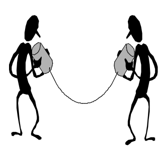
IEEE INFOCOM 2001

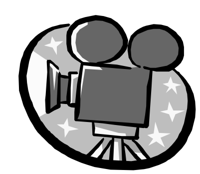
Tutorial T5

Video Over IP

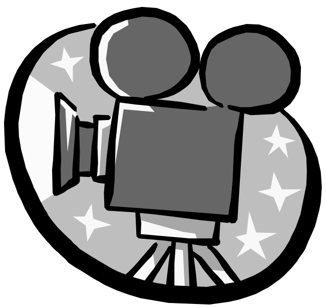
Magda El-Zarki (University of California at Irvine)

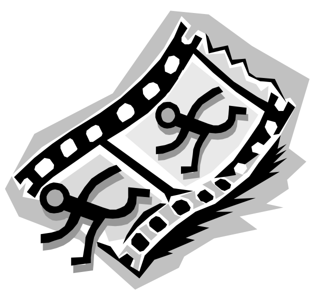
Monday, 23 April, 2001 - Morning

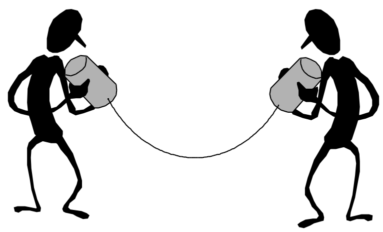




MPEG-4 over IP - Part 3



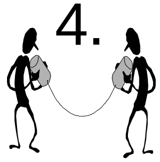


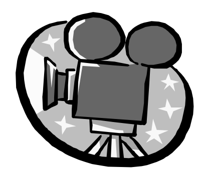


Magda El Zarki
Dept. of ICS
UC, Irvine
elzarki@uci.edu

Outline of Tutorial

1. Part 1:
    1. Overview of Video Compression
    2. The MPEG suite
    3. Video Quality
2. Part 2:
    1. MPEG-4
3. Part 3:
    1. MPEG-4 Delivery over IP
4. Conclusions



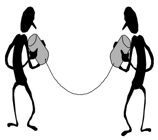


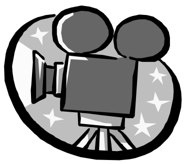
MPEG-4 Delivery over Internet

- TCP/IP
- SL-packet over UDP
- RTP/UDP/IP
- Signaling
- MPEG-4 Delivery Architecture

TCP/IP

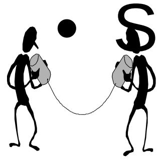
- Not suitable for real-time
  - Retransmissions can lead to high delay and cause delay jitter
- Does not support multicast
- Congestion control mechanism (slow start) not suitable for AV media

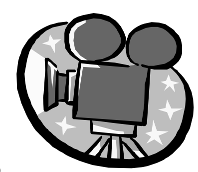




SL-packet over UDP/IP

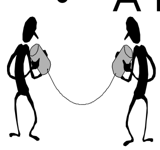
- SL provides:
  - Timing
  - Sequence numbering
- UDP provides:
  - Multiplexing
  - Length field
  - Checksum service
- SL+UDP as transport protocol?

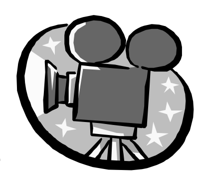




Some of the Problems when using SL/UDP/IP

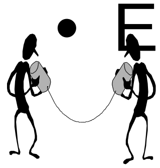
- No other media stream can be synchronized with MPEG-4 data carried directly over UDP
- The dynamic scene and session control concepts cannot be extended to non-MPEG-4 data
- No defined technique for synchronizing MPEG-4 streams from different servers
- Streams from different servers may collide (their sources may become unresolvable at the destination) in a multicast session
- Mechanisms need to be defined to protect sensitive MPEG-4 data (RTP supports FEC)
- A feedback channel must be defined for quality control

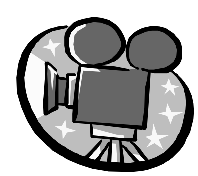




RTP & RTCP

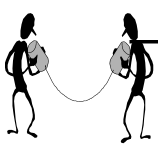
- Real time Transport Protocol
- Real time Transport Control Protocol
- A session consists of an RTP/RTCP pair of channels
- Usually works over UDP/IP
  - Can work with other protocols
- End-to-end protocol

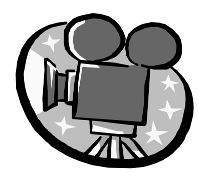




RTP Supports:

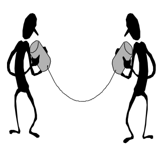
- Multicasting
- Payload type identification
- Time stamping
- Sequence numbering
- Delivery monitoring
- Underlying UDP supports:
  - Multiplexing
  - Length field
  - Checksum service

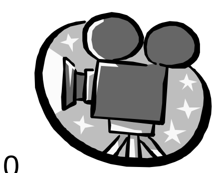




RTP Problems:

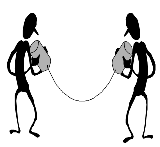
- It does not support the timely delivery of data or any other QoS guarantees
  – In-time delivery requires lower layers that have control over resources in switches and routers (e.g. RSVP)
- It does not guarantee delivery, so packets may be delivered out of order or get lost
  – No mechanism to recover from packet loss

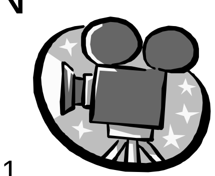




RTP Time Stamp & Sequence Number

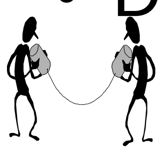
- Time Stamp (TS)
  - Place incoming packet in correct timing order
    - Initial values are picked randomly and independently for each RTP stream
    - Increase in time indicated by each packet
- Sequence Number (SN)
  - Detect packet loss
  - Increase by one for each packet
- For a video frame that is split into multiple RTP packets: they share same TS but use different SN

