



# Università degli Studi di Bergamo



## DIPARTIMENTO DI INGEGNERIA E SCIENZE APPLICATE

# RETI INTERNET MULTIMEDIALI

---

## VoIP

### Introduzione e Codifica della voce

*Il documento è adattato da materiale cortesemente messo a disposizione dai Prof. Stefano Paris, Vittorio Trecordi, Antonio Capone e Flaminio Borgonovo*

# **VOIP**

---

## **Introduzione**

# Generalità

---

- Voice-over-IP è una modalità di gestione delle chiamate telefoniche utilizzando la rete Internet, invece della tradizionale rete telefonica (PSTN)
  - La rete Internet è una rete a commutazione di pacchetto, mentre la rete PSTN è una rete a commutazione di circuito
- In VoIP, la voce viene interamente inviata in formato digitale
  - Nella rete PSTN sull'*ultimo miglio* avviene tutt'oggi una trasmissione della voce analogica, mentre il resto della rete è interamente digitale

# Generalità

---

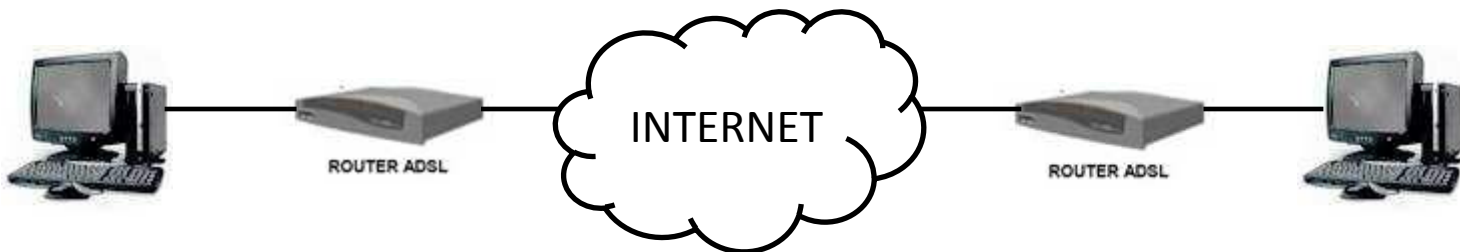
- Esistono due diverse tipologie di operatori VoIP
  - Operatori over-the-top VoIP (OTT VoIP)
  - Operatori di rete che adottano tecnologia VoIP
- Diverse tipologie di chiamate telefoniche possono essere instaurate
  - VoIP-VoIP
  - VoIP-PSTN/PSTN-VoIP
  - VoIP-Rete Mobile/Rete Mobile-VoIP

# Generalità

## VoIP-VoIP – VoIP puro

---

- La comunicazione è tra due terminali intrinsecamente VoIP (es. PC o telefono VoIP)
- La rete PSTN è completamente bypassata
- I pacchetti voce sono instradati tra sorgente e destinazione utilizzando interamente la rete Internet

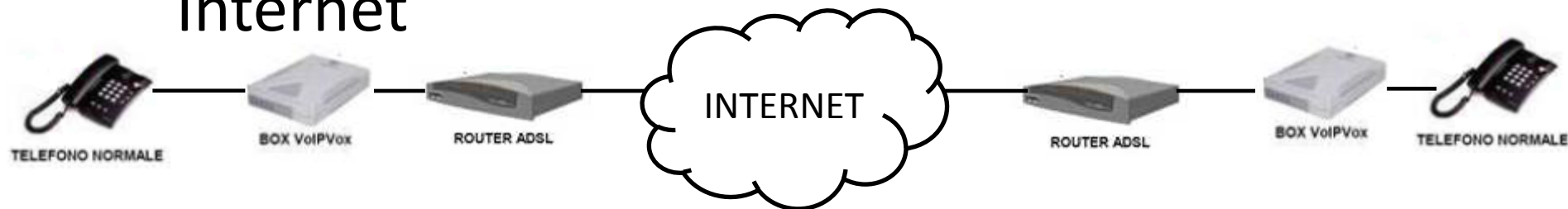


# Generalità

## VoIP-VoIP – Terminali adattati

---

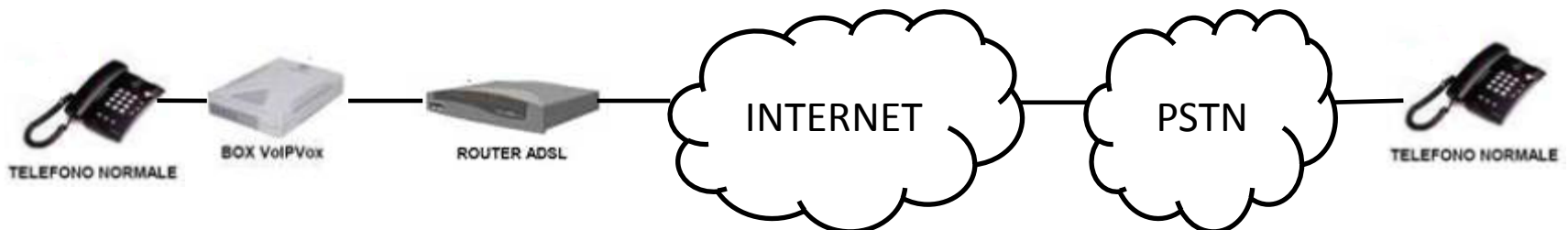
- La comunicazione è tra terminali usati dalla rete PSTN (solitamente telefoni) che sono stati adattati al VoIP
- La rete PSTN è completamente bypassata
- I pacchetti voce sono instradati tra sorgente e destinazione utilizzando interamente la rete Internet



# Generalità VoIP-PSTN

---

- Comunicazione tra un dispositivo VoIP (PC, telefono VoIP o telefono adattato) e un normale telefono PSTN
- La comunicazione è ibrida, ed è necessario che sia garantito *interworking* tra la rete Internet e la rete PSTN

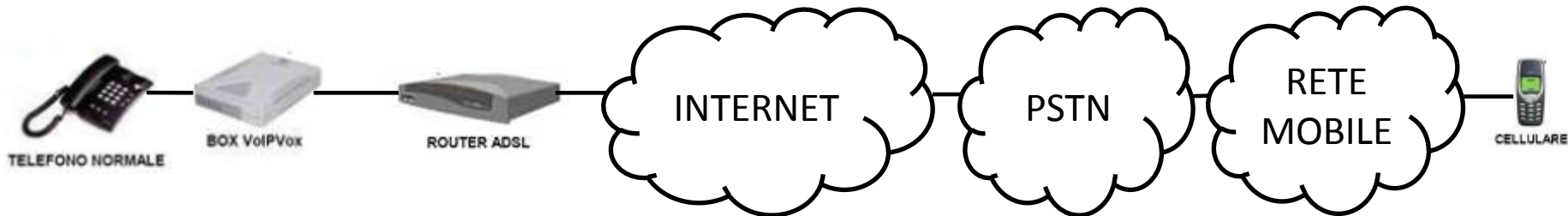


# Generalità

## VoIP-Rete mobile

---

- Comunicazione tra un dispositivo VoIP (PC, telefono VoIP o telefono adattato) e un telefono cellulare
- La comunicazione è ibrida, ed è necessario che sia garantito *interworking* tra la rete Internet e la rete PSTN (oltre che l'interworking tra la rete PSTN e la rete mobile)





# Perché VoIP?

---

- L'integrazione di telefonia e IP rappresenta l'opportunità di realizzare un sistema globale di comunicazione (multimediale e non)
  - L'unica rete dati universale è basata su IP
- L'integrazione, in generale, porta a dei vantaggi
  - Gestione unificata di problematiche e risorse/servizi di comunicazione
  - Metodo di trasporto e interfaccia utente universali
  - Servizi realizzati con protocolli end-to-end che utilizzano piattaforme aperte e che aprono la strada a molteplici vendors e a forti innovazioni a basso costo

# Vantaggi per l'utente

---

- Utilizzo di apparati intelligenti che permettono di programmare direttamente servizi di chiamata e risposta
  - Es. Programmazione automatica/schedulata di «call forwarding»
    - «In funzione dell'ora diretta a questo terminale e, per mancata risposta, diretta alla segreteria telefonica»
- Mobilità e portabilità del numero telefonico/identità
  - L'identità dell'utente è sganciata dall'indirizzo di rete (associato alla terminazione della linea)
    - E' possibile indirizzare una persona indipendentemente dalla sua posizione fisica

# Vantaggi per l'utente

---

- Messaggistica/Comunicazione unificata
  - Integrazione E-mail, voice mail, SMS
  - Scambio di dati durante la chiamata, sessioni in video
  - Multiconferenza
- Accesso remoto alla Intranet
  - Completo utilizzo dell'Intranet, compresa la telefonia
    - Es. Telelavoro da casa

# Vantaggi per l'utente

---

## ■ Qualità regolabile

- Le chiamate possono essere scelte o programmate con qualità scelta dall'utente in modo da massimizzare il rapporto costo/prestazioni
- Le chiamate interne possono essere di alta qualità, mentre quelle esterne, o lunga distanza, possono essere a minima banda (minimo costo)

# Vantaggi per il gestore di rete

---

- Unificazione della gestione fonia/dati
- Gestione semplificata della fonia con interfacce standard e “user friendly”
- Facilità di offrire servizi su reti IP
  - Unificazione di servizi voce dati (call centers, WEB, security...)
  - Facile espandibilità di servizi
  - Minor costo di servizi ed hardware

# Enti di standardizzazione

---

- International Telecommunication Union (ITU)
- Internet Engineering Task Force (IETF)
  - AVT: RTP
  - SIP: SIP per IP Telephony
  - IPTEL: CPL and GLP
  - SIPPING: estensioni SIP
  - SIMPLE: SIP per Instant Messagging
  - PINT (PIN): PSTN-Internet services
  - MEGACO: Media gateway control
  - SIGTRAN: Telephony signaling transport
- European Telecommunication Standards Institute (ETSI)
  - Typhon

# **VOIP**

---

## **Codifica della voce**

# Qualità del segnale audio

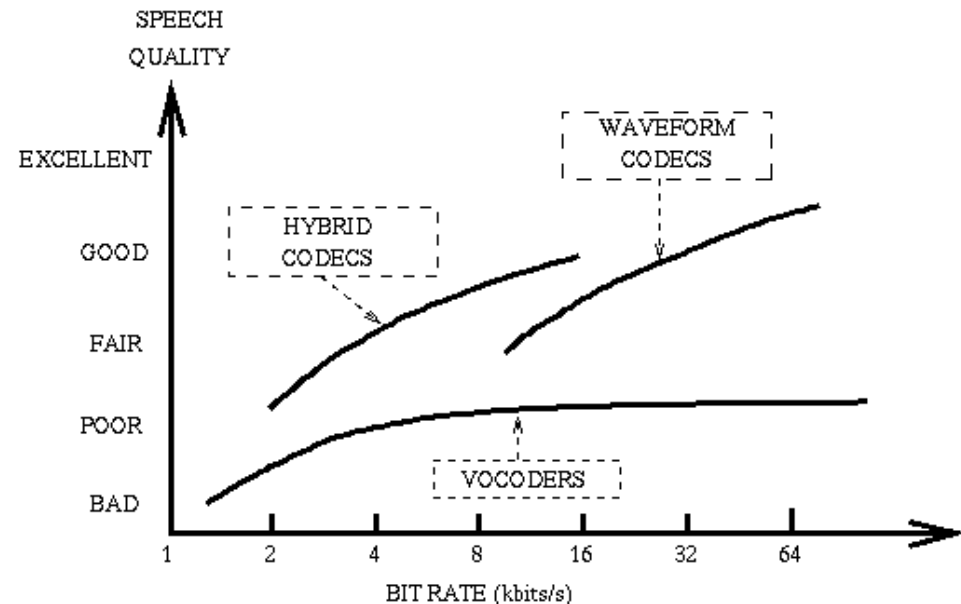
- La codifica avviene a partire dalla rappresentazione digitale del segnale audio
- La rappresentazione di partenza può essere più o meno fedele, in relazione all'uso che si intende fare del segnale audio
- Il segnale vocale può avere componenti fino a una ventina di kHz
  - Per una riproduzione perfetta è necessario campionarlo ad una frequenza di 40 kHz (teorema del campionamento)
  - La maggior parte dell'energia è concentrata nei primi 4 kHz
- Usualmente si identificano tre livelli di qualità in funzione della finalità d'uso del segnale audio

Frequenza di campionamento	Qualità	Caratteristiche	Applicazioni
44 kHz	CD	Riproduzione perfetta	Multimedia
14 kHz	Audio	Riproduzione quasi perfetta	Videoconferenza
8 kHz	Telefonia	Riproduzione buona	Telefonia



