UNIVERSITA' DEGLI STUDI DI PARMA
Dipartimento di Ingegneria dell'Informazione

# Real-Time Multimedia Communications over IP

Luca Veltri
(mail.to: luca.veltri@unipr.it)

Corso di Reti di Telecomunicazioni C, a.a. 2008/2009

http://www.tlc.unipr.it/veltri



## Reti verticali o orizzontali

- Il rapido sviluppo di reti e servizi ha portato alla nascita di una moltitudine di soluzioni, definite "verticali" (reti dedicate)
  - **dedicate a singoli servizi**
  - **focalizzate sulle caratteristiche che si vogliono offrire e sulla gestibilità della singola applicazione**
  - **svantaggiose quando si desideri integrare due o più servizi**
    - maggiori costi, tempi, e complessità
    - minore qualità del servizio finale
  - **difficilmente sostenibili a lungo termine**
    - crescente complessità di gestione ed interlavoro, e delle inefficienze provocate dalla duplicazione di funzionalità simili
- Al contrario, sono possibili soluzioni "orizzontali" (reti integrate)
  - **condividono maggior parte degli elementi di rete per vasta gamma di servizi**
  - **permettono maggiore facilità nell'integrazione dei servizi**
  - **riduzione dei tempi e costi**
  - **separazione dei livelli di trasporto e controllo**





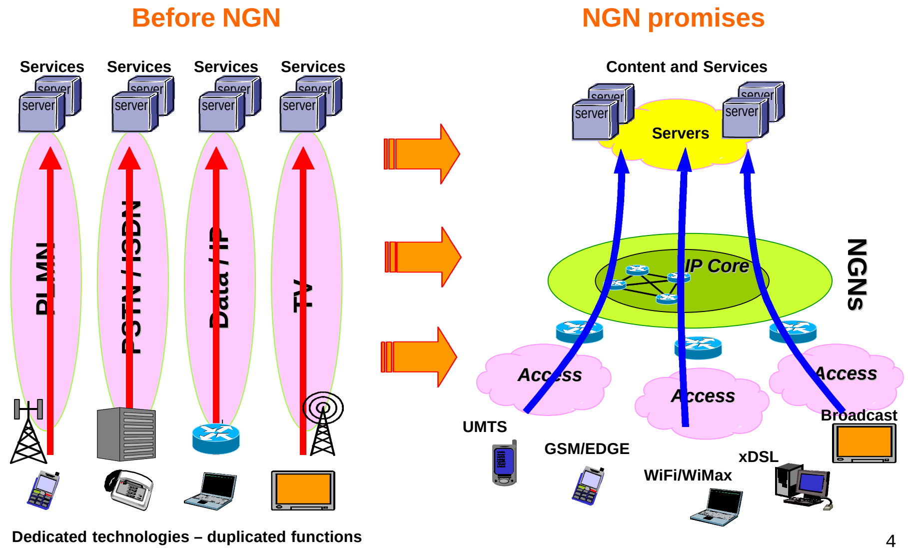
## Verso le NGN

- Next Generation Networks (Reti di nuova generazione)
  - **termine molto ampio per indicare l'insieme delle reti di TLC che sono previste per i prossimi 5-10 anni e che prevedono alcune evoluzioni chiave come la convergenza dei servizi e il trasporto su pacchetti IP**
    - una unica rete (IP) integrata che trasporta servizi differenti (voce, dati, video, etc.)
    - per questo motivo viene spesso utilizzato il termine All IP per descrivere le evoluzioni verso le NGN
  - **In pratica la rete fissa, quella mobile e quella broadcast tenderanno ad un processo di integrazione sempre più ampio fino alla completa convergenza sotto il paradigma IP**
- Questo concetto tiene conto delle nuove realtà nel mondo delle TLC quali:
  - **l'utilizzo sempre più intenso di Internet**
  - **diffusione di accesso a larga banda e wireless**
  - **la domanda crescente di multimedialità, pervasività e mobilità attraverso reti eterogenee**
  - **possibilità di introduzione di nuovi servizi, personalizzabili**
  - **nuovi terminali (PDA, smartphone, WiFi-phone, etc.)**





## Verso le NGN



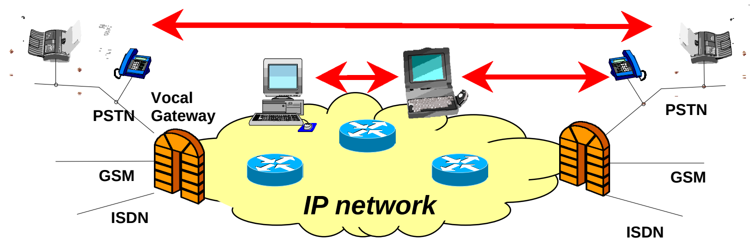
# Separazione dei livelli funzionali

- Separazione dell'architettura in livelli funzionali diversi e indipendenti:
  - **livello di servizio**
    - piattaforma software per la creazione di servizi a partire da logiche elementari comuni
  - **livello di controllo**
    - permette introduzione di nuovi servizi senza intervento su tale strato
    - se necessario prevede la centralizzazione logica delle informazioni di utente (identità, tipo di accesso, servizi sottoscritti e profilo di servizio)
  - **livello di rete (trasporto e accesso)**
    - il più possibile comune tra piano dati (traffico voce, dati, video) e controllo
    - basato su una infrastruttura IP comune per tutti i tipi di servizio
    - include sia le reti di accesso di varia natura (xDSL, Ethernet, WiFi, WiMAX, HDSPA, etc.), sia la rete core di trasporto (MPLS/SDH, etc.)





