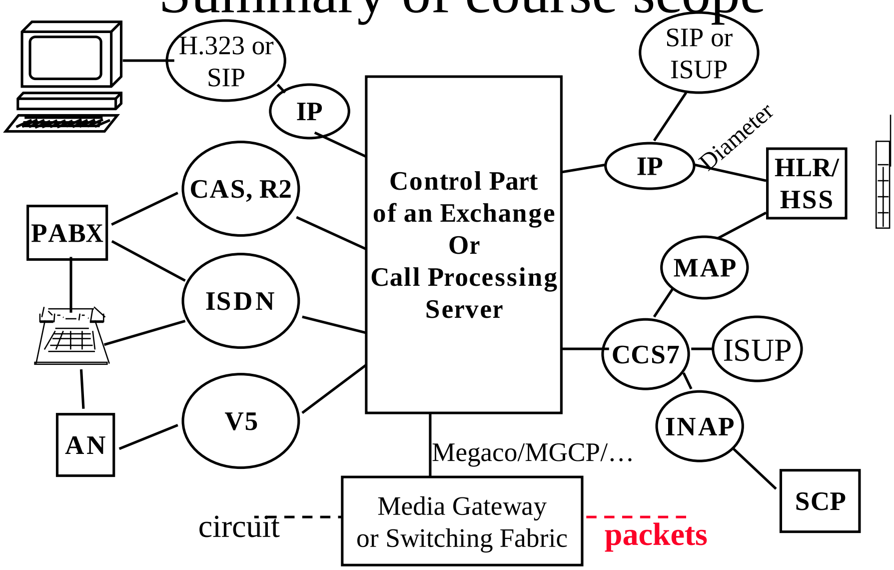
IP Telephony

Overview of IP Telephony

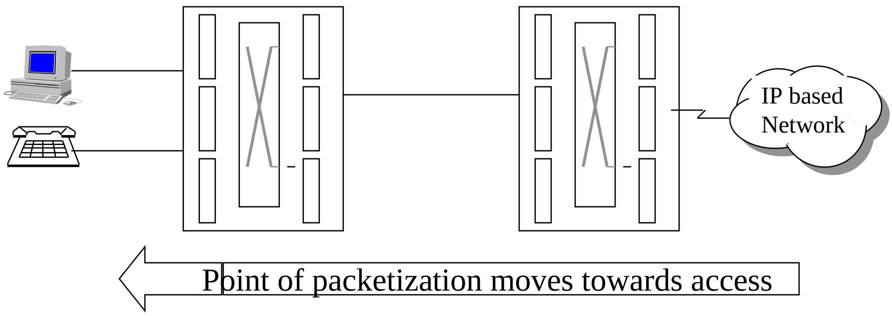
Media processing, RTP, RTCP

Quality of Service

Summary of course scope



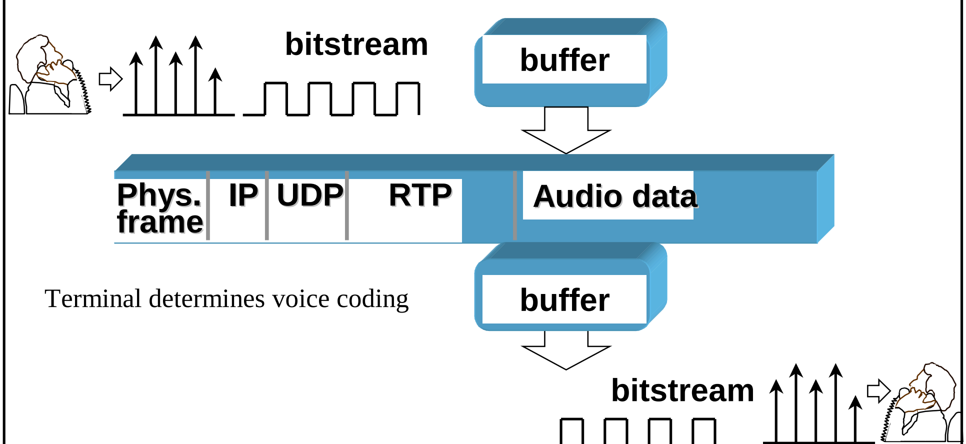
Data traffic already dominates voice in volume. Therefore Data will drive the Network Architecture.



