IEEE Globecom 2009

November 30, 2009

Broadband Video Networking

Benny Bing
Georgia Institute of Technology
http://users.ece.gatech.edu/~benny
bennybing@ieee.org

Outline

- Introduction to broadband Internet
- HD wireless interfaces
  - 802.15
  - 802.11
  - Wireless HD
- Broadband video
  - Switched digital video
  - Wireless video
  - H.264 and VC-1 compression standards
  - Loss resilience and error containment
  - Video quality assessment
  - Video transport protocols
  - Video bandwidth conservation and management
  - Next-generation video
- Summary

Broadband Internet

- Economic driver for 21st century
  - Globally, over 300 million households subscribe to broadband Internet
  - Expected to increase to 525 million in 2011
- Broadband Internet can bring significant economic/social benefits
  - Improved healthcare and education to enhanced public-safety programs
    - Improved healthcare through telemedicine and electronic healthcare records
  - Can also bring efficiencies by ushering smart grids, smart homes, and smart transportation
- Broadband in the U.S.
  - FCC task force estimates total cost of broadband deployments in the U.S. between $20 billion and $350 billion
    - Assumes services provided 100 Mbit/s or faster
  - Actual broadband speeds lag advertised speeds by at least 50%
    - Possibly more during busy hours
    - Peak usage hours (e.g., 7 to 10 pm) create network congestion and speed degradation
    - About 1% of users drive 20% of traffic while 20% of users drive up to 80% of traffic
  - Much more wireless spectrum needed
    - Smartphone sales to make up majority of wireless device sales by 2011

Broadband Internet

- The Internet is trumping digital video recorders for on-demand TV
  - Over 80% of Internet users watch video, 20% of these users watch TV
    - Online TV content projected to increase from 2.7 to over 20 billion streams in 5 years
  - Number of Americans watching online TV shows has doubled in last 2 years
    - Either free or cheaper than payTV, ideal in difficult economic times
  - In 2009, some 40 million households worldwide watch online video regularly on their TV sets
    - As a consequence, these households will watch less broadcast payTV
- Over-the-top (OTT) online TV providers trumping payTV providers
  - Dramatic increase in subscription and advertising revenue for online providers
    - Netflix boasting strong growth in both revenue and subscribers during economic downturn
  - Walt Disney reported a 32% drop in quarterly net income in Dec 08
    - Primarily due to a huge decline in DVD sales
- Internet poised to support increasing video traffic load
  - CNN.com Live served 1.3 million concurrent live streams in the moments leading up to President Barack Obama's inaugural address on Jan 20, 2009
  - Served a record-breaking 26.9 million live streams during President's speech
  - Shatters previous record of 5.3 million live streams set on 2008 Election Day

High-Definition (HD) Video

- **Growing in importance and popularity**
  - Intense competition between satellite, cable, and telcos to offer highest number of HD channels
- **Emerging personal wireless network standards geared toward HD**
  - IEEE 802.15 Task Group 3c
  - IEEE 802.11ac/ad
  - IEEE 802.11aa
  - Wireless HD

IEEE 802.15 Task Group 3c

- **Formed in March 2005 (www.ieee802.org/15/pub/TG3c.html)**
  - Developing millimeter-wave WPAN standard
  - Providing data rates extending beyond 2 Gbit/s
- **Five usage models (UMs) define 60 GHz applications and environments**
  - UM1: Single-set *uncompressed* video streaming (implications?)
  - UM2: Multi-set *uncompressed* video streaming
  - UM3: Office desktop data transfer
  - UM4: Conference ad-hoc data transfer
  - UM5: Kiosk file downloading
  - Note: Storage and DRM dictate need for compression, not bandwidth of transmission

New IEEE 802.11 Task Groups

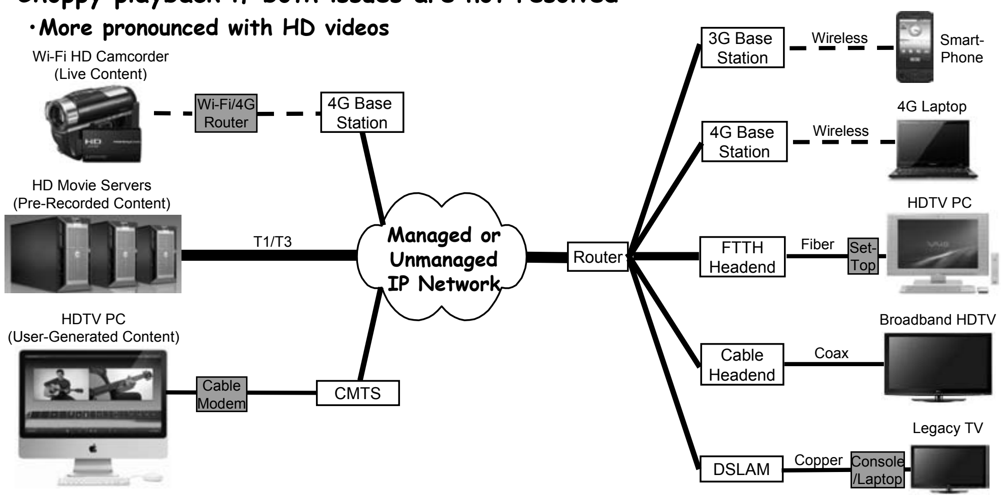
- Two multi-gigabit task groups formed
  - Evolved from the Very High Throughput (VHT) study group formed in 2007
    - Two frequency bands considered: Under 6 GHz and 60 GHz
  - 802.11ac (under 6 GHz) formed in Sept 2008
  - 802.11ad (60 GHz) formed in Dec 2008
  - Will be backward compatible to legacy 2.4 and 5 GHz devices
    - Seamless handoff between 60 GHz and 2.4/5 GHz connections
  - Data rates in excess of 1 Gbit/s
    - Maximum mandatory data rate for a single link may exceed 500 Mbit/s
- 802.11aa task group
  - Focuses on video streaming
- Gigabit Wireless Alliance (WiGig) (http://wirelessgigabitalliance.org)
  - To achieve a data rate of up to 6 Gbit/s
  - Maximum throughput just over 5 Gbit/s
  - Low power option to have a minimum throughput of 1 Gbit/s

Wireless HD

- Consortium promoted by Matsushita, Samsung, Sony
- Version 1.0 completed in April 2008
  - Converts High Definition Multimedia Interface (HDMI) to wireless
    - Handles HD video streams between audio/visual equipment without high-efficiency coding
  - First compliant equipment appeared in Jan 2009
    - LG and Panasonic HDTVs
  - Employs 60 GHz band, streams up to 4 Gbit/s at up to 30 feet
    - CMOS RF IC technology lowers cost of transceiver circuits
  - Tens of antenna arrays
    - "Beam steering" devices used with dynamic adjustment of voltage input to each to adaptively control signal radiation angle
  - Transceiver circuit developed by SiBEAM uses a ceramic package measuring about 20mm square as antenna module
    - Surface of the module is covered with an array of about 36 antenna elements
    - Voltage supplied to each is adjusted to control radiation angle
    - Utilizes OFDM to improve performance in non-line-of-sight (NLOS) use

Multi-Mode Video Connections

- End-to-end management needed for delivering high-quality video
  - Random packet losses on Internet backhaul
  - Bandwidth restrictions on the access network
  - Choppy playback if both issues are not resolved
    - More pronounced with HD videos



Video Content Distribution

- Cable and satellite providers employ closed "walled gardens" systems
  - Offer only selected video content and mostly appointment-based viewing
  - Compare open Internet model (an unmanaged network)
    - Users can access any content they choose and watch videos whenever they want to
  - With online videos becoming popular, satellite providers may lose out more
    - Typically no Internet service available, unlike cable/telco providers
- Emergence of OTT devices and service providers
  - Offers more video choices to consumer
    - Seamlessly integrate live TV with stored video, on-demand movies, online Internet video
  - Replacement or supplementary TV services
    - Providers like Apple TV, Hulu, Netflix, Amazon, Sling, Sezmi may complement or compete with existing payTV providers
- PayTV providers pushing content beyond TV to PC and mobile devices
  - 3-screen bundled service
    - Video content on any video-enabled device, any location, at anytime
  - Single offering
    - One price, one point of customer contact, one integrated electronic program guide
    - Subscribers do not sign contracts with three different providers, receiving three different packages of content, and paying three different fees

Video on Demand (VoD)

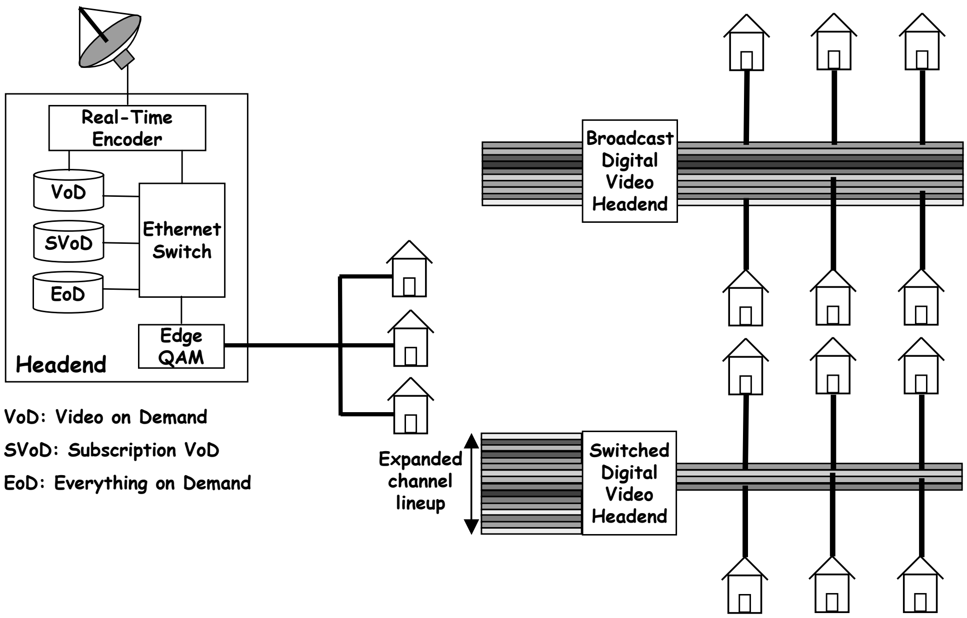
- Allows subscribers to view selected movies or TV channels on demand
  - May be the best form of video service
    - Subscribers watch TV on their own time and at their convenience
    - Movies or channels can be customized to subscriber's preference
    - For example, subscriber will not pay for "unwanted" bundled channels
  - Primarily focused on pre-recorded video
    - High channel latency

- Small payTV VoD traffic compared to broadcast payTV
  - May use substantial amounts of bandwidth
    - VoD streams are normally unicast
  - As such, VoD load and utilization need to be monitored and appropriately sized
- Online TV is basically VoD Internet service
  - Supports both pre-recorded and live video
  - Intrinsically more bandwidth efficient than broadcast payTV
    - Users actually watch videos (in broadcast payTV, channels are broadcast continuously)
  - Cheaper to deploy
    - Bandwidth reclamation equipment such as switched digital video equipment not needed

Switched Digital Video (SDV) over Cable

- Improves bandwidth utilization
  - Allows unused bandwidth to be reclaimed when inactive channels are not accessed by users
- Allows expanded channel line-up without sacrificing existing channels
  - Over 1,000 HD channels can potentially be supported
- Narrowcast approach
  - Popular channels are broadcast continuously
  - Less popular channels are dynamically activated as subscribers view them
    - Channels sent only to set-top boxes (STBs) that tune in to them
    - Saves network bandwidth by not broadcasting channels to all STBs all the time
  - Allows for fast channel change and facilitates multicast operation
  - Significant cost benefits from bandwidth sharing and optimization through high density video processing
    - Allows each node or region to operate with level of programming complexity once reserved for main distribution center
    - Moves complex processing of channel lineup far closer to subscriber (e.g., at network edge), placing heavy demands on edge video processing equipment

VoD and SDV Architectures



VoD: Video on Demand

SVoD: Subscription VoD

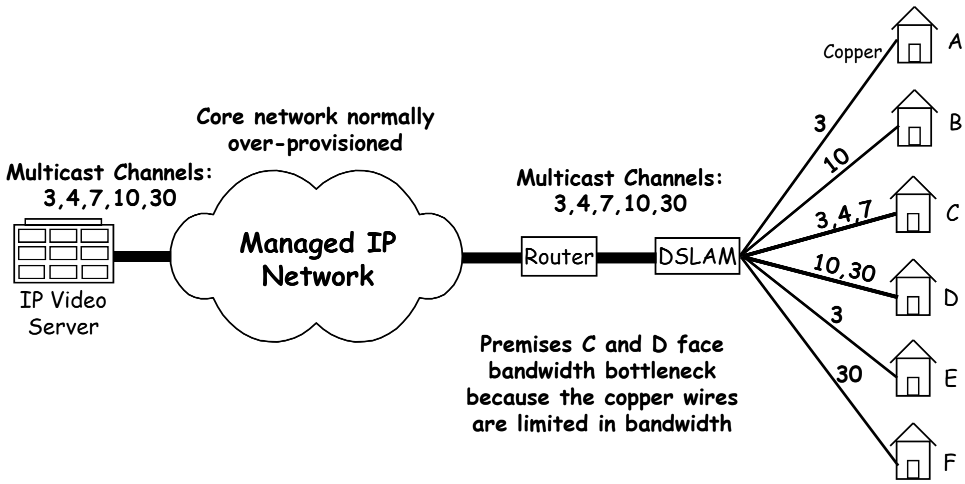
EoD: Everything on Demand

IPTV over DSL

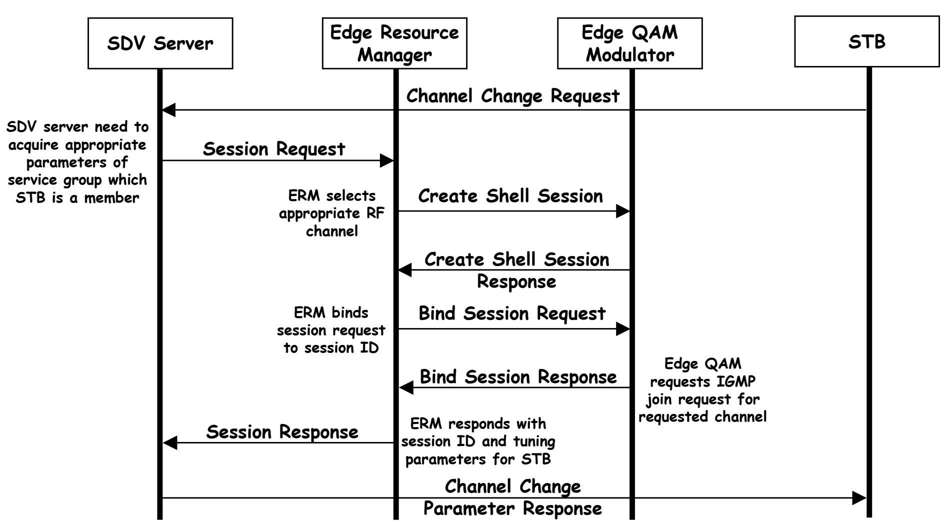
- Switched digital video over a managed IP infrastructure
  - Dedicated DSL connections in access network
- Unicast approach
  - High channel change latency
  - Bandwidth consumption on access link becomes a significant bottleneck in homes or businesses with simultaneous users
  - Channel change latency can be improved with complex buffer management and video playback solutions
    - Inevitably lead to increased network overheads and set-top complexity
- Multicast approach
  - Core technology driving IPTV deployments with switched video
  - Every channel maps to a multicast address
    - Flipping to new channel results in joining multicast group corresponding to channel
  - Multicast DSLAM approach
    - Bandwidth bottleneck in DSLAM backhaul, router, video headend is solved
    - Access link bottleneck remains