# Multimedia Communications: VoIP and IP Telephony




- **The two terms are often used as synonymous...**
  - "Voice over IP": support for voice communications using the IP Protocol over Internet/Intranet
    - **The phone path may be entirely or partially over IP network**
  - "IP telephony": usually when the service offered to the user is IP-based
    - **at least one terminal should be IP (the service is not PSTN)**
    - **is not only Telephony over IP: includes conferencing, video communication, new services**
  - "Internet telephony": the service directly over Internet (e.g. Skype)





# VoIP motivations

**Integration (network/transport aspect)**
- **the IP data traffic is dramatically growing with reference to voice traffic: in the backbones, the present/future highways are built for IP traffic!**
- **Only one network for all services (real-time and non-real-time)**
- **ISPs may become ITSP (Internet Telephony Service Provider)**

**Integration (operation and management aspect)**
- **an integrated Voice/Data network can represent a large cost saving from the operational point of view**
- **for example within an enterprise the same people could handle the Data and Voice network: the technology and the knowledge will be in common**

**Costs of network devices**
- **VoIP Gateways (GWs), VoIP PBXs and softswitches are cheaper**
- **Availability of very small GWs (i.e. 2 ports)**

**Services**
- **the Voice/Data integration over IP networks opens the door to new advanced services**

**Regulations**
- **Laws and authorities... deregulation**





# The main players

***The traditional manufactures***
**Ericsson, Nokia, Lucent, Nortel, ...**

***Data manufactures***
**Cisco, 3Com, ...**

***New Comers***
**Vocaltec, Dialogic, Digium, ...**

***PSTN Operators***
**AT&T, TelecomItalia, Vodafone, Wind, H3G, ...**

***New Comers***
**Skype, Google, ...**

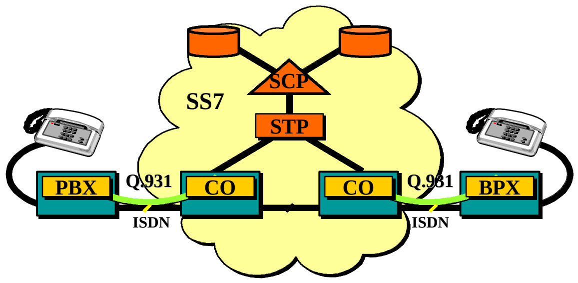
***ISP and ITSP***
**AOL, tiscali, fastweb, ...**

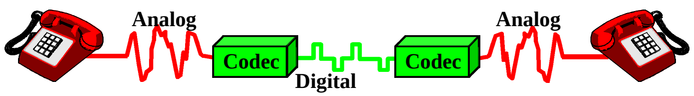
# The PSTN – a reminder



## Once upon a time… the PSTN

- Circuit switching
  - **PSTN (Public Switched Telephone Network), ISDN (Integrated Services Digital Network)**
- Main characteristics:
  - **circuit setup**
  - **fixed bandwidth**
  - **fixed delay**








## PSTN services

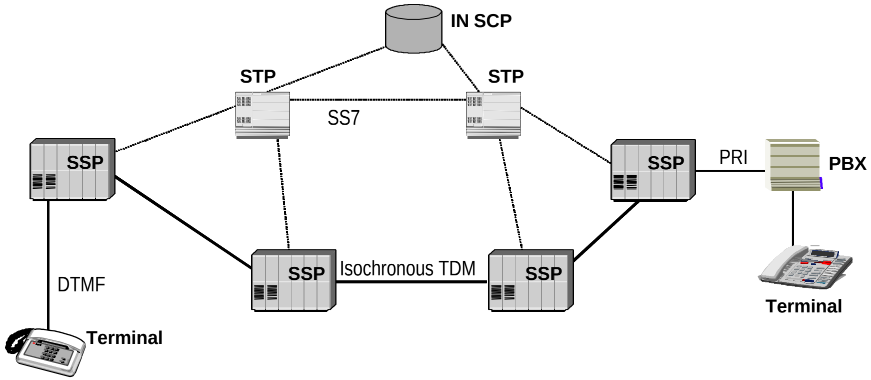
- Basic POTS (Plain Ordinary Telephone Service)
  - **augmented with additional (and value-added) services**
- Additional services
  - **Call back when free, Three-way calling, Voicemail, etc.**
  - **Number translation services**
    - (800-xxx) freephone, local call, premium call
  - **Virtual Private Networking**
    - aimed at corporates, integration with PBX
- Other value-added aervices
  - **Calling card**
- The voice community has understood that the voice business lies in advanced services, not just the bearer service of voice connectivity (POTS)





## PSTN Generic Architecture

- Dumb terminals (don't run IP - so can't be hosts)
- Two networks - a connection-oriented one for TDM voice and a separate connectionless network for signalling
  - **Service switching point (SSPs)**
  - **signal transfer point (STPs)**
  - **Service Control Point (SCPs)**



# PSTN characteristics

- The Public Switched Telephone Network is a global network which has evolved over the decades to be a highly-engineered real-time system of great reliability
- It is, however a network which is specialised to do only a few things very well – highly vertically integrated
- The IP challenge is to replicate and transcend the engineering requirements of PSTN-grade voice in an open and extensible IP-based architecture





# SS7 performance

- Less than 1 in 10 million messages lost (<= 1 in 10E+7)
- Less than 1 in 10 billion messages out of sequence, or duplicated (<= 1 in 10E+10)
- Less than 1 in a billion message errors should go undetected (<= 1 in 10E+9 undetected)
- End-to-end route availability more than 99.998% (less than ten minute downtime per year)
- Example end-to-end permitted delay for an “average” country: simple message = 390 ms; complex message = 600 ms
- This is quite challenging for an IP network