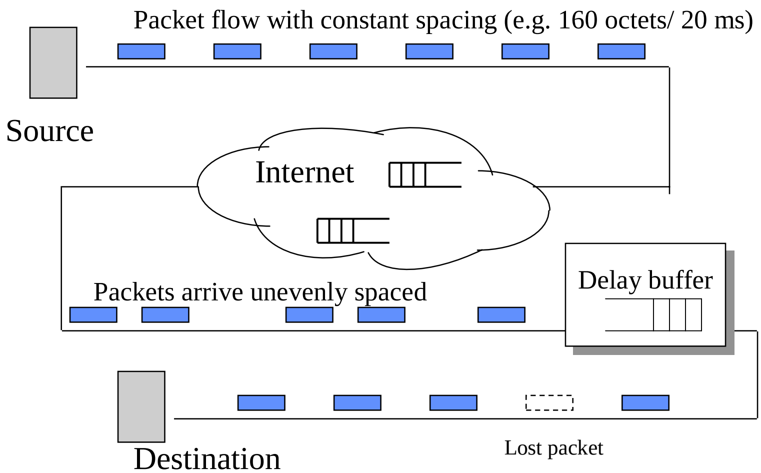
- Broadband Networks will be based on packet switching
- BB network emerges from the existing Internet
- Each step of Development pays for itself.

VoIP in action

Coded samples (G.711, G.729B, G.723.1)



Delay variance is compensated at reception by buffering



IP Telephony Standardization is active on de-jure and de-facto fora

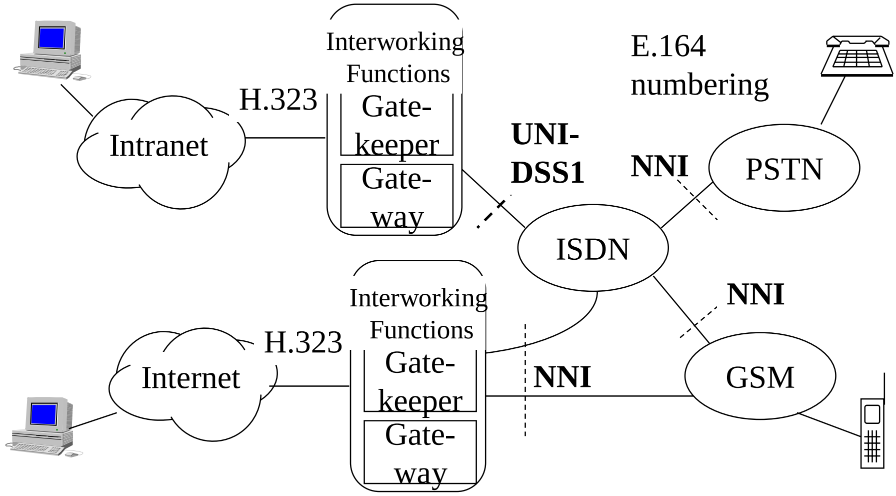
- ITU-T - H.3xx, H.2xx series
- ETSI - TISPAN (NGN – next generation networks… took over from TIPHON which was a project)

IETF working Groups

- IPTEL (IP telephony) and PINT
- MMUSIC (Multiparty Multimedia Session Control)
- SIMPLE – SIP for Instant Messaging and Presence Leveraging Extensions
- AVT – audio video transport
- ENUM – tElephone NUmber Mapping
- Megaco – media gateway control
- Rohc – robust header compression
- SIP – Session Initiation Protocol
- SIPPING – Session initiation Protocol Investigation
- SIGTRAN (ISUP and other CCS7 over IP)

