Multimedia Applications

Internet Technologies and Applications

Aims and Contents

- Aims
  - Define multimedia applications
  - Introduce technologies for delivering multimedia applications in the Internet
  - Raise the issues in delivering multimedia applications

- Contents
  - Application characteristics and requirements
    - Voice, audio, video; Performance requirements
  - Voice over IP
  - Streaming Stored Audio/Video
  - IPTV

  - (and Multimedia and QoS)

Application Characteristics and Requirements

Types of Multimedia Applications

- One-way (unidirectional) communications
  - Listening to radio/music; viewing recorded or live video
  - Also referred to as Streaming
    - Stored audio/video
    - Live audio/video
- Two-way (bidirectional) communications
  - Voice calls, video-conferencing
  - Also referred to as Interactive
- Sometimes multimedia applications are referred to as real-time applications
  - Real-time communications: sender and receiver communicate as if they were at the same physical location
    - Very small delay and/or jitter between sender and receiver

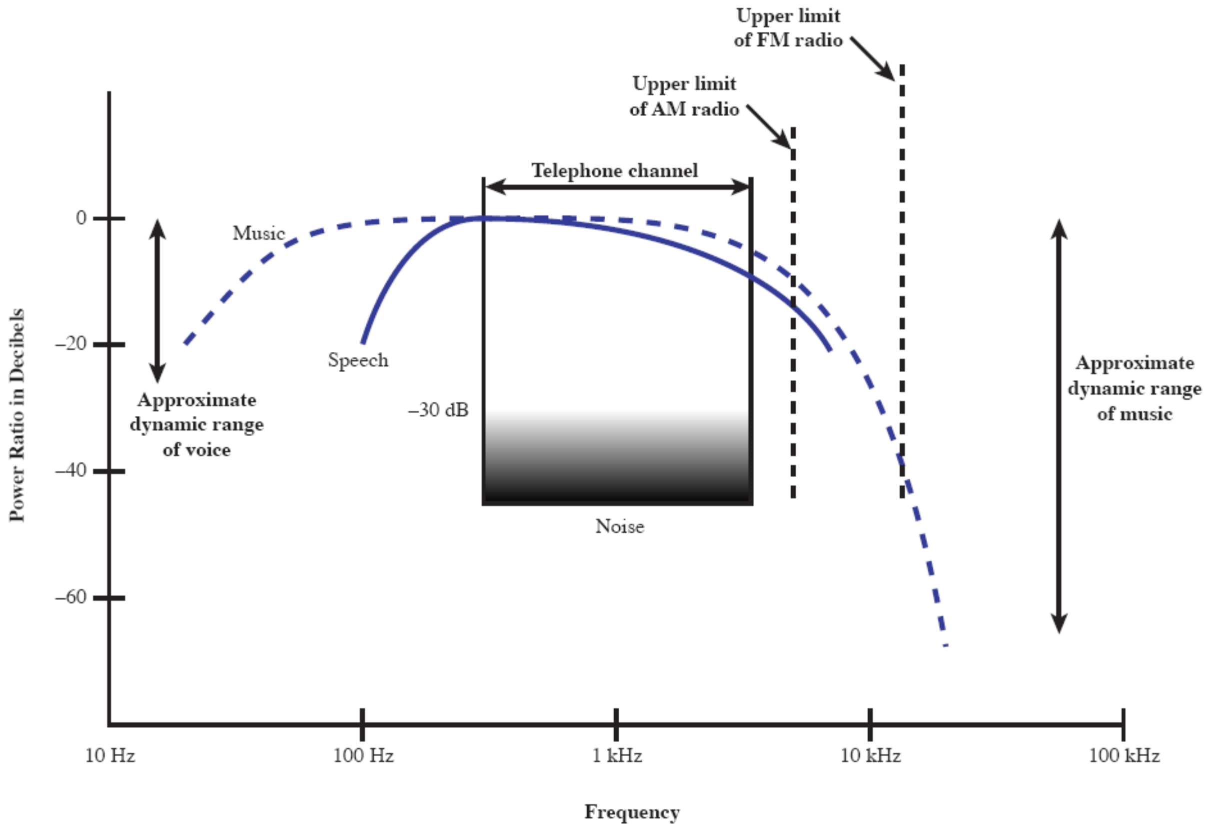
Multimedia vs Other Applications

- “Other” Applications
  - Web browsing, file download, email, database access, …
  - Reliability is essential
  - Large and/or varying delays can be tolerated

- Multimedia applications
  - Delay-sensitive: large delays or jitter make the application un-useable
  - Loss-tolerant: if some data is lost, the application is still usable

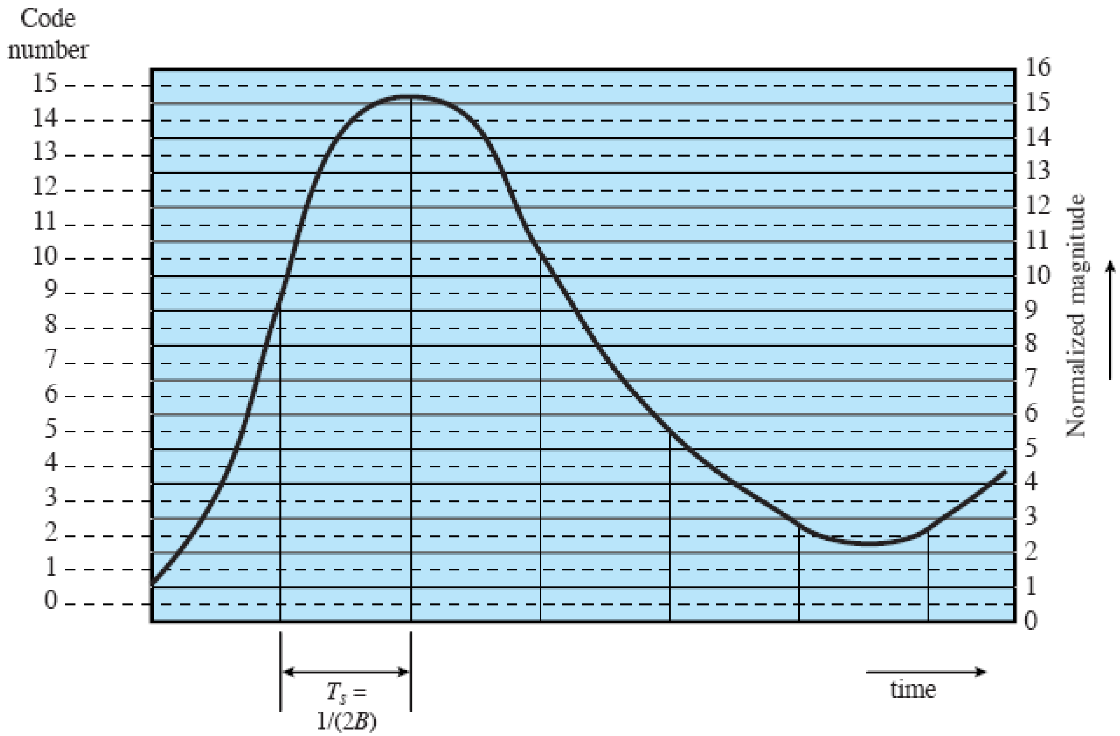
Voice Communications

- The human voice uses frequencies in the range of 100Hz to <10KHz
  - Majority of voice communications is 300HZ to 3400Hz



