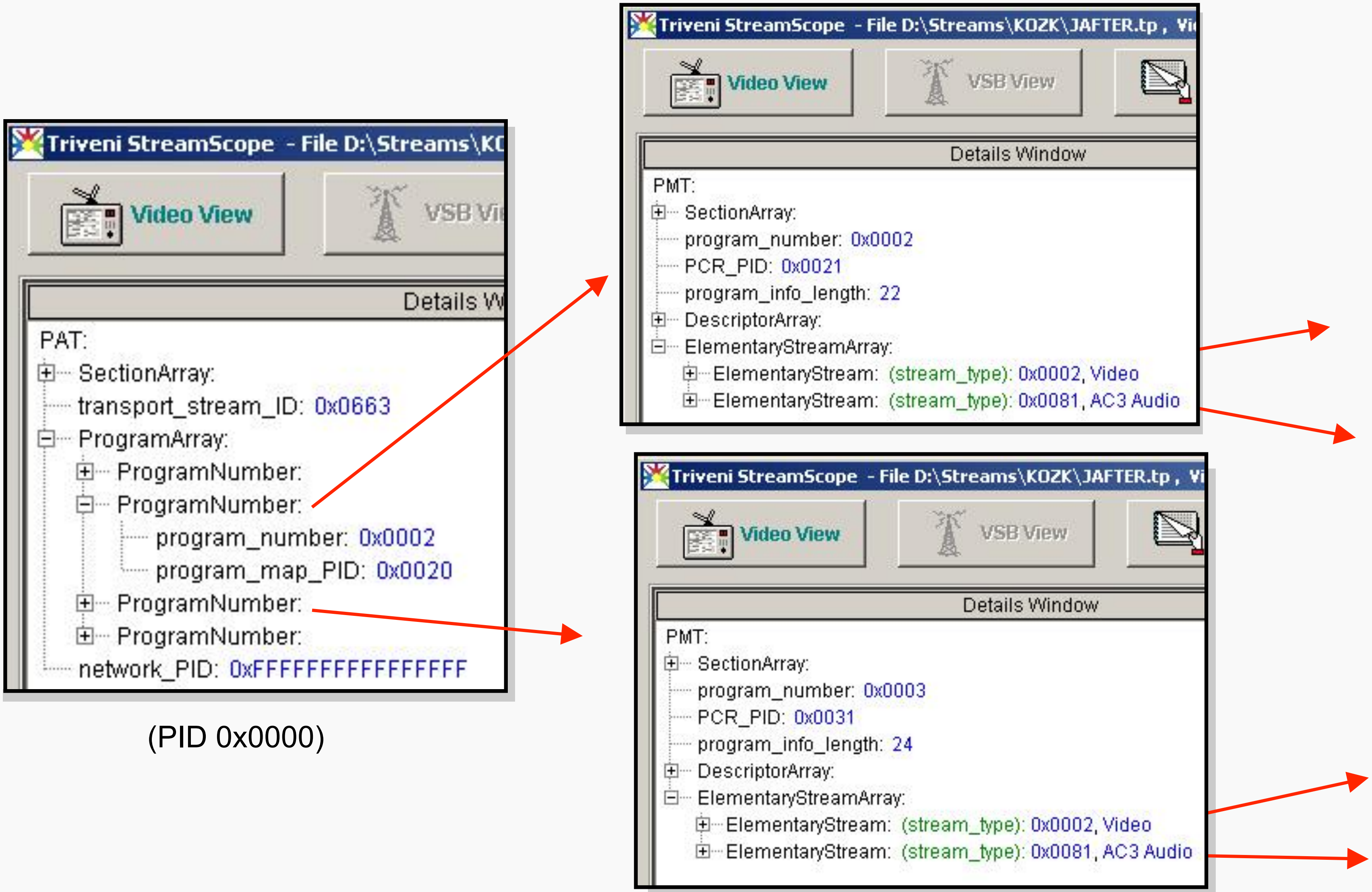
Signaling vs Announcement

- Signaling
  - Information about what is "on now"
  - Used to assemble program elements into whole
    - Provides linkages
    - Ex: PMT and/or VCT used to link different components of television program (I.e. video and audio)
  - Used to define characteristics of current program
    - Captioning, ratings, redistribution…
- Announcement
  - Information about what will be available in the future
  - Program Guide information (name, description schedule)
  - Characteristics of future programs (captioning, ratings, redistribution…)
  - Typically does not provide linkages between program elements

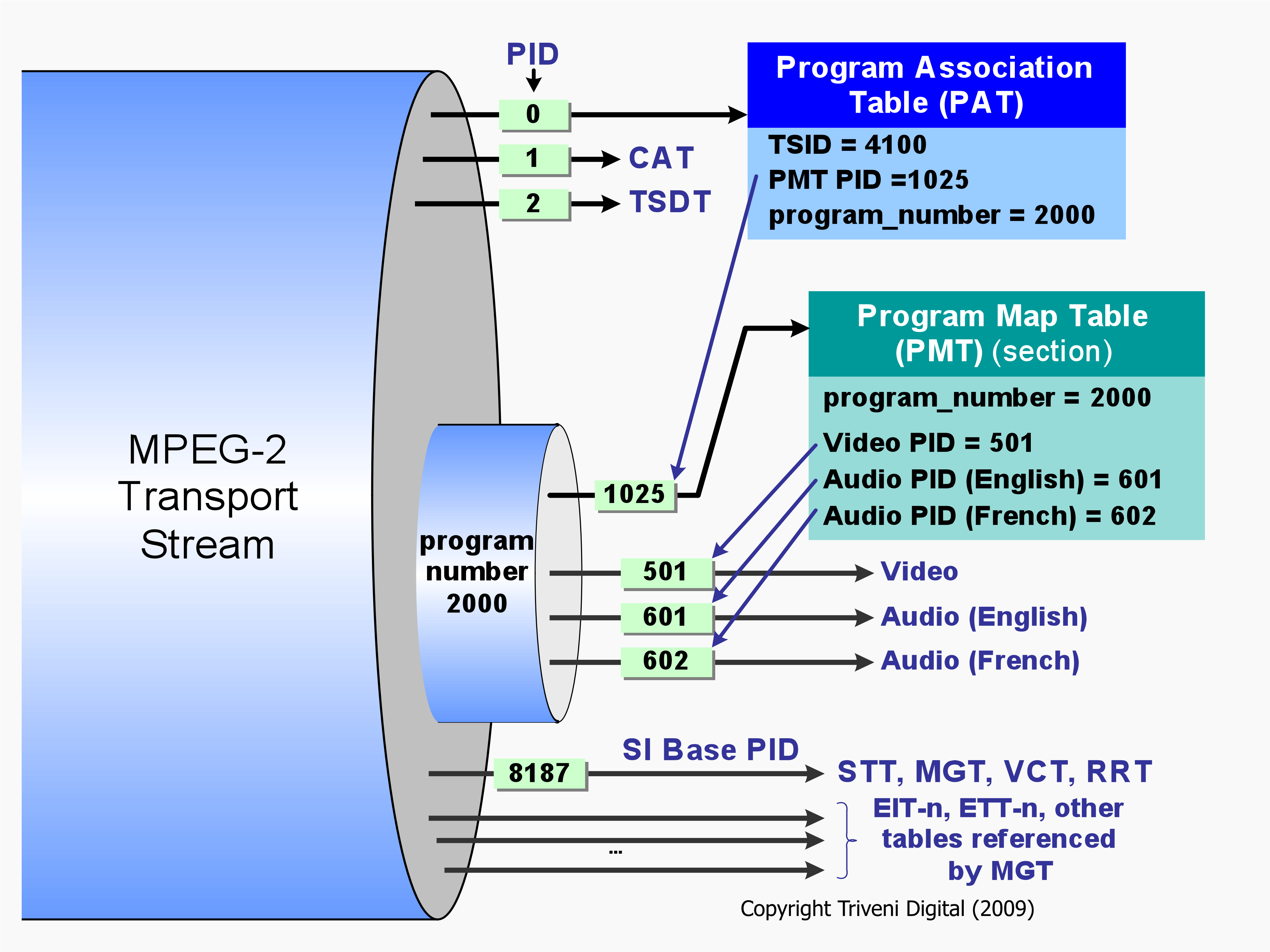
MPEG-2 Program Specific Information (PSI)

- Gives very basic tuning information:
  - PAT (Program Association Table: one for entire transport stream; identifies "programs" (virtual channels) in stream and gives PIDs for PMTs.
  - PMT (Program Map Table): one per "program"; identifies elementary streams in "program" and gives their types (audio, video, etc.) and PIDs.
- Supports tuning by physical channel number and MPEG-2 program number.

Graphical View of PAT/PMT



(PID 0x0000)



PID
0
1
2
CAT
TSDT
Program Association Table (PAT)
TSID = 4100
PMT PID =1025
program_number = 2000
Program Map Table (PMT) (section)
program_number = 2000
Video PID = 501
Audio PID (English) = 601
Audio PID (French) = 602
MPEG-2 Transport Stream
program number 2000
1025
501
601
602
Video
Audio (English)
Audio (French)
SI Base PID
8187
STT, MGT, VCT, RRT
...
EIT-n, ETT-n, other tables referenced by MGT
Copyright Triveni Digital (2009)

Agenda

- ✓ *DTV Overview*
- ✓ *Transport Basics*
- ➢ **Timing**
- Coding
- Loudness
- Future

Audio-Video Synchronization

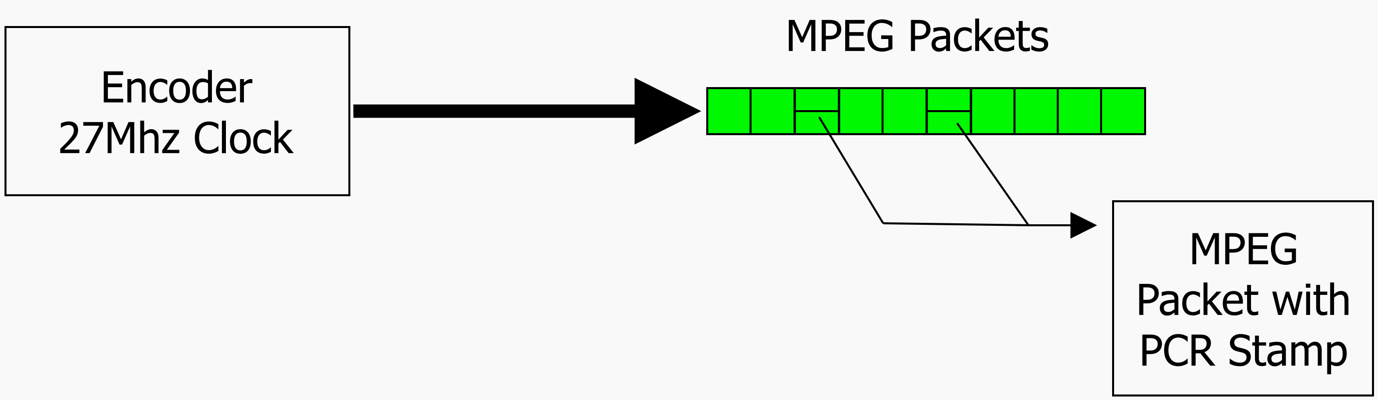
- Audio, video are encoded independently, must be synchronized during play.
  - If not, lip-sync issues ensue
- MPEG has to allow for great distances between the Encoder and Decoder, and still allow for Correct Decode of the transport stream
- MPEG has defined a strict timing & buffer model
  - Assumes constant delay between encoding and decoding
  - Allows low cost decoders, places expensive items at send side
  - This model has held up to the test of time
    - It works, but isn't always implemented correctly

How to Assure Audio/Video Sync?

- In order for the audio and video Elementary Streams to remain in Sync, the Encoder Clock and the Decoder Clock must remain in sync
  - Must work in pure broadcast mode
  - 1 way communication
  - 1 to many
- The next few slides will demonstrate how this happens

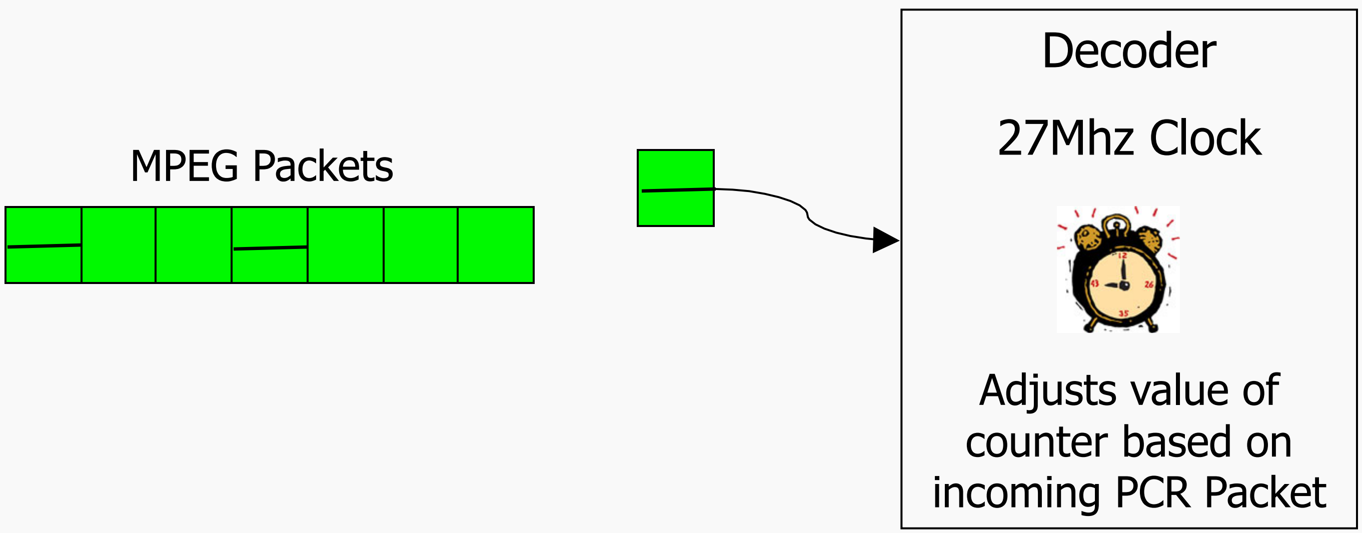
Encoder Inserts PCR

- When the encoder creates packets, it embeds the current value of its 27 MHz clock into the stream
- This time reference is called the **PCR**: **Program Clock Reference**
- MPEG demands that one PCR packet appear in the stream every 100ms
  - “User” standards more stringent



Decoder Consumes PCR

- When the decoder gets a packet containing a PCR timestamp, it adjusts its 27Mhz clock accordingly
  - Digital Phase Lock Loop circuit



PCR Jitter Defined

PCR Jitter is:

- Difference between

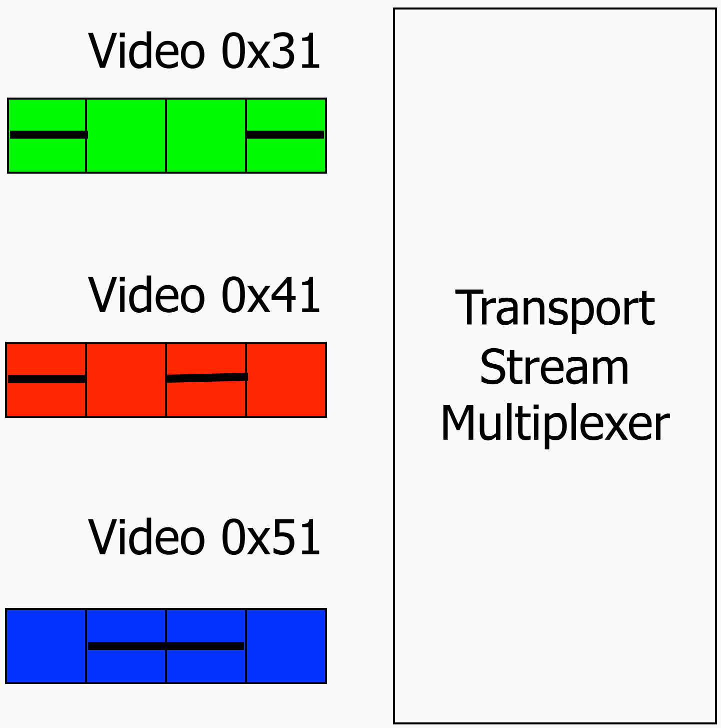
  the Actual Value of the PCR time stamped by encoder

  and

  the Expected Value of the PCR as calculated by decoder based on the clock rate and the time at which the PCR value is received.
- PCR Jitter spec: 500ns

PCR Packet Spacing Before Muxing

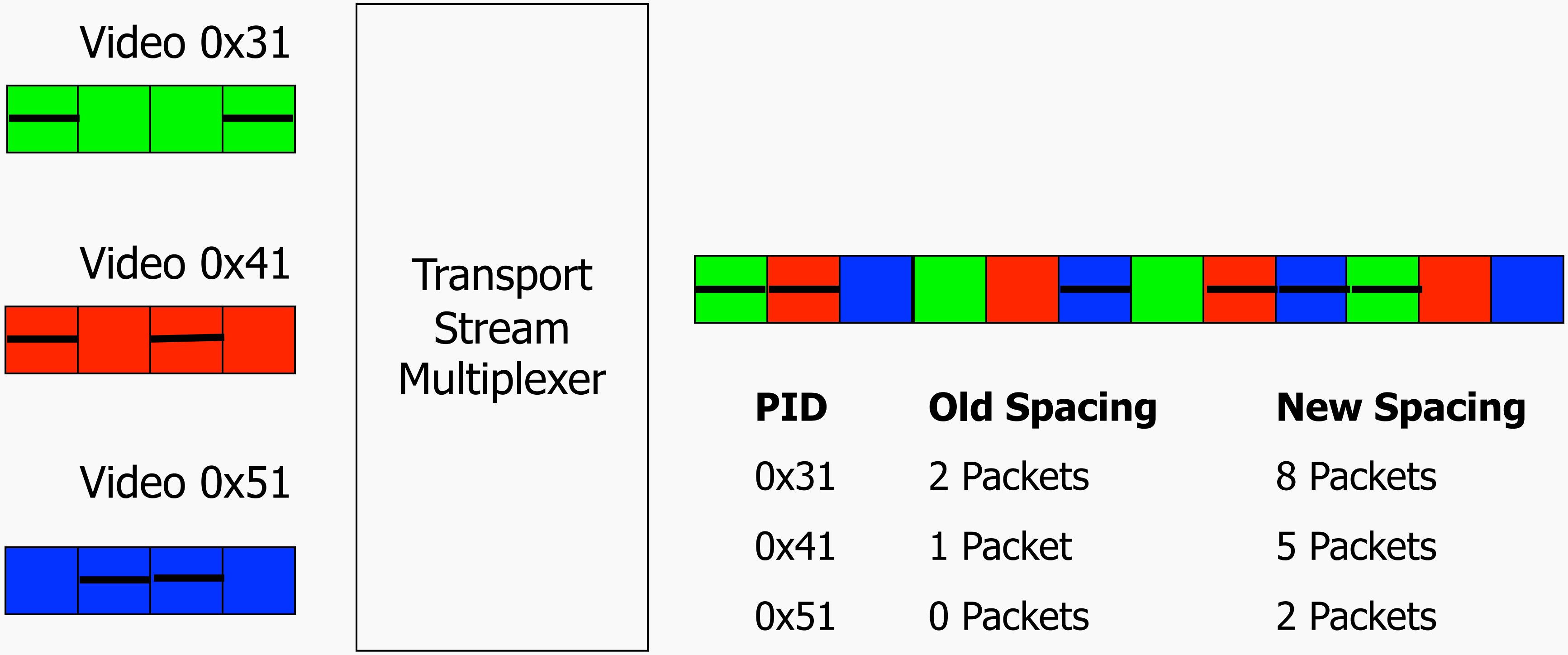
- Note the Number of Packets between each PCR packet in each Input Stream



PID	**Packet Spacing**
0x31	2 Packets
0x41	1 Packet
0x51	0 Packets

PCR Packet Spacing After Muxing

- Note that the PCR packet spacing has changed!



PID	Old Spacing	New Spacing
0x31	2 Packets	8 Packets
0x41	1 Packet	5 Packets
0x51	0 Packets	2 Packets

Muxing Causes PCR Jitter (Contd.)

- The MUX has to RESTAMP all the PCR values to correct for the change in the packet spacing – THIS IS VERY HARD TO DO
  - The more services on the output, the harder it is to restamp
  - The fewer 'null' packets at the output, the harder it is to restamp

Audio and Video Buffers

- Receiver must buffer audio and video frame data until presentation time.
- If data appears too late in the transport stream, buffer underflow results.
- If data appears too early in the transport stream, buffer overflow results.
- Either condition results in garbled play or incorrect synchronization.
- Different set top boxes may respond differently to the same underlying buffer violations

The Encoder and Decoder Clock

- The MPEG encoder and MPEG decoder use a 27Mhz ‘clock’ to encode/decode incoming audio and video
- The clock is actually a ‘counter’ which advances every 1/27000000 seconds

Encoder

27Mhz Clock

1,2,3,4…..

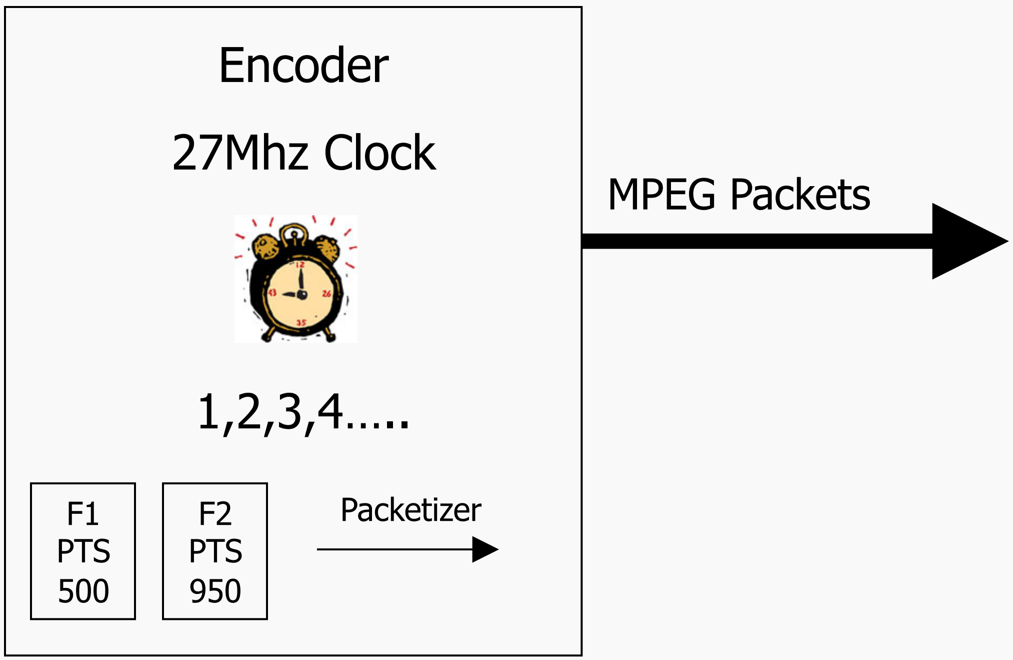
Decoder

27Mhz Clock

1,2,3,4…..

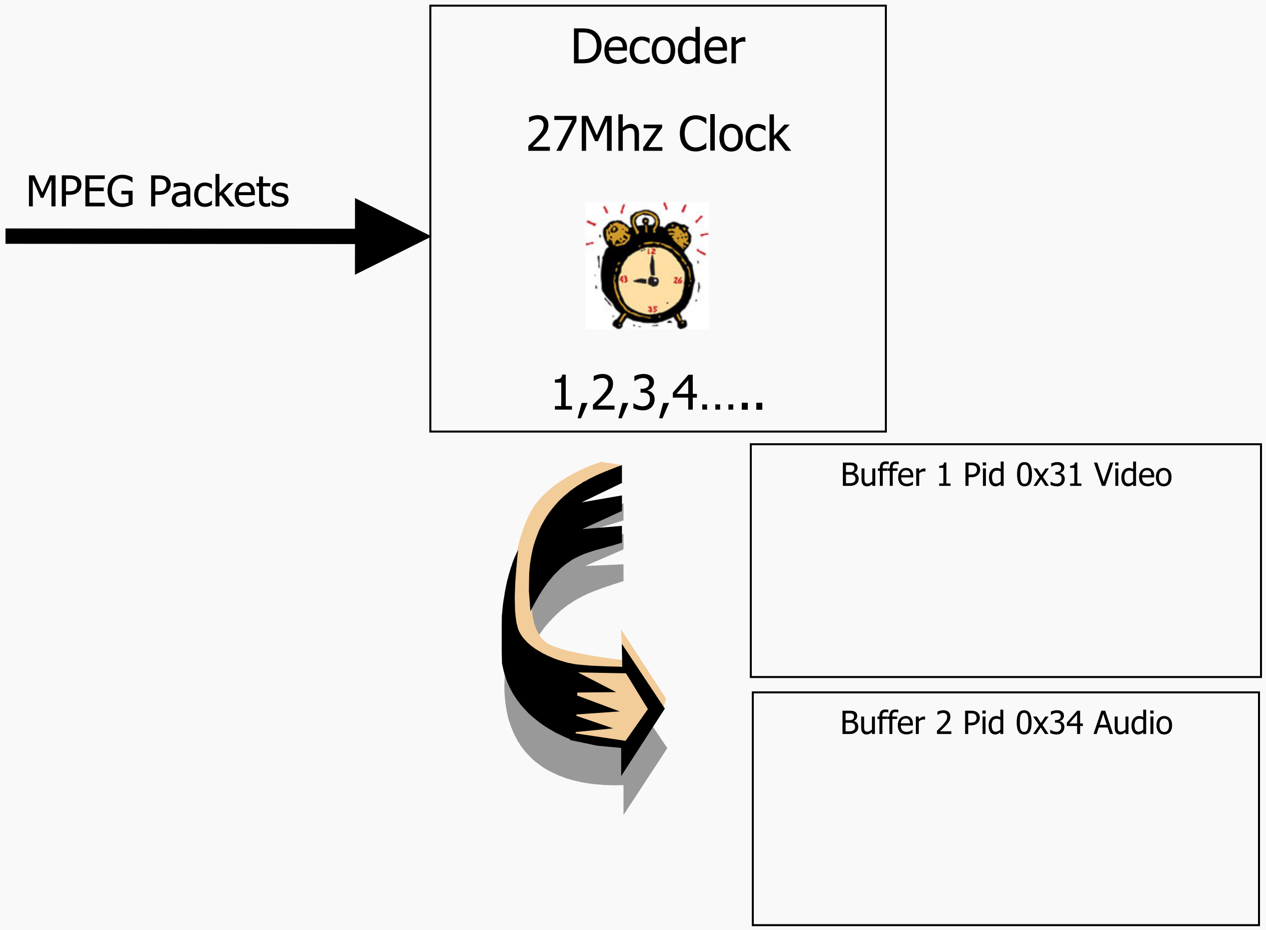
Presentation Time Stamp - PTS

- Each Frame is marked with a PTS – "Presentation Time Stamp" – a positive number
- The value of the PTS is set to the value of the Encoder Clock when the frame is encoded



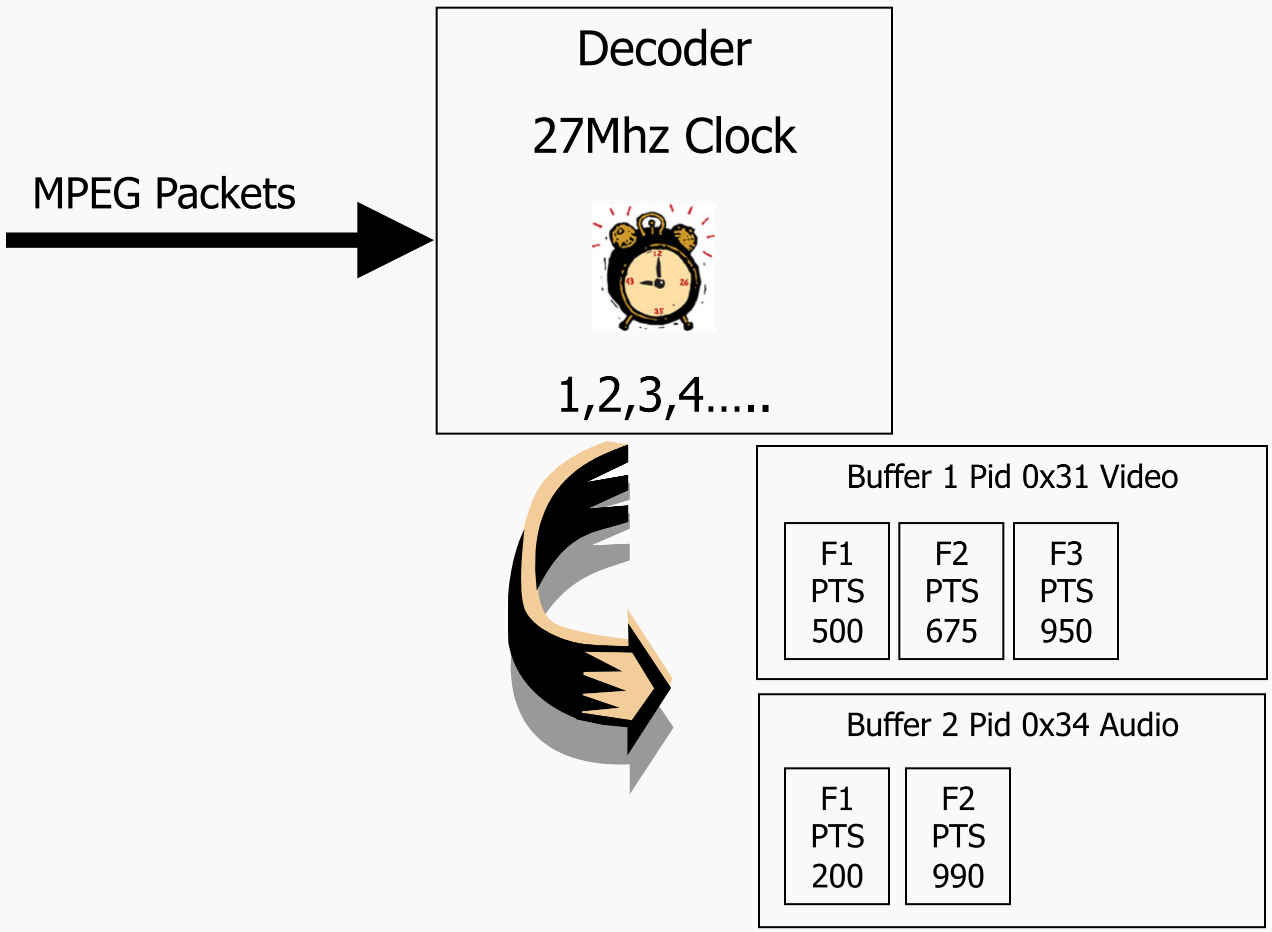
Packets Assigned to Decode Buffer

- As packets flow into the Decoder, a space in memory is set aside for them, one buffer for each PID.