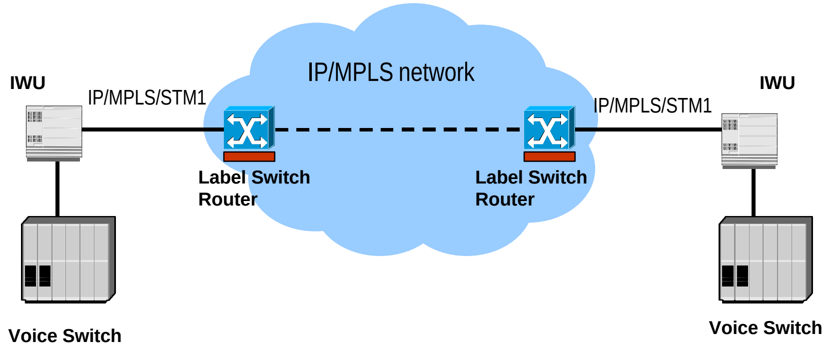


# Different VoIP views



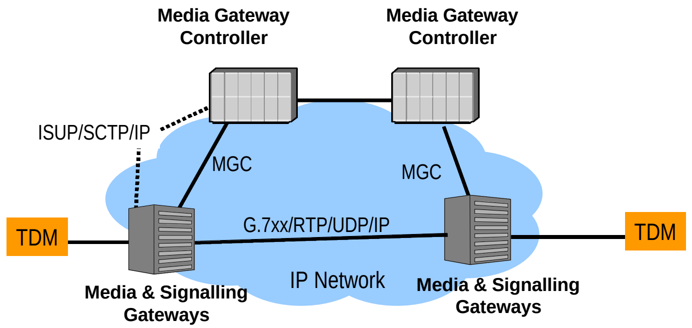
# Voice Transport

- Voice DS0 and signalling flows are carried transparently on PVCs (Permanent Virtual Circuits)
  - **e.g. through MPLS LSPs (Label Switched Paths)**
- This view was most popular a few years ago



# Voice peering to the TDM world

- The IP network emulates a PSTN switching peer
- It is the way IP-carriers are going to handle traditional TDM voice customers
- Interconnects partner services and enterprise (PBX) voice

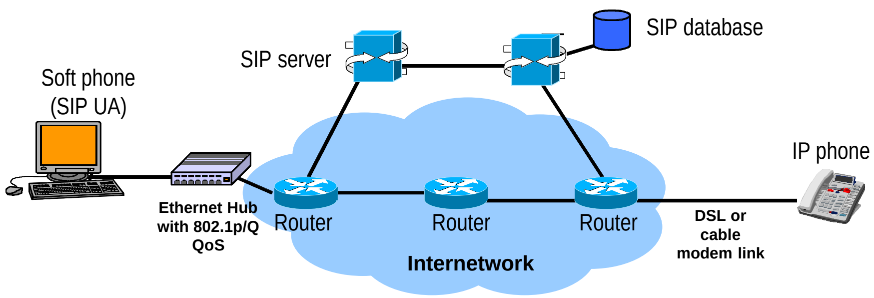







# IETF view of voice - Session Initiation Protocol

- Call model runs in intelligent end-devices
- Network Services run on network servers
- SIP protocol makes it all work (leveraging other protocols)
- Voice is special case of multiparty, multimedia session
- Similar to H.323 but faster & more flexible & within IETF paradigm






# Voice Transport over IP



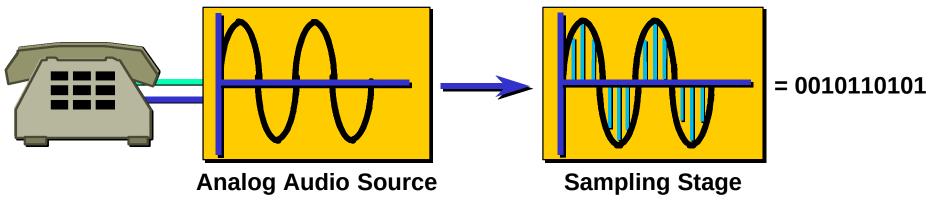
# Introduction: Voice packetization

- Once voice has been digitized, a number of processing and transmission steps occur before it reaches the eventual user, at the other end of the transmission link
  - **digitization**
  - **coding, compression methods**
  - **streaming and transmission**
- Most of these steps result in delay, delay variation and loss

# Analog to Digital

**Pulse Code Modulation—"Nyquist Theorem" (Sample at twice the Frequency)**

Voice Bandwidth =
300 Hz to 3400 Hz




**Codec Technique**

= Sample
8 kHz (8,000 8-bit samples/Sec)





# Voice Transport Requirements

- The impairments in the transport of Voice over IP are:
  - **loss**
  - **delay**
  - **delay variation (jitter)**
- This transport impairments impact the perceived voice quality
- A subjective measurement of voice quality is still used, called Mean Opinion Score (MOS)





# Speech quality and Tradeoff

- Each CODEC provides a certain quality of speech
  - **The quality of transmitted speech is a subjective response of the listener**
- A common benchmark used to determine the quality of sound produced by specific CODECs is the mean opinion score (MOS)
  - **wide range of listeners judge the quality of a voice sample (corresponding to a particular CODEC) on a scale of 1 (bad) to 5 (excellent)**
  - **the scores are averaged to provide the mean opinion score for that sample**
- From a financial standpoint it might seem logical to convert all calls to low-bit rate CODECs
- One of the drawbacks compressing voice is signal distortion due to multiple encodings (called tandem encodings)
  - **For example, when a G.729 voice signal is tandem encoded three times, the MOS score drops from 3.92 (very good) to 2.68 (unacceptable)**
- Another drawback is CODEC-induced delay with low bit-rate CODECs





# Mean Opinion Score (MOS)

| Rating | Speech Quality | Level of Distortion |
|---|---|---|
| 5 | Excellent | Imperceptible |
| 4 | Good | Just perceptible, not annoying |
| 3 | Fair | Perceptible, slightly annoying |
| 2 | Poor | Annoying but not objectionable |
| 1 | Unsatisfactory | Very annoying, objectionable |

**MOS of 4.0 = Toll Quality**

# ITU-T G-series Voice coding

- The ITU-T G-series recommendations standardized different coding techniques
- The most popular coding standards are:
  - **G.711---Describes the 64-kbps PCM voice coding technique. In G.711, encoded voice is already in the correct format for digital voice delivery in the public switched telephone network (PSTN) or through PBXes**
  - **G.723.1---Describes a compression technique that can be used for compressing speech or audio signal components at a very low bit rate as part of the H.324 family of standards. This CODEC has two bit rates associated with it: 5.3 and 6.3 kbps. The higher bit rate is based on ML-MLQ technology and provides a somewhat higher quality of sound. The lower bit rate is based on CELP and provides system designers with additional flexibility**
  - **G.726---Describes ADPCM coding at 40, 32, 24, and 16 kbps**
  - **G.728---Describes a 16-kbps low-delay variation of CELP voice compression**
  - **G.729---Describes CELP compression where voice is coded into 8-kbps streams. There are two variations of this standard (G.729 and G.729 Annex A) that differ mainly in computational complexity; both provide speech quality similar to 32-kbps ADPCM**