VOIP - Voice over IP by IMTC - Int'l Multimedia Teleconferencing Consortium

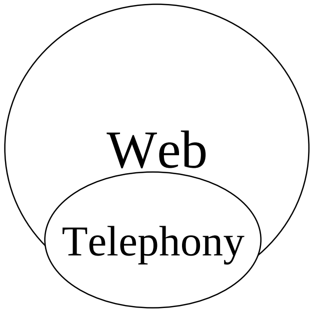
TIPHON specifies IP Voice to PSTN/ISDN/GSM Interworking



H.323 products are available

- ITSPs are committed to H.323
- MS Netmeeting, Intel Videophone, Netscape Conference are examples of H.323 clients
- H.323 version 2 products are available
- Gateways and Gatekeepers/Call managers are available
- SIP has been taking the lead over past 2…4 years but takes time

IETF alternatives to H.323 pursue Integration of Telephony to the Web

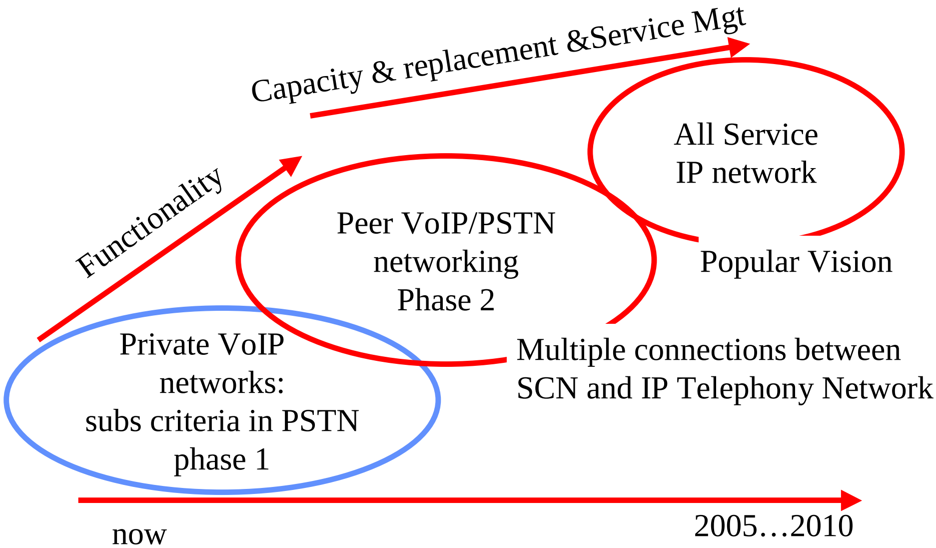


- AVT - Audio Video transport (…RTP)
- PINT worked on Click-to-Dial, Click-to-Fax, Click-to-Fax-Back “www-buttons”. The idea is to integrate www to IN
- Mmusic (now SIP group) works on SIP - idea is to use web-technology to absorb signaling
  - SIP has been adopted by 3GPP for 3G packet telephony
- Media Gateway Control (Megaco)
- SIGTRAN works on C7 over IP
- ENUM - numbering info in DNS

Latest move is the emergence of Peer-to-Peer VOIP

- Designers of KaZaA have released SKYPE – a peer-to-peer VOIP application
  - No network based servers are needed
  - Has node/supernode structure like KaZaA
- In p2p search of callee is integrated in the application
- Everything in SKYPE is secret, even monitoring of signaling is forbidden by licence conditions!

Roadmap to the Future



Interoperability Issues

- Signaling and Call control
- Quality of Service

Phase 1
--->

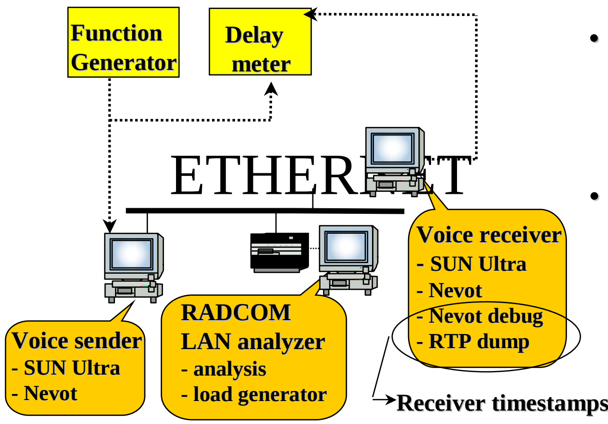
- Telephony Routing and addressing
  - Input Information gathering
  - Alternative routing over IP

