TUT SYSTEMS

Transforming Broadband into Services

Advanced Compression at the Edge

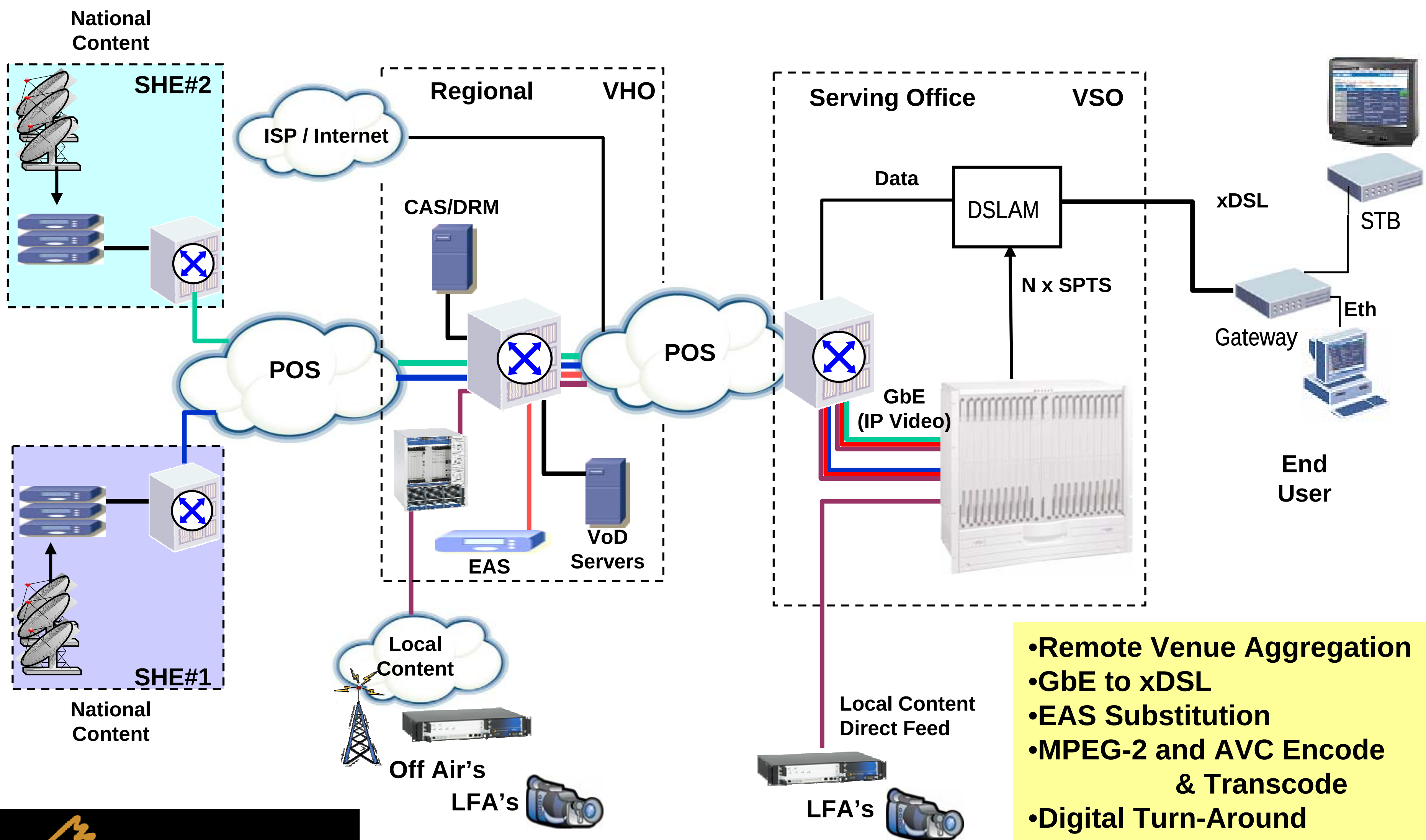
Chuck Van Dusen, CTO

MPEG Industry Forum: NAB April 17, 2005

Hierarchical architectures for IPTV

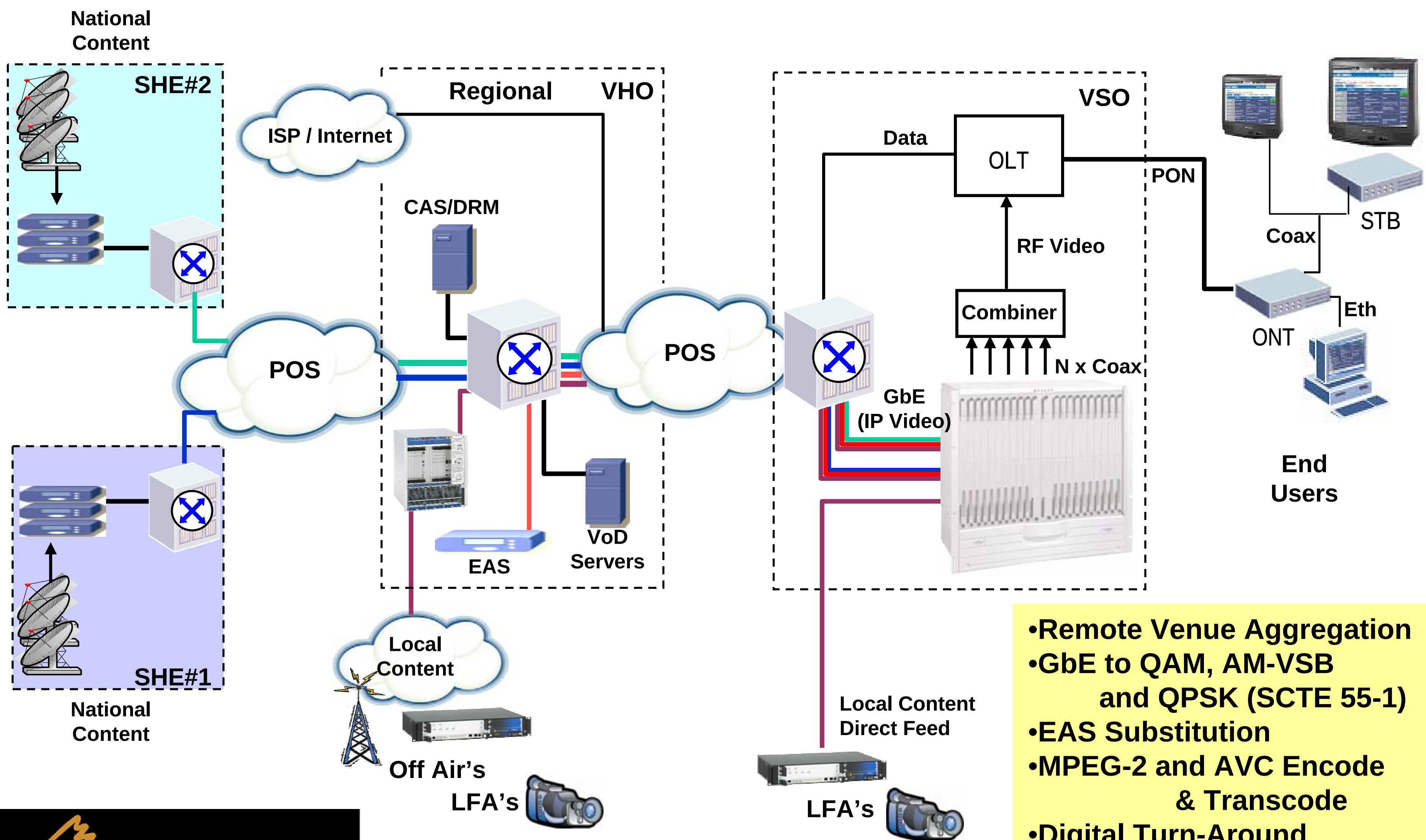
- **Each layer has compression requirements, including:**
  - **Prime encoding in the Super Head-End (SHE)**
  - **Local broadcasters in the Regional Head-End (VHO)**
  - **PEG content in the serving office (VSO)**
- **SHE, VHO and VSO can each have digital turn-around and transcoding requirements.**
- **Local program aggregation and insertion simplifies the backbone network.**
- **Local Franchise Authority may require multiple PEG channels. Several remote venues may be involved.**