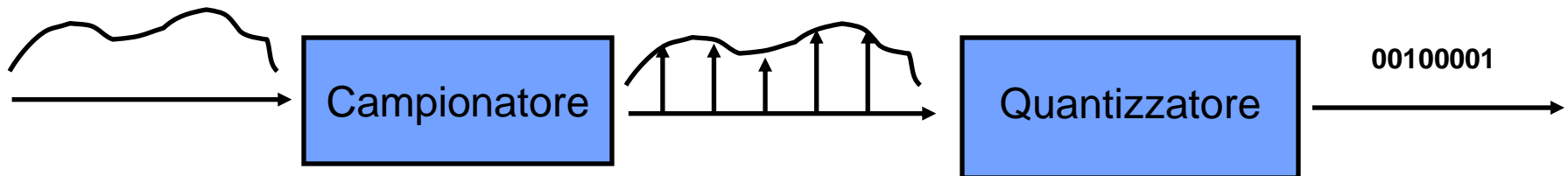
# Codificatori Vocali

- Permettono di trasformare la voce in un flusso di bit
- Tipologie
  - Waveform codecs
    - Sistemi a codifica della forma d'onda (PCM, DPCM)
  - Source codecs (vocoders)
    - Sistemi a codifica di parametri del modello della sorgente (LPC)
  - Hybrid codecs
    - Sistemi misti



# Waveform Codecs

- Codificano il segnale di banda  $B$  senza perdite (al netto dell'errore di quantizzazione)
  - Campionamento: si passa da una rappresentazione tempo-continua ad un segnale tempo-discreto
    - Teorema di Nyquist: campionamento a  $2B$
    - Telefonia classica:  $B = 4 \text{ kHz}$
- Quantizzazione: si rappresentano i valori dei campioni con un numero di livelli discreto e finito
  - Operazione irreversibile
  - Uniforme o non-uniforme per addensare i livelli ove si necessita maggiore accuratezza
  - Telefonia base:  $8 \text{ bit/campione} = 64 \text{ kb/s}$



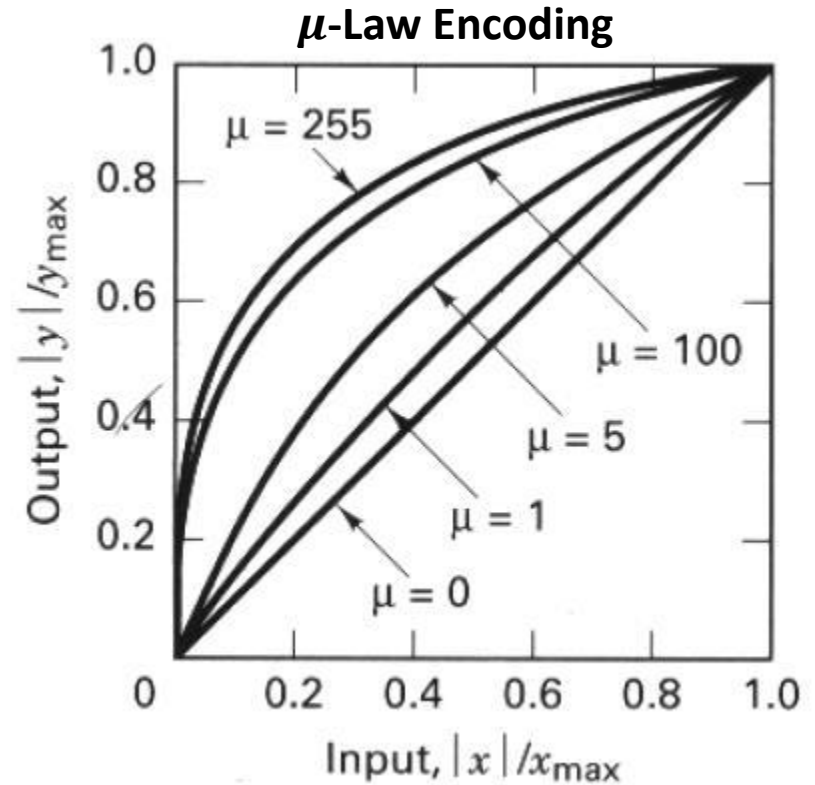
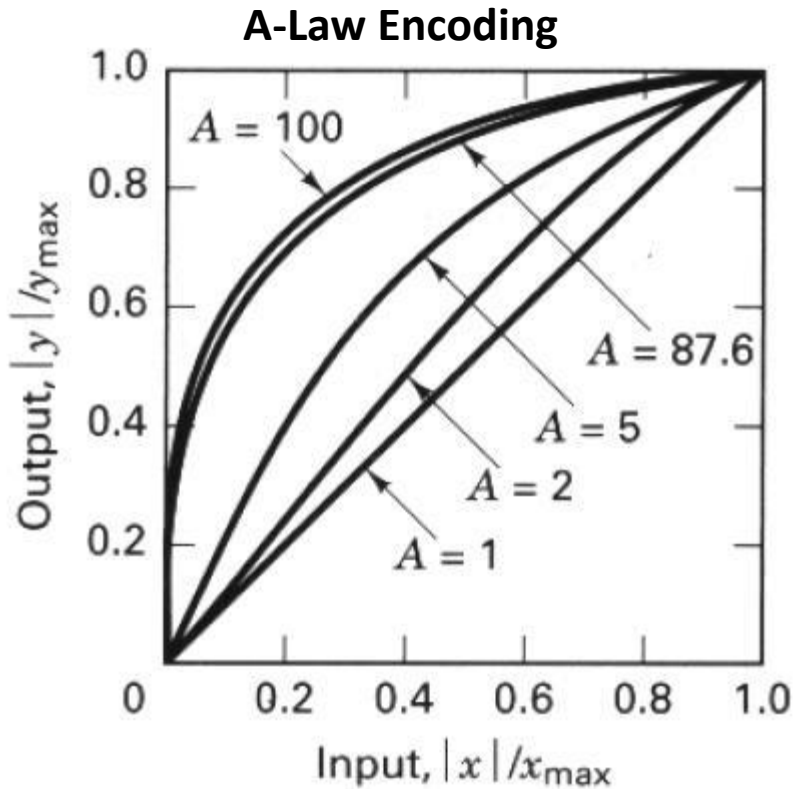
Alta qualità, bassa complessità, basso ritardo, robustezza agli errori e al rumore di fondo

# Pulse Code Modulation (PCM)

---

- Standardizzato da ITU nel 1960: G.711
  - Si assume  $B=4$  kHz, e la frequenza di campionamento  $B_c=8$  kHz, 8 bit/campione, 64 kb/s
- Due differenti regole di quantizzazione (logaritmica) a raffinare la granularità ove le ampiezze dei segnali sono più basse
  - Regole di conversione standard
    - America:  $\mu$ -law
    - Europa: A-law
- Quantizzazione che si adatta alla dinamica dei di *loudness* percepita dall'orecchio umano

# Quantizzazione

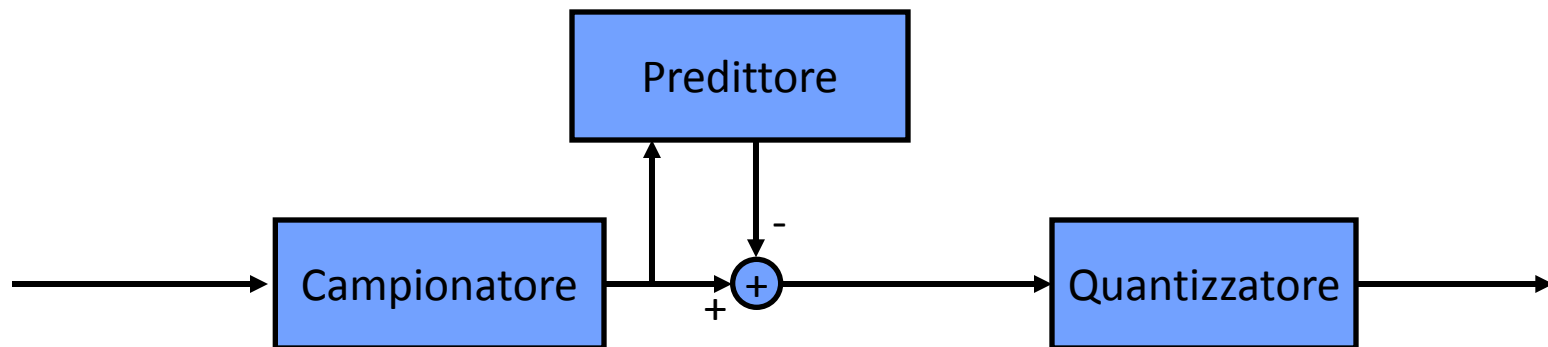


$$y = \begin{cases} y_{\max} \frac{A(|x|/x_{\max})}{1 + \log_e A} \operatorname{sgn} x & 0 < \frac{|x|}{x_{\max}} \leq \frac{1}{A} \\ y_{\max} \frac{1 + \log_e [A(|x|/x_{\max})]}{1 + \log_e A} \operatorname{sgn} x & \frac{1}{A} < \frac{|x|}{x_{\max}} < 1 \end{cases}$$

$$y = y_{\max} \frac{\log_e [1 + \mu(|x|/x_{\max})]}{\log_e (1 + \mu)} \operatorname{sgn} x$$

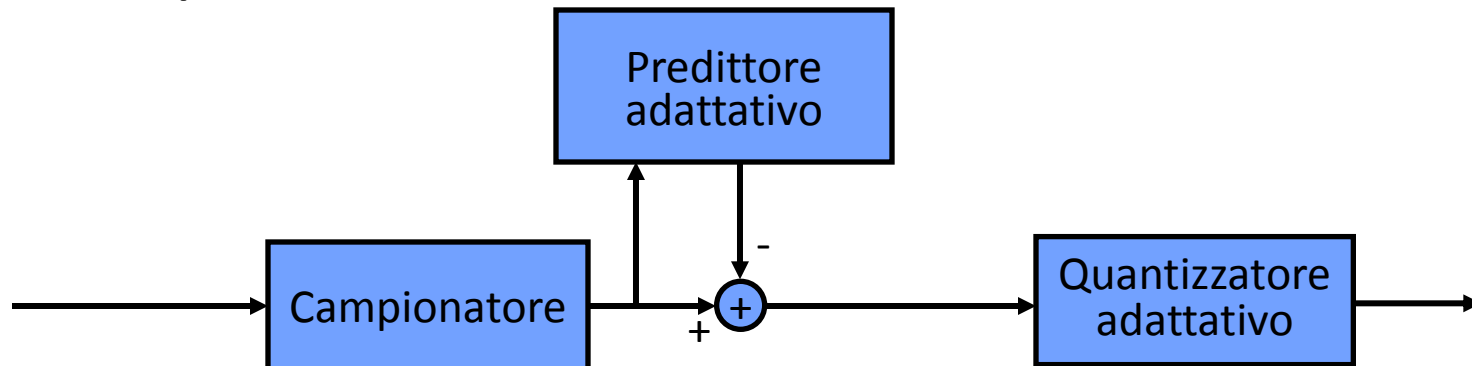
# Differential PCM (DPCM)

- I campioni vocali successivi presentano della correlazione
- E' possibile utilizzare metodi di predizione per valutare il campione successivo noti i precedenti
- Si trasmette solo la differenza tra valore predetto e valore reale
- A causa della correlazione la varianza della differenza è minore ed è possibile codificarla con un minor numero di bit
- Viene utilizzata una quantizzazione uniforme per la codifica della differenza



# Adaptive DPCM (ADPCM)

- Le prestazioni migliorano se predittore e quantizzatore sono adattativi
  - Ciò che viene adattato è la dimensione dello step di quantizzazione
- Standardizzato nel 1980 da ITU ADPCM a 32 kbit/s: G.721
- Successivamente ADPCM a 40, 32, 24, 16 kbit/s: G.726 e G.727

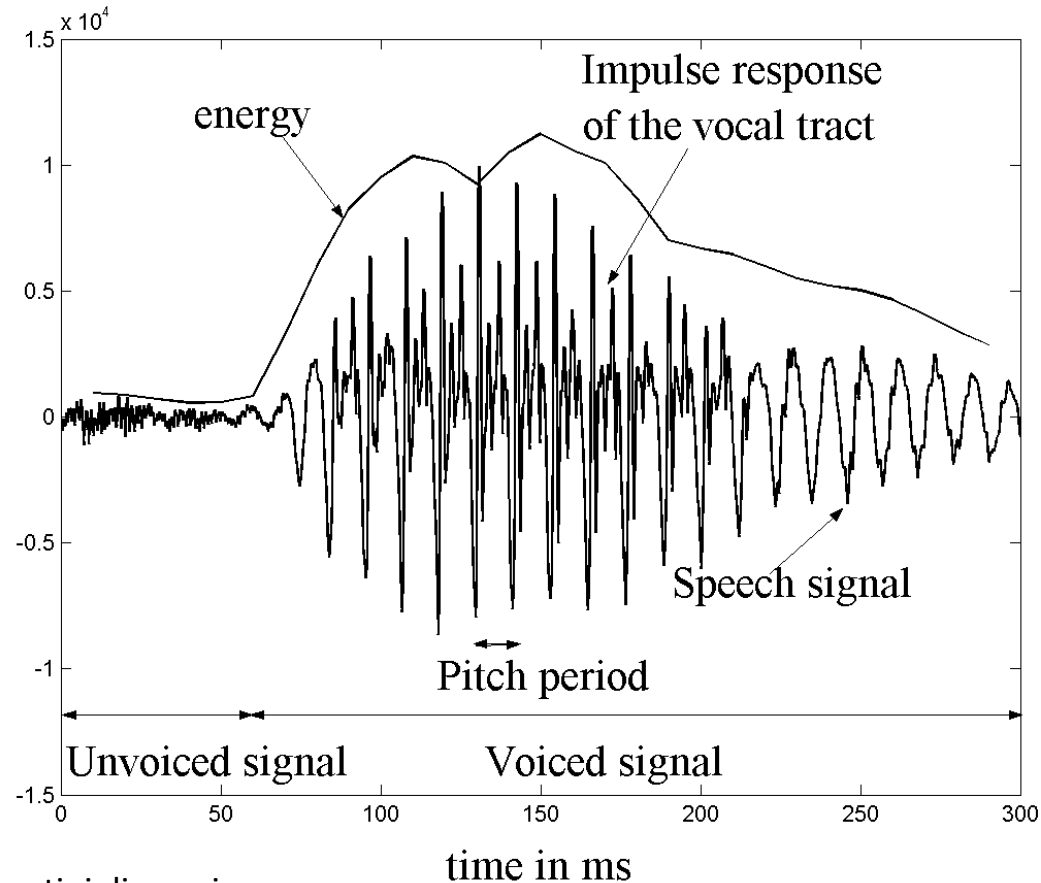


# Source Codecs (vocoders)

---

- Si basano su modelli che permettono di generare/riprodurre la voce umana sulla base delle sue caratteristiche intrinseche
- Tali modelli permettono di rimuovere la ridondanza dalla voce fino ad ottenere la minima informazione sufficiente a riprodurre la voce
- Caratteristiche
  - Altissima efficienza (pochi Kb/s per la voce)
  - Elevata complessità
  - Ritardi mediamente elevati
  - Sensibili a errori, rumore di fondo e suoni non prodotti da esseri umani