Access Link Bottleneck in Multicast IPTV over DSL

- Multichannel viewing in single household remains a challenge

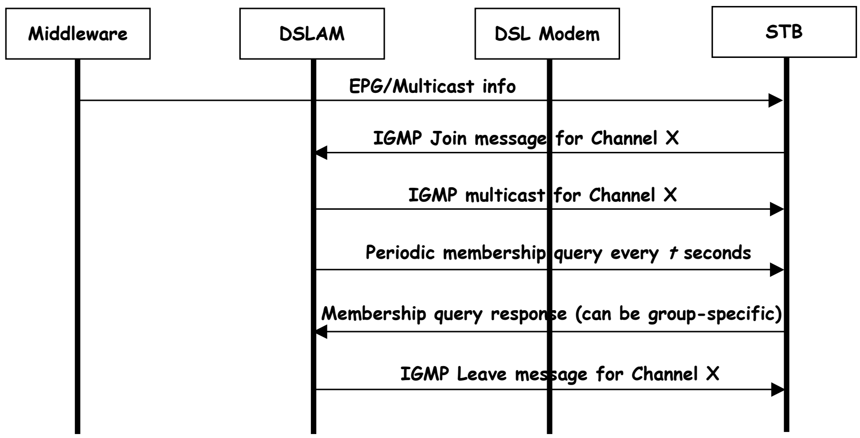


Channel Change in SDV



Channel Change in IPTV

- IGMP Join time is network dependent
  - System will glitch if latency Join < latency Leave



Online Video Delivery

- Key competitor to payTV
  - IPTV has not been a cost-effective competitor to payTV
- Delivery modes
  - Buy it
    - Use over and over
  - Rent it
    - Use it once
  - Subscribe to it
    - Access included with monthly fee
    - May need to purchase console
  - Watch it for free
    - Full episodes, TV shows, and selected movies
    - Have to put up with commercials
    - A lot less commercials compared to payTV, for now...

Broadband Video Delivery Issues

- Transcoding
  - Equivalent to network streaming except that output is sent to a file instead
  - Allows selection of appropriate codec/bit rate for delivery network
  - Key to anywhere, anytime, any device delivery
    - Difficult to maintain 15 different versions of the same movie
- Digital rights management (DRM)
  - Ownership, control, and distribution of stored media
- Various delivery platforms
  - Video set-top, game console, network media player, Internet-TVs, PC, laptop
  - Some typical functions
    - Record, archive, and play back video and music
    - Store and organize digital photos from various sources
    - Store and play video games
    - Distribute digital media around the home

Broadband Video Delivery Issues

- PC or laptop may cause jerkiness and stalling playback
  - May not related to transmission, good video card may aid HD playback
- Game consoles TV-friendly and easily connected to the Internet
  - Currently used by more than 90 million households worldwide
  - Favored by younger consumers
- Internet-enabled TVs
  - Additional features help both cable and online providers e.g., ad management
    - Companies: ActiveVideo Networks, AnySource Media, GridNetworks, Sezmi, TiVo, etc.
  - Widget channel extends Web-based services and applications directly to TV
    - Designed to pull selective content from the Internet to complement TV watching
    - Users can buy products advertised on TV from online stores
    - LG, Roku, Samsung, Toshiba to introduce Widget Channel-enabled hardware
- Proprietary STBs to become a relic of the 20th century?

Online Video Portals Requiring Consoles or Broadband TVs

	Console	TV Episode	Movie Rental	Movie Purchase	Comments
NETFLIX	$99	Starting $4.99/month			10 million users, over 17,000 movies/episodes.
XBOX	$199	$2	$4 (SD), $6 (HD)		Video game console.
	$299	$1.99-$2.99	$2.99	$9.99	Video game console.
tv	$229	$1.99	$3.99	$14.99	Online video streaming and retail. Sold 200 million TV programs, over 32,000 movies.
vudu	$299	$1.99	$3.99	$19.99	2,000 1080p HD and 16,000 movies.
BLOCKBUSTER	$99	$1.99	$3.99	$9.99	Supported by Yahoo!/Intel Widget Channel.
amazon.com	Optional	$1.99	$2.99	$14.99	Online video streaming and retail, over 50,000 titles. Supported by TiVo, Sony's Bravia, Xbox 360, Windows Media Center, and Roku.

Online Video Portals with Direct Viewing

	Free Episodes and Movies	Comments
veoh	Mostly free, subscription needed for premium channels.	Open-platform supports user-generated video.
CRACKLE	Free, ad-supported episodes and movies.	Supported by Sony Entertainment.
ADOBE TV	Free, need to install Adobe Media Player.	Blip.tv, CBS, Comedy Central, Epicurious, KQED, MTV, MyToons.com, Nickelodeon, and more.
hulu	Free, ad-supported episodes and movies.	Partners Sling Media and Disney. Owned by NBC Universal and News Corp. Began HD videos since Aug 2008. **Nearly 500 million views a month!**
Joost	Free browser-based streaming.	WiFi video streaming via iPhone.
FANCAST	Free, ad-supported episodes and movies.	Supported by Comcast.
tv.com	Free, ad-supported episodes and movies.	Owned by CBS.
CBS	Free, ad-supported episodes and movies.	Owned by CBS.

Online Video Portals with Direct Viewing

- YouTube
  - Over 10 billion videos viewed in a single month
    - Evolve from user-generated video and video sharing portal to an aggregator of premium video
    - Started offering HD-quality videos since Dec 08
  - On April 19, 2009, announced new destination for TV shows and improved destination for movies (http://www.youtube.com/shows)
    - Partners Crackle, Lionsgate, Starz, CBS, MGM, National Geographic, etc.
    - TV episodes and movies include music videos, full-length films/TV shows
  - Published an expanded set of APIs
    - Allows third-party TiVo devices, STBs, mobile handsets, Web sites, etc, to gain easier access to YouTube's content

Online Streaming Rates

- Most videos in standard definition (360p/480p)
  - Very few HD movies although HD clips and trailers are available
- Connection rates
  - See http://www.hulu.com/hd and http://www.cbs.com/hd
    - Typical *minimum* rates: 480p (1.5 Mbit/s), 720p (2.5 Mbit/s), 1080p (3.5 Mbit/s)
    - Employ H.264 and Flash 9.0+ player
    - Freeze frames are common even with these rates
    - Unclear what frame rate some sites are employing
    - May be 24 frame/s or lower because fast movement may appear to be in slow-motion
  - Cannot watch HD videos over 3G, challenging even with 4G!

Potential of Online Video

- $400 million invested over 30 companies in last 4 quarters
  - $180M in Q3/09, $64M in Q2/09, $75M in Q1/09, $80M in Q4/08
- Rise of Hulu
  - Free viewing of full-length episodes and TV shows, and movies
- Dominance of YouTube
  - Free viewing of user-generated and premium videos
- Resounding success for Netflix's Watch Instantly streaming service
  - Recession-proof: $51M in cash flow in Q4/08 alone, more than in all of 2007
- Entry of cable operators, telcos, and networks with TV Everywhere
  - Comcast's On Demand Online trial: HBO, Cinemax, Starz, TNT, TBS, CBS
  - Bell Canada offering TMN online TV servce (bell.ca/tvonline)
- Entry of 4G broadband wireless networks with mobile Internet
  - On-the-go HD video streaming on laptops
- Entry of consumer electronics vendors with Internet-enabled HDTVs
  - Shipments may top 6 million by 2013
  - Such HDTVs may become the norm, just like digital tuners

Potential of Online Video

- Nintendo to launch online video streaming services in U.S. and Europe
  - 30+ million Wii users may watch Internet video on their TV sets
- Social media network giant Facebook launches VoD on Oct 16, 2008
  - Partners Kyte channel (http://www.kyte.tv)
  - Partners CNN and Amazon
- Providers exploring solutions to counter threat of Internet TV boom
  - Plan to offer large numbers of TV shows online (e.g., Discovery and Disney TV networks) but accessible only to subscribers
    - Anytime, anyplace access to content already paid by subscribers
  - Contradictory requirements
    - Positioning this new online "entitlement" service
    - Restricting or "managing" bandwidth usage among subscribers via tiered bandwidth caps
  - Focus on improving live sports
    - More natural movement with 120/240 frame/s HD video
    - May not be a compelling differentiator – human vision may not distinguish between high frame/s and high motion, this why slow-motion replays are needed
    - Fixed channel bandwidth (e.g., 6 MHz) and MPEG-2 STBs may have restrictions

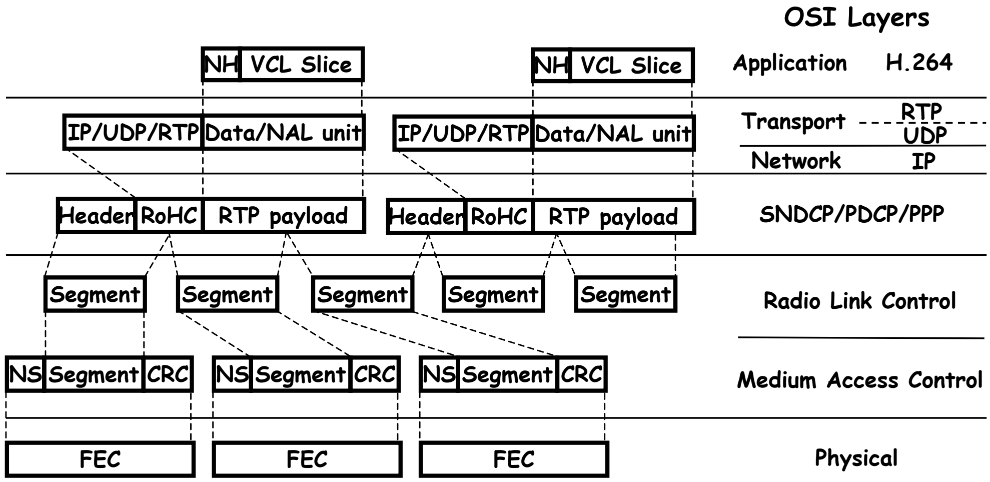
Key Differences Between Online TV and payTV Viewing

- **Downloading latency versus instant access**
  - Multiple videos can be pre-loaded on a webpage while user decides on selection
    - Reduces channel change latency
  - Play commercial before video
    - Creates delay for initial buffering of selected video
    - Commercial is streamed from a separate server
- **Mouse selection versus TV widgets versus TV remote**
  - In payTV, users are presented with a single channel
  - Channel change may be more frequent than online TV
    - Fast channel change critical for payTV

Wireless Video

- **Video encoder generates data units containing compressed video stream, possibly stored in an encoder buffer before transmission**
  - Wireless medium might delay, lose or corrupt individual data units
  - Overflows or underflows may occur in receiving client device buffer
  - Upstream contention among multiple client devices for channel bandwidth, congestion, and retransmission may lead to losses or delays
  - May have significant impact on perceived video quality due to spatio-temporal error propagation

Wireless H.264 Packetization in 3GPP Framework



FEC: Forward Error Correction

NH: Network Abstraction Layer (NAL) header

PDCP: Packet Data Convergence Protocol

PPP: Point to Point Protocol

RoHC: Robust header compression

SNDCP: Sub Network Dependent Convergence Protocol

VCL: Video Coding Layer

Wireless 3GPP Video Applications

Application	3GPP	Max. delay	Encoder buffering requirements	Transport feedback	CSI	Encoding
Download and play	MMS	NA	None	Yes	NA	Offline
On-demand, pre-encoded streaming	PSS	1 s	Yes	Yes	Partly	Offline
Live streaming	PSS	200 ms	Yes	Partly	Partly	Online
Multicast	MBMS	1 s	Limited	Limited	Limited	Both
Broadcast	MBMS	2 s	None	None	None	Both
Conferencing	PSC	250 ms	Limited	None	Limited	Online
Telephony	PSC	200 ms	Yes	Limited	Partly	Online

CSI: Customized Applications for Mobile network Enhanced Logic (CAMEL) Subscription Information

MBMS: Multimedia Broadcast/Multicast Service

MMS: Multimedia Messaging Service

PSC: Primary Synchronization Code

PSS: Packet Switched Stream

H.264 and VC-1 Compression Standards

- Two powerful codecs that can support efficient video delivery
  - Achieve higher compression ratios than legacy codecs such as MPEG-2
  - H.264 typically provides a two-fold improvement in compression efficiency over MPEG-2
- Both codecs are becoming widely adopted in consumer electronics as well as in narrowband and broadband network transport
  - In general, a higher compression efficiency can be achieved for high-definition than standard-definition videos
- Standardization
  - H.264 advanced video codec (AVC) is standardized by the International Telecommunications Union (ITU)
    - http://www.itu.int/rec/T-REC-H.264/e
    - http://www.itu.int/itudoc/gs/promo/tsb/87066.pdf
  - VC-1 specification is standardized by the Society of Motion Picture and Television Engineers (SMPTE)
    - Implemented by Microsoft as Windows Media Video (WMV) 9
    - Website: www.smpte.org

H.264/MPEG-4 Part 10 Advanced Video Codec (AVC)

- Powerful set of video compression/decompression formats
  - Improves compression efficiency over prior standards
  - Future-proof
    - Codecs with new extensions can be added as technology improves
- Widely adopted
  - Consumer electronics e.g., camcorders, surveillance cameras, phones
    - Apple's iPhone and iPod Touch support H.264 Baseline profile
  - Recommended codec for all 3GPP video services
- Adaptable to different applications, client devices, networks
  - Necessary when transmitting interactive media over heterogeneous networks
    - Video quality may be prioritized over compression efficiency when bandwidth is abundant
  - Useful for applications when client device is not capable of displaying full resolution or full quality video
- Jumpstarts OTT video services
  - Allows such services to achieve video quality superior to that delivered over managed networks and legacy MPEG-2 set-top devices

