18-796

Multimedia Communications:
Coding, Systems, and Networking

Prof. Tsuhan Chen
tsuhan@ece.cmu.edu

What We Have Done...

- VQ and Subband Coding
- Motion Estimation
- JPEG, JPEG-2000
- H.261, H.263, H.263 Version 2
- MPEG-1,2,4,7
- MPEG Audio
- Networking Issues
  - Network characteristics and error resilience
  - RTP, RTCP, RTSP, RSVP
  - Multimedia over ATM
- Systems: H.32x, MPEG-1/2/4

Lecture 15

ITU-T Systems

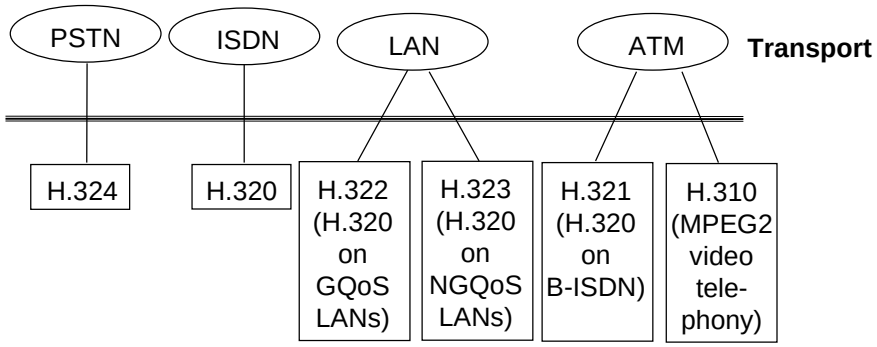
Outline

- Fundamentals
- H.320: N-ISDN
- H.324: PSTN, Wireless Mobile, N-ISDN
- H.323: Packet-based networks

Purposes of System Standards

- Media multiplexing
  - Video, audio, data, and control streams
- Capability negotiation
  - Coding algorithms, bit rate, frame rate, data capability, network capability, encryption, etc.
- System control
  - Media channel signaling, flow control, master-slave determination

ITU-T System Standards



PSTN: Public Switched Telephone Network
ISDN: Integrated Switched Digital Network
LAN: Local Area Network
ATM: Asynchronous Transfer Mode

GQoS: Guaranteed Quality of Service
NGQoS: Non-Guaranteed QoS

Components of Audiovisual Systems

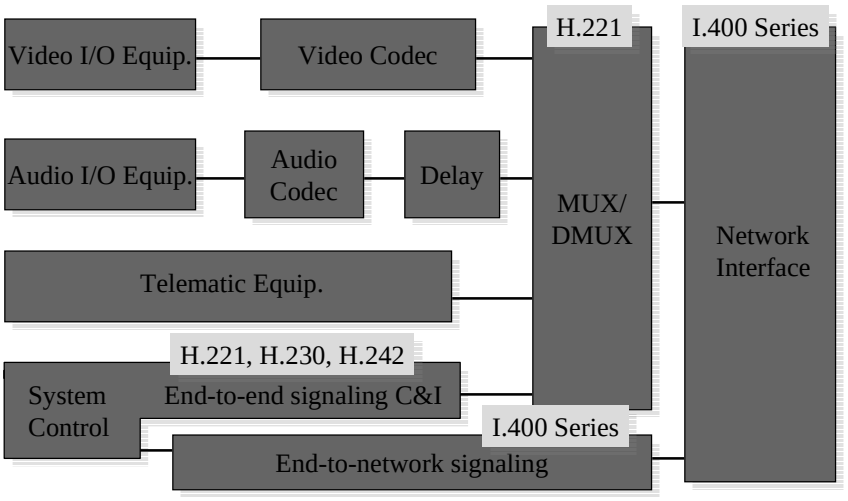
Network		System	Video	Audio	Mux	Control/ Signaling	Comm. Interface
WAN	PSTN,N-ISDN Mobile	H.324	H.261, H.263	G.723.1	H.223	H.245	V.34
	N-ISDN	H.320	H.261	G.7xx*	H.221	H.242	I.400
	B-ISDN (ATM)	H.321	H.261	G.7xx*	H.221	Q.2931	I.361/363 I.400
		H.310	H.261/ H.262	G.7xx*/ MPEG-1	H.222	H.245	I.361/363 I.432
LAN	QoS-Guaranteed	H.322	H.261	G.7xx*	H.221	H.242	TCP/UDP IP
	Non QoS-Guaranteed	H.323	H.261, H.263	G.7xx* G.723.1	H.225.0	H.245	

G.7xx*: G.711, G.722, G.728

H.320 Suite

- First determined in 1990
- Intended for use in N-ISDN
  - Bit rate 64 ~ 1920 kbits/s
  - Audio: G.711 is required, G.722, G.723.1, G.728, G.729, and MPEG1 audio are optional
  - Video: H.261 is required, H.262 (MPEG-2 video) H.263 are optional
  - T.120 is optional

H.320 Architecture



C&I: Control and Indication

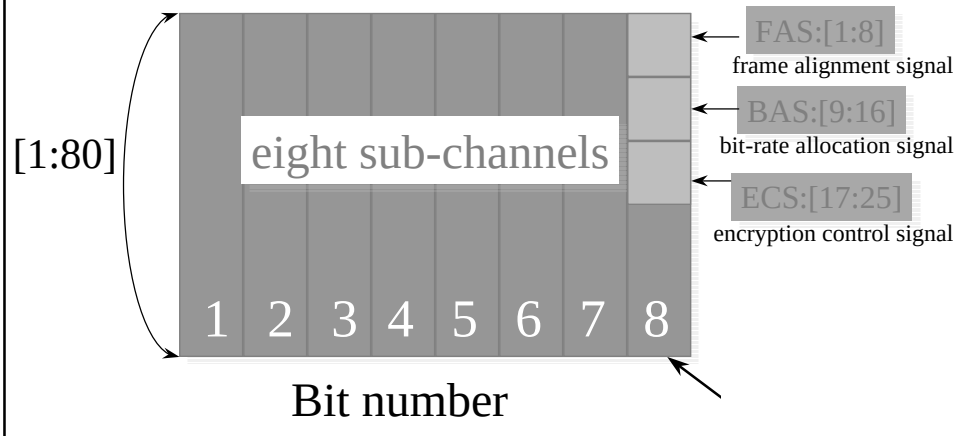
ISDN

- Narrowband ISDN
  - B (64 kbps), H_0 (384 kbps), H_{11} (1.536 Mbps) , H_{12} (1.920 Mbps), D (16 or 64 kbps)
  - BRI (basic rate ISDN) 2B + 1D
  - PRI (primary rate ISDN) 1472 kbps (1856 kbps in Europe) + 64 kbps D-channel
  - PRI can be in several combinations of B and H_0 channels