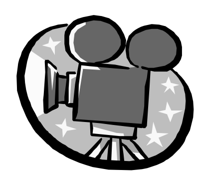




RTP characteristics

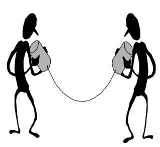
- Only can map one source stream onto one RTP session
  – Multiplexing causes problems
  – For scalable coding, each layer must have its own RTP session and multicast group
- Within each RTP session, source can change its data format over time
- Does allow for FEC

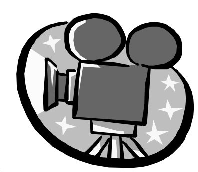




RTCP

- Synchronize across different media streams
  - NTP timestamp in RTCP Sender Report
- Provide feedback on the quality of data using lost packet counts
- Identify and keep track of participants



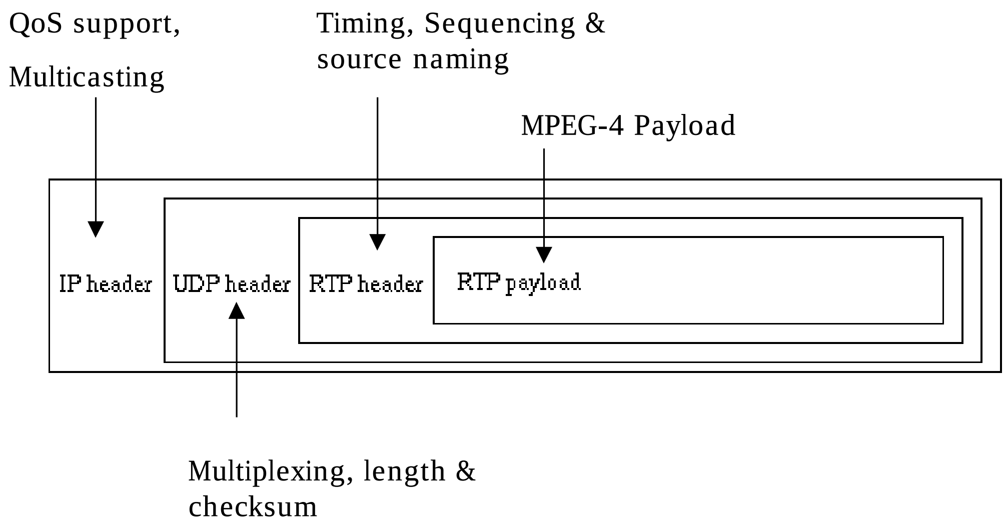


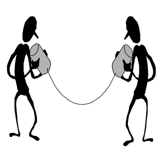
MPEG-4 over RTP/UDP/IP: direct wrap

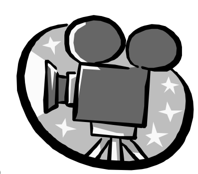
- Easiest is to wrap the MPEG-4 SL packet in an RTP packet
  - High overhead: two full headers
  - RTCP may not provide enough control for the MPEG-4 stream

MPEG-4 over RTP/UDP/IP: Payload per ES

- Several types of MPEG-4 payloads are being defined by the IETF for different ESs

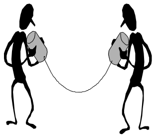


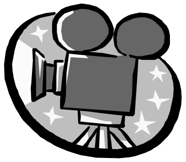




Elementary Stream payload

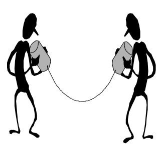
- Can transmit an audio/video ES without SL & FlexMux headers
- Mapping of ES_ID onto SSRC - not custom
- New RTP payload needs to be defined for every ES

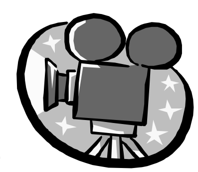




RTP ES payload restrictions

- RTP packetization is media-aware
  - No unique scheme can be defined, need a payload definition per payload type, not practical.
  - This may require the definition of new session and scene description mechanisms to deal with all the flows.
- Common restrictions
  - RTP timestamp corresponds to composition time stamp (CTS) of MPEG-4 stream
  - Packets should be sent in decoding order.
  - Streams can be synchronized using RTCP



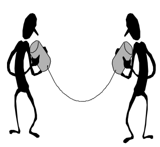


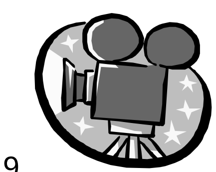
RTP SL Stream Payload

- Map SL stream onto RTP session
  - SL header is optional
- Reduced SL header does not include:
  - Sequence number (mapped to RTP header)
  - Composition Time Stamp (mapped to RTP header)
- Single SL packet map into single RTP packet
  - RTP packet SHOULD be no larger than the path-MTU
  - If it is, then fragmentation at a lower layer will take place that could cause problems

Multiple Streams in MPEG-4

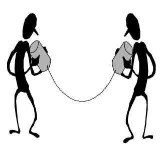
- Port consuming
  - Each AVO contains one or more streams
  - Each stream needs one RTP session
- Potential Solution
  - Selective bundling of Ess - FlexMux -> define a multiplexed MPEG-4 payload
    - May have to be defined for multiple payload types
  - Generic RTP multiplexing for use with MPEG-4, under development by IETF (mixers)

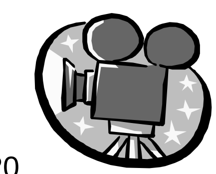




RTP Generic Multiplexing Scheme

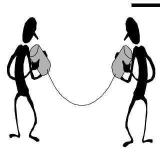
- Must take the following into consideration:
  – ESs multiplexed in one stream can change frequently during a session -> the coding type, individual packet size, and temporal relationships between the multiplexed data units must be handled dynamically
  – Must have a mechanism to determine the ES identifier (ES_ID) for each multiplexed stream. Note that the ES_ID is not part of the SL header.
  – An SL packet does not generally contain any size info, when multiplexing several packets into one payload, you must be able to delineate them.