H.264 Architecture

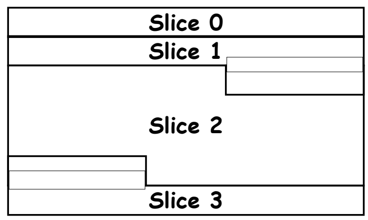
- Comprises two conceptually different layers
  - –Video coding layer (VCL) and network abstraction layer (NAL)
  - –VCL defines core video compression engines that perform basic functions such as motion compensation, transform coding of coefficients, and entropy coding
    - VCL is transport unaware and its highest data structure is the video slice, an integer set of macroblocks (MBs) coded in raster scan order
  - –NAL is an interface between codec and transport network
    - Is therefore responsible for encapsulation of coded slices into transport entities namely transport protocols (e.g., MPEG-2 transport stream, real-time transport protocol or RTP) and file formats (e.g., MPEG-4)
  - –NAL operates on NAL units (data packets)
    - Each unit comprises a one-byte header and a bit string that represents the bits constituting the MBs of a video slice
  - –RTP payload supports 3 modes
    - Single NAL unit transported in single RTP packet
    - Non-interleaved mode: NAL units of same picture are packetized into single RTP packet
    - Interleaved mode: NAL units from different pictures are packetized into single RTP packet, not necessarily in their decoding order

H.264 Video Hierarchy

• **Sequence**, consisting of • **Pictures**, consisting of	} NAL
• **Slices and slice groups**, consisting of • **Macroblocks**, consisting of • **Blocks**, consisting of • **Pixels/Pels**	} VCL

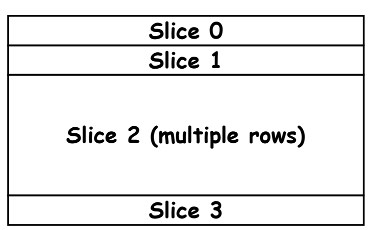
Slice Coding in H.264

- Provides bit rate scalability
  - Each video picture is subdivided into one or more slices
    - Order can be modified when an error-resilient method such as flexible macroblock ordering (FMO) is used
  - Slice is given increased importance if it is the basic spatial segment that is independent from its neighbors
    - Errors or missing data from one slice cannot propagate to any other slice within picture
    - Increases flexibility to extend frame types (I, P, B) down to level of slice types
    - Redundant slices are permitted



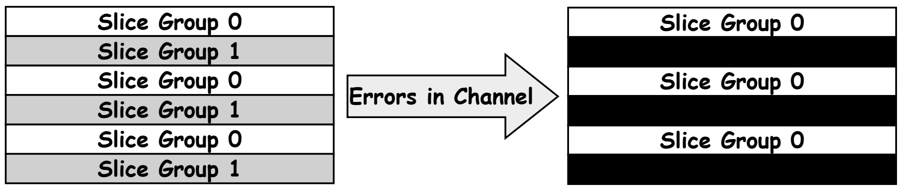
Slice Coding in VC-1

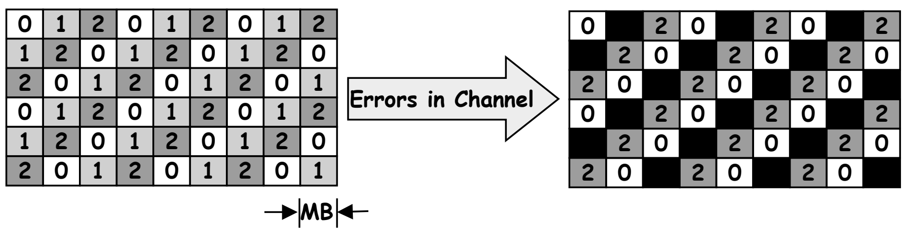
- Allows only one slice group
- Several slices can be implemented but length of each slice is restricted to a single row of MBs in the picture
  - Slice takes only rectangular shape
  - Less flexible than H.264 where several slice groups can be used
    - In addition, length of slices can be less than a row of MBs
    - Granularly can be a single MB if desired



Spreading Errors using Multiple Slice Groups

- Interlaced or dispersed map improves error immunity
  - Enhances performance of error concealment





H.264 Profiles

- Specify restrictions on H.264 bitstreams
  - Limits the capabilities needed to decode the bitstreams
  - If the same parameters are chosen for encoding, profile choice has no bearing on video quality, encoding time or compressed video size
    - If a different set of parameters is chosen, then profile may provide additional capabilities for bitstream
  - Decoder conforming to High 4:4:4 profile is capable of decoding a bitstream encoded with High 4:2:2, High 10, High, and Main profiles
    - Similarly, High 4:2:2 profile decoder is capable of decoding High 10, High, Main profiles
    - These profiles do not offer tools for loss robustness or resilience and are mainly designed for storage or for broadcasting in loss-free environments
    - They provide capabilities for higher compression efficiency such as use of weighted prediction for P slices, 8x8 transform coding, etc
- Higher profiles target higher quality videos
  - Employs more chroma samples per luminance sample (4:4:4 versus 4:2:2) or finer quantization parameter values (up to 14 bits per sample for High 4:4:4)
  - Use of higher profiles is therefore justified if video is already in good quality

H.264 Profiles

- **3 basic profiles are defined: Baseline, Extended, and Main**
  - **Baseline profile is a subset of Extended profile**
    - **Both profiles address problems in loss-prone environments**
- **Extended profile**
  - **Reduces temporal correlation using B-frames**
  - **Offers error resilience capabilities (e.g., data partitioning)**
  - **Computationally more complex**
- **Baseline profile**
  - **Used when short encoding and decoding time are desired**
    - **For example, with applications such as videoconferencing or video streaming**
    - **For this reason, B-slices are not allowed whereas they are allowed in extended profile**

Profile	Typical Application	Additional Decoder Complexity over MPEG-2	Typical Efficiency over MPEG-2
Baseline	Low delay applications	2.5 times	1.5 times
Extended	Mobile streaming	3.5 times	1.75 times
Main	Broadcast video	4 times	2 times

H.264 Profiles

Parameter	Baseline	Main	Extended	High	High 10	High 4:2:2	High 4:4:4
Use of B-slices		×	×	×	×	×	×
Data Partitioning			×				
Interlaced Fields		×	×	×	×	×	×
Arbitrary Slice Ordering	×		×				
Multiple Slice Groups	×		×				
8-bit Depth for Samples	×	×	×	×	×	×	×
9 to 10 bit Depth for Samples					×	×	×
11 to 14 bit Depth for Samples							×
Transform Bypass Operation							×
Redundant Pictures	×		×				
Quantization Scaling Matrices				×	×	×	×
Weighted Prediction for P and SP Slices		×	×	×	×	×	×
Context-Based Adaptive Binary Arithmetic Coding		×		×	×	×	×
8x8 Transform Decoding				×	×	×	×
Separate Picture Scaling				×	×	×	×
Second QP Chroma Index Changeable				×	×	×	×
Monochrome Format				×	×	×	×
Chroma Format 4:2:0	×	×	×	×	×	×	×
Chroma Format 4:2:2						×	×
Chroma Format 4:4:4							×

VC-1 Profiles and Rates

Profile	Level	Maximum Bit Rate	Resolutions by Frame Rate
Simple	Low	96 Kbit/s	176 x 144 @ 15 Hz (QCIF)
	Medium	384 Kbit/s	320 x 240 @ 24 Hz (QVGA) 352 x 288 @ 15 Hz (CIF)
Main	Low	2 Mbit/s	320 x 240 @ 24 Hz (QVGA) 352 x 288 @ 30 Hz (CIF)
	Medium	10 Mbit/s	720 x 480 @ 30 Hz (480p) 720 x 576 @ 25 Hz (576p)
	High	20 Mbit/s	1920 x 1080 @ 25 Hz (1080p) 1920 x 1080 @ 30 Hz (1080p)
Advanced	L0	2 Mbit/s	352 x 288 @ 25 Hz (CIF) 352 x 288 @ 30 Hz (CIF) 352 x 240 @ 30 Hz (SIF)
	L1	10 Mbit/s	704 x 480 @ 30 Hz (NTSC-SD) 720 x 576 @ 25 Hz (PAL-SD)
	L2	20 Mbit/s	704 x 480 @ 60 Hz (480p) 1280 x 720 @ 25 Hz (720p) 1280 x 720 @ 30 Hz (720p)
	L3	45 Mbit/s	1280 x 720 @ 50 Hz (720p) 1280 x 720 @ 60 Hz (720p) 1920 x 1080 @ 25 Hz (1080i) 1920 x 1080 @ 30 Hz (1080i) 1920 x 1080 @ 25 Hz (1080p) 1920 x 1080 @ 30 Hz (1080p) 2048 x 1024 @ 30 Hz
	L4	135 Mbit/s	1920 x 1080 @ 50 Hz (1080p) 1920 x 1080 @ 60 Hz (1080p) 2048 x 1536 @ 24 Hz (Digital Cinema) 2048 x 2048 @ 30 Hz

Entropy Coding Methods in H.264

- Two methods applied to transform coefficients
  - Context-Adaptive Binary Arithmetic Coding (CABAC)
  - Context-Adaptive Variable-Length Coding (CAVLC)
- CABAC achieves better efficiency using variable bit-rate encoding
  - Considerable bit-rate savings
  - Encodes entire bitstream
  - Only supported in main and higher profiles
  - Requires higher amount of processing to decode compared to other algorithms
- CAVLC is combined with another entropy coding technique, universal variable length coding (UVLC)
  - Encodes only the headers
  - Less complex than CABAC
    - Used to improve performance of slower playback devices

CABAC and CAVLC under VBR Mode

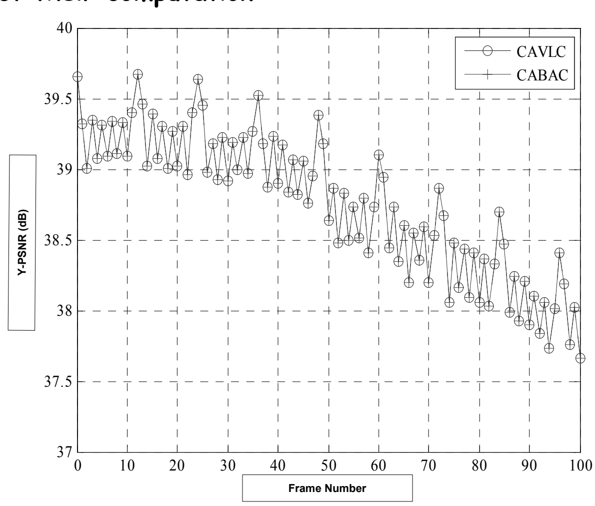
- CABAC achieves higher encoding efficiency than CAVLC
  - Independent of video format (i.e., QCIF, CIF, HD)
  - More efficient for higher quality videos
    - Video file size is 5.3% smaller for Foreman QCIF, 7.7% smaller for Foreman CIF, 10.4% smaller for Coastguard, 8.3% smaller for Mobile, 11.1% smaller for Blue Sky
  - Decoding times are almost the same
    - For encoded videos with the same size, decoding time for CABAC becomes longer
    - Higher amount of processing required by CABAC is compensated by smaller bitstreams

Video	Entropy Coding	Size (bytes)	Decoding Time (sec)
Foreman QCIF (400 frames)	CAVLC	235,649	27.564
	CABAC	223,133	27.362
Foreman CIF (300 frames)	CAVLC	554,249	47.330
	CABAC	511,711	48.155
Coastguard CIF (300 frames)	CAVLC	1,052,730	49.828
	CABAC	943,413	49.860
Mobile CIF (300 frames)	CAVLC	7,575,666	69.204
	CABAC	7,031,726	69.221
Blue Sky 1080p HD (217 frames)	CAVLC	7,372,239	583.613
	CABAC	6,551,358	584.332

Videos were encoded with IBPBPBPBPBPB GOP structure and quantization parameter (QP) value of 30 for I, P, and B frames

CABAC and CAVLC under VBR Mode

- Entropy coding method has no impact on decoded video quality
  - Y-PSNRs of raw video (YUV) files are identical
  - Entropy coding only changes coding of quantized transform coefficients but does not affect their computation



Y-PSNR for CABAC and CAVLC under VBR mode (Blue Sky HD video)

CABAC and CAVLC under CBR Mode

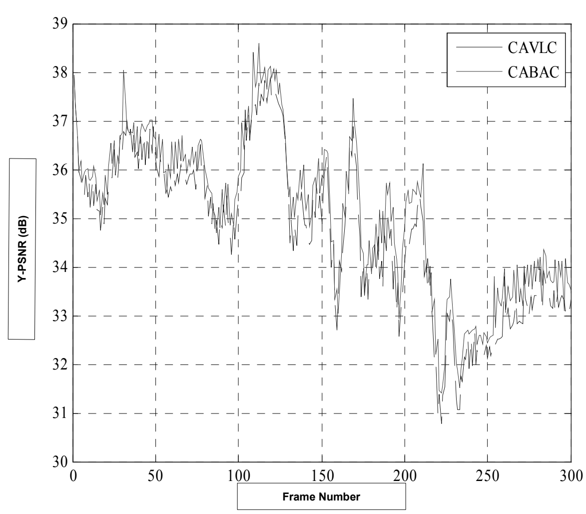
- QP of each encoded frame is adapted to fit targeted bit rate
- Encoding time significantly shorter for CABAC
  - Increases with specified bit rate
  - 21 sec less with 300 Kbit/s and 43 sec less with 1 Mbit/s
- Decoding time slightly reduced with CAVLC
  - Unlike the VBR case, difference in decoding complexity is visible
    - Difference is more pronounced with higher quality, higher resolution videos
- Average Y-PSNR slightly better with CABAC

Performance of CABAC and CAVLC under CBR Mode (Foreman CIF video)

Entropy Coding	Encoding Time (sec)	Decoding Time (sec)	Average Y-PSNR (dB)
	300 Kbit/s		
CAVLC	1040	41.138	34.69
CABAC	1019	42.343	35.13
	1 Mbit/s		
CAVLC	1045	50.922	39.69
CABAC	1002	52.438	40.12

CABAC and CAVLC under CBR Mode

- Y-PSNR of each frame encoded using CABAC is better than CAVLC
  - Due to bit-rate savings provided by more efficient encoding, a lower QP value can be used for CABAC, which improves video quality



Y-PSNR for CABAC and CAVLC with a constant bit rate of 300 Kbit/s (Foreman CIF video)

Rate Distortion Optimization (RDO)

- H.264 MB coding modes
  - Intra coding (spatial) and inter coding (temporal) modes
  - In intra coding, luminance component of each MB is uniformly predicted or is subdivided into 4x4 blocks
    - Latter case is useful for highly detailed regions of the pictures
    - In both cases, several prediction techniques are available
  - Inter coding modes also subdivide MBs for motion-compensated prediction
  - RDO algorithms essentially choose the coded modes to achieve the best tradeoff between low distortion and low bit-rate, based on specific metrics
- Two RDO modes: Fast High Complexity and High Complexity
  - Fast High Complexity employs a simplified algorithm
  - Requires fewer computations and thus, reduces the encoding time slightly
    - This is done at the expense of the PSNR
- Provides better visual quality for P-frames
  - Encoding time and frame size are reduced for B-frames
- Increases size of encoded video as well as encoding time
  - Encoding time may not be important for pre-recorded videos

