# Multimedia

## COMP90007
## Internet Technologies

Chien Aun Chan

# Outline: Application Layer

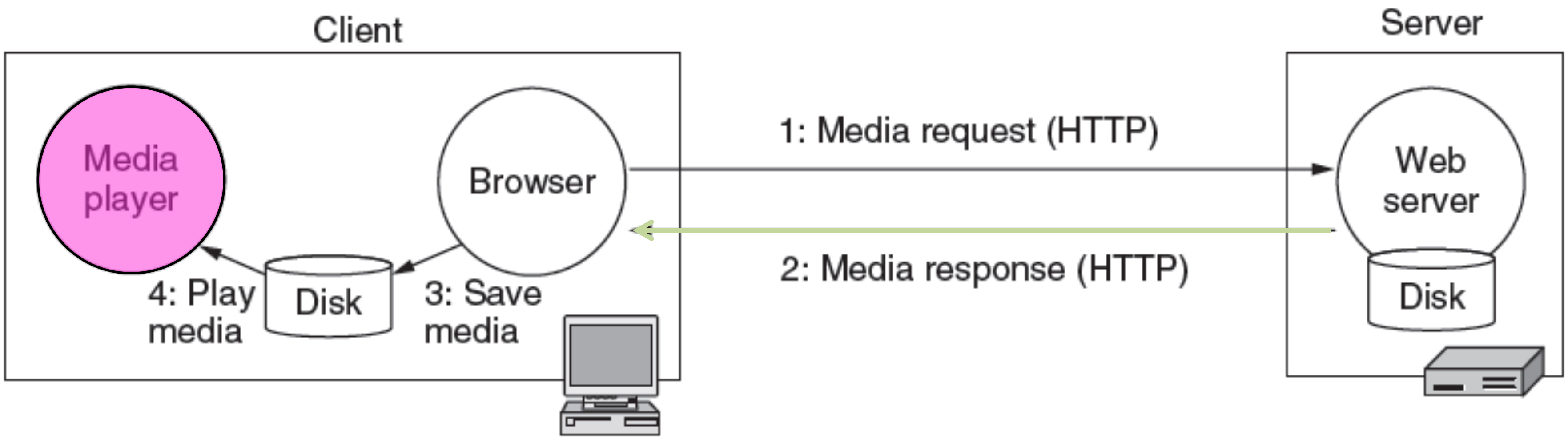
| |
|---|
| Application |
| Transport |
| Network |
| Link |
| Physical |

- Multimedia Networks
  - Audio
  - Video
- Audio and video have become key types of traffic, e.g., voice over IP, and video streaming.

# Characteristics of Multimedia Networks

- Higher bandwidth requirements
- Higher QoS requirement (i.e., delay sensitive)
- New infrastructure models
  - Need separate multimedia servers from web servers
- New service providers
  - Streaming multimedia service providers are often separated and highly specialised, compared to traditional web hosts

# A Basic Model for Multimedia on the Web



# Problems with the Basic Model

- The entire media file must be transmitted over the network before playback starts, causing delay in user experience (e.g., to transmit a  5Mb file over a 56Kbps link takes about 5 minutes)
- Basic model assumes mainly point-to-point media distribution rather than a point-to-multipoint (broadcast) distribution model
- Does not scale
- Basic model assumes the browser/plugin/helper integration and traditional service types

# Streaming Media Protocols

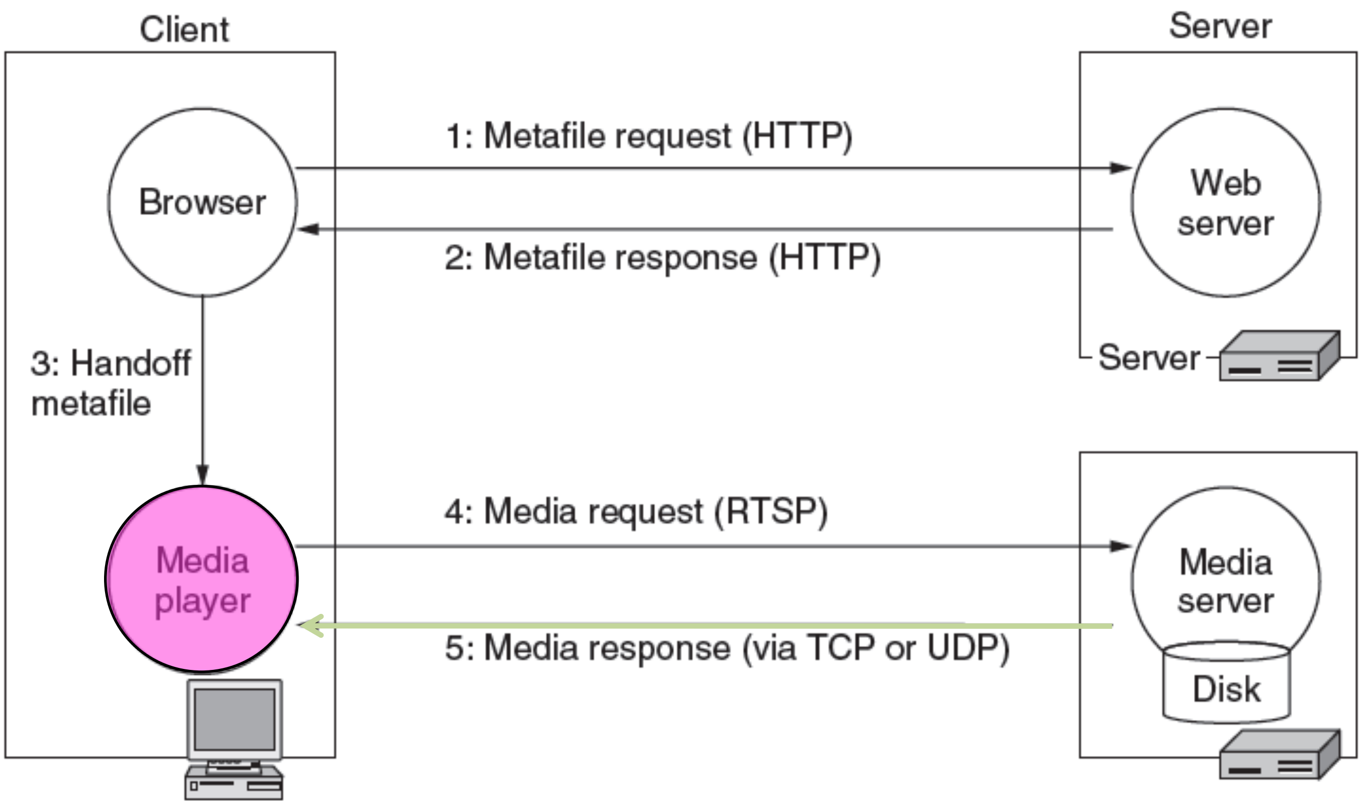
- Transport Protocols
  - TCP
- Open Protocols
  - HTTP
  - RTP - Real-time Transport Protocol (RFC 1889)
  - RTSP - Real Time Streaming Protocol (RFC 2326)
  - MPEG-4 (ISO)
- Closed Protocols
  - Real Networks' RealAudio
  - Microsoft's Windows Media
  - Apple's QuickTime

