SCOM/ SRSI
Multimedia Delivery

Ana Aguiar
DEEC, FEUP

2018-19

Contents

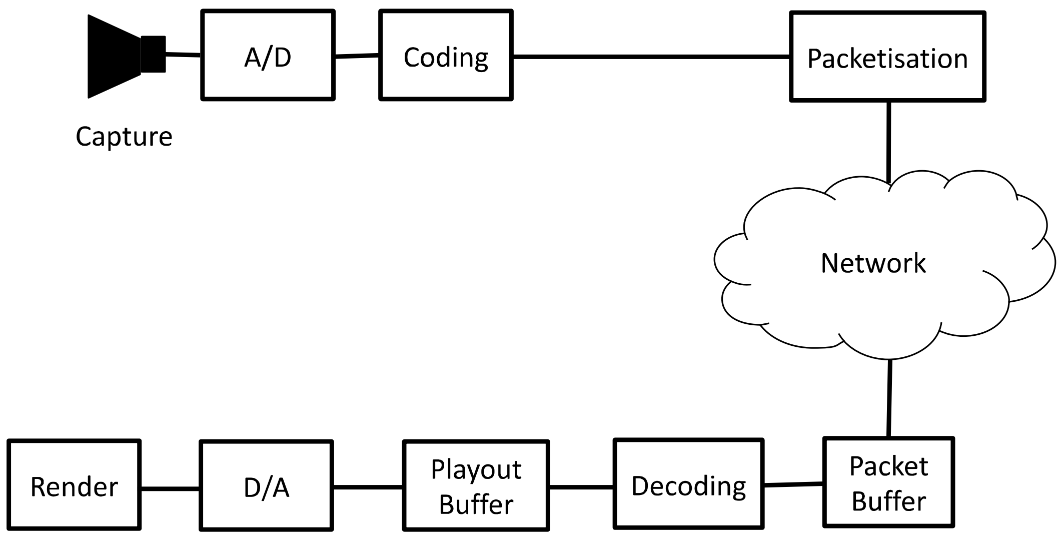
- Why is multimedia different from other content?
- Requirements of multimedia traffic
- Impact of network QoS on multimedia QoE

Multimedia = Multiple + Media

- Audio + video (what we usually call video)
- Audio + video + presentation
- Audio + presentation

- Human to human communication
  - Filtered by human perception => QoE
  - Sometimes interactive

Multimedia Flow



Audio Coding (Info only)

- Temporal redundancy in adjacent samples
  - Pulse Coded Modulation (PCM) and variants thereof, like DPCM, ADPCM
  - codecs: ITU-T G.711, G.721, G.723, G.729 (used in UMTS)
- Perceptual redundancy: some sounds are more perceived than others, after some sounds the ear cannot perceive others => psychoacoustic models
  - MPEG type of coding
  - codec: MP3

Audio Coding (Info only)

- Additional mechanisms for low bandwidth
  - VAD: voice activity detector
  - DTX: discontinuous transmission
  - When no voice is detected, different DTX mechanisms can be used:
    - No transmission
    - Comfort noise, according to RFC 3389 or codec built-in

Image and Video Redundancy (Info only)

The image part with relationship ID rId5 was not found in the file.

Images have 2 types of redundancy

- Spectral: redundancy in the colour components
- Spatial: redundancy in adjacent pixels
- Statistical: after spectral and spatial encoding, more non-zero coefficients exist close to 0, giving room for statistical encoding

Video has 1 additional dimension

- Temporal: redundancy in adjacent frames

MPEG Video Coding

- Video is a sequence of still images (frames)
- Consecutive images contain similar information
- MPEG uses temporal redundancy
  - I-frame: only intra-picture encoding
  - P-frame: predicted frame, difference from previous I-frames
  - B-frame: bi-directional predicted frame, interpolation between previous and next I or P-frames

MPEG Video Stream Syntax

The image part with relationship ID rId3 was not found in the file.