# Caratteristiche del segnale vocale

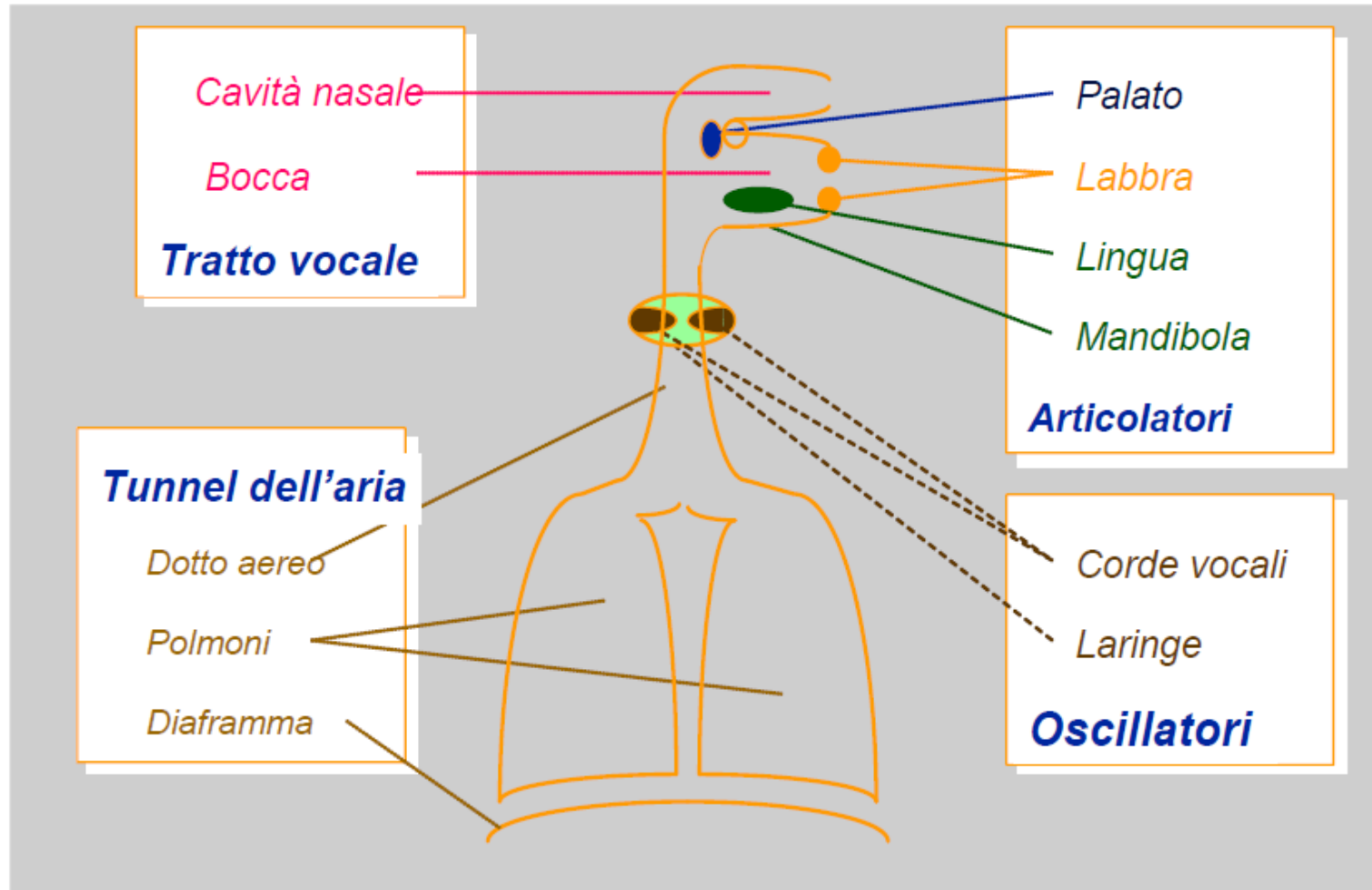


La voce è composta da due tipi di suoni:

- Suoni sonori (voiced), tipici nelle vocali, caratterizzati da un andamento periodico secondo una frequenza di pitch
- Suoni sordi (unvoiced), tipici delle consonanti, di ampiezza ridotta e caratterizzati da alte frequenze



# Produzione vocale



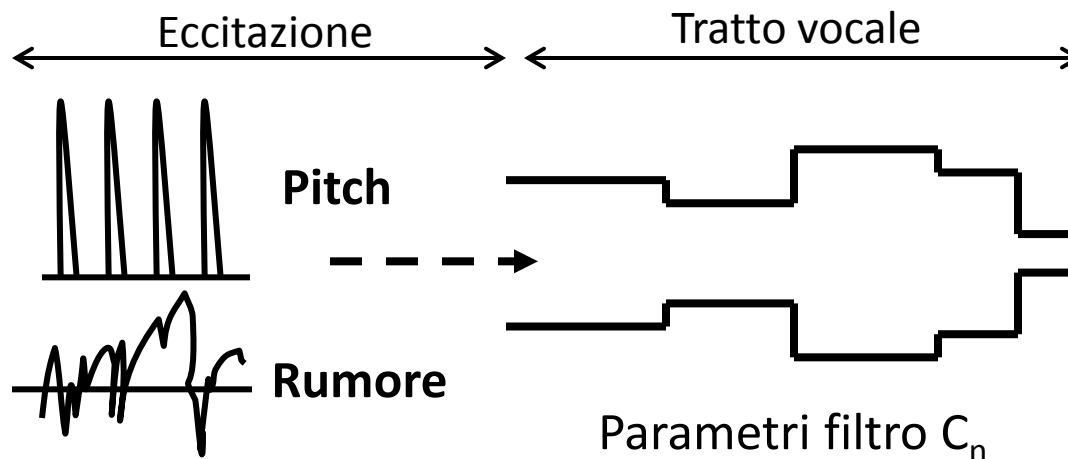
# Produzione vocale

---

- Il processo di produzione vocale può essere scomposto in tre fasi
  1. Produzione del fiato (aria espirata dai polmoni)
  2. Generazione del suono (vibrazione delle corde vocali)
  3. Modulazione del suono (risonanze del canale vocale, composto da faringe e bocca)
- Tale processo viene modellato da un *filtro riverberante a parametri discreti*

# Modello a fonemi

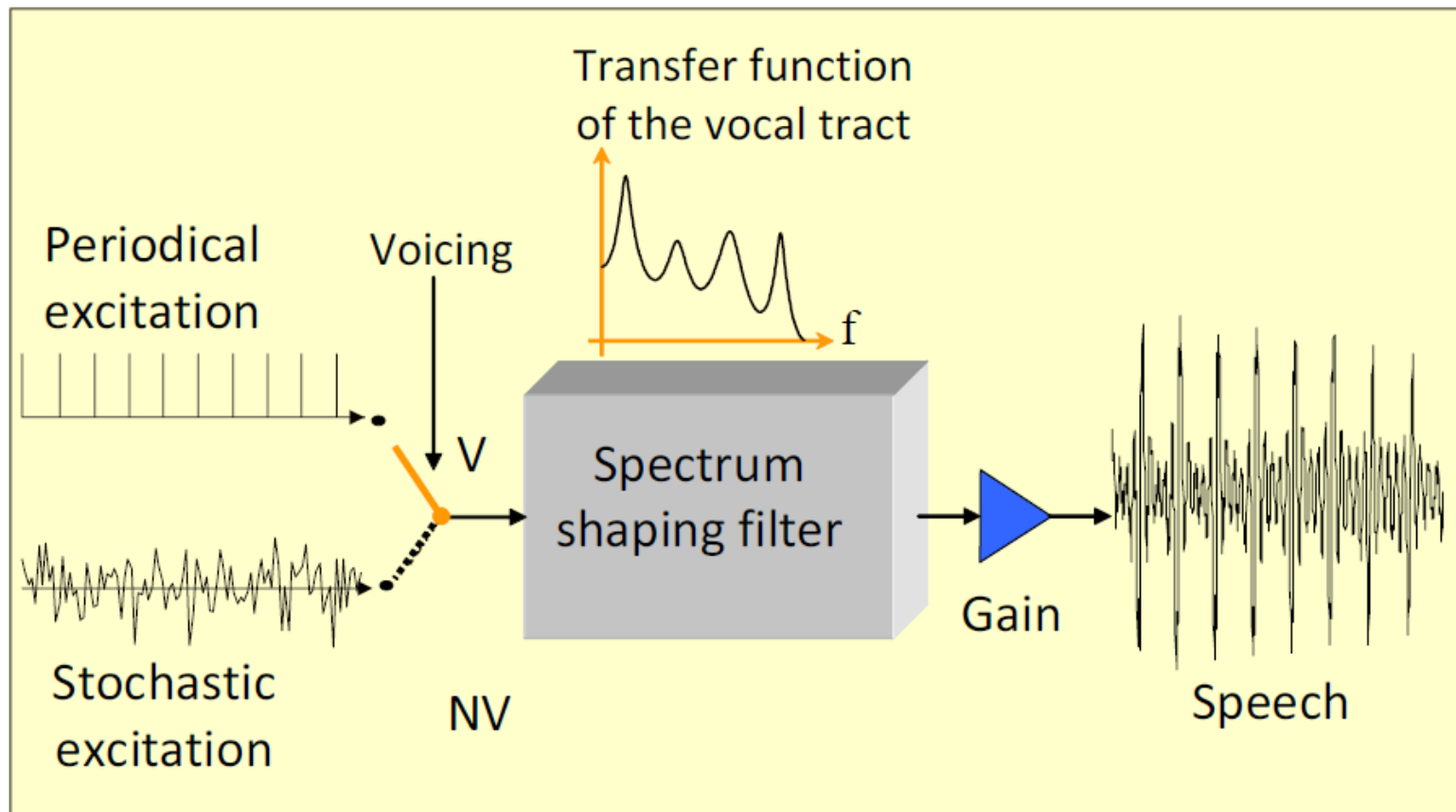
- Filtro riverberante a parametri discreti
- Segnale in ingresso (treno di impulsi o rumore bianco)
- Parametri del filtro variati periodicamente (10-20 ms)