Voice Communications

- Analog voice data is converted to digital data using pulse modulation techniques, e.g.
  - Pulse Code Modulation (PCM), Delta Modulation



PAM value	1.1	9.2	15.2	10.8	5.6	2.8	2.7
quantized code number	1	9	15	10	5	2	2
PCM code	0001	1001	1111	1010	0101	0010	0010

Voice Communications

- How much information is needed in digital data to accurately reproduce analog data at receiver?
  - Sampling Rate: how often is the analog data sampled?
    - Units: Samples per second or Hertz (Hz)
    - Nyquist's Sampling Theorem tells us if we sample twice the highest frequency signal component then can make a perfect reproduction
      - If highest frequency component is B, then sampling rate should be 2B
  - Sample Size: how many different levels can a sample represent?
    - Units: bits
  - Low sample rates and small sample sizes can lead to poor voice reproduction at the receiver
  - High sample rates and large sample sizes require high transmission data rates

Voice Communications

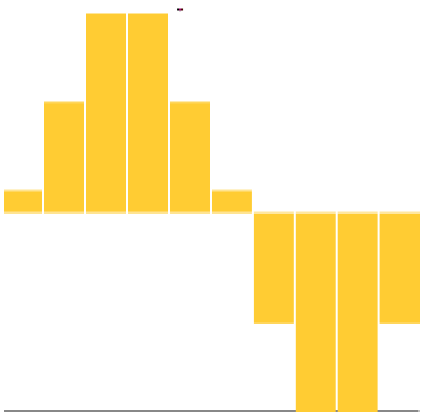
1HZ sine wave

50 samples per second
100 levels

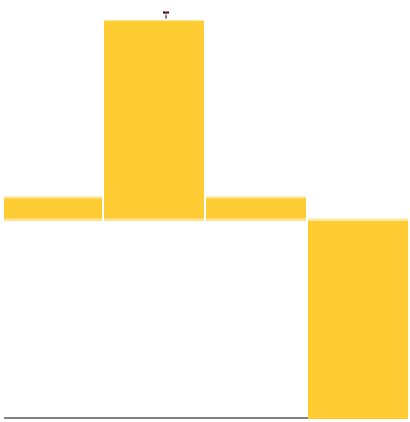
50 samples per second
10 levels

50 samples per second
4 levels

50 samples per second
10 levels



10 samples per second
10 levels



4 samples per second
10 levels

Audio Communications

- Voice communications is a type of audio communications
- Non-voice audio communications: e.g. music
- Same concepts apply as for voice communications
  - Music generally has a larger bandwidth than voice
  - Users often desire higher quality output

Video Communications

- Video: still images representing scenes in motion
  - Still images are called *frames*
  - Video is often accompanied by audio
- How much information is contained in video?
  - Frame size (or resolution):
    - How large (width x height) is each frame? Number of pixels
    - How much depth in the colour? Bits per pixel
  - Frame rate: how often are frames generated?
    - About 15 frames per second needed to make illusion of motion

Compression

- Raw data for audio/video:
  - PCM voice: 8kHz sampling rate, 8 bits per sample: 64kb/s
  - PCM audio (e.g. music): 44kHz, 16 bits per channel, 2 channels: 1.4Mb/s
  - Standard Definition Digital TV: 720x576 pixels, 24 bits per pixel, 25 frames per second: 248Mb/s
  - High Definition TV: 1920x1080 pixels: 1244Mb/s

- Compression is often used to reduce capacity needed for:
  - Storage
  - Transmission

Compression

- Lossy compression
  - Reduces the quality (lose information)
  - Most commonly used today; reduce amount of data from 5% to 25% of original size
- Lossless compression
  - No loss of information, hence quality is maintained
  - Compress to 50% to 70% of original size

- Amount of compression depends on algorithm and data source
  - Example: high motion video cannot be compressed as much as low motion video

Codecs

- A general term referring to the software/hardware/standard that perform modulation, compression, formatting of source into digital formats
- Trade-off in: bit rate, quality, complexity (processing time)

- Voice codecs
  - ITU G.711: 64kb/s PCM
  - ITU G.722: 64kb/s ADPCM
  - ITU G.726: 16-40kb/s ADPCM
  - ITU G.728: 16kb/s CELP
  - ITU G.729: 8kb/s ACELP
  - GSM: 14kb/s
- Audio codecs (lossy)
  - MPEG: MP3, AAC
  - Dolby Digital AC-3
  - DTS
  - Vorbis
  - WMA
- Audio codecs (lossless)
  - FLAC
  - Shorten
  - WMA Lossless
  - MPEG-4 Lossless
- Video codecs
  - MPEG-1: VCD (1.5Mb/s)
  - MPEG-2: DVD, digital TV (3-6Mb/s)
  - MPEG-4: DivX, Xvid, FFmpeg
  - H.264 (MPEG-4 AVC): Bluray
  - WMV
- (Container formats:
  - AVI, Ogg, WAV, MOV, MPEG4 TS/PS, …)

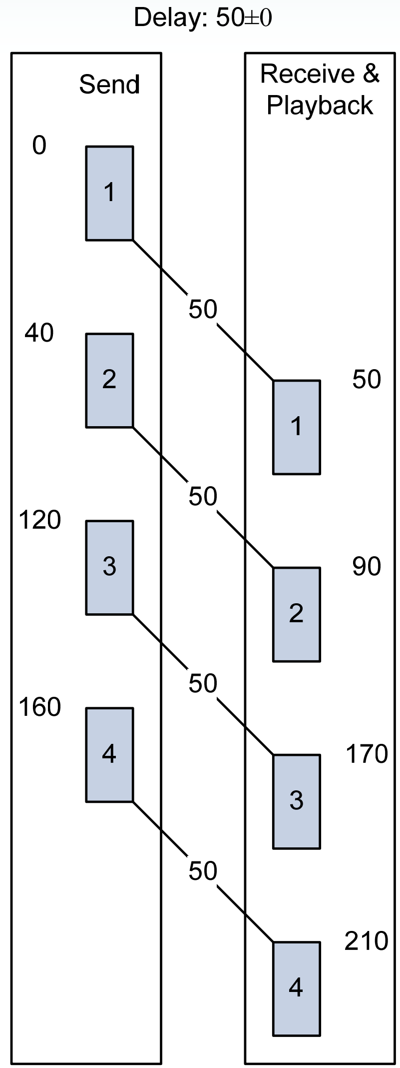
Performance Requirements for Applications

- Data Rates
  - Voice/audio applications require 10’s to 100’s kb/s
  - Video applications require 100’s kb/s to 10’s Mb/s
- Errors
  - Most applications can tolerate small number of errors (i.e. loss of data)
    - Can use Forward Error Correction (FEC) to reduce errors
    - Re-transmission schemes are avoided because of the extra delay they incur
  - Errors result in drop in quality at receiver
- Delay
  - Interactive (or conversational) applications have strict delay requirements
    - Voice call: <150ms is unnoticeable; 150-400ms is tolerable; >400ms is unusable
  - Streaming applications can tolerate large delays by using buffers
- Jitter
  - Most applications require low jitter for smooth playback of audio/video