MPEG Predictive Encoding (Info only)

- Macroblock in B (or P) frame
  - Coordinate in picture
  - Motion vector relative to previous frame
  - Motion vector relative to next frame
  - Difference from reference frames for each pixel (encoded as pixels in I-frame)

MPEG Video Parameters (Info only)

Implications

- Large GOP: higher compression, but errors take more time to be corrected
- Small GOP: errors are corrected quickly, but less compression, so higher bandwidth needs or lower picture resolution

User Experience Guidelines for Real-Time Services

- Frame rate vs Resolution

Consequences of Video Syntax

- Different frames have different importance
  - Different packets too
- One lost frame can imply several other frames lost
  - E.g. lost I-frame means loss of entire GOP
- Errors in a frame produce errors in dependent frames
  - Error propagation
- Frame transmission order differs from play order

  Example MPEG GOP, P-frame must arrive before the previous B frames that depend on it to be decoded

Multimedia Requirements

- Delay dependencies between sender and receiver
- Delay requirements depend on type of media
  - E.g. for VoIP http://www.voip-info.org/wiki/view/QoS
  - Not so unanimous for video
- Time correlation between multiple media
- Interactivity with stream source
- Source and receiver must be able to adapt to application-specific details

Synchronisation between Streams

<Video delay> - <Audio delay>

[Technical Committee on Human Factors - Guidelines for real-time communication services]

Multimedia Support

Multimedia Traffic

- Highly variable traffic characteristics and QoS requirements
- Strict timing requirements (QoS, interactivity, synch)
- Traffic requires specific agreement between sender and receiver

Network

- Varying throughput
- No guarantees
- Delay variations
- Traffic agnostic

ENABLE and ADAPT!

Playout Buffer

- Buffer on the receiver side where frames are stored after re-assembly and before being played
    - Turns variable into constant delay
    - Delay < playout delay + propagation delay
    - Playout size dimensioned to limit packet loss
- Application waits for the buffer to fill up to a certain threshold before starting to play => <u>playout delay</u>
    - Relevant for live streaming, like Internet radio, football

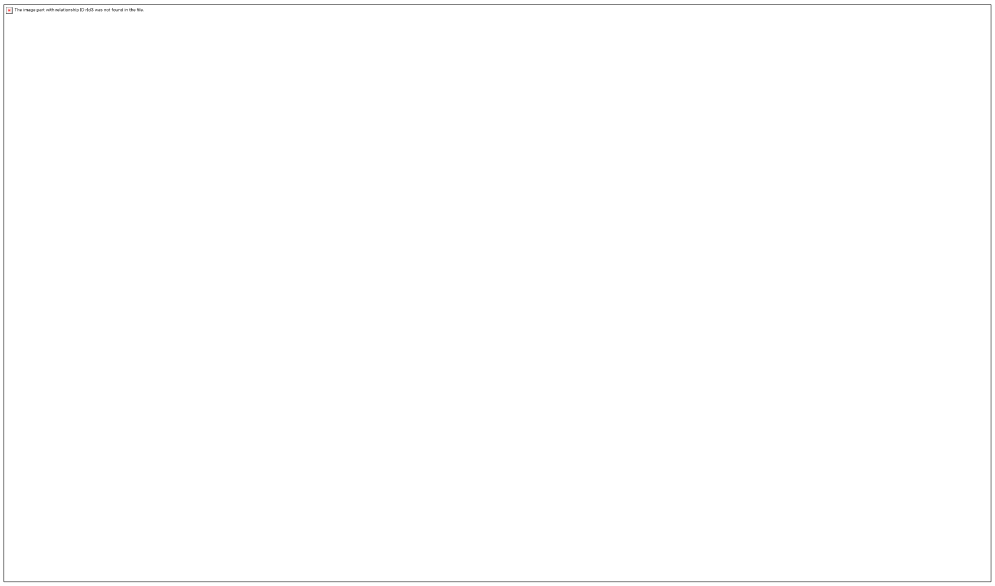
Image in “Computer Networks: a Systems Approach”, L. Peterson, B. Davie

Playout Buffer

- Decoder reads from buffer
  - Frame to be played at time n must arrive earlier
  - Packet not on time may lead to frame not on time
- N^{th} frame should be played at time $t_n=t_0+n/r$
  - t_0: initial playout delay
  - r: frame rate

- Q: What may cause an empty playout buffer?

How does QoS influence QoE?



[Technical Committee on Human Factors - Guidelines for real-time communication services]

Transport and Signalling

- What support does data from such applications need?

  Transport

- What must sender and receiver agree on before, during and after engaging in communication?