Typical Architecture - IP Video to xDSL



- •Remote Venue Aggregation
- •GbE to xDSL
- •EAS Substitution
- •MPEG-2 and AVC Encode & Transcode
- •Digital Turn-Around
- •Ad Insertion

Typical Architecture - IP Video to RF Fiber



- •Remote Venue Aggregation
- •GbE to QAM, AM-VSB and QPSK (SCTE 55-1)
- •EAS Substitution
- •MPEG-2 and AVC Encode & Transcode
- •Digital Turn-Around
- •Ad Insertion

IPTV Edge Processors and Encoders

- **Encoders with IP interfaces can use existing data networks with simple FEC per Pro-MPEG COP3 to achieve economical and robust remote venue connections.**
- **Edge Processors with DC power and NEBS compliance meet the strict Telco requirements for the serving office .**
- **Both the encoders and edge processors can be managed by a single centralized EMS.**
- **Localization and regional branding of the service offering is simplified.**
- **A single Edge Processor platform can serve FTTH, ADSL, VDSL and HFC access networks simultaneously.**

Thank You !

See the Tut Systems
Encoding and Edge Processing solutions
at Booth # SU 9956

charles.vandusen@tutsys.com

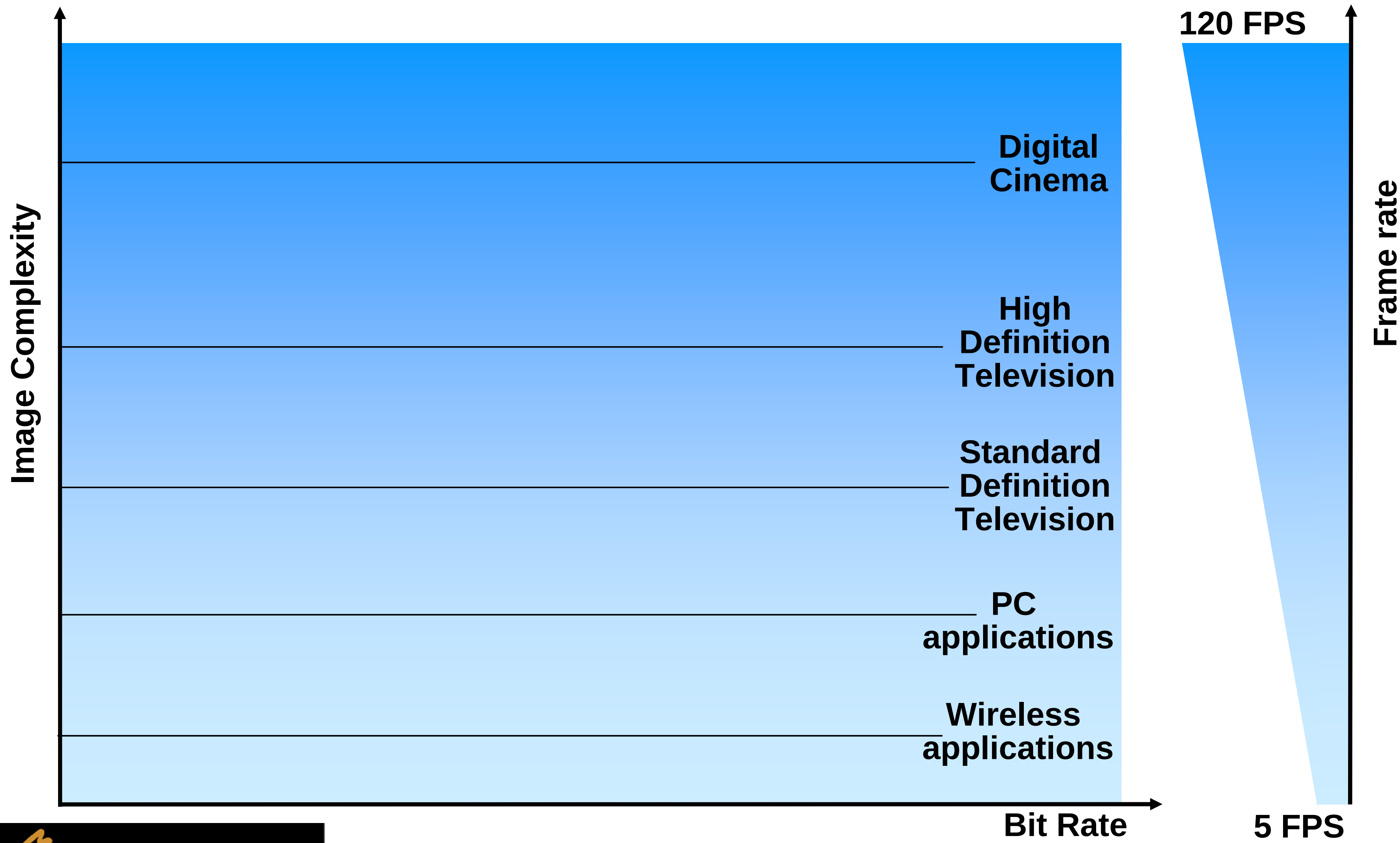
Additional slides for reference:

Advanced Compression and the Future of Digital TV

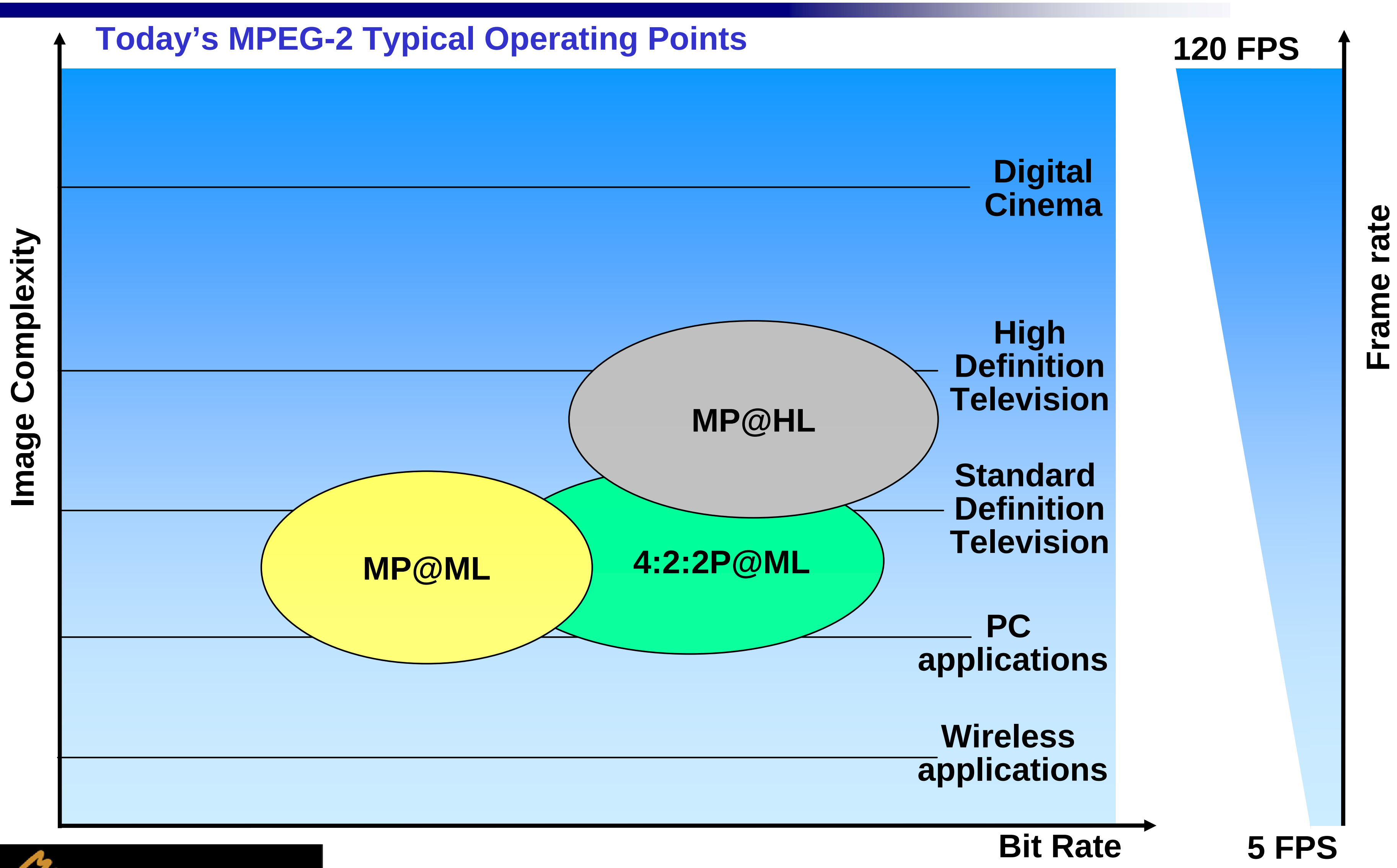
Chuck Van Dusen, CTO

TUT SYSTEMS

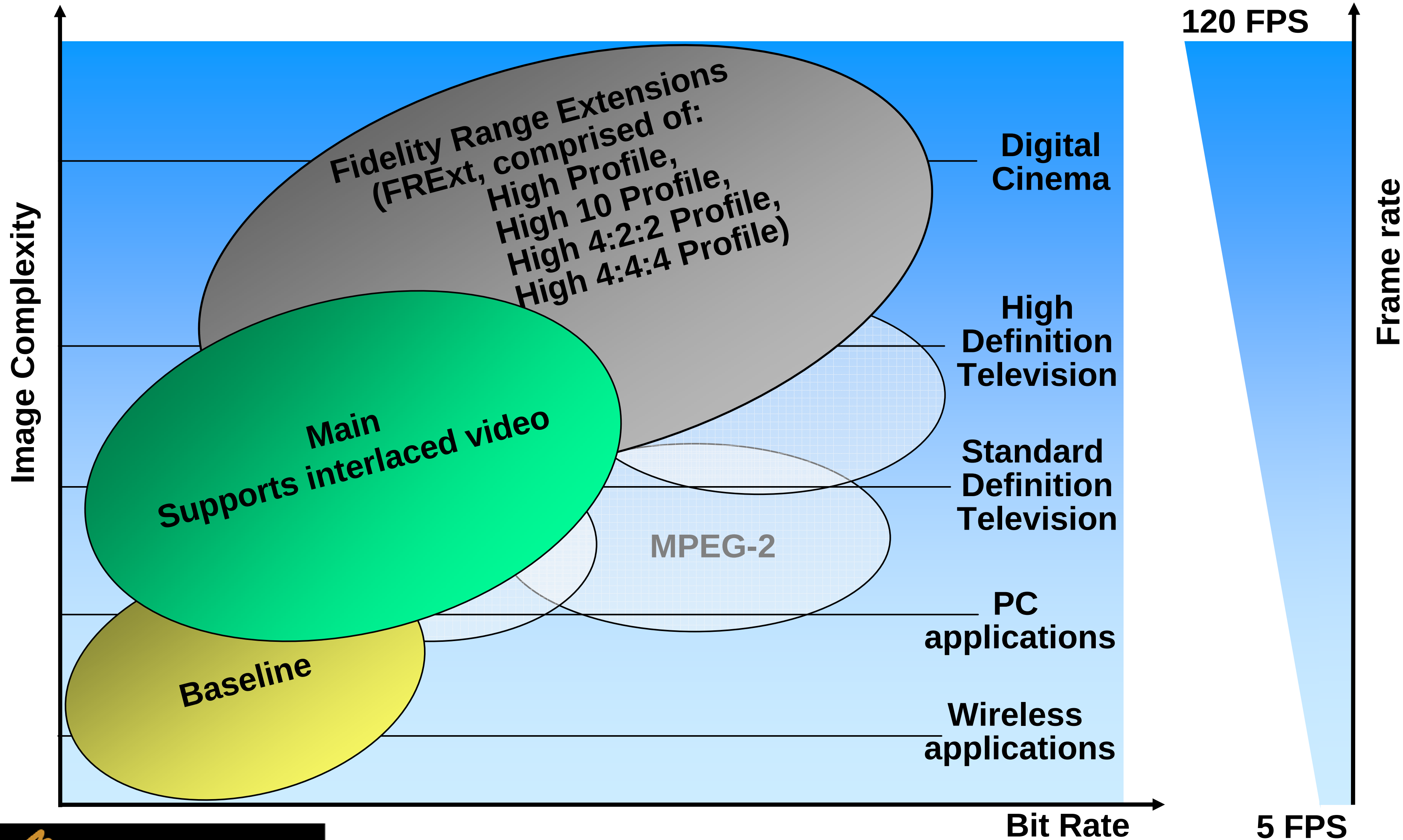
The Video Application Space



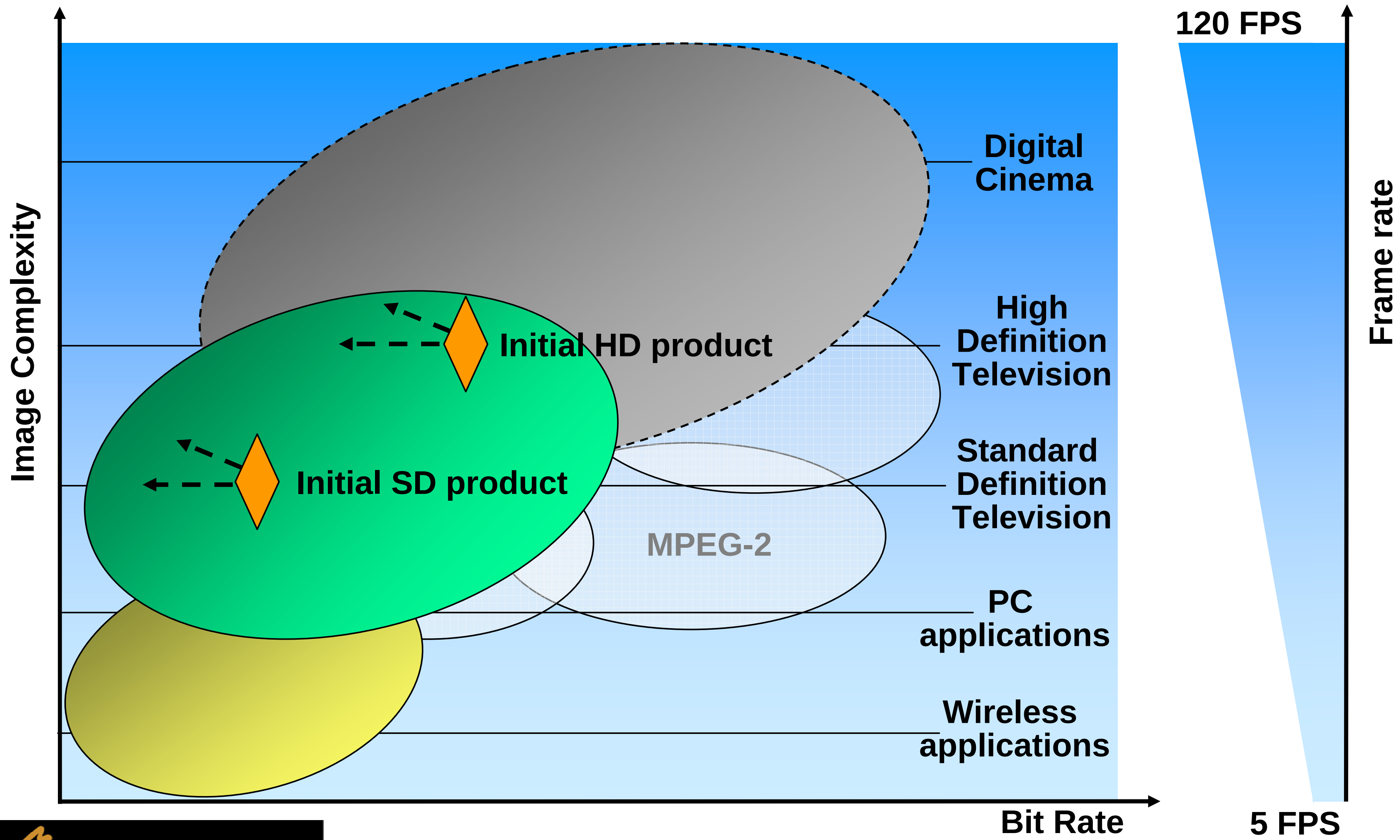
Processing Content with MPEG-2



H.264/MPEG-4 AVC Profiles

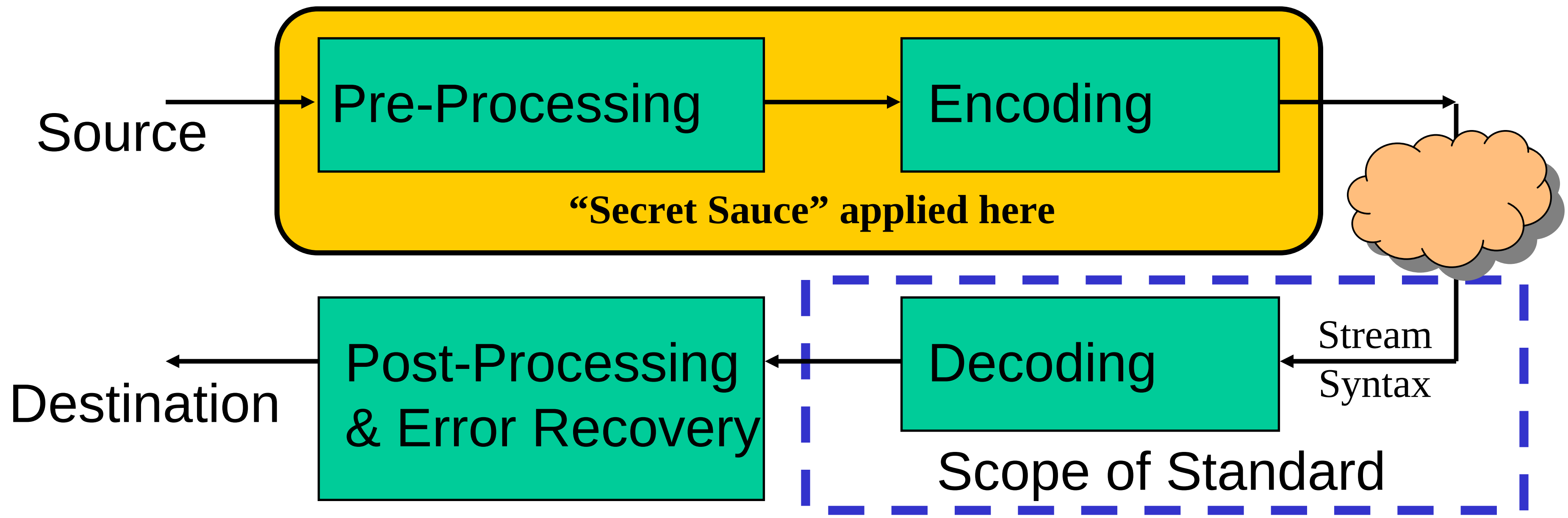


Tut Product Focus



The *Scope* of Coding Standardization

- **Only the *Syntax* and *Decoder* are standardized. This approach then-**
  - **Permits optimization beyond the obvious**
  - **Permits complexity reduction for assisting implementations**
  - **Allows improvements over time without change-out of CPE**
  - **Provides *no* guarantees of Quality**