H.221 Frame Structure

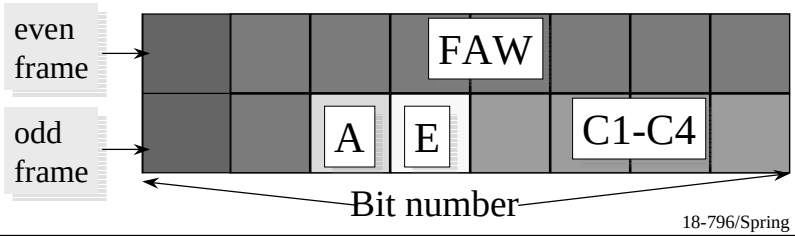
- Basic Principle of Frame Structure
  - Frame structure was designed for 1 to 6 B-connections , 1 to 5 H_0-connections, or an H_{11} or H_{12} connection.
  - A connection is first divided into a number of 64 kbps channels (time slot).
  - Each channel is further divided into eight 8 kbps sub-channels.
  - The 8-th channel is called Service Channel (SC).

H.221 64 kbits/s Frame Structure



Frame Alignment Signal (FAS)

- Frame-level synchronization
  - Frame alignment word (FAW): 00110111
  - A-bit: indication of frame alignment
- Error correction capability
  - E-bit: indication of transmission error
  - C1-C4: redundant bit for cyclic redundancy check



Multiframe Structure

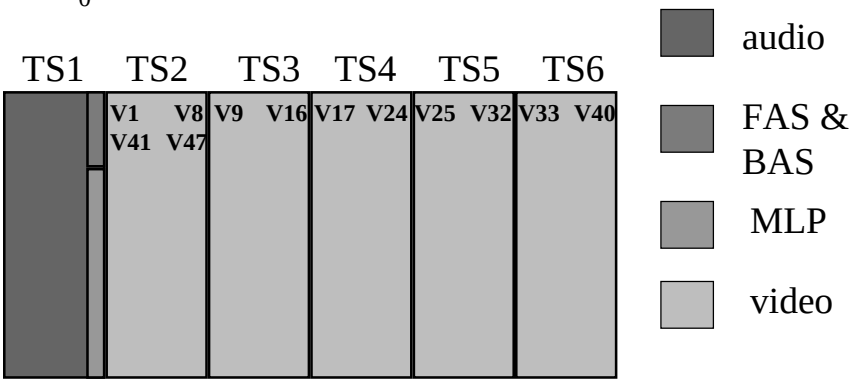
- For connections higher than 64 kbps
  - Each multiframe consists of 16 consecutive frames.
  - Within a multiframe, the ensemble of the first bit of FAS provide *channel number information* (L3-L1) *multiframe number information* (N4-N1).
  - *Multiframe number* is used for synchronization of multiple B or H_0 connections.
  - *Channel number* is used for connections with multiple 64 kbps channels to correctly insert FAS into frames every 125 msec.

Bit-Rate Allocation Signal (BAS)

- Functions and Structure of BAS
  - Used for signaling system capabilities and commands.
  - Encoded by an attribute method
    - the first 3 bits identify the *general* commands or cap. and the other 5 bits indicates the *specific* command or cap.
  - Each BAS is sent with a 8-bit error correction bits generated by (16,8) cyclic code.
    - BAS is in even-numbered frames, the associated error correction bits is in odd-numbered frames.

Example of a Frame Structure

- 56 kbps G.722 audio, MLP 6.4 kbps, and video on H_0



MLP: Multi Layer Protocol

H.242 Function

- Main function
  - Specify how to set video, audio, data, and transfer rate capabilities of a terminal.
  - Specify procedures to conduct capabilities exchange, dynamic mode switching, fault recovery.
  - Provides procedures for call connection, disconnection, and call transfer.

BAS Capability Set

- CapSet can include
  - audio, video, transfer rate, restricted network, single/multiple channel compatibility, LSD, HSD, LS-MLP, HS-MLP, Application in data channel, Cap. in H.230, Audio-ISO Cap, Encryption, MBE
- CapSet Structure
  - {<Cap. Mark> <CapSet>} {<Cap. Mark> <Repeat CapSet>} … <Cap. Mark>

LSD: Low Speed Data
HSD: High Speed Data

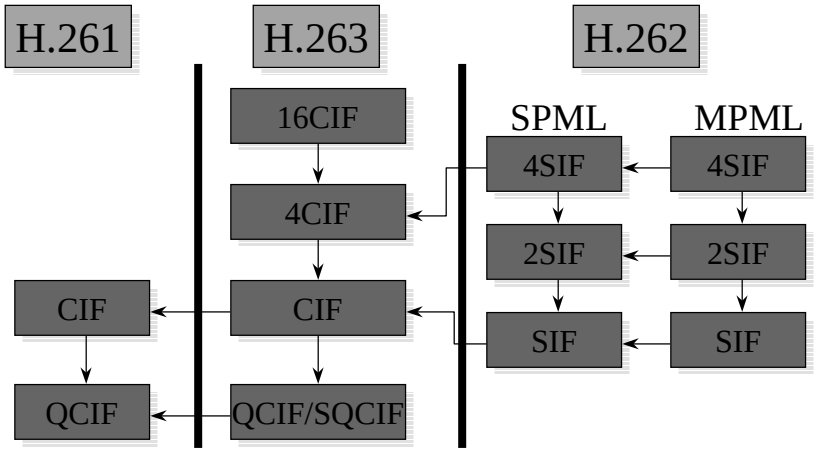
SBE and MBE

- To allow for more number of commands and capabilities to be specified, introduce
  - SBE (Single-byte extension) - two consecutive BAS codes
  - MBE (Multiple-byte extension) - <Start MBE/ N bytes/…/…/…>
- Example of H.263 MBE
  - {Start-MBE/ 7 bytes/ <H.262/H.263> / H.263_4CIF + MPI_4 + Options/AC/ H.263_CIF + MPI_3 + Options/AC+UMV/ H.263_QCIF/SQCIF + MPI_2/AC + PB + UMV}

