

# Audio and Video Over IP

Paul Streffon CTS-D CTS-I  
Staff Instructor  
InfoComm International

# Presenter



Paul Streffon, MBA, CTS-D, CTS-I, is an InfoComm Academy Staff Instructor. With more than 25 years in the AV industry, he has worked in many audiovisual areas: production, technical, sales and marketing, management and education.

# About InfoComm International®

- Leading non-profit association serving the professional AV communications industry worldwide
- Member-driven
- Everything Audiovisual



# CTS<sup>®</sup> Certification

- Certified Technology Specialist
  - CTS
  - CTS-I
  - CTS-D
- Renewal Units – 2 Units for this seminar
  - Check [www.infocomm.org/cts](http://www.infocomm.org/cts) for details



# Compression and Encoding

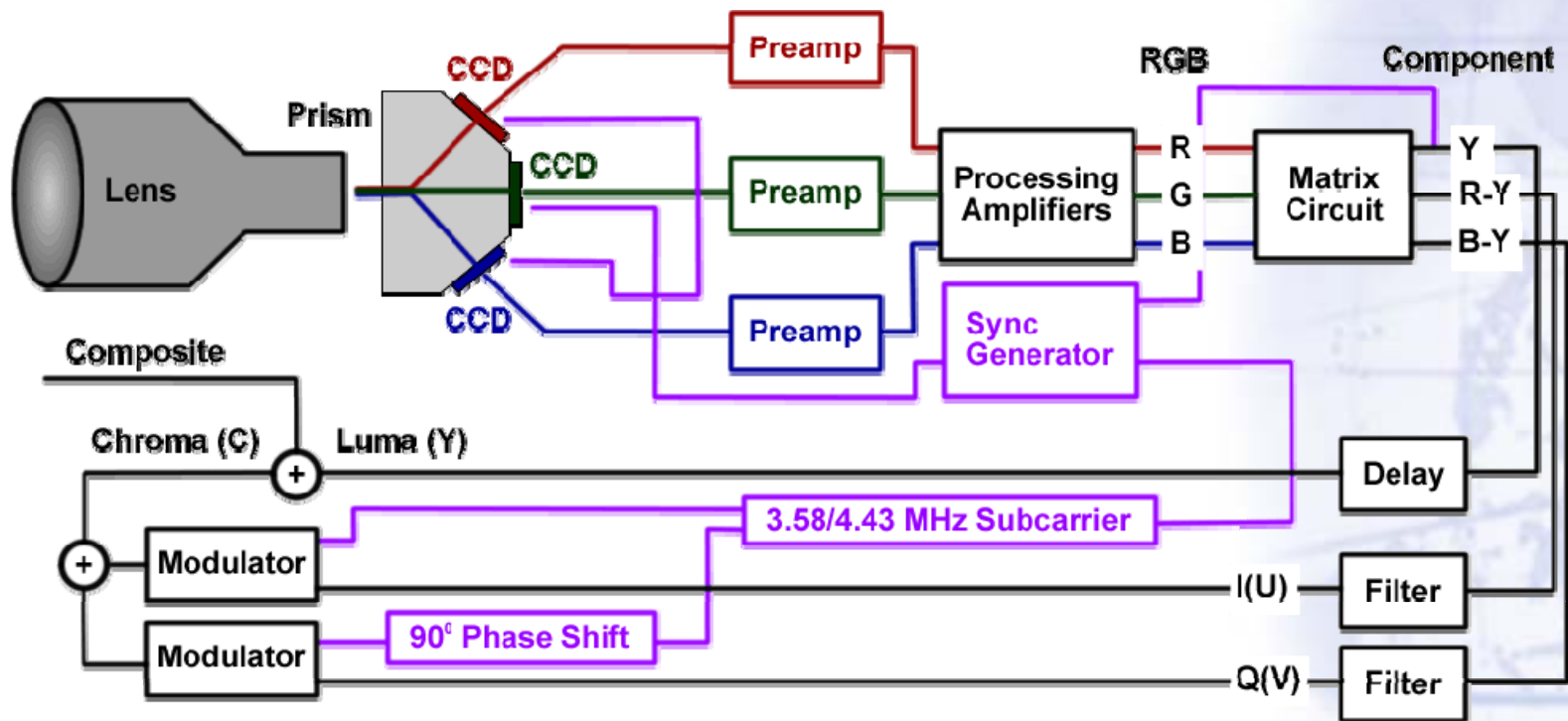
- AV Refresher
- Sampling
- Compressing
- Calculating  
bandwidth/file size
- Codecs



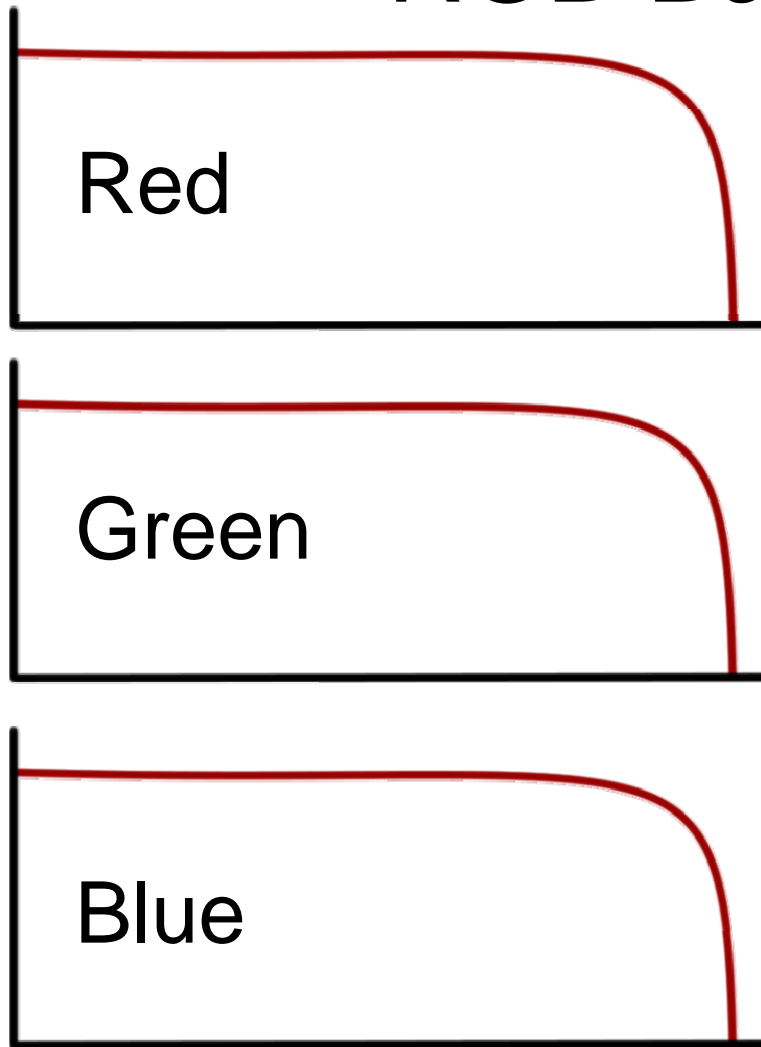
# AV Refresher

- Camera Block Diagram
- Video Signal Waveforms
  - RGB
  - Component
  - Composite
- Computer RGB Signals