Phase 2
-->

- Service Management in the hybrid network

Phase 3

IP Voice in Ethernet - Delay is in the Workstation (IPANA -97)



- **Terminal delay:**

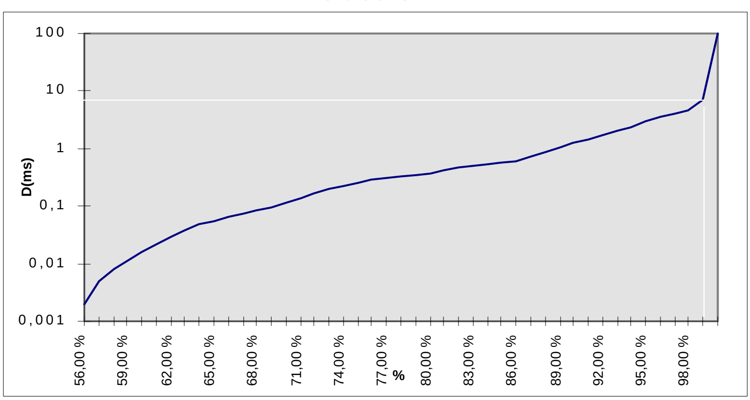
	Delay
HW	8.9 ms
VoIP Client	103.9 ms

- **End-to-end delay**

Packet length	Delay
0.02 s	104.5 ms

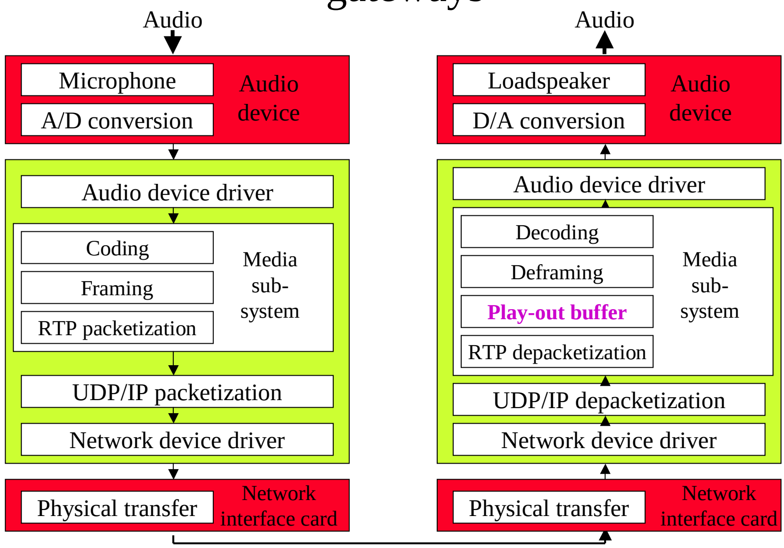
Difference = network delay

Packet spacing difference in a campus network



- In the public Internet lack of bandwidth, congested routes/links and underdeveloped charging are blockers to IP Voice.