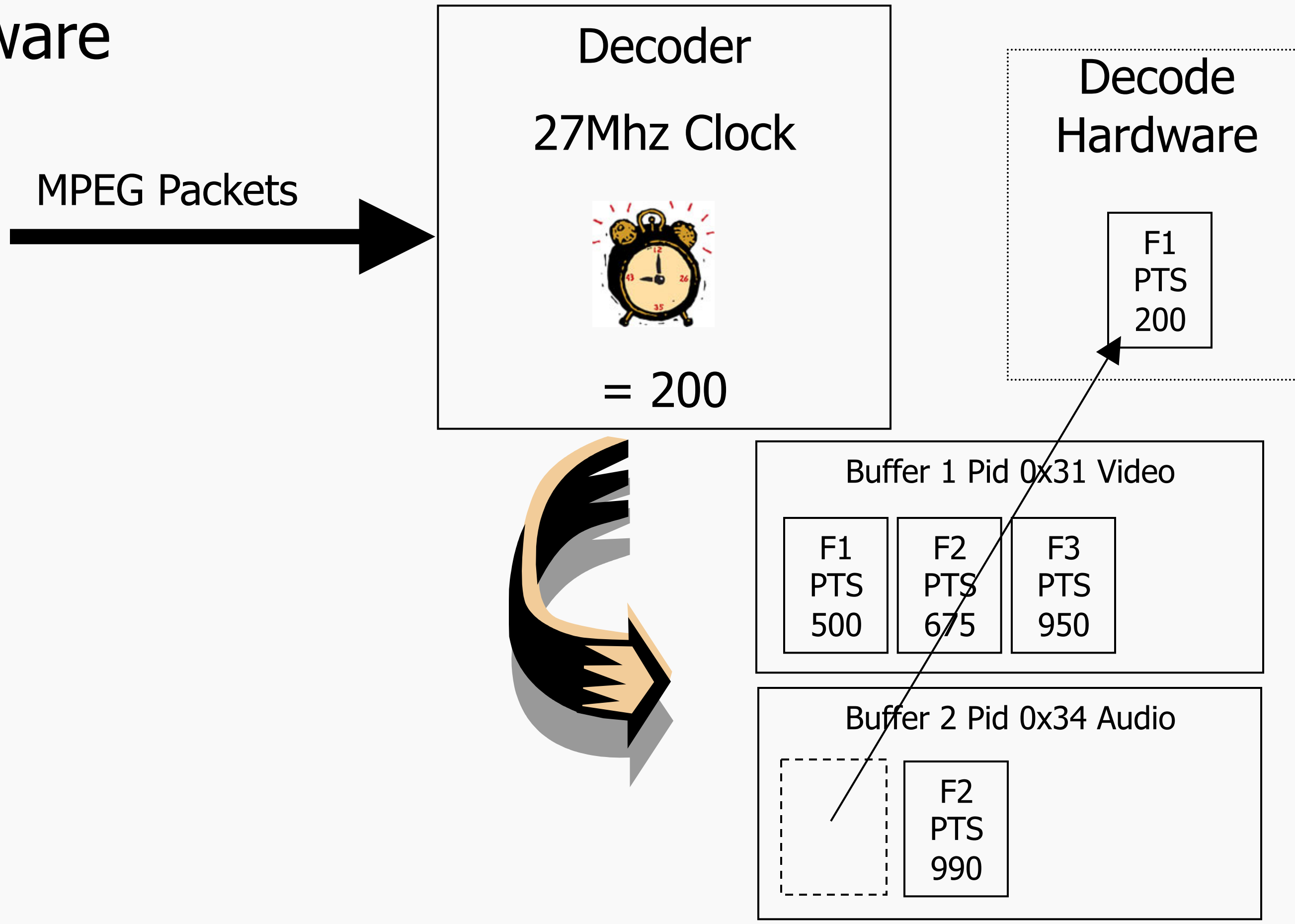
Reconstruction of Frames From Buffer

- Packets form Video and Audio Frames in the buffer



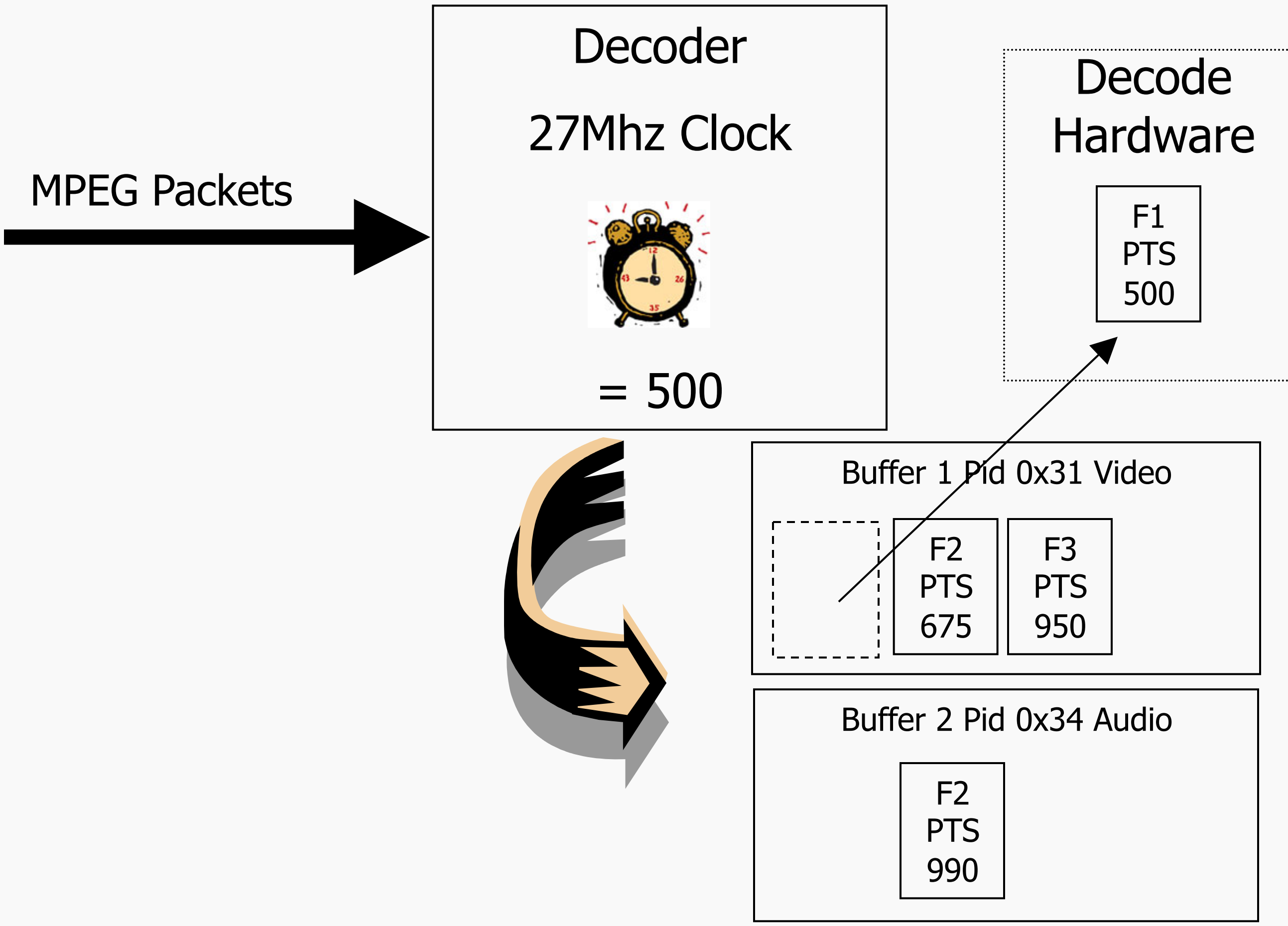
The Magic of Decode

- When the value of the Decode clock MATCHES the PTS on the frame, that frame is sent to the decode hardware



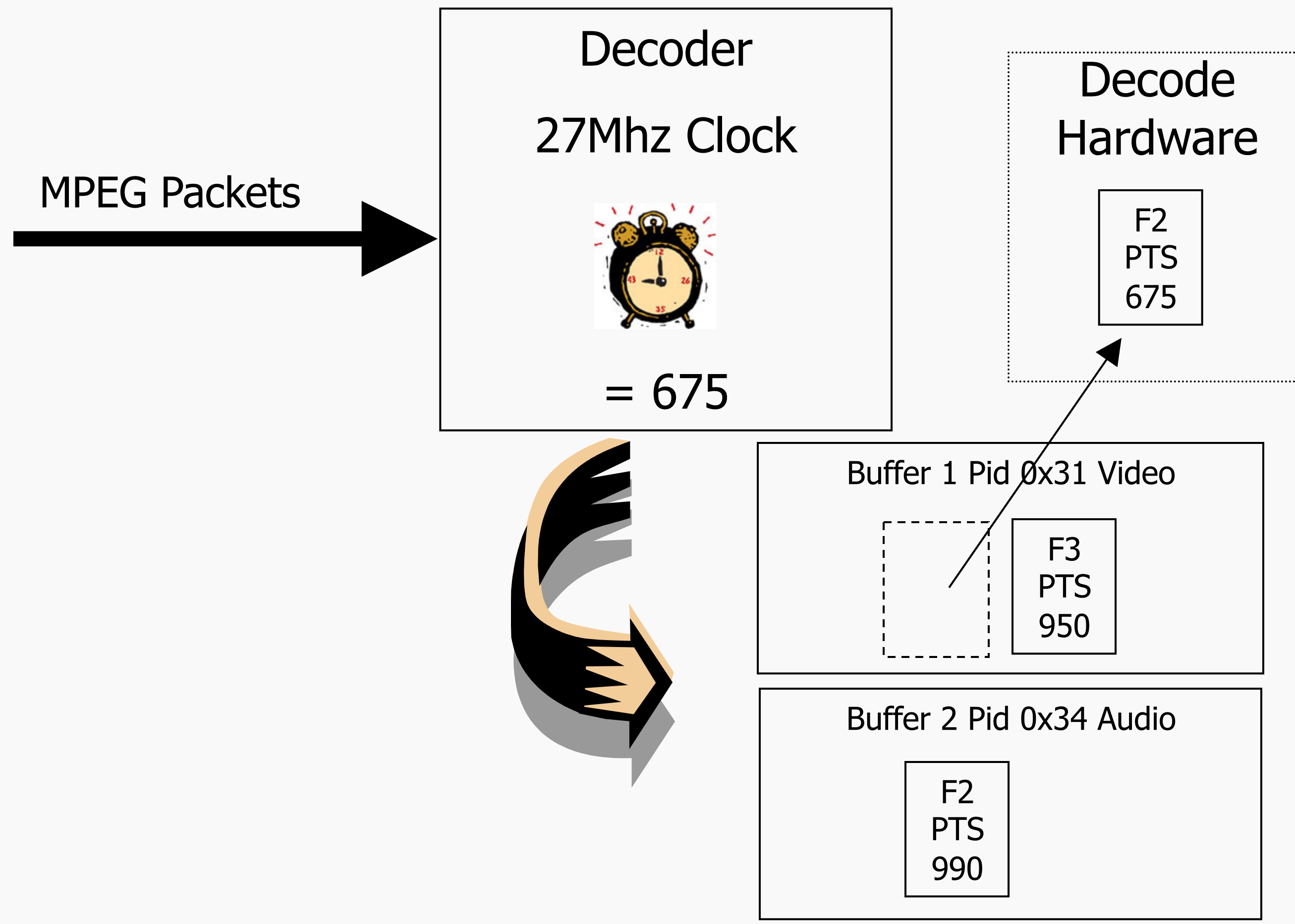
Another Frame Goes to Decode Hardware

- Next Frame



One More Frame Heads to Decode

- And the next frame…



Summary: Audio/Video Sync

- PCR values allow the Encoder Clock and the Decoder Clock to remain in sync
- PCR jitter can cause synchronization problems for elementary streams
- Ensure
  - PCR jitter and frequency offsets are within standard limits
  - Elementary stream buffers limits are NOT violated
- Large PCR jitter values can cause “Lip sync” error
- Buffer over- or underflow problems may cause “tiling”, “pixelization”/”macroblocking” errors

Extensions past MPEG-2 TS

- MPEG has defined a strict timing & buffer model
  - IT WORKS!!!!! (*if used at both ends*)
- Analogous model needs to be used for other transports – such as IP
  - Without established model, "broadcast" quality can not be assured
- Example: ATSC Mobile DTV used IP transport, not MPEG
  - Utilizes RTP over UDP
  - RTP timestamps relative to RTCP provides PTS equivalent
  - RTCP timelines normalized via reference to NTP stream
  - Clock resolution, jitter, drift… specified
    - Borrowed from MPEG-2 model

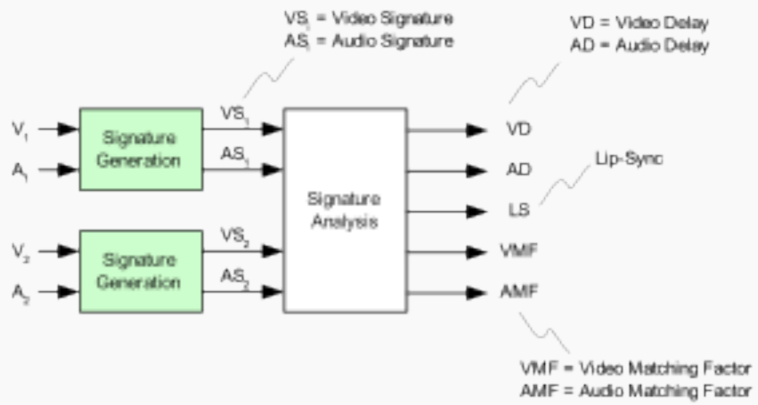
Lip Sync – what went wrong?