# The Role of Multimedia Playback Software

- 4 main tasks of the multimedia playback software
  - Manage the user interface (e.g., volume, playback, next, etc..)
  - Handle transmission errors in conjunction with transport protocols
    - Using RTP/UDP errors will likely occur, playback software must manage them gracefully
  - Decompress the multimedia files (codecs)
  - Eliminate jitter
    - Small buffer, quick playback but susceptible to jitter/delay
    - Large buffer, delay at start of playback while buffer fills, but less susceptible to delay/jitter

# Manage the user interface

Effective streaming starts the playout during transport



# Handling Errors: Streaming Stored Media

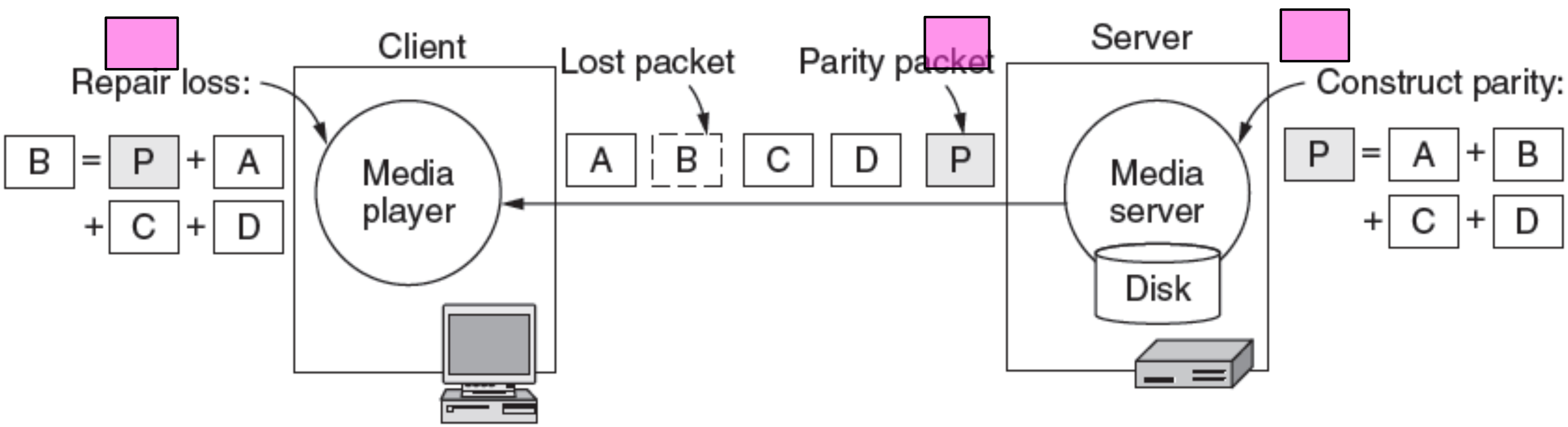
| **Strategy** | **Advantage** | **Disadvantage** |
|---|---|---|
| Use reliable transport (TCP) | Repairs all errors | Increases jitter significantly |
| Add FEC (e.g., parity) | Repairs most errors | Increases overhead, decoding complexity and jitter |
| Interleave media | Masks most errors | Slightly increases decoding complexity and jitter |

Forward Error Correction (FEC) is simply the error-correcting coding. For every four data packets a fifth  The parity packet, P, contains redundant bits that are the parity or exclusive-OR sums of the bits in each of the four data packets.