Example: Delay, No Jitter

Source sends a packet every 40ms

The network has delay of 50ms. The jitter is 0 (every packet experiences a delay of 50ms)

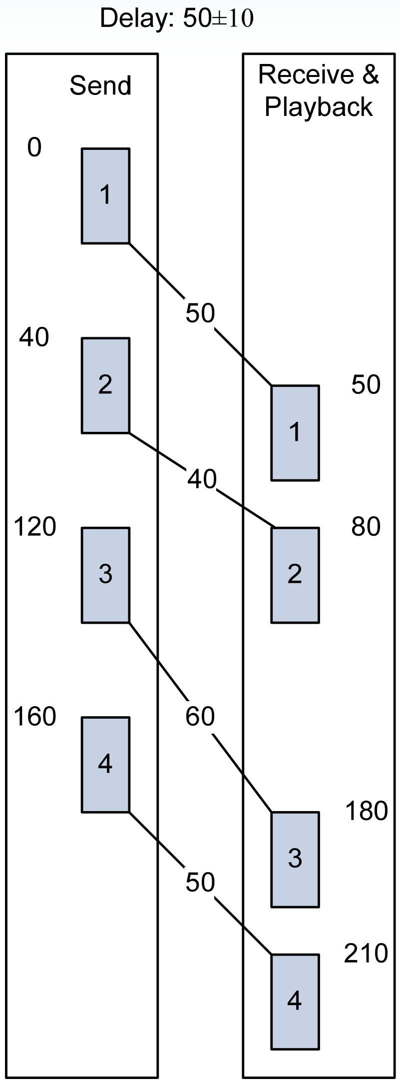


When a packet is received, the data is played back.

There is 50ms delay between when the source starts sending to when the playback starts at receiver.

With no jitter, the playback is smooth.

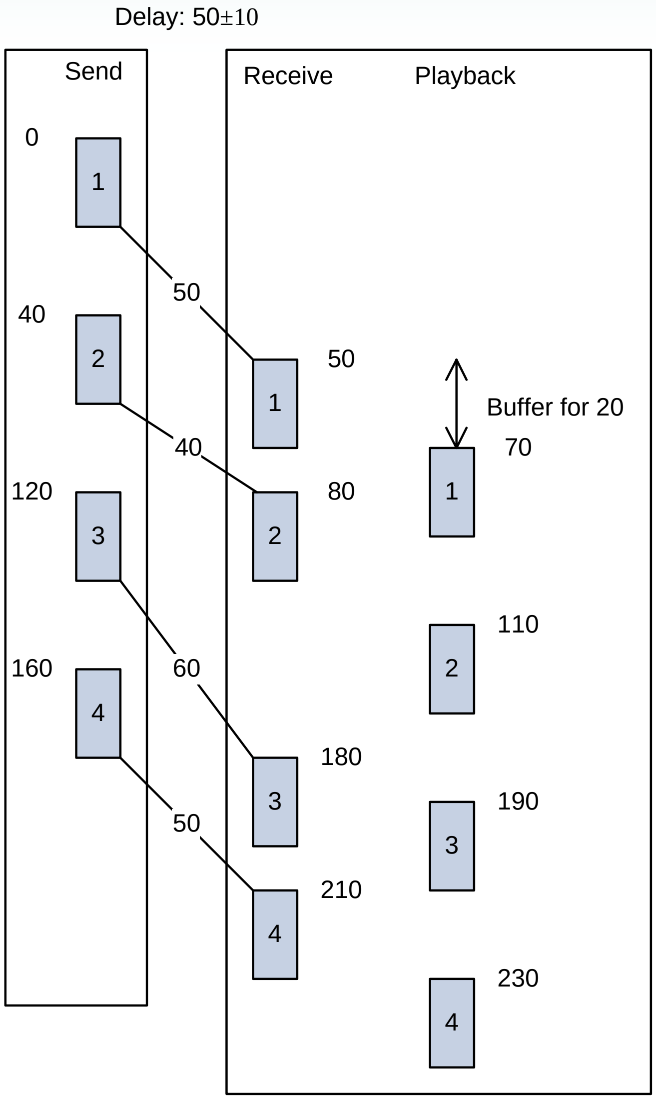
Example: Delay and Jitter



There is 50ms delay between when the source starts sending to when the playback starts at receiver.

With jitter, the playback is no longer smooth. There is a large delay between the second and third packet. E.g. the frame in packet 2 will be frozen until the packet 3 arrives

Example: Delay, Jitter and Buffering



The receiver stores the each packet in a buffer (not played immediately).

The first packet is buffered for 20ms, and then the frame is played back.

Packets are buffered when necessary and played back at a constant rate. Hence the playback is smooth.

Note there is now a 70ms delay between when the source starts sending to when the playback starts at receiver.

And the receiver needs extra complexity and memory for buffering.

Dealing with Jitter

- Playback buffers are the main mechanism to deal with jitter
- Packets must include sequence numbers and timestamps
- Receiver buffers packets until they are ready to be played
- How long to delay before starting playback?
  - For stored audio/video streaming, several seconds is ok
  - For live audio/video streaming, 1-3 seconds may be tolerable
  - For interactive applications, must be less than tolerated delay (100's of milliseconds)

Multimedia Applications in the Internet

The Internet offers Best-Effort Service

- IP:
  - Unreliable, best-effort delivery of datagrams
  - No guarantee of delivery, no timing guarantees
  - No priority for different applications
  - Datagrams from an application may be processed in different ways by routers, and even take different routes
- TCP:
  - Provides reliability using a retransmission scheme
    - Adds considerable extra delay if errors occur
  - At start of TCP connection, throughput is low (to avoid congestion)

- The Internet, and TCP/IP, do not have built-in mechanisms to support multimedia applications
- Yet, with additional supporting mechanisms, multimedia applications in the Internet are successful

Voice over IP Networks

Terminology

- Remember, *the Internet* refers to a specific IP network
- Voice over the Internet Protocol, VoIP, IP Telephony, IP phone
  - Technology using IP (and other protocols) to make voice calls
- Internet telephony, Internet phone, Voice over the Internet
  - Using the Internet for voice calls (using VoIP technology)
- Example:
  - Skype is VoIP software; it allows voice calls over the Internet
  - NTT (Japan) over a VoIP service to customers on the NTT network
  - Thammasat Uni may deploy their own VoIP network
  - In the case of NTT and Thammasat, they may use their private IP network, separate from their network attached to the Internet