# Basic Camera Block Diagram



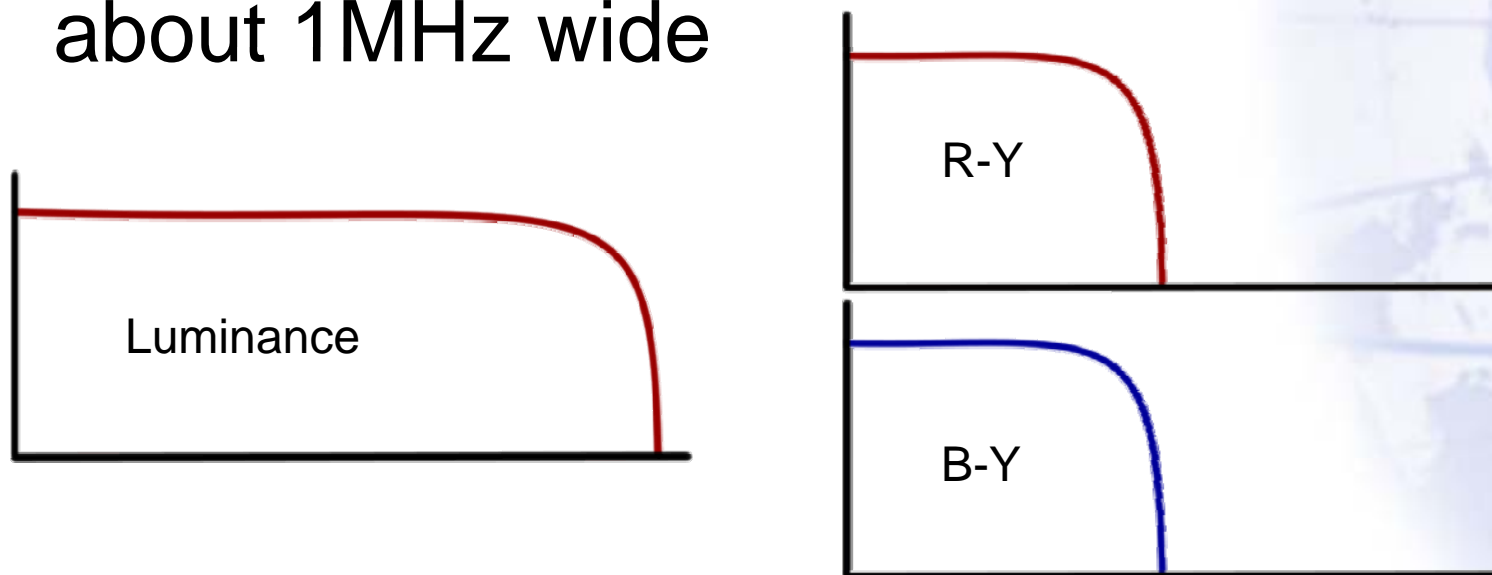
# RGB Bandwidths



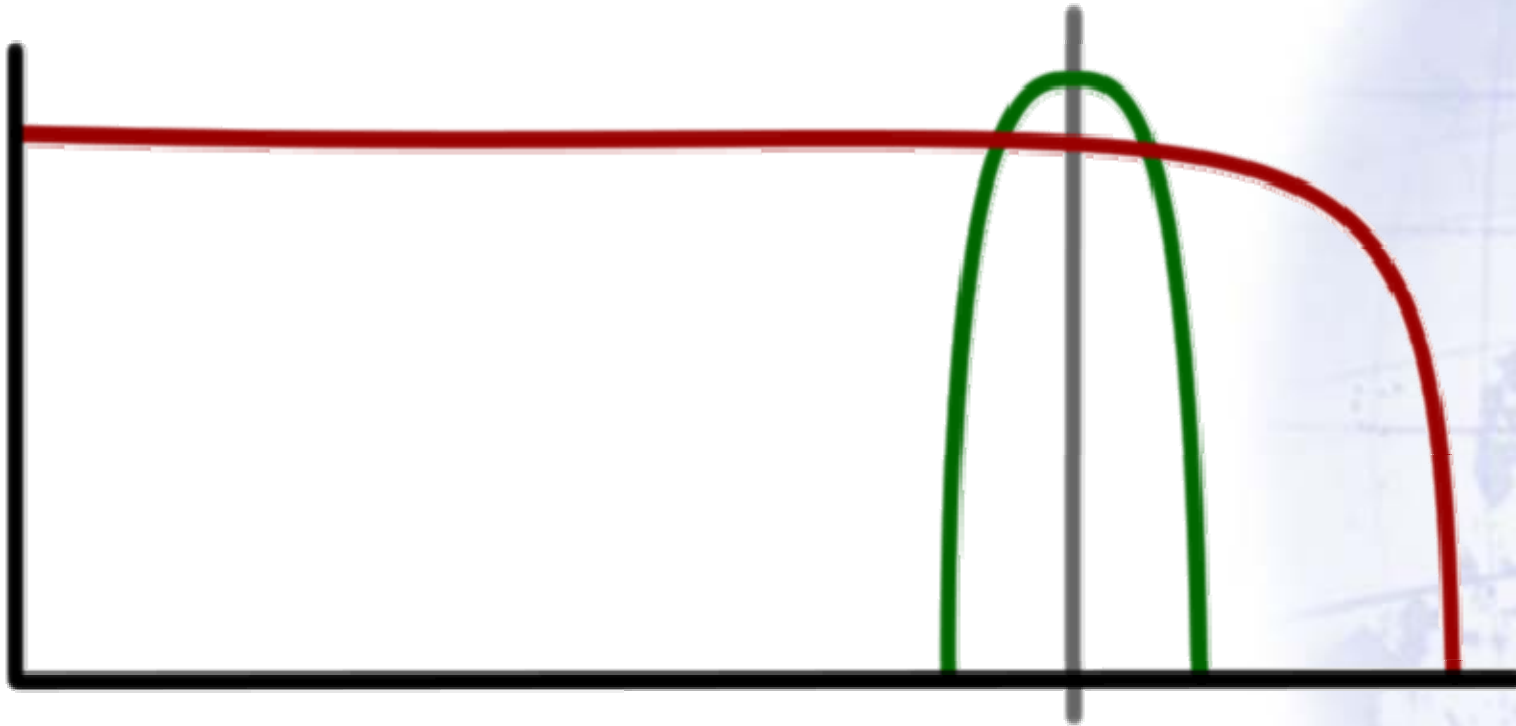
The RGB signals each have a equally wide bandwidths