Media processing path in terminals and gateways



Delay in practical IP voice systems

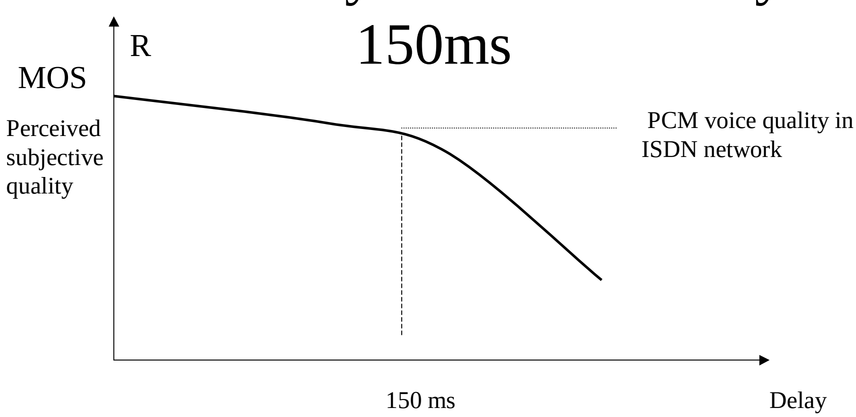
Delay component	ms	Explanation
Audio HW &device driver	0-100	Buffering
Algorithm	20-37.5	Sample length + lookahead time
Operating system	0 - 30	Depends on load and implementation
Coder	<5	Predictable delay in coding algorithm
Decoding	<1	Typically an easy process
Framing and packetization	<1	A small software delay
NIC and device driver	<5	Has some signifigance especially in WLAN
Network	0 - 500	In LAN about 1 ms, Dimensioning Issue!
Play-out buffer	0 - 100	At reception,depends on the state of the network
Synchronization	0 - 30	Audio device requests for data at constant intervals that can not be synchronized with packet arrivals. Avg = half a packet time

Source: M.Sc thesis by Jari Selin

Voice coding for IP networks

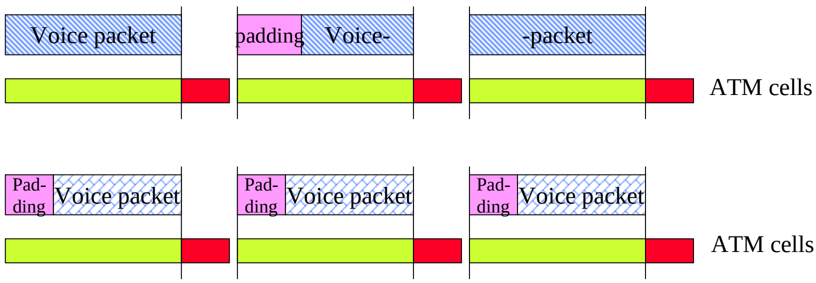
- IP networks are characterised by packet loss
  - coders that have dependies between packet do not perform well (such as Cellular etc.)
  - even 5% packet loss may seriously degrade quality
- Higher than PSTN quality can be targeted:
  - Coding can be done at e.g. 16kHz (not 8 kHz like in PSTN),
  - packets can be variable length
  - in BB environment bitrate can be increased
- E.g. GIPS (Global IP Sound) provides proprietary codecs specifically designed for packet loss networks. E.g. sound quality stays good even at 30% packet loss (at avg 80kbit/s + packet overhead).

Voice quality starts to degrade, when one way end-to-end delay > 150ms



Quality can be measured e.g. based on the E-model or using MOS –measurements. MOS - Mean Opinion Score.

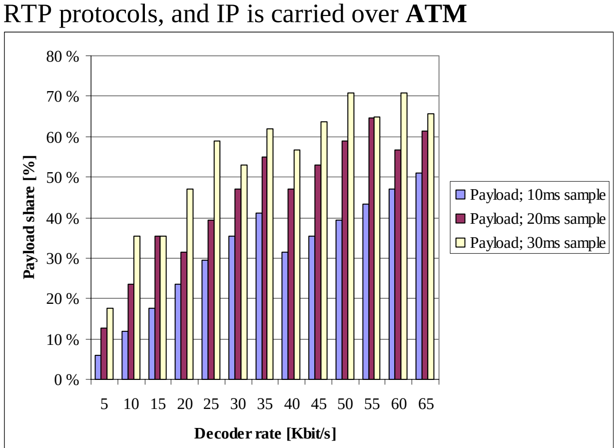
Packetization of voice samples in ATM cells



- Length of voice packet depends on coding method and the length of voice frame
- Packet overhead includes ATM headers and padding, which is needed in order not to increase packet delay.