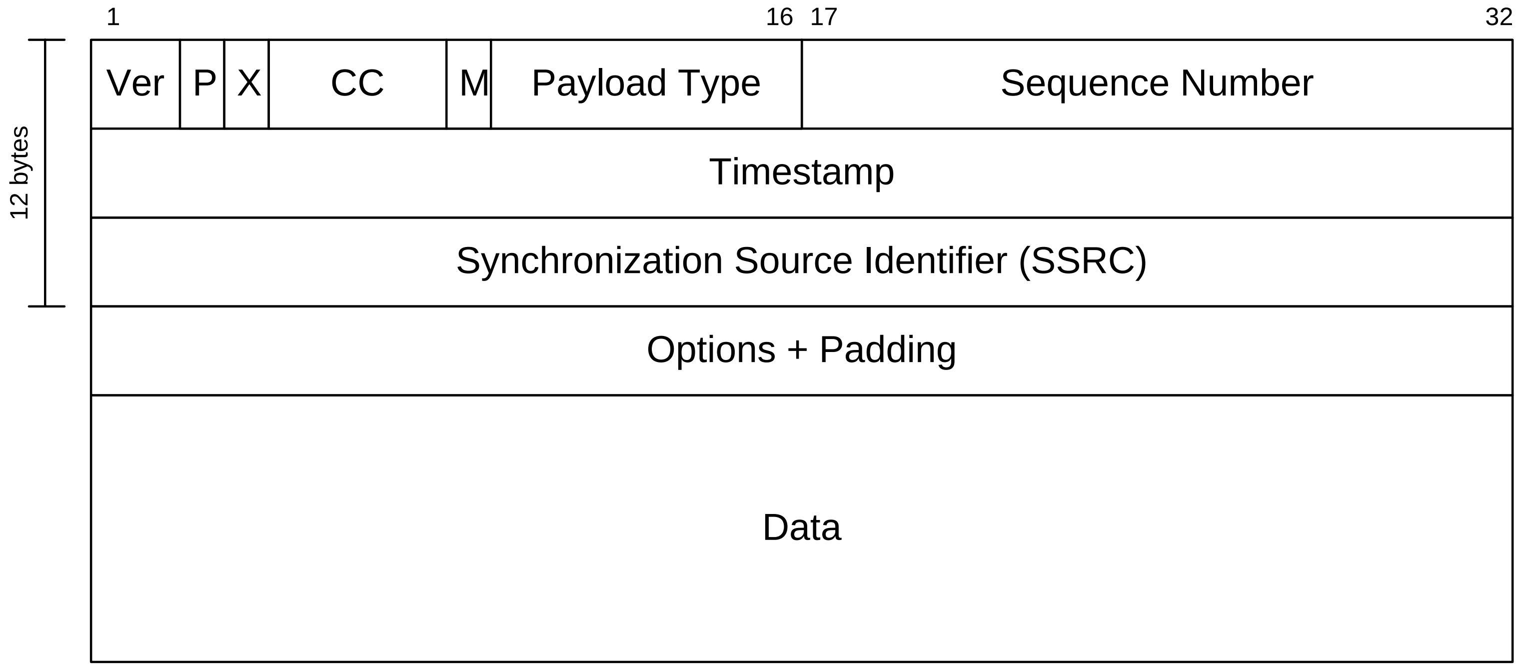
Real-time Transport Protocol

- RTP: Transmit digitized audio/video signals over an IP network
- Uses UDP
  - Consider RTP as a transport protocol
  - Since TCP is not well suited to transfer of multimedia communications, RTP was designed
- Main functionality of RTP
  - Allow any type of media (voice, video using any codec) to be transferred
  - Adds a sequence number to each block of media
  - Adds a timestamp to each block of media

- RTP does not provide any guarantees of reliability, timeliness or priority mechanisms

RTP Packet Format

- Minimum size of header is 12 bytes



RTP Packet Format

- Sequence number is used for each packet (initial value chosen randomly by application)
- Optional fields may be included; if so, the X bit is set to 1
- P bit is set if no padding is needed after payload
- M bit is used by applications to indicate if markers are included
- Payload type indicates the format of the data (payload), e.g.
  - 0: PCM, 8kHz, 64kb/s
  - 3: GSM, 8kHz, 13kb/s
  - 14: MPEG audio, 90kHz
  - 26: Motion JPEG
  - 33: MPEG2 video
- Timestamp indicates time when data was sampled at source
- SSRC is a unique ID for the source
- CC field indicates the number of sources contributing to the stream
- Optional fields include a Contributing Source ID

RTP Translation and Mixing

- An application may change the payload type in the middle of a session
  - E.g. change encoding to achieve better quality or lower data rate
  - The Payload Type field makes such *translation* possible
  - Translation may be performed by intermediate devices
- Multiple sources may contribute to a session
  - E.g. a tele-conference between group of people at one location to a group at another location
    - Each person at location A is a source; their stream of data may go to a central mixer, which combines them together into a single stream to be sent to the other location
  - SSRC for each person would be different
    - The mixer uses a new SSRC, but includes the original SSRC's in the optional Contributing Source ID fields and sets CC accordingly
  - Combined with multicast, mixing can lead to significant performance improvements

RTP Control Protocol (RTCP)

- RTP is for sending audio/video data streams
- RTCP is used for exchanging information between senders/receivers about the streams and users
- There are 5 types of RTCP messages:
  - Sender report: sender periodically sends a report to receivers; includes at least an absolute timestamp so receivers can synchronise different streams
  - Receiver report: receiver periodically sends a report to senders; indicate the conditions of the reception (e.g. congestion, buffer size) allowing senders to adapt their sending rates/quality
  - Source description message: sender may send information describing the owner of the stream
  - Bye message: sender sends this when ending the stream
  - Application specific message: applications may use this for their own purpose, e.g. closed captions or subtitles for a video stream

IP Telephony and Signalling

- With voice communications, *signalling* refers to the process of establishing a telephone call
- In the PSTN, signalling is performed using a protocol called Signalling System 7 (SS7)
  - Given a destination phone number, forms a circuit between source and destination
  - Handles call forwarding, error reporting, busy signals, …
- In IP telephony, an equivalent protocol is needed
  - Must be able to translate between PSTN and IP network using a gateway device
  - Two sets of standards proposed for IP telephony signalling:
    - Session Initiation Protocol (SIP) by IETF
    - H.323 by ITU