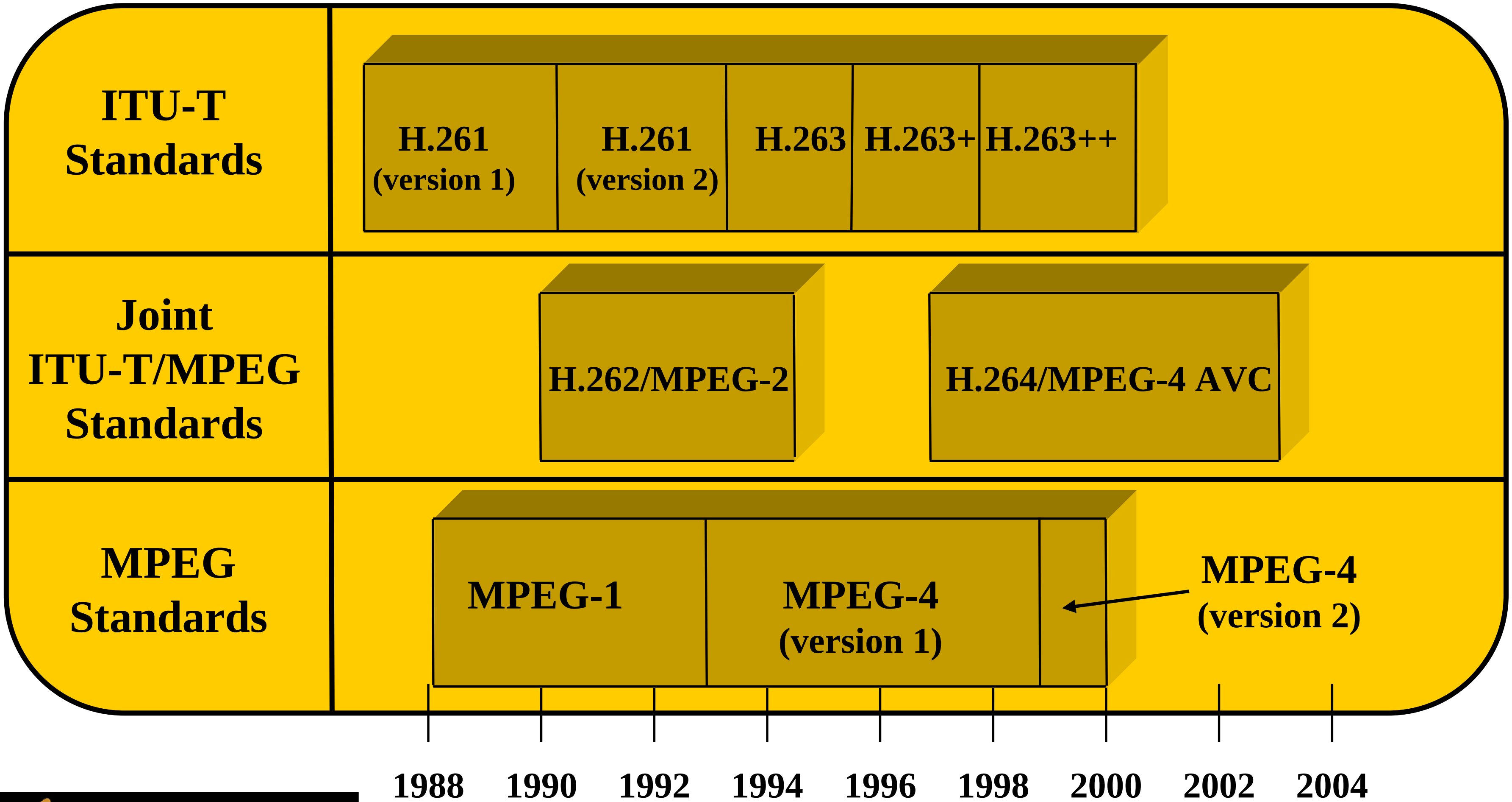
Advanced Encoding Toolsets

	MPEG-2	MPEG-4 AVC (H.264)	VC-1 (WM9)
Motion Vectors	cannot cross picture boundary	constant extension across picture boundary	constant extension across picture boundary
Vector Resolution	half-pel bi-linear	quarter-pel via 6-tap half-pel filter, then bilinear to 1/4	either half-pel bilinear or quarter-pel via 4-tap bicubic filters
Vector Blocks	16x16 or 16x8	16x16, 16x8, 8x16, 8x8, 8x4, 4x8, or 4x4	16x16 or 8x8
Vector Coding	vector predicted from one neighbor block	vector predicted from three neighbor blocks	vector predicted from three neighbor blocks
Vector References	P pictures have one frame or 2 field refs	P pictures have a choice of many reference frames or fields for each macroblock	P pictures have one frame or 2 field references
	B pictures have 2 frames or 4 field refs	B pictures have a choice of many reference frames or fields for each macroblock, 2 per vector	B pictures have 2 frames or 4 field refs
Macroblock skipped	P picture vector is 0,0 predict B pic vector	P picture vector is predicted B skip uses either spatial or temporal direct mode predict	P picture vector is predicted B skip uses temporal direct predict
B-prediction weighting	1/2, 1/2 of 2 refs	can do cross-fades with more general weights	Just 1/2, 1/2
B-pictures	used except in low-delay profile	used in Main & FRExt Profiles, not in Baseline Profile	used in both Main and Advanced profiles

Advanced Encoding Toolsets - continued

	MPEG-2	MPEG-4 AVC (H.264)	VC-1 (WM9)
Intra Prediction	predict DC from one neighbor block	multiple types of 16x16 and 4x4 spatial prediction	some intra prediction,