  Signalling

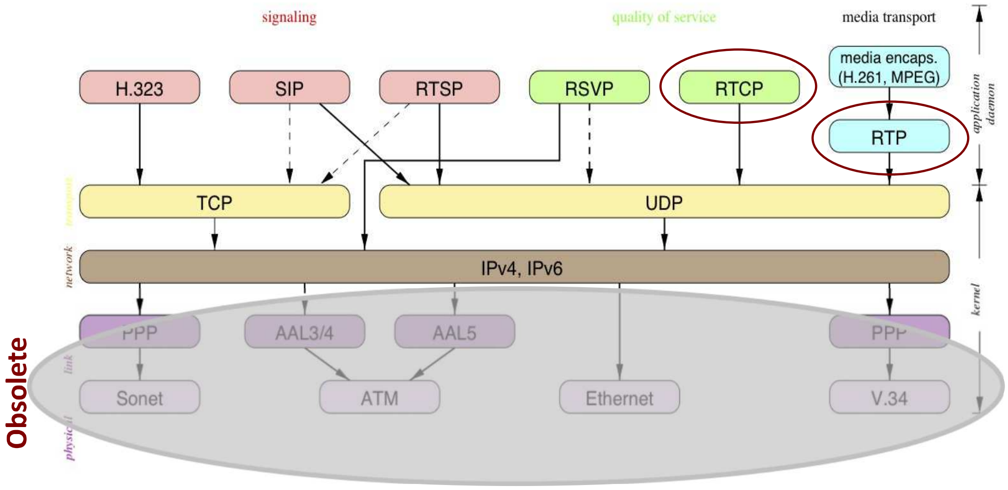
- What information must sender and receiver exchange during communication?

  Signalling

RTP, RTCP, and friends

- Q: Is TCP adequate for streaming multimedia or for interactive multimedia?
  - Why/ why not?

IETF Protocol Suite



In “Internet Telephony: Architecture and Protocols -- an IETF Perspective”, H. Schulzrinne, J. Rosenberg, IEEE Network, 1999

Real Time Protocol (RTP)

- Provides mechanisms for streaming data end-to-end in real-time among multiple clients
- Flexible
  - Supports varied media formats and transport protocols
- Out-of-band signalling
  - data plane: RTP
  - control plane: RTCP

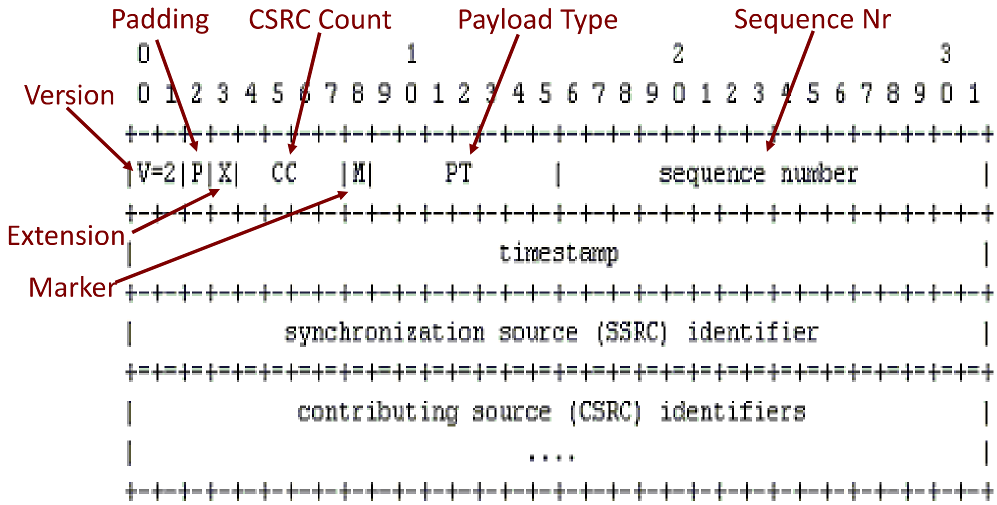
Real Time Protocol (RTP)

- Provides timestamping
  - for playout buffer
- Provides sequence numbers
  - Identify losses, re-order packets
- Identifies source of transported media streams
  - Useful for conferencing
- Enables multiplexing different content

RTP Control Plane: RTCP

- Synchronisation of different media (lip synchronisation)
  - NTP timestamp
- Inform of network issues, e.g. congestion information
- Negotiate and adapt media specific characteristics
  - Coding, compression, dynamic frame boundaries

RTP Header



SSRC: Synchronisation source identifier

CSRC: Contributing sources identifier

CSRC Count: number contributing sources

Detailed description of header: RFC 3550 or http://www.siptutorial.net/RTP/header.html