# ITU-T G-series Voice coding

| Compression Method | Bit Rate (kbps) | Framing Size (ms) | MOS Score |
|---|---|---|---|
| **G.711 PCM** | **64** | **0.125** | **4.1** |
| **G.726 ADPCM** | **32** | **0.125** | **3.85** |
| **G.728 LD-CELP** | **16** | **0.625** | **3.61** |
| **G.729 CS-ACELP** | **8** | **10** | **3.92** |
| **G.729 x 2 Encodings** | **8** | **10** | **2.68** |
| **G.729 x 3 Encodings** | **8** | **10** | **2.68** |
| **G.729a CS-ACELP** | **8** | **10** | **3.7** |
| **G.723.1 MP-MLQ** | **6.3** | **30** | **3.9** |
| **G.723.1 ACELP** | **5.3** | **30** | **3.65** |

Table taken from Cisco Systems:
http://www.cisco.com/univercd/cc/td/doc/product/software/ios113ed/113t/113t_1/voip/voipover.htm





# G.711 - 64000 bit/s

| frame [bytes] | frame [ms] | frames/ packet | packets/ sec | payload [bytes] | packet [bytes] | bits/sec | % optimal | latency [ms] |
|---|---|---|---|---|---|---|---|---|
| 240 | 30 | 1 | 33,33 | 240 | 280 | 74667 | 117% | 30 |
| 240 | 30 | 2 | 16,67 | 480 | 520 | 69333 | 108% | 60 |
| 240 | 30 | 3 | 11,11 | 720 | 760 | 67556 | 106% | 90 |
| 240 | 30 | 4 | 8,33 | 960 | 1000 | 66667 | 104% | 120 |
| 240 | 30 | 5 | 6,67 | 1200 | 1240 | 66133 | 103% | 150 |
| 240 | 30 | 6 | 5,56 | 1440 | 1480 | 65778 | 103% | 180 |
| 240 | 30 | 7 | 4,76 | 1680 | 1720 | 65524 | 102% | 210 |
| 240 | 30 | 8 | 4,17 | 1920 | 1960 | 65333 | 102% | 240 |
| 240 | 30 | 9 | 3,70 | 2160 | 2200 | 65185 | 102% | 270 |
| 240 | 30 | 10 | 3,33 | 2400 | 2440 | 65067 | 102% | 300 |
| 240 | 30 | 11 | 3,03 | 2640 | 2680 | 64970 | 102% | 330 |
| 240 | 30 | 12 | 2,78 | 2880 | 2920 | 64889 | 101% | 360 |





# G723.1 - 6300 bit/s

| frame [bytes] | frame [ms] | frames/ packet | packets/ sec | payload [bytes] | packet [bytes] | bits/sec | % optimal | latency [ms] |
|---|---|---|---|---|---|---|---|---|
| 24 | 30 | 1 | 33,33 | 24 | 64 | 17067 | 267% | 30 |
| 24 | 30 | 2 | 16,67 | 48 | 88 | 11733 | 183% | 60 |
| 24 | 30 | 3 | 11,11 | 72 | 112 | 9956 | 156% | 90 |
| 24 | 30 | 4 | 8,33 | 96 | 136 | 9067 | 142% | 120 |
| 24 | 30 | 5 | 6,67 | 120 | 160 | 8533 | 133% | 150 |
| 24 | 30 | 6 | 5,56 | 144 | 184 | 8178 | 128% | 180 |
| 24 | 30 | 7 | 4,76 | 168 | 208 | 7924 | 124% | 210 |
| 24 | 30 | 8 | 4,17 | 192 | 232 | 7733 | 121% | 240 |
| 24 | 30 | 9 | 3,70 | 216 | 256 | 7585 | 119% | 270 |
| 24 | 30 | 10 | 3,33 | 240 | 280 | 7467 | 117% | 300 |
| 24 | 30 | 11 | 3,03 | 264 | 304 | 7370 | 115% | 330 |
| 24 | 30 | 12 | 2,78 | 288 | 328 | 7289 | 114% | 360 |

# G723.1 - 5333 bit/s

| frame [bytes] | frame [ms] | frames/ packet | packets/ sec | payload [bytes] | packet [bytes] | bits/sec | % optimal | latency [ms] |
|---|---|---|---|---|---|---|---|---|
| 20 | 30 | 1 | 33,33 | 20 | 60 | 16000 | 300% | 30 |
| 20 | 30 | 2 | 16,67 | 40 | 80 | 10667 | 200% | 60 |
| 20 | 30 | 3 | 11,11 | 60 | 100 | 8889 | 167% | 90 |
| 20 | 30 | 4 | 8,33 | 80 | 120 | 8000 | 150% | 120 |
| 20 | 30 | 5 | 6,67 | 100 | 140 | 7467 | 140% | 150 |
| 20 | 30 | 6 | 5,56 | 120 | 160 | 7111 | 133% | 180 |
| 20 | 30 | 7 | 4,76 | 140 | 180 | 6857 | 129% | 210 |
| 20 | 30 | 8 | 4,17 | 160 | 200 | 6667 | 125% | 240 |
| 20 | 30 | 9 | 3,70 | 180 | 220 | 6519 | 122% | 270 |
| 20 | 30 | 10 | 3,33 | 200 | 240 | 6400 | 120% | 300 |
| 20 | 30 | 11 | 3,03 | 220 | 260 | 6303 | 118% | 330 |
| 20 | 30 | 12 | 2,78 | 240 | 280 | 6222 | 117% | 360 |