# Component Bandwidths

- The luminance bandwidth remains wide to display detail, about 6MHz wide
- R-Y and B-Y have reduced bandwidth, about 1MHz wide

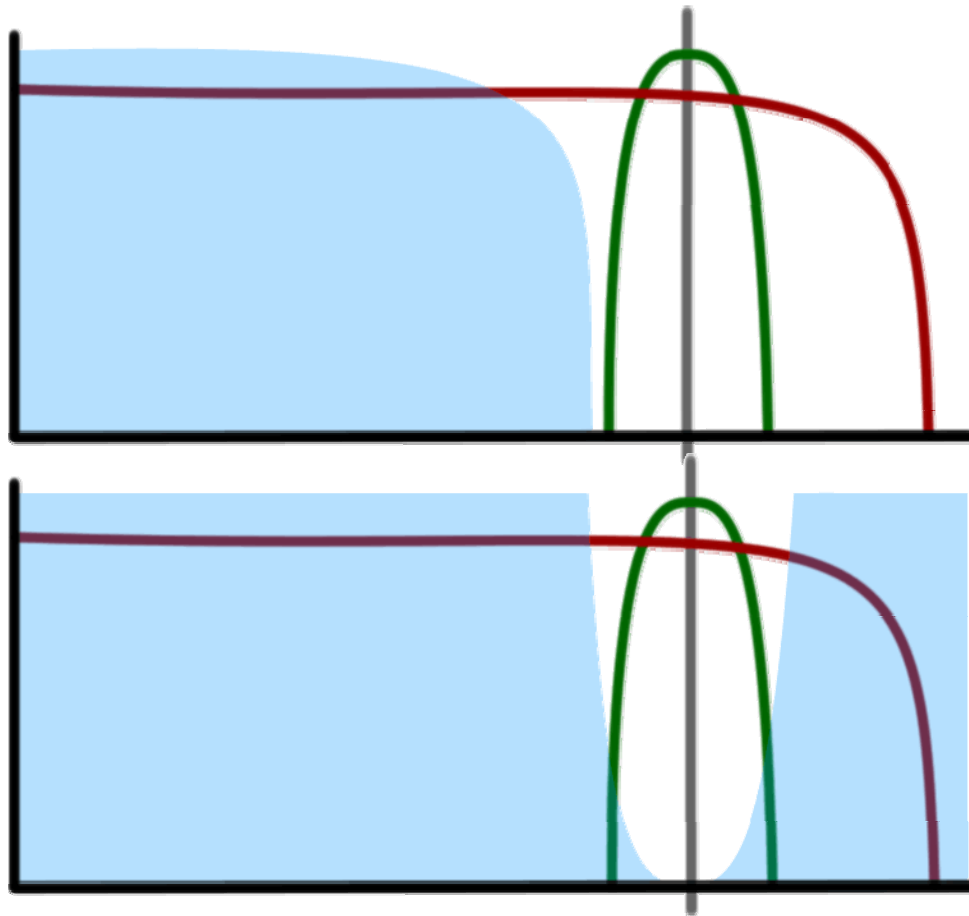


# Luminance Bandwidth



- Luminance is wide
- Chrominance is limited to about 1MHz

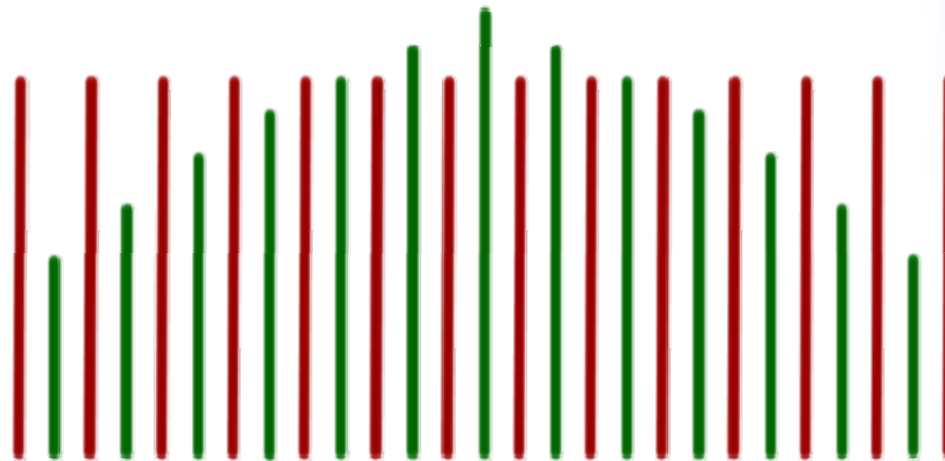
# Chroma Removal



- Low pass filter – losses high frequencies
- Notch filter – losses some mid-frequencies