Rate Distortion Optimization (RDO) under CBR

- CBR and RDO can be considered as complementary tools
  - If bit-rate is fixed and QP values are determined by rate-control algorithm, RDO simply determines best prediction mode
  - Before providing a QP value, rate control algorithm requires information which is only available after RDO algorithm has completed its prior computations
  - To simplify the encoding process, RDO results are predicted based on complexity of previous pictures in the sequence
- Encoding time increases with bit-rate when RDO is used
  - QP values derived by rate control algorithm decrease when bit-rate increases
  - Encoding time is longer for smaller QP values

Loss Resilience and Error Containment in H.264

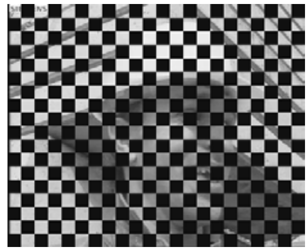
- Targets video bitstream rather than PHY layer bitstream
  - More effective than forward error correction (FEC)
- Network abstraction layer (NAL)
  - Allows same video syntax to be used in many network environments
  - Sequence parameter set (SPS) and picture parameter set (PPS) in each NAL unit provide more robustness and flexibility than prior MPEG designs
- Data partitioning (DP)
  - Partitions compressed data units into different levels of importance
    - Allow higher priority syntax elements (e.g., sequence headers) to be separated from lower priority data (e.g., B-picture transform coefficients)
    - May minimize loss rates for important data with unequal error protection (UEP)
- Redundant slices (RS)
  - Allow encoder to send an extra representation of a MB (typically at lower fidelity) that can be used if primary representation is corrupted or lost
  - Since a MB is typically contained only in one slice, RS allows more representations of a MB to be coded in the bitstream
- Multiple reference frames
  - Used for improved motion estimation
  - Also allow for partial motion compensation for a P-frame when one of its reference frames is missing or corrupted

Flexible Macroblock Ordering (FMO)

- Partitions picture into several slice groups
  - Allows restructuring and reordering of MBs in pictures
  - MBs no longer assigned to slices in raster scan order
  - Each MB can be assigned freely to a specific slice group using a MB allocation map (MBA map)
  - Up to 8 slice groups in one picture
  - Within a slice group, MBs are coded in scan order
  - When only one slice group is activated, FMO is deactivated
- Bits associated with adjoining MBs can be scattered more randomly throughout bit stream
  - Reduces probability that a packet loss will affect a large picture region
    - For example, slice groups can be constructed in such a way that if one slice group is not available, each missing MB can be surrounded by MBs of other slice groups
  - Enhances error concealment
    - Ensures that neighboring MBs will be available for predicting a missing MB
- An example of multiple description coding
  - Each slice group represents a description and is independently decoded

FMO Types with 2 Slice Groups

Type 0



Type 1

Type 2

Type 3

Type 4

Type 5

Increased FMO Overheads

- **Reduction of coding efficiency when more SGs or more slices are used**
  - –Each H.264 slice is designed to be decodable without other slices of picture
    - Implies no intra prediction between different slices, which degrades coding efficiency
  - –Overhead is present for all FMO types because they require creation of several slices
- **Coding efficiency of FMO type 1 is the worst**
  - –No neighboring MBs in original picture are encoded in the same slice
  - –With 4 SGs, difference is even more pronounced because MBs of the same slice are farther from each other compared to the case when 2 SGs are used
- **CABAC or CAVLC efficiency**
  - –Depend on elements which have already been coded
    - When they are correlated to the next elements, entropy coding is more efficient
  - –Use of several slices breaks the entropy coding

Increased FMO Overheads

- Extra information such as slice headers and picture parameter sets
- 4 FMO types do not require extra PPSs
  - In FMO type 1, a MB at a given location always belongs to the same SG
  - For FMO types 3, 4, and 5 with a steady cycle of change, a single PPS is sufficient
- Other FMO types may require extra PPSs
  - For type 0, location of first MB of each SG is coded in PPS
  - For type 2, location of top left MB and bottom right MB of each SG is coded in PPS
  - For type 6, entire MBA map is encoded in PPS
    - If MBA map changes for different frames, new PPSs are coded in bitstream

Video	No FMO	With FMO	Overhead
720p, QP 25	4.20 MB	4.56 MB	8.5 %
1080p, QP 20	14.95 MB	16.24 MB	8.6 %

Error Concealment

- Can improve video quality and deal with packet losses effectively
  - Does not rely on channel feedback
  - Activation is optional and can be made adaptive
    - Not needed with sufficient bandwidth or good channel
- Changes types of video artifacts
  - Significantly different from MPEG-2
  - Frequency of occurrence of artifacts substantially lower compared to MPEG-2
- Several basic methods
  - Copy previously-received uncorrupted picture or MB
  - Employ motion compensation
  - Employ spatial/temporal interpolation from adjacent areas of the same frame or of the previous frame
    - Requires detection of missing MBs after decoding to locate damaged areas of picture
- Effectiveness improves when combined with loss resilient methods
  - Fairly effective even when loss resilient method is not activated i.e., error resilience is not mandatory for error concealment to function
    - However, if error resilience is activated, than error concealment is mandatory
  - Typically incurs more processing time than error resilience
  - FMO type 1 provides best visual quality with EC in a lossy environment
    - Improves robustness at the expense of a relative overhead two times greater

Evaluation of Error Concealment Methods

- Disadvantages of methods that copy previous frame
  - Cannot protect first frame or frames with scene changes
  - May be prone to error propagation
  - Performance depend on scene complexity
    - Scenes with fast motion or rapid changes are difficult to conceal with this method
    - Fortunately, fast scene change and high motion do not usually go together!
- Methods based on interpolation do not have these disadvantages
  - Pixel values of received or concealed neighboring MBs are interpolated
  - Employed when frame to conceal does not resemble previous frame e.g., scene change, I-frame
  - May not always perform well
- Actions performed after error concealment
  - Deblocking filter may change the value of the pixels of the edges of the MB depending on the smoothness of the edge between the MB and its neighbor
  - Performing deblocking filtering with concealed blocks may corrupt the pixels of the correctly received MB

Examples of Error Concealment Implementation

- For scene change and I-frame
  - Error in frame can be concealed by weighted average of luminance and chroma values of pixels surrounding missing MB
  - By default, only the correctly received pixels are used
  - When there are consecutive rows of missing MBs, pixels from the concealed MBs may be used for concealing other neighboring MBs
- For P-frame
  - Each missing MB can be concealed by "guessing" its motion vector (MV)
  - First compute all MVs of surrounding blocks, their average, their median, and the zero MV
  - Then choose MV minimizing the difference of the luma pixels at the edge of the MB to be concealed
  - When there are consecutive rows of missing MBs, same process can employed but MVs for the concealed neighboring MBs may be used

Order of MB Concealment

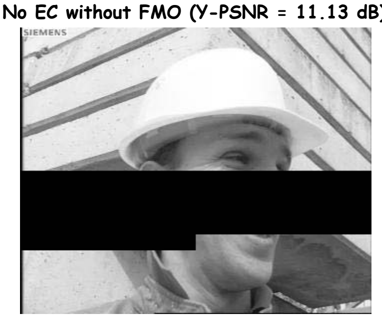
- **Can be done column-wise or alternating column**
  - **First method starts from leftmost column and moves to rightmost column**
  - **Second method alternates between leftmost and rightmost column**
    - **Last column to be concealed is center column**
- **Complexity of methods is identical, performance may vary**
  - **Alternating column method may give better results on the average**
  - **Center of the frame is usually more difficult to conceal**
  - **Depending on video content, may allow easiest MBs to be concealed first and then use them to conceal more difficult MBs located in center of the frame**

Error Concealment Performance

No EC with FMO (Y-PSNR = 11.52 dB)

JM EC with FMO (Y-PSNR = 29.82 dB)

No EC without FMO (Y-PSNR = 11.13 dB)



JM EC without FMO (Y-PSNR = 24.05 dB)

Error Concealment Performance

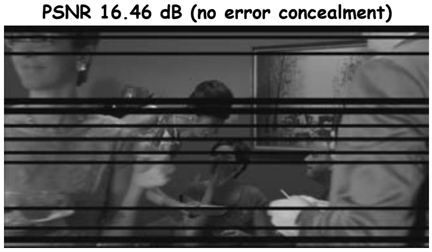
480p SD video: An improvement of *nearly 18 dB* with no observable artifacts

PSNR 9.30 dB (no error concealment)

PSNR 27.12 dB (with error concealment)

1080p HD video: An improvement of *over 24 dB* with no observable artifacts

PSNR 16.46 dB (no error concealment)



PSNR 40.54 dB (with error concealment)

H.264 Summary

H.264 Mode		Results
CABAC (when compared to CAVLC)	**VBR**	Smaller files Shorter encoding time Slightly longer decoding time
	CBR	Better video quality (VQ) Shorter encoding time Slightly longer decoding time
RDO (High Complexity)	**VBR**	Larger files Relative overhead increases with low QP values (not desirable for high quality video) Encoding time increases with low QP values (not desirable for high quality video) Better VQ with low QP values compared to the case without RDO VQ improvement is achieved for each encoded frame
	CBR	Better VQ with higher relative gain for low bit-rates Longer encoding time when compared to the case without RDO
FMO Type 1	**VBR**	Bigger files Relative overhead increases with number of SGs Relative overhead decreases with low QP values (desirable for high quality videos) Better video quality with packet losses than Types 3, 4, and 5
	CBR	VQ decreases in lossless environments compared to the case without FMO

Video Quality (VQ) Assessment

- Measure video quality as perceived by user
  - Not easy since reconstructed image is not meant to be identical to the original
    - Perceptually irrelevant information will be discarded in compression/decompression process
    - What counts as "irrelevant" depends on the viewer's subjective response
  - Three measures: resolution, noise, and overall impression
- Basic metrics can be classified under subjective and objective
  - ITU recommends longer sequences (10 sec where possible) for subjective viewing
  - For subjective comparisons, sequence under test should be presented side-by-side with a sequence generated by JM reference software

Display Resolution versus Encoding Quality

- Higher resolution involves a higher density of pixels
  - 720p and 1080p HD videos have a higher resolution over CIF videos
  - Impact of resolution on VQ dictated by screen size of user device
    - Maximum screen resolution can be set by user device or computer
    - 15-inch laptop monitor may be more suited for 720p than 1080p or QCIF video playback
- Encoding quality is dictated by number of bits representing each pixel
  - A HD video may suffer in VQ if less bits are used for each quantization level
  - A CIF video encoded with a fine quantization level can achieve good VQ on a handheld device

	Resolution
Digital Cinema	2048 × 1536 pixels
High-Definition 1080p	1920 x 1080 pixels
High-Definition 720p	1280 x 720 pixels
Standard Definition 480p	720 x 480 pixels (NTSC), 720 x 576 (PAL)
CIF	352 x 240 pixels (NTSC), 352 x 288 pixels (PAL)
Wide QVGA (used in PSP)	368 x 208 pixels
QVGA (used in iPod)	320 x 240 pixels
QCIF	176 x 120 pixels (NTSC), 176 x 144 pixels (PAL)

Subjective Video Quality Metrics

- Assess actual distortions perceived by viewer
  - –E.g., blockiness, blurriness, ringing artifacts, added high frequency content
  - –Common artifacts: jerky playback, frozen picture
  - –Experiments are performed in a controlled environment
- Cannot be measured easily using quantitative measures
  - –Take into account sensitivity of human perceptual system, which is complex
  - –May not be consistent across all video displays, resolutions, human subjects
    - Children, young adults, seniors may have wide differences in visual perceptions
  - –Averaging subjective ratings of a panel of viewers via a single mean opinion score is clearly restrictive and requires proper calibration
  - –Even if an accurate human visual and perceptual system can be modeled, human intelligence is still required to prevent false alarms/error propagation
    - Humans can decipher the age of the movie by simply checking out the name of the actors/actresses - an old movie is expected to have poorer quality
    - Humans can decipher between deliberate slow motion and problems in playback
    - Humans can decipher between deliberate blurring in background versus blurring on subject
  - –This led to some analysts predicting that great video quality with high-definition TV may not be as important as good choice in video content

Objective Video Quality Metrics

- Attempt to quantify observed video distortions
  - –Can be classified under content-independent and content-dependent metrics
  - –Content-independent metrics
    - Have the same impact on all videos (e.g., number of frames affected by loss of a P-frame is the same irrespective video content)
    - Play an important role in optimizing network transport of videos
  - –Content-dependent metrics
    - Provide measures whose values depend on actual video under consideration
    - Many content-dependent metrics require a reference for their computation

Content-Independent Metrics

TMDR	Time duration for which video is affected during a frame loss
FRAME SIZE	Size of the frame
FRAME TYPE	Type of frame (I, P, or B) that is lost/dropped

Content-Dependent Metrics

MSE	Mean squared error in a frame
PSNR*	Peak Signal to Noise Ratio in a frame
RELATIVE PSNR*	PSNR as compared to highest quality video
SSIM*	Structural Similarity between two images

*Full reference metric that measures image quality based on an uncompressed or distortion-free image as reference

Peak Signal to Noise Ratio (PSNR)

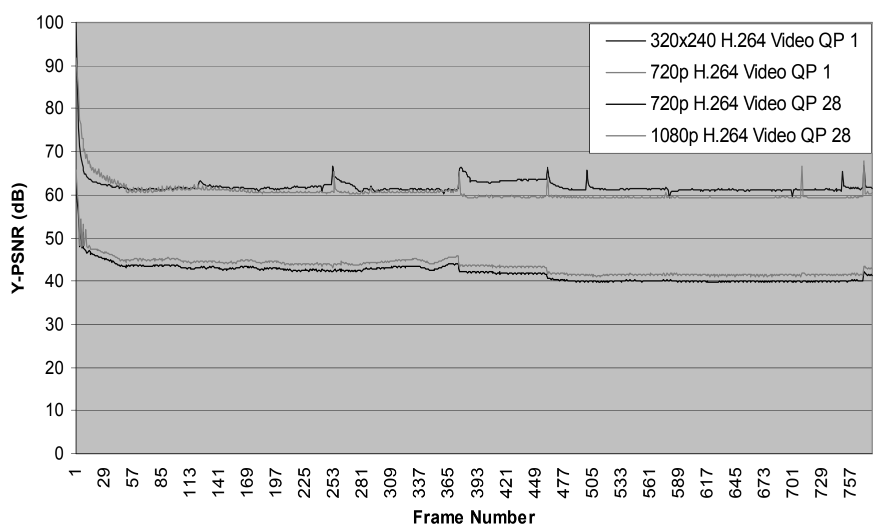
- **A higher PSNR indicates a less noisy signal**
  - In general, a video picture with significant details lowers the PSNR since it is more difficult for the encoder to replicate the original frame
- **Governed by mean squared error (MSE) and number of bits/sample (*B*)**
  - For two *mxn* monochrome images *I* and *K, PSNR* is given by following equation
  - For RGB color images, same equation is valid but MSE is sum over all squared value differences divided by image size and by three
  - Typical values: 30 to 50 dB, where a higher value is better
    - Acceptable values for wireless transmission are between 20 to 25 dB
  - Infinite PSNR: when two images are identical, MSE will be zero
  - Zero PSNR: *I* is completely white and *K* is completely black (or vice versa)
  - Luminance or Y-PSNR: Visual perception most sensitive to luminance

$$PSNR = 20\log_{10}\left(\frac{P_I}{\sqrt{MSE}}\right)$$

$$\text{where } MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left\|I(i,j) - K(i,j)\right\|^2 \text{ and } P_I = 2^B - 1$$