*Un fonema è la rappresentazione astratta di un suono*

# Modello del vocoder

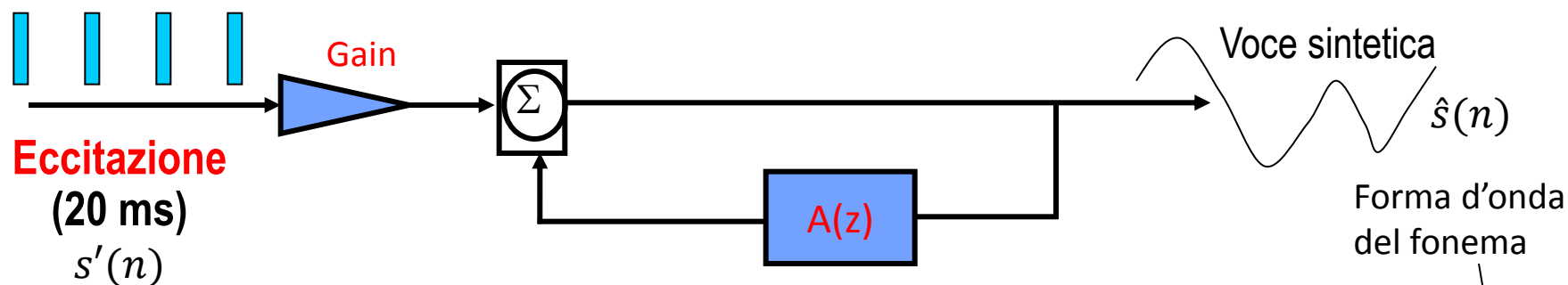
Modello elettrico equivalente



# Vocoder lineari (LPC)

- Modello della voce basato su un filtro lineare a parametri discreti
- Ad intervalli regolari (10-20 ms) vengono stimati e trasmessi i parametri del modello
  - Coefficienti del filtro  $a_i$
  - Voiced/unvoiced flag, varianza del segnale di eccitazione/pitch, guadagno

$$\hat{s}(n) = \sum_{i=1}^p a_i s'(n-i)$$



La stima è effettuata minimizzando la varianza dell'errore

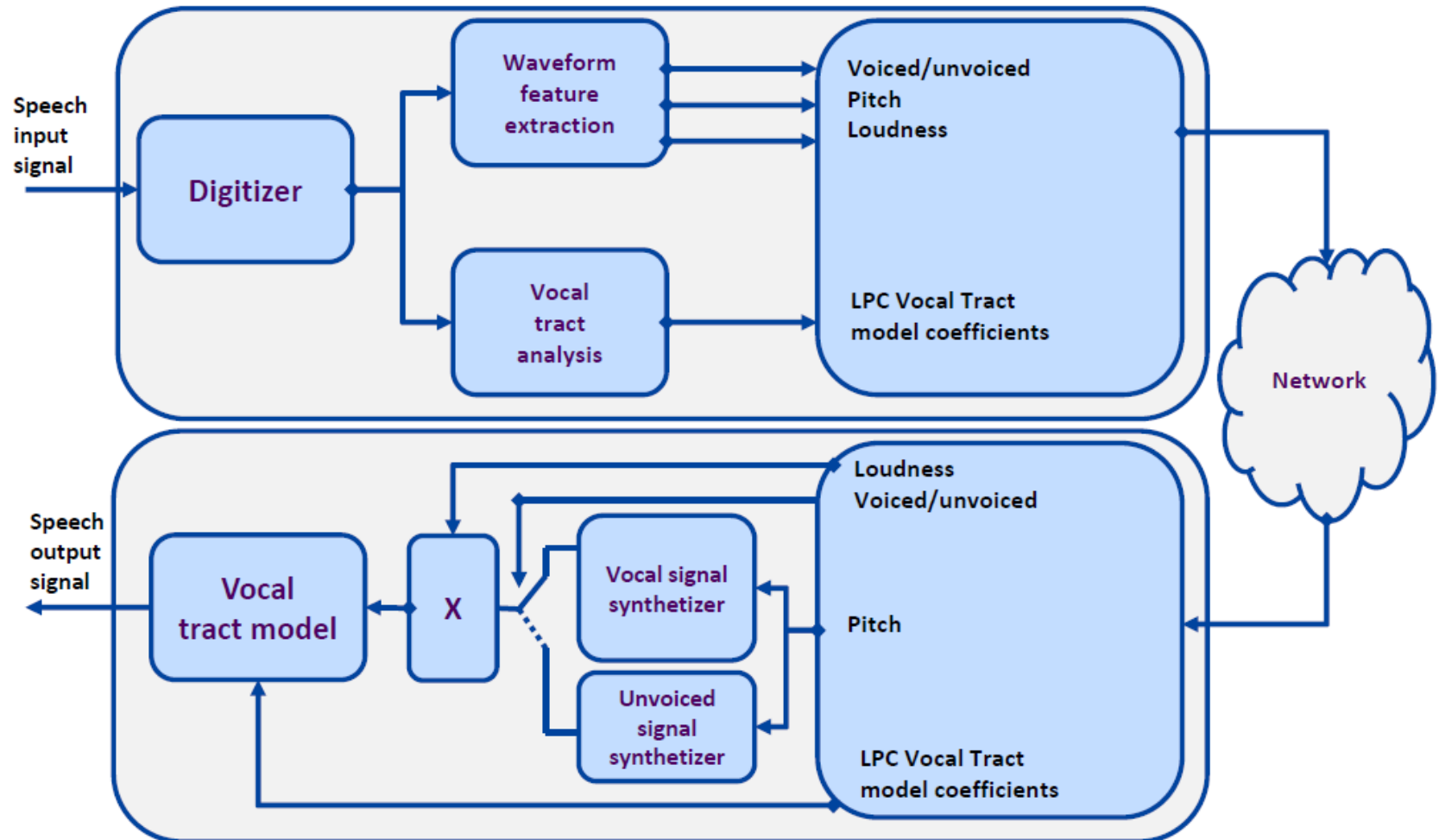
$$e(n) = \hat{s}(n) - s(n)$$

# Vocoder lineari (LPC)

---

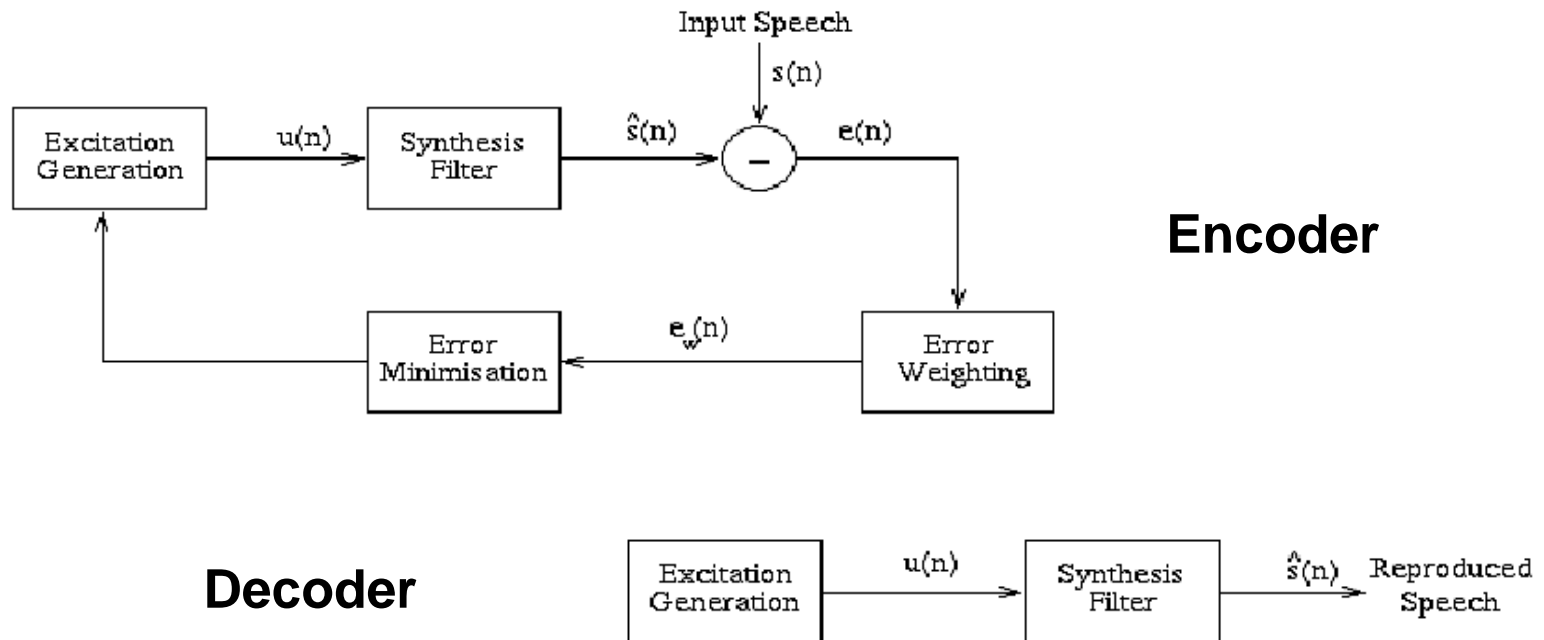
- Esempio
  - Campionamento a 8 kHz con buffer da 240 campioni → Stima ogni circa 30 ms
  - Si trasmette:  $G = 5$  bit,  $a_i = 8$  bit ciascuno ( $i = 1, \dots, 10$ ), voiced/unvoiced decision = 1 bit, pitch/varianza rumore eccitazione = 6 bit →  $92 \text{ bit} / 30 \text{ ms} = 3,067 \text{ kbit/s}$
- In decodifica un sintetizzatore utilizza i parametri ricevuti per riprodurre il segnale
- Ritardi elevati: segmentazione, analisi, sintesi
- Qualità: intelligibile ma non naturale (limiti modello + problemi con rumori di fondo)
- Bit rate basso:  $< 2.4 \text{ kbit/s}$

# Vocoder lineari (LPC)



# Hybrid Codecs

- Cercano di colmare il vuoto tra vocoder e waveform codec
- I più diffusi usano le stesse tecniche dei vocoder (LPC), ma ottimizzano alcuni parametri (segnale di eccitazione) minimizzando il segnale errore





# Hybrid Codecs

---

- **Multipulse-Excited Linear Prediction (MPELP)**
  - Sequenza di  $N$  impulsi non equi-spaziati e di ampiezze variabili
  - Posizione e ampiezza di ciascun impulso sono determinate con procedura iterativa che minimizza una funzione dell'errore
- **Regular Pulse Excitation (RPE)**
  - Sequenza di  $N$  impulsi equi-spaziati
  - Parametri: posizione primo impulso (fase), periodo e ampiezza
  - Es. LTP-RPE GSM 13 kbit/s

# Hybrid Codecs

---

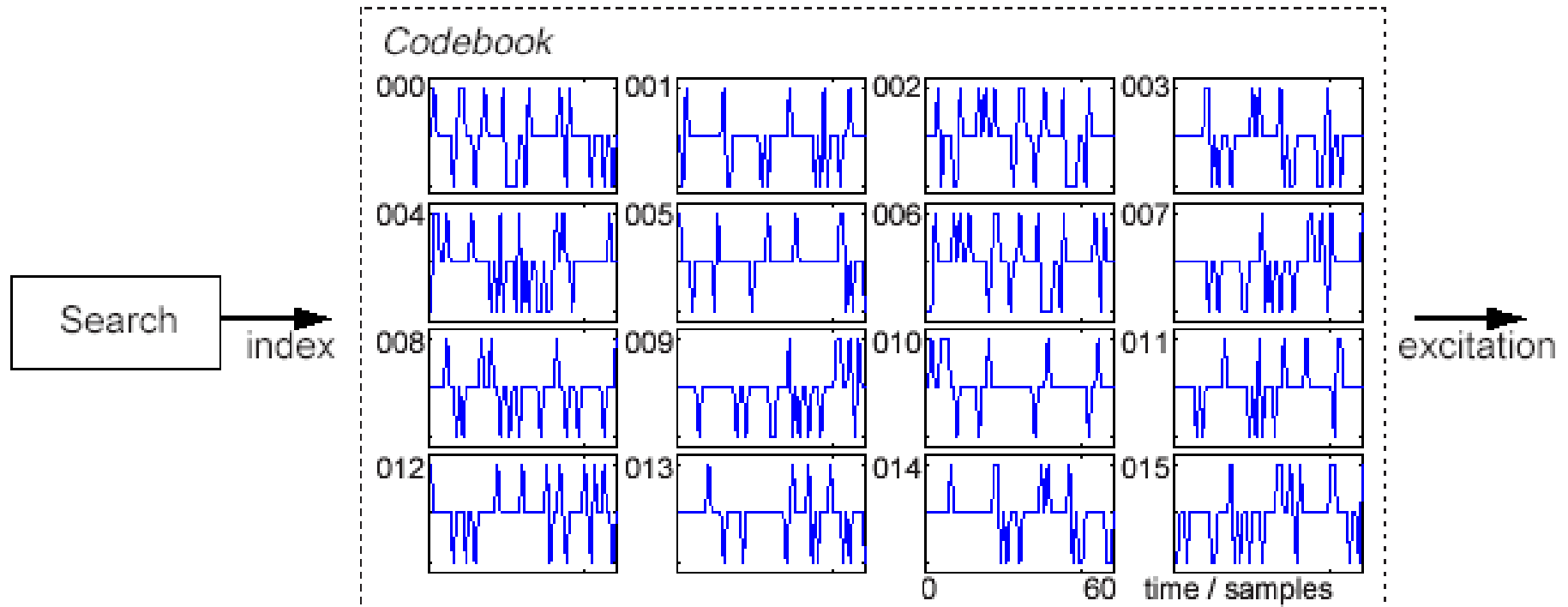
- Codificatore GSM: codificatore LPC-LTP (Linear Prediction Coding- Long Term Prediction) con RPE (Regular Pulse Excitation)
- Standard ETSI
  - Blocchi da 20 ms che producono 260 bit l'uno raggruppati in 3 livelli di importanza
    - 50 richiedono massima protezione
    - 132 richiedono una protezione media
    - 78 non sono protetti per nulla
- Totale: 13 kbit/s

# Hybrid Codecs - CELP

---

- Code Excited Linear Prediction (CELP)
  - La sequenza di eccitazione viene scelta tra un insieme di sequenze (*code-book*) in modo da minimizzare l'errore
- Problema: lunghi ritardi dovuti alla ricerca della sequenza ottima nel code-book
- Semplificazione dell'algoritmo mediante metodi efficienti di ricerca e modifica del code-book
  - G.728 low delay CELP codec 16 kbit/s
  - G.729 CS-ACELP codec 8 kbit/s
  - G.723.1 ACELP 5.3 kbit/s

