INTRODUZIONE AL VOICE OVER IP

Dr. Ing. Maurizio Casoni
mcasoni@unimo.it

Dipartimento di Ingegneria dell'Informazione
Università degli Studi di Modena e Reggio Emilia

WHY VOICE OVER PACKET NETWORKS?

- Voice traffic will be in the near future a small fraction of the total telecom traffic moved around the world
- Network operators are building high-capacity packet switched infrastructures
- By providing telephony on this infrastructure they lower costs

Toll by-pass

•Usage of leased lines

•Implemented in PBXs

→

Consolidation

•Integration of voice and data

•Effective sharing of BW

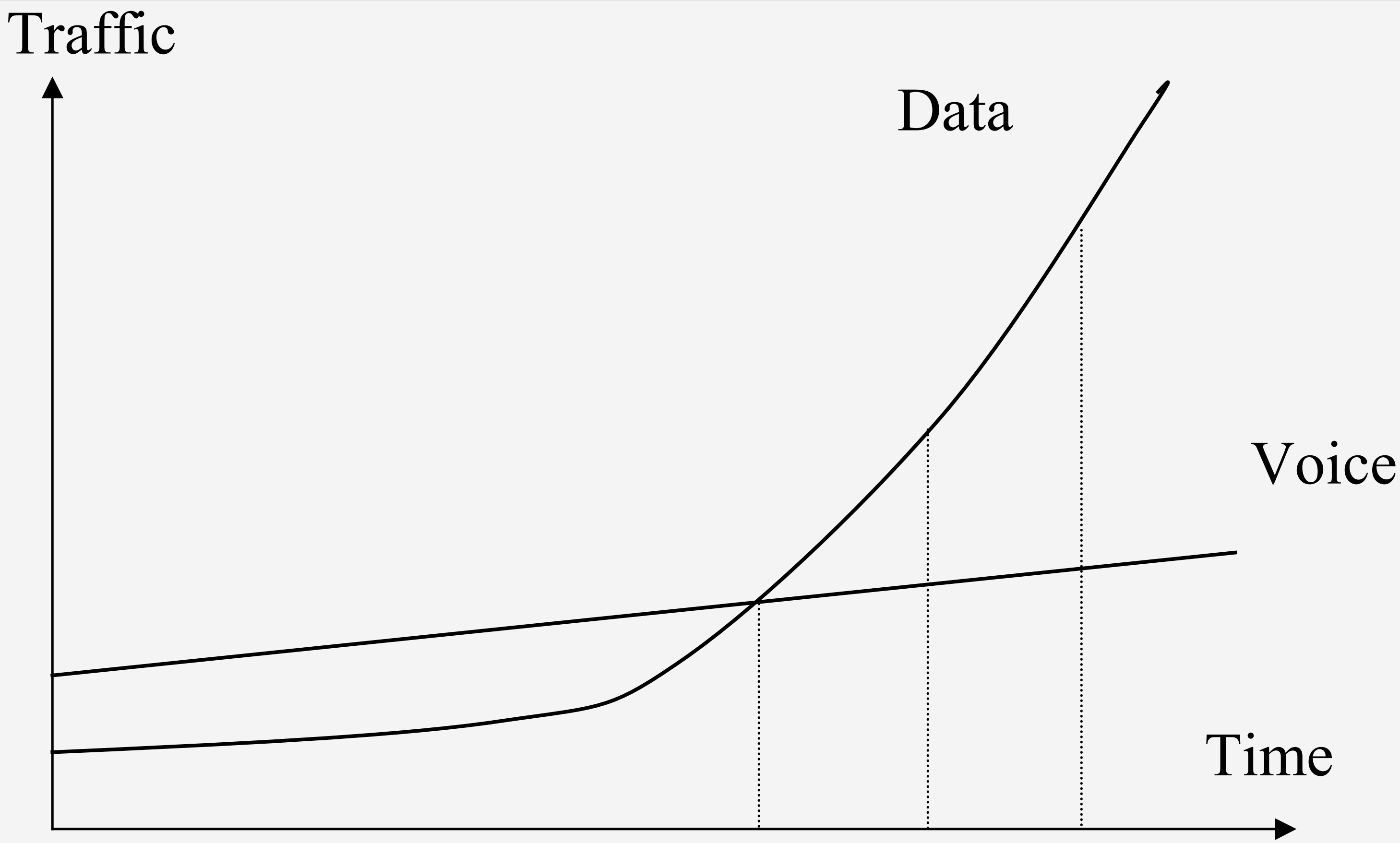
•Some savings

→

Convergence

•Full mm integration

COMPARISON OF VOICE AND DATA NETWORKS GROWTH



- 1998/2000: data networks increase 5 times
- 2000/2005: data growth expected to increase 23 times

TECHNICAL ISSUES

Before expecting a widespread deployment of Internet Telephony some issues must be solved

- Internetneeds to provide some assurance regarding QoS ➔ mechanisms for providing QoS
- Control architectures and protocols are needed to locate users and manage calls (setup, tear down,…)
- Security mechanisms to provide authentication of users and confidentiality for conversation
- Mechanisms for charging (but QoS must be defined before customersare billed…)

VOICE OVER PACKET NETWORKS

Two main approaches

- ☞ **New backbone from the “telephony world”**
  - **Voice over ATM**
    - **mechanisms for providing different QoS levels**
    - **AAL1, AAL2, AAL5**
  - **Voice over Frame Relay**
    - **Many proprietary implementations**
- ☞ **Existing Internet “data networking world”**
  - **VoIP**
    - **traditionally “store&forward”**
    - **QoS has not been considered a basic goal (best effort delivery system)**