Timing Information

- Timestamp
  - Measured in units of samples of the media being carried
    - So they can be used to calculate playout time instant
    - $t_s = f_s * \Delta t$, f_s: sampling frequency, Δt: packetisation interval
  - Playout delay compensation at receiver
    - Playout time instant: $p = b * t + c$

    b = position in buffer

    t = time

    c = offset, can change to accommodate jitter variations

Timestamp Example

- Audio
  - Sampled @ 8KHz
  - Packets sent every 20ms
  - Timestamp increment = 160
- Video
  - Clock rate fixed at 90KHz
  - Fps = 30
  - Timestamp increment = 3000
  - All packets carrying the same video frame carry the same timestamp

RTP Entities

- Transport address
  - Combination of IP and port (endpoint)
- RTP media type
  - Collection of payload types within a session
- RTP session
  - Communication between a pair of transport addresses
- RTP multimedia session
  - Set of sessions among a group of participants
- Synchronisation Source (SSRC)
  - Source of a synchronised RTP stream
- Contribution Source (CSRC)
  - Source of stream contributing to combined stream produced by mixer

RTP Mixer

- Intermediate system receives and combines RTP packets from different sources
- Needs to establish synchronisation relationship between mixed streams
- Resulting stream goes to a new multicast address

RTP Translator

- Different participants in a multimedia session may use different data standards (formats, compression, ...)
- Intermediate systems that does not change SSRC
- Can change payloads: transcode, change compression

RTP Media Profiles

- Header fields interpretation defined by media profile
- Payload type: identifies format and interpretation of payload
  - E.g. RTP AVP: Audio/ Video Profile with minimal session control (RFC 3551)
- "It provides interpretations of generic fields within the RTP specification suitable for audio and video conferences. In particular, this document defines a set of default mappings from payload type numbers to encodings."
- Marker: interpretation is defined by profile, e.g. frame boundaries
- Extension headers: may be defined to carry additional information, e.g. MPEG-4 transport streams

RTP as a Framework

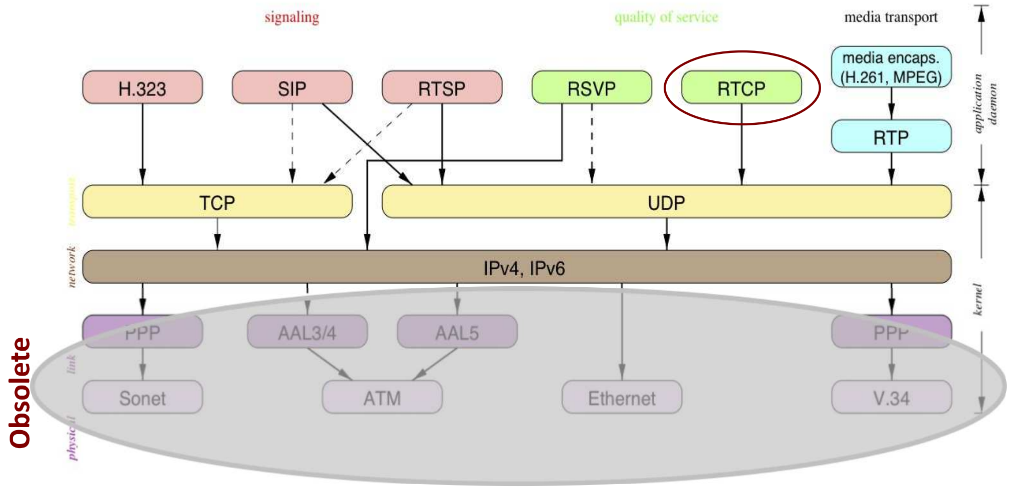
- RTP offers protocol mechanisms to signal relevant information between sender and receiver

- Meaning of fields defined by media profile
  - Easy to define new functionality

- Separation of protocol from meaning of information!