Limitations of PSNR

- **Applies only to videos with the same resolution**
  - Identical videos with different resolutions but same QP value may lead to similar PSNR



Structural Similarity Index (SSIM)

- **Accounts for higher-level structural information**
  - **Some extensions**
    - **Structural Texture Similarity Index**
    - **Color Structural Texture Similarity Index**

$$l(x,y) = \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

Luminance comparison, function of *mean* intensity

$$c(x,y) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

Contrast comparison, function of *variance*

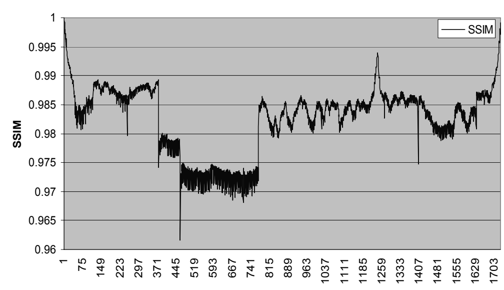
$$s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3}$$

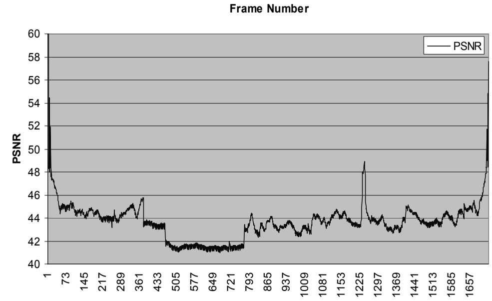
Structure comparison, function of *covariance*

$$\mathrm{SSIM}(x,y) = l(x,y)^{\alpha}\, c(x,y)^{\beta}\, s(x,y)^{\gamma}$$

PSNR versus SSIM

- **Similar characteristics**





Impact of H.264 Information Loss on Video Quality

Foreman H.264 Video, 352 x 288 CIF Resolution

Subjective quality varies from undecodable (X), very bad (1), bad (2), OK (3), good (4)

Test sequences downloadable from http://users.ece.gatech.edu/~benny/foreman*x*.264 (*x* = 0, 1, …, 8)

Coded Bitstream	Description	Subjective Quality
Foreman0	Original H.264 coded bitstream	4
Foreman1	Remove first SPS at 0x557	X
Foreman2	Remove first PPS at 0x587	X
Foreman3	Remove one coded slice IDR (I-slice) at 0x595 (Error concealment not activated)	1
Foreman4	Random cuts within coded slice IDR (I-slice) (Error concealment activated but fails to conceal errors)	1
Foreman5	Remove one coded slice non-IDR (P-slice) at 0x26892 (Error concealment not activated)	2
Foreman6	Remove one coded slice non-IDR (B-slice) at 0x630 (Error concealment not activated)	4
Foreman7	Replace bytes from 0x30080 to 0x30160 with zeros (Error concealment activated but fails to conceal errors)	1
Foreman8	Remove bytes from 0x30080 to 0x30160 (Error concealment not activated)	3

Observable versus Perceptual Visual Artifacts

- **Observable or visible artifacts more quantitative**
  - **Unlike perceptual artifacts, not prone to false alarms**
- **Temporal aspect needs to be accounted for**
  - **Superposition of video frames 1 and 3 reveals obvious artifact in background**

1. Original Video Frame

3. Video frame with *one* artifact on background, PSNR 32.09, QP = 30

2. Video frame with *two* artifacts on subject, PSNR 32.10, QP = 30

4. Video frame with no artifact, PSNR 32.14, QP = 37

Observable versus Perceptual Visual Artifacts

- **In 2 and 3, visible artifacts affect only one frame**
  - In 4, quality for all frames are degraded with lower QP value
  - But there are no observable artifacts
  - Some online streaming systems adjust QP value and resolution dynamically to prevent observable artifacts (VQ and resolution degrade with losses)

1. Original Video Frame

3. Video frame with *two* artifacts on background, PSNR 35.10, QP = 30

2. Video frame with *one* artifact on subject, PSNR 35.03, QP = 30

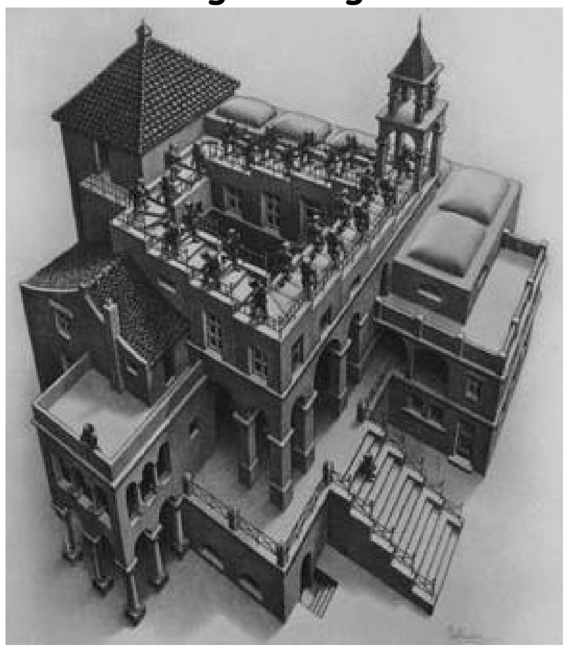
4. Video frame with no artifact, PSNR 35.3, QP = 32

Artifact may well be a censored part!

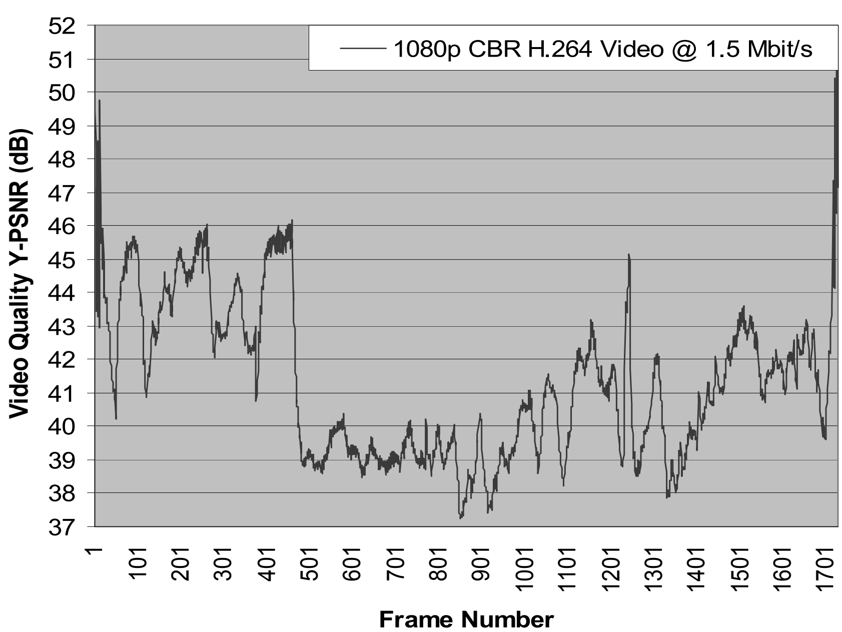
Video Quality (VQ) Assessment

- Practical measurement of video quality may not be accurate or precise
  - Also difficult to quantify
    - Overall picture quality is poor but there are no observable artifacts
    - Subtle changes in frame rate to adapt to network condition
    - Observable artifacts but overall picture quality is good
- CBR encoding
  - Not possible to measure overall VQ
    - Video quality is variable may change dramatically for each frame due to changes in quantization level and content
    - More noticeable with HD
  - Service providers now migrating from CBR to VBR encoding
- VBR encoding
  - Video quality more or less constant
  - High bit rate variability for HD videos
- Artifacts for online streaming and payTV systems vary greatly
  - Freeze frames are common in online streaming
  - Video frame breakup more common in cable/satellite systems
  - Codec plays a part: H.264 versus MPEG-2

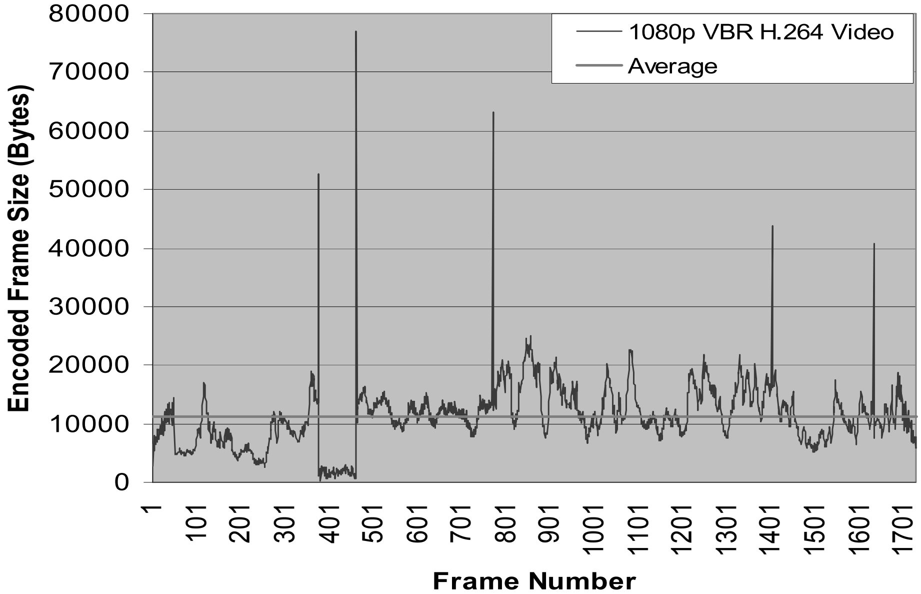
Visual or Engineering Artifact?



Variable Video Quality with CBR HD Encoding

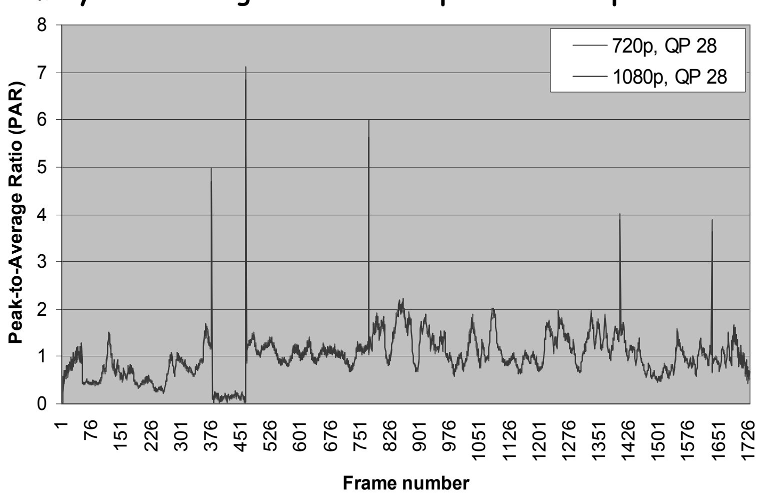


Encoded HD Frame Sizes



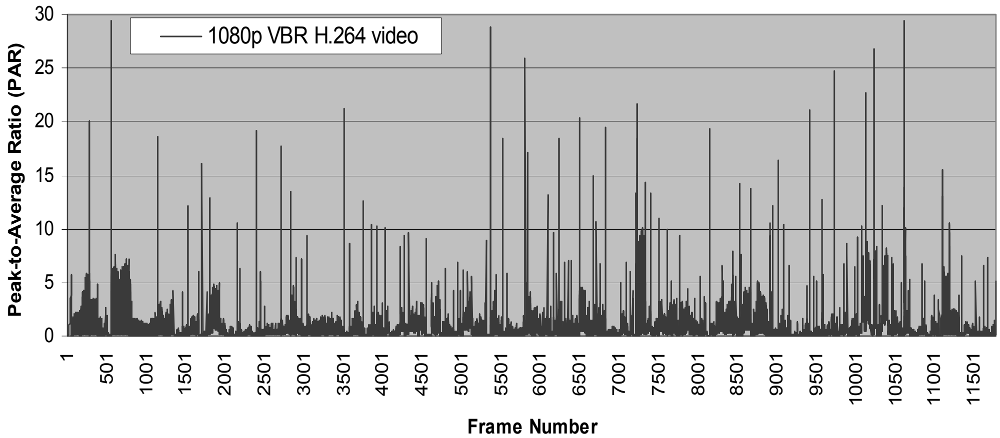
Peak to Average Ratio (PAR)

- **Similar PAR for video with different resolutions but encoded with the same QP value**
  - **–Peaks are normally caused by scene changes (which are intra-coded)**
  - **–Motion normally leads to high PAR but not peaks due to prediction**



High Bit Rate Variability for VBR HD Encoding

- PAR as high as 30

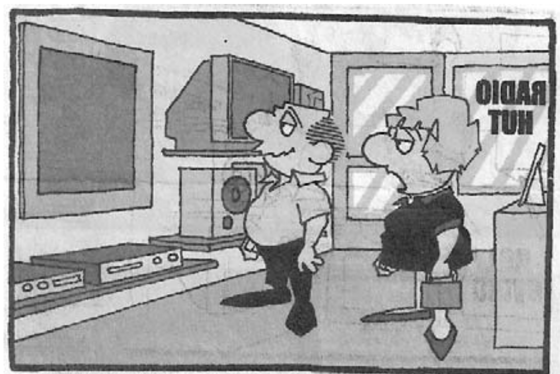


Detecting Scene Changes

- Scene change typically leads to a significant change in frame sizes
  - Can be larger or smaller
  - Can be classified under sudden and gradual scene change
    - Abrupt changes easy to detect as two successive frames are completely uncorrelated
    - Gradual changes are used to enhance quality of video production and are more difficult to detect as difference between frames corresponding to two successive scenes is small
    - Majority of scene change detection methods are proposed for pre-recorded videos
- May use a simple metric to detect scene change (e.g., I-frame size)
  - Threshold of scene change detector crucial to performance
  - High threshold lowers impact because encoder will only detect a few scene changes, although these changes tend to be abrupt scene changes
  - If threshold is too low, encoder starts over more often
    - Prevents real-time video transmission but enables detection of both gradual and abrupt scene changes

Content Quality versus Video Quality

- Content, price, convenience, then quality
  - A 500-channel payTV channel lineup seems limited when compared to the amount of video available on the Internet
  - Many online TV providers allow users to rate the video content, including the commercials
    - Video quality is not rated
    - With CBR systems, video quality cannot be rated consistently
  - Digital cinema provides the best quality but does not enjoy the biggest audience
    - More people are watching online videos even if the video quality is not the best at times



"WHY DO WE NEED A HIGH-DEFINITION TV FOR LOWEST-COMMON-DENOMINATOR PROGRAMS?"