STANDARD BODIES

- IETF
  - RTP: audio & video transport
  - SIP and SDP
  - Megaco - H.248
- ATM Forum
  - VTOA: Voice and Telephony over ATM
  - RMOA: Realtime Multimedia over ATM
  - MPOA: Multiprotocol over ATM
- ITU-T
  - H.323, H.248

APPLICATIONS

- Real Time
  - Interactive (two-way)
    - telephony
  - Streaming (one-way)
    - radio-TV broadcast (consumed live)
  - Recording (one-way)
    - replay (stored at the receiver)

Connection Oriented Nets

- Non Real Time
  - Short transfers
    - e_mail
  - Long transfers
    - large image retrieval

Connection less

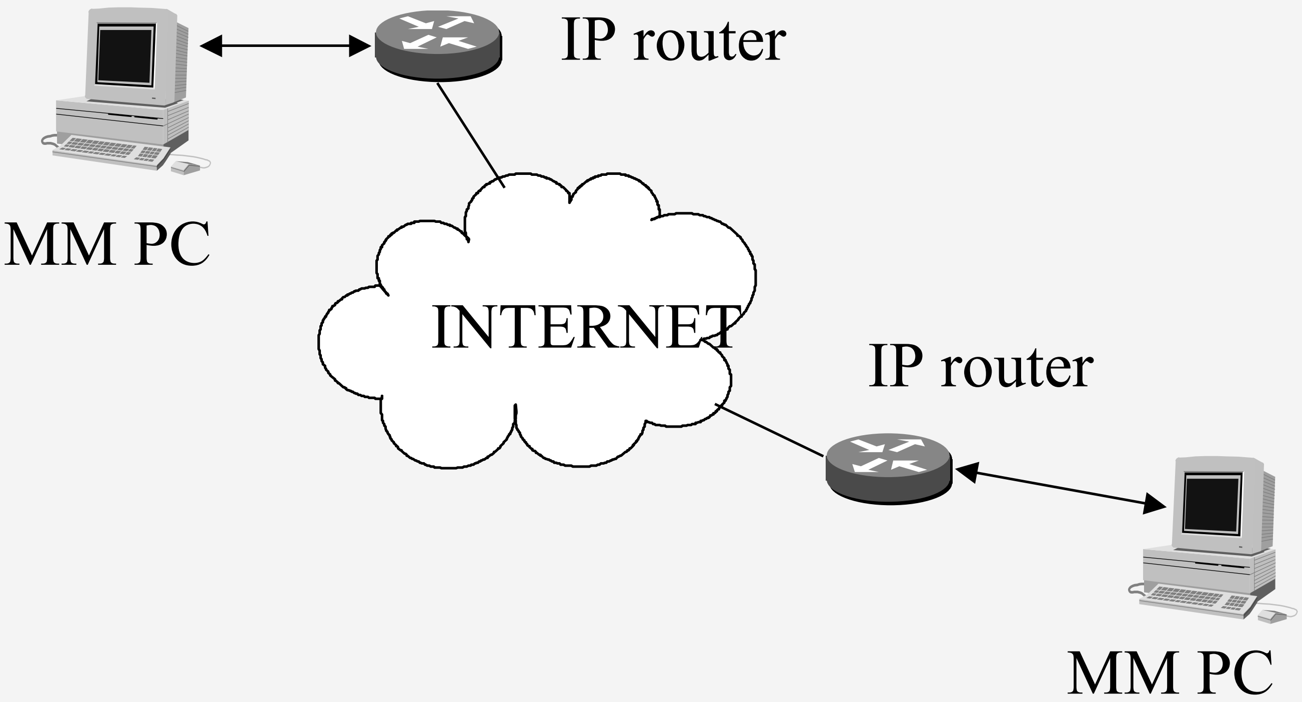
Circuit Switched

TELEPHONY QoS REQUIREMENTS

- ITU-T G.114 specifies that one-way delay for voice
  - **less than 25 ms without echo cancellers (RTT=50ms)**
  - **less than 150 ms with echo cancellers for good quality voice**
  - **less than 400 ms with echo cancellers for acceptable quality voice**
  - **above400 ms unacceptable for most applications**
- This is not valid for one-way streamed traffic
  - **real audio streaming**