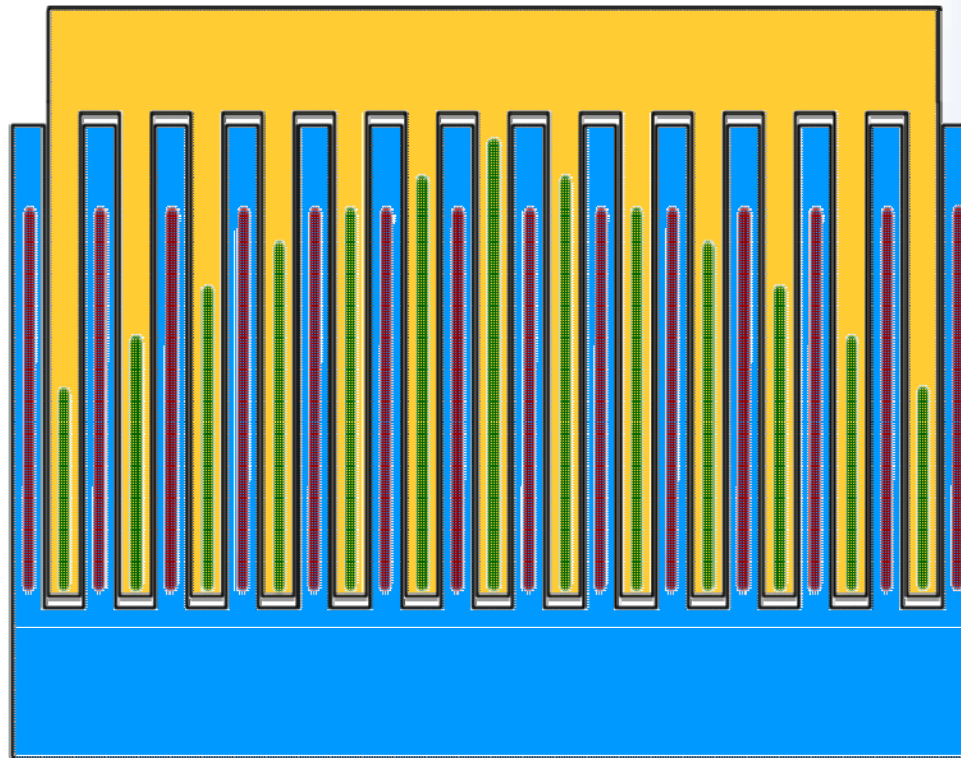
# Chroma Interleave

Green information shows how chroma signals are interleaved with luma signals shown in red.



# Comb Filter

The luma and chroma information are “combed” apart resulting in very good signal separation.



# Audiovisual Content on the Network

- Besides control signals, we wish to convert audio and video signals for network transport
- This requires conversion, and uses more bandwidth than control signals

# Digital Audio: AES/EBU

- Audio only format
- AES: Audio Engineering Society (AES3, 1992)
- EBU: European Broadcasting Union
- The most commonly used method for digitally interconnecting audio equipment worldwide using physically separate cables for each stereo audio connection
- Based on Pulse Code Modulation (PCM)
- Three types:
  - Balanced (3 conductors, twisted pair, XLR connector)
  - Unbalanced (Coaxial cable, BNC or 8-pin modular [RCA] connector): precursor was the Sony/Philips Digital Interface (S/PDIF)
  - F05 connectors, 5mm connectors for plastic optical fiber (Toshiba TOSLINK)

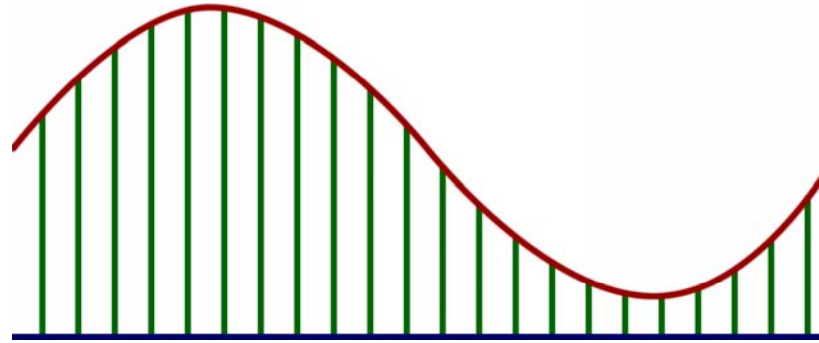
# Digital Audio: MP3

- A group of audio and video coding standards agreed upon by MPEG (Moving Picture Experts Group)
- MPEG-1 Audio Layer 3
- Variable compression ratios possible, e.g., 128 Kbps

# Digital Audiovisual: MPEG

- A group of audio and video coding standards agreed upon by MPEG (Moving Picture Experts Group)
- Remove as much as 90% of original data
- MPEG-2 was re-designed to handle HDTV signals in the range of 20 to 40 Mbits

# Audio Sampling



- Sampling Frequency Per Second  
X
- Quantizing Level or Bit Depth Per Sample  
X
- Channels To Be Encoded  
=  
Data Rate Per Second in bits

# Sampling Rates

- Telephone 8 kHz
- Compact disc (CD) 44.1 kHz
- DVD (audio portion) 48 kHz
- HD DVD 96 or 192 kHz
- Super Audio CD (SACD) 2.8224 MHz
  
- D1 Video 13.5 MHz

## Signal Acquisition/Digitization: AVI, MPEG-#, etc.

- Conversion from analog to digital (A/D)
- Conversion from digital to analog (D/A)
- All-digital acquisition (digital camera)
- All-digital transmission and display (LCD, DLP display devices)

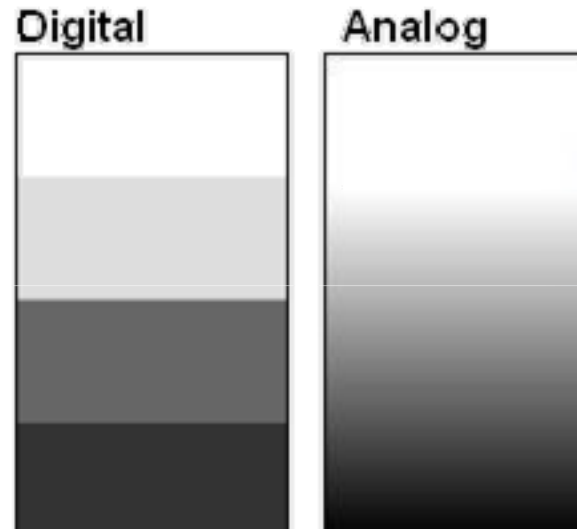
# High Quality Streaming Video

- Compression Schemes are used:
  - MPEG-1: “System Stream” over UDP
  - MPEG-2: “Transport Stream” over UDP
  - ~~MPEG-3: No longer used~~
  - MPEG-4: Independent Audio and Video Streams over UDP. Most adjustable format.
- It is necessary to configure the router/firewall ports to allow transport over UDP, otherwise http port will be used (tunneling)