MPI: Minimum Picture Interval

Video Hierarchy

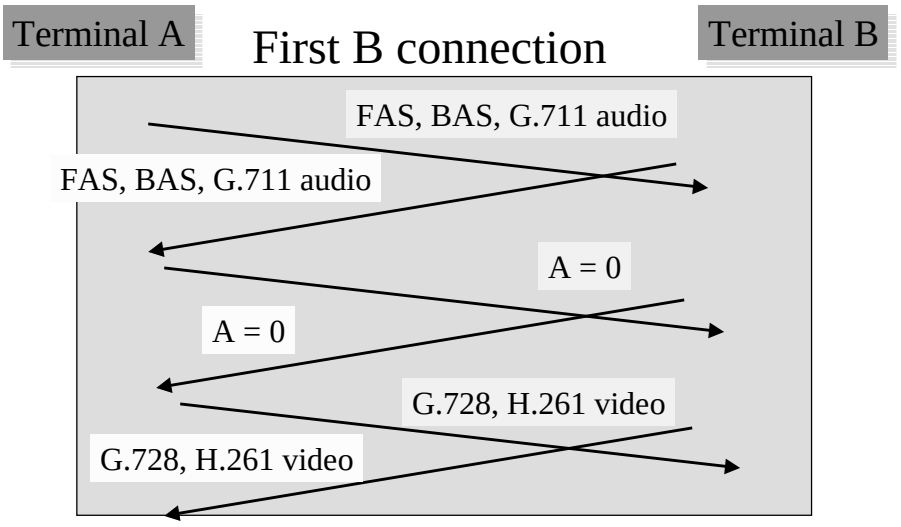
- Enforce a guaranteed level of interoperability



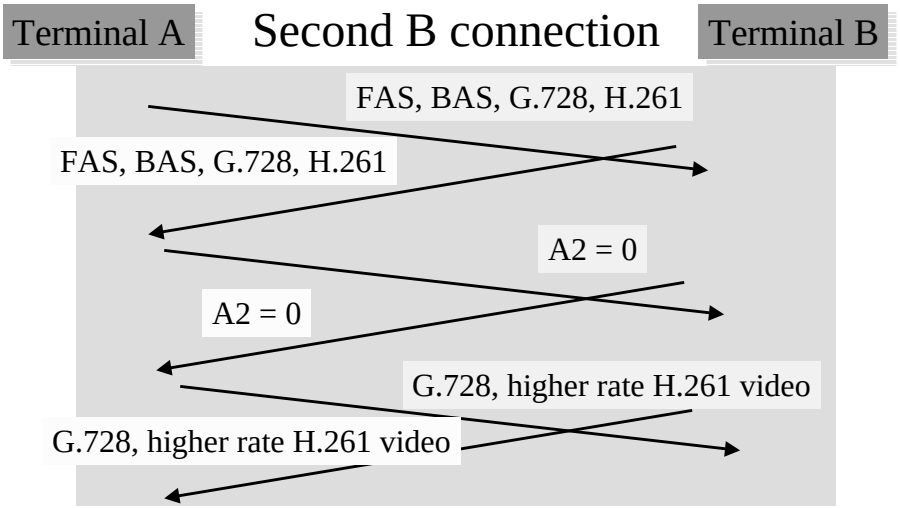
Other Hierarchy

- Audio
  - G.722 (48kbps) > G.722 (64kbps) > G.711
  - G.728 > G.711
- Video
  - higher frame rate cap. indicates lower frame rate cap.
- Transfer rate
  - 6B > 5B > 4B > 3B > 2B > 1B
  - $5H_0 > 4H_0 > 3H_0 > 2H_0 > 1H_0$

Example of a 2B ISDN Call



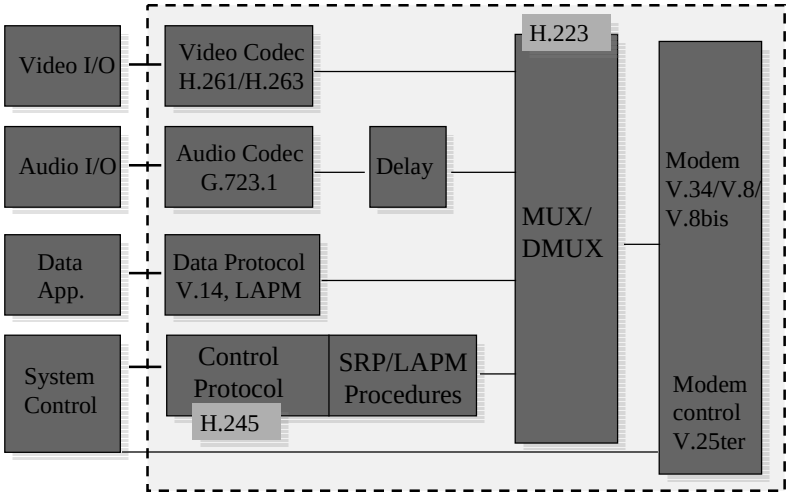
Example of a 2B ISDN Call (Cont.)