Transport Protocols

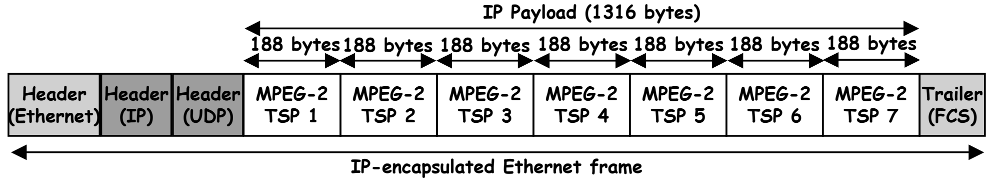
- Real-time Transport Protocol (RTP)
  - Designed to send real-time media (e.g., voice and video) over UDP/IP
  - Supplies information to allow receiver to re-synchronize media
    - For lip syncing or for having text appear at correct time in relation to an image or word
  - Can be configured for low latency
    - Useful for interactive conversations as well as streaming media
  - Data can be encrypted for improved privacy against eavesdropping
  - Can be enhanced for better monitoring, streaming capabilities, code support
    - RFC 3550, "RTP: A Transport Protocol for Real-Time Applications", July 2003
    - RFC 3551, "RTP Profile for Audio and Video Conferences with Minimal Control", July 2003
    - RFC 3984, "RTP Payload Format for H.264 Video", February 2005
- Real-time Transport Control Protocol (RTCP)
  - Companion protocol to RTP that collects statistics on media connection (e.g., bytes or packets sent, lost packets, jitter, round trip delay)
  - Application can use information to judge connection quality and make adjustments as required (e.g., changing from low to high compression)
- Transport Control Protocol (TCP)
  - Loss-free protocol may be better suited for compressed video that is more sensitive to information loss

Real-Time Transport Streaming Protocol (RTSP)

- Client-server protocol for "multimedia remote control over IP"
  - Defined by IETF RFC 2326
  - IP application-level protocol for controlling delivery of multimedia content, similar to SIP or H.323
  - Enables client device to support live or stored web content streaming
- Takes advantage of lower-level protocols to provide complete streaming service over the Internet
  - Complementary protocols include RTP for streaming, RSVP for QoS assurance
- Most suitable for IPTV Multimedia-On-Demand services
  - Provides DVR remote control functions for audio and video streams services
    - For example, content navigation (e.g., Pause and Fast Forward), absolute positioning and programs for later operations
  - Provides means for choosing delivery methods e.g., UDP, multicast UDP, RTP
  - Highly beneficial for both large audience multicasting and real-time Multimedia-On-Demand unicasting

MPEG-2 Transport Stream (TS)

- Provides strict sender/receiver synchronization
  - Program clock reference (PCR) written in TS header
    - Based on local clock timing information at sender
  - Each User Datagram Protocol (UDP) packet can encapsulate up to seven 188 byte transport stream packets (TSPs)
  - Implies that losing 1 UDP packet can result in 7 media packets being lost
- Minimum overhead required to carry a TSP over IP is about 3.4%
  - IP/Ethernet encapsulation adds 46 bytes of overhead per Ethernet frame
    - With 7 TSPs giving 188 x 7 or 1316 bytes, % overhead = 46/1362 x 100%
  - Number of TSPs per IP packet may vary
    - Encapsulating more TSPs per IP packet reduces % overheads but increases network jitter
    - Jumbo packets with up to 47 TSPs per IP packet may increase network utilization to 99%



Media Delivery Index (MDI)

- Quality indicator of video transport performance
  - Employs network level (IP) measurements to identify and measure jitter and packet loss
  - Independent of video encoding scheme
  - Lightweight and scalable alternative to measurements that decode and examine the video itself
  - Specified in RFC 4445
- Media Loss Rate (MLR)
  - Relates to MPEG packets being lost and their rate
- Delay Factor (DF)
  - Relates to current buffer size required for a flow at that point in the network
  - A value in milliseconds and is dependant on the bit rate of the stream being monitored
  - Since network jitter has a different impact on streams with different bit rates, value decreases for higher bit rates
  - Can warn impairments that result in unacceptable video delivery and on conditions that result in unacceptable network margin before VQ is impacted

Link Quality Enhancement for H.264 Video Transport

- Delay Jitter (DJ) improves as delivery rate increases
  - Buffer size does not influence DJ as much as delivery rate
  - Moderate-sized buffers (greater than average frame size) reduce DJ to a reasonable value while maintaining losslessness of transmission
- Packet Loss Rate (PLR) is affected by small buffers
  - PLR indicates proportion of lost frames (e.g., frames that are delivered late)
  - As delivery rate increases, more frames are expected for delivery but small buffer restricts availability of these frames, thus PLR suffers

Receiver Buffer Size = 20 KB		
Delivery Rate (Mbps)	DJ (bytes/sec)	PLR (Ratio of lost frames to total frames)
9.6	15.39	0.048
8.0	24.27	0.048
6.4	37.59	0.045
4.8	59.79	0.021
3.2	104.18	0.003
1.6	237.39	0
Receiver Buffer Size = 60 KB (average frame size = 59.2 KB)		
Delivery Rate (Mbps)	DJ (bytes/sec)	PLR
9.6	15.22	0
8.0	24.07	0
6.4	37.39	0
4.8	59.49	0
3.2	103.68	0
1.6	236.38	0

Delivery Rate Computation

- Important for maintaining smooth playback
  - Prevents receive buffer underflow or overflow
- Useful when streaming with TCP
  - Transport timing synchronization are absent in these protocols
- Requires number of frames in video file or segment to be determined
  - Average frame size or segment = file size or segment/number of frames
  - Suppose video file size or segment is 3 Mbytes with 300 frames
    - Then average frame size or segment = 3 Mbytes/300 frames or 10 Kbytes
    - Frame rate = 30 frame/s (or at least 25 frame/s for uninterrupted playback)
    - Desired average receiving rate = 10,000 x 8 x 30 = 2.4 Mbit/s or 300 Kbytes/sec

TCP Optimization

- Counters high network latency
  - One-way Internet latency can be as high as 500 ms
  - Need to modify TCP window size and use window scaling
    - Network throughput will improve 100-500% on WAN links
    - Less impact on performance in LAN environments
- Two methods for Linux
  - Modify parameters on a running system by modifying values in /proc/sys/net/core/ and /proc/sys/net/ipv4/
  - Modify parameters permanently by changing values in Linux kernel sources and compiling the kernel
- Registry settings for Windows
  - Modify SystemKey
    - [HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\Tcpip\Parameters]
    - Value Name: TcpWindowSize
    - Data Type: REG_DWORD (DWORD value)
    - Value Data: 0 - 0xFFFF (default = 8760 or 0x00002238 for Ethernet)
  - Modify GlobalMaxTcpWindowSize value
    - Value Data: 0–0x3FFFFFFF (0x00007FFF = 32767)

TCP Optimization

- **Receive TCP window**
  - Specifies number of bytes a sender may transmit without receiving an acknowledgment (equivalent to amount of bytes in receiver's memory buffer)
  - Reducing the TCP window size effectively causes an acknowledgment to be sent to the sender for data received in a shorter period of time
  - Reduces probability that sender will time out while waiting for an acknowledgment
  - However it will also increase amount of backlog traffic at sender, thereby lowering throughput
  - In general, larger receive windows will improve performance over high delay, high bandwidth networks
- **For greatest efficiency, receive window should be an even multiple of TCP Maximum Segment Size (MSS)**
  - Default setting of 64 KB fine for most LANs, too low for Internet connections
  - Value should be set to 256 KB for T1 lines or lower and 2 to 4 MB for T3, OC-3 or even faster connections
  - Optimal buffer size = 2 x bandwidth x delay

Enhanced TCP

- **Microsoft's Compound TCP**
  - Next Generation TCP/IP stack that optimizes sender-side throughput
  - Together with receive window auto-tuning, can increase link utilization and improve performance for large bandwidth-delay product connections
    - Optimized for TCP connections with large receive window size
    - Aggressively increases amount of data sent at a time, yet ensures that its behavior does not negatively impact other TCP connections
- **Enhancements for high-loss environments**
  - RFC 2582: The New Reno Modification to TCP's Fast Recovery Algorithm
  - RFC 2883: An Extension to the Selective Acknowledgement (SACK) Option for TCP (defined in RFC 2018)
    - Reduces number of retransmissions to improve overall throughput
  - RFC 3517: A Conservative SACK-based Loss Recovery Algorithm for TCP
    - Performs loss recovery when duplicate acknowledgements have been received
  - RFC 4138: Forward RTO-Recovery (F-RTO): An Algorithm for Detecting Spurious Retransmission Timeouts with TCP and the Stream Control Transmission Protocol (SCTP)
    - Prevents unnecessary retransmission of TCP segments when there is a sudden or temporary increase in the round-trip time (RTT)

Multicast Video Streaming

- Allows efficient delivery of streaming video to thousands of receivers by replicating packets throughout network
  - Problems arise when node is located far away from multicast publishing points
    - Streaming video that uses interframe compression require a reference frame
    - Out-of-order video packets or missing reference frame may cause video to freeze
  - To deal with this problem, one can reproduce the multicast closer to the user
    - A much better solution is to employ peer-to-peer multicast streaming
    - See IEEE JSAC Vol. 25, No. 9, "Advances in Peer to Peer Streaming Systems", Dec 07

Factor	Influence on Performance
Number of Customer Movie Requests	Initial increase in request blocking can be reduced with multicast because requests are grouped
Buffering at Set-top	More required for multicast compared to unicast
On-Demand Nature	Near on-demand, instead of true on-demand for unicast
System complexity	Higher compared to unicast
Total number of channels	Multicast increases capacity further because each channel can satisfy multiple requests

Peer-to-Peer (P2P) Streaming

- Accounts for over 80% of Internet backbone traffic
  - Disruptive technology, just like wireless access
  - Overcomes current lack of IP multicasting support by major ISPs
  - Napstar is the first popular peer-to-peer file sharing platform
    - Invented by Shawn Fanning in 1998 while he was a college student in Northeastern University in Boston
- P2p voice applications such as Skype are challenging traditional VoIP
  - Skype captured a significant portion of international voice calling
    - 2 million users first 3 months, 1 million simultaneous subscribers 1 year later
    - FCC considering regulating VoIP but Skype remains unregulated
  - Wireless Skype now emerging
    - Partnered Boingo to provide voice over Wi-Fi service for 18,000 hotspots
  - Skype now provides p2p video conferencing
- Online p2p TV
  - My p2p TV (http://www.myp2p.eu)
    - Free live sports programs supported by multiple video players and streaming platforms: TVAnts, Sopcast (uses p2p), Mediaplayer, VLC, Ustream
  - Typical CIF resolution data rates range from 300 to 500 Kbit/s

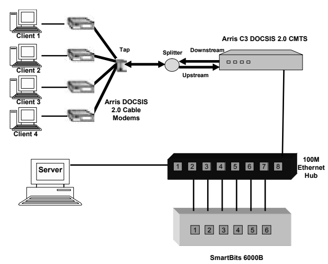
Peer-to-Peer (P2P) Streaming

- Well suited for long-distance (e.g., transatlantic) transmission
  - –Higher rates are needed but highly salable to increasing number of users
- Peer nodes need to contribute upload bandwidth
  - –Relieves bandwidth bottleneck at video source
- Spread of worm viruses can be very rapid
  - –Exponential data dissemination capability can be exploited to halt spreading
  - –However, dynamic participation of peer nodes can reduce effectiveness
- Port-hopping capabilities
  - –Bypasses port controls (e.g., port 80 for http, port 20 for ftp)
  - –Makes it difficult to manage traffic
- Prevalent star topology of access networks creates a local bottleneck
  - –All traffic from end-users directed to hub or central office
    - Cable/telco operators may filter p2p traffic if there is evidence of oversubscription
  - –Problem aggravated if bandwidth for upstream or downstream links is limited
- Mesh architectures provide a better match for p2p applications
  - –Very robust to failures as well as high churn rate of participating peers

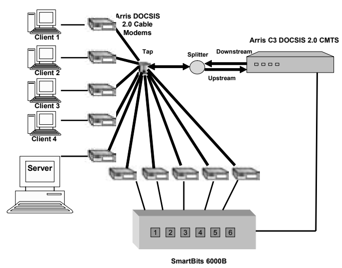
BitTorrent P2P Protocol

- Leading P2P distribution protocol
  - –Bitlet launches BitTorrent video streaming
  - –Employs two kinds of files
    - –data file (content file)
    - –.torrent metainfo file (small file that provides tracker location and data file description such as data file length, piece ID, hashing information)
- Basic mechanisms
  - –Publish
    - Generate a BitTorrent file and run a tracker server
  - –Join
    - Contact a centralized tracker server, obtain list of peers
  - –Piece
    - Data file is broken down into smaller pieces with fixed size
    - Each downloaded piece is reported by all participating peers
  - –Piece selection
    - Rarest first, if not available, then random first
  - –Fetch
    - Download file pieces from peers
    - Upload file pieces to peers

BitTorrent P2P File Sharing

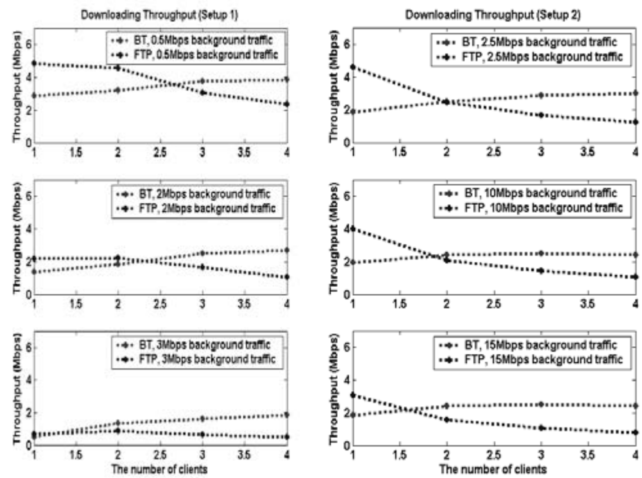


Experimental Setup 1: Background Traffic on the Backbone



Experimental Setup 2: Background Traffic on the Upstream

BitTorrent (BT) peer-to-peer file distribution performs better than the conventional file transfer protocol (ftp) when (1) the background traffic increases (2) the number of clients (users) increases



Throughput is computed by averaging the overall traffic load with the number of clients. Individual traffic load variance is negligible. Overall traffic load for BT and ftp are the same.