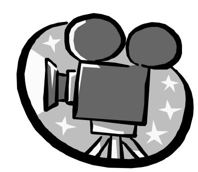




Header compression for low speed links

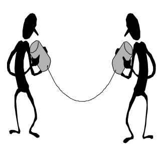
- For low bandwidth links, large packets tend to be segmented into several small ones to reduce delay
- Header overhead for one RTP session
  - RTP+UDP+IP>=40bytes
- Header compression
  - Link-by-link basis
  - 2 bytes without UDP checksum
  - 4 bytes with UDP checksum

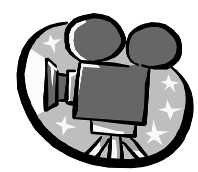




MPEG-4 Media Control

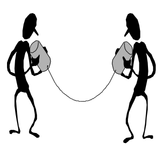
- Remote interactivity: add or delete a stream, modify C/Cs of a stream, etc.
- Media control channel allows renegotiating in time the available network and processing resources.
- Must have an efficient signaling protocol that can handle such messages

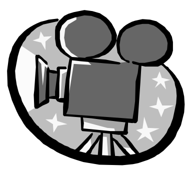




Media Control Framework

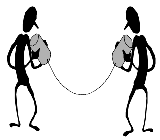
- To allow a client and one or more servers to exchange different types of control messages and also allow for peer to peer exchange between 2 or more clients, the framework requires several components:
  - A description of the stored or live presentation
  - A set of protocols that can provide proper services for the back channel message delivery
  - A set of protocols that can allocate resources for the involved hosts and networks

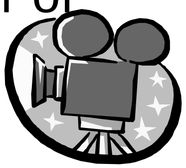




Three components to media control (1)

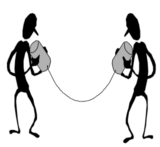
- Presentation Description:
  - The client needs to refer to the description of a presentation that expresses the temporal and static properties of the presentation itself
  - Must include information about the media, location of the media, etc.
  - Should provide multiple description instances of the same presentation so that the client can specify a given combination that fits its needs/capabilities - the client is the orchestrator of the presentation and the server is participating

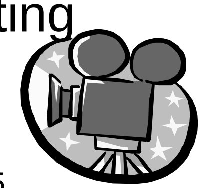




Three components to media control (2)

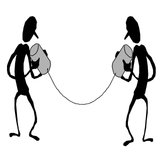
- Client and Server State Representations:
  – Out of band signaling is used: the data streams and the control information are carried over separate channels using different protocols, each best suited to their needs and modes of operation
  – As the media streams may be modified by the end user, the server needs to a state of the stream(s) status for each client it is serving
  – The client has to keep track of all the participating streams

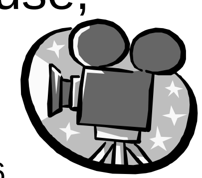




Three components to media control (3)

- Basic Media Control Messages:
  - A multimedia system should have access to control messages ranging from remote VCR functions such as stop, play, fast forward and fast reverse, to messages generated in response to user actions to modify the presentation of a given object stream such as add or remove or alter, etc.
  - The basic control functionality relates to presentation and stream set-up; play, stop, pause, teardown and recording



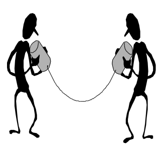


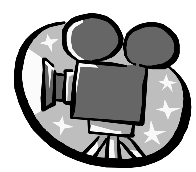
Real Time Streaming Protocol (RTSP)

- RFC 2326
- It is an application level protocol that provides an extensible framework to enable controlled delivery of real-time data, such as audio & video.
- Sources can include both live and stored content
- It can be transported over UDP, TCP and is designed to work with RTP and HTTP.
- Provides support for bidirectional communications

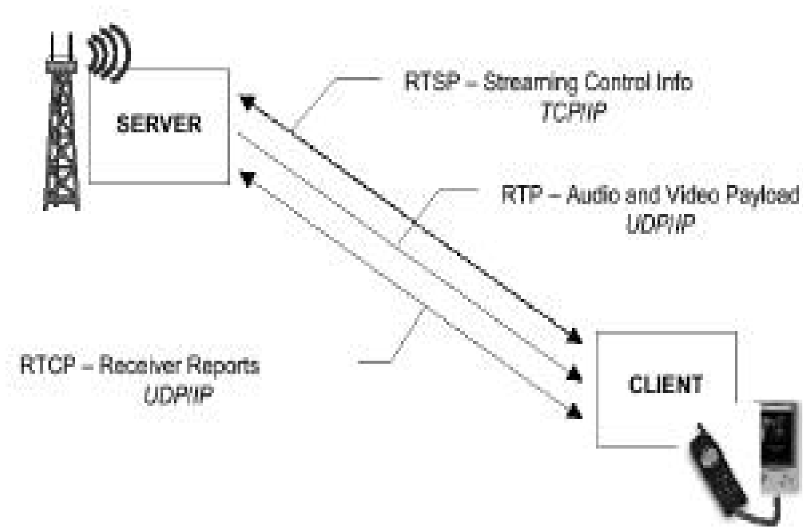
What RTSP can do:

- Allows:
  - retrieval of media from a given server,
  - Invitation of a media server to a conference
  - Addition of media to a given presentation
  - Negotiation of transport method
  - Negotiation of capabilities of the server
- It is transport independent and multiserver capable
- It provides frame level timing accuracy to allow for remote video editing



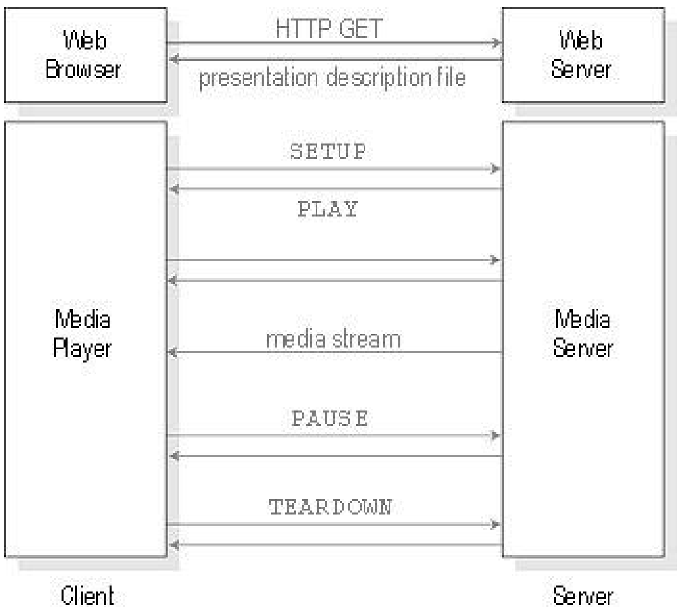


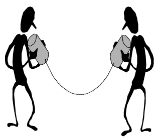
Using RTSP for application control

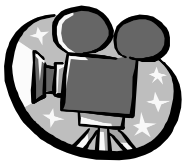


Application Control Commands	Audio Data, Video Data, Receiver Reports
RTSP	RTP/RTCP
TCP	UDP
IP	
Radio Link/Data Link	
Physical Layer	