# Hybrid Codecs - CELP



# Principali Codifiche

---

<b>Compressione</b>	<b>Year</b>	<b>Bit rate (kbit/s)</b>	<b>Frame size (ms)</b>	<b>Look ahead (ms)</b>
G.711 PCM	1972	64	0.125	0
G.726 ADPCM	1990	32	1	0
G.722 Subband ADPCM	1988	48-64	0.125	1.5
G.728 LD-CELP	1992-94	16	0.625	0
G.729 CS-ACELP	1995	8	10	5
G.723.1 ACELP	1996	5.3	30	5
RPE-LTP (GSM)	1987	13	20	0

# Valutazioni di Qualità

---

## ■ Mean Opinion Score (MOS)

- E' una misura soggettiva della qualità della voce
  - Misura la sua chiarezza
- Basato sull'opinione di un gran numero di ascoltatori
- Standardizzato nella raccomandazione P.800 (ITU-T)

### Score Quality of Speech

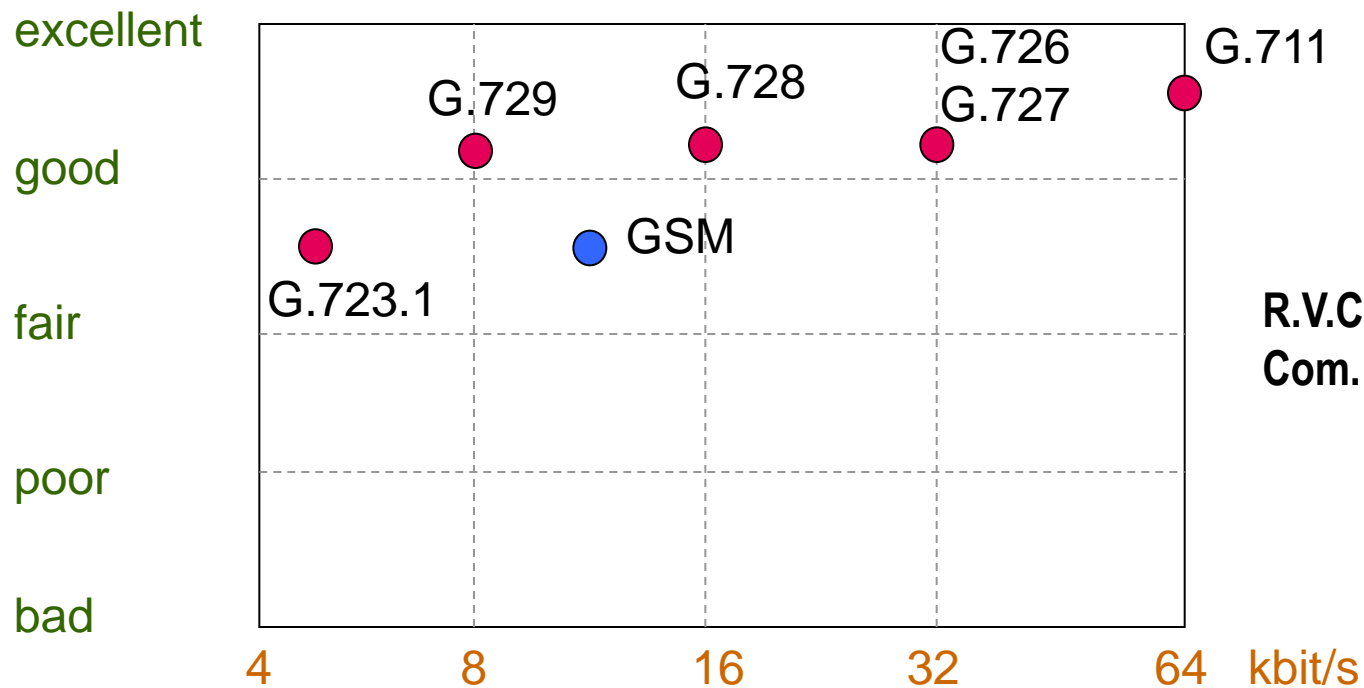
5 Excellent  
4 Good  
3 Fair  
2 Poor  
1 Bad

### Difetti:

- Metodo complesso e costoso
- Non adatto a misure continue

# Valutazioni di Qualità

## ■ Mean Opinion Score (MOS)



R.V.Cox, AT&T  
Com. Mag, Sept 97

# Valutazioni di Qualità

---

- Perceptual Speech Quality Measurement (PSQM)
  - Raccomandazione P.861 (ITU-T) del 1996
  - Misura oggettiva della qualità dei codificatori
  - Utilizza un confronto matematico fra segnale non elaborato (pre-codifica) e segnale degradato (post-codifica) operato attraverso un modello che tiene conto delle modalità percettive/psicoacustiche dell'essere umano
  - I risultati sono tarati in modo da predire il MOS

## Difetti:

- Pensato esclusivamente per i codec
- Non riscontra i difetti tipici delle reti (errori del canale, perdita di pacchetti...)



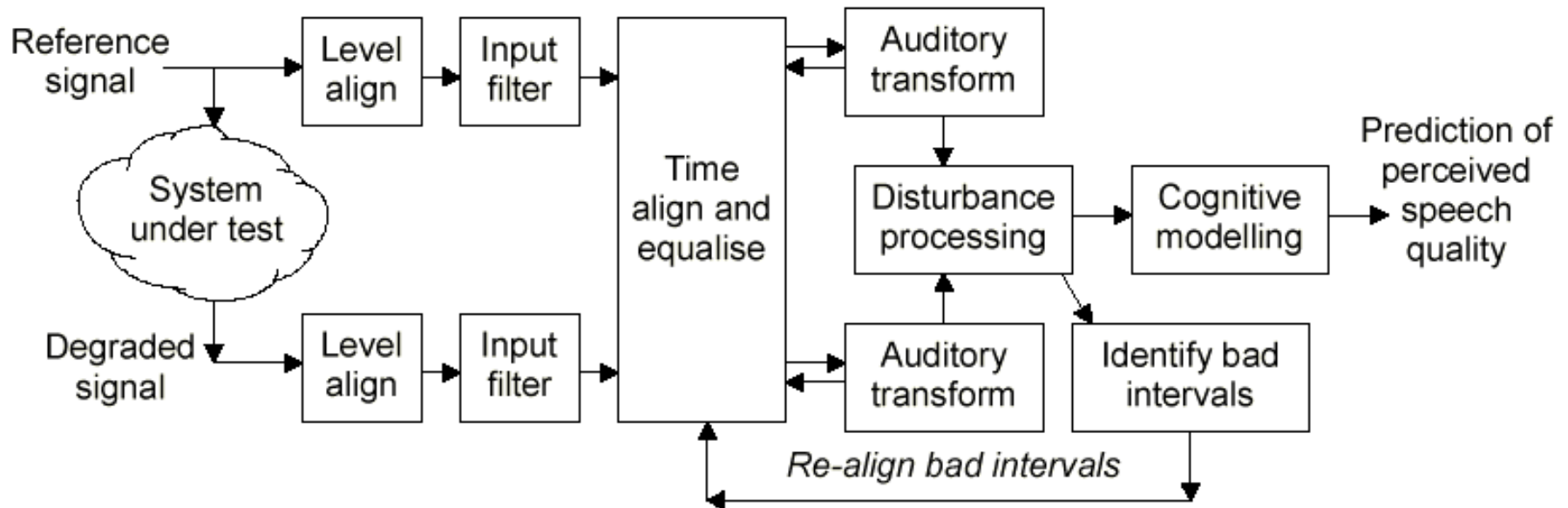
# Valutazioni di Qualità

---

- Perceptual Analysis Measurement System (PAMS)
  - Sviluppato da British Telecom a partire dal 1998 per ovviare alle mancanze di PSQM
  - Primo sistema adatto a testare in modo completo sistemi VoIP

# Valutazioni di Qualità

- Perceptual Evaluation of Speech Quality (PESQ)
  - Raccomandazione P.862 (ITU-T) del 2001



# Valutazioni di Qualità

## E-Model

---

- L'E-model è stato standardizzato dall'ITU-T (G.107) per prevedere, in modo approssimato, la qualità del segnale vocale su un collegamento IP
- Dati il codec e misure oggettive di ritardo e perdita di pacchetti (oltre ad altri fattori di degrado della comunicazione), con l'E-Model si determina una previsione approssimata del MOS che si otterrà su un collegamento telefonico IP
- L'E-model è un utile strumento impiegato largamente nell'industria per la progettazione preliminare e la pianificazione, anche se non è riconosciuto come strumento ufficiale da ITU-T

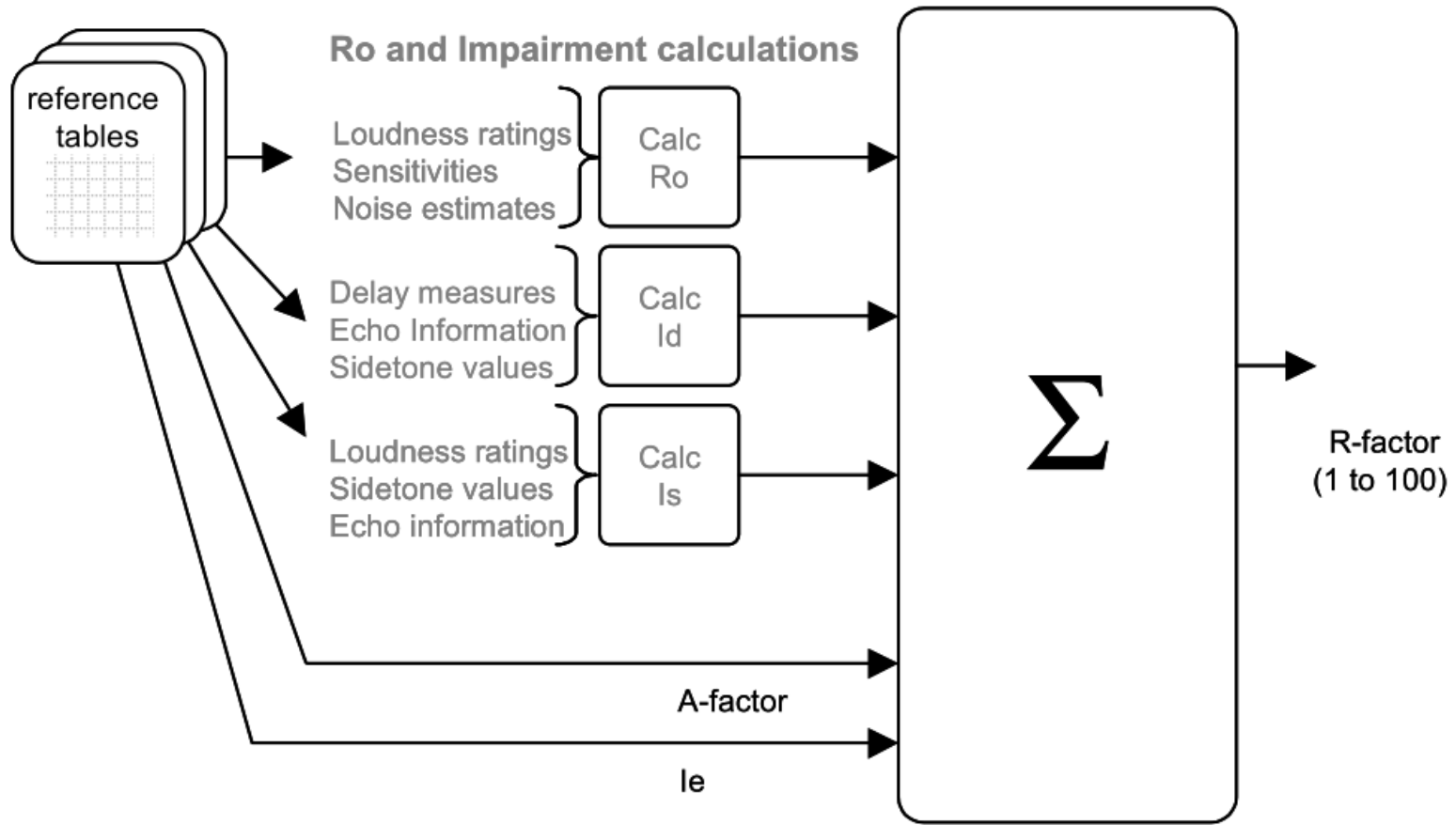
# Valutazioni di Qualità

## E-Model

- Si definisce una metrica R “rating factor” (compresa tra 1 e 100)
- La qualità R è definita come
- $R = R_0 - I_s - I_d - I_e + A$ 
  - $R_0$ : rapporto base segnale-rumore (es. rumore del circuito e dell’ambiente)
  - $I_s$ : fattore di degrado simultaneo alla trasmissione del segnale vocale (es. quantizzazione)
  - $I_d$ : degrado imputabile ai ritardi del segnale vocale (eco parlatore, eco ascoltatore, ritardo assoluto)
  - $I_e$ : degrado imputabile al dispositivo (equipment/codec)
  - $A$ : fattore di vantaggio dell’utente (l’utente è sempre proiettato a compensare degni prestazionali e a non riconoscerli tanto più se si trova in condizioni meno favorevoli)