Block Transforms	8x8 DCT and IDCT	4x4 integer transform (and in new High Profile also 8x8)	choice of 8x8, 4x8, 8x4, and 4x4 integer transforms, optional overlap smoothing at intra block edges
Entropy Coding	simple variable-length coding VLC	choice of enhanced context adaptive VLC, or for even more bit savings (~10%) context adaptive arithmetic coding	context adaptive VLC; no arithmetic coding
Deblocking Filter	none	deblocking filter in-loop before the reconstructed is used as reference for another frame	deblocking filter in-loop before the reconstructed is used as reference for another frame
Post Decoding Filter	none specified	none specified	selectable post-decoder deblock and deringing filters
Rate Control	none	leaky bucket model to support VBR, CBR	leaky bucket model to support VBR, CBR
Interlace Support	both picture-adaptive and macroblock-adapt field/frame encoding in Main Profile	both picture-adaptive and macroblock-adaptive field/frame encoding in Main FRExt Profiles, not in Baseline	both picture-adaptive and macroblock-adaptive field/frame encoding in Advanced Profile, not in Main Profile

History of Related Standards



There are 18 Digital TV Formats

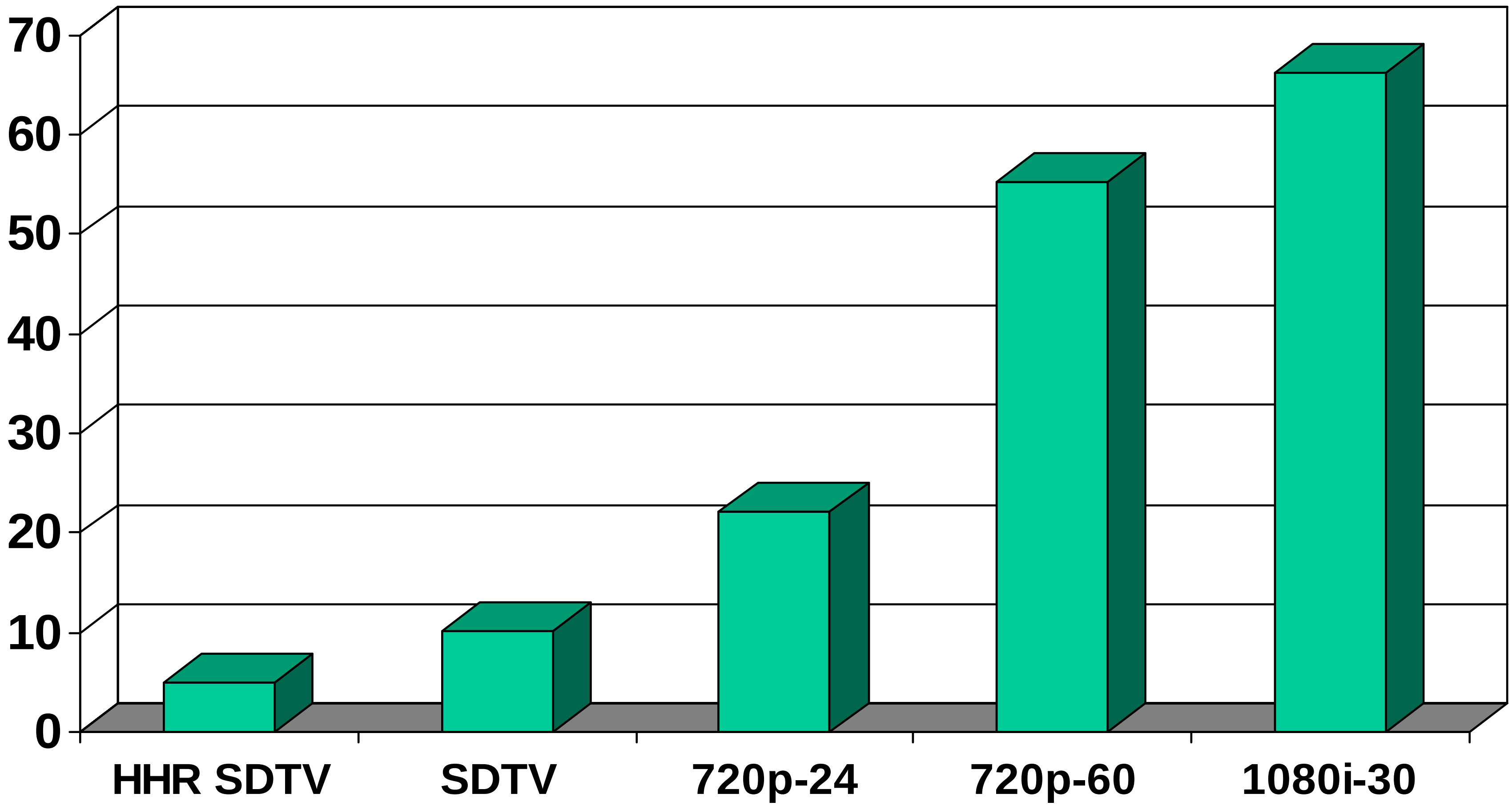
Format Index	Vertical Resolution	Horizontal Resolution	Aspect Ratio	Scan Type	Refresh Rate [Hz]	Type
1	480	640	4:3	interlaced	30	SDTV
2				progressive	24	
3					30	
4					60	
5		704	4:3	interlaced	30	
6				progressive	24	EDTV
7					30	
8					60	
9		704	16:9	interlaced	30	
10				progressive	24	
11					30	
12					60	
13	720	1280	16:9	progressive	24	HDTV
14					30	
15					60	
16	1080	1960	16:9	interlaced	30	
17				progressive	24	
18					30	