# Streaming Stored Media

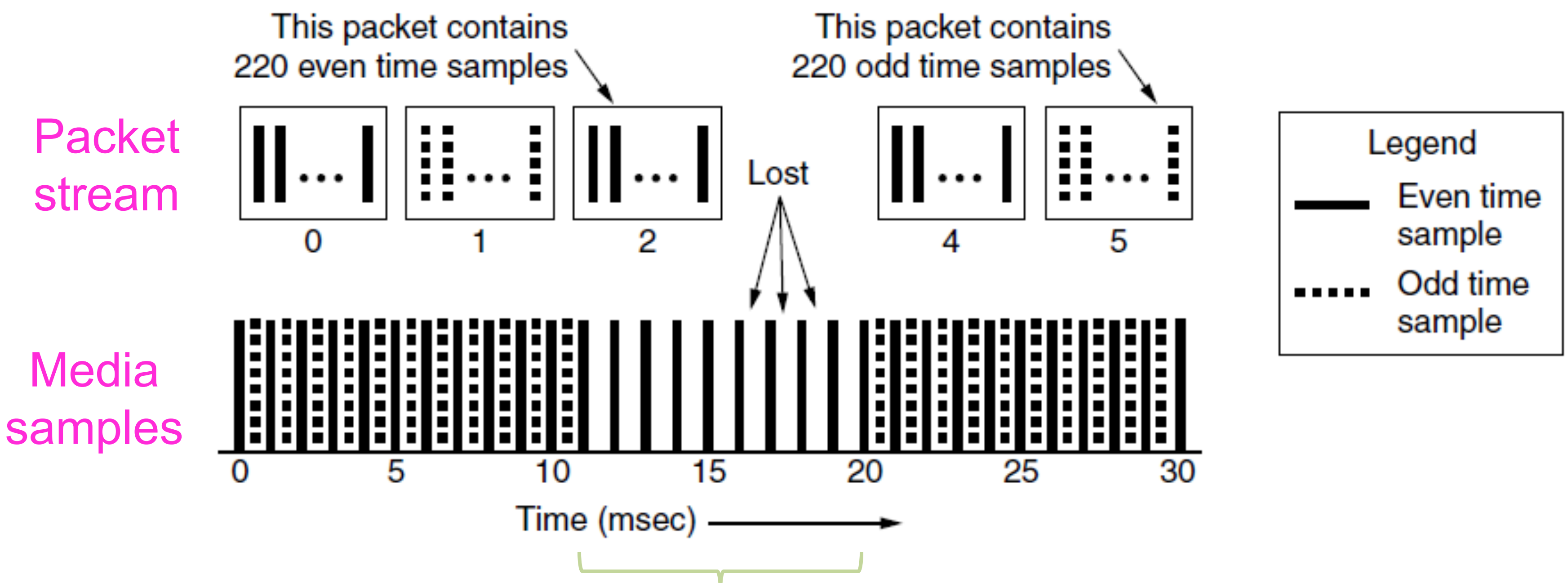
Use of Parity packets can repair one lost packet in a group of N

- Cons: Decoding is delayed for N packets



# Streaming Stored Media

Interleaving spreads nearby media samples over different transmissions to reduce the impact of loss



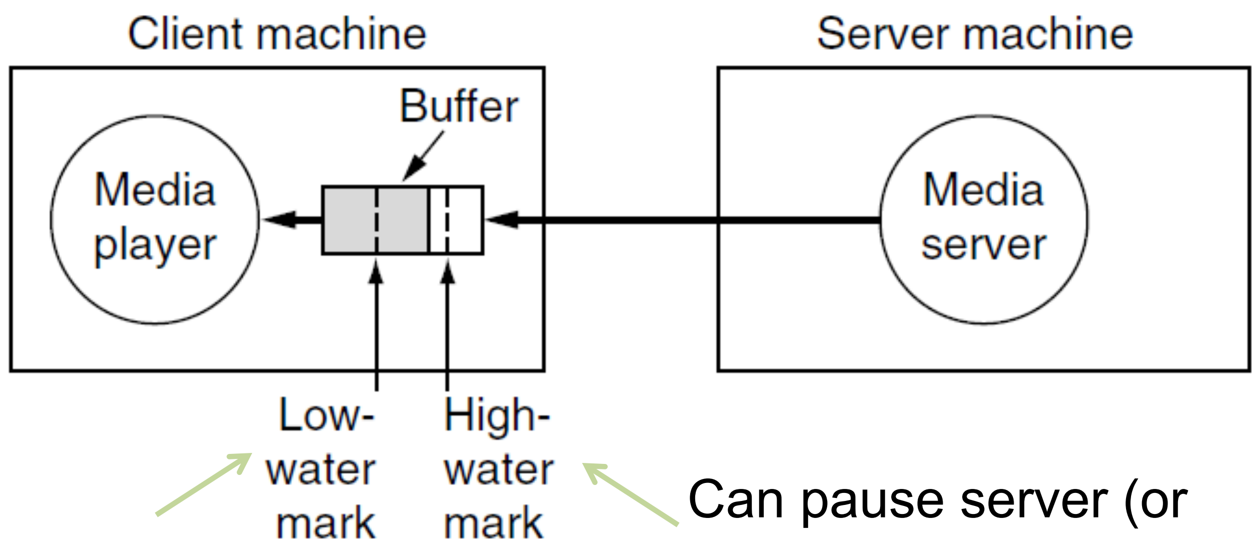

Loss reduces temporal resolution; doesn't leave a gap

# Jitter Management

- Similarly to router buffering, multimedia software buffers streamed media sources prior to transmission
- Buffering is a defensive mechanism to reduce jitter (variance in average packet arrival times)
- Ideally the stream buffer will continue to be filled at the same rate the stream is played back to the user