# Digital vs. Analog Signals

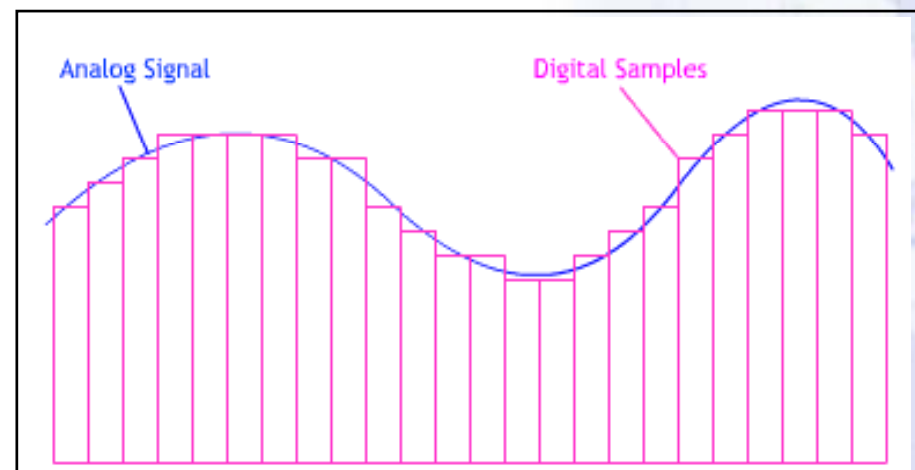
- **Digital**

- On or off
- Represented as ones and zeros
- Cleaner transmission
- More reliable storage
- Exact replication

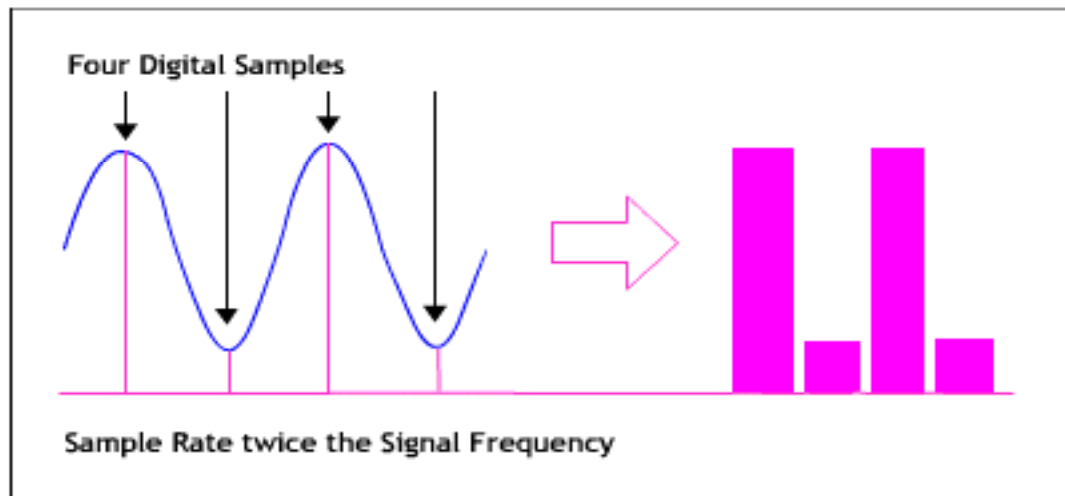
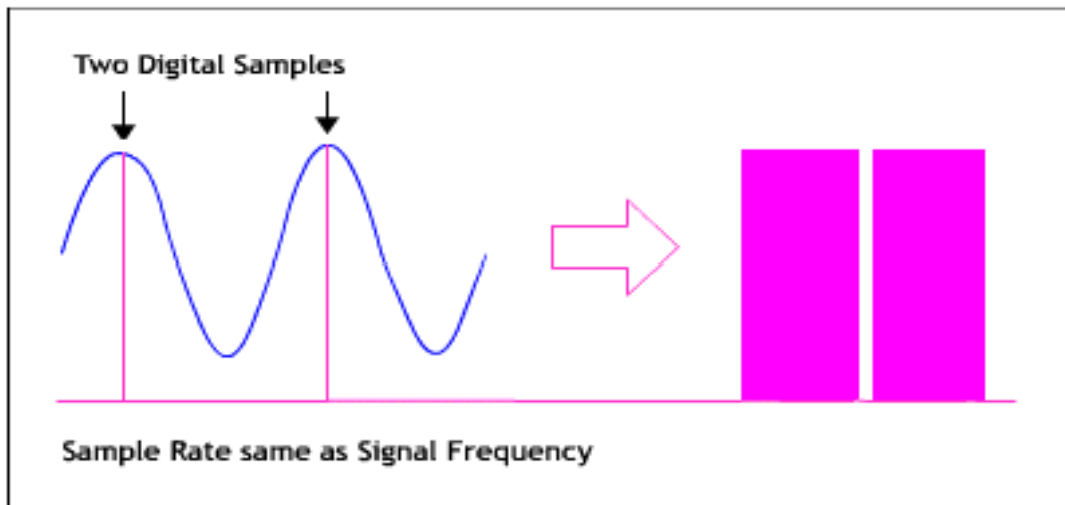


- **Analog**

- Infinitely variable
- Usually represented as a wave
- Represents reality well because it accommodates all gradations



# Sampling Rate of a Digital Signal



## A – D Conversion

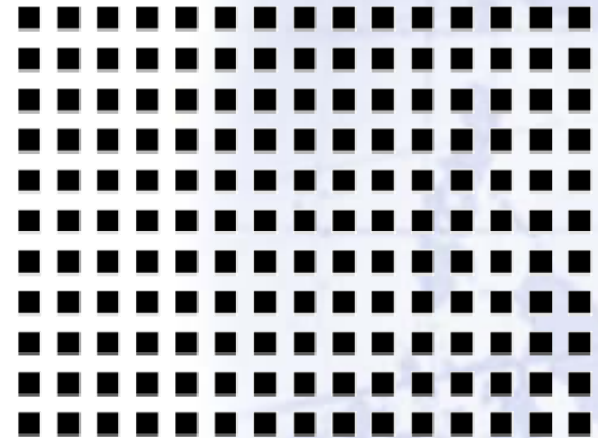
**Nyquist Criterion**  
Sampling rate is *at least* twice the highest frequency of the signal being sampled