Session Initiation Protocol

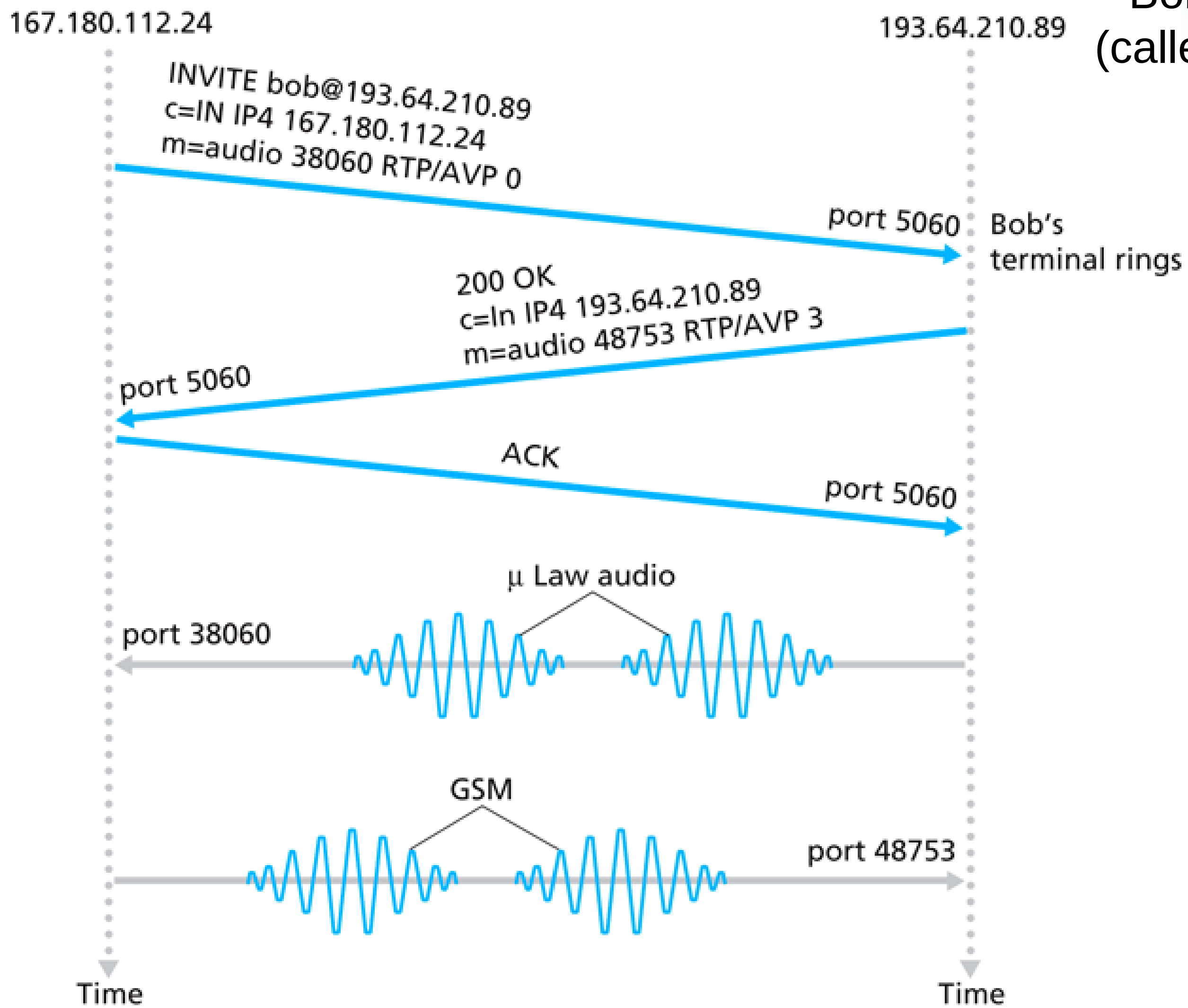
- SIP provides following mechanisms for an IP network:
  - *Caller* notifies a *callee* that it wants to start a call; allows participants to agree upon codecs and end calls
  - Caller determines the IP address of the callee
  - Changing codecs during a call; inviting new participants to a call; call transfer; call holding; …

- SIP is an application level protocol
  - Uses UDP (or TCP in special cases)
  - SIP uses port 5060
  - Sends text-based messages in a format similar to HTTP
  - Uses addresses similar to email address, e.g. sip:steve@siit.tu.ac.th
  - Does not specify the data transfer mechanism (RTP or others can be used)

SIP Example

Alice
(caller)

Bob
(callee)

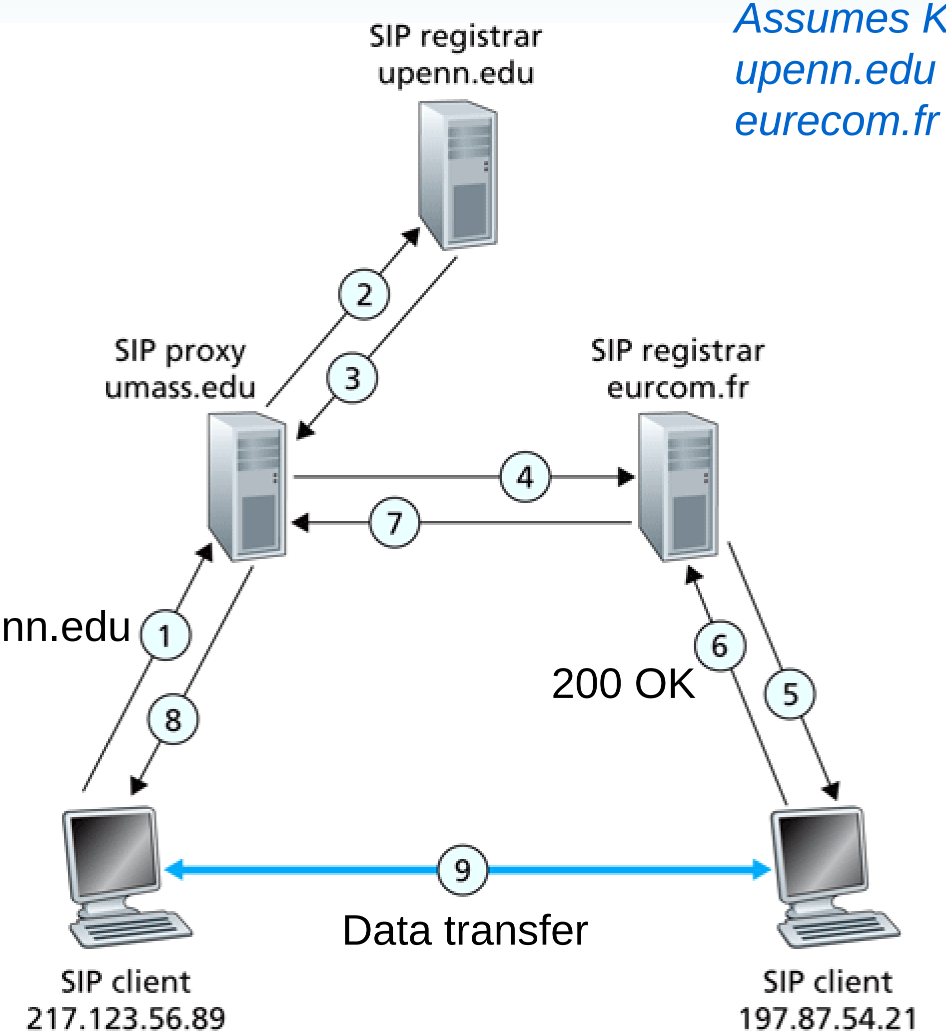


SIP Names and IP Addresses

- In IP networks, DNS maps domain names to IP addresses
- DNS works because servers are normally associated with a single fixed IP address
- But users are often associated with multiple, dynamic IP addresses
  - Static IP for PC at work; dynamic IP for PDA; dynamic IP for PC at home
- SIP uses:
  - Registrar Servers to keep track of a users current IP address
    - Each user has an associated Registrar
    - When a user starts their SIP client, the client informs the Registrar Server of the IP address
  - Proxy Servers to handle SIP INVITE’s on behalf of users
    - A caller sends an INVITE to a Proxy. The Proxy may:
      - Find the IP address of the callee via the Registrar Server, and initiate the call
      - Redirect the caller to another location (e.g. voicemail or website)

SIP Example

Assumes Keith has informed upenn.edu that he is at eurecom.fr



SIP for Voice, Video and Data

- SIP is a general protocol for initiating and managing sessions
  - Voice calls
  - Video calls
  - Data sessions, especially instant messaging