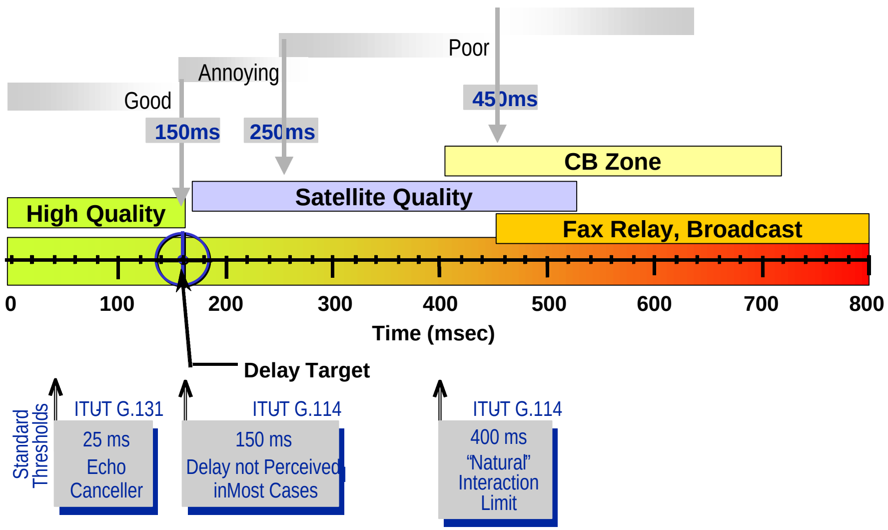
# Voice coding (ITU-T + ETSI)

| Standards Body | ITU-T | ITU-T | ITU-T | ITU-T | ITU-T | ETSI | ETSI | ETSI |
|---|---|---|---|---|---|---|---|---|
| Recc. | G.726 | G.728 | G.729 | G.729A | G.723 | GSM (FR) | GSM (HR) | GSM (EFR) |
| Year | 1990 | 1992 | 1995 | 1996 | 1995 | 1987 | 1994 | 1995 |
| Bit Rate Kbit/s | 16-40 | 16 | 8 | 8 | 6.3 & 5.3 | 13 | 5.6 | 12.2 |
| Quality | ≤PSTN | PSTN | PSTN | PSTN | ≤ PSTN | < PSTN | =GSM | PSTN |
| Frame Size | 0.125 ms | 0.625 ms | 10 ms | 10 ms | 30 ms | 20 ms | 20 ms | 20 ms |
| Look Ahead | 0 | 0 | 5 ms | 5 ms | 7.5 ms | 0 | 4.4 ms | 0 |
| MOS | 4.1 | 4.0 | 3.9 | < 3.9 | 3.8 | 3.5 | 3.5 | > 4 |





# Cumulative One-Way Transmission Delay

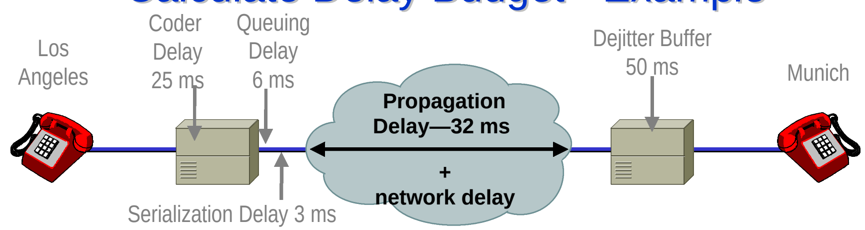



**ITU's G.114 Recommendation = 0 – 150 msec 1-way delay**





# Calculate Delay Budget - Example




| | Fixed Delay | Variable Delay |
|---|---|---|
| Coder Delay G.729 (10 ms + 5 ms look ahead) | 15 ms | |
| Packetization Delay—Included in Coder Delay | ~ 0 | |
| Queuing Delay 64 kbps Trunk | | 6 ms |
| Packet Serialization Delay 64 kbps Trunk (*) | 3 ms | |
| Propagation Delay | 32 ms | |
| Network Delay | | ? |
| Dejitter Buffer | 50 ms | |
| Total | 110 ms | |

Delay budget:
if the maximum tolerable end to end delay is 150 ms, the fixed network delay must be < 40 ms the variable network delay must be < dejitter delay

(*) compressed packet header (8 bytes) and 20 byte payload

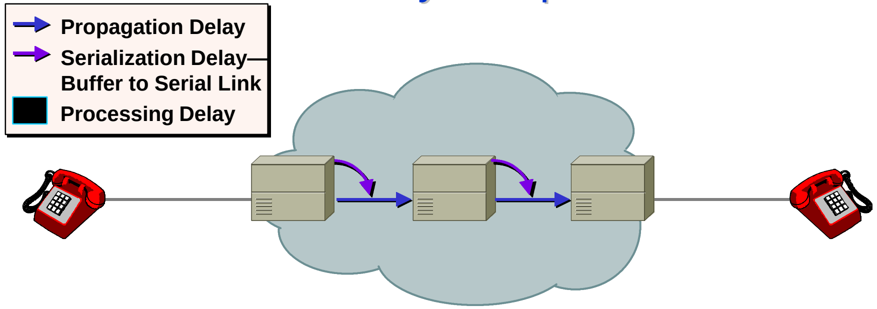
# End-to-end delay components

- End-to-end delay is given by:
  - **Fixed component:**
    - voice processing delay (voice coding and packetization, with fixed size packets)
    - propagation delay
    - serialization delay (fixed size packets)
  - **Variable component**
    - Delay introduced by the network (queuing delay, packet processing)
    - Variable packet sizes
- “Dejitter” buffers compensate the variable component





# Fixed Delay Components




- Processing
  - Coding/compression/decompression/decoding
  - Packetization
- Propagation—six microseconds per kilometer
- Serialization