- RTP FAQ: http://www.cs.columbia.edu/~hgs/rtp/faq.html

IETF Protocol Suite



In “Internet Telephony: Architecture and Protocols -- an IETF Perspective”, H. Schulzrinne, J. Rosenberg, IEEE Network, 1999

Real Time Control Protocol (RTCP)

- Functionalities:
  - feedback receiver quality
  - provide means for adaptation: flow control, adaptive media coding
- RTCP provides periodic feedback to all members of an RTP session using the same distribution as media data
  - reception statistics on quality (loss, delay, jitter)
  - faults for network problem diagnosis
  - distribution properties
- De-multiplexing of RTP and RTCP is achieved by using different ports
  - RTCP uses the port directly above the RTP port

RTCP Packets

- SR
  - Sender report, for transmission and reception statistics from participants that are active senders
- RR
  - Receiver report, for reception statistics from participants that are not active senders
- SDES
  - Source description items, including CNAME
- BYE
  - Indicates end of participation
- APP
  - Application specific functions

RTCP Sender Reports

- NTP timestamp
- RTP timestamp
- Sender total packet count
- Sender total octet count
- Receiver reports of receivers

RTCP Receiver Reports

Receivers report to sources

- Fraction lost blocks since last report
- Total number of lost packets
- Highest sequence number received
- Jitter estimate
- Time of last report
- Time since last report

These reports are re-distributed by the source

RTCP Performance Issues

- Signalling protocol: overhead vs performance trade-off
  - Too many reports => occupies too much bandwidth
  - Too few reports => low reaction capability
  - RFC suggests RTCP bandwidth limitation to 5%
    - 1.25% for sources
    - 3.75% for receivers
- RTCP packets from different session participants should be de-synchronised
  - If synchronised, RTCP packet bursts can occur

Performance Issues

1. Calculate the frequency of RTCP reports to be sent by receivers in a multimedia conference with 10 participants and 64kbps data rate each. Re-calculate the frequency for 256kbps each.

2. Calculate the overhead for a G.729 (8kbps, 10ms packetisation) VoIP packet using RTP over IPv4 and IPv6.

3. Why can the use of RTP+UDP seriously interfere with TCP flows?

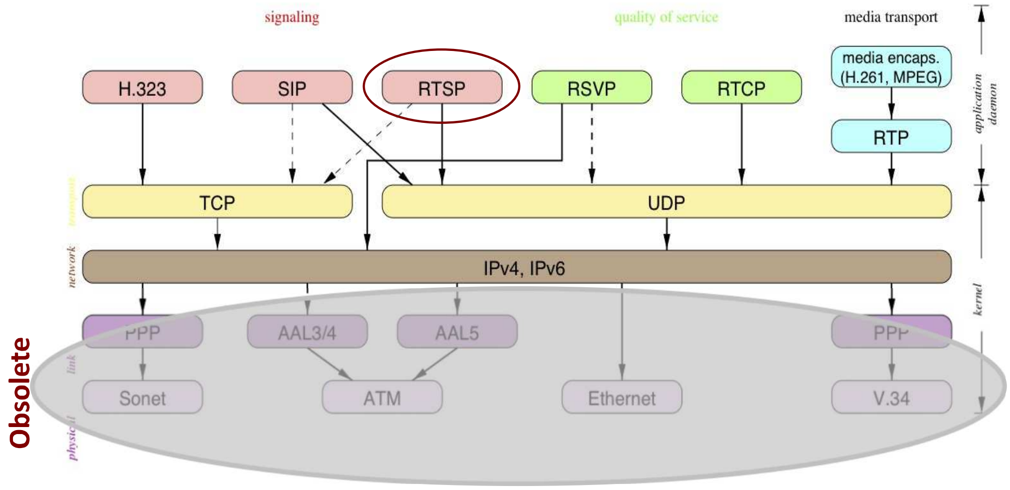
Session Signalling

Purpose of Signalling

- Connection management
  - Establishment, termination, forwarding, etc
- Negotiation of endpoint features or settings
  - Before and/or during data exchange
- Resource reservation
  - Between endpoints and routers

- Multimedia signalling protocols
  - RTSP (Real-Time Streaming Protocol)
  - SIP (Session Initiation Protocol), H.263
- QoS Signalling Protools
  - RSVP (Resource reSerVation Protocol)
  - Next Steps in Signalling (NSIS)
    - GIST: General Internet Signalling Transport
    - NLSP: NSIS signalling Layer Protocol

IETF Protocol Suite



In “Internet Telephony: Architecture and Protocols -- an IETF Perspective”, H. Schulzrinne, J. Rosenberg, IEEE Network, 1999

RTSP

RFC 2326: Real Time Stream Control Protocol

“The Real-Time Streaming Protocol (RTSP) establishes and controls either a single or several time-synchronized streams of continuous media such as audio and video. It does not typically deliver the continuous streams itself, although interleaving of the continuous media stream with the control stream is possible (see Section 10.12). In other words, RTSP acts as a "network remote control" for multimedia servers.”,