H.324 Suite

- First determined in 1995
- Design for circuit-switched network
  - Originally for low-bit-rate PSTN at 9.6~33.6 kbps
  - New revision will cover ISDN (H.324/I) and Mobile (H.324/M)
  - H.261/H.263 video, G.723.1 audio, T.120 data

H.324 Structure

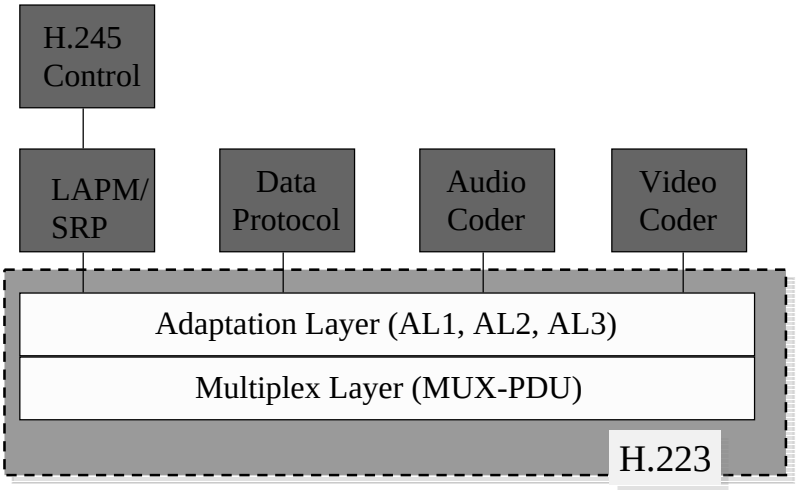


SRP: Simple Retransmission Protocol
LAPM: Link Access Procedure for Modems

Reasons for a New Standard

- H.221 64 kbps frame structure cannot be applied to low-bit-rate PSTN.
- Difficult to upgrade H.242 BAS scheme for meeting the increasingly need of complicated system functionality.
- The less reliable PSTN channel requires more flexible media multiplex and error correction schemes than those H.221 offers.

H.223 Architecture



MUX-PDU: Multiplex-Protocol Data Unit

H.223 Adaptation Layer

- Purpose
  - Add information for error detection/correction, sequence numbering, and retransmission.
- Type
  - AL1: designed for *data* and *control* information. No error protection is provided.
  - AL2: designed for *audio.* 8-bit CRC and optional sequence numbering.
  - AL3: designed for video. 16-bit CRC and optional sequence numbering and retransmission.

Logical Channel (LC)

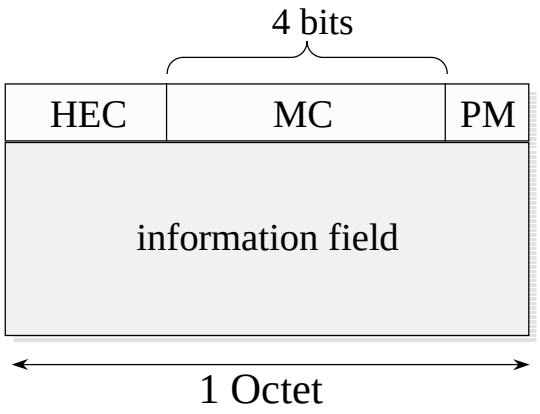
- Each LC can be specified to carry exactly one channel of video, audio, data, and control signal.
- LC number 0 is permanently assigned to the H.245 control channel.
- *Uni-* or *Bi-direction* LCs. LCs can be *dynamically* opened or closed.
- *Segmentable* and *Non-segmenbtable* LC concept, allows one to temporarily suspend transmission of long segmentable LC information, e.g., video, for non-segmentable audio.

H.223 Multiplex Layer

- MUX-PDU
  - H.223 multiplexes media streams from LCs into a variable-length MUX-PDU in a *byte-oriented* fashion.
  - Multiplex table specifies how information from various LCs are mixed in a MUX-PDU.
  - Each MUX-PDU is delimited by HDLC flags (01111110).
  - HDLC zero-bit insertion is needed for media streams to ensure HDLC flag unique
    - Insert 0 after every five 1's

High-level Data Link Control

MUX-PDU Format



- **MC** (Multiplex Code): specify the multiplex table entry
- **HEC** (Header Error Control): CRC error coded
- **PM** (Packet Marker): signal the end of a segmentable LC

Example of an MUX-PDU

- MUX table
  - LCN1: audio, LCN2: video, LCN3: control
  - MUX descriptor:
    {LCN1, RC4}, {{LCN2 RC1}, {LCN3, RC2} RC UCF}
    RC: repeatCount, UCF: untilClosingFlag
- MUX-PDU
  - {HDLC flag}{MUX-PDU header}{LCN1}{LCN1} {LCN1}{LCN1}{LCN2}{LCN3}{LCN3}{LCN2} {LCN3}{LCN3}{HDLC flag}

H.245 Control

- Control Function
  - includes capability exchange, opening and closing logical channels, mode preference request, multiplex table transmission, master-slave determination, flow control commands, general commands and indication.
- Syntax
  - uses ASN.1 (Abstract Syntax Notation version 1) and binary encoded using X.691 PER (packed encoding rules)

Example of H.245 Syntax

```
VideoCapability                 ::CHOICE
{
   nonStandard                  NonStandardParameter,
   h261VideoCapability          H261VideoCapability,
   h262VideoCapability          H262VideoCapability,
   h263VideoCapability          H263VideoCapability,
   is11172VideoCapability       IS11172VideoCapability,
   ...
}
```

Example of H.245 Syntax (Cont.)

```
H261VideoCapability             ::SEQUENCE
{
   qcifMPI                      INTEGER (1..4) OPTIONAL,
   cifMPI                       INTEGER (1..4) OPTIONAL,
   tempSpatTradeOffCap          BOOLEAN,
   maxBitRate                   INTEGER (1..19200),
   stillImageTransmission       BOOLEAN,
   …,
   encryptionCap                EncryptionCap OPTIONAL,
   authenticationCap            AuthenticationCap OPTIONAL,
   integrityCap                 IntegrityCap OPTIONAL
}
```


H.245 Capability Features

- CapSet is indicated by Capability Descriptor which uses
  - AlternativeCapabilitySet
    - terminal capable of operating one mode in the set, e.g., {G.711, G.728, G.722}
  - SimultaneousCapabilitySet
    - grouped by AlternativeCapabilitySet structures
    - terminal capable of operating a set of modes simultaneously, e.g., {{H.261, H.263}, {G.711, G.728}} and {{H.262}, {G.711}}

H.245 Capability Features (Cont.)