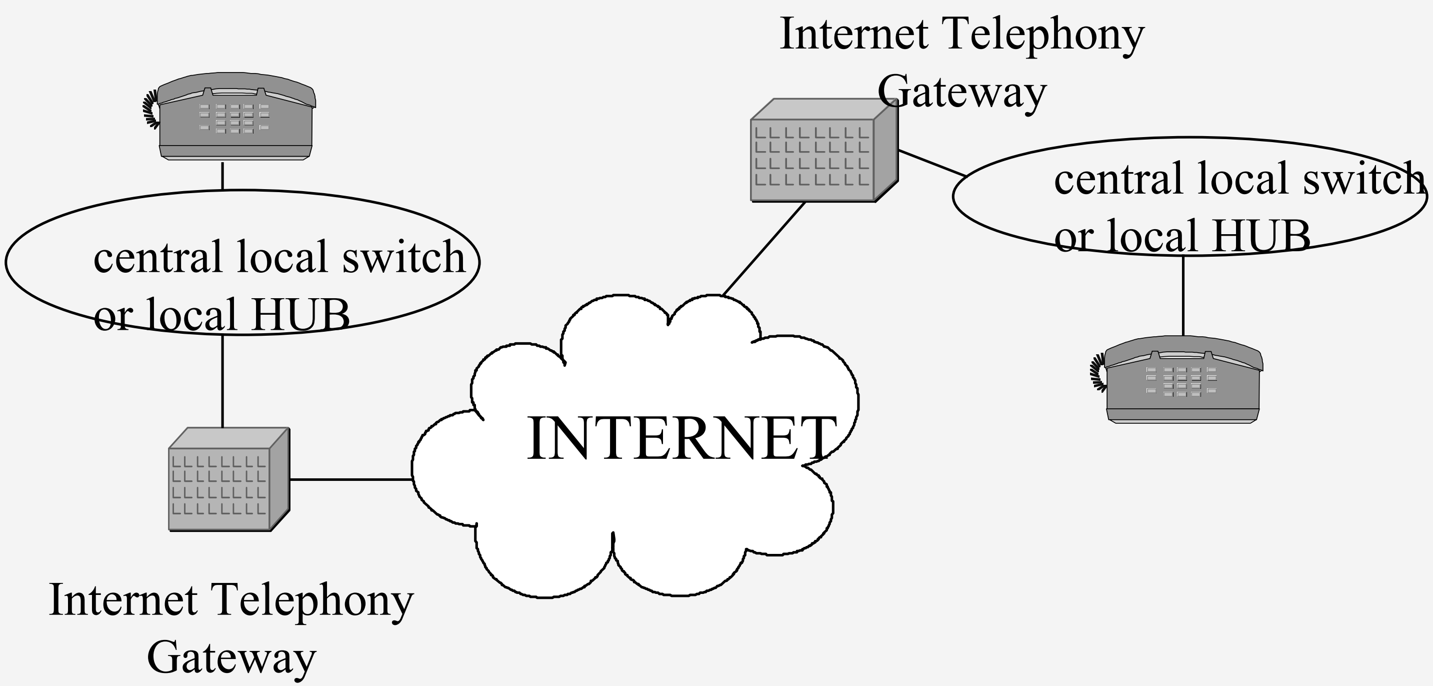
INTERNET TELEPHONY ARCHITECTURE (1)

- Users have access to MM computers connected to Internet by LAN or ISP
- PC-to-PC architecture



INTERNET TELEPHONY ARCHITECTURE (2)

- **Standard telephones to make an receive calls over Internet**
- **User calls an Internet telephony gateway located near a central office switch or local hub**



FACTORS AFFECTING QoS

To transport audio over a NON guaranteed packet switched network, audio samples must be:

- Coded, with some form of compression
- Inserted into packets with
  - sequence numbers
  - creation timestamps
- Transported by the network
- Received in a playout buffer
- Decoded in sequencial order
- Played back

All real-time transport scheme use this scheme

FACTORS AFFECTING QoS

Barriers to the operation of previous schemes, requirements for:

- Codecs
- Bandwidth
- Delay and Jitter
- Losses

PERFORMANCE OF THE INTERNET

- Many studies have been carried on
- Hard task
  - Complexity and variability
  - Mainly based on PING (i.e. round trip time)
  - RTT is NOT one way delay
  - Delay is related to hop counts and number of autonomous systems crossed rather than geographical distance
  - Great variance of packet loss (<1% to 10%)
  - Within the home network RTT < 100 ms

PERFORMANCE OF THE INTERNET

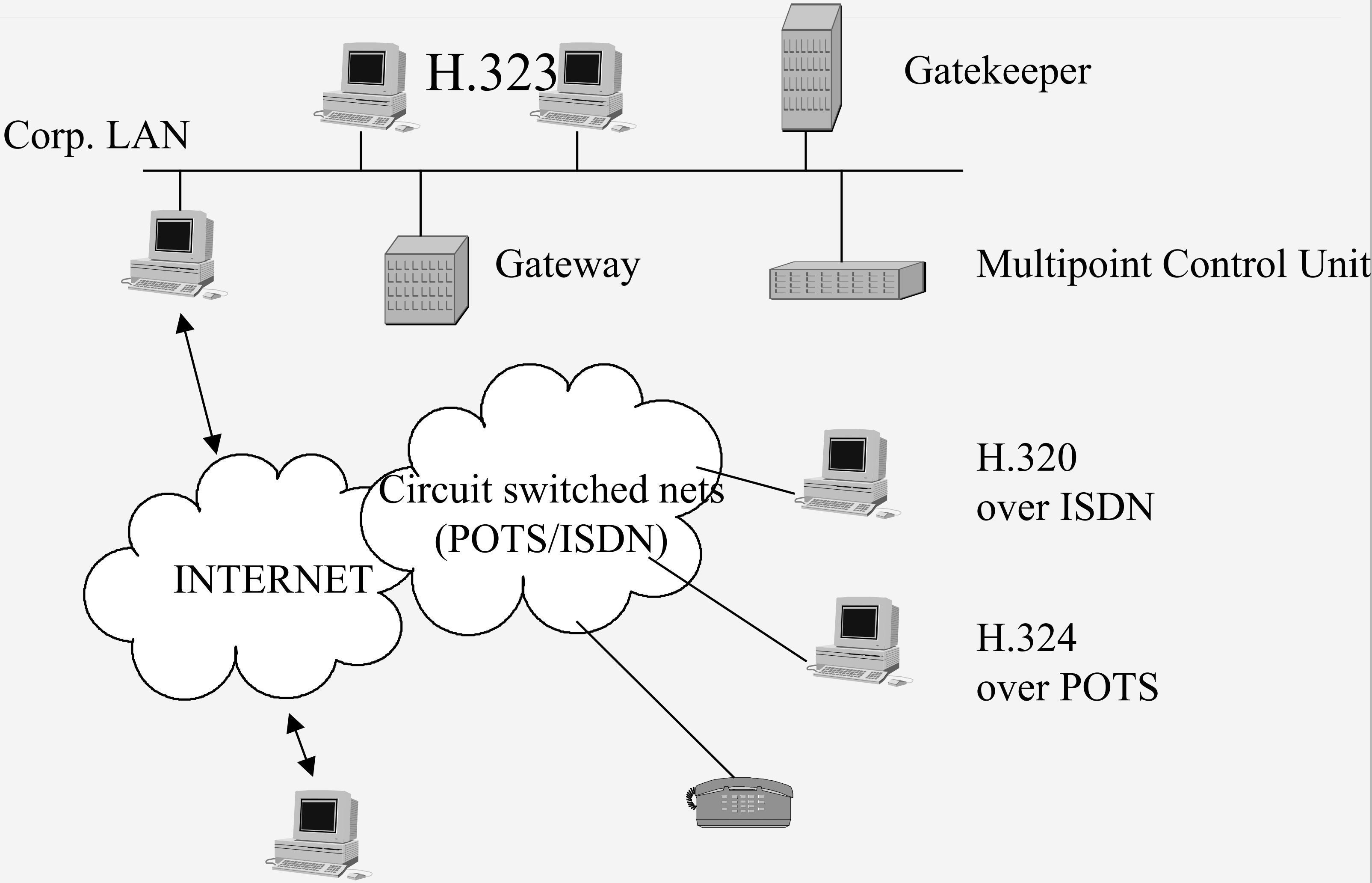
Infocom and Globecom 2002:

"Despite heroine end-system efforts, the Internet (not IP !) is currently incapable of carrying real-time or delay-sensitive traffic"

OVERVIEW OF THE H.323 SYSTEM

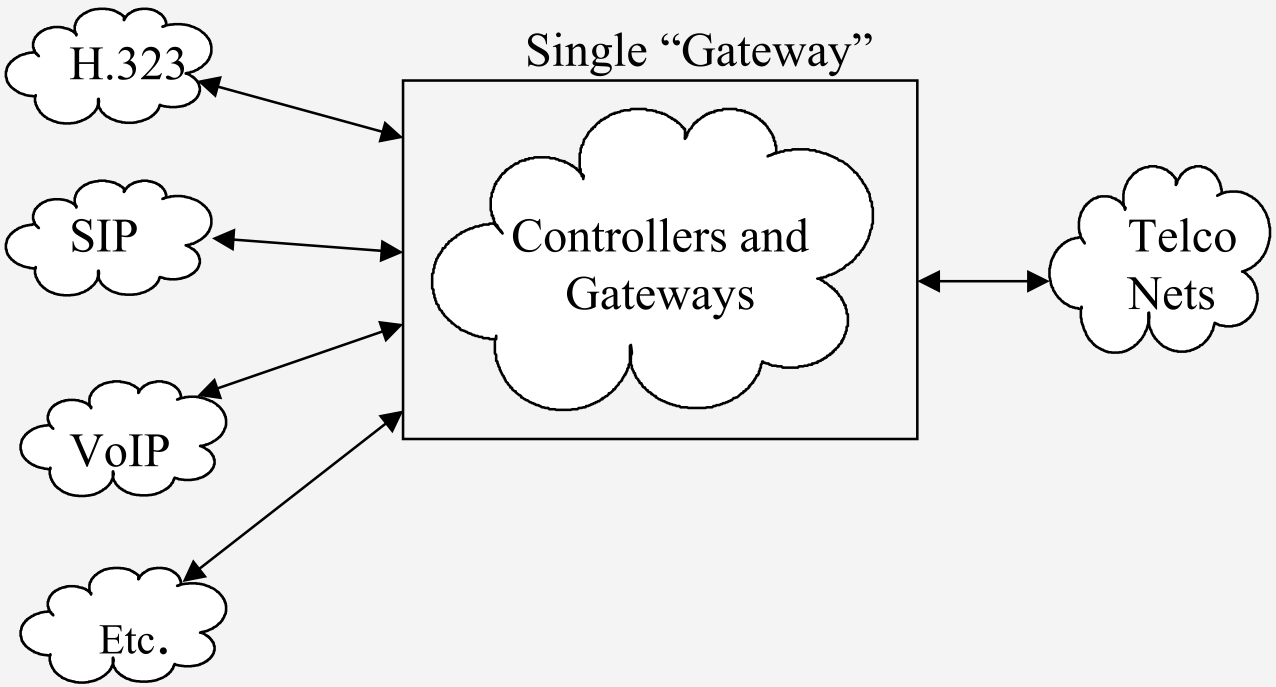
- ITU-T has developed a series of recommendations to support multimedia (audio, video and data) communications in packet based networks
- The H.323 system:
  - describes types and functions of H.323 terminals and other devices and their interactions
  - minimally requires only an audio stream to be supported
  - is THE standard for IP telephony
    - currently the most widely implemented control protocol for VoIP
    - interoperability with ISDN and PSTN networks