Packet header and padding overhead is significant

Percentage of voice payload when samples are carried over IP, UDP and RTP protocols, and IP is carried over **ATM**



Source: Veikko Brax (Lic thesis)

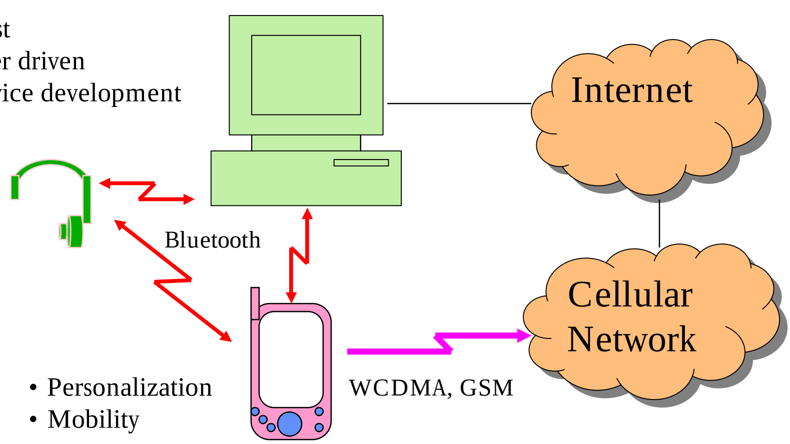
Why voice over IP, when ISDN/GSM work perfectly well?

NB: Voice brings currently ca. 90% of operator revenues!

- Integration of voice and data networks creates new services.
- Maintaining two networks is expensive.
- Data traffic grows >30%/year, voice ≈ 5%/y, volumes were approximately equal 2002. If trend continues, in 2010 share of voice will be < 10%, data will be 90% of all traffic.
- Cost of transmission is in free fall: xDSL, SDH, WDM - this trend is difficult to take advantage of using circuit switching: only one sample (8 bits) can be switched at a time cmp. E.g. 20 ms sample => 1 Gbit router is less expensive than an exchange with a 1 Gbit switch fabric.
- Terminals can do more -> consumer market economy helps.

User view on VOIP

- Cost
- User driven service development



- Personalization
- Mobility

Real time Services in IP

RTP (RFC 1889)

RTCP - “ -

Telephony over IP

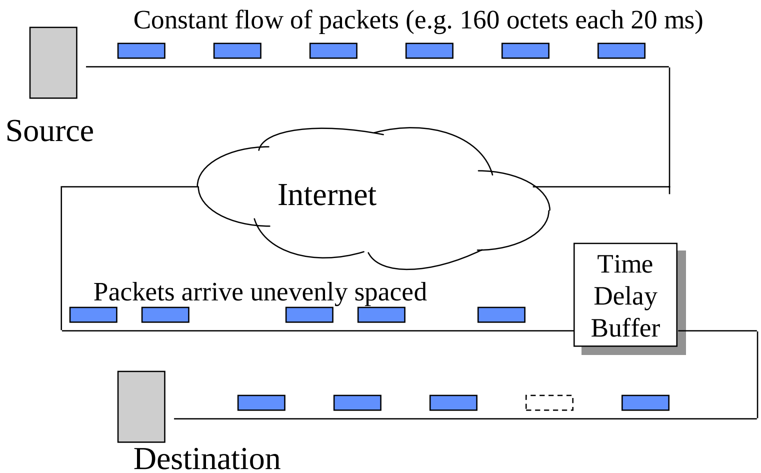
TCP is not suitable for real time services

Applications include

- Audio and video conferencing
- Shared workspaces
- Telephony
- Games
- Remote medicine
- ...

- TCP is point-to-point - not suitable for multicast
- TCP has retransmission for lost segments --> out of order delivery
- No mechanism for associating timing info with segments

Variable delay has to be compensated at reception by delay buffer



Packet arrival process is characterised by delay jitter and packet spacing difference

Delay jitter = Maximum variance in packet delay in a session

Example: fastest packet arrive in 1 ms
slowest arrive in 8 ms.
Delay jitter is 7 ms.