- MPEG has defined a necessary and sufficient timing & buffer model
- However, Lip Sync issues have been experienced in DTV systems worldwide

- MPEG & Other standards have defined "bits on the wire"
- None have said things like:
  - "The receiver shall continuously synchronize its clock using PCR"
  - "The receiver shall utilize transmitted PTS values to determine when to display audio or video"
- Implementers may not understand how the system works & take shortcuts

Sidebar – SMPTE Lip Sync activity

- Detecting Lip Sync issues is difficult
- SMPTE is in the process of standardizing a detection system
  - Generating matching Audio and Video Signatures (Fingerprints) at a known good location
  - Signature matching downstream – are the same Audio and Video frames still aligned?



Agenda

- ✓ *DTV Overview*
- ✓ *Transport Basics*
- ✓ *Timing*
- ➢ **Coding**
- Loudness
- Future

Compression

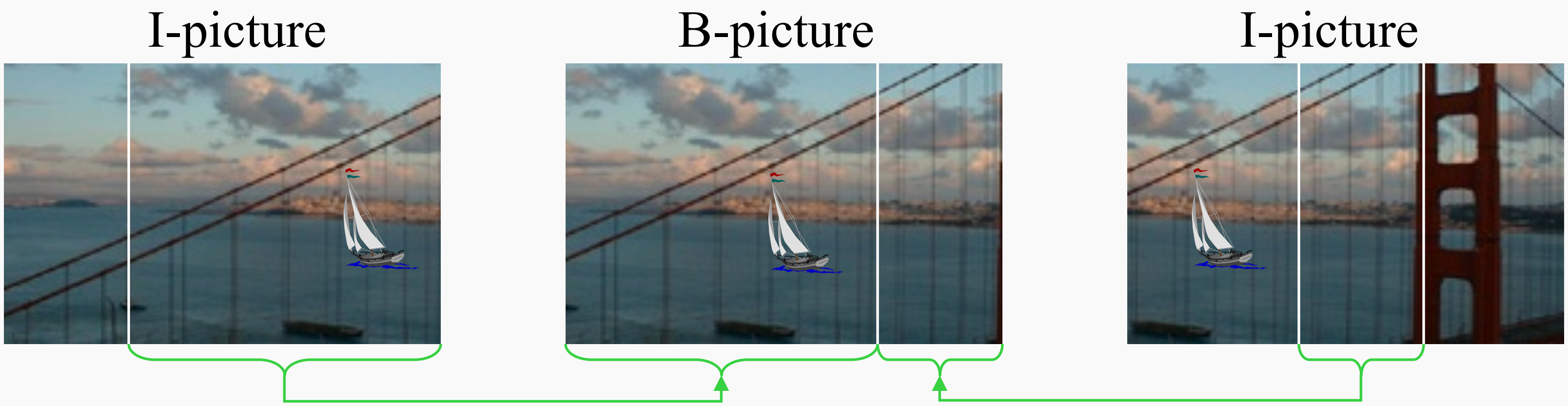
- Why Compress Video?
  - Sending raw Digital Video takes up very large bandwidth, even in SD
  - Full motion 525 line video needs over 200Mb/s without compression
  - It is not possible to send 200Mb/s down a 8MHz Pipe without further compression
- Use techniques that take advantage of human perception – throw away information that people won’t miss
- Utilize redundancies in “moving” pictures

MPEG-2 Video Encoding

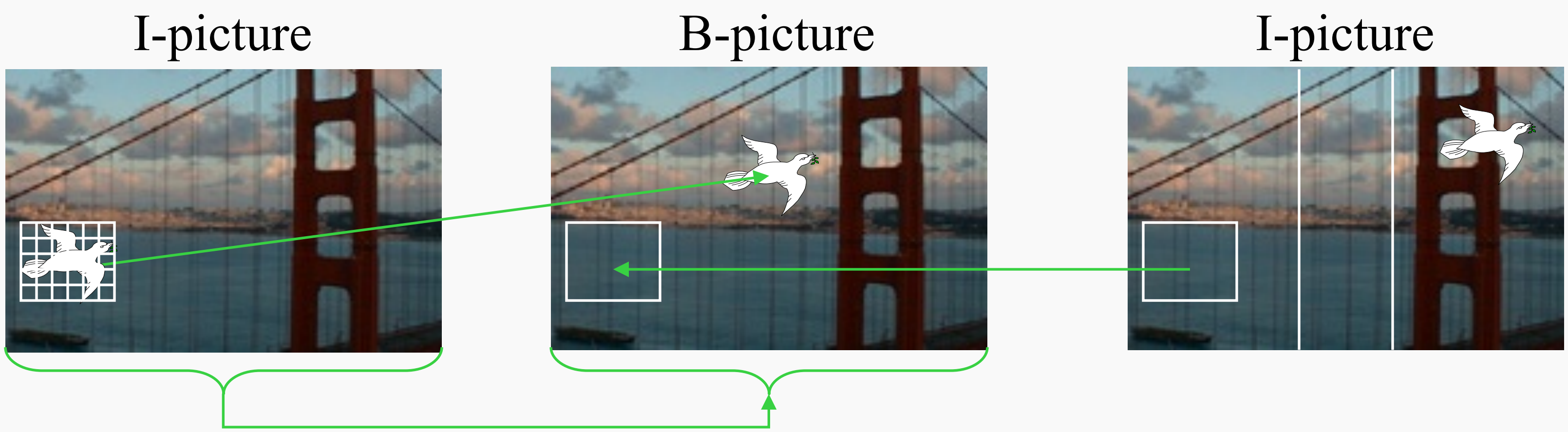
- Video is sequence of ***frames.***
- Each frame is encoded in one of three ways:
  - ***I-picture: intra-picture*** encoding, similar to jpeg encoding (exploiting spatial redundancy).
  - ***B-picture***: ***bi-directional*** encoding, using motion adjusted deltas from a previous and a future frame (exploiting temporal redundancy).
  - ***P-picture***: ***predictive*** encoding, using motion adjusted deltas from a previous reference frame (exploiting temporal redundancy).

MPEG-2 Video Encoding (Contd.)

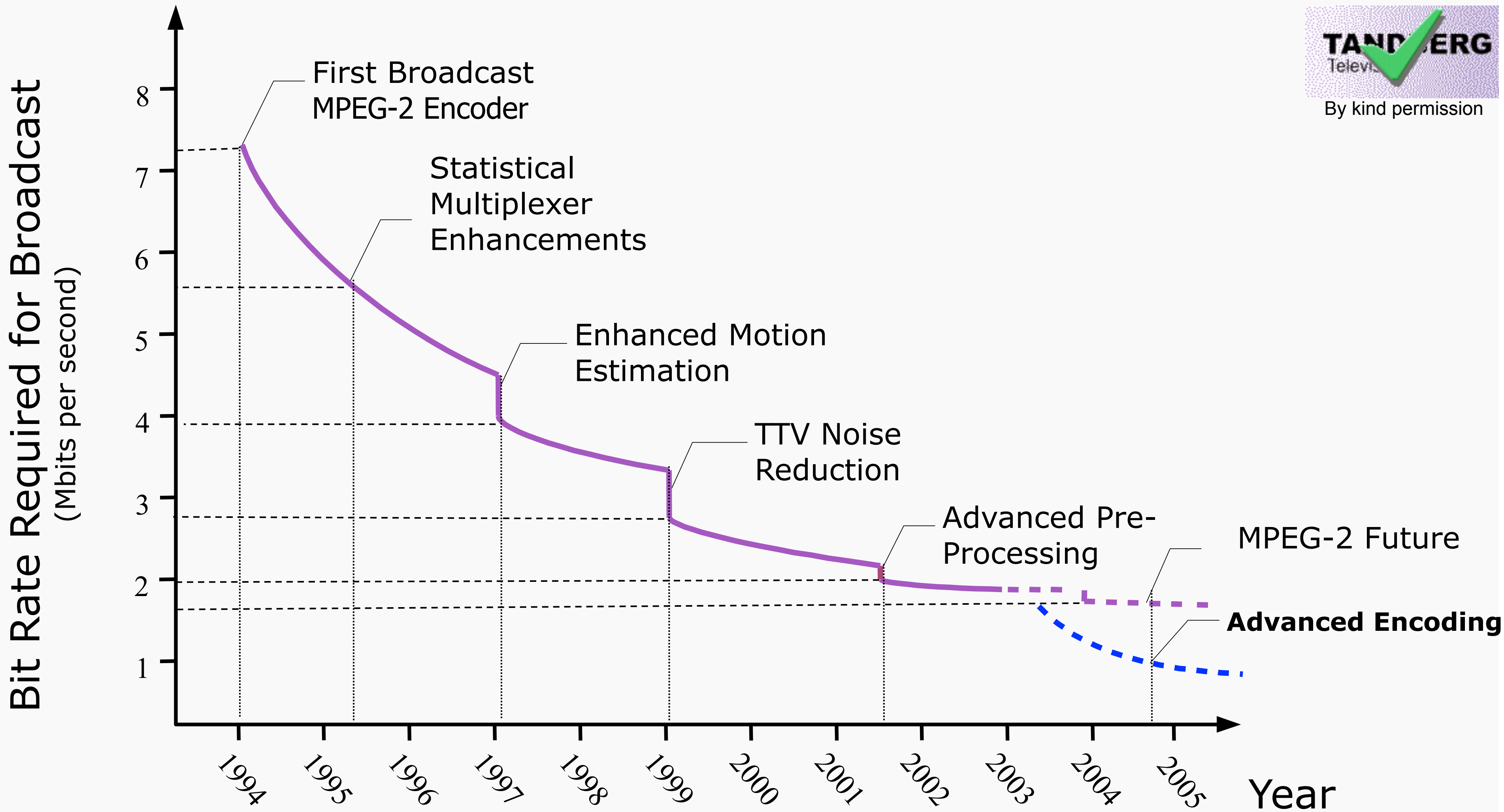
Example 1. Panning Camera



Example 2. Moving Object



Video coding trends - SD



MPEG-2 Video Basic Tools

- Take advantage of spatial/temporal redundancy & psychovisual redundancy
- Intraframe DCT coding
  - 8x8 pixel macroblocks, followed by:
  - Quantization (reduce number of values to be transmitted)
  - Entropy coding (use variable length codes to further reduce bits needed)
- Motion inter-frame prediction
  - Using model of decode, predict motion of blocks between frames
  - Convey only the motion vectors, not the pixels
- GOP structures
  - Choice of grouping of I, P & B frames
  - Adapt to content difficulty & quality desired

Tool Comparison: MPEG-2 vs. AVC (1 of 3)

Tool	*MPEG-2 Video (H.262)*	*MPEG-4 AVC (H.264)*
Intra Prediction	-None: MB encoded -DC predictors	- 4x4 Spatial - 16x16 Spatial - I_PCM
Picture Coding Type	- Frame - Field - Picture AFF	- Frame - Field - Picture AFF - MB AFF
Motion Compensation Block Size	- 16x16 - 16x8, 8x16	- 16x16 - 16x8, 8x16 - 8x8 - 8x4, 4x8 - 4x4

Tool Comparison: MPEG-2 vs. AVC (2 of 3)

Tool	*MPEG-2 Video (H.262)*	*MPEG-4 AVC (H.264)*
Motion Vector Precision	- Full Pel - Half Pel	- Full Pel - Half Pel - Quarter Pel
P Frame Feature	- Single Reference	- Single Reference - Multiple Reference
B Frame Feature	- 1 Reference Each Way	- 1 Reference Each Way - Multiple Reference - Direct & Spatial Direct Modes - Weighted Prediction

Tool Comparison: MPEG-2 vs. AVC (3 of 3)

Tool	*MPEG-2 Video (H.262)*	*MPEG-4 AVC (H.264)*
In-Loop Filters	**- None**	**- De-Blocking**
Entropy Coding	**- VLC**	**- CAVLC** **- CABAC**
Transform	**- 8x8 DCT**	**- 4x4 Integer "DCT"** **- 8x8 Integer "DCT"**
Other	**- Quantization Scaling Matrices**	**- Quantization Scaling Matrices**

Tool comparison: AVC vs. HEVC

AVC High Profile	**HEVC** "High Efficiency 10" Config	**HEVC** Main Profile
16x16 Macroblock	Coding Unit quadtree structure (64x64 down to 8x8)	
Partitions down to 4x4	Prediction Units (64x64 to 4x4) with asymmetric motion partition	Prediction Units (64x64 to 8x8) with asymmetric motion partition
8x8 and 4x4 transforms	Transform Units (32x32, 16x16, 8x8, 4x4 + non-square transforms)	Transform Units (32x32, 16x16, 8x8, 4x4)
Intra prediction (9 modes)	Intra prediction (35 modes)	
Inter prediction luma 6-tap + 2-tap to ¼ pel	Inter prediction luma 8-tap to ¼ pel	
Inter prediction chroma bi-linear interpolation	Inter prediction chroma 4-tap to 1/8 pel	
Motion vector prediction	Advanced motion vector prediction (spatial + temporal)	
8b/sample storage & output	10b/sample storage & output	8b/sample storage & output
In-loop deblocking filter	In-loop deblocking filter, Adaptive Loop Filter (ALF) & Sample Adaptive Offset (SAO) filter	In-loop deblocking filter & Sample Adaptive Offset (SAO) filter
CABAC or CAVLC (CAVLC = Context Adaptive Variable Length Coding)	CABAC with Wavefront Parallel Processing (CABAC = Context Adaptive Binary Arithmetic Coding)	