# Jitter Management

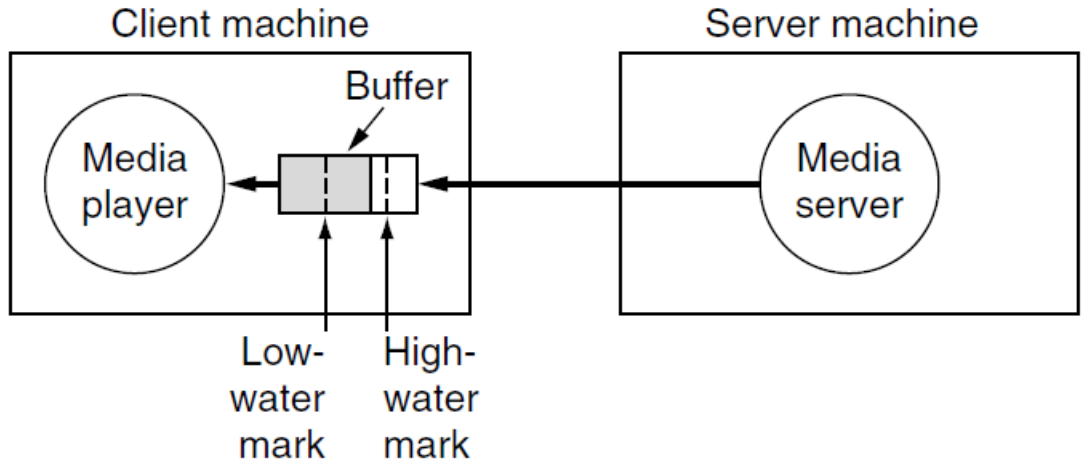
Jitters happens because of variable bandwidth and loss/retransmissions




Safety margin, to avoid a stall

Can pause server (or go ahead and save to disk)

# Example




An video streaming server takes 10 ms to communicate with a client media player. If the media player has a 2 MB buffer, what can you say about the position of the low and high-water mark? Assume the media being streamed has a bitrate of 1 Mbps and the server sends data at a fix rate of 2 Mbps.

*Ans: It takes 10 ms to send a pause command to the server and another 10 ms for the data already in the network to drain after the server stops sending.*

***HWM**: Bandwidth delay product= 20 ms × 2 Mbps = 5,000 bytes will arrive, so the high-water mark should be at least 5,000 bytes below the top to avoid buffer overflow. To be absolutely safe, set the mark to be 20,000 bytes from the top.*

***LWM**: 20 ms × 1 Mbps = 2,500 bytes for the client to request data from the server and start receiving it. 2,500 bytes will be played during this time. Hence, the low-water mark should be at least 2,500 bytes and probably 15,000 bytes to be safe.*

# Real Time Streaming (RTSP)

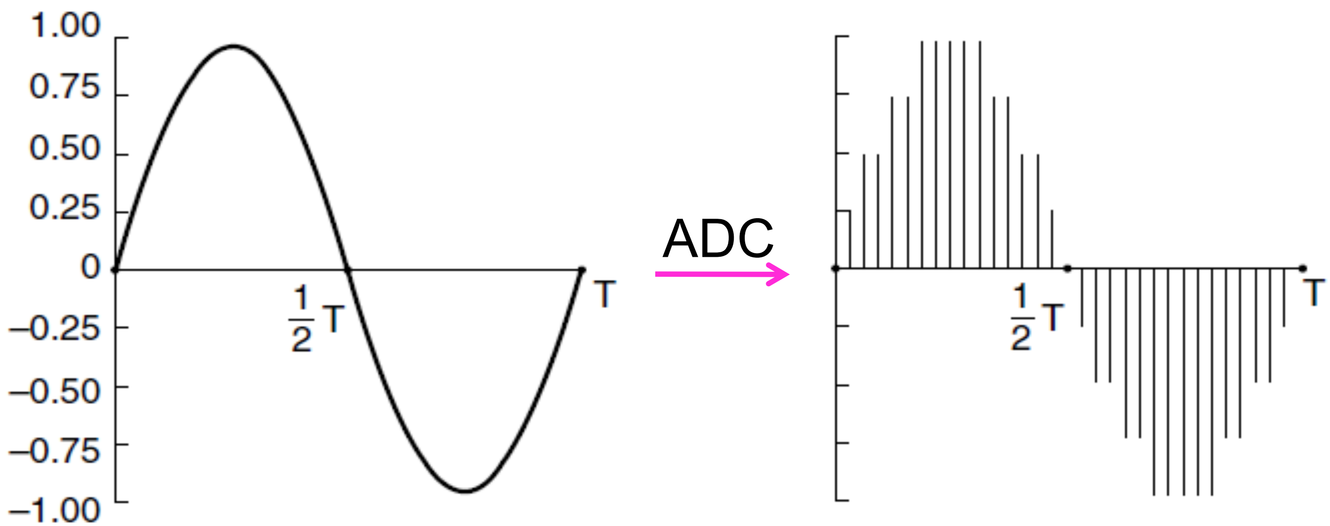
- RTSP has mechanisms to start and stop  the flow of streaming media:RTSP commands are sent from player to server to adjust streaming
- Defined in RFC 2326;
- For Data stream, RTP over UDP or RTP over HTTP over TCP
- RTSP commands are sent from player to server to adjust streaming

| Command | Server action |
|---|---|
| DESCRIBE | List media parameters |
| SETUP | Establish a logical channel between the player and the server |
| PLAY | Start sending data to the client |
| RECORD | Start accepting data from the client |
| PAUSE | Temporarily stop sending data |
| TEARDOWN | Release the logical channel |