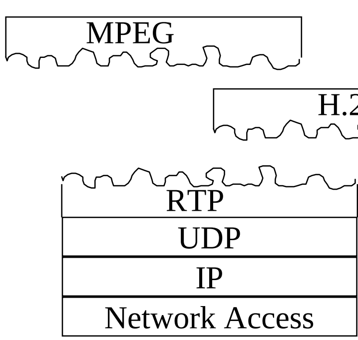
Packet spacing difference is measured based on receiver clock only:

Spacing difference = $[(t_i - t_{i-1}) - (t_j - t_{j-1})]$

Soft real time communications tolerate some loss but need the following

- Low jitter and Low latency
- Ability to integrate real-time and non-real-time services
- Adaptability to changing network and traffic conditions
- Performance for large networks and large nrof connections
- Modest buffering requirements in the network
- Effective capacity utilization
- Low processing overhead per packet

RTP - Real time protocol is a “sub-layer” library on top of UDP

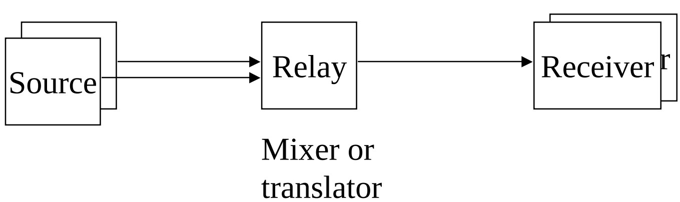


- RTP leaves recovery from loss to the application
- Instead of retransmission e.g. more compact coding may be chosen
- RTP provides sequencing

RTP supports the transfer of real time data among participants of a session

- Session is defined by
  - RTP port number (dest port in UDP header of all receivers)
  - RTCP - Real time control protocol port number
  - Participant IP addresses - multicast address or a set of unicast addresses
- For session set-up e.g H.323 or SIP - Session Initiation Protocol can be used

RTP transport model includes sources, relays and receivers



- A mixer will combine sources - e.g. add voice signals from all conference participants
- A translator may translate from one video format to another
- The relay will mark itself as the synchronisation source

RTP header

V=2 = version

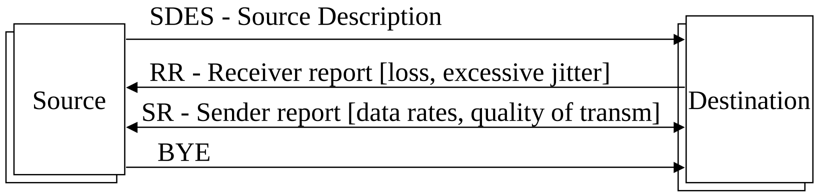
0			4	8	9	16 – 31
V	P	X	CC	M	Payload type	Sequence number
Timestamp						
Synchronisation source (SSRC) identifier						
Contributing source (CSRC) identifier						
Contributing source (CSRC) identifier						

P - Padding - indicates that last octet of payload = nrof preceeding padding octets
X - Extension - there is an experimental extension header
CC - CSRC count - Nrof CSRC identifiers following the fixed header
M - Marker - e.g. End of video frame, Beginning of talk spurt
Payload type - format of RTP payload.
Seq. nr - each source starts at a random nr and =+1 for each packet - determines order of packets with the same timestamp
Timestamp - value of local clock at source at generation of first octet of payload
SSRC and CSRC identifiers are generated at random

Main RTP functions are ordering of received packets and timely playout

- Sequence number gives the order of packets
  - say one in sequence is missing – when the playout time of the missing packet comes, e.g. the previous packet can be played out again to concel the error
- The order is not enough, the receiver must know the time difference between the playout times of two consequtive packets – timestamp gives exactly this as measured by the source of the packet