An RTSP Example

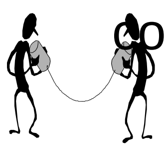


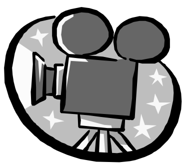




MEPG-4 and RTSP

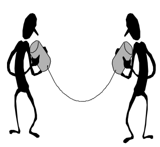
- From DMIF’s perspective, RTSP is an application alongside MPEG-4 systems
- The RTSP client and server interact with the MPEG-4 systems
- The RTSP client and server control the streams through the DAI by an RTSP-DMIF interface
- The interface is kept very simple by limiting it to field mapping between the RTSP fields and the DAI primitive parameters.
- The RTSP client server interactions are used to control the MPEG-4 elementary streams.

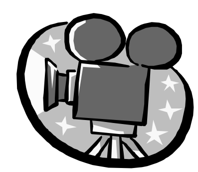




What RTSP Does and Does not do

- RTSP does:
  - Control the transmission of media stream
  - Use out-of-band signaling
- RTSP does not:
  - Define compression schemes
  - Define how AV is encapsulated
  - Define how to buffer

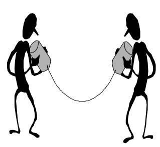


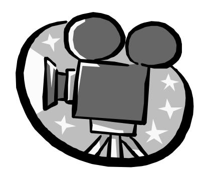


Can We Use RTSP?

- Normally one stream per session
- Can aggregate control for multiple streams
- Cannot provide different controls for different streams in one session

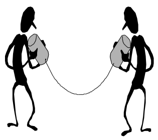
------> in other words it does not have all the functionality that is needed for MPEG-4 interactivity

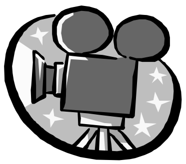




IETF Family of Session Protocols

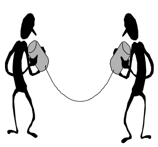
- Session Description Protocol (SDP)
- Session Announcement Protocol (SAP)
- Session Initiation Protocol (SIP)

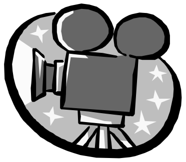




IETF Protocols for Setting Up Sessions

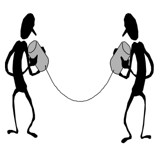
- Two ways: announce / invite
  - SAP → Announce
  - SIP → Invite
- Both of these protocols carry the description of the session in SDP format

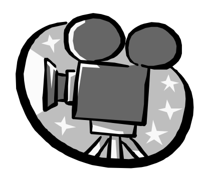




Session Description Protocol (SDP)

- RFC 2327
- Short structured textual description (name, purpose,media, protocols, codec formats, timing, transport)
- Can be carried by different transport protocols, such as SAP, SIP, RTSP, HTTP and even email using the MIME extensions



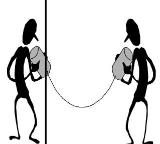


An Example of SDP Session Description

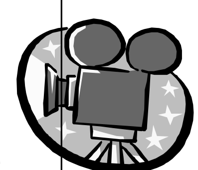
Session description
- v= (protocol version)
- o= (owner/creator and session identifier).
- s= (session name)
- i=* (session information)
- u=* (URI of description)
- e=* (email address)
- p=* (phone number)
- c=* (connection information - not required if included in all media)
- b=* (bandwidth information)
- z=* (time zone adjustments)
- k=* (encryption key)
- a=* (zero or more session attribute lines)

Time description
- t= (time the session is active)
- r=* (zero or more repeat times)

Media description
- m= (media name and transport address)
- i=* (media title)
- c=* (connection information - optional if included at session-level)
- b=* (bandwidth information)
- k=* (encryption key)
- a=* (zero or more media attribute lines)

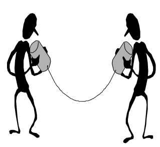


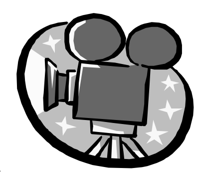
```
v=0
o=llynch 3117798252 3117798459 IN
    IP4 128.223.214.23
s=UO Presents KKNU New Country
i=MacKenzie River Broadcasting
    Company's New Country Radio
    KKNU 93.1
u=http://darkwing.uoregon.edu/~uoco
    mm/
e=UO Multicasters
    multicast@lists.uoregon.edu
p=Lucy Lynch (University of Oregon)
    (541) 346-1774
c=IN IP4 224.2.163.188/127
t=0 0
a=tool:sdr v2.4a6
a=type:broadcast
m=audio 23824 RTP/AVP 0
c=IN IP4 224.2.163.188/127
a=ptime:40
```



SDP summary

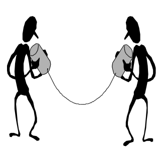
- Originated for announcing multicast sessions
- Not originally for unicast sessions, but works well enough in the context of SIP and RTSP.
- Can't negotiate content with application peers
- Widespread use