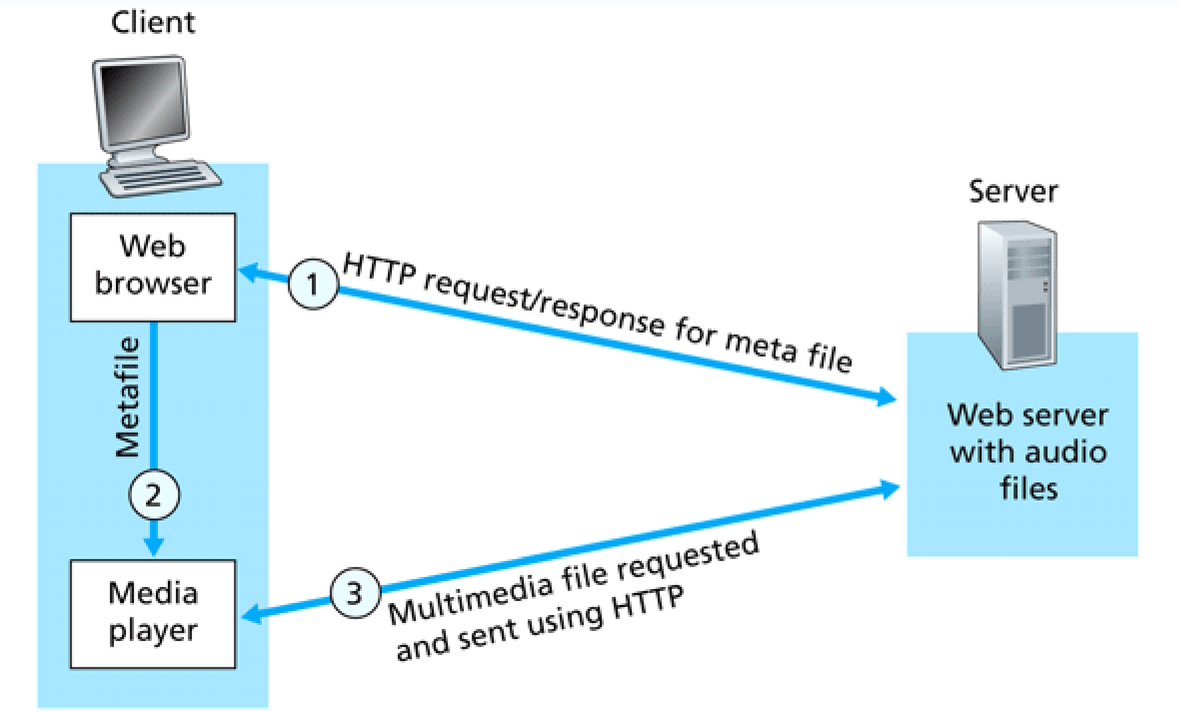
- Most IP phones today will support SIP, RTP and RCTP
  - Softphone: software that implement these protocols; run on normal computers
    - Note that Skype uses its own proprietary protocol, not SIP or RTP
  - Standalone phone: hardware built for the purpose of an IP phone
  - Adapters for PSTN phones

Streaming Stored Audio/Video

Streaming Audio/Video

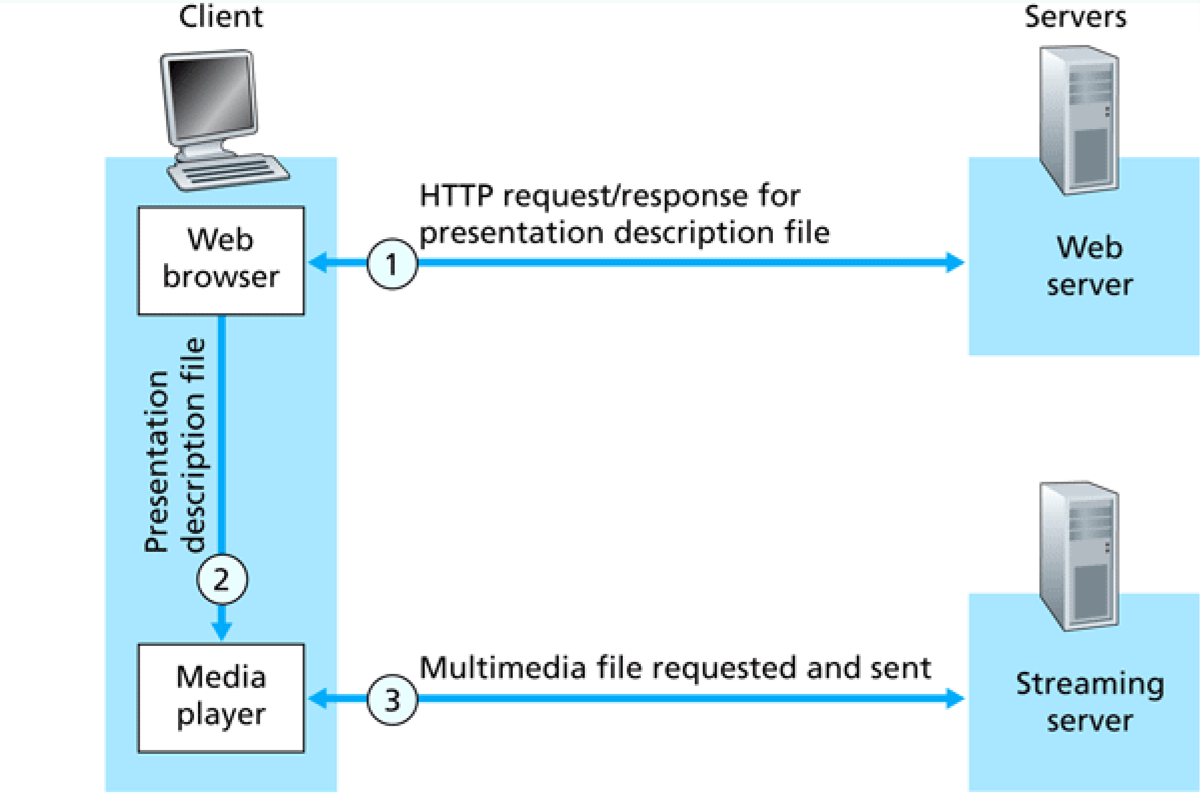
- Examples:
  - Internet Radio
  - Web-based Video: Youtube
- Approach:
  - Content (audio/video) is stored on a streaming server
    - For live content, it is generated and stored on the server
  - Users request content
    - Clients are typically web browsers and media players
      - Media player may be standalone (e.g. Windows Media Player, WinAmp, …) or embedded in web pages (e.g. Flash Media Player)
    - Requests are either direct to streaming server or via a separate web server
  - Content is sent from streaming server to client
    - Using standard (RTP, HTTP) or proprietary protocols

Accessing Content on Web Server



1. Metafile describes the content: location, encoding, name, …
   - Examples: ASX, RAM, PLS, SMIL, …
2. Web browser launches the media player and sends the metafile to player
3. Media player uses HTTP to request the content

Accessing Content from Streaming Server



- Metafile accessed via web server, whereas content is accessed via streaming server
- HTTP is not needed for content delivery
  - Can use RTP or other protocols

Real-Time Streaming Protocol

- RTSP can be used for controlling the stream playback
  - Start, Pause, Describe streams