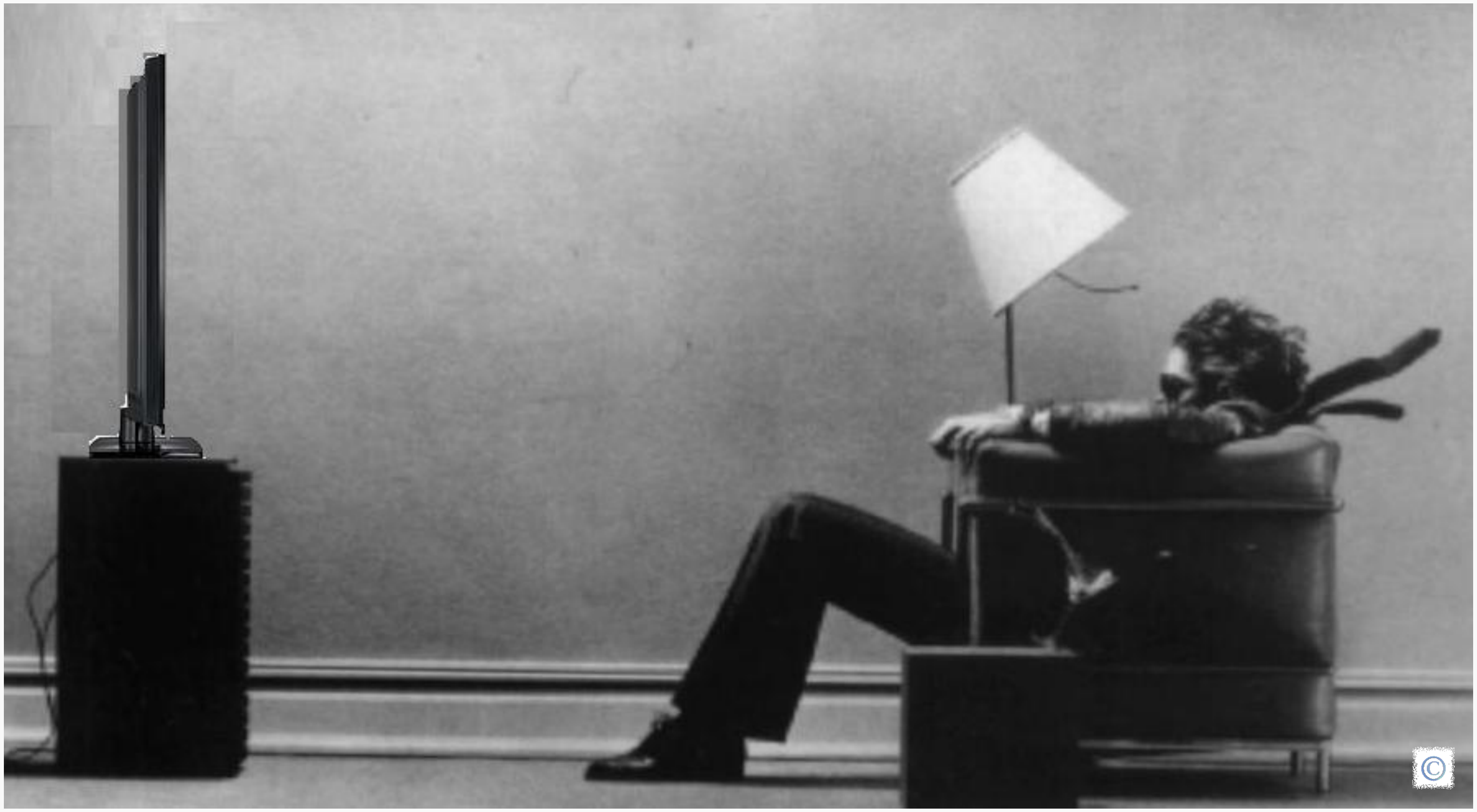
Agenda

- ✓ *DTV Overview*
- ✓ *Transport Basics*
- ✓ *Timing*
- ✓ *Coding*
- ➢ **Loudness**
- Future

Are we CALM yet?: Audio Loudness and its Side Effects

A Bit of CALM History (Commercial Advertising Loudness Mitigation)

- A member of US Congress was annoyed by commercials blaring loudly out of the TV!



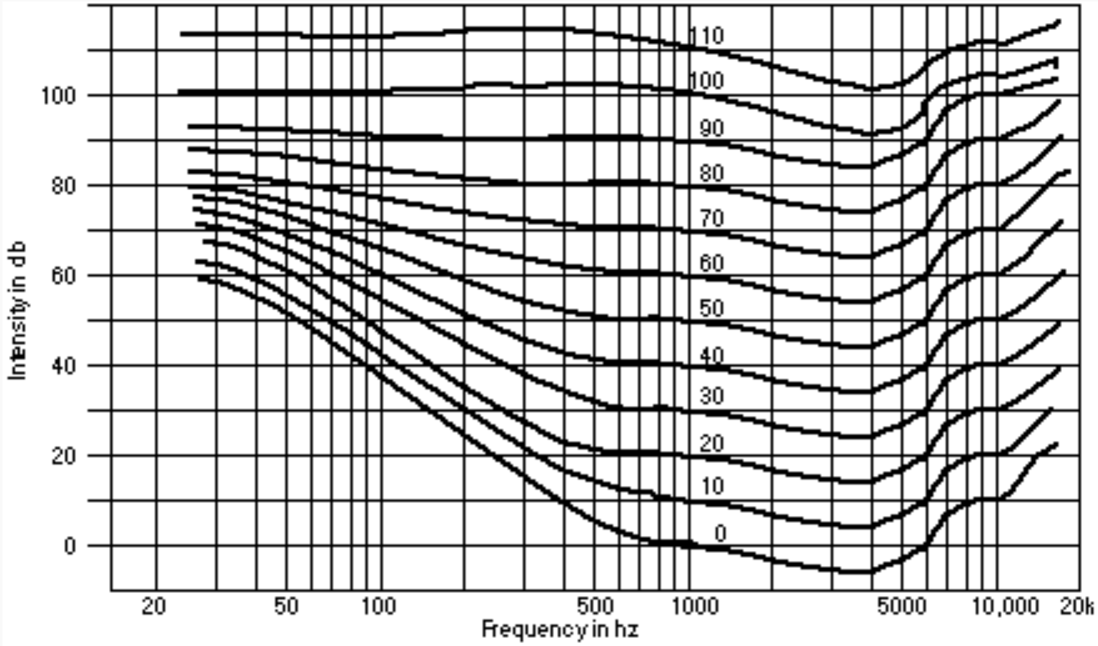
Audio Signals

- Audio is a very complex waveform (to think of it in its electrical representation)
  - Our ears typically hear a very wide frequency range (subsonic to almost ultrasonic – or .01 Hz through close to 20 kHz)
    - We would also feel the subsonic as well as hear it
  - While ears may lose that full range of detection over time, most humans start life able to hear like that
  - The dynamic range is enormous, thus the use of the decibel as the unit of measure
    - Bel is a diminutive of Bell, as in Alexander Graham Bell
  - Microphones typically deliver microvolts or less
  - Amplified signals may deliver volts to hectovolts
    - We are talking about a span from 10^{-6} to 10^{2} volts

Do we hear different frequencies at the same loudness?

- No
- Similar to our eye's heightened sensitivity to green and lack of sensitivity to blue, our ears have different sensitivities to different frequencies
- Measured first in the 1930's by two Bell Labs researchers, Fletcher and Munson
- They tested subjects wearing headphones, not thinking of the possible effects of the outer ear on loudness perception

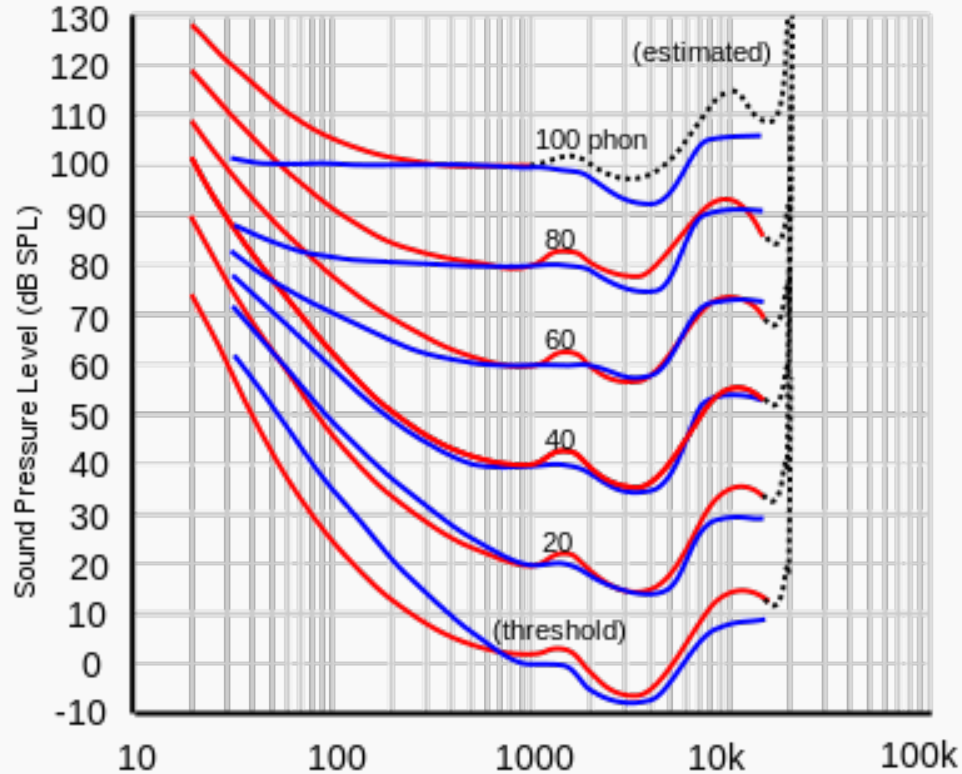
Human Ear's Frequency Response



Fletcher-Munson (1933) "Loudness, its definition, measurement and calculation," Journal of the Acoustic Society of America

These "curves of equal loudness" were measured by listeners wearing headphones

Fletcher-Munson Updated



Equal-loudness contours (red) (from ISO 226:2003 revision)
Fletcher–Munson curves shown (blue) for comparison

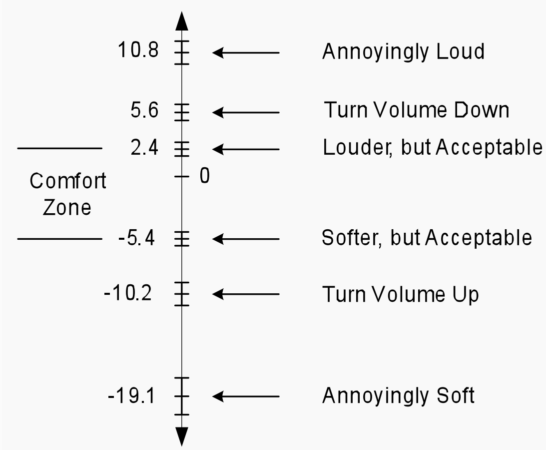
Subjects were tested in an anechoic chamber (free space)

The digital transition's (unintended) consequences

- Digital audio provides over 100 dB of dynamic range in sharp contrast to analog audio's less than 40 dB
  - has lead to widely divergent operational choices
    - Since the noise floor is no longer coupled to signal level
  - consumer frustration with those choices
- World-wide deployment of "all digital" content delivery systems
  - Compounded by the lack of industry standards for audio operating levels
- "That darn digital…"

Two Types of Listener Frustration

- Channel to channel level variation
- Intra-channel level variation
  - "Loud commercials"
  - Program to program variations
- Listeners can cope with variations, but only if they remain in the "comfort zone"
  - Comfortable variations range:
    - Up to 2.4 dB louder
    - Up to 5.4 dB softer
  - Uncomfortable, adjust volume range:
    - At 5.6 dB louder
    - At 10.2 dB softer



Relative Loudness(in dB) of the Listening levels investigated, with 95% confidence intervals

Riedmiller, Lyman and Robinson, AES115, 2003

Volume is not Loudness…

- These are two related but different parameters:
  - Volume - a measure of the levels of sound at any given instant in time
  - Loudness - a measure of the intensity with which we perceive a sound
- This is important because traditional audio metering only reports volume measurements
  - They were essentially voltmeters (with some dampening)
- Loudness measurements are NOT instantaneous!
  - Rather, they are time averaged
  - This remains a disconnect point for operators…

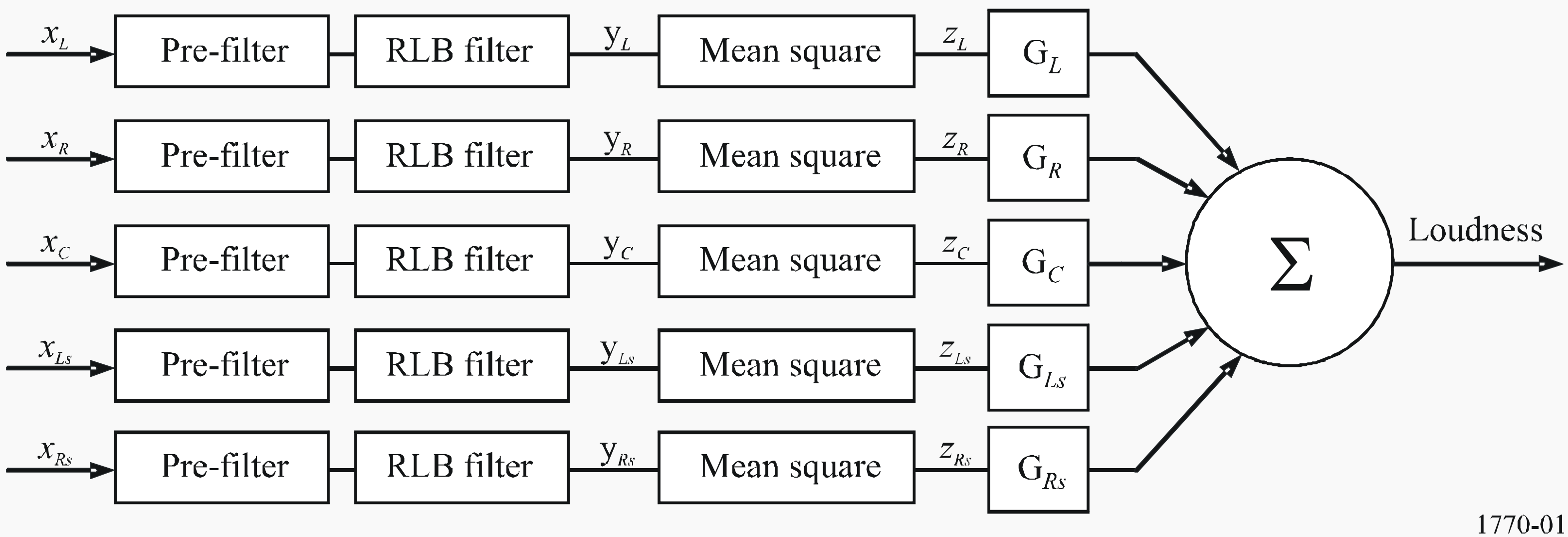
Volume vs. Loudness

- Traditional audio metering only reports volume measurements
  - The realization of this lead the ITU-R to develop BS.1770 which reports loudness, not volume
  - This is a time-average (long term) measurement
- The ITU worked on developing a "correct" loudness measurement method for over 15 years
  - This wasn't an obvious thing
  - Human hearing response isn't flat across all frequencies
- The ITU document also addresses measurement of "true peak" values

ITU-R BS.1770 (2007)