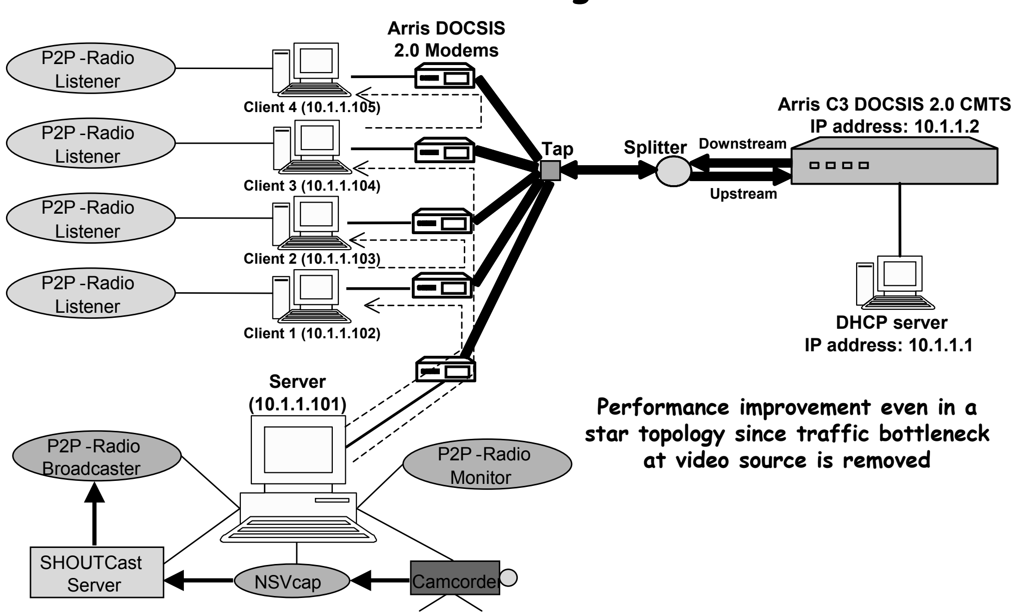
Microsoft Research Live P2P Broadcast System

- **Components**
  - –p2p real-time communication library with a unified programming interface, a distributed NAT traversal module and a congestion control module
  - –p2p broadcast service that includes a video publish server, user registration server, and watching client module
  - –Performance monitor system that can report online statistics and analyze performance of p2p overlay
  - –Management system for managing number servers and video channels

P2P Radio Streaming

- Audio/video p2p Internet "broadcast"
  - Similar to a radio station broadcast
  - Can deliver audio or video live or on-demand
  - Supported formats
    - MPEG Layer 3 (MP3), Ogg Vorbis, Nullsoft Streaming Video (NSV)
- Websites
  - http://www.theora.org
  - http://p2p-radio.sourceforge.net

Peer-to-Peer Video Streaming Communication Paths

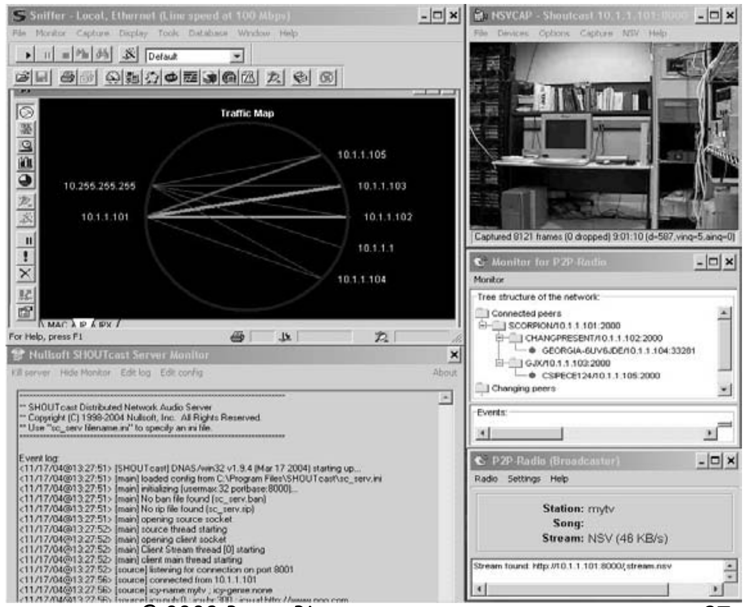


Performance improvement even in a star topology since traffic bottleneck at video source is removed

Live Peer-to-Peer Video Streaming Snapshot

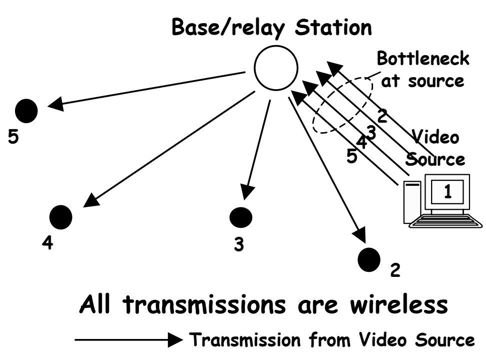
- **Experiment performed on cable network with symmetrical bandwidth**

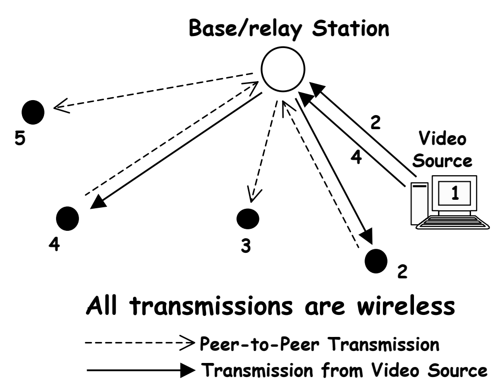
Computer	IP Address
Peer 1 (Sender)	10.1.1.101
Peer 2	10.1.1.102
Peer 3	10.1.1.103
Peer 4	10.1.1.104
Peer 5	10.1.1.105



Wireless P2P Video Streaming

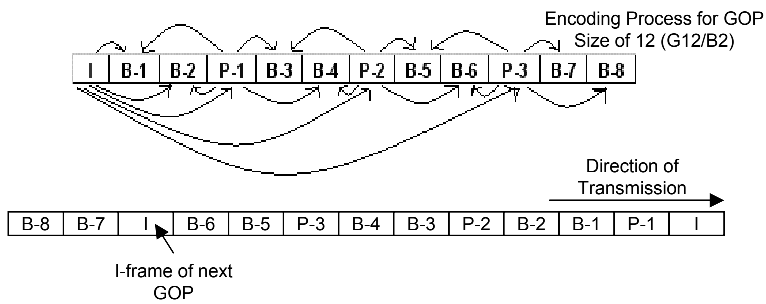
- **Relieves traffic bottleneck at video source, even in a star topology**
  - **In a multihop mesh topology, bottleneck at video source *and* base station are removed**





MPEG Encoding and Transmission

- Video encoded as a stream of I, P, and B frames
  - Each frame contains varying amount of information regarding video content
  - I-frames are still images that contain maximum amount of information in video representation
  - P-frames contain predictive information that is used to reconstruct B-frames
  - B-frames are smallest of all 3 frame types
    - Contain least amount of information
    - In a group of pictures (GOP), there are more B-frames than P and I frames
    - When present, tends to cause higher bit rate variability



B-Frame Dropping

- B-frames can sometimes be dropped to conserve bandwidth resources and reduce bit rate variability
  - Dropping B frames less harmful because subsequent frames transmitted following a B-frame are not dependent on that B frame
    - Predictive information from P-frames (in both MPEG-2 and H.264) can be used to reconstruct dropped B-frames, effective for small number of dropped B-frames
    - For a high number of B-frame loss, error concealment using H.264 can compensate for loss in motion vectors
  - B-frames only contain temporal information and so their loss only causes motion artifacts
    - May be difficult to notice unless the loss rate is very high
    - In some cases (e.g., low motion video), all or a large number of B-frames in entire video can be removed without introducing visible artifacts
  - Random frame loss can cause artifacts randomly in both temporal and spatial domains
    - More observable at lower loss rates
  - Reducing bit rate variability equivalent to smoothing encoded video bit stream
    - Can also mitigate impact of abrupt scene change
  - MPEG encoder cannot delete all P and B frames, otherwise no compression

Video Smoothing

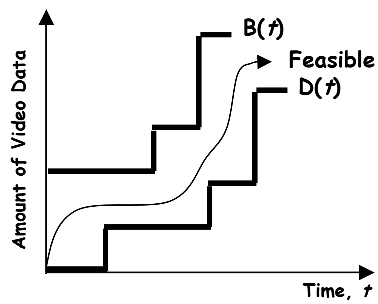
- Reduces bit rate variability of VBR video stream
  - Can be online (live) or offline
  - Can be performed with or without rate control
- Different from network traffic shaping
  - Scheduling is performed to ensure smooth video playback
- Smoothed video segments must be demarcated properly in order for decoder to recover original frames
  - Smoothed segments can contain one or more individual video frames if the sizes of the frames are small
  - Segments can be fragmented to multiple segments if frame sizes of are large
  - Size of smoothed segment must be compatible with packet size of transport protocol (e.g., RTP/UDP, TCP)

Video Smoothing with Rate Control

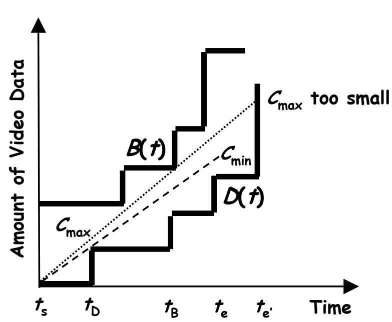
- Factors affecting peak rate of smoothed video transmission
  - Startup latency, transmission delay, buffer size
- Most critical factor is client device (e.g., STB) buffer size
  - Large buffers may be required since VBR video can be very lengthy and bursty (both short-term and long-term)
  - With small buffers, smoothed video streams continue to exhibit long-term, slow-time rate variability
  - An optimal video slice transmission rate is required for both online and offline smoothing
- Start-up latency
  - First frame of MPEG compressed video (an I frame) is much larger than immediately subsequent B and P frames
    - Ratio of size of I to B slices in H.264 can be as high as 10
  - Start-up latency needed to reduce peak rate of initial segment
    - Some online video systems mitigate the problem by playing a commercial at the start, which typically contains more text than regular TV episodes or movies
- Video pausing
  - Used when transmission exceeds rate limit

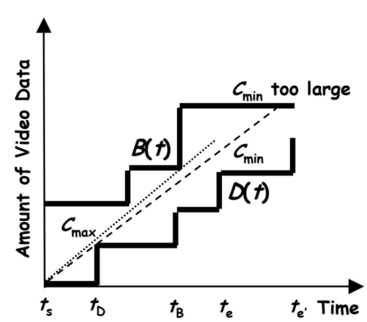
Video Smoothing with Rate Control

- **Optimum rate based on**
  - **Upper bound B(*t*)**
    - **Maximum cumulative data that can be received by client over time *t***
    - **Value depends on buffer space at client device, available video frames at server, and desired channel change latency**
  - **Lower bound D(*t*)**
    - **Cumulative data that must be sent by server over time *t***
    - **Ensures continuous playback at client device**
    - **Value depends on desired minimum buffer utilization for ensuring low packet losses**



Video Smoothing with Rate Control



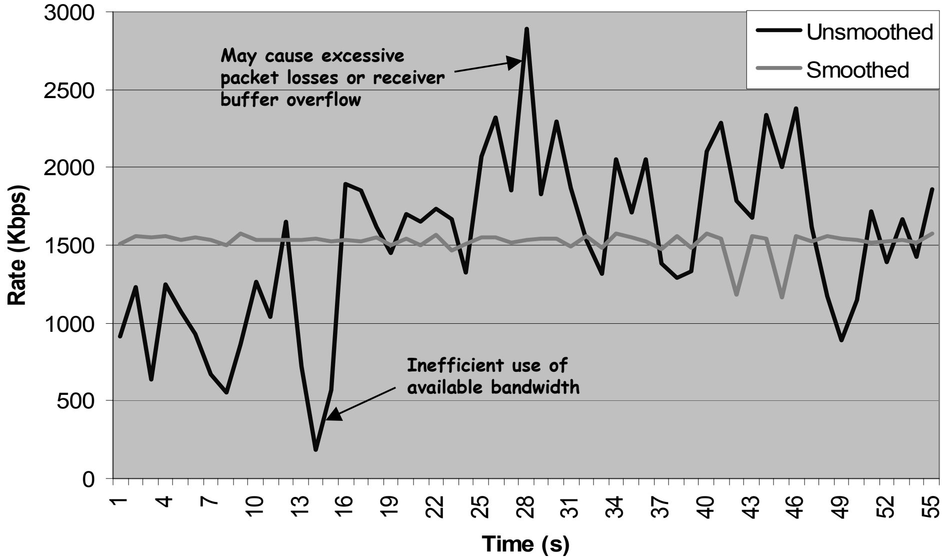


$$C_{max}(t) = \frac{B(t) - [D(a) + q]}{t - a} \qquad C_{min}(t) = \frac{D(t) - [D(a) + q]}{t - a}$$

t_B: Latest time at which the STB buffer is full when server transmits at C_{max} over [*a*, *b*] starting with an initial buffer level *q*.

t_D: Latest time at which the STB buffer is empty when server transmits at C_{min} over [*a*, *b*] starting with initial buffer level *q*.