ENVIRONMENT OF H.323

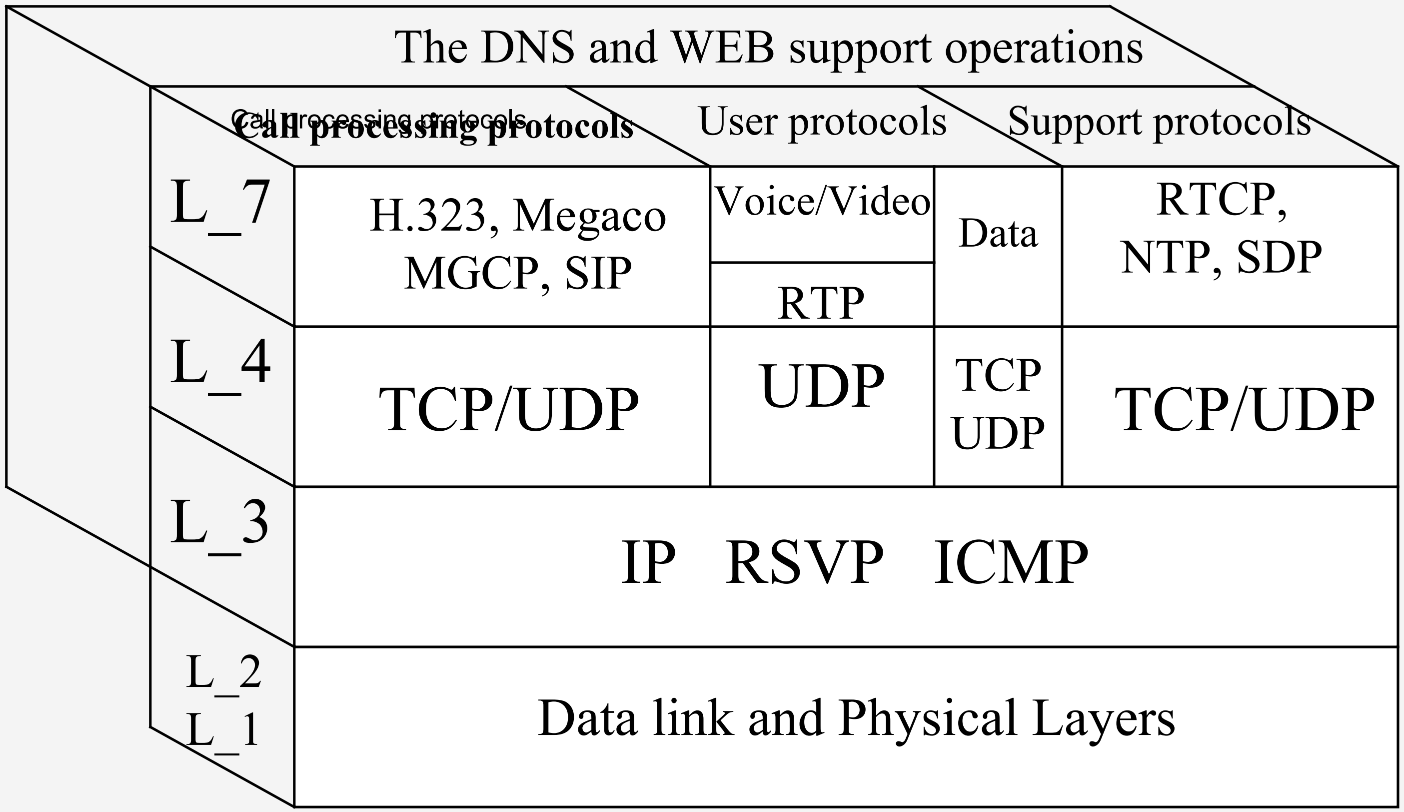


UNIFIED VIEW of the INTERNET TELEPHONY GATEWAYS

- **The Internet Gateway model must work both with "conventional" telephony protocol (ISDN User Part) and with packet telephony protocols (H.323)**



THE INTERNET CALL PROCESSING LAYERED MODEL



COMPONENTS AFFECTING VOICE QUALITY

- Codec
  - Analog-to-digital conversion
  - Digital-to-analog conversion
  - Coding algorithm (signal distorsion)

- Transport Network
  - Loss
  - Delay
  - Jitter (delay may vary)

SPEECH CODECS (Coder/Decoder)

Algorithm	Bit Rate
G.711 PCM	64 kbps
G.726 ADPCM	16,24,32,40 kbps
G.728 LD-CELP	16 kbps
G.729 CS-ACELP	8kbps
G.723.1	5.3/6.4 kbps

PCM: **Pulse Code Modulation**

ADPCM: **Adaptive Differential PCM**

LD-CELP: **Low Delay Code Excited Linear Prediction**

CS-ACELP: **Conjugate Structure Algebraic CELP**

SPEECH CODECS (cont.)

CODEC	G.723.1	G.729	G.729A
Bit rate	5.3/6.4 kbps	8 kbps	8 kbps
Frame size	30 ms	10 ms	10 ms
Processing delay	30 ms	10 ms	10 ms
Lookahead delay	7.5 ms	5 ms	5 ms
Frame length	20/24 bytes	10 bytes	10 bytes
DSP MI PS	16	20	10.5
RAM (16-bit words)	2200	3000	2000

Processing delay: 3-5 ms (G.728), 1 ms (G.726), 0.125 (G.711)

ENCODING OF AUDIO SAMPLES

- G.711 PCM
  - RTP payload is 160 bytes (160 samples every 20 ms)
- G.723.1
  - 5.3 kbps=20 bytes every 30 ms; 6.3 kbps=24 bytes every 30 ms
  - 4 bytes SID (Silence Insertion Descriptor) frames
- G.726
  - 16 kbps=40 bytes every 20 ms; 32 kbps=80 bytes every 20 ms
- G.728
  - 40 bytes every 20 ms
- G.729
  - 20 bytes (2 frames every 20 ms)
  - 2 bytes comfort noise frame

SAMPLE FORMAT (G.711)