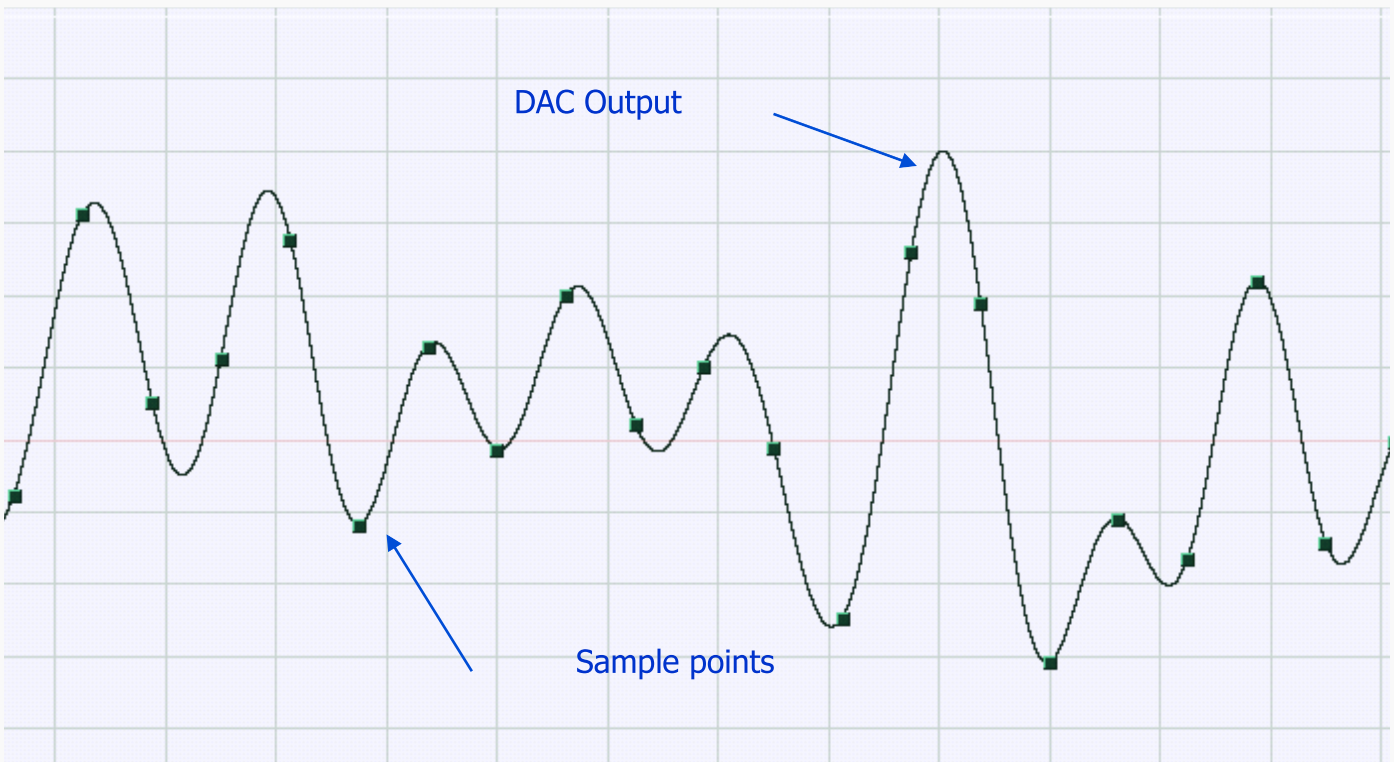
- First international standard designed to measure audio loudness, not industrial sound pressure levels
  - Documents two critical audio measurements
    - Objective multichannel loudness measurements
    - Accurate measurement of “true-peak” levels
  - This new method is a giant leap forward
- Earlier methods (ANSI S1.4-1983 “LAeq”) used industrial noise measurement techniques which did not closely track human hearing
- BS.1770 uses a measurement unit expressed as “LKFS” (Level, K-weighted, against Full Scale)
  - Excludes the LFE channel from a multi-channel measurement

The BS. 1770 loudness measurement block diagram

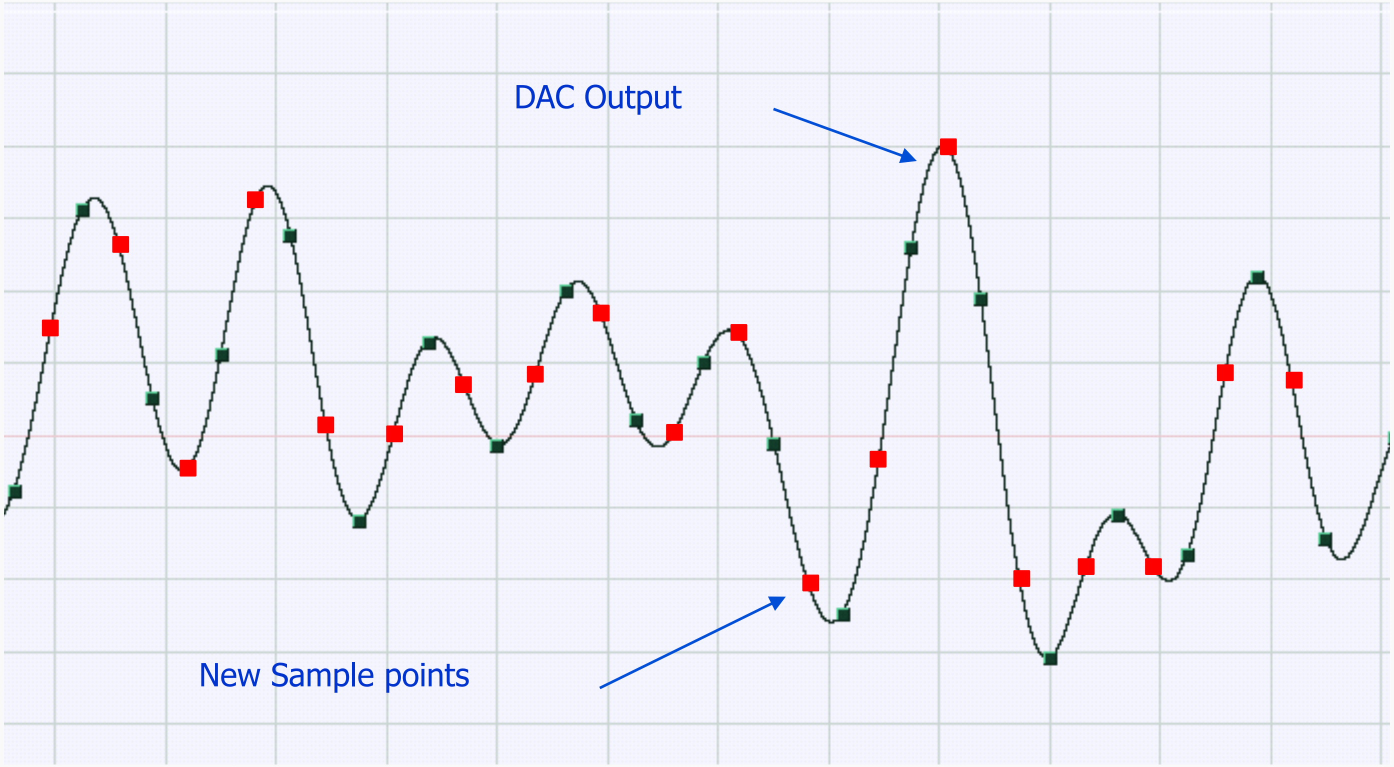


RECOMMENDATION ITU-R BS.1770 - Algorithms to measure audio programme loudness and true-peak audio level

Peak measurement in the Digital domain



Oversample to fill in the gaps and measure something closer to the true analog signal peaks



True Peak Measurements

- With reliable measurement of true peaks in the digital domain, operators can operate much more reliably in the upper range of digital levels
  - Many operators were running -30 dBFS or lower to “play it safe”
  - Digital clipping is audible and sounds really bad
- Reliable true peak measurements mean that a suggested operating point of -24 LKFS is safe
  - As suggested by A/85 and BS.1864

Developing A/85

- The ATSC Board of Directors (in 2006) recognized that:
  - "Audio implementation issues continue to dog the DTV transition, with widespread problems that tend to involve levels and time sync with video."
- Remember, this was at a point in the US digital transition where many broadcasters seemed to be ignoring their ATSC services
  - "dialnorm, what's that? why do I need to set it? and where in these complex controls can I do that?"
- Led to the creation of TSG/S6-3 AHG
  - They worked for 18 months reaching consensus on the technical issues and writing what became ATSC A/85

S6-3’s Mixing Test (“Is BS.1770 accurate?”)

- Tom Holman (USC) and Jim DeFilippis (Fox) also assembled a group of experienced audio mixers to conduct a “blind” listening test
  - Typical TV program audio material
  - Used one of the mixing rooms at USC
    - Meters covered
  - Each mixer, in turn, was asked to normalize the levels on the collection of content
    - Fader positions were saved
    - Fader positions were then converted to BS.1770 equivalent levels
- The results: the humans matched the BS.1770 algorithm's results
  - This gave everyone a good “gut feel” that indeed this is the tool to use for measuring program loudness

ATSC Recommended Practice A/85

- ATSC Recommended Practice A/85
  "Techniques for Establishing and Maintaining Audio Loudness for Digital Television"
  - Originally published 4 November 2009 , Annex J added 25 May 2011, Annex K added 25 July 2011
  - www.atsc.org
- ATSC A/85 summarizes
  - Despite the conclusion of the DTV transition, many broadcasters and the production community have been slow to effectively adapt to the changes required to transition from analog NTSC audio techniques to contemporary digital audio practices.
  - Consumers do not expect large changes in audio loudness from program to interstitials and from channel to channel.
    Inappropriate use of the available wide dynamic range has led to complaints from consumers and the need to keep their remote controls at hand to adjust the volume for their own listening comfort.
- ATSC A/85 relies upon the ITU-R BS.1770 loudness measurement method
- ATSC A/85 recommends the adoption of the ITU-R BS.1864 program interchange level of -24 LKFS

Who is the A/85 target audience?

- Primarily the content creation community
- Secondarily the distribution chain
  - TV stations
  - MVPDs

- Includes several "Quick Reference Guides" aimed at the "staff in the trenches"

Monitoring

- Film loudness is consistent because the mixing facilities and monitor levels are consistent
  - TV loudness is all over the map because TV mixing facilities and monitor levels are all over the map
  - Monitor at a consistent loudness to make programs with consistent loudness (DUH !)
  - A/85 Section 10 gives the basic design requirements for TV mix rooms and monitoring levels
- Rooms, SPL and Monitor Calibration
  - The larger the volume of the space, the higher the SPL for the same perceived loudness. (Table 10.2)
  - The ATSC web site has tone, pink noise and speech samples you can download (for free).
    - Use the 440 Hz tone to set unity gain through the system, up to the output of the mixing desk
    - Use the pink noise to set the sensitivity of each monitor amp to give the SPL shown in Table 10.2
    - The loudness of the speech sample is -24 LKFS. Use it to calibrate your ears.

Which edition of BS.1770?

- As is true of many Standards, the ITU-R has produced several versions of BS.1770
  - Original (version zero in 2005), Version one in 2007, Version two in 2011
  - A/85 normatively references BS.1770-1 (2007)
  - In March 2011, the ITU released BS.1770-2 (remember, A/85 was published in November 2009…)
    - This version added "gating" to the measurement algorithm (to exclude low level passages from the measured value)

What is new to BS.1770-2??

- The Revision of BS.1770-1 was driven by input from the EBU P/LOUD group
  - P/LOUD is comparable to S6-3 but without the due process that the ATSC brings
  - The group also included a number of manufacturers, each pushing their own technology (including Jünger)
  - The members who represented operators were also worried about having automatic systems that could run unattended
- This lead to the concept of “gating”
  - The measurement is stopped when the audio level falls below a certain threshold
  - Some folks wanted a fixed threshold, others a relative one

Gating

- First, the gates are specified in terms of measured levels AND in terms of 400 ms blocks (overlapping by 75%)
- There are two gate levels:
  - An absolute one at -70 LKFS
  - A relative to the currently measured level at -10 below that value
  - On startup the relative gate would not be applied

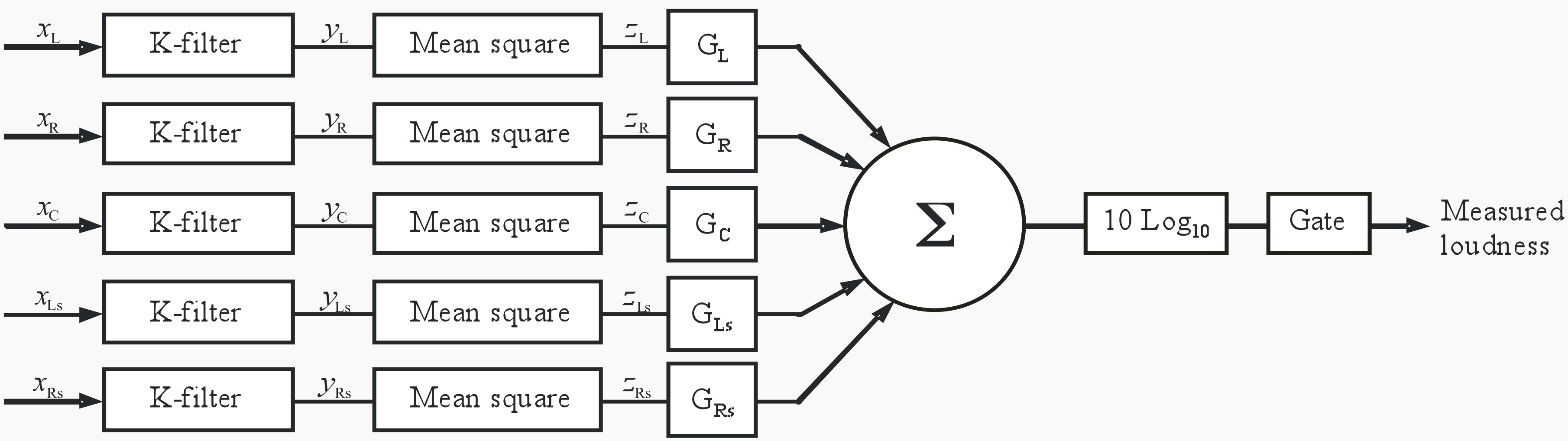
- The gate block is placed at the output of the overall measurement block

Reasons for Gating

- BS.1770-1 is an integrating measurement
  - Periods of silence or low-level signals may result in a result that is lower than what the perceived loudness should be
- The BS.1770-2 does not change any of the original algorithm other than adding a gating block after the channels are summed
  - The gating algorithm is meant as a refinement to BS.1770-1, and NOT meant as an anchor element/speech detection algorithm
  - Dolby has made the “Dialog Intelligence” algorithm public
    - But it is not standardized …