Of which 6 are High Definition

What Makes HDTV a Challenge??

- 1080i-30 (1920h x 1080v x 30 frames a second) is 62,208,000 picture elements, or pixels per second. (SMPTE 274M)
- 720p-60 (1280h x 720v x 60 fields a second) is 55,296,000 pixels per second. (SMPTE 296M)
- 720p-24 (1280h x 720v x 24 fields a second, as used in film sourced material…) is considerably less at 22,118,400 pixels per second.
  - For comparison traditional digital television at standard definition is 480i-30 with a maximum of 704h x 480v x 30 frames a second, which is a sample density of 10,137,600 pels per second. Some content on direct to home satellite feeds use ½ this horizontal resolution which is only 5,058,800 pixels per second.
  - Most DVDs use ¾ of SDTV horizontal resolution which is 528h x 480v x 30 frames a second and is 7,603,200 pixels per second.
- The computational cost to implement any encoder is directly proportional to the number of source pixels processed.

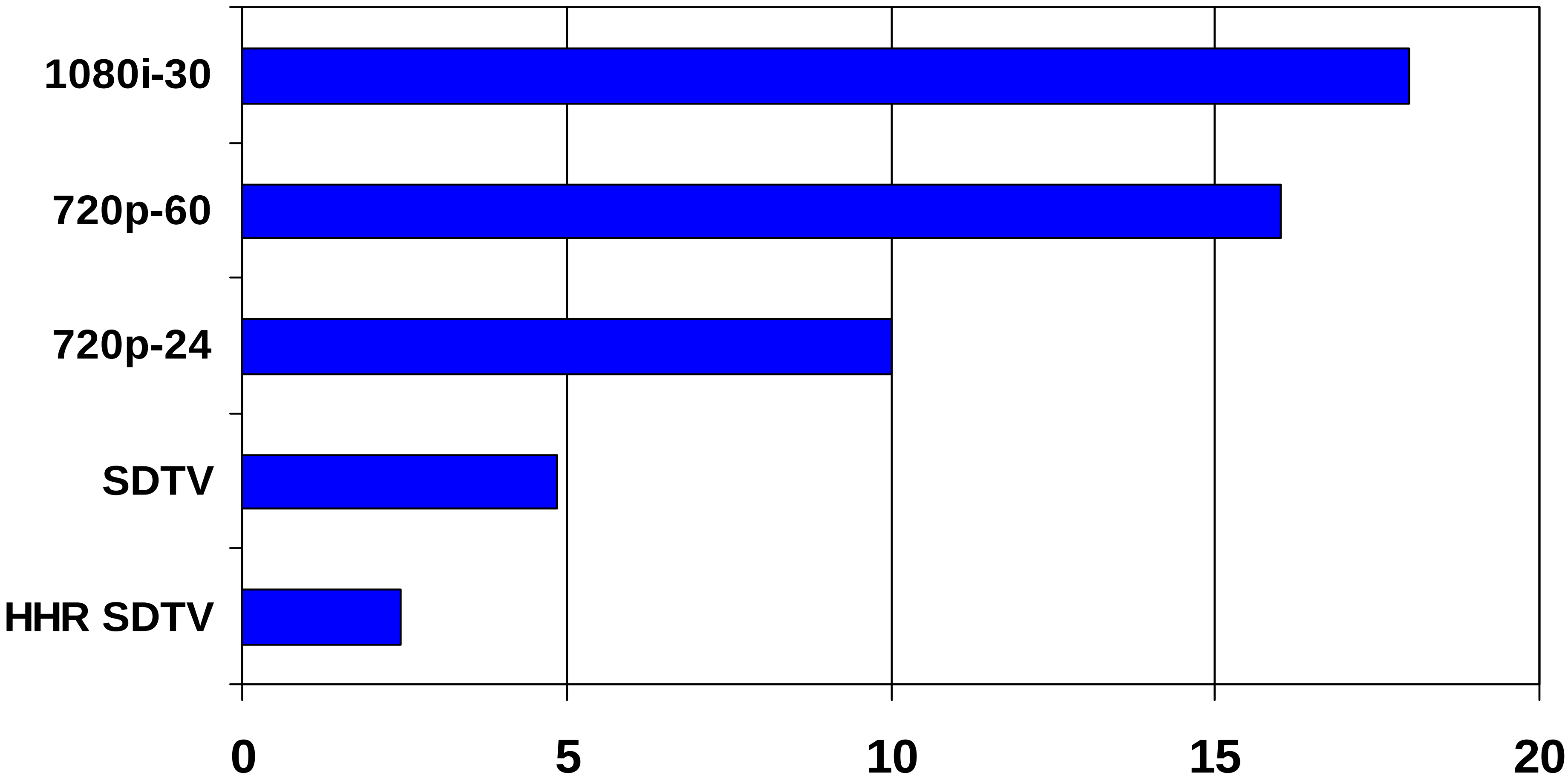
Sample Density in Megapixels/sec



How much bandwidth is needed?