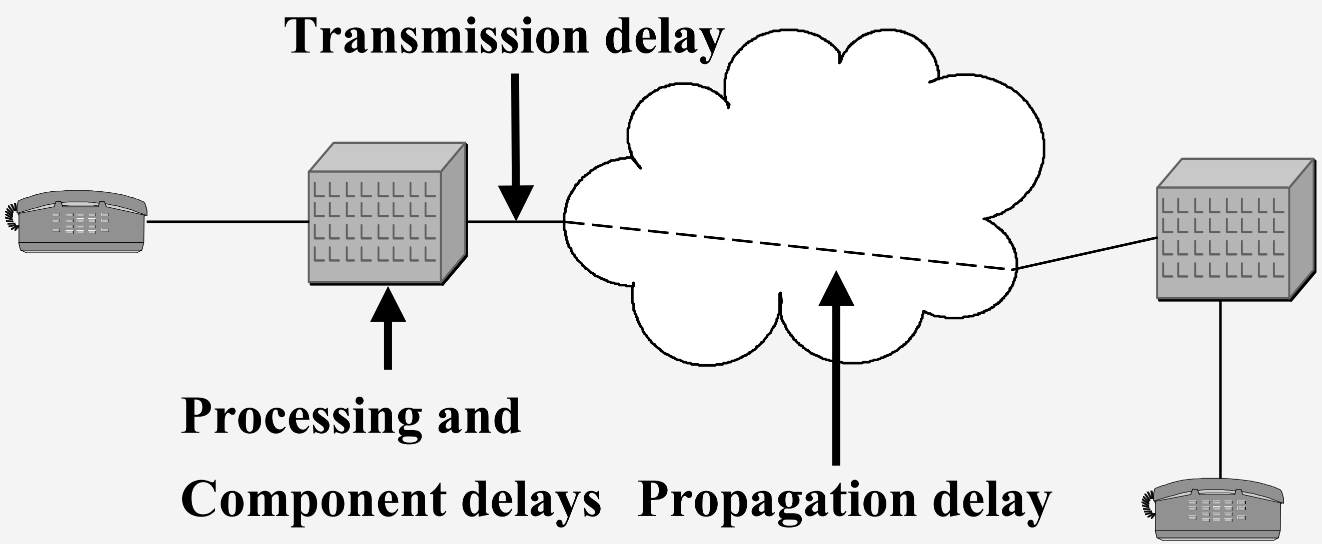
- Total frame length
  - 206 bytes using PPP encapsulation (WAN)
    - actually 82.4 kbps per connection
  - 218 bytes using Ethernet encapsulation (LAN)
    - actually 87.2 kbps per connection
  - with silence suppression: roughly 40 kbps

PPP	IP	UDP	RTP	G.711 payload	FCS
4	20	8	12	160	2

Ethernet	IP	UDP	RTP	G.711 payload	FCS
14	20	8	12	160	4

SOURCES OF FIXED DELAYS

- Processing delay
- Transmission delay
- Propagation delay
- Component delay



SOURCES OF VARIABLE DELAYS (JITTER)

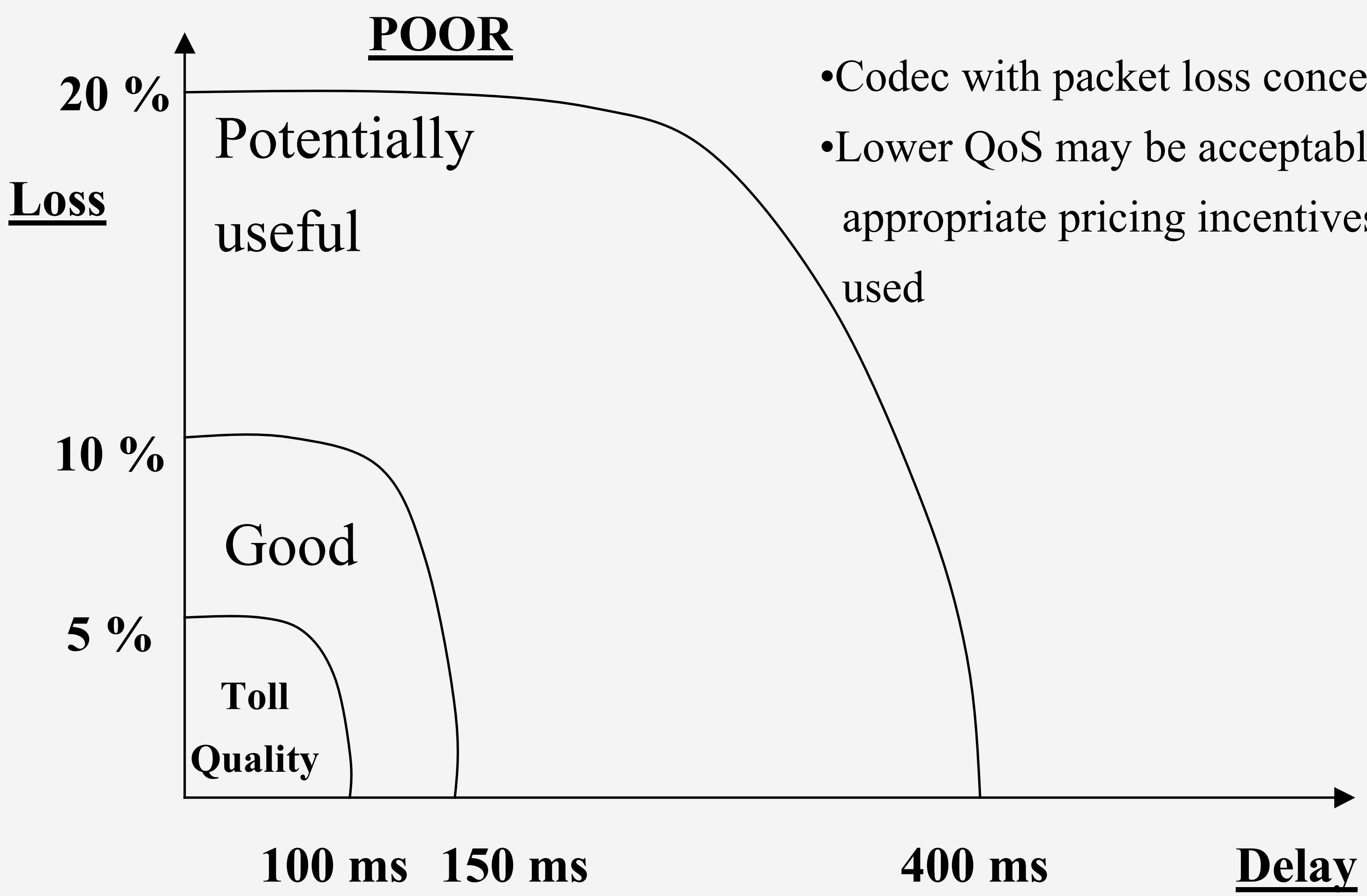
- Delay due to transport network is non-deterministic in nature
  - Variable processing delay
    - busy router or switch will take longer to lookup the address table
  - Queuing delay
    - network congestion
- Poor network conditions: average packet delay and packet delay variance (jitter) are high (75-300 ms)
- With jitter some packets arrive early while others arrive late
- Receive buffers can hide jitter at the cost of additional delay but beyond the playout point packets are effectively lost