Communication system example	Maximum value of A
Conventional (wirebound)	0
Mobility by cellular networks in a building	5
Mobility in a geographical area or moving in a vehicle	10
Access to hard-to-reach locations, e.g., via multi-hop satellite connections	20

# Valutazioni di Qualità E-Model



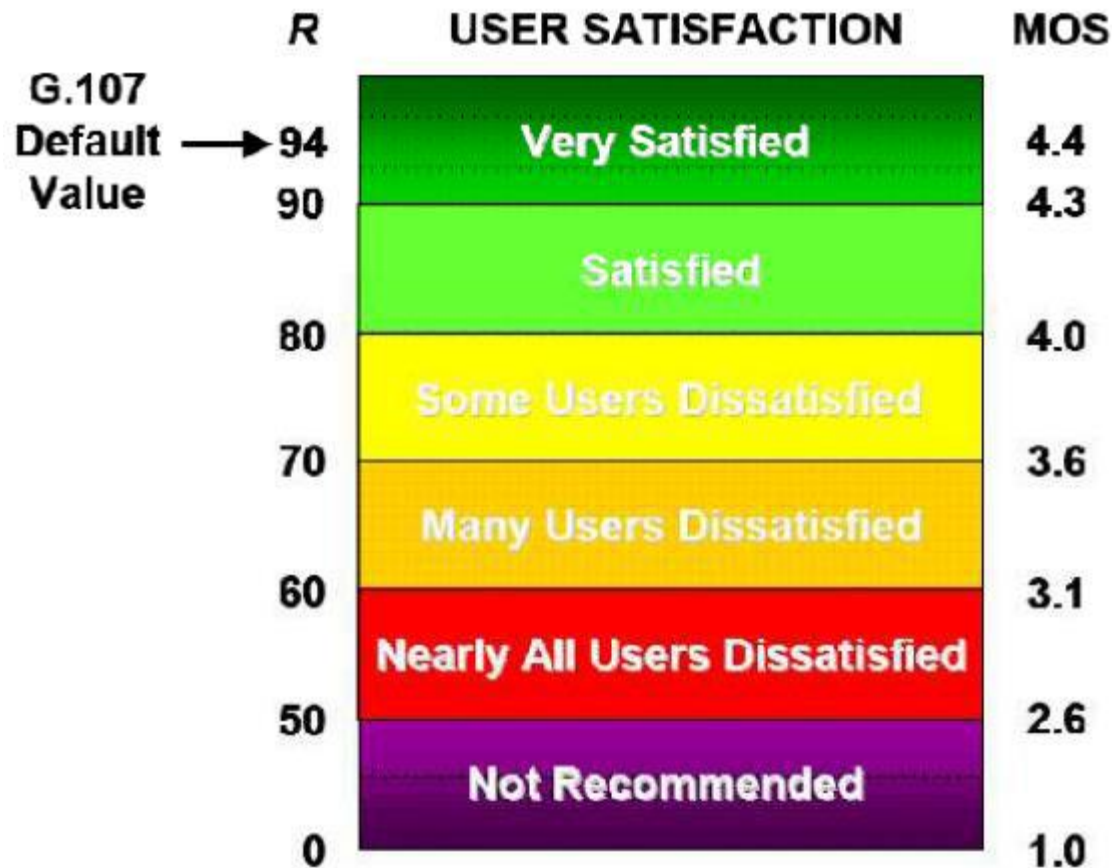
# Valutazioni di Qualità E-Model

Valori di  $I_E$  sulla base delle caratteristiche intrinseche dei codec

Codec Type	Reference	Operating Rate kbit/s	I <sub>E</sub> Value
PCM	G.711	64	0
ADPCM	G.726, G.727	40	2
	G.721, G.726, G.727	32	7
	G.726, G.727	24	25
	G.726, G.727	16	50
LD-CELP	G.728	16	7
		12.8	20
CS-ACELP	G.729	8	10
	G.729-A + VAD	8	11
ACELP	GSM 06.60, EFR	12.2	5
ACELP	G.723.1	5.3	19
MP-MLQ	G.723.1	6.3	15

# Valutazioni di Qualità E-Model

---



# VOIP

---

## Degrado della voce

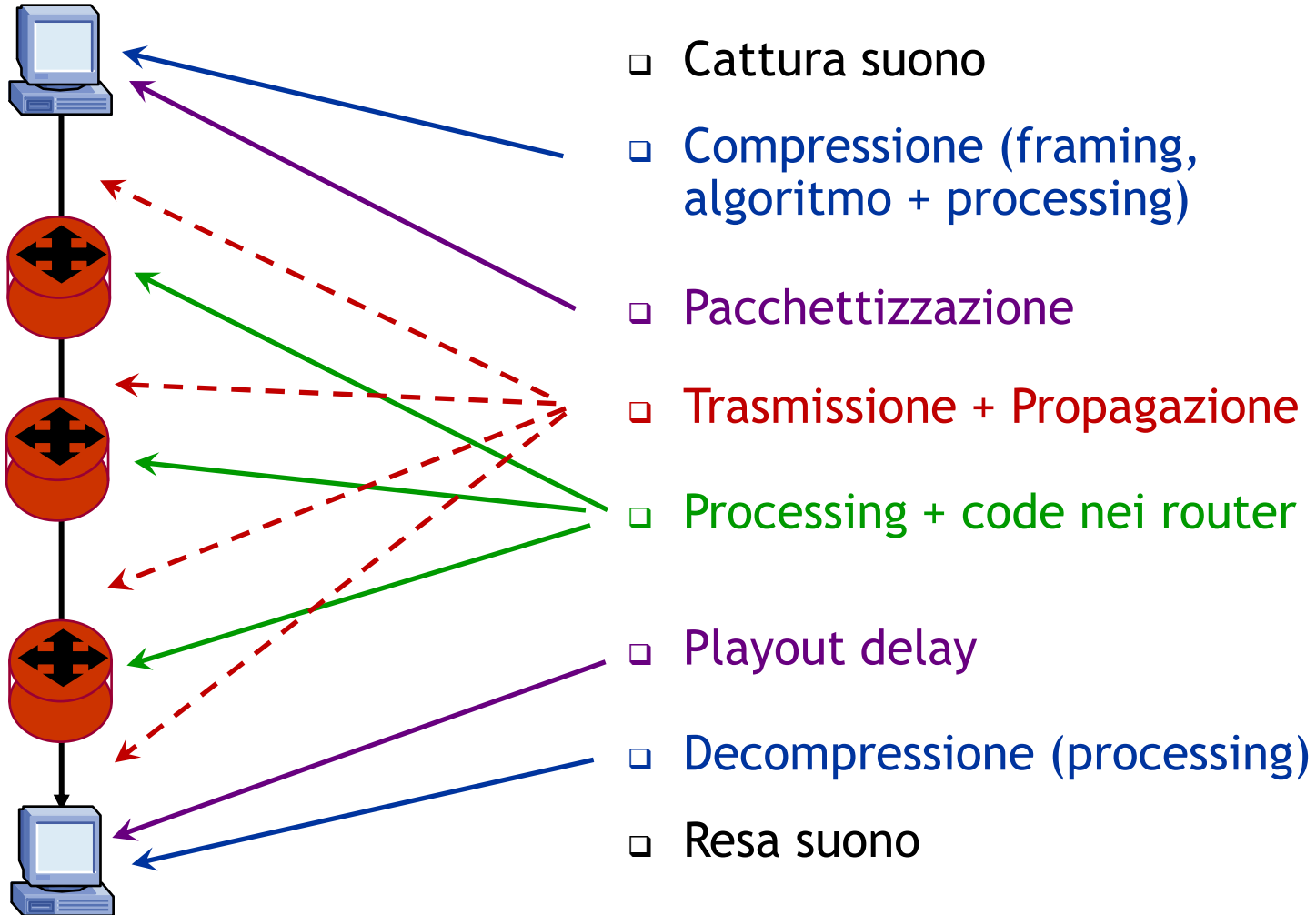


# Qualità della voce ricostruita

---

- Oltre alle caratteristiche intrinseche della codifica, nelle reti trasmissive la qualità della riproduzione è influenzata anche da
  - Ritardi di trasmissione/riproduzione eccessivi
  - Eco
  - Perdita totale o parziale di segmenti vocali codificati

# Cause di ritardi in VoIP



# End-to-End Quality of Service (QoS)

---

	<i>Best</i>	<i>High</i>	<i>Medium</i>	<i>Best Effort</i>
<b>MOS Quality:</b>	4.0 - 5.0	3.8 - 4.2	2.9 - 3.8	2.0 - 2.9
<b>Mouth-to-Ear Delay:</b>	0ms - 150ms	150ms - 250ms	250ms - 450ms	450ms and above
<b>Call Setup:</b>	0 sec - 1 sec	1 sec - 3 sec	3 sec - 5 sec	5 sec and above

# Esempio di un delay budget VoIP

---

	Fixed Delay	Variable Delay
Coder Delay G.729 (5 ms Look Ahead)	5 ms	
Coder Delay G.729 (10 ms Per Frame)	20 ms	
Packetization Delay Included in Coder Delay		
Queuing Delay 64 kbps Trunk		6 ms
Serialization Delay 64 kbps Trunk	3 ms	
Propagation Delay (Private Lines)	32 ms	
Network Delay (For Example, Public Frame Relay Svc)		
Dejitter Buffer		2-200 ms
<b>Total - Assuming 50 ms Jitter Buffer</b>	<b>110 ms</b>	

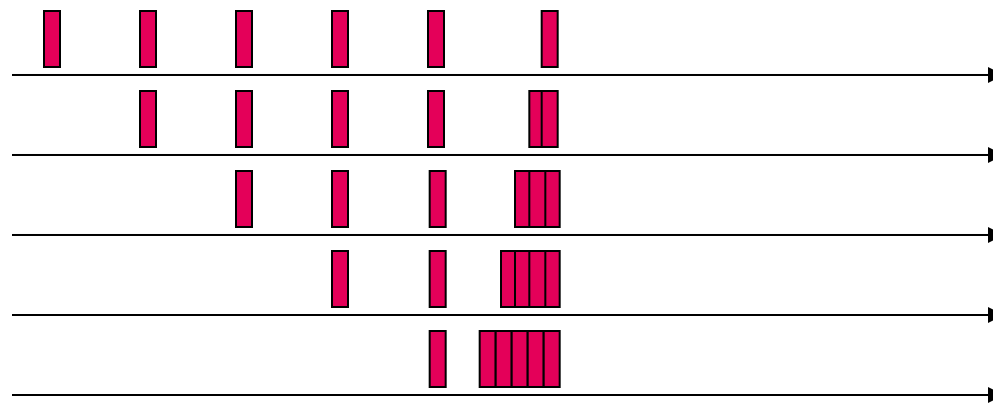
# Ritardo di compressione

---

<b>Compressione</b>	<b>Bit rate (kbit/s)</b>	<b>Framing (ms)</b>	<b>Look ahead (ms)</b>	<b>Processing (ms)</b>
<b>G.711 PCM</b>	<b>64</b>	<b>0.125</b>	<b>0</b>	<b>0</b>
<b>G.726 ADPCM</b>	<b>32</b>	<b>1</b>	<b>0</b>	<b>0</b>
<b>G.728 LD-CELP</b>	<b>16</b>	<b>0.625</b>	<b>0</b>	<b>0</b>
<b>G.729 CS-ACELP</b>	<b>8</b>	<b>10</b>	<b>5</b>	<b>0-10</b>
<b>G.723.1 ACELP</b>	<b>5.3</b>	<b>30</b>	<b>5</b>	<b>0-30</b>
<b>RPE-LTP (GSM)</b>	<b>13</b>	<b>20</b>	<b>0</b>	<b>0-30</b>

# Pacchettizzazione

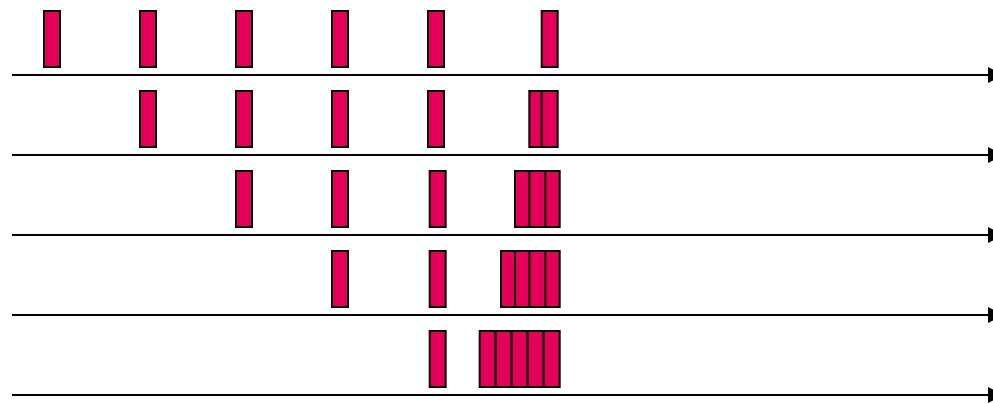
- E' la procedura per cui uno o più segmenti vocali vengono raggruppati per la trasmissione su reti a pacchetto
- E' causa di ulteriori ritardi qualora occorra l'accumulo di più segmenti codificati