# Digital Audio: Basics

ADC (Analog-to-Digital Converter) produces digital audio from a microphone

- Telephone: 8000 8-bit samples/second (64 Kbps); computer audio is usually better quality




Continuous audio (sine wave)

Digital audio (sampled, 4-bit quantized)

# Audio Compression

- Audio CD can represent frequencies up to 22.05kHz, hence Nyquist (sample) rate is 44.1 kHz, Stereo channels: 44100 samples/sec, 16 bits/sample = 2*44100*16 = 1,411,200 bits/s

- The key property of perceptual coding is that some sounds can mask other sounds and at any point those sounds are identified and encoded for transmission
- Frequency masking: Some sounds mask/hide others so there is no point encoding them.
- Temporal masking: Human ears can miss soft sounds immediately after loud sounds, takes time for the ear to adjust, no need

# Digital Video

- Video is digitized as pixels (sampled, quantized)
  - TV quality: 640x480 pixels, 24-bit color, 30 times/sec ~ 200Mbs uncompressed
- Video is sent compressed due to its large bandwidth
  - Lossy compression exploits human perception
    - E.g., JPEG for still images, MPEG, H.264 for video
  - Large compression ratios (often 50X for video)

# Compression with JPEG

- JPEG lossy compression sequence for one image
- JPEG often provides compression ratios of 20:1
- JPEG compression is symmetric, decoding takes as long as encoding

# MPEG Standard

- MPEG - Motion Picture Experts Group
- MPEG can compress both audio and video together (using synchronised streams)
- The evolution of MPEG
  - MPEG-1: VCR quality at 1.2 Mbps (40:1)
  - MPEG-2: Broadcast quality at 4-6Mbps (200:1)
  - MPEG-4: DVD quality at 10Mbps (1200:1)

# Video over the Web

- Embedded with web content
  - Streaming servers required to deliver content over ordinary connection
- Integrated with other consumer services eg cable television (video on demand)
  - Content stored on central video servers
  - Delivered via a shared network to consumers

# Outline

- Voice over IP (VOIP)
  - VOIP Benefits
  - VOIP Technologies
  - Protocols

# The Emergence of VOIP

- Voice services becoming applications on top of data networks are being driven by a number of factors:
  - Data has overtaken voice as the primary traffic on many networks originally built for voice
  - PSTN infrastructure is not flexible enough for the rapid deployment of new features
  - PSTN technologies are largely incompatible with the convergence of data/voice/video
  - The architecture built primarily for voice is not flexible enough to carry data
  - Network providers are increasingly looking to leverage investment in network infrastructure by bringing new services to data networks
- Where suitable data networks are already in existence, the evolution of audio encoding technologies has allowed voice to be transmitted over data links - hence VOIP

# Benefits of VOIP

- Financial savings
- Consolidated infrastructure
- Flexible infrastructure
- Standards based voice and data