- Multiple video/audio streams CapSet
  - {{H.261}, {H.261, H.263}, {G.728}, {G.722, G.728}}
- Receive Capability
  - inform transmitter receivers' ability to process incoming data.
- Transmit Capability
  - offer receivers a choice of possible modes of operations.

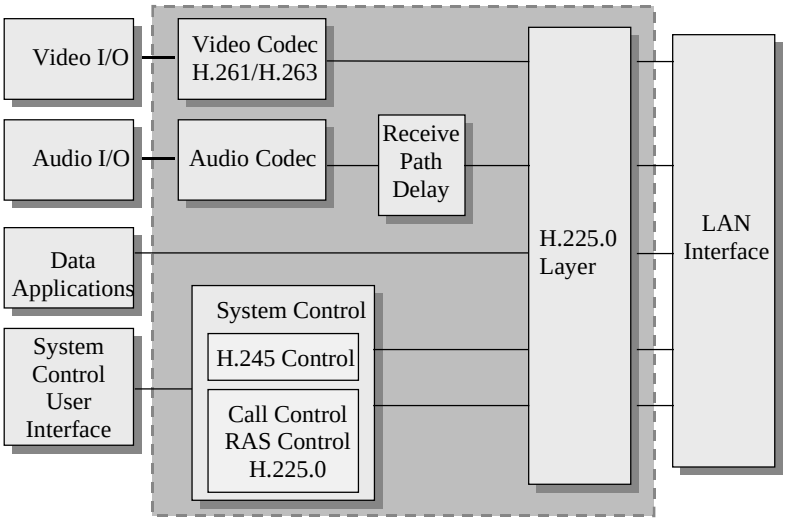
Extension of H.324

- ISDN (Annex D) ← **H.324/I**
  - **H.Mutilink** is developed for channel aggregation.
  - **H.Dispatch** is developed to
    - determine network end-to-end connectivity
    - automatically negotiate a selected mode among H.324/I, H.324, H.320, and voiceband mode
  - Support of H.320 of an H.324/I terminal is mandatory for an interoperability reason.
- Wireless Mobile (Annex C) ← **H.324/M**
  - H.223/A

H.323 Suite

- First determined in 1996
- Targeting at packet-switched network
  - LAN, enterprise area network, Metropolitan Area Network (MAN), intranet and Internet
  - Also works on PPP connection over PSTN and ISDN
- Non Guaranteed QoS networks

H.323 Structure

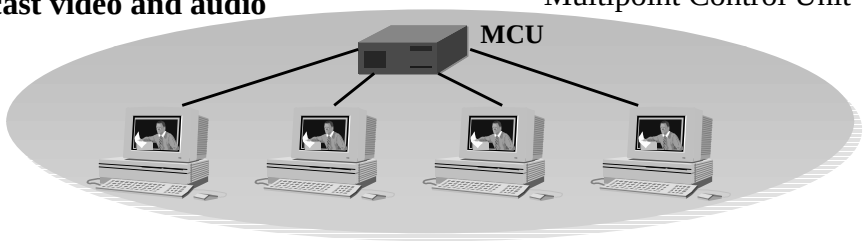


RAS: Registration, Admission, and Status

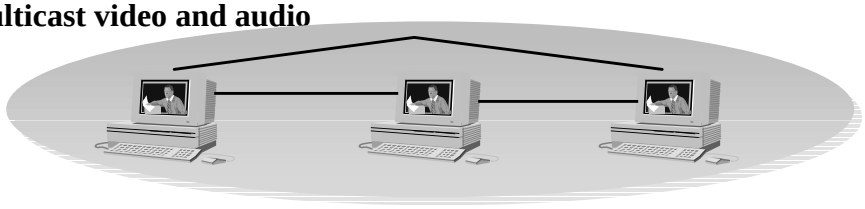
Multipoint Conference Types

Centralized unicast video and audio

Multipoint Control Unit

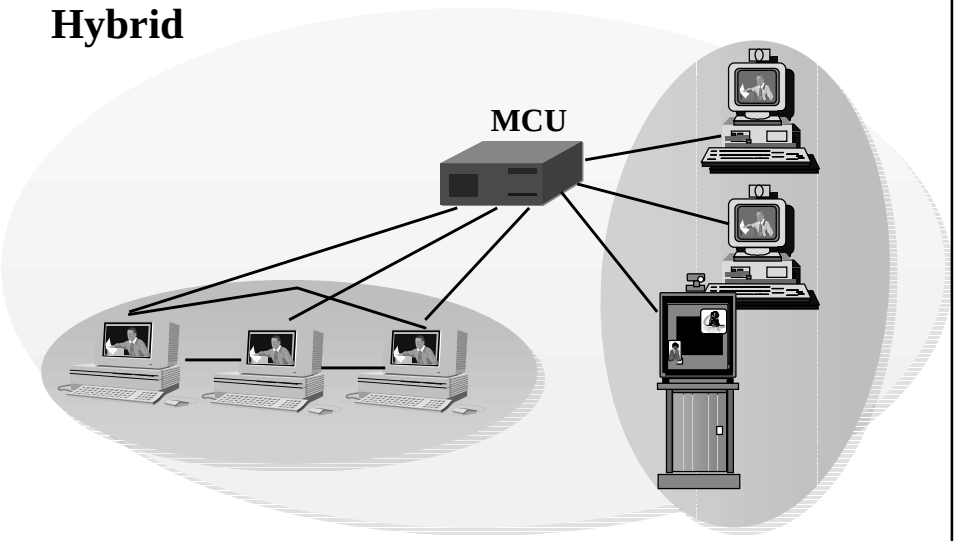


Decentralized multicast video and audio



Multipoint Conference Type (cont.)