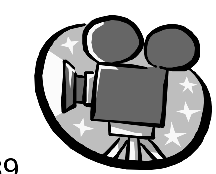




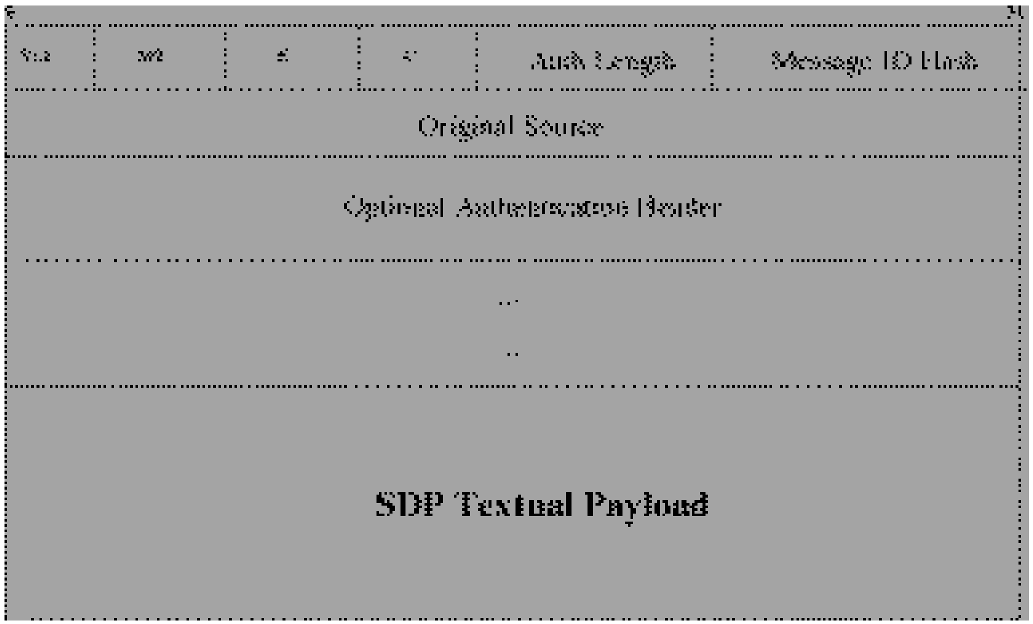
Session Announcement Protocol (SAP)

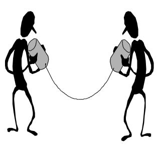
- RFC 2974
- Used to create/modify/terminate sessions
- Bootstrap mechanism using IP multicast
- Contains SDP as payload
- Takes security into consideration

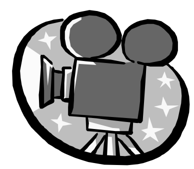




Format of SAP Packets

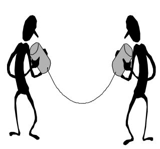


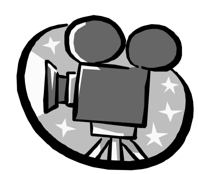




Session Initiation Protocol (SIP)

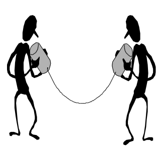
- RFC 2543
- A "control" protocol that supports:
  - Internet multimedia conferences, Internet telephony and multimedia distribution
  - Communication via multicast or a mesh of unicast relations (or a combination)
  - Negotiation
  - "User mobility'' by proxying and redirecting
- Independent of lower layer protocols
- Extensible to be application specific

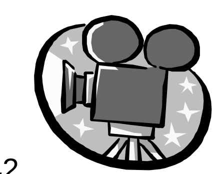




SIP Terminology

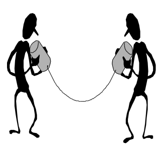
- Initiator, calling party, caller
- Invitee, invited user, called party, callee
- Invitation, provisional response, ring back
- User agent client (UAC), user agent server (UAS)
- Location server, proxy server, redirect server

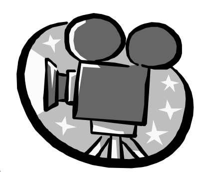




SIP Architecture

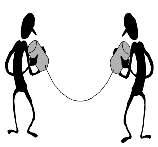
- Two basic components : the SIP user agent and the SIP network server.
- The user agent is the end system component for the call, includes two elements: UAC and UAS
- The SIP server is the network device that handles the signaling procedures, includes: location/redirect/proxy server

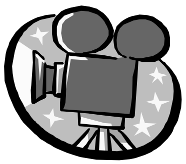




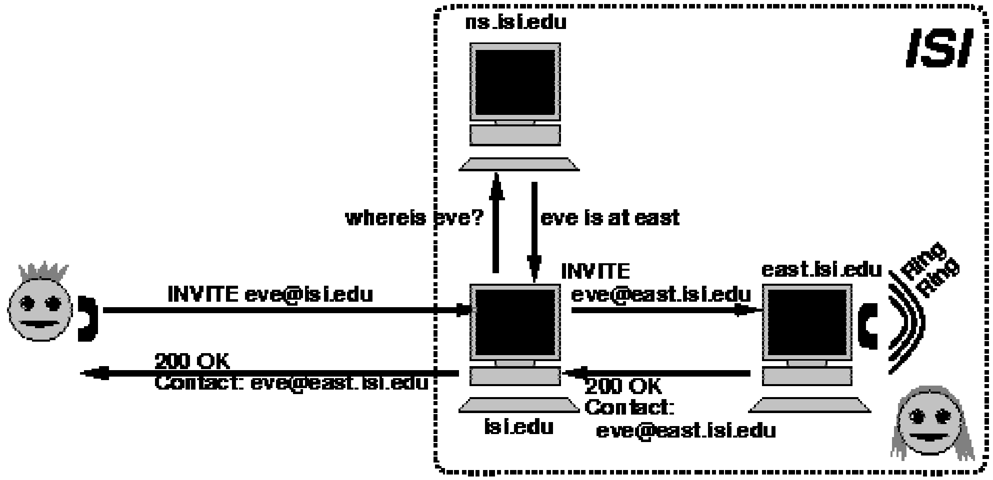
SIP Signaling

- Initiates sessions / invites members to sessions (Sessions can be advertised using SAP, email, etc.)
- Supports name mapping, redirecting and personal mobility
- Manages session: user location, user capabilities, user availability, call setup, call handling
- Modifies session