IPTV

Viewing TV and Videos in Networks

- Note: not limited to traditional TV programming; also includes video-on-demand (VoD) and other content
- Three main approaches:
  - Internet or Web-based Television/Video
    - Using the public Internet (especially WWW) to view video
    - Small image (post card sized) on PC
    - Speeds less than 1Mb/s required for acceptable quality
  - File Based TV/Video Distribution
    - Viewed on a PC or TV
    - Non-real-time (i.e. download entire file, watch at any time), quality depends on coding
    - Accessed from normal Internet, usually using P2P file sharing
  - IPTV
    - High quality image, real-time reception on large TV display
    - Transfer requires “network in network” (much more control than normal Internet)
      - Multicasting, QoS, caching
    - Separate network than Internet

Customers Equipment for IPTV

Example IPTV Applications

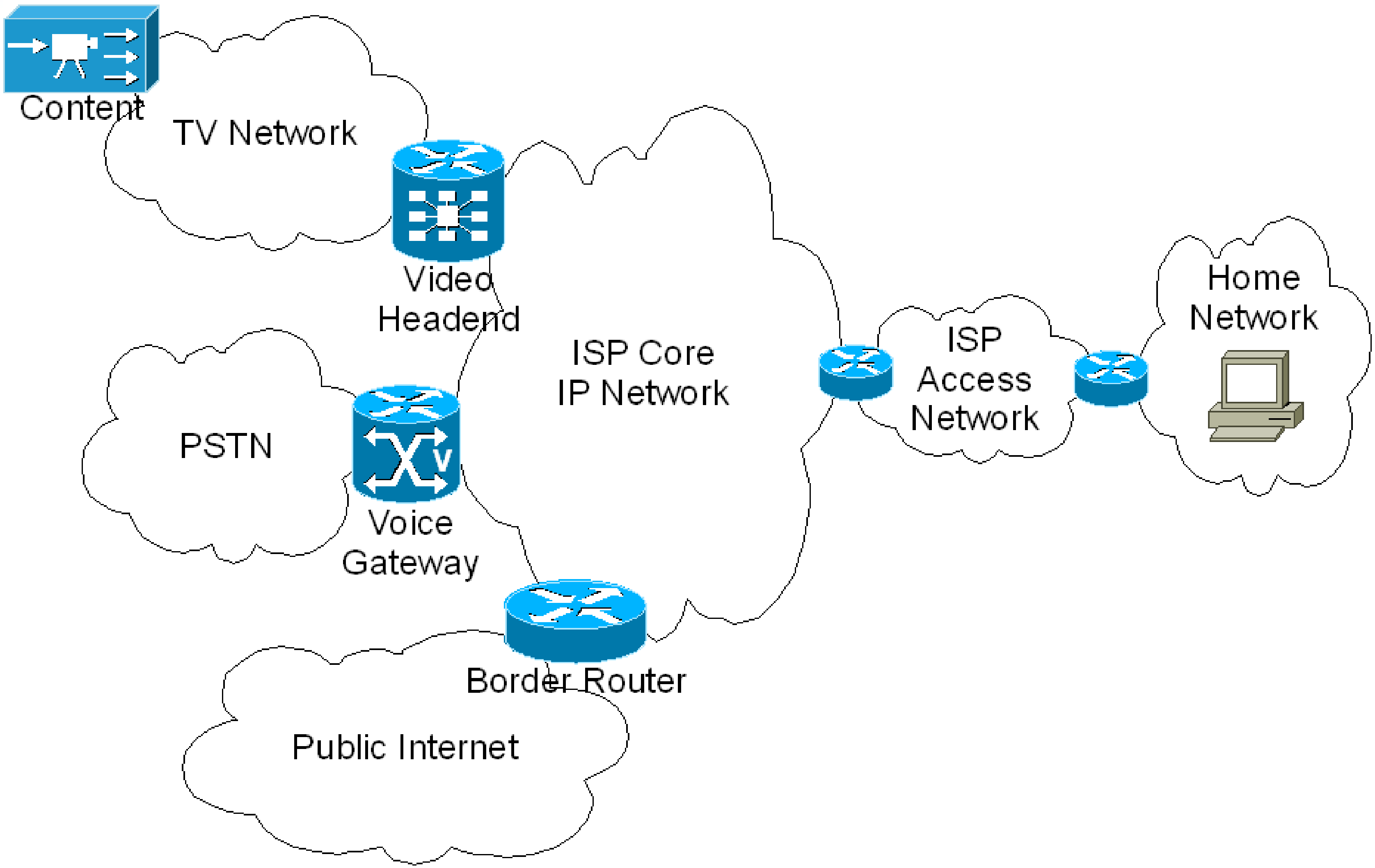
- Digital Television
  - Delivering existing and new digital TV content to consumers
- Video on Demand (VoD)
  - Users can select specific video content, usually for a fee (similar to “pay-per-view”)
- Business TV to Desktop
  - E.g. employees view news channels or financial reporting
- Distance Learning
  - Although traditional teleconference systems support lectures, IPTV will deliver content to the individuals (rather than conference rooms)
- Corporate Communications
  - Director or CEO delivering speeches to employees
- Mobile Phone TV
  - With high-speed wireless data networks, the most practical way of delivering TV to mobiles
- Video Chat

IPTV using Private Networks

- Many companies are looking to deliver IPTV over private IP networks
  - Either existing IP networks for Internet access, or separate IP networks

- Why a separate IP network?
  - To deliver the quality expected for standard TV (including high definition digital TV), require a high level of control over network operation
  - If a company (ISP, TV network, Cable company) owns/operates the entire IP network, they can control the performance delivered to applications

IPTV Network



Technologies for IPTV

- Devices
  - Video Headend: converts audio/video into appropriate digital format for transmission (e.g. MPEG2, MPEG4)
  - Set Top Box (STB): manage IPTV content within customers network
- Protocols
  - Video delivery: RTP, RTSP etc.
- Network Management and Control
  - Multicast
  - QoS control
  - Authentication, authorisation, accounting, …
- Network Technologies
  - Core Networks: SDH, optical fibre
  - Access Networks: ADSL2, optical fibre, coaxial cable, Ethernet
  - Home Networks: Ethernet, wireless LAN