PACKET LOSSES

- IP networks do not guarantee delivery of packets
- Stringent delay requirements => TCP cannot be used
- Packet loss is unavoidable but can be compensated for by codec loss-concealment schemes
- While single packet losses are easily managed, loss bursts, like those by Internet, can remarkably degrade the voice quality
- Relation between voice quality and losses depends on the coding algorithm
  - Forward error correction schemes have been proposed
  - FEC introduces additional delays which may cause the recovered packet to arrive too late (then lost anyway)
  - FEC => receiver buffer depth of several packets
  - G.711 has an optional feature for Packet Loss Concealment (PLC)

DELAY VS. LOSS: QoS MAPPING

- Unidirectional Internet delay and loss

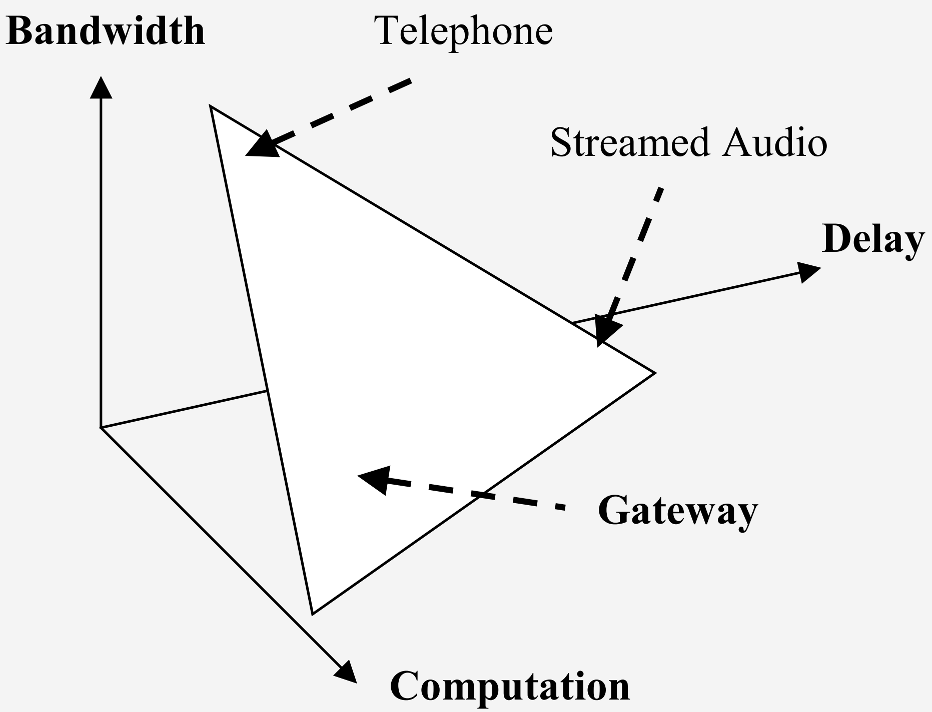


•Codec with packet loss concealment

•Lower QoS may be acceptable if appropriate pricing incentives are used

TRADE-OFF: QUALITY SURFACE

- Any solution for packetized audio can be characterized by its required
  - bandwidth
  - end-to-end delay
  - computational complexity