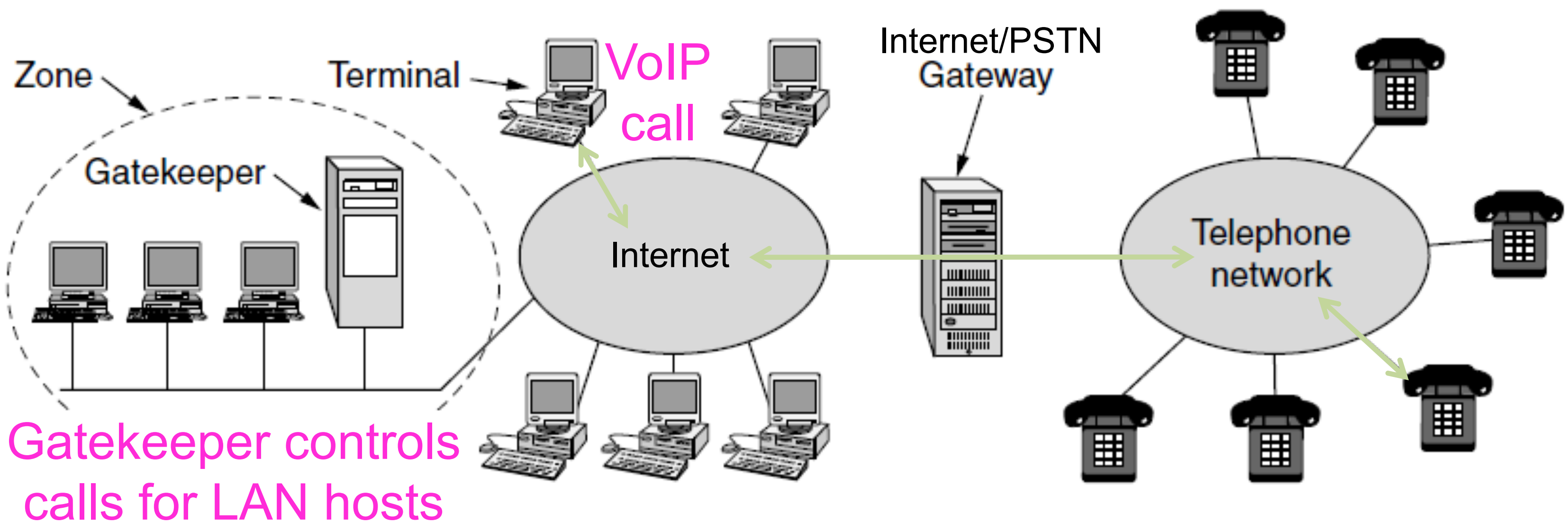
# VOIP Technologies

- Within the VOIP domain, there are 3 distinct models for service provision
  - infrastructural - PSTN/PABX integration
  - virtual - media gateways, virtual directories
  - value-added - voice mail
- Alternative VOIP Technologies
  - H.323
  - SGCP (Simple Gateway Control Protocol) and MGCP
  - SIP

PABX - private automatic branch exchange

# H.323 architecture for Internet telephony

supports calls between Internet computers and PSTN phones.

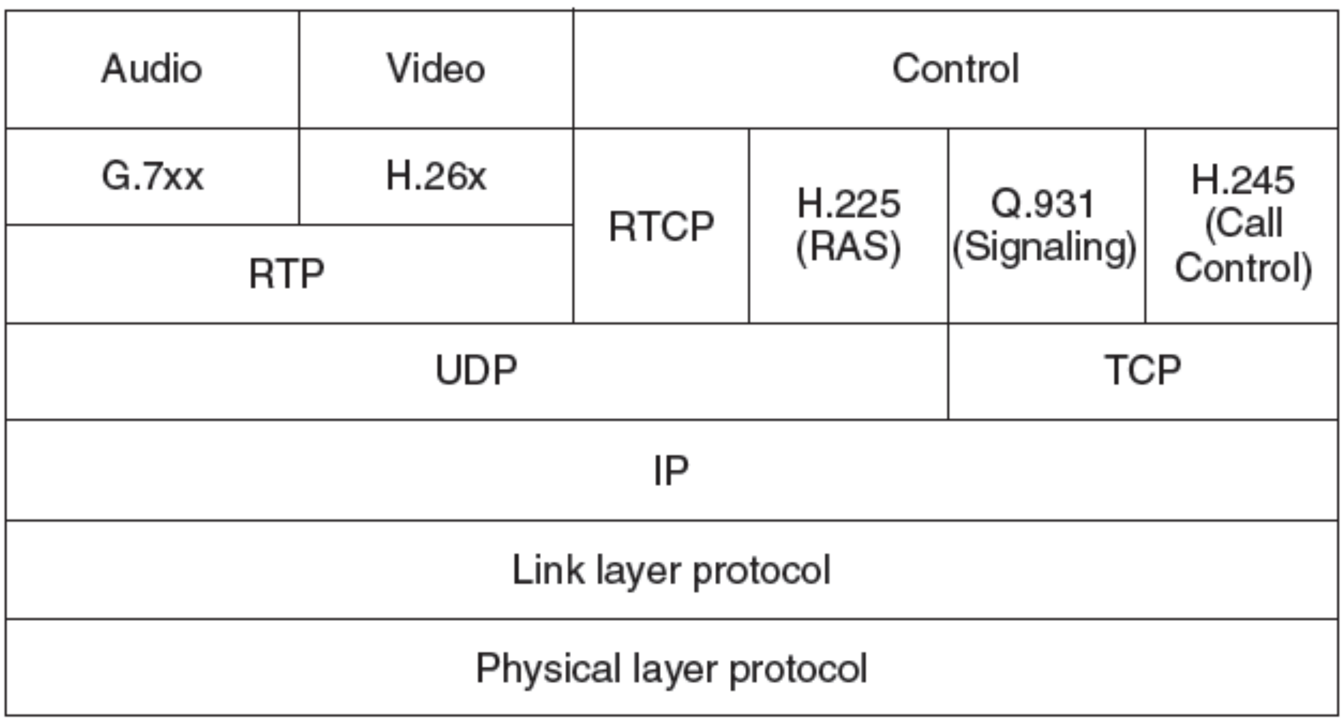


# H.323

- H.323 is an international standard that specifies how multimedia traffic is carried over packet networks
- H.323 integrates existing standards to provide a layered protocol stack
- H.323 was originally developed to enable multimedia applications to run over unreliable data networks - includes support for audio
- (including VOIP), video and data sharing services within a single protocol stack