# Bit Depth of a Digital Signal

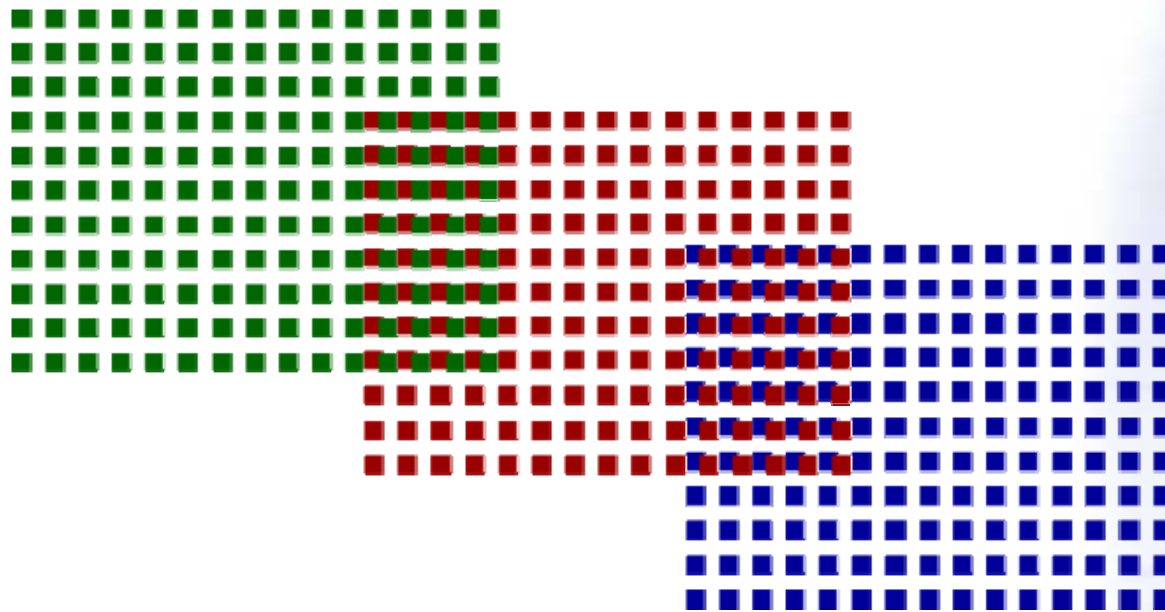
- Bit depth - the number of bits you have in which to describe the value
- Determines the digital signal's precision
- 1 bit = 2 states: 1,0
- 2 bit = 4 states: 00,01,10,11
- 3 bit = 8 states:  
000,001,010,100,011,110,101,111

# Pixel Count

- Horizontal Pixels  
X
- Vertical Pixels  
X
- Quantizing Level or Bit Depth  
X
- Frames Per Second  
=
- Bits Per Second

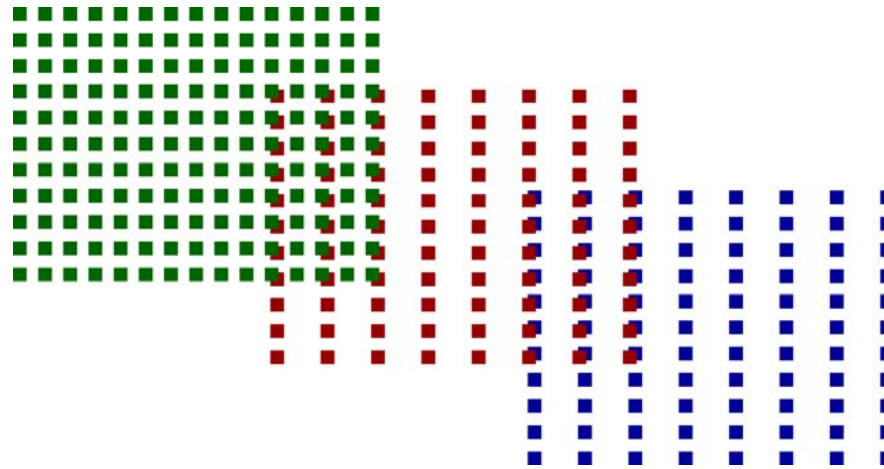


# 4:4:4 Sampling



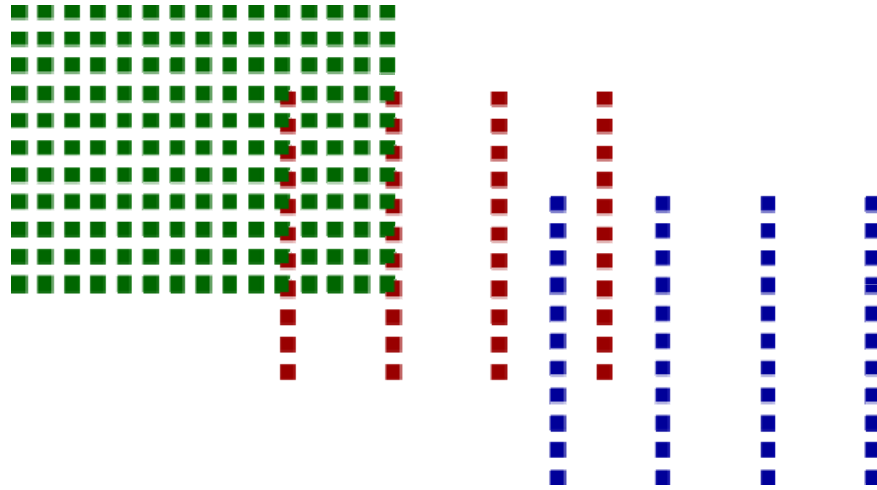
- Equal Quality of Red, Green, and Blue
- What's the data rate per second?