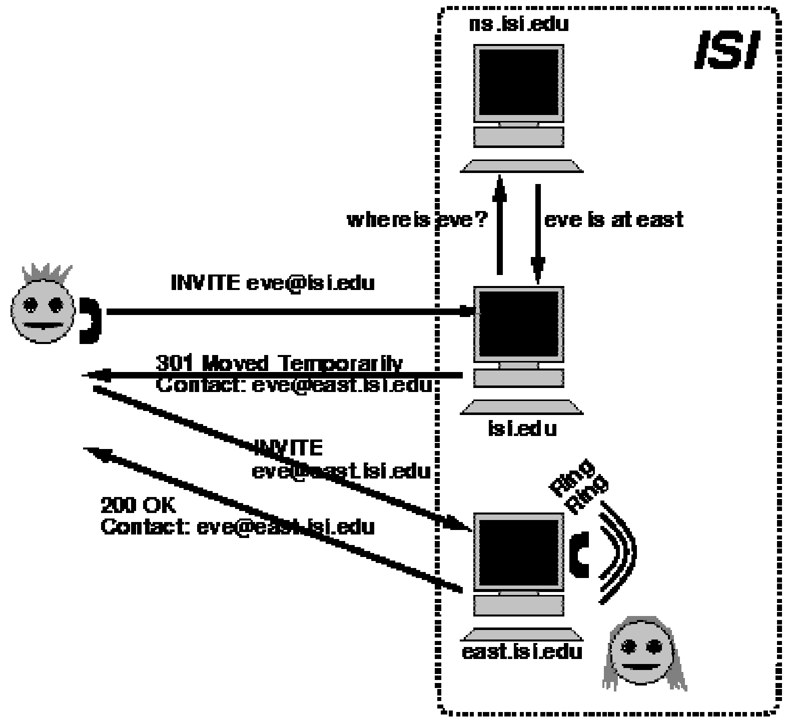




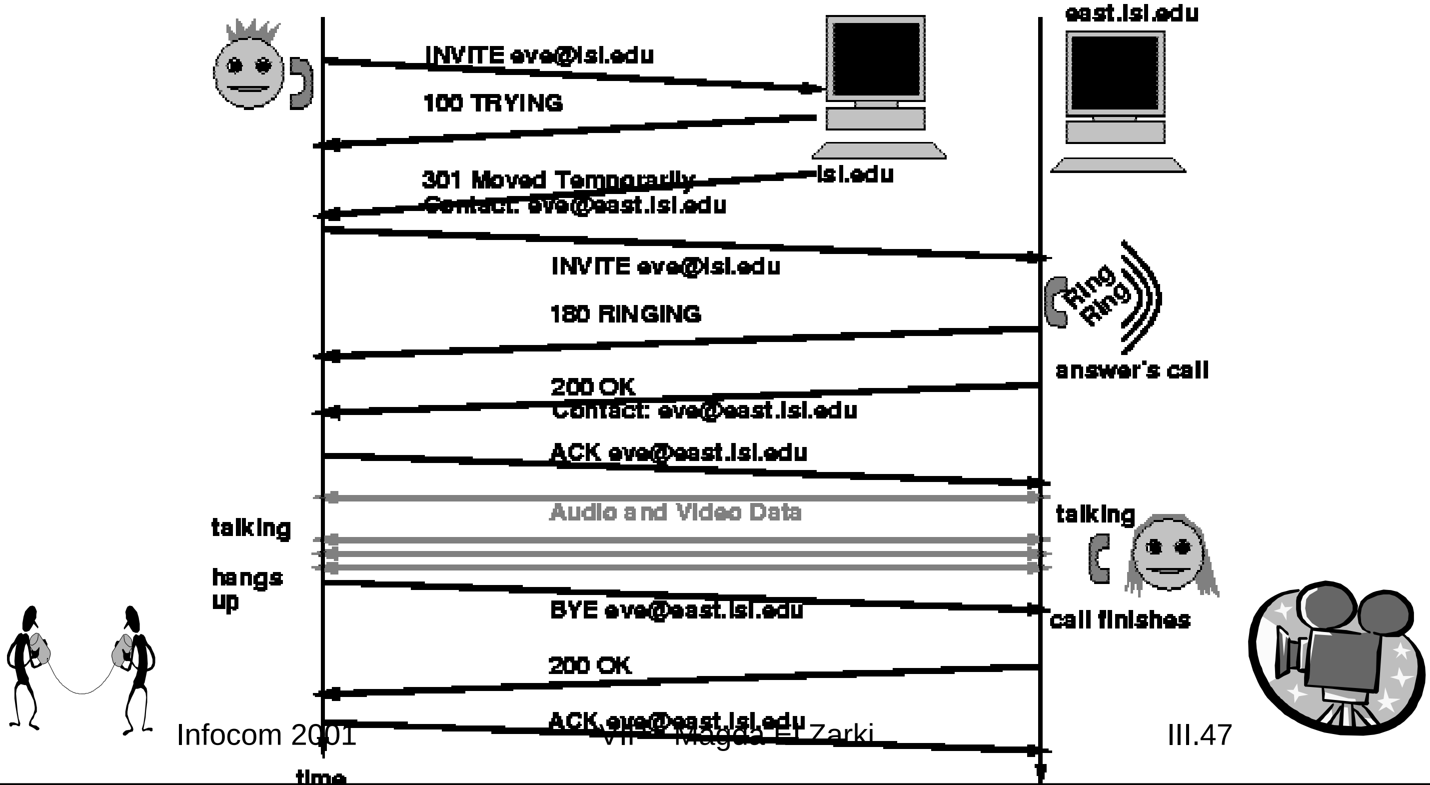
Example I (Relay)



Example II (Redirect)

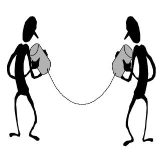


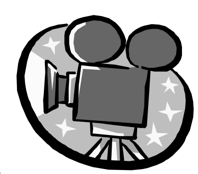
Example III (A call)



SIP Summary

- Text-based, uses MIME
- Light-weight (compared to ITU-T H.323)
- Gaining widespread use in IP telephony