Gateway

•Internet audio apps

•low to moderate BW

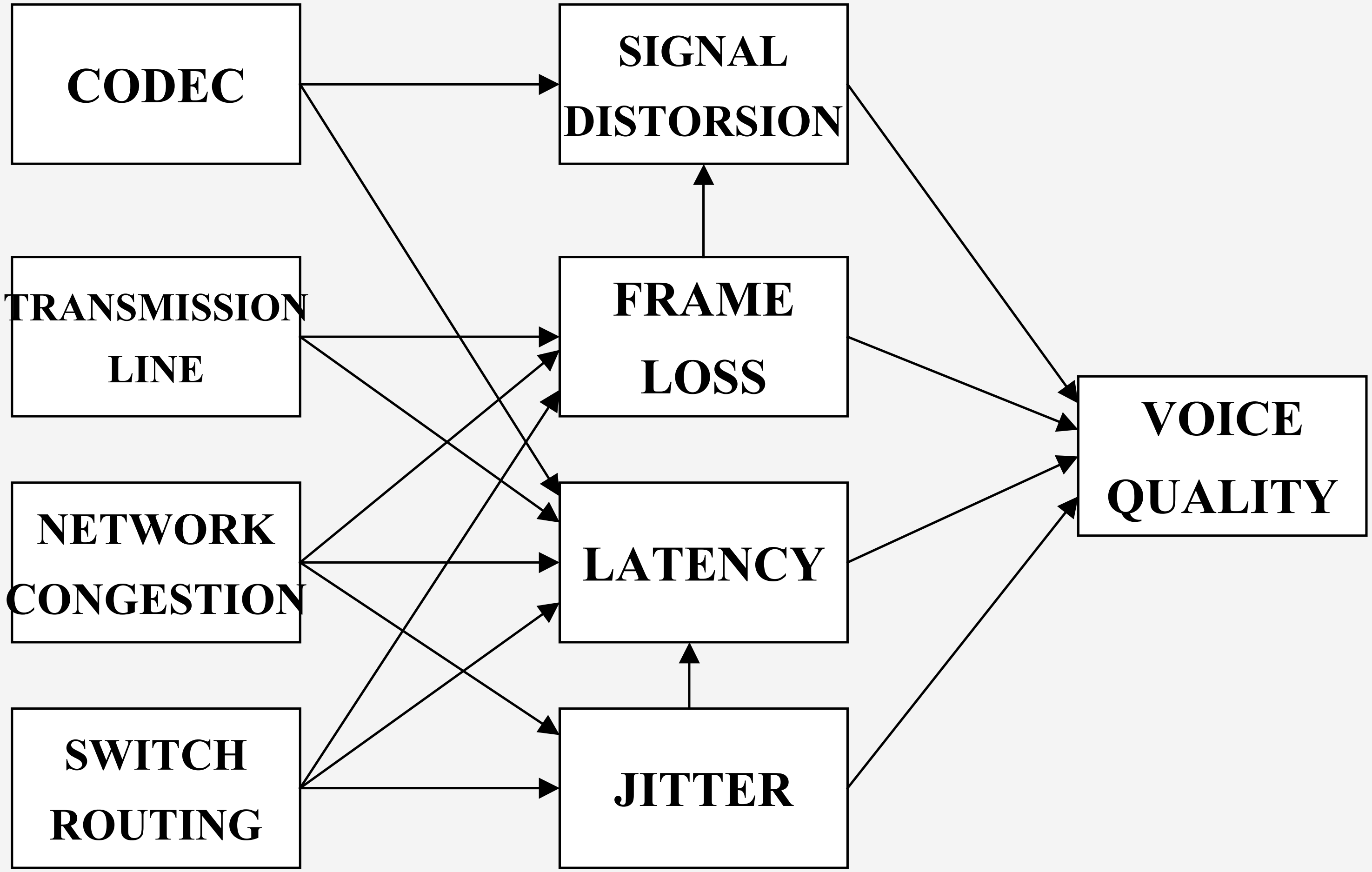
•intermediate delay

•high computational comlpexity

•based on DSP HW for

•lower-bit-rate coding

SUMMARY OF VOICE QUALITY



HOW TO TEST VOICE QUALITY

There are different ways to measure voice quality

- Mean Opinion Score (MOS)
- Perceptual Speech Quality Measure (PSQM & PSQM+)
- Perceptual Evaluation of Speech Quality (PESQ)
- Measuring Normalizing Blocks (MNB)
- Perceptual Analysis Measurements System (PAMS)
- E-Model

MEAN OPINION SCORE (MOS)

- ITU P.800
- Test performed by people
  - difficult to do regularly
- Subjective measure of voice quality
- Score ranges from 5 to 1
  - Excellent = 5; Good = 4; Fair = 3; Poor = 2; Bad = 1
  - Toll quality ≥ 4
- An objective test method is required

TCP/IP NETWORKS

- ✓ No QoS guarantee
- ✓ “Best Effort” delivery system
- ✓ Indeterminate level of packet loss, delay, out-of-sequence
- ✓ Multimedia applications have pushed for proposals for QoS
  - Priority marking (DiffServ)
  - Service marking (IPv4 ToS)
  - Label switching (MPLS)
  - Integrated services/RSVP

GENERAL QOS STRATEGIES

- Implementing QoS means separate delay sensitive traffic such as VoIP from other data traffic
  - all traffic is carried within IP packets
  - guaranteed vs. differentiated QoS
    - ✓ IP precedence in routers, gateways
  - identify RTP traffic by means of UDP range
  - identify the physical port where VoIP enters the net
  - route VoIP traffic differently
- QoS on link-by-link basis
  - scheduling algorithms
  - buffer management schemes
- QoS on end-to-end basis
  - IntServ/RSVP
  - DiffServ

CONCLUSIONS

- Designing a VoIP network is not easy…
- Mantaining acceptable voice quality requires careful planning
  - coding algorithm
  - packet loss
  - transmission delay
  - jitter
    - ☞ acceptable = one way delay ≤ 150 ms and packet loss ≤ 2%
- QoS can be implemented
  - Accurate Internet service model
  - appropriate scheduling algorithm
- Many, many,…., many network tests are necessary to verify that required performance can be achieved

PROGRAMMA

- VoIP sulla rete GARR fra le sedi di Cesena, Modena e Bologna
  Franco Callegati, Università di Bologna
- Introduzione del VoIP nella rete di centralini dell'Università di Bologna
  Valerio Mattioli, Università di Bologna
- Sperimentazioni di qualità di servizio su Internet
  CarlaRaffaelli , Università di Bologna

- IP Telephony l 'integrazione dei servizi
  Alessandro Boschetti , VEM Sistemi
- Soluzioni per la convergenza video voce dati e la telefonia IP
  Cisco Systems
- Dalla voce su IP alla telefonia su IP
  NextiraOne Italia
- Misura della qualità di una rete VoIP
  Massimo Bruni, ATS

THANK YOU FOR YOUR ATTENTION

mcasoni@unimo .it

http://www.dii.unimo.it/casoni

… suggestions are welcome