Ritardo di  
pacchettizzazione

# Ritardo di pacchettizzazione

- Dipende dalla lunghezza dei segmenti vocali (ms di voce) e dal numero di segmenti vocali che vengono pacchettizzati insieme
- Si compone in modo approssimativamente additivo al ritardo di compressione



Ritardo di  
pacchettizzazione

# Ritardo di trasmissione/propagazione

## Ritardo di trasmissione

velocità del link	40 Byte (voce)
64 kb/s	5 ms
128 kb/s	2.5 ms
512 kb/s	0.625 ms
2 Mb/s	0.16 ms
10 Mb/s	0.033 ms

**moltiplicato  
per il n° di link  
attraversati**

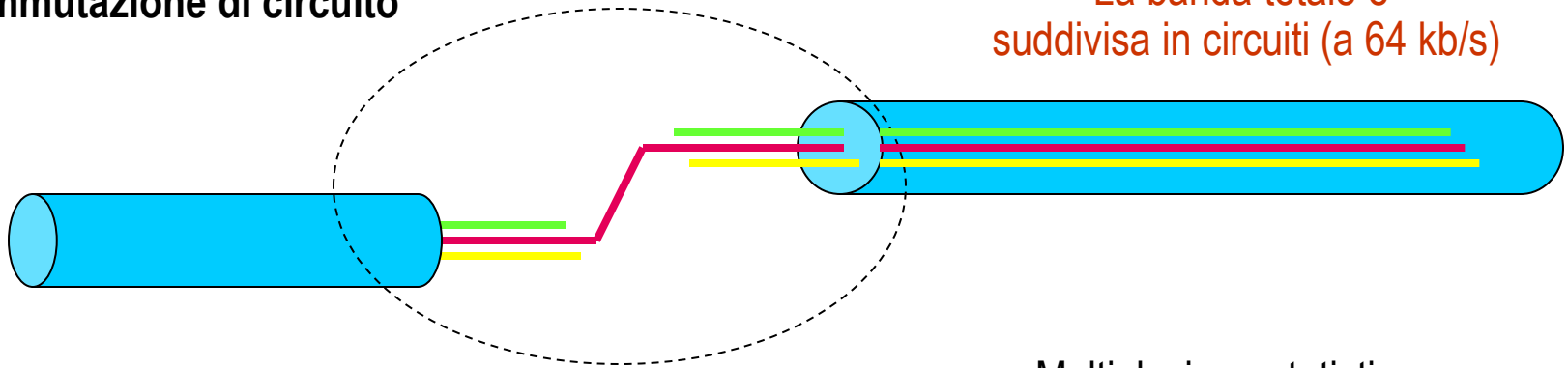


Ritardo di propagazione:  $5 \mu\text{s} / 1 \text{ km}$

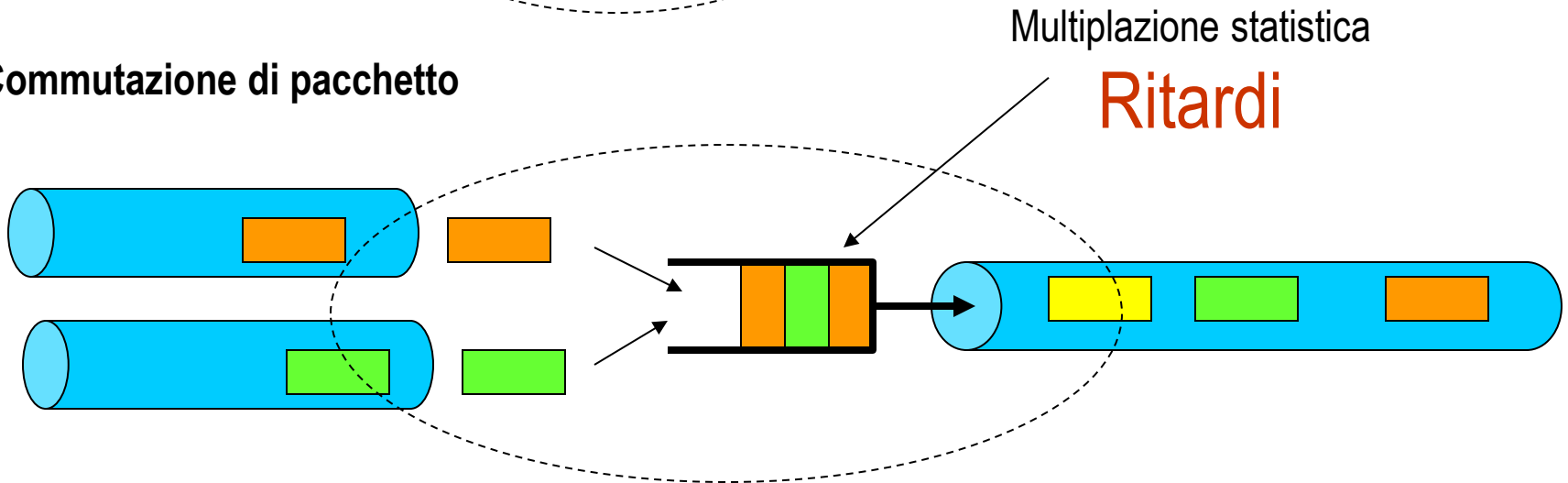


# Ritardo in coda

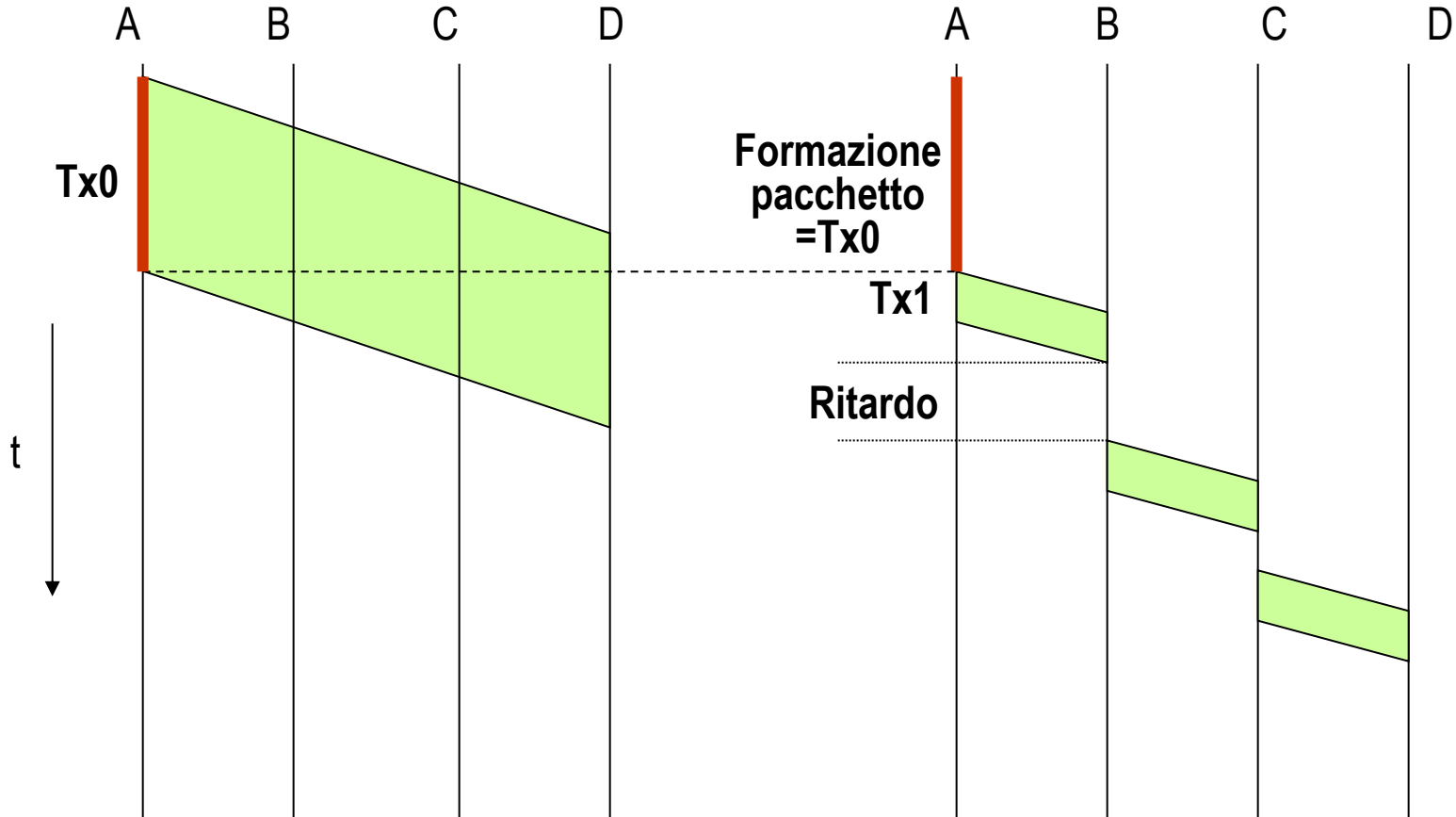
## Commutazione di circuito



## Commutazione di pacchetto



# Ritardo in coda



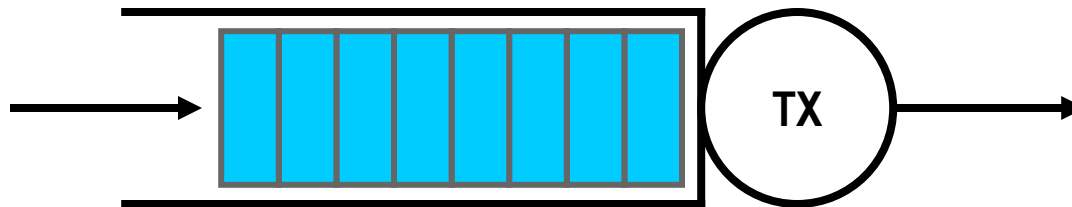
Circuiti a 64 kb/s: pipeline  
(No ritardo in coda)

Store and forward nei nodi  
(Ritardo in coda)

# Ritardo in coda

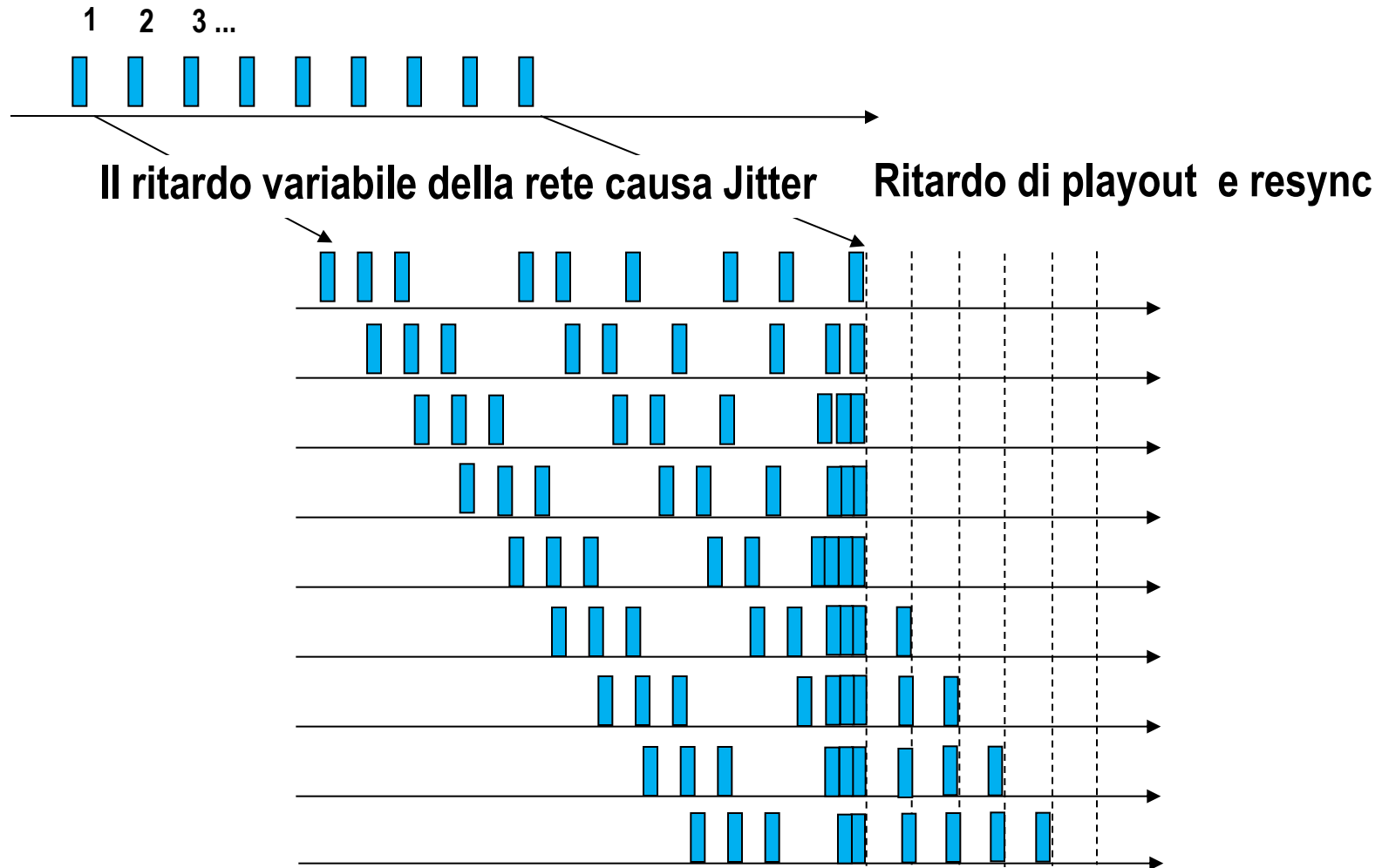
---

- La teoria prova che il ritardo medio nella coda di trasmissione è proporzionale al tempo di trasmissione  $T$
- Un carico accettabile produce un ritardo pari a circa 1-2 volte  $T$