Hybrid

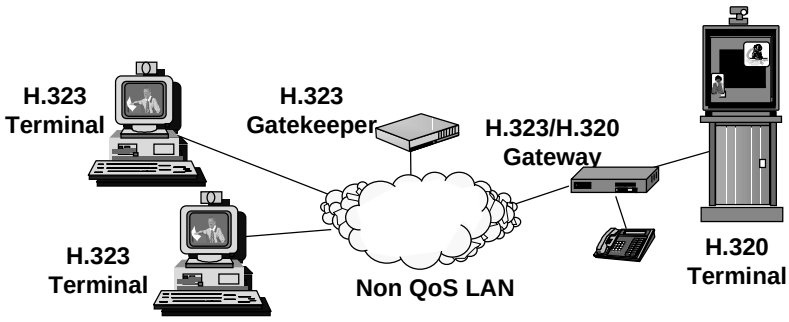


H.323 MCU

- MCU is divided into two parts: multipoint processor (MP) and multipoint controller (MC)
- MP
  - Provides audio mixing, video mixing, and video/audio switching
  - Not needed for decentralized multipoint conferencing
- MC
  - Provides conference control, establishment of common communications mode, and of media channels
  - Can be embedded into a terminal, a gateway, or a gatekeeper

H.323 Gatekeeper

- Performs registration, admissions control, calling bandwidth control for a shared network
- Provides address translation (alias address to transport address)



H.323 Gateway

- Provide interoperability with other ITU-T terminals on circuit-switched networks (H.320, H.324, H.310, H.321)
- Provide translation of call signaling, control channel messages, and media multiplexing, e.g.,
  - Call signaling: Q.931 on ISDN to H.225.0 on LAN
  - Control: H.242 on ISDN to H.245 on LAN
  - Multiplex: H.221 on ISDN to H.225.0 on LAN
- Optionally provide video/audio transcoding, e.g.,
  - G.723.1 on H.324 to G.711 on H.323

H.225.0 Protocol Stack

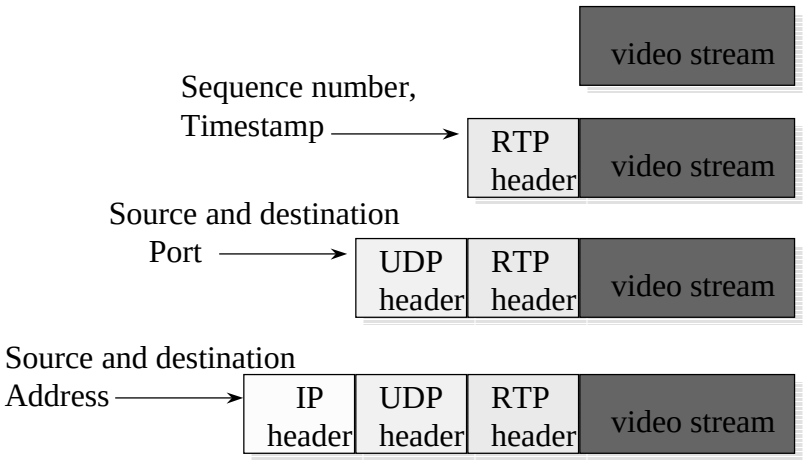
<table>
<tr><td>Audio app.</td><td>Video app.</td><td colspan="4">Terminal control and management</td><td>Data app.</td></tr>
<tr><td>G-series</td><td>H.261/ H.263</td><td rowspan="2">RTCP</td><td rowspan="2">H.225.0 RAS Channel</td><td rowspan="2">H.225.0 Call Sig. Channel</td><td rowspan="2">H.245 Control Channel</td><td>T.124</td></tr>
<tr><td colspan="2">RTP</td><td>T.125</td></tr>
<tr><td colspan="4">Unreliable transport (UDP)</td><td colspan="2">Reliable transport (TCP)</td><td rowspan="4">T.123</td></tr>
<tr><td colspan="6">Network layer (IP)</td></tr>
<tr><td colspan="6">Link layer (IEEE 802.3)</td></tr>
<tr><td colspan="6">Physical layer (IEEE 802.3)</td></tr>
</table>

IEEE 802.3: Ethernet

H.225.0 Media Transport

- Data of each logical channel is transmitted to a separate transport address
- No need to perform MUX/DMUX at the H.225.0 layer
- Audio and video information is formatted into packets based on IETF RTP (real-time protocol)
- Audio, video, RTCP, and RAS are transmitted using unreliable transport (UDP)
- Call signaling, control signal, data are transmitted using reliable transport (TCP)

Data Encapsulation Example



Packet Loss

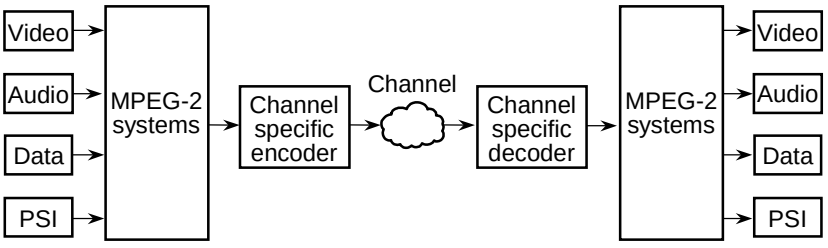
- Packet loss can happen due to
  - Packet collision on shared-LAN (Ethernet)
  - Late arrival of packets (too many hops)
  - Buffer overflow in routers
  - Noise in transmission link

References

- H.320: ftp://standard.pictel.com/avc-site/DraftRec
- H.324: ftp://standard.pictel.com/h324-site/
- H.323, H.310, H.32: ftp://standard.pictel.com/avc-site/9801_Gen/
- Special issue on Multimedia Modem, IEEE Communication Magazine, December 1996
- R. Schaphorst, Videoconferencing and Videotelephony: Technology and Standards, Artech House, 1996