HD Video Smoothing

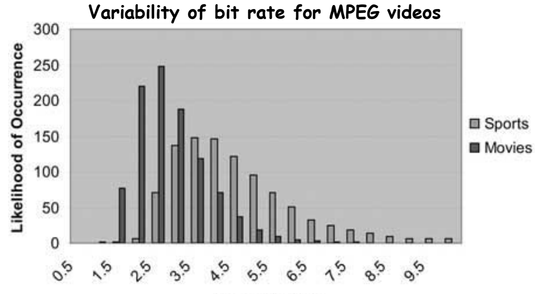


Smoothed and unsmoothed VBR HD videos are encoded using H.264 and have the same file size

Statistical Multiplexing

- Can increase number of video channels within a fixed channel by exploiting inherent variations in instantaneous bit rates
  - Picture quality is kept constant
  - If one channel is demanding high bit rate, it is likely that other channels have capacity to spare (see figure below)
  - A large number of aggregated streams tends to "smooth" to a normal distribution
    - Based on central limit theorem
    - Unlike per stream buffer-based smoothing, does not introduce delay

Variability of bit rate for MPEG videos



Coefficient of Variability for Multiplexed H.264 Videos

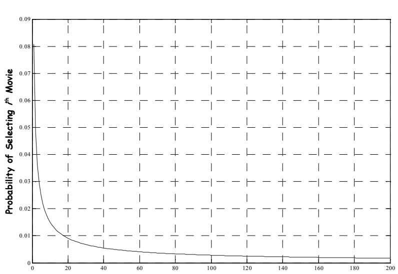
- Total number of frames for multiplexed videos equals individual videos
  - *Multiplexed 1* = Terminator 2 + Sony (QP = 10)
  - *Multiplexed 2* = Terminator 2 + Sony (QP = 28)
  - *Multiplexed 3* = Terminator 2 + Sony + From Mars to China + Horizon Talk show (QP = 28)

Movie	Standard	QP	GOP		I-frames		P-frames		B-frames	
			CoV	Std. Dev.	CoV	Std. Dev.	CoV	Std. Dev.	CoV	Std. Dev.
Terminator 2	MPEG-4 AVC	10	0.3074	3.52E+06	0.2710	4.15E+05	0.2913	3.68E+05	0.3698	2.83E+05
		28	0.5442	4.82E+05	0.4374	8.96E+05	0.5602	6.84E+04	0.9159	3.55E+04
	MEPG-2	10	0.4493	5.80E+05	0.5291	9.28E+04	0.5524	8.87E+04	0.6769	5.37E+04
		15	0.4159	4.15E+05	0.5362	7.58E+04	0.5718	7.41E+04	0.6484	3.80E+04
Sony	MPEG-4 AVC	10	0.4102	3.74E+06	0.4478	9.18E+04	0.4253	4.82E+05	0.5607	2.61E+05
		28	0.5472	5.38E+05	0.5622	2.43E+05	0.7302	9.43E+04	1.1888	2.42E+04
	MEPG-2	10	0.4722	5.86E+05	0.7111	2.15E+05	0.8772	1.28E+05	0.6441	4.03E+04
		15	0.4194	3.67E+05	0.7147	1.65E+05	0.9849	9.93E+04	0.6188	2.65E+04
From Mars to China	MPEG-4 AVC	28	0.5722	1.11E+06	0.5194	3.84E+05	0.6641	1.66E+05	0.9365	5.27E+04
Horizon Talk show	MPEG-4 AVC	28	0.3429	2.11E+05	0.2796	9.37E+04	0.5239	3.31E+04	0.8872	9.93E+03
Multiplexed 1 : Terminator 2 + Sony with MPEG-4 AVC, QP=10			0.5472	5.34E+06	0.5622	9.76E+05	0.7302	6.16E+05	1.1888	3.96E+05
Multiplexed 2 : Terminator 2 + Sony with MPEG-4 AVC, QP=28			0.3926	7.34E+05	0.3867	2.48E+05	0.4612	1.16E+05	0.7204	4.27E+04
Multiplexed 3 : Terminator 2 + Sony + From Mars to China + Horizon Talk show with MPEG-4 AVC, QP=28			0.2678	1.03E+06	0.3264	5.08E+05	0.3312	1.59E+05	0.4556	4.80E+04

$$CoV = \frac{\sigma_{frame}}{\overline{X}_{frame}}$$

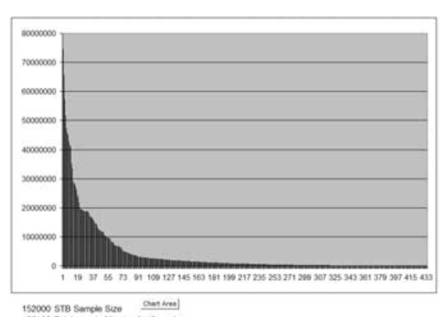
Modeling Popularity of Movies and Video Channels

- Zipf-like (long-tail) distribution
  - Can be used to characterize popularity of different movies and video channels
  - Access pattern predictability removes need for a more complex algorithm



Zipf function with 200 movies and skew factor q = 0.271. Can be easily scaled to fit closely with actual data. Based on the fetch-at-most-once model. It can be shown that this model fits the data more accurately than the fetch-repeatedly model.

Service ID	Aggregate Viewing Minutes	Min per Day
1	74,391,837	34.96
2	65,636,334	30.84
3	57,167,771	26.86
4	51,591,374	24.24
5	47,214,135	22.19
6	45,827,624	21.54
7	44,971,682	21.13
8	42,719,368	20.07
9	41,623,165	19.56
10	40,795,084	19.17
11	35,242,590	16.56
12	33,486,559	15.74
13	29,028,231	13.64
14	28,409,579	13.35
15	28,202,088	13.25
16	27,103,379	12.74
17	25,969,624	12.20
18	24,437,261	11.48
19	23,540,362	11.06
20	21,740,764	10.22
21	20,129,669	9.46
22	19,448,868	9.14
23	19,263,275	9.05
24	19,199,877	9.02
25	18,963,005	8.91
26	18,850,166	8.86
27	18,783,985	8.83



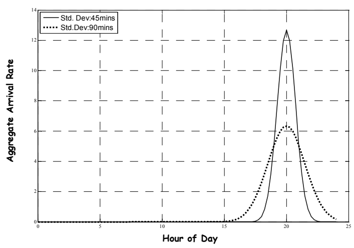
Source: Mark Davis, BigBand Networks, VP Engineering (Network Solutions Group), March 2007.

$$p_i = \frac{1}{i^{1-q} \sum_{j=1}^{L} \frac{1}{j^{1-q}}}$$

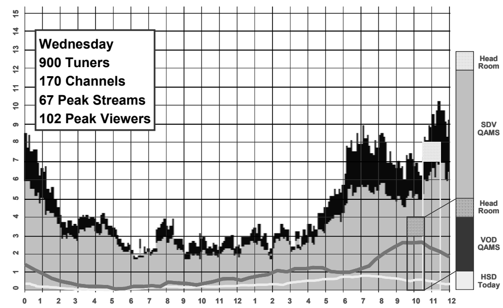
q = skew factor (q = 1 => uniform distribution)
L = number of selectable video programs

Time-Dependent Request Rate

- **Predictability of request rate can be exploited**
  - **Aggregated video streams at time *t* can be modeled using Normal distribution**
    - **Movie duration uniformly distributed between 90 and 120 min, ~60 min for episodes**
  - **Removes need for a more complex algorithm**



Multiple peaks (shifted in time) can be generated to fit actual profile below although this may not be necessary – working solely on the *highest* peak (corresponding to the highest demand) may be sufficient since this will determine the maximum bandwidth requirements.



Source: Mark Davis, BigBand Networks, VP Engineering (Network Solutions Group), March 2007.

$$\lambda_{agg}(t) = \frac{N\frac{\lambda}{7}}{\sigma\sqrt{2\pi}} e^{-(t-\mu)^2/2\sigma^2}$$

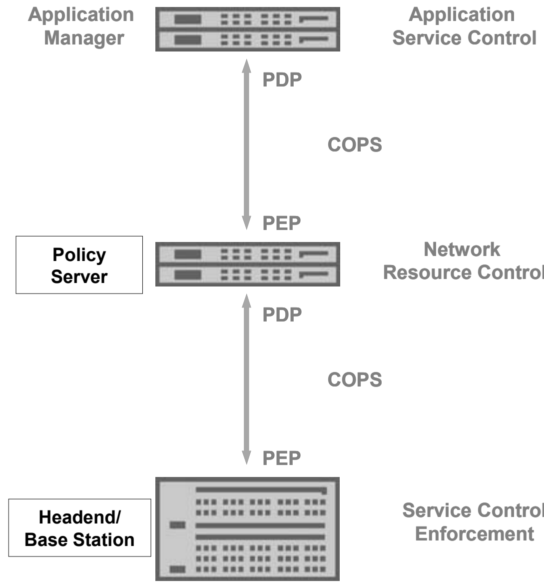
λ (user request rate per week) = 2,
σ = 45 minutes,
μ = 20 hours (or 8 p.m.),
N (number of subscribers) = 5,000.

Self-Similarity (Long Range Dependence) of H.264 Videos

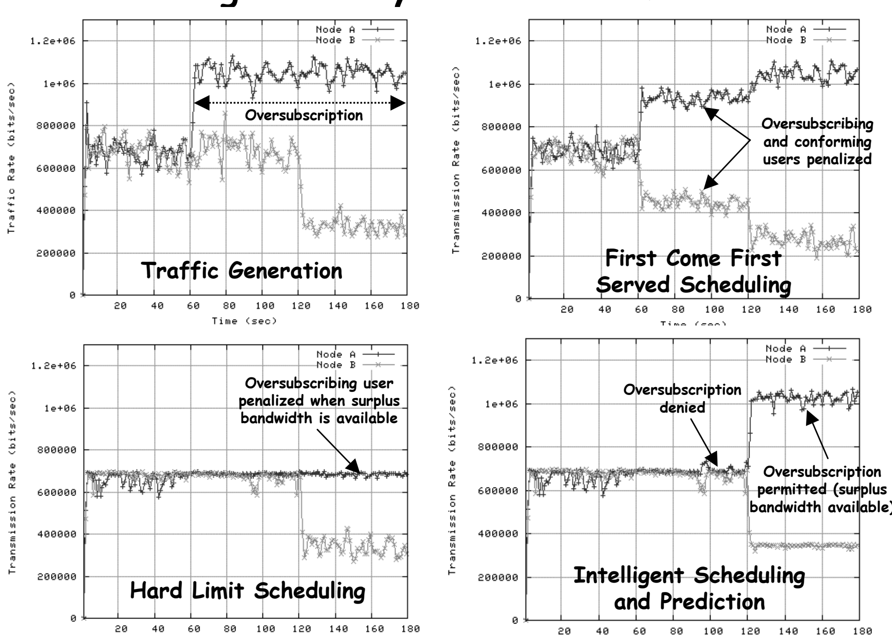
- **Significant impact on network performance**
  - **Losses and delays considerably larger for video traffic with high degree of long range dependence due to its burstiness over a wide range of timescales**
- **Hurst parameter**
  - **Metric for measuring degree of long-range dependence and burstiness**
  - **Long-range dependence properties appear strong for all quality levels of H.264 encoded videos**
    - **Appear weaker for MPEG-2 videos due to lower frame size variability compared to H.264**
- **Compressed video traffic may tend towards clustering and becomes less predictable as number of video streams increases**
  - **Compare Poisson distributions that become smoother as volume increases**
    - **Thus, for multiplexed video streams, larger buffers may be needed to cater for more extreme traffic-burst scenarios**
  - **Turns out that the coefficient of variability (CoV) reduces as more H.264 video streams are multiplexed**
    - **A larger number of aggregated H.264 videos actually tends to "smooth" due to weaker long-range dependency**

Bandwidth Management

- Key elements
  - Collect, store, and manage incoming traffic
  - Employ intelligent algorithms to schedule and predict bandwidth resources via COPS policy server
    - Contrast bandwidth caps that hard limits bandwidth usage
- Policies can be based on time, volume, and location
  - Intelligently control bandwidth-intensive applications
  - Mitigate serious security problems such as worms viruses, DoS attacks



Intelligent Policy-Based Resource Allocation



Next-Generation Video

- Super Hi-Vision
  - Experimental digital video format proposed by NHK, BBC, RAI
  - Also known as Ultra High Definition Video (UHDV)
- Main specifications:
  - Resolution: 7680 × 4320 pixels (16:9) -> 33 million pixels
    - 16 times higher resolution than 1080p HD video (4 x 1920 x 4 x 1080 pixels)
  - Frame rate: 60 frame/s
  - Bandwidth: 600 MHz, 500 - 6600 Mbit/s
- On Dec 31, 2006, NHK demonstrated a live relay over IP for display over a 450 inch (11.4 m) screen
  - Video was compressed from 24 Gbit/s to 180 – 600 Mbit/s
  - Audio was compressed from 28 Mbit/s to 7 – 28 Mbit/s

Summary

- Broadband video a key application
  - 80 million U.S. households subscribe to services of cable TV companies, telcos and satellite TV providers at the end of 2008
- For broadband video to succeed
  - Proprietary set-top boxes should not be the only mode of TV connection
    - Content providers and consumer electronics vendors teaming to provide OTT services to deliver content directly to the TV, reducing or removing reliance on STBs
    - Partnership yields a highly differentiated "product"
    - Zero upfront costs to consumer to watch video
  - Content must be portable
    - Convenient anytime, anyplace access to content
      - With the exception of live content, content should be accessible without a predetermined schedule
    - OTT video content suppliers to play an important role in wireless access networks
    - Networked digital video recording
      - Allows subscribers to record programs in a central data center, rather than in their set-top boxes
      - Reduces CAPEX, fewer truckrolls, and more storage capacity for consumers
  - Video delivery must achieve high-quality user experience
    - Accessing digital movies should be as simple as flipping a channel

Summary

- Video is driving the growth in Internet traffic
  - Proliferation of video and peer to peer applications
  - Continued, or even accelerated growth in demand for video downloads
    - Equally as exciting is demand for being able to enjoy these applications while on the go
- Netflix versus HBO: Which is the better value?
  - Online video versus payTV video on demand
  - Online video versus digital video recording
- Existing cable network TV programs are pushed online
  - Internet-only content and bonus material are pushed directly to the TV
- Migration from payTV to online video
  - Both consumers and advertisers are moving in the same direction
- Problem with live shows e.g., sports programs
  - Free p2p video websites offer a solution
  - Sports channels such as ESPN are offering shows online with subscription
  - More websites expected to offer subscription in future

Summary

- Open access is emerging
  - Managed networks may become become underutilized in the same way as the public switched telephone network
    - Managed networks are limited in reach, the Internet is global
    - Will erode the business model for payTV
  - Simplifies TV connectivity
    - Open broadband STBs
    - A unified standard for authoring a TV experience from the Internet
  - Virtual operator services
- HD video transport over the Internet is a huge challenge
  - End-to-end one-way delays range from 20 to 500 ms
    - Losses dues to congestion can be as high as 20%
    - FEC for individual packets may not be effective
    - Feedback control may not be responsive (just like in satellite communications)
  - More challenging than terrestrial wireless
    - Typically deals with channel errors in local access
    - Delay not significant since packet transmission is local
      - Errors can be corrected quickly using FEC or packet retransmission

Summary

- Improving the quality of Internet video transmission
  - Manage packet losses and video artifacts
  - H.264 a key standard
    - Reduces bandwidth requirements
    - No channel feedback required with error concealment
      - Applicable to broadcast, multicast, unicast networks, and end-to-end Internet streaming
    - Built-in ability to detect and track video artifacts, and collect statistics
      - Variable length decoder can detect missing MBs with no false alarms
  - Bandwidth conservation
    - Reduces occurrence of packet losses
  - Error concealment
    - Maintains video encoding rate and video quality in the presence of burst or random losses
    - Conceals errors caused by packet losses during video decoding
    - Many commercial systems do not employ error concealment
- Bandwidth management challenges
  - End-to-end QoS-guarantees, seamless connectivity, effective policy/traffic management
  - Applications must not be discriminated
  - Fixed bandwidth caps not a scalable solution

References

- Compression standards
  - Advanced Video Coding for Audiovisual Services, ITU-T Rec. H.264 | ISO/IEC 14496-10 Version 8, 2007
  - VC-1 Compressed Video Bitstream Format and Decoding Process, SMPTE 421M, 2006
- Reference software
  - http://iphome.hhi.de/suehring/tml (H.264)
  - http://wiki.multimedia.cx/index.php?title=Understanding_VC-1 or buy from SMPTE (VC-1)
- Network transport
  - RFC 3984 (H.264) and RFC 4425 (VC-1)
- 3D Television
  - http://en.wikipedia.org/wiki/3D_television
- Streaming media
  - L. Bouthillie, "Streaming versus Downloading Video", July 2003 (http://www.streamingmedia.com/article.asp?id=8456&page=2&c=11)
- B. Bing, *Broadband Video Networking*, Artech House, 2010, to appear
- Tutorial will be expanded, recorded, and posted in www.comsoc.org