- Each pixel contains both intensity information and color information and for the 4:2:0 type of sampling used in MPEG-2 this is 12 bits per pixel (8 for the luminance and 4 for the chrominance.)
- MPEG-2 can attain compression ratios of up to 25:1 with good preservation of quality, so that:
  - ½ horizontal resolution SDTV needs ~ 2.43 Mbps.
  - Full resolution SDTV needs ~ 4.86 Mbps
- HDTV images have more spatial and temporal redundancy and ratios of perhaps 30:1 to 45:1 are useable so that:
  - 1080i-30 and 720p-60 HD both need ~ 14 to 18 Mbps then adding the audio, and in some cases a simulcast SD content yields the ATSC transport rate of 19.4 Mbps.
  - 720p-24 (HD movies on DVD) would need ~ 8 to 10 Mbps.

Required MPEG-2 Bit Rate in Mbps

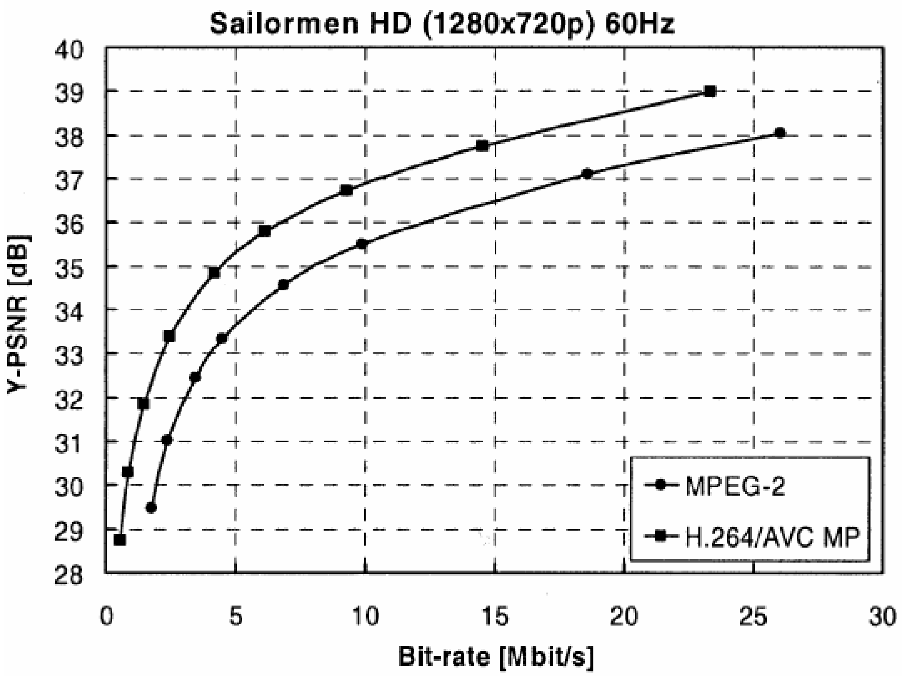


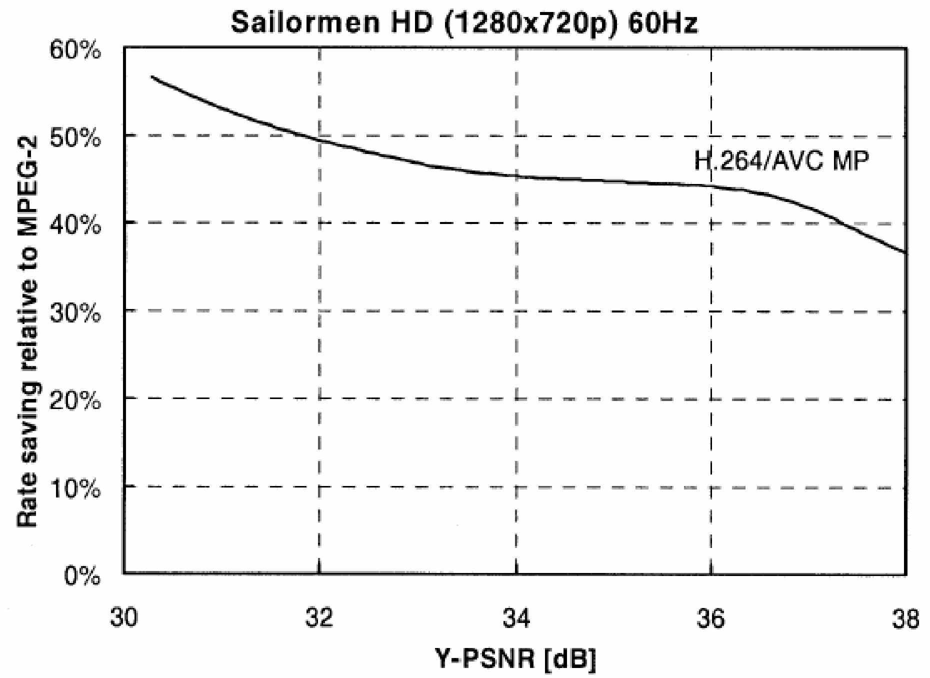
Can we continue to reduce bit rates?

- **Advanced coding techniques can reduce the requirements for both SD and HD bandwidth.**
  - **MPEG-4 AVC (also known as H.264) can reduce the bit rate by a factor of 50% over MPEG-2.**
  - **Windows Media Technology (specifically VC-1) can be applied to both SD and HD (particularly for progressive images) and reduce the bit rate by 35% and up over MPEG-2.**
  - **One or both of these will be used for HD-DVD movie content publishing on the 9 Gigabyte version of the removable optical disk.**

- **Using "intermediate" resolutions as is done for current digital tier cable, satellite and some DVDs for SDTV may also enable delivery of HDTV-like content over constrained bandwidth connections.**

HD Performance of H.264 vs. MPEG-2





From a recent IEEE paper on H.264

Ensuring A High Quality User Experience

Results of Improved Perceptual Quality

- Informal perceptual tests only done so far
- At the same objective measurement (PSNR,) people generally prefer JVT
- This is likely attributable to improvements from:
  - Small motion compensation block size
    (breaks up block structure)
  - Small transform block size
    (breaks up block structure, reduces ringing)
  - In-loop deblocking filter
- By how much?
  - Needs further study
  - No rigorous testing reported
  - 10-15% might be a good guess

Perceptual Improvements: an example

When equally ‘bit starved’ the perceived differences are large