# Ritardo di playout

## Compensazione del jitter



# Ritardo di playout

## Compensazione del jitter

---

- Il ritardo di playout del primo pacchetto deve essere settato per compensare il jitter
- $T_{out}$  costante: settato a priori al valore di jitter massimo tollerato
- $T_{out}$  adattativo: variato sulla base del jitter osservato nel periodo di attività precedente

# Misure di ritardo

---

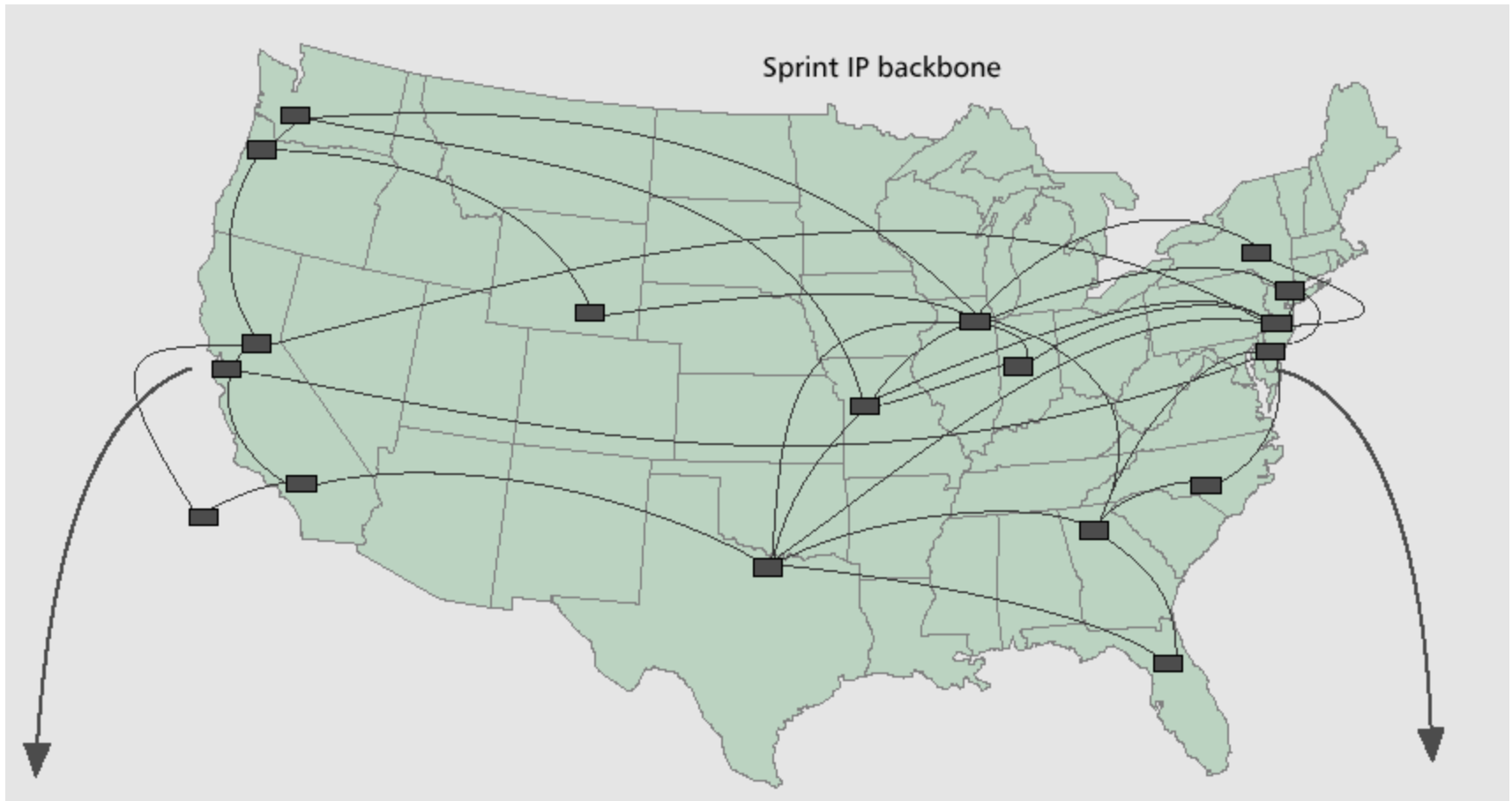
**Chuck Fraleigh, NetVMG - Sue Moon, KAIST - Bryan Lyles, Independent Consultant - Chase Cotton, Mujahid Khan, Deb Moll, Rob Rockell, and Ted Seely, Sprint - Christophe Diot, Intel Research**

*Packet-Level Traffic Measurements from the Sprint IP Backbone*

*IEEE Network Nov. Dec. 2003*

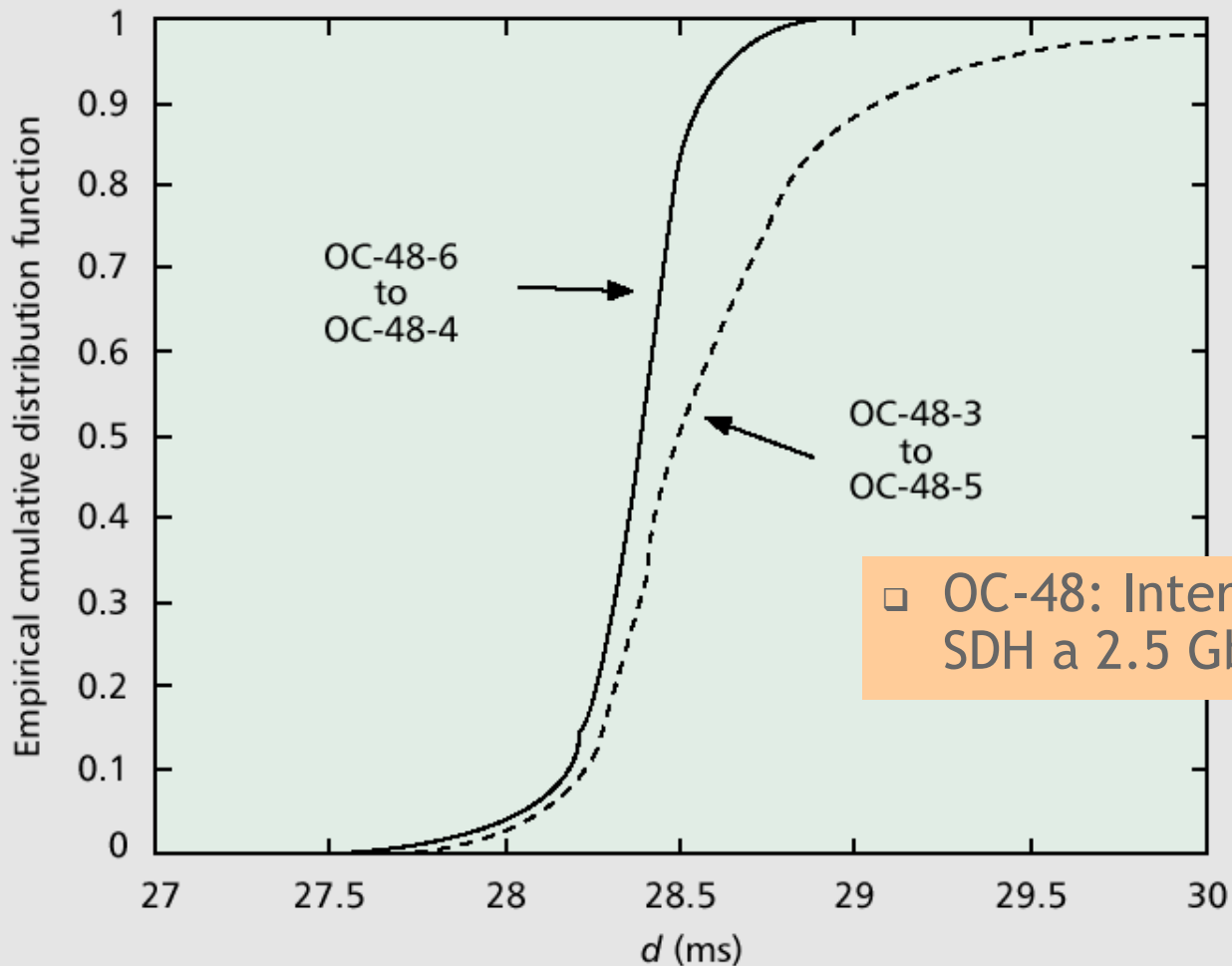
- U.S. transcontinental delay distributions are obtained between San Jose and New York, and reflect 200 million packet matches in a 1 h period
- Packets identified in these delay distributions crossed five POPs and eight core routers

# Backbone Sprint



# Distribuzione del ritardo

## San Jose – New York





# Distribuzione del ritardo

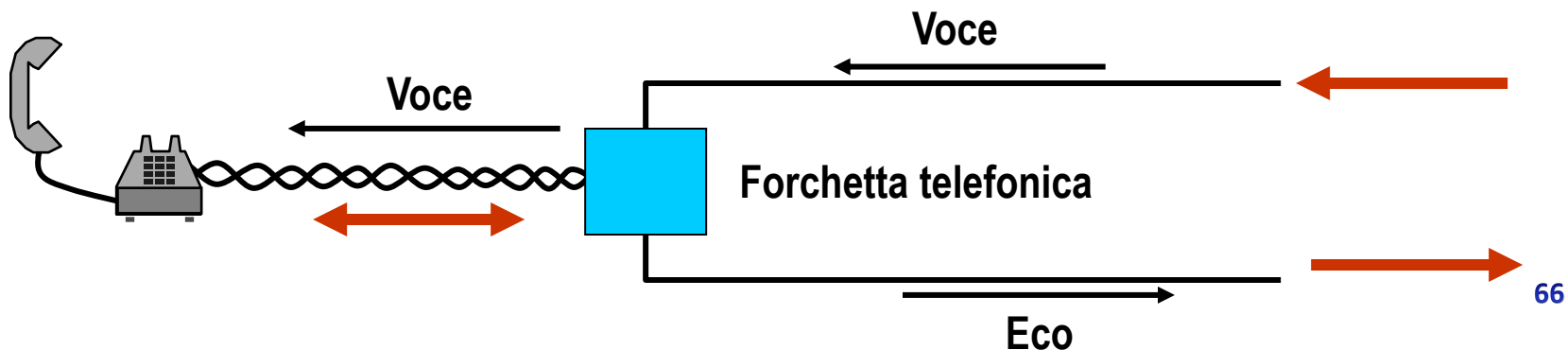
## San Jose – New York

---

- The minimum delays are 27.58 ms from OC-48-6 to OC-48-4 (from San Jose to New York), and 27.34 ms from OC-48-3 to OC-48-5 (from New York to San Jose)
- The average delays are 28.37 ms and 28.58, and the 99.9 percent delays are 28.99 ms and 31 ms, respectively
- The jitter on these paths is limited to less than 3 ms. This amount of jitter is not sufficient to impact the performance of delay-constrained applications such as media streaming or VoIP
- While we observe a very small number of packets that experienced delay above 100 ms. Router idiosyncrasies are identified as a cause of large delays
- The main contributing factor in network delay is the speed of light

# Il problema dell'eco

- L'eco può essere dovuto a
  - Accoppiamento tra speaker e microfono (lato trasmettitore)
  - Riflessioni del segnale dovute alla forchetta telefonica (solo in caso di chiamate su rete PSTN)
  - Rientro del segnale ricevuto nel microfono (lato ricevitore)
- Con ritardi maggiori di 25 ms l'eco va cancellato



# Il problema dell'eco

---

- Approcci per risolvere il problema dell'eco
  - *Eliminazione alla sorgente*: bilanciamento della forchetta telefonica (chiamate su PSTN)
  - *Soppressione d'eco*: sfrutta una tecnica di soppressione dei silenzi
  - *Cancellazione d'eco*: approccio matematico con sottrazione del segnale trasmesso dal segnale di ritorno

# Soppressione dei silenzi

---

- Nelle conversazioni duplex il canale è usato in media per il 50% del tempo e in più ci sono le pause del parlato
- Utilizzabile in reti a pacchetto per ridurre la banda media occupata e aumentare la capacità a disposizione
- I silenzi devono essere soppressi senza dare la sensazione di «caduta della linea» (introduzione di un rumore di fondo)
- Esistono molti metodi di *silence-detection*
  - Decisione presa per ogni segmento vocale sulla base del livello di energia del segmento vocale stesso
  - Decisione con memoria (es. almeno due segmenti consecutivi sotto soglia per l'attivazione)
    - Soglie adattative