In RFC2326, Schulzrinne et al, 1998

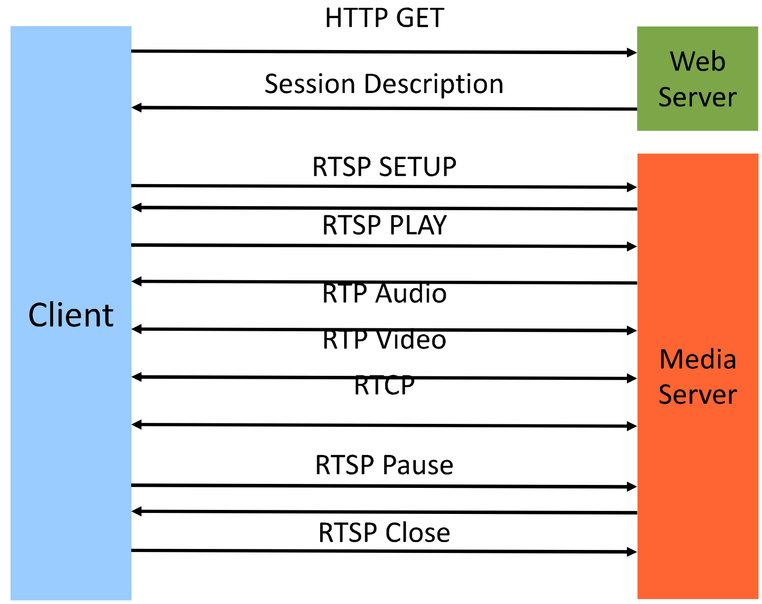
RTSP Features

- Supports:
  - Retrieval of media from a server
  - Invitation of a media server to a conference
  - Recording of a conference
- Similar to HTTP, but with differences:
  - Stateful
  - Both client and server can issue requests
  - Data carried by another protocol
- Independent of description protocol (SDP, SMIL, …)
- Works with unicast and multicast

RTSP Methods

• OPTIONS	Get available methods
• SETUP	Establish transport
• PLAY	Start playback
• RECORD	Start recording
• REDIRECT	Redirect client to new server
• PAUSE	Halt delivery but keep state
• TEARDOWN	Remove state
• ANNOUNCE	Change description of media object
• DESCRIBE	Get description of media object
• SET_PARAMETER	Device or encoding control

Interplay with RTP



Session Description Protocol

- RFC 4566
- “SDP is intended for describing multimedia sessions for the purposes of session announcement, session invitation, and other forms of multimedia session initiation.”

- What describes a multimedia session?

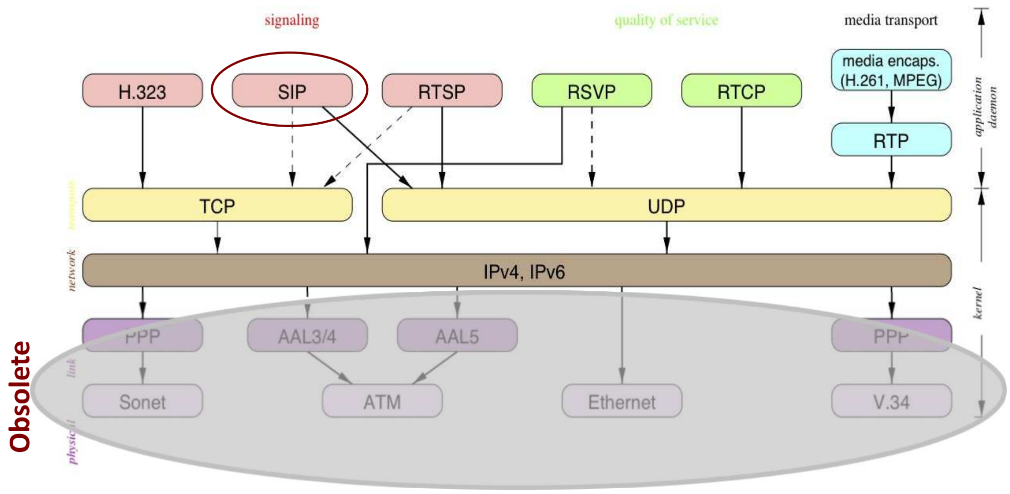
SDP Format

- Session name
- Time boundaries
- Type of media
- Transport protocol
- Media format
- Remote address and port