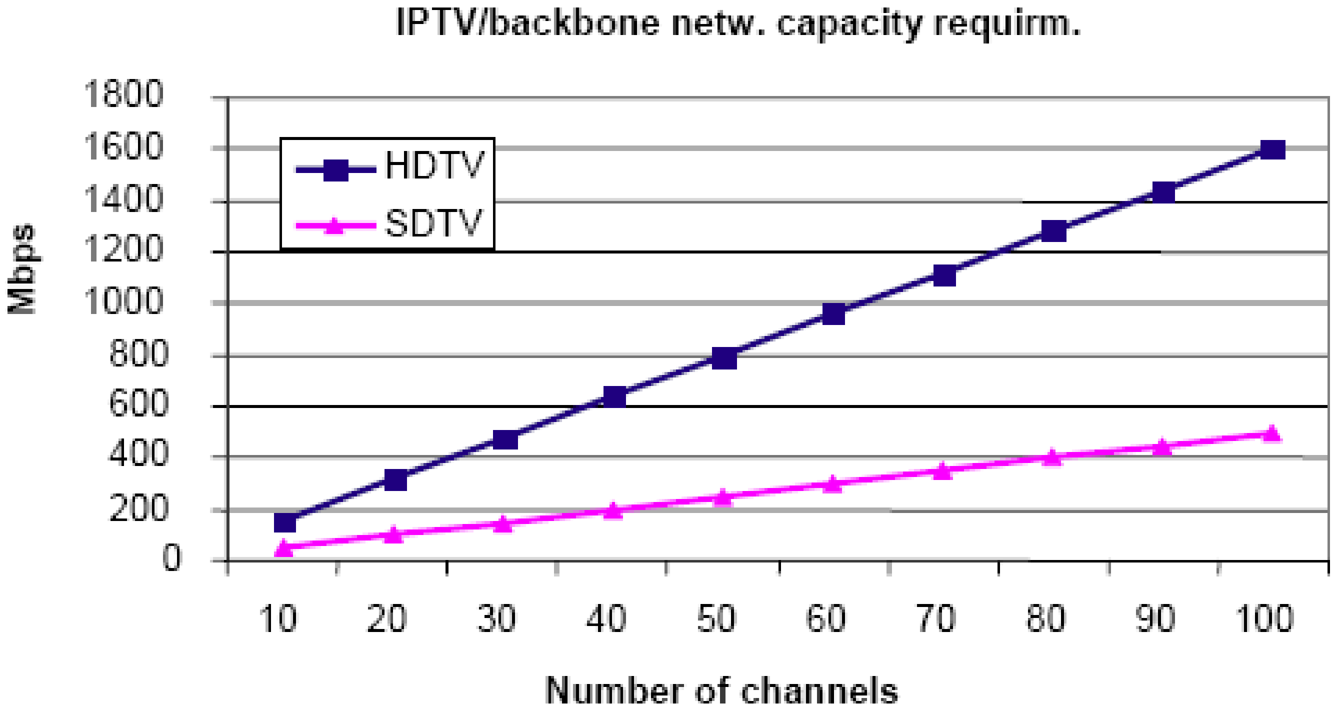
IPTV Bandwidth Requirements

- Lets consider example scenario in a home:
  - Digitized voice: 64kb/s (per voice call)
  - High speed data access: 2 to 4Mb/s (per user)
  - Standard Definition TV (SDTV): 2 to 4Mb/s (per channel)
    - 720 x 576 (width x height) pixels
    - Analog TV, Digital TV, SVCD, DVD, DV
  - High Definition TV (HDTV): 8 to 10Mb/s (per channel)
    - 1080 x 720, 1260 x 1080, …
    - 1920 x 1080 (Full HD)
    - HDTV, Blueray Discs, HD DVD
- Then a house may require 15Mb/s to 30Mb/s

- The “bottleneck” is usually the “last mile”: Service Provider Access Network

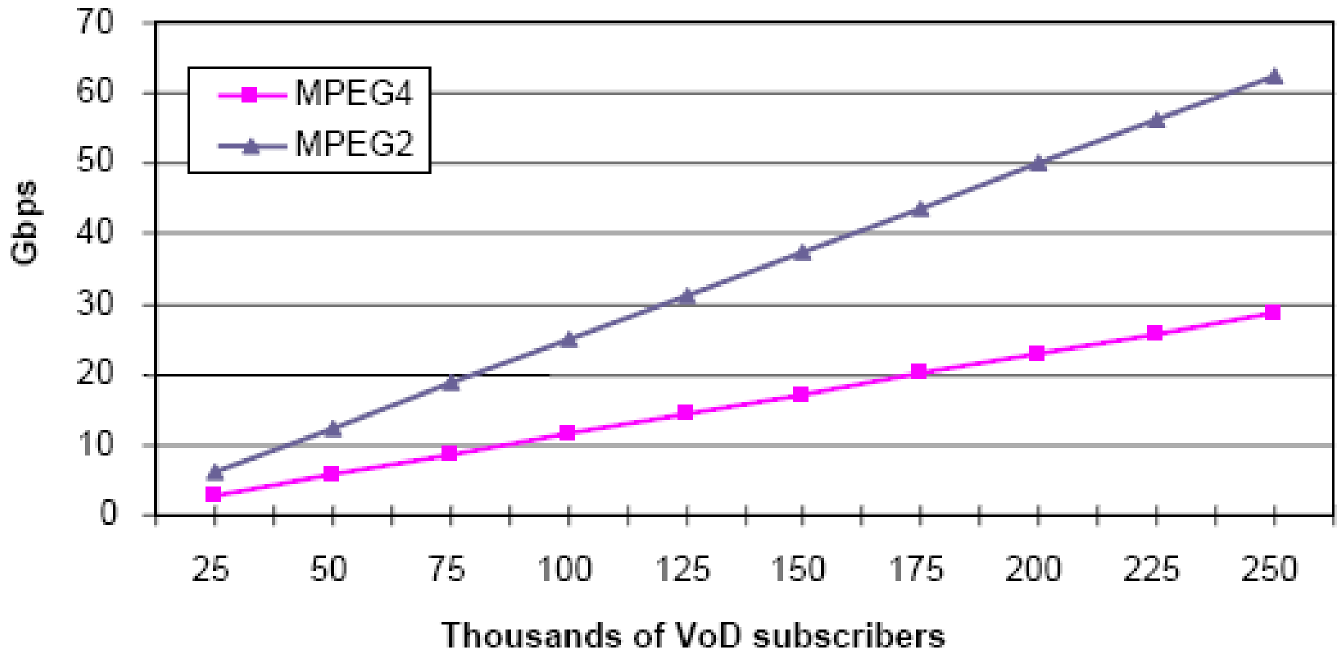
Example: Core Network Requirements

- Service Provider IP Network



Example: Core Requirements for Video on Demand

- With true VoD, need to use unicast (send separate stream to individual subscribers)



Technologies for Service Provider Access Network

- ADSL and ADSL2+
  - Uses existing copper telephone lines
  - Download speeds depend on distance from telephone exchange

Distance (km)	ADSL (Mb/s)	ADSL2+ (Mb/s)
0.3	12.5	26.0
1	12.5	25.5
2	11.0	15.5
3	7.5	7.5

  - ADSL2+ (and similar DSL technologies) are only suitable if the termination point is close to the home (distance is short)
  - Hence, fibre installations are typically need to either:
    - Bring the termination point closer to the home
    - Connect directly to the home (removing the need for copper/ADSL)

Technologies for Service Provider Access Network

- Fibre-to-the-Node:
  - Optical fibre connects to nodes or cabinets in a neighbourhood (100’s to 1000’s of homes)
  - Existing copper (ADSL) or coaxial cables (HFC) are then use from the node to the home
- Fibre-to-the-Curb:
  - Usually to the street-level, support several or 10’s of users
  - Again, copper or coaxial to the home
- Fibre-to-the-Home:
  - Fibre runs direct to each home (or business, building), directly connecting to the home network
  - No need for ADSL, HFC or other (much slower) alternatives
- Summary:
  - Optical fibre can support speeds of Gb/s+
  - The closer the fibre gets to home, the better (however usually very expensive to install!)
  - Other options: wireless (IEEE 802.11n), Ethernet (especially for businesses)…