MPEG System

MPEG-2 System



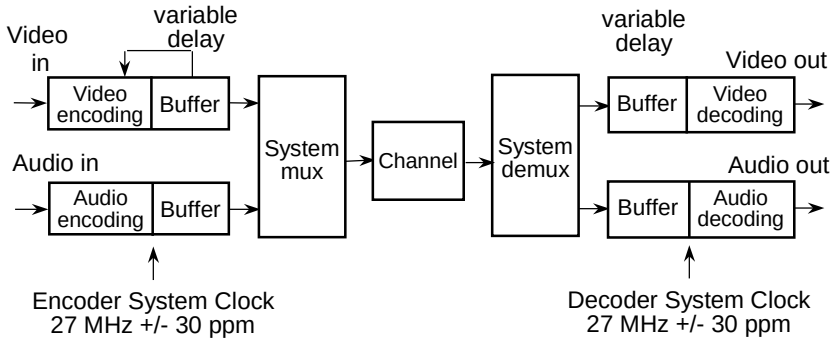
PSI: Program Specific Information

Purposes

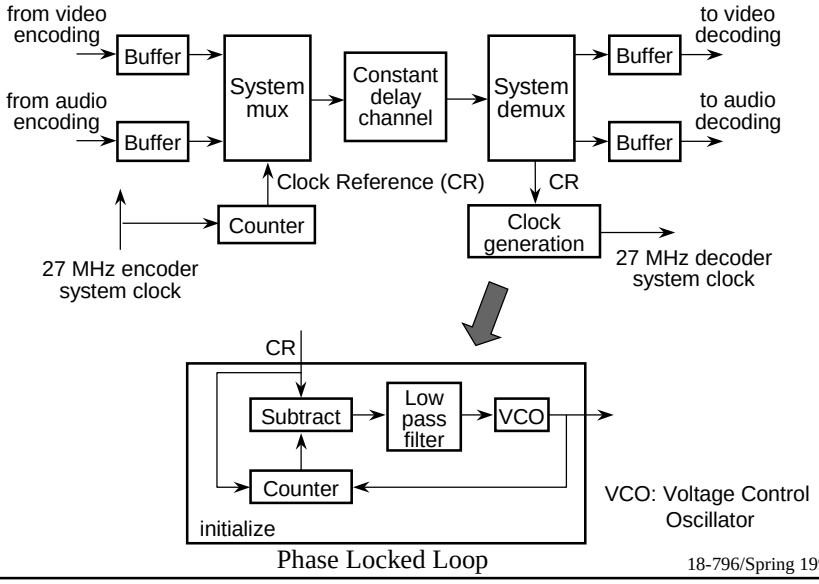
- Providing time stamps
  - Timebase recovery, buffer initialization and management, synchronized playback of audio and video
- Multiplexing of Elementary Streams
  - Audio, video, data, control, ...
- Program Specific Information
  - For the decoder to find easily and rapidly the data for a desired program
- Others
  - Error resilience, conditional access, random access, digital storage control, ...

Timing Problems

- Timebase recovery
- Receiver Buffer overflow/underflow
- Synchronized playback of audio and video

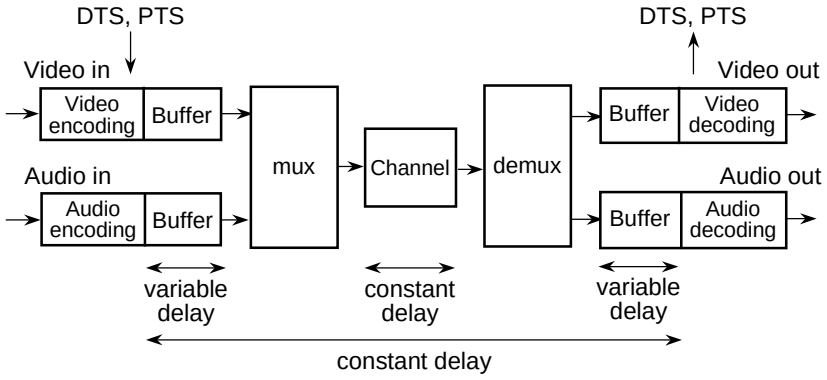


Timebase Recovery Using PCR



Phase Locked Loop

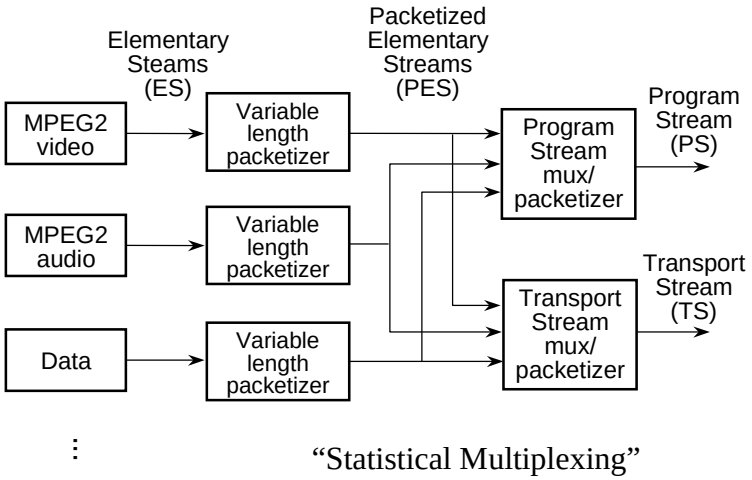
Buffer Management And Lip Sync



DTS: Decoding TimeStamp

PTS: Presentation TimeStamp

Multiplexing of Elementary Streams

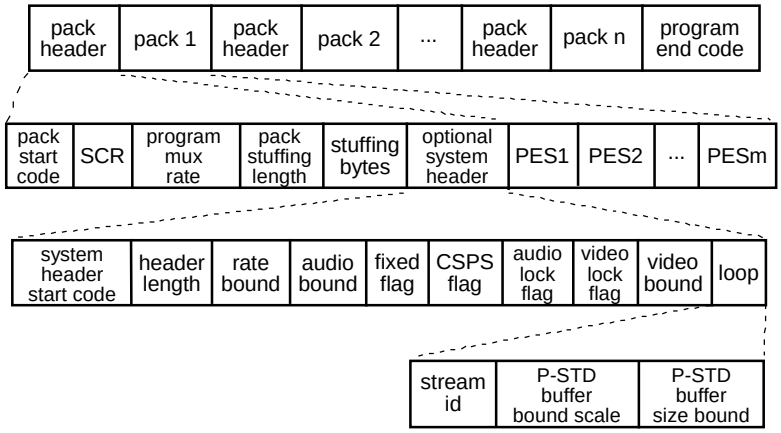


"Statistical Multiplexing"