# The H.323 Protocol Stack

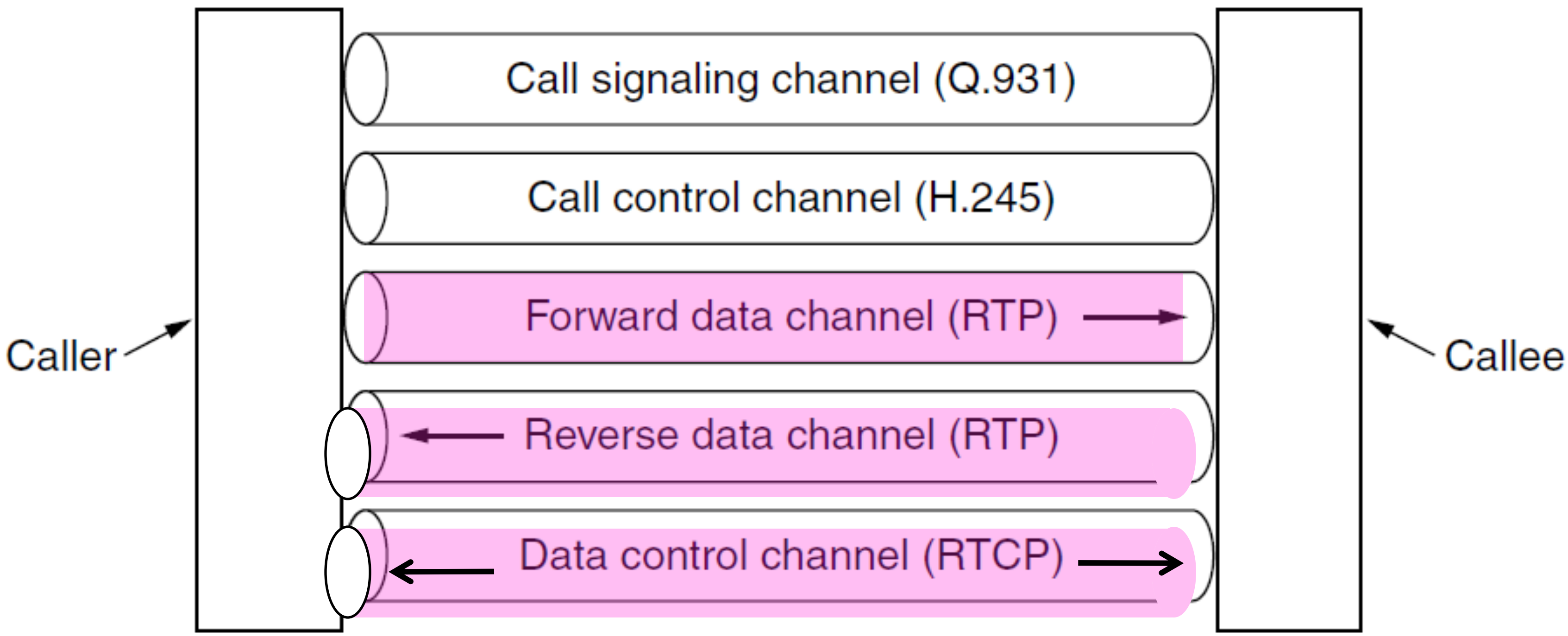
- Call is digital audio/video over RTP/UDP/IP
- Call setup is handled by other protocols (Q.931 etc.)



# The H.323 Protocol Stack

- G.7xx e.g. G.711 encodes a voice channel by sampling 8000 times/sec with 8 bit sample rate, giving uncompressed 64kbps speech
  - Other G.7xx compression protocols exist that have different quality/bandwidth tradeoffs
- H.245 protocol handles negotiation to decide which compression algorithm to use, and the bit-rate.
- Q.931 protocol handles connection establishment and release, provides dial-tones, ringing sounds and other telephony functions.
- H.225 (RAS = Registration/Admission/Status) protocol is needed to talk to the gatekeeper, if the client is behind one. e.g. LAN.
- RTP is responsible for data flow once negotiation has been completed.
- RTCP (RTP Control Protocol) manages congestion control, and if the call contains video as well as audio, RTMP handles the synchronisation

# Logical channels in H.323 VOIP



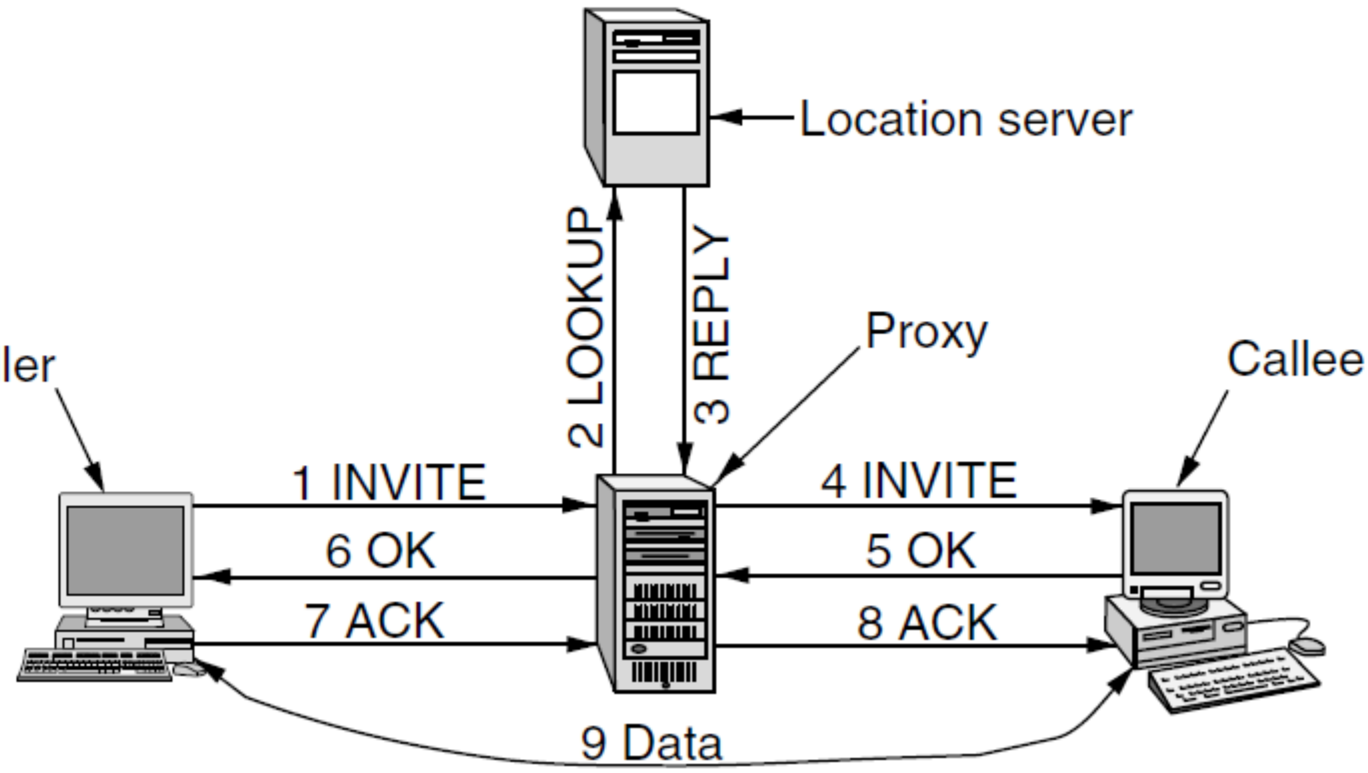
# Session Initiation Protocol (SIP)