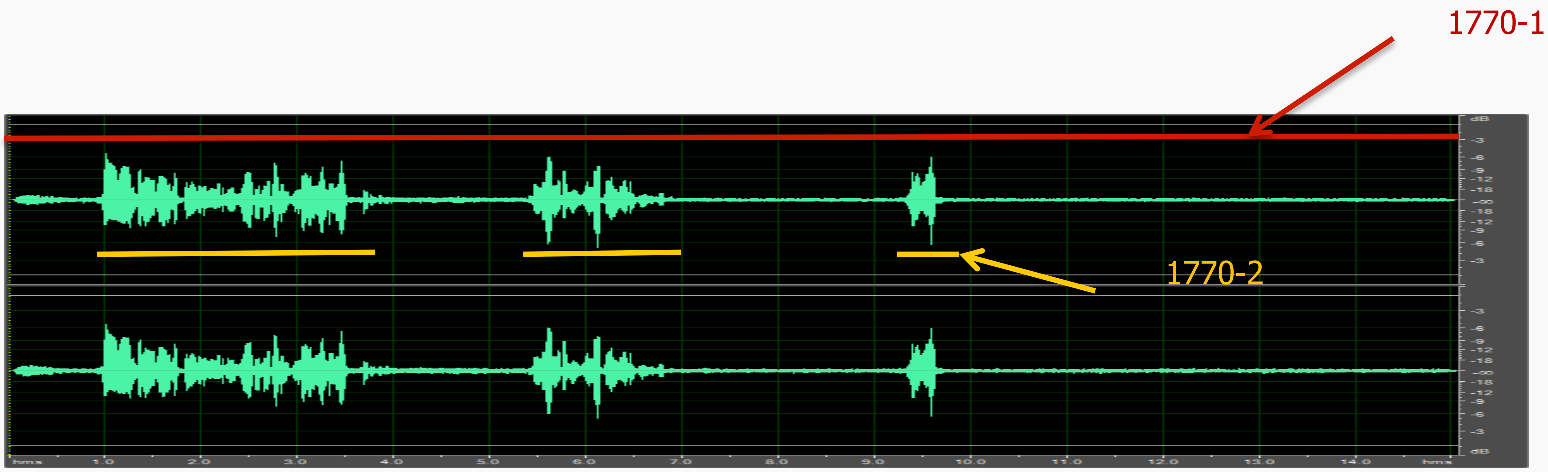
The 1770-2 Algorithm

- -70 LKFS absolute gate to remove silence or background noise
- -10 dB relative gate (relative to loudness with absolute gate applied) to determine foreground loudness
- Operates on 400 ms blocks of audio updated every 100 ms



Comparison of 1770-1 & 1770-2 Using Short-Form Content

- Per A/85, the entire content is measured



- 1770-1 = -27.2 LKFS, 1770-2 = -23.1 LKFS

- 4.1 dB difference (approx. 65% background noise)

The CALM Act

- A seemingly simple, short, and appealing Bill
  - Full name: "Commercial Advertisement Loudness Mitigation Act"
- Adopts (as "mandatory") an ATSC Recommended Practice: A/85
  - A/85:2011 "Techniques for Establishing and Maintaining Audio Loudness for Digital Television"
- Requires the FCC to adopt Rules within 1 year "… only insofar as such recommended practice concerns the transmission of commercial advertisements by …" (almost everybody)

- Unintended consequences:
  - Giving the ATSC the capability to produce "any successor thereto" is highly unusual and possibly might result in litigation aimed at the ATSC …
  - Not a position an SDO wants to be placed in! (and thus adds an extra level of deliberation regarding each proposed change to A/85)

CALM: The text

- Only 400 words long …
  - "Within 1 year after the date of enactment … the FCC … shall prescribe … a regulation … incorporating by reference and making mandatory … the RP … A/85, and any successor thereto …"
  - "… only insofar as such RP concerns the transmission of commercial advertisements by a television broadcast station, cable operator, or other multichannel video programming distributor. " [MVPD]
- Sec. 2(c) COMPLIANCE:
  - "Any … operator … that installs, utilizes, and maintains in a commercially reasonable manner the equipment and associated software in compliance with the regulations issued by the FCC in accordance with subsection (a) shall be deemed to be in compliance with such regulations"

New FCC Rules (TV station parallels MVPD words)

- Expects that program providers will “certify” their content
  - Assumes that operators want to (and will) “do the right thing”
  - Tracks public’s complaints, if any, to require deeper inspection
  - This section covers “embedded commercials” (inserted by network)
- The certification is widely available by website or other means
  - television broadcast station has no reason to believe that the certification is false
  - television broadcast station performs a spot check … response to an enforcement inquiry concerning a pattern or trend of complaints
- If transmitting any programming that is not certified
  - A television broadcast station … must perform annual spot checks … of all the non-certified commercial programming

Meanwhile, what's a content provider to do?

- If you haven't already, move all audio mixing from VU or PPM meters to BS.1770 loudness meters and True Peak meters
  - Make sure you can tell the meter "which version" [BS.1770-1 versus BS.1770-2]
  - Make sure your provider will keep the software current as both BS.1770 and A/85 evolve

- If you haven't already, make sure all your mix rooms are properly aligned per the A/85 recommendations
  - Make sure they stay that way!
  - While individual mix engineers may want to vary the standard level, this makes for problems for the staff who moves between rooms…
  - Even the "torture chambers" should be properly setup (at least headphone levels!)

The Golden Rule

- dialnorm shall equal your measured loudness in LKFS (this is from A/53 Part 5)
  - Per A/85 with a focus on the "anchor element" for long form
  - Per A/85 all channels for short form
- Understand that no one "downstream" can exactly replicate your readings!
  - They typically cannot isolate the anchor element
  - Depending upon the distribution of audio in the surrounds, their readings may be quite a bit off

Management of commercial loudness

- You can either
  - Ensure all ads to be inserted are measured and leveled to the dialnorm value of the channel into which they will be inserted

    or

  - Pre-compress them with a value of dialnorm equal to the measured loudness of the ad

- The choice depends upon your insertion infrastructure

- See A/85 Section 8

REMEMBER …

Agenda

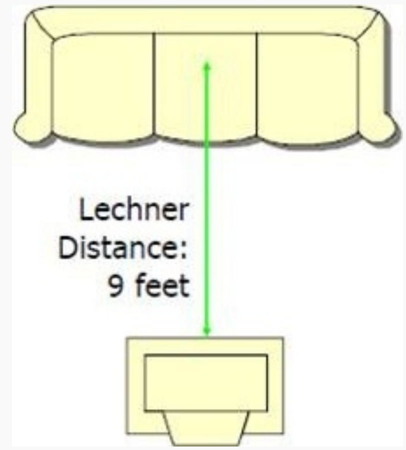
- ✓ *DTV Overview*
- ✓ *Transport Basics*
- ✓ *Timing*
- ✓ *Coding*
- ✓ *Loudness*
- ➢ **Future**

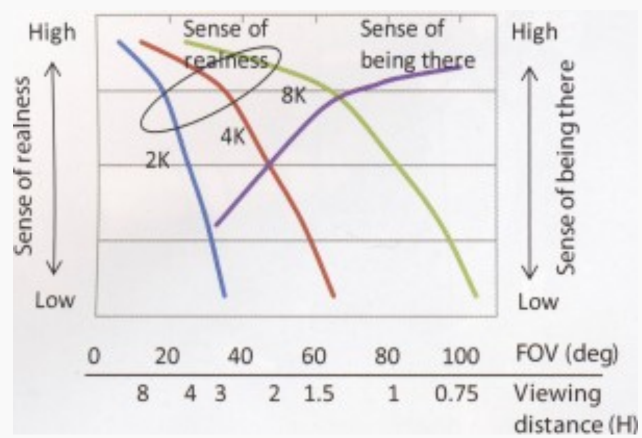
Future DTV – 4K and 8K

- "The killer app for TV is … more TV" – Ben Keen (iSuppli)
- Today:
  - SD 704x480
  - HD 1280x720 or 1920x1080 (known as "2K")
- Next:
  - 4K – 3480x2160 (UHDTV-1 (EBU))
  - Used in Digital Cinema today
  - Home use possible in near future (Consumer 4K displays on market – US 20-25K$)
- Followed by:
  - 8K – 7680x4320 (UHDTV-2 (EBU))

Why >2K?

- Most People watch TV at 3 meters
  - Lechner distance
  - Mostly due to room size/furniture
- "Sense of Realness" increases with decreasing viewing distance
- "Sense of Being There" increases with increasing viewing distance
- Academic debate over visual acuity, viewing conditions & distance:
  - Is higher resolution useful at home?
- Practical answer is probably "Yes"





Will UHDTV fit?

- HDMI 1.4 - 10.2Gbps at 340MHz clock rate
  - Will fit 4K UHDTV resolution with 10 or 12 bits of color @24 Frames per second
- Transport
  - Variables include color bit depth, 4:2:0 or 4:2:2 or 4:4:4 Y'CbCr Chroma subsampling schemes, wider color gamut, higher frame rates, higher screen refresh rates
- Current projections
  - 4K/10 bit ~10Mb/s HEVC
  - 8K/10 bit ~30Mb/s HEVC

NDS "Surfaces"

triveni DIGITAL
An LG Electronics Company

BTS IEEE Broadcast Technology Society

NDS "Surfaces"

ULTRA-HD

SUPER Hi-VISION

Glimpse the Future of Television

Comcast/NBCUniversal has partnered with NHK and the BBC to host the nation's only Ultra-HD/Super Hi-Vision demonstration for the 2012 Summer Olympic Games. Washington, D.C. is the sole North American site where this next-generation media technology can be viewed and the only location outside the UK and Japan. Ultra-HD features 16 times the resolution of today's HD and 3D audio, creating the feeling of sitting front row at the 2012 Games. This demonstration is part of a long standing tradition at both Comcast and NBCUniversal of leading the way in television technology.

Technical Specifications:

- Display: World's first Ultra-HD direct-view LCD display
- Quality: 16xHD (7680 x 4320 pixels)
- Aspect Ratio: 16x9
- Scanning: 60Hz progressive Target: 120Hz progressive
- Audio: 22.2 channel
- Viewing Angle: 100°
- Optimum Viewing Distance: 0.75 display height
- Data Rate: TS 280 Mbps (audio & video)
  TS 360 Mbps (including FEC packets)
- Retail: -2020

Comcast | NBCUniversal

ULTRA-HD
SUPER Hi-VISION

7680 PIXELS

HORIZONTAL VIEWING ANGLE OF 100°

4320 PIXELS

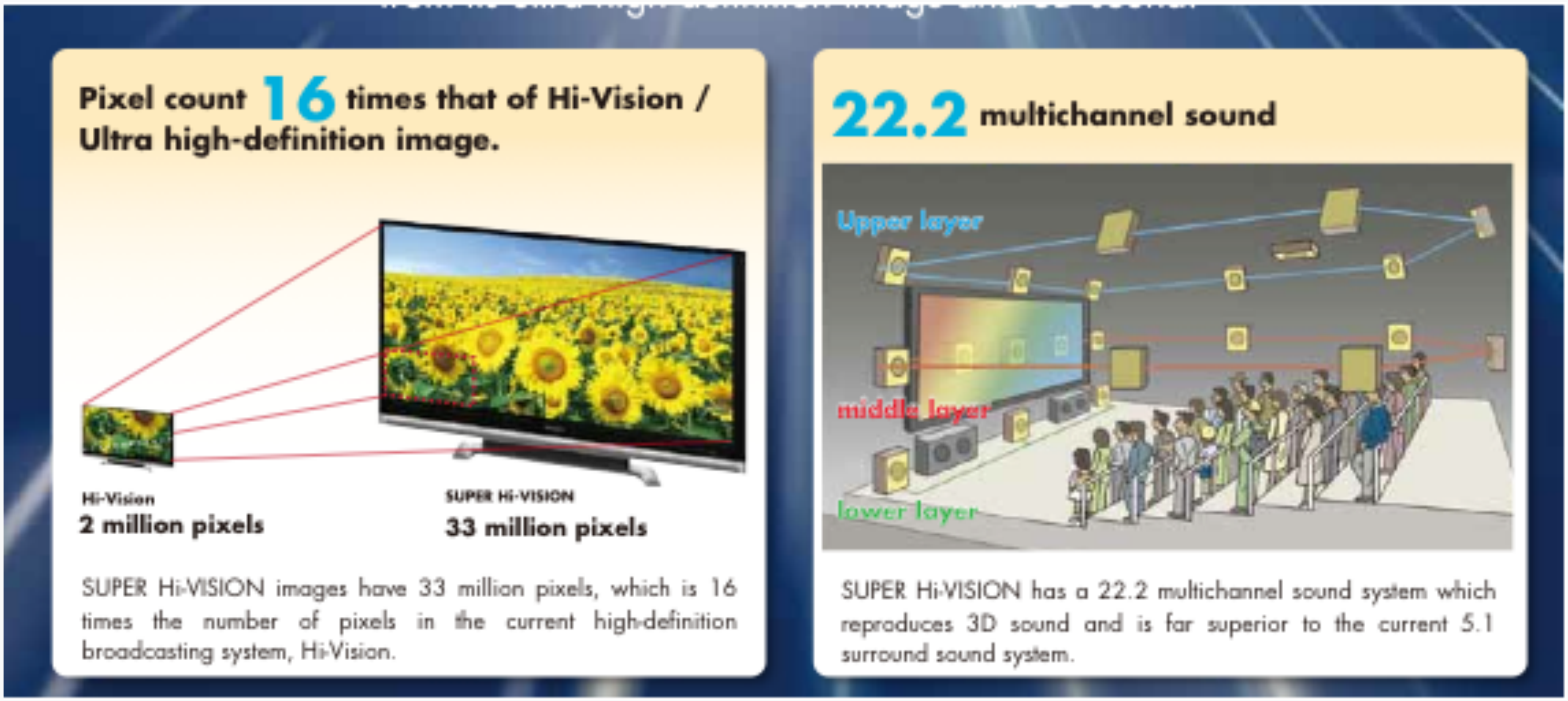
Glimpse the Future of Television

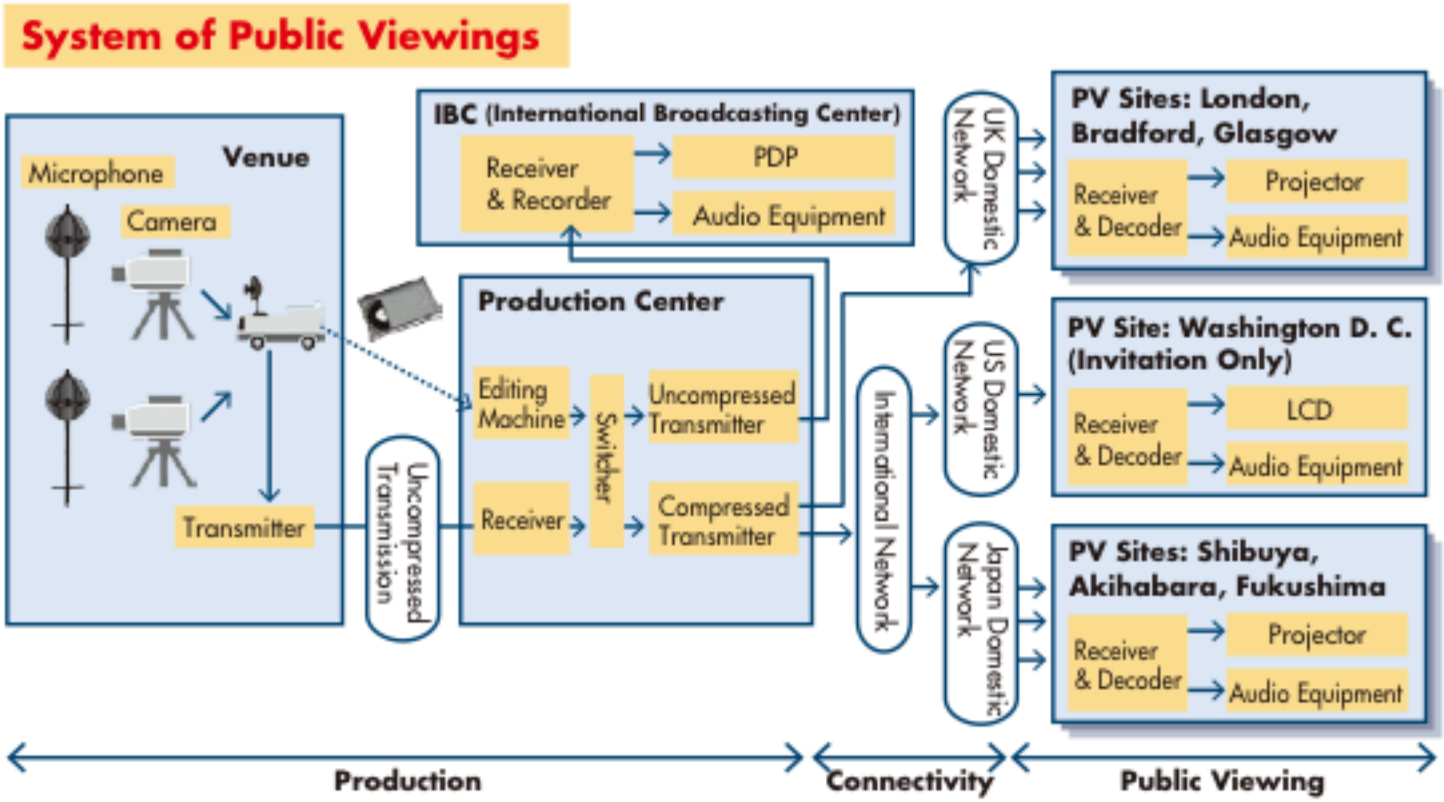
Comcast/NBCUniversal has partnered with NHK and the BBC to host the nation's only Ultra-HD/Super Hi-Vision demonstration for the 2012 Summer Olympic Games. Washington, D.C. is the sole North American site where this next-generation media technology can be viewed and the only location outside the UK and Japan. Ultra-HD features 16 times the resolution of today's HD and 3D audio, creating the feeling of sitting front row at the 2012 Games. This demonstration is part of a long standing tradition at both Comcast and NBCUniversal of leading the way in television technology.