RTCP - RTP Control Protocol provides feedback among participants of the session



- RTCP packets may be multicast in parallel to RTP using another UDP port
- RTCP source is identified by plain text
- Few participants: RTCP reports are sent once in 5s Rate of reports is reduced to max 5% of session traffic if there are more participants

RTCP fixed header is

0		4	8	16 31
V	P	RC/SC	PT	Length
SSRC of sender (or CSRC)				

V =2 = version, P - Padding, same as RTP

RC - Reception report block count in SR or RR

SC - Source item count in SDES or BYE

PT - RTCP packet type [RR, SR, SDES, BYE]

Length - lenght of this packet in 32 bit words - 1

SSRC - same as in RTP

Sender Report carries sender info and reception report blocks

Sender information is

NTP timestamp (MS word)
NTP timestamp (LS word)
RTP timestamp
Sender's packet count
Sender's octet count

NTP is the wall-clock time when sending this report (used for round-trip time measurement)

RTP timestamp lets relate this report to RTP stream

Packet and octet counts run from beginning of session

Reception report block

SSRC_i (SSRC of source)	
Fraction lost	Cum nrof packets lost
Ext highest seq nr received	
Interarrival jitter	
Time of last sender report	
Delay since last sender report	

SSRC identifies source

Fraction lost since last SR or RR, Cum loss is for the whole session

16 LS bits= highest RTP seq nr. 16 MS bits= nrof times seq nr has wrapped back to zero

SR is sent by party who is both sender and receiver !

Average inter-arrival jitter for a source is estimated as follows

S(i) = Timestamp from RTP data packet i
R(i) = Time of arrival of data packet i in RTP timestamp units
D(i) = (R(i) - R(i -1)) - (S(i) - S(i -1))
J(i) = Estimate of Inter-arrival jitter up to the receipt of RTP packet i

$$J(i) = 15/16 \ * J(i-1) + 1/16 * | D(i) |$$

- Receivers use the estimate of Jitter to adjust the play-out delay
- According to measurements the above exponential average is not always optimal

RTCP other packets

- RR are made of the fixed header + reception report blocks (see SR format lower part)
- SDES can carry
  - CNAME - Canonical Name
  - NAME - Real user name of the source
  - Email address of the source
  - Phone number of the source
  - TOOL - name of the tool used by the source

How to provide SCN-like QoS over IP?

- Integrated Services (use RSVP to make reservations in routers for each call!) changes Routers into SCN-Exchange -like systems. Does not scale well.
- DiffServ
  - mark voice packets with higher than BE priority at ingress
  - priority queuing in transit nodes
  - How to prevent voice from blocking BE traffic?
  - How to do Service Management?
  - Voice packets have high overhead - how to minimize?
- Overprovisioning

How is IP Telephony different from Circuit switched telephony?

Circuit Telephony

- Voice sample = 8 bits
- A- and μ -law PCM voice standard
- Reference connection gives network design guidelines => end-to-end delay is under control
- Wire-line telephones are dumb. Cellular phones are pretty smart
- Call control is tied to the voice path - IN is used to add service processing on the side.

IP Telephony

- Voice in 10…40 ms samples, Bits in a sample can be switched in parallel
- No single coding standard
- End-to-End delay is big challenge
- Terminals are intelligent - consumer market economics
- Call control is separate from voice path - first find out whether parties want and can talk, if yes, set-up the voice path

Note: Using todays technology IP Telephony is not less expensive in replacement nor green field investments in Corporate networks!

How realistic is the idea of replacing CSN with VOIP?

- There is more data traffic now than Voice traffic.
  - Data is growing much faster than Voice
  - Voice revenue is still 90% of operator’s networking revenue.
  - Voice is fast becoming mobile
- CSN networking product development has stopped. All R&D effort in telephony goes to VOIP telephony
- Replacement Scenario in Finland: PSTN can be replaced, required max link capacities are 2,5…10Gbit/s. Present FUNET upgrade is to links of 10 Gbit/s and FUNET is just the University Network!
- PCs are still lousy phones!