SDP Example

```
v=0
o=jdoe 2890844526 2890842807 IN IP4 10.47.16.5
s=SDP Seminar
i=A Seminar on the session description protocol
u=http://www.example.com/seminars/sdp.pdf
e=j.doe@example.com (Jane Doe)
c=IN IP4 224.2.17.12/127
t=2873397496 2873404696
a=recvonly
m=audio 49170 RTP/AVP 0
m=video 51372 RTP/AVP 99
a=rtpmap:99 h263-1998/90000
```

IETF Protocol Suite



In "Internet Telephony: Architecture and Protocols -- an IETF Perspective", H. Schulzrinne, J. Rosenberg, IEEE Network, 1999

Session Initiation Protocol (SIP)

- RFC 3261

- “This document describes Session Initiation Protocol (SIP), an application-layer control (signaling) protocol for creating, modifying, and terminating sessions with one or more participants. These sessions include Internet telephone calls, multimedia distribution, and multimedia conferences.”

SIP Functionalities

- Signalling for multimedia sessions
- Call setup
  - Ringing and establishment
- Call handling
  - Sustaining, transferring, terminating
- User location
  - Discovery of user presence
- User availability
- User capabilities
  - Parameters for communication

SIP Components

- SIP Adresses
  - URI
- SIP Messages
  - Request-response transactions, format similar to SMTP
- User Agent Server
  - Session requests, service register and control, AAA, proxy, location
- User Agent Client
  - Initiates session
- SIP Protocol
  - Rules for transactions between user agents

SIP Message Example

```
INVITE sip:tpu@hp.com SIP/2.0
Via: SIP/2.0/UDP local.hp.com
From: OC <sip:OpenCall.SIP@hp.com>
To: TPU <sip:tpu@hp.com>
Subject: Confcall
Call-ID: 132059753@local.hp.com
Content-Type: application/sdp
CSeq: 1 INVITE
Contact: <sip:telecom@16.188.155.140>
Content-Length: 187
```

Header

```
v=0
o=user1 51633745 1348648134 IN IP4 16.188.155.140
s=Interactive Conference
c=IN IP4 224.2.4.4/127
t=0 0
m=audio 3456 RTP/AVP 0 22
a=rtpmap:22 application/g723.1
```