# Serialization Delay

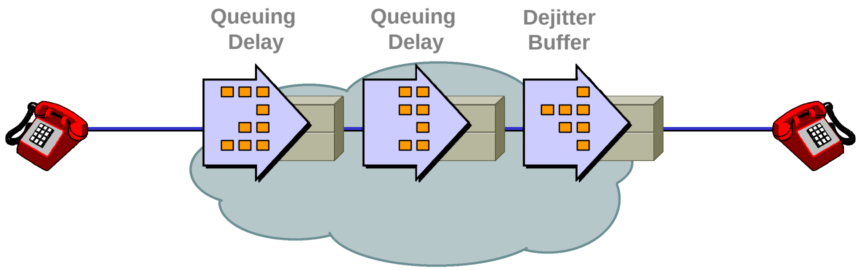
**Frame Size**

| Link Speed | 1 Byte | 64 Bytes | 128 Bytes | 256 Bytes | 512 Bytes | 1024 Bytes | 1500 Bytes |
|---|---|---|---|---|---|---|---|
| 56kbps | 143us | 9ms | 18ms | 36ms | 72ms | 144ms | 214ms |
| 64kbps | 125us | 8ms | 16ms | 32ms | 64ms | 128ms | 187ms |
| 128kbps | 62.5us | 4ms | 8ms | 16ms | 32ms | 64ms | 93ms |
| 256kbps | 31us | 2ms | 4ms | 8ms | 16ms | 32ms | 46ms |
| 512kbps | 15.5us | 1ms | 2ms | 4ms | 8ms | 16ms | 23ms |
| 768kbps | 10us | 640us | 1.28ms | 2.56ms | 5.12ms | 10.24ms | 15ms |
| 1536kbs | 5us | 320us | 640us | 1.28ms | 2.56ms | 5.12ms | 7.5ms |



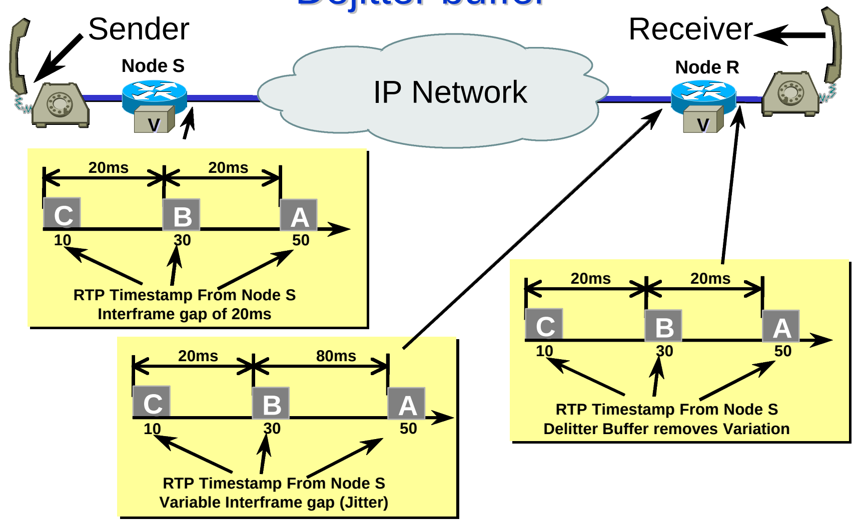


# Variable Delay Components




- Queuing delay
- Variable packet sizes

# Dejitter buffer

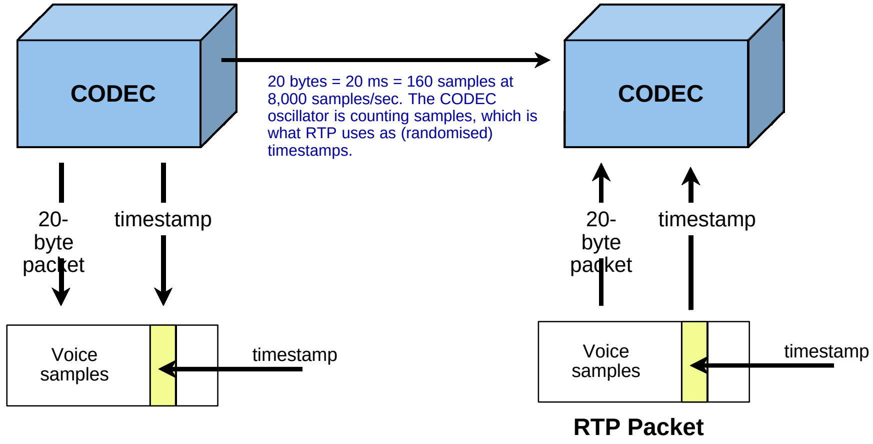



- Large packets can cause playback buffer underrun, resulting in slight voice degradation
- Jitter or playback buffer can accommodate some delay/delay variation





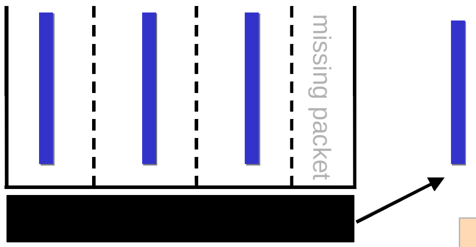
# The mechanism of RTP "timing"








# VoIP and Packet Loss




| *Quality* | *Loss (%)* | *Jitter (ms)* |
|---|---|---|
| Perfect | 0 | 0 |
| Good | 3 | 75 |
| Medium | 10 | 125 |
| Poor | 25 | 225 |

- The friendly "retransmission" (e.g. TCP) is of no use in the voice world… late is as good as never
- VoIP coding is tolerant to occasional loss
- Complex strategy in algorithm can interpolate lost packets based on context





# QoS parameters in the global IP network

- As an IP packet enters a Service Provider IP network, the major impairments it can suffer are the following:
  - **Unavailability of service**
    - In the case of a fibre breakage, or equipment failure, throughput can vanish until recovery mechanisms complete.
  - **Delay**
    - Packets are delayed when they are processed (e.g. through encryption gateways), when they encounter non-empty queues in routing devices, and in playout buffers (jitter-removal).
  - **Jitter**
    - Packets accumulate jitter when they encounter varying router queue occupancies on their path through the network . As a result, they incur a different overall delay than that of their predecessors or successors. This skews the timing relationships between successive packets.
  - **Packet Loss**
    - Packets are lost when they encounter a queue in a router which is completely full, or when they are subject to policy-based discard - e.g. they are out-of-profile of their SLA.