# 4:2:2 Sampling



- Our eyes can't see color detail as well as B&W detail; we limit the color information
- Betacam and DVD formats use 4:2:2 sampling
- There are half as many samples used for color

## 4:1:1 Sampling



- We limit the color information even more
- DV and some MPEG formats use 4:1:1 sampling
- There are 1/4 as many samples used for color

# 4:2:0 Sampling

- Mathematically the same as 4:1:1
- 4:1:1 Samples four sub-samples in a row
- 4:2:0 Samples two sub-samples in a row and two-samples in the row below

**4:1:1**



**4:2:0**



## SDTV and HDTV

<b>SDTV 4:3 Aspect Ratio</b>	<b>HDTV 16:9 Aspect Ratio</b>
480p – 480 progressive scan lines	720p – 720 progressive scan lines: 1280 X 720
480i – 480 interlaced scan lines	1080i – 1080 interlaced scan lines 1920 X 1080

Computer Video: 4:3 Aspect Ratio 1024 X 768

16:9 Aspect Ratio 1365(6) X 768

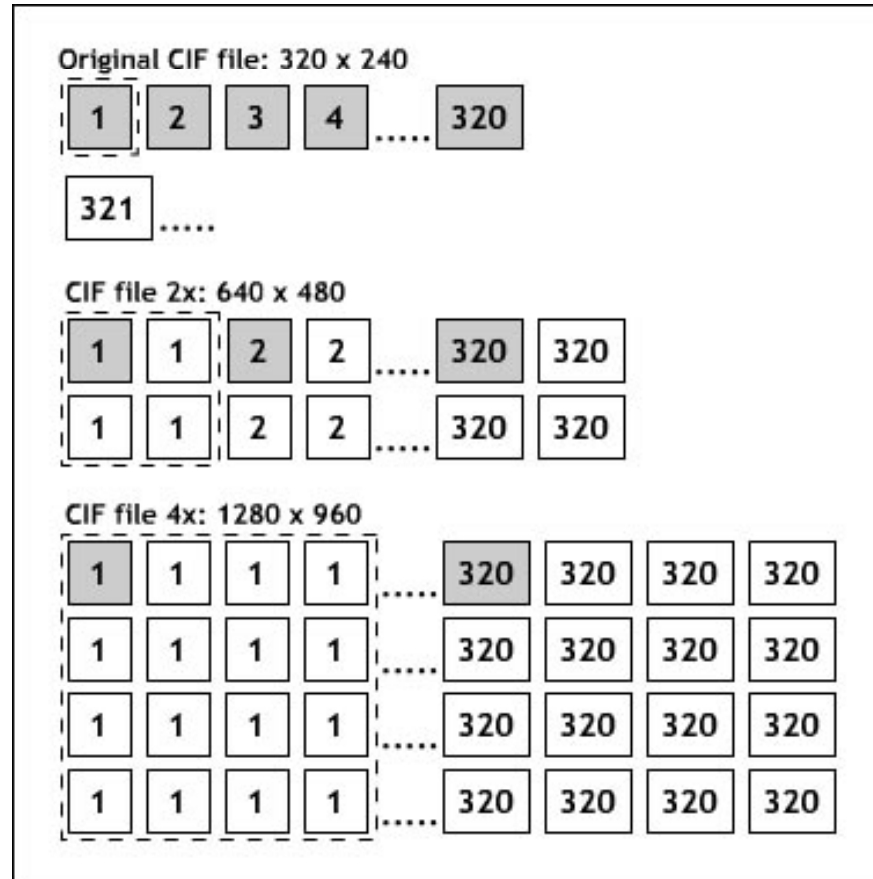
# Picture Size

- **CIF - Common Intermediate Format** – *a set of standard video formats used in videoconferencing, defined by their resolution*
  - The original CIF is also known as Full CIF (FCIF)
  - QCIF - Quarter CIF (resolution 176x144)
  - SQCIF - Sub quarter CIF (resolution 128x96)
  - 4CIF - 4 x CIF (resolution 704x576)
  - 16CIF - 16 x CIF (resolution 1408x1152)

CIF Video resolutions  
(in pixels)

Format	NTSC-based	PAL-based
SQCIF		128 × 96
QCIF	176 × 120	176 × 144
QCIF+	176 × 220	176 × 220
CIF	352 × 240	352 × 288
2CIF	704 × 240	704 × 288
4CIF	704 × 480	704 × 576
9CIF	1056 × 720	1056 × 864
16CIF	1408 × 960	1408 × 1152

# Video Scaling Process



# Transmitting The Program

- Total Program Size  
Video Bit Rate + Audio Bit Rate + Over Head
- Overhead Are the Extra Bits of Information That Add Error Control and Routing Information to the Signal

## Bandwidth and Speed: Matching Content to Network

- Determining usable network bandwidth
- Only about 70% of the “rated” network capacity is really available
  - Base this on the “bottleneck” point speed
  - Allows for collision avoidance, even in a switched network
- About 30% of this “available” network is to be used for streaming media.
  - Otherwise other traffic has to wait
  - QoS (Quality of Service) is required
  - Bandwidth management (throttling)

## Bandwidth and Speed: Matching Content to Network

- Example: Calculate Required Bandwidth
- 100 Mbit network
- 70% Available
  - Available network is 70 Mbits
- 30% Usable for Streaming
  - Streaming capacity 21 Mbits per second

# Compression

- Reducing Data Size Of Video and Audio
- Lossless
  - Compresses with a mathematically exact recovery of the original file
  - PKZIP is lossless
- Loss-y
  - Compresses and discards “unneeded” data yet returns an acceptable quality
  - MPEG, JPEG, and wavelet compression



# Improving Compressed Images

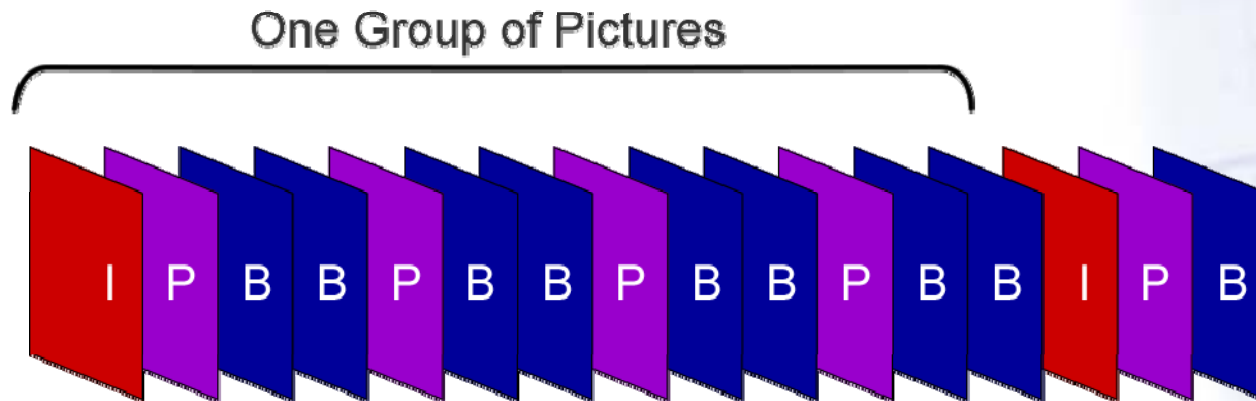
- Start With a Frame Size that Matches Your Output Size
- Reduce Noise in the Image
- Select the Proper Audio Quality
- Select the Proper Codec