- SIP codified in RFC 3261
- Modelled on HTTP, simple ASCII based protocol, method + parameters, similar headers to MIME
- SIP Architecture
  - A single network module rather than a complete protocol suite
- SIP only handles the setup, management and termination of sessions - requires other protocols such as RTP/RTCP for data transport. SIP is an application layer protocol and can run over UDP or TCP
- SIP Functionality
  - Two party, multiparty and multicast
  - Callee location, callee capabilities, call setup, call termination
- SIP Addressing: addressing based on a URL type schema
  - sip:badenh@estragon.cs.mu.oz.au
  - SIP URLs can contain IPv4 or IPv6 addresses or actual telephone numbers

# SIP Methods

| Method | Description |
| --- | --- |
| INVITE | Request initiation of a session |
| ACK | Confirm that a session has been initiated |
| BYE | Request termination of a session |
| OPTIONS | Query a host about its capabilities |
| CANCEL | Cancel a pending request |
| REGISTER | Inform a redirection server about the user's current location |

# SIP Methods


Location server
2 LOOKUP
3 REPLY
Proxy
Caller
Callee
1 INVITE
6 OK
7 ACK
4 INVITE
5 OK
8 ACK
9 Data

# H.323 and SIP Compared

| Item | H.323 | SIP |
|---|---|---|
| Designed by | ITU | IETF |
| Compatibility with PSTN | Yes | Largely |
| Compatibility with Internet | Yes, over time | Yes |
| Architecture | Monolithic | Modular |
| Completeness | Full protocol stack | SIP just handles setup |
| Parameter negotiation | Yes | Yes |
| Call signaling | Q.931 over TCP | SIP over TCP or UDP |
| Message format | Binary | ASCII |
| Media transport | RTP/RTCP | RTP/RTCP |
| Multiparty calls | Yes | Yes |
| Multimedia conferences | Yes | No |
| Addressing | URL or phone number | URL |
| Call termination | Explicit or TCP release | Explicit or timeout |
| Instant messaging | No | Yes |
| Encryption | Yes | Yes |
| Size of standards | 1400 pages | 250 pages |
| Implementation | Large and complex | Moderate, but issues |
| Status | Widespread, esp. video | Alternative, esp. voice |

# Evolving Voice over Data Solutions

- First voice-data integration technologies were designed to eliminate long distance phone toll charges by providing tie lines between PABXs over a WAN infrastructure
- Later support for analogue telephony devices was introduced to allow off-premises extensions to PABX
- As data networks expanded, enterprise wide call handling began to migrate towards the data network - shorter call forwarding distances between PABX and WAN gateways
- With increased voice traffic, connection admission control became a more significant issue
- With larger networks, dial plans and directory services became essential
- Centralised call control through the use of H.323 gatekeeper functions allowed voice and data transmissions to be regulated at a single point

# Future of Network-Centric Telephony Applications

- As integration increases, new solutions are emerging – typically allowing for packetised voice technologies to replace PABX with end to end solutions
- Generally 2 categories of technologies and application architectures
  - UnPABX: Server based call routing, with direct connections to data network, trunk telephony network and analogue telephony networks
  - LAN-PABX: Telephony on the desktop - rather than actual telephone handsets, the telephony function is in software on a workstation

# Incentives for Packet Based Telephony

- Un-PABX systems typically cost less than systems they replace
- LAN-PABX systems bring much greater flexibility to end users
- Both leverage existing infrastructure - either the convergence of communications infrastructure or wired/wireless connections to the workstation
- Fully integrated applications, which allow manipulation of complex software services via telephony interfaces will likely drive the convergence of voice and data even further

# Summary - Multimedia

- Multimedia networks
  - General operations of audio and video streaming
  - Describe techniques for jitter management
- VoIP
  - Contrast H.323 and SIP
  - Explain steps in SIP call establishment