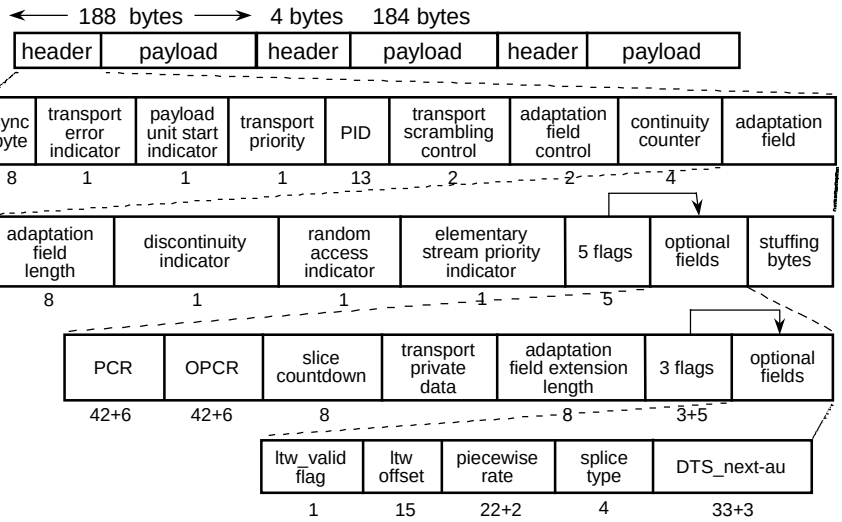
MPEG-2 Systems Streams

- Program Streams (PS):
  - Similar to MPEG-1 Systems
  - For error free environment
  - Relatively long variable-length packets (e.g., 2,000 bytes)
  - A common timebase for all elementary streams
- Transport Steams (TS):
  - For error prone environment
  - Short packets (188 bytes)
  - May contain multiple programs with independent timebases
- PS can be converted to TS, and vice versa

Program Stream Syntax



Transport Stream Syntax



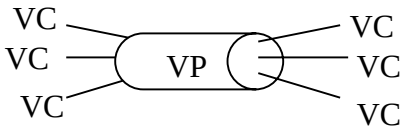
References

- Barry G. Haskell, Atul Puri, Arun N. Netravali, Chap. 3, *Digital Video : An Introduction to MPEG-2*, Chapman & Hall, New York, NY

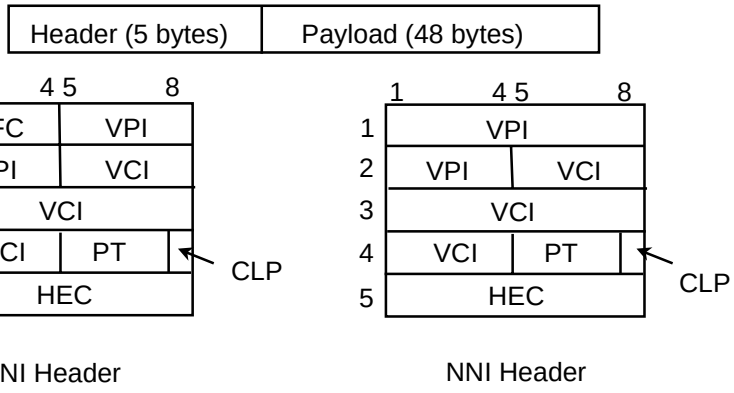
Multimedia over ATM

ATM

- Asynchronous Transfer Mode
- International standard for B-ISDN
- Supports Guaranteed QoS, bandwidth on demand
- High-speed switching
- 5-octet header (ATM layer) and 48-octet payload
- Connection specified by VCI (Virtual Channel Identifier) and VPI (Virtual Path Identifier)



ATM Cell Format



UNI : User-to-Network Interface
NNI : Network-to-Network Interface
GFC : General Flow Control
VPI : Virtual Path Identifier

VCI : Virtual Channel Identifier
PT : Payload Type
CLP : Cell Loss Priority
HEC : Header Error Control

ATM Adaptation Layer (AAL)

- AAL acts as the interface between user applications and the ATM layer
- Different types of AALs support different types of services
  - AAL1 for constant bit-rate circuit emulation
  - AAL2 for very low bitrate delay sensitive applications
  - AAL3/4 for connection-oriented data services
  - AAL5 for connectionless data services

ATM for Multimedia Services

Service Type	Traffic Descriptors	QoS Parameters
CBR (Constant Bit Rate)	PCR	MCTD, PCDV, CLR
VBR (Variable Bit Rate)	PCR, SCR, MBS	MCTD, PCDV, CLR
ABR (Available Bit Rate)	PCR, MCR	CLR
UBR (Unspecified Bit Rate) (best effort service)	PCR	-

PCR: Peak Cell Rate
SCR: Sustainable Cell Rate
MBS: Max. Burst Size
MCR: Min. Cell Rate

MCTD: Max. Cell Transfer Delay
PCDV: Peak-to-peak Cell Delay Variation
CLR: Cell Loss Ratio

Multimedia Over ATM vs. IP

ATM	IP
• Multimedia data into cells	• Packetization
• Guaranteed QoS	• Non QoS
• VBR coding	• VBR coding
• Scalable coding	• Scalable coding
• Cell-loss concealment	• Packet-loss concealment
• Choice of AAL	• Very low bit-rate coding
• Effect of delay variation	• Low delay video coding
• Statistical multiplexing	• Software codec

References

– Michael Orzessek and Peter Sommer, *ATM & MPEG-2: Integrating Digital Video Into Broadband Networks*, Prentice Hall, 1998