# Digital Audio - Video: Compression

- Compression can be “Intra-Frame”
  - Within a frame by itself
  - Motion-JPEG algorithm
- Compression can be “Inter-Frame”
  - Use information of how consecutive frames change: the differences between them
  - MPEG algorithms
  - Most videoconferencing codec’s

# Group of Pictures (GoP)

- A series of frames that are required to display a complete sequence
- I, P, B Frames



# Signal Acquisition: Codecs

- COder/DECoder
- May be one-way or two-way
- Hardware codecs
- Software codecs
- Specialized codecs for videoconferencing
- Sometimes compressed, sometimes not

# MPEG-2

- International standard ISO/IEC 3818
- Most common digital audio/video format
- DVD Players, digital cable, satellite TV
- Allows text and other data (e.g., a program guide for TV viewers) to be added to the video and audio data streams
- Many variables are possible to achieve different quality levels/file sizes

# MPEG-4

- International standard ISO/IEC 14496
- Separates audio, video, data, subtitles, compresses them separately
- Newer, potentially higher quality format than MPEG-2
- Lower data rates, smaller files sizes than MPEG-2
- Better interoperability between devices

# Levels and Profiles

- Profiles are quality groupings within a compression scheme
- Levels are specific image sizes and frame rates of a profile
- Provide interoperability or standards within a compression format

# SMPTE 421M or VC-1 Encoding Format

<b>Profile</b>	<b>Level</b>	<b>Maximum Bit Rate</b>	<b>Representative Resolutions by Frame Rate (Format)</b>
Simple	Low	96 kilobits per second (Kbps)	176 x 144 @ 15 Hz (QCIF)
	Medium	384 Kbps	240 x 176 @ 30 Hz 352 x 288 @ 15 Hz (CIF)
Main	Low	2 megabits per second (Mbps)	320 x 240 @ 24 Hz (QVGA)
	Medium	10 Mbps	720 x 480 @ 30 Hz (480p) 720 x 576 @ 25 Hz (576p)
	High	20 Mbps	1920 x 1080 @ 30 Hz (1080p)*

# VC-1 Advanced Profile

Profile	Level	Maximum Bit Rate	Representative Resolutions by Frame Rate (Format)
Advanced	L0	2 Mbps	352 x 288 @ 30 Hz (CIF)
	L1	10 Mbps	720 x 480 @ 30 Hz (NTSC-SD) 720 x 576 @ 25 Hz (PAL-SD)
	L2	20 Mbps	720 x 480 @ 60 Hz (480p) 1280 x 720 @ 30 Hz (720p)
	L3	45 Mbps	1920 x 1080 @ 24 Hz (1080p)* 1920 x 1080 @ 30 Hz (1080i) 1280 x 720 @ 60 Hz (720p)
	L4	135 Mbps	1920 x 1080 @ 60 Hz (1080p)* 2048 x 1536 @ 24 Hz

# Digital Audiovisual: AVI

- Microsoft Standard Audio/Video File Format
- Audio Video Interleave (AVI)
- Packaging Format: relies on different codec's according to application
- Video Format Variables:
  - Size, or resolution (e.g., 320 X 240)
  - Data per pixel (e.g., 16 Bits)
  - Total number of frames (e.g., 50)
  - Number of frames per second (e.g., 10 fps)
  - Data rate (e.g., 1521 KB/sec)
  - Video compression format (e.g., “uncompressed”)

# Digital Audiovisual: Windows<sup>®</sup> Media

- Codec Technology from Microsoft<sup>®</sup>
- No License Required as in MPEG-2 and MPEG-4
- Also a SMPTE Standard: VC1
  - (Society of Motion Picture and Television Engineers)
- Windows<sup>®</sup> Media Player: ubiquitous file format (.wmv)
- Multicast Capable
- Multi-vendor Support
- Digital Rights Capable
- Challenge: Delay

# Digital Audiovisual: Apple® QuickTime®



**QuickTime 7** For Windows

**H.264** Advanced video with pristine quality

- Amazingly responsive live player resize
- Up to 24 channels of surround sound
- All new playback controls

The advertisement features a central video player window showing a scene from the movie 'Ratatouille' with the Eiffel Tower in the background and the character Remy in the foreground. The player has a progress bar and playback controls at the bottom.

From the Apple website: [www.apple.com](http://www.apple.com)

# Digital Audio Visual: Real Media

- .rm file extension
- Offers compression codec
- Offers streaming products
- Offers stream hosting



# Digital Audio Visual: DIVX

- Codec Engine
- Playback on specific hardware devices
- VoD applications
- Mobile device applications



# Flash Encoding

- Adobe format mostly used on web applications
- Good scripting support
- Generally highly compressed format
- Plays quickly – low latency
- Encoder available for Quicktime
- FLV format

# Codec Determination

- [www.fourcc.org](http://www.fourcc.org)
- Information on common codecs
- Applications to discover codec versions
- Codec download links

# Exercise: File Comparisons

- Use software to analyze audio video files for size vs. quality differences



# Compression and Transcoding

- It is possible to convert a file from one type to another
- Sometimes this introduces artifacts, distortions, etc.
- Usually requires software, but may also require a hardware device

# Streaming

- Software
- Hardware
- Unicast
- Multicast



# Delay/Latency

- Internet is “best-effort” delivery
- Problems:
  - Packet loss, delay, reordering, jitter
- Desktop players (WMP, RealOne, Quicktime) experience delay when playing Internet-based media
- Varying degrees of delay are acceptable:
  - Videoconferencing: 200 msec
  - High fidelity audio: approx 50  $\mu$ sec
  - Desktop video (streaming): 1 sec!

# Intelligent Streaming

- .wmv files may be created with multiple resolutions
- Intelligent streaming feature of servers and players determines optimal transmission resolution for file
- This is transparent to the user, but must be defined by the file creator
- Streaming server is required

# Software/PC Encoding

- Easy to update codecs and authoring programs
- Processor must divide its time between encoding and other PC duties/programs
- IT must be concerned about security and virus attacks
- Higher hardware requirements than typical office PC

Questions