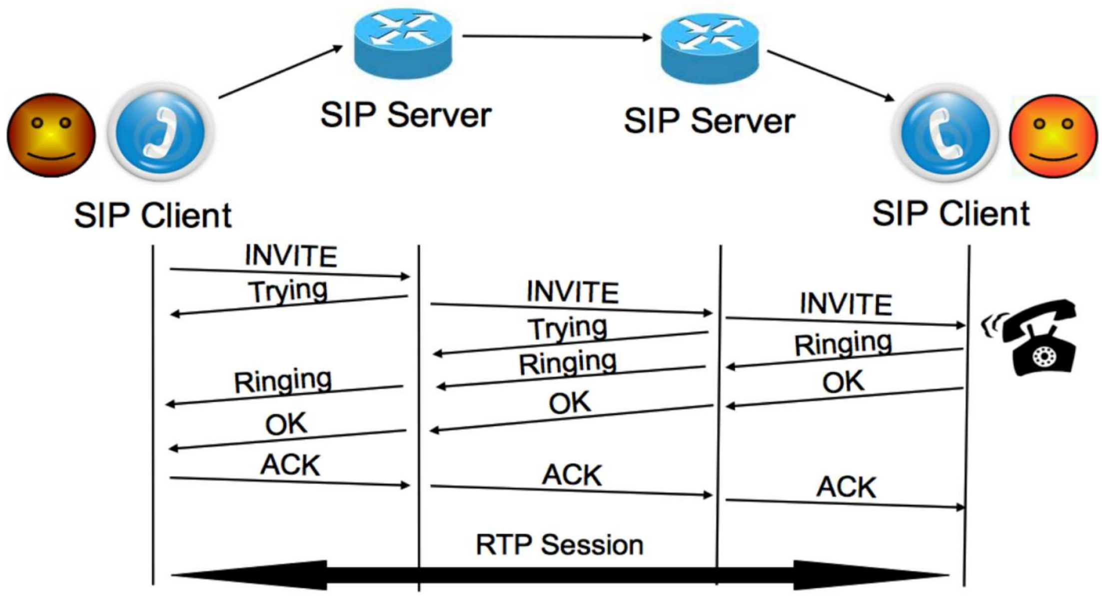
Body

http://docs.hp.com/en/5992-1950/ch01s04.html

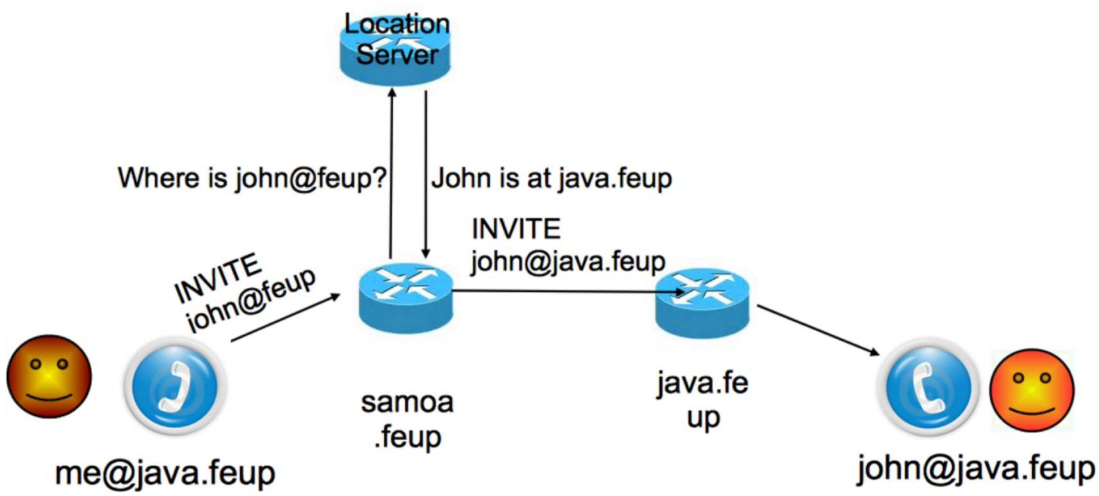
SIP Methods

- INVITE: Indicates a user or service is being invited to participate in a call session.
- ACK: Confirms that the client has received a final response to an INVITE request.
- BYE: Terminates a call and can be sent by either the caller or the callee.
- CANCEL: Cancels any pending searches but does not terminate a call that has already been accepted.
- OPTIONS: Queries the capabilities of servers.
- REGISTER: Registers the address listed in the To header field with a SIP server.

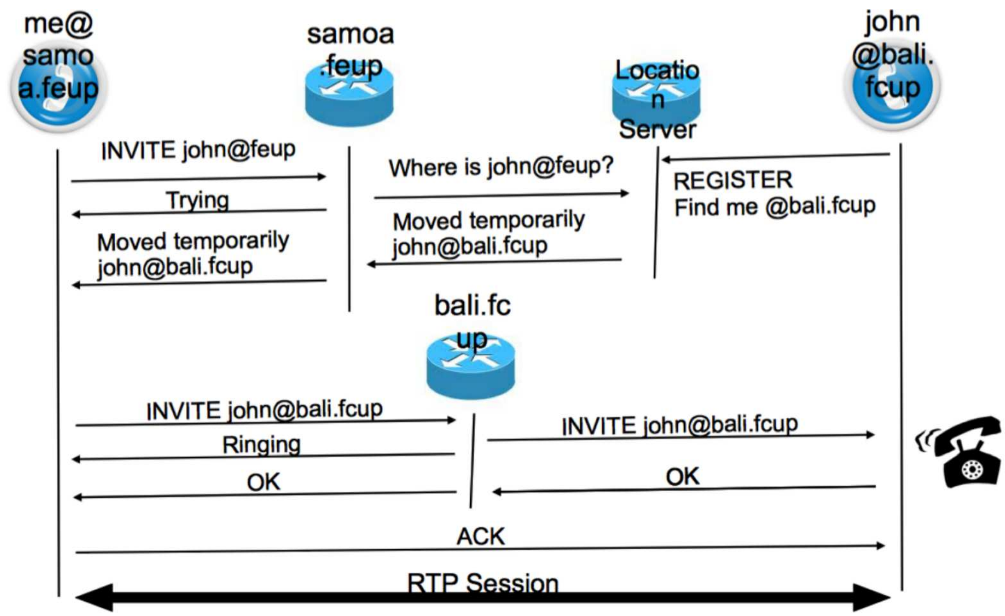
SIP Session Setup Transaction



SIP User Location & Call Relay



SIP Call Re-direct



NEXT: MPEG DASH