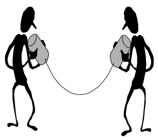


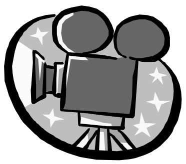
MPEG-4 and SIP

- Unique ability to control different media types within a single session → Multiple stream transmission in one network session
- User Agent model fits inwell with an MPEG-4 Client/Server model (point-to-point communication)

Resource ReServation Protocol (RSVP)

- RFC 2205, 2208, 2209
- Deemed as part of IntServ (RFC 2210)
- Originally Designed for providing reservation for bandwidth in multicast trees

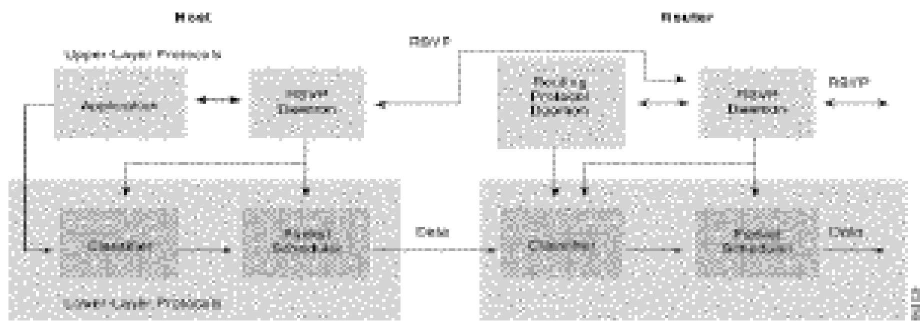




What RSVP Does and Does not do

- RSVP does:
  - exchange info.
    - Reservation information
    - Path message
- RSVP does not:
  - specify how the network provides the reserved bandwidth.
  - do routing

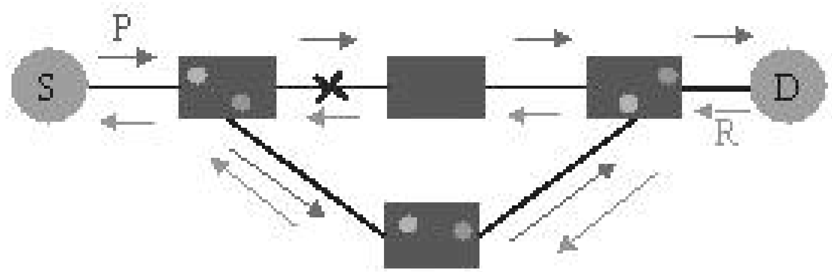
How RSVP Works

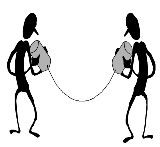


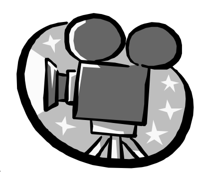
- RSVP daemon communicates with two local decision modules, ***admission control*** and ***policy control*** to decide whether enough resource available and whether a reservation can be made
- If succeed, RSVP daemon sets parameters in ***packet classifier*** and ***packet scheduler*** to obtain the desired QoS.

RSVP Features

- Soft-state (timer-associated), needs to be updated periodically.
- Works at transport layer, directly over IP (port # 46)
- Receiver-oriented

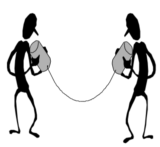


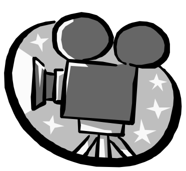




Problems with RSVP

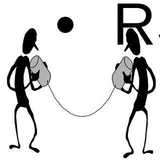
- Scalability?
  - Maintain per-flow state for ***each*** flow at core router?
- Flexible service models?
  - Need more qualitative or relative definitions of service distinctions

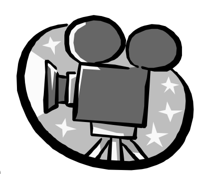




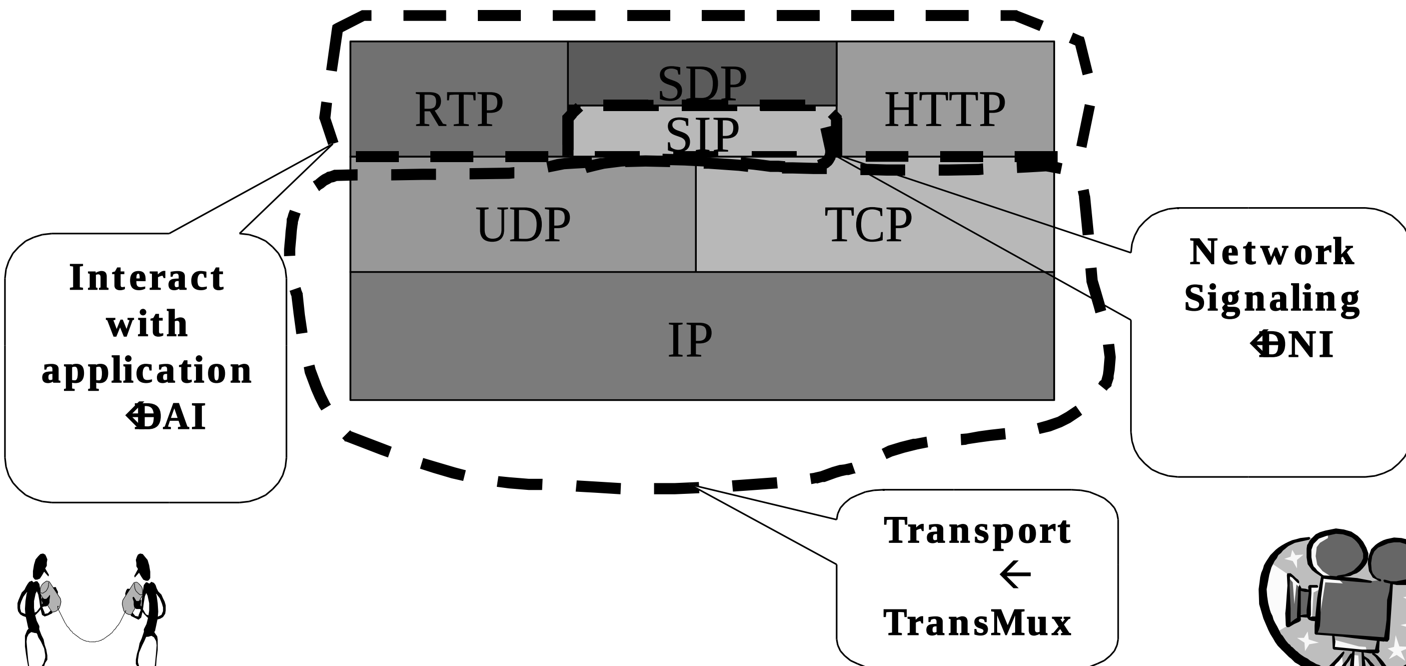
Toolboxes Available for transmitting MPEG-4 over the Internet