Technical Specifications:

- Display: World's first Ultra-HD direct-view LCD display
- Quality: 16xHD (7680 x 4320 pixels)
- Aspect Ratio: 16x9
- Scanning: 60Hz progressive Target: 120Hz progressive
- Audio: 22.2 channel
- Viewing Angle: 100°
- Optimum Viewing Distance: 0.75 display height
- Data Rate: TS 280 Mbps (audio & video)
  TS 360 Mbps (including FEC packets)
- Retail: ~2020

Comcast | NBCUniversal


Pixel count 16 times that of Hi-Vision / Ultra high-definition image.
Hi-Vision
2 million pixels
SUPER Hi-VISION
33 million pixels
SUPER Hi-VISION images have 33 million pixels, which is 16 times the number of pixels in the current high-definition broadcasting system, Hi-Vision.
22.2 multichannel sound
Upper layer
middle layer
lower layer
SUPER Hi-VISION has a 22.2 multichannel sound system which reproduces 3D sound and is far superior to the current 5.1 surround sound system.



SHV camera

22.2 multichannel sound microphone

SHV compact projector

85-inch liquid crystal display (LCD)

145-inch plasma display panel (PDP)

Audience reaction to Olympics UHDTV

- "…most common reaction was in fact dead silence, with jaws slightly ajar…"

FOBTV

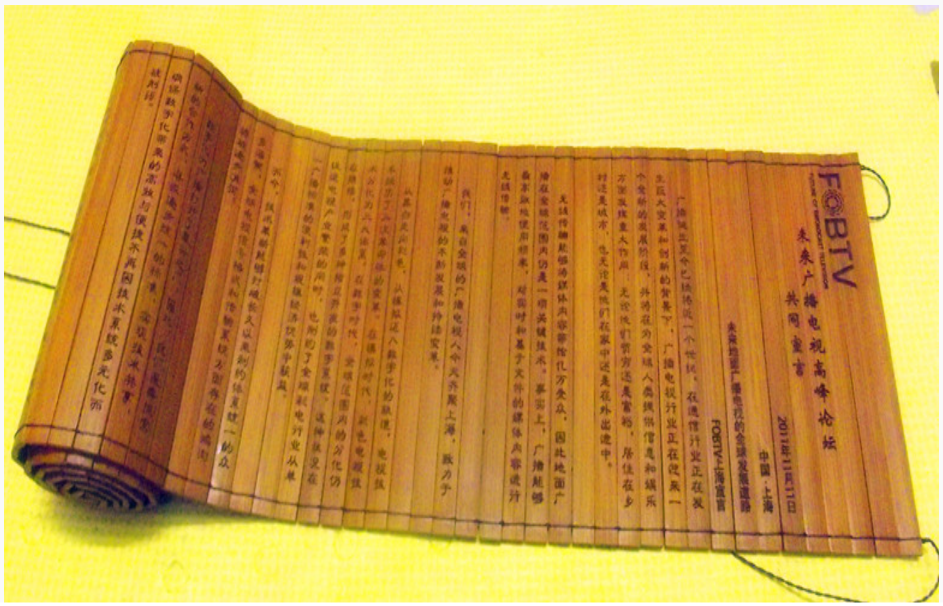
FUTURE OF BROADCAST TELEVISION

FOBTV Summit

- November 10 &11, 2011
- Shanghai, China
- Hosted by NERC-DTV

FOBTV
FUTURE OF BROADCAST TELEVISION

Joint Declaration on 11-11-11 at 11:11:11



Joint Declaration

- **Objectives**
  - Define the requirements of future terrestrial broadcast systems.
  - Explore unified terrestrial broadcast standards.
  - Promote global technology sharing

FOBTV
FUTURE OF BROADCAST TELEVISION

Joint Declaration

- **Why Broadcasting?**
  - Broadcasting is the most spectrum-efficient wireless delivery means for popular real-time and file-based media content
    - Infinite scalability: Wireless delivery of media content to an unlimited number of receivers makes terrestrial broadcasting a vital technology

Joint Declaration

- **Why a global standard?**
  - Mass production drives down cost
  - Show the commitment of the broadcasting industry to embrace new technologies and leverage interest from global technology developers to develop next generation products
  - Handheld and mobile devices travel across borders

FOBTV
FUTURE OF BROADCAST TELEVISION

Joint Declaration

- **Why now?**
  - Industries are driven by new technology
  - Spectrum is being sought after for broadband putting pressure on the broadcast industry
    - Next Generation Technology can be a part of the solution

FOBTV
FUTURE OF BROADCAST TELEVISION

Memorandum of Understanding (MOU)

- **April 2012**
  - MOU signed by thirteen founding organizations
    - ATSC, CBC,CRC DVB, EBU, ETRI, Globo-TV, IEEE-BTS, NAB, NERC-DTV, NHK-STRL, PBS, and SET

FOBTV
FUTURE OF BROADCAST TELEVISION

Goals of FOBTV

- Develop future ecosystem models for terrestrial broadcasting taking into account business, regulatory and technical environments.
- Develop requirements for next generation terrestrial broadcast systems
- Foster collaboration of development laboratories

FOBTV
FUTURE OF BROADCAST TELEVISION

Goals of FOBTV

- Select major technologies to be used as the basis for new standards
- Request standardization of selected technologies (layers) by existing SDOs such as ATSC, ARIB, DVB and TTA
  - FoBTV is not a standards development organization (SDO)

Structure of FOBTV

- **Management Committee**
  - Chairman: Mark Richer (ATSC)
  - Vice-Chairman: Phil Laven(DVB)
- **Technical Committee**
  - Chairman: Wenjun Zhang (NERC-DTV)
  - Vice Chairman: Yiyan Wu (CRC)
  - Vice Chairman: Namho Hur (ETRI)
  - Vice Chairman: Toru Kuroda (NHK)
- **Secretariat**
  - NERC-DTV

FUTURE OF BROADCAST TELEVISION

Founding Organizations

Participant List

FOBTV

FUTURE OF BROADCAST TELEVISION

Use Case Scenarios

- Ad Hoc Group Co-Chaired by Peter Siebert, DVB and Jim Kutzner, PBS
- Use cases compiled into preliminary summary document to be basis of Technical Commitee work

FOBTV Technical Committee

- Tasks:
  - Create a vision paper for a global end-to-end model for the future of broadcast television
  - Develop technical requirements based on use case scenarios
  - Issue Request for Proposals (RFP) based upon technical requirements
  - Evaluate proposals
  - Reporting the results of analysis to the MC
  - Recommend major technologies to be used as the basis for new standards
  - Develop minimum specifications of technologies recommended to be the basis for new standards
  - Recommend path for standardization

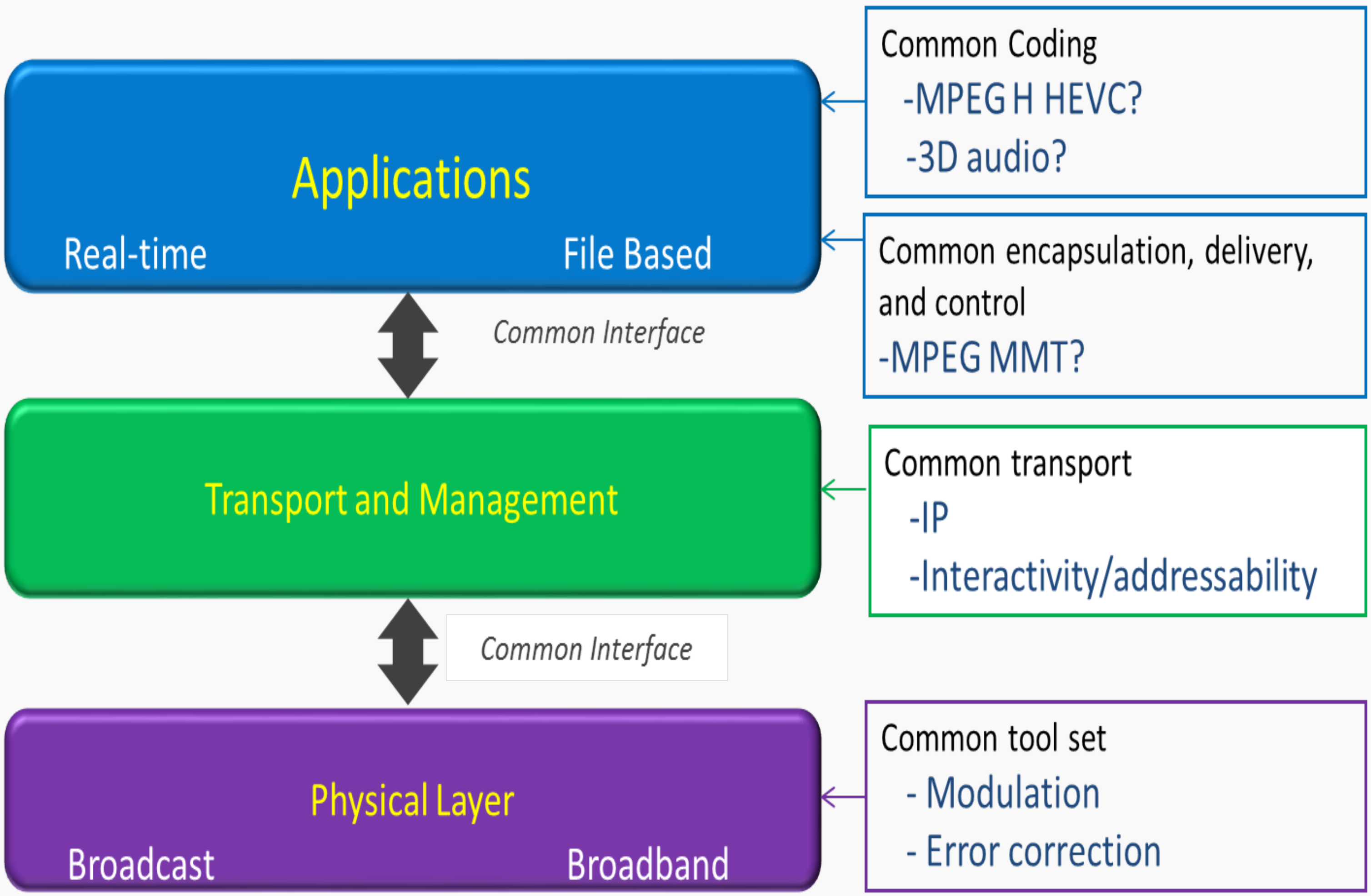
FOBTV
FUTURE OF BROADCAST TELEVISION

TC Working Groups

- Content Production and High Quality Service Management
- High Efficiency and Robustness Transmission
- Smart Screen and Data Layer Convergence
- Intelligent or Collaborative Technical Structure

Some thoughts
on
Next Generation
System requirements

Independent Layer Structure



System Requirements

- **Improved Coverage & Service**
  - Highly robust
  - Support various network architectures
    - High Power Central Transmitter
    - Single Frequency Networks
    - Hybrid

System Requirements

- **Efficiency**
  - Number of bits available for a given bandwidth (bits/hertz)
  - Improved compression technologies

Example Target: Minimum of 6x improvement compared to first generation systems

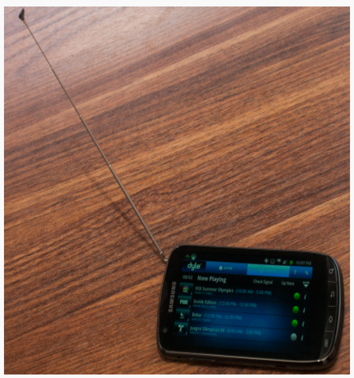
System Requirements

- **Flexibility/Scalability**
  - Bandwidth Utilization
    - Can not assume fixed bandwidth channels such as 5, 6, or 7MHz

Example Target: System dynamically configurable for .25 MHz to 25 MHz channels

System Requirements

- **Next Generation Systems should focus on broadcasting to devices that are on-the-move…**



System Requirements

- **Ultra High Definition**
  - 3840 x 2160 “4K”
  - 7680 x 4320 “8K”

System Requirements

- **Audio**
  - 22.2 Channel?
  - Metadata to support different listening environments
- **Accessibility**
  - Inherent support to meet needs of sensory impaired

System Requirements

- Define a **baseline receiver profile** for handheld, portable and mobile applications:
  - Intelligent receivers that can adapt to changes in transmission parameters
  - Software Definable Applications
    - Download Applications
    - Conditional Access, tools against piracy and authentication of applications
  - Seamless integration with Internet provided services
    - Multi-screen integration
  - Transparent to end user
- **Advanced receiver profiles** will process new services such as UHDTV
  - Will include baseline profile capabilities

Thanks

Rich Chernock

rchernock@trivenidigital.com

Special thanks to:
Matthew Goldman – Ericsson
Sheau Ng – NBC Universal
Pat Waddell – Harmonic
Mark Richer - ATSC