## Example of expected QoS parameters

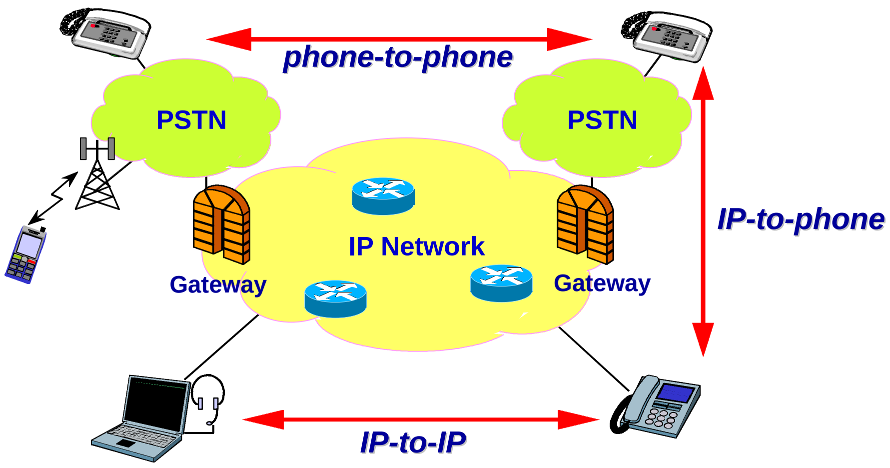
- Availability:
  - 99.99% = 4 mins downtime/month; 99.999% = 25 seconds downtime/month
  - Enterprise voice: 99.999% desktop-to-desktop. For enterprise customers, planned downtime in slack periods may be permissable
  - Public telephony voice: 99.999% Media Gateway-to-Media Gateway. No planned outages in service
- Resilience:
  - Node or link failure recovered within 50 ms to avoid impairments to voice calls and Internet TV
- Delay, or Latency (maximum):
  - 70 ms PoP-PoP network RTT delay for IP transit services
  - 100 ms desktop to network to server RTT for transactional services
  - 150 ms network and application one-way delay for voice/multimedia (G.114)
- Jitter:
  - Less than 2% of one-way delay as seen by the customer
- Packet loss:
  - < 1.5% “usually” but we need to specify time-dynamics
  - TCP traffic (web, email) is resilient to episodic packet loss (triggering rate-adaptive behaviour)
  - RTP/UDP traffic (voice, video) is resilient to small packet loss, and brittle to larger packet loss



# VoIP and IP Telephony



## VoIP Scenarios

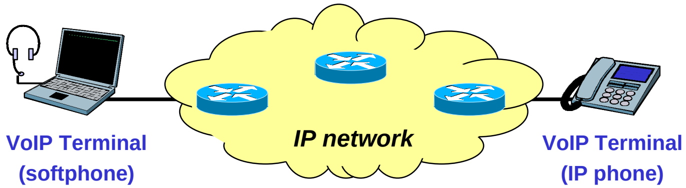






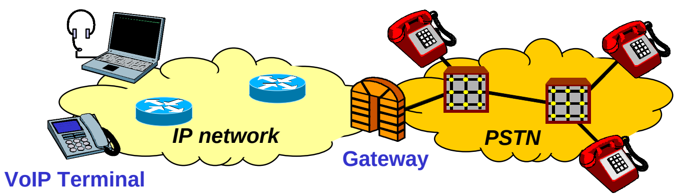

## VoIP Scenarios: IP-to-IP

- IP-to-IP
  - **The PC-to-PC architecture was historically the first to be implemented (Vocaltec, “Internet Phone”, 1995)**
  - **Need a host with audio card, mic, + telephony software (Netmeeting, etc)**
  - **Examples: Internet phone, corporate telephony service**
  - **Optional services: video, chat, data sharing…**



# Scenarios: IP to Phone

- IP-to-PSTN
  - ➢ **Need a gateway that connects IP network to phone network (also called IWU InterWorking Unit)**
  - ➢ **It realizes the interworking: voice signal, service control (signaling, routing, ...)**

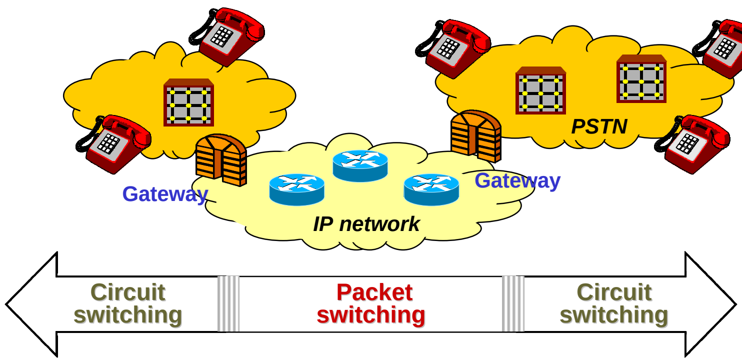







# VoIP Scenarios

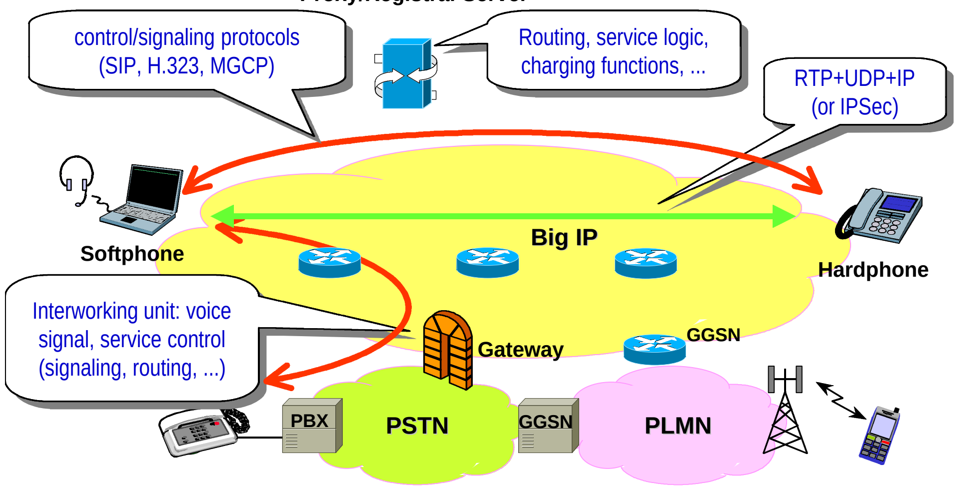
- PSTN-to-IP-to-PSTN (phone-to-phone via an IP network)
  - ➢ **Need more gateways**
  - ➢ **The IP network could be dedicated intranets or the Internet**
  - ➢ **The phone networks could be intra-company PBXs or the carrier switches**