- RTP → ***transport*** of audio/video/... data, quality-of-service feedback
- RTSP →very simple media ***control*** of streams
- SIP → ***inviting*** people, media servers to sessions - telephony and streaming audio/video
- HTTP, SDP → ***retrieve*** media descriptions
- RSVP -> resource reservation

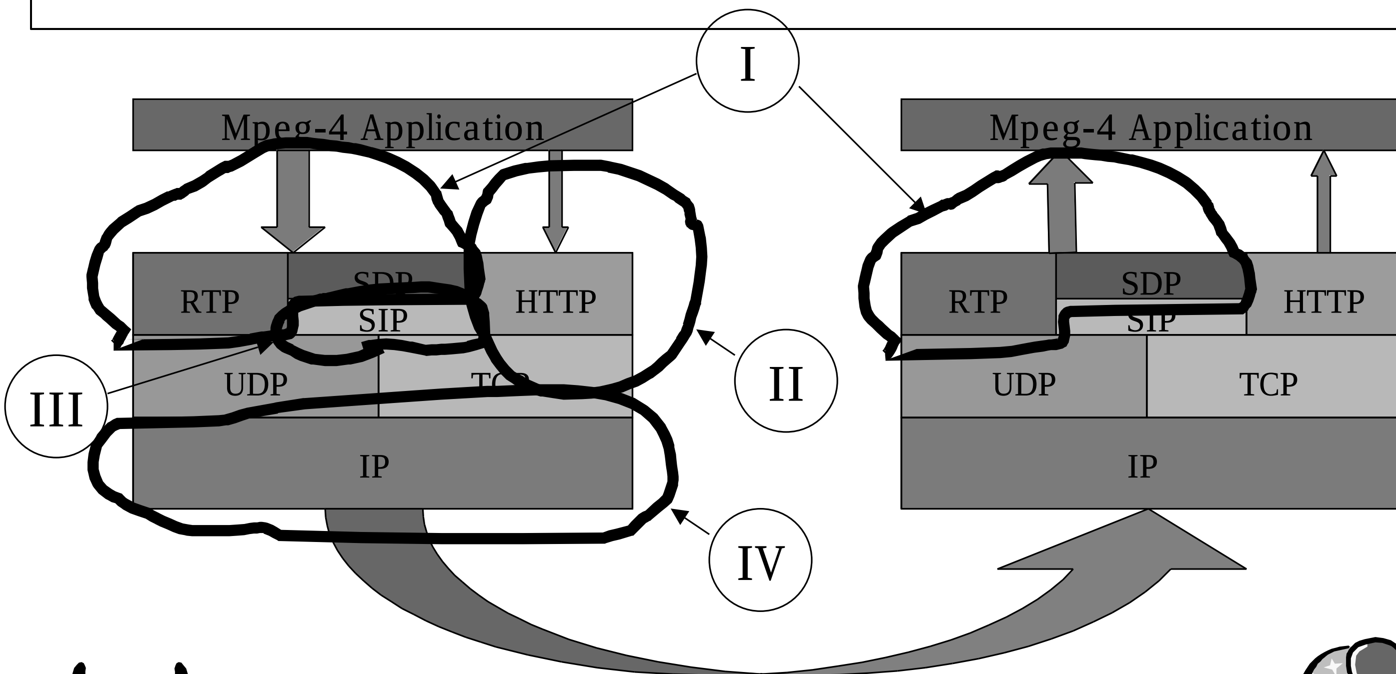




Protocol Stack for MPEG-4 over IP



Client/Server Model: Open Issues



Issues (1)

1. Encapsulation of MPEG-4 Sync layer packetized stream
   - IEC/ISO 14496-8, framework still in revision
   - Lots of issues still remain: time axis, buffer management, packet size, payload definition, SDP syntax, etc…

Issues (2)

2. Interactivity between application and End User
   - Description of MPEG-4 content
   - Initialization of an MPEG-4 session

Issues (3)

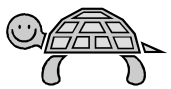
3. SIP and IPC (Inter Process communication)
   - How to describe the dynamic process of channel (stream) setup and release?
   - What control information is necessary and how to transport it? Need to take into account client/server interactivity

Issues (4)

4. Transport and IP QoS
   - Must define a mapping mechanism among the different QoS mechanisms: transport QoS (not yet available, how to define?), network QoS (provided, but is it sufficient?), CAR, etc.

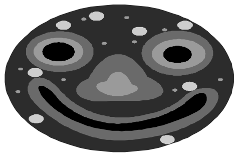
Issues (5)

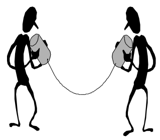
5. System Integration
   - Client/Server design
   - Coordination of and interaction between different layers (need to design specific modules)
6. And much more…

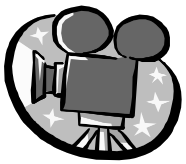


Conclusions

- MPEG-4 is here in a limited fashion
- We will see a marked growth in MPEG-4 products as chip sets become more readily available
- Simple basic MPEG-4 service is not a problem
- The extended version still requires a lot more work







Reference Text

- Multimedia Systems, Standards, and Networks, edited by Atul Puri and Tsuhan Chen, publ. Marcel Dekker, Signal Processing and Communications Series