# Basic network and control components








# Basic network and control components (cont.)

- User/data plane
  - ➢ **Codec/decodec (audio, video)**
  - ➢ **RTP/RTCP/UDP/IP**
  - ➢ **IP Multicast**
- Control plane
  - ➢ **Signaling - Call Setup, Tear-down**
    - Session Initiation Protocol (SIP) (IETF)
    - H323 (ITU-T standard includes H.225.0, H.245, Q.931)
    - others (e.g. MGC..)
  - ➢ **Gateways (H.323-to-PSTN, SIP-to-PSTN, etc.)**
  - ➢ **RTSP - Media Server Control**
    - VCR like controls
      - – Fast-forward, play, rewind, pause, record
  - ➢ **AAA (Authentication, Authorization, Accounting)**
    - RADIUS
    - Diameter

# Basic network and control components (cont.)

- QoS: Resource Reservation
  - **RSVP**
  - **Differntiated Services**
- Others:
  - **SDP - Session Descrition**
    - Describes the characteristics of a multimedia session such as time, format, address/ports, purpose, etc.
    - SDP may distributed using SAP, e-mail or the Web
  - **Call processing scripts**
    - Specifies behaviour of a server or terminal in responding to incoming and outgoing calls
  - **SAP - Session Announcement**
    - announcement protocol for multimedia sessions
    - uses IP multicast to a well-known address/port
    - SAP/SDP are used to advertise multimedia session information to a set of potential participants





# VoIP protocol stack

| *media* | *session control* |
|---|---|
| audio/video codec | H.323, SIP, SDP, alt. |
| RTP/RTCP | H.323, SIP, SDP, alt. |
| UDP | TCP or UDP |
| IP ( + QoS/RSVP/Diffserv ) | |
| layer 2 technologies (MPLS, ATM, SDH, ...) | |





# The Gateway

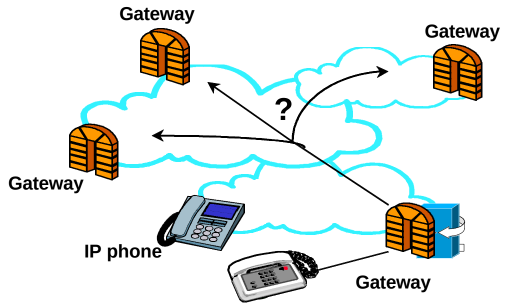
- It is the IWU (inter-working unit) between an IP network and a Switched Circuit Network (SCN) (ex.: PBX, ISDN, POTS)
- A Gateway takes a phone call from the SCN, digitizes it if not already digital, compresses and packetizes it into IP datagrams
- Reverse for the other way (full-duplex)
- It acts as gateway for both:
  - **user plane (data)**
  - **control plane (signaling)**
- Gateways are often based on workstation platforms





# Gateway Location Problem

- When
  - **PC to PSTN Calling**
  - **PSTN to IP to PSTN calling**
- Must locate a remote IP-PSTN gateway
  - **Need to “learn” about remote gateways**
  - **Gateways can be also within another administrative domain**



# The phone works — why bother with VoIP?

*User prospective*

- Voice over IP
  - **variable compression/quality**
  - **low cost calls**
  - **no local access fees**
- security through encryption
- caller, talker identification
- better user interface
- video, whiteboard, and new services

*Carrier prospective*

- voice and silence suppression → low traffic
- integration of data and voice switching (only one network)
- operational advantages
- new services





# VoIP technical issues

- Large delay and delay jitter
- Packet length
- Voice compression and silence suppression
- Lost of packets
  - **replacement of lost packets by silence, extrapolate previous waveform, etc.**
- Echo cancellation
- Address translation: Phone # to IP; directory servers
- Telephony signaling: Different PBXs may use different signaling methods
- Multiplexing
  - **Subchannel multiplexing → Multiple voice calls in one packet**
- Security
  - **Firewalls may not allow incoming IP traffic**
  - **Insecurity of Internet**
- Charging
  - **Really, not a technical issue**





# Standardization

- Main IETF Working Groups (WGs)
  - **AVT - Audio/Video Transport**
    - formed to specify a protocol for real-time transmission of audio and video (RTP)
    - the associated profile for audio/video conferences and payload formats (e.g. MPEG-4, PureVoice)
  - **SIGTRAN - Signaling Transport**
    - transport of packet-based PSTN signaling (such as Q.931 or SS7 ISUP) over IP Networks, between IP nodes such as a Signaling Gateway and Media Gateway Controller or Media Gateway
  - **MEGACO - Media Gateway Control**
    - develops an architecture for controlling Media Gateways (MGW) from external control elements such as a MGW Controller
    - ex MEGACO/H.248, ex MGCP
  - **MMUSIC - Multiparty Multimedia Session Control**
    - to develop Internet standards track protocols to support Internet teleconferencing sessions





# Standardization (cont.)

- Main IETF WGs (cont.)
  - **IPTEL - IP Telephony**
    - to develop protocols for Internet telephony (ex. signaling and capabilities exchange)
    - e.g. specification of the Call Processing syntax, specification of the service model, protocols, etc.
  - **SIP - Session Initiation Protocol**
    - Sviluppo ed estensione di SIP
- Others
  - **ITU-T**
  - **3GPP**
  - **IEEE**