



Università degli Studi di Bergamo



**DIP. DI INGEGNERIA DELL'INFORMAZIONE E
METODI MATEMATICI**

RETI INTERNET MULTIMEDIALI

VoIP Introduzione, Codifica della voce

Il materiale presentato in questa sequenza di diapositive è stato cortesemente messo a disposizione dal Prof. Antonio Capone e dal Prof. Flaminio Borgonovo

VOIP

Introduzione

Perché IP Telephony

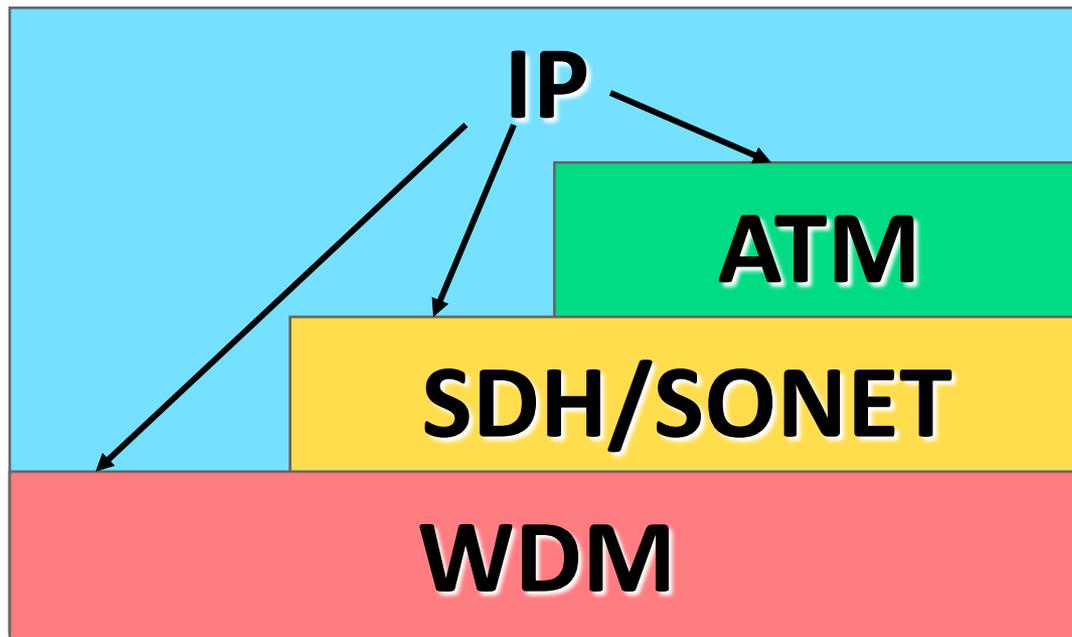
- L'integrazione di telefonia e IP rappresenta l'opportunità di realizzare un sistema globale di comunicazione multimediale
- Infatti l'aggiunta di video o altri segnali "stream" non cambia di molto il problema
- Riconosciuto il fallimento
 - dapprima di ISDN come strumento di integrazione end-to-end di dati su reti fonia ...
 - e successivamente di ATM ...
- rimane solo la possibilità di integrare la fonia su reti dati
- L'unica rete dati universale è basata su IP

Perché IP Telephony

- I vantaggi dell'integrazione (gestione unificata di problematiche e risorse e servizi di comunicazione) si sommano ai vantaggi dell'utilizzo di IP:
 - Crescita
 - Ubiquità
- Metodo di trasporto e interfaccia utente universali
- Metodi di scoperta delle risorse scalabili
- 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

Architetture di Backbone IP

- L' uso di IP su lunga distanza diventa sempre più efficiente



Vantaggi per l'utente

- Utilizzo di apparati intelligenti (PC) che permettono di programmare direttamente servizi di chiamata e risposta
 - Programmazione di call forwarding.
 - Es: in funzione dell'ora diretta a questo numero e, per mancata risposta, diretta su voice mail
- Mobilità e portabilità del numero
 - L'identità dell'utente sganciata dall'indirizzo di rete consente l'indirizzamento alla persona ovunque esso sia

Vantaggi per l'utente

- Messaggistica unificata
 - Integrazione E-mail, voice mail, short messages,
 - Scambio di dati durante la chiamata, sessioni in video
 - Multiconferenza, numeri personali
- Accesso remoto alla Intranet
 - Completo utilizzo dell'Intranet, compresa la telefonia
 - Sganciamento dallo strato fisico
 - Possibilità di connessioni diverse contemporaneamente
 - Da casa per telelavoro

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 privato

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

Vantaggi per l'utente pubblico

- Toll bypass
- Linee multiple
- Click to dial
- Servizi avanzati di reperimento risorse
- Messaggistica universale
- Integrazione servizi di rete intelligente (IN) e Internet

Esempi di vantaggi per i carrier

- Integrazione delle reti PSTN, Wireless, IP
- Offerta di nuovi servizi
- Roaming
- Servizi avanzati di reperimento risorse
- Servizi alla persona
- Messaggistica universale
- Integrazione servizi di rete intelligente (IN) e Internet

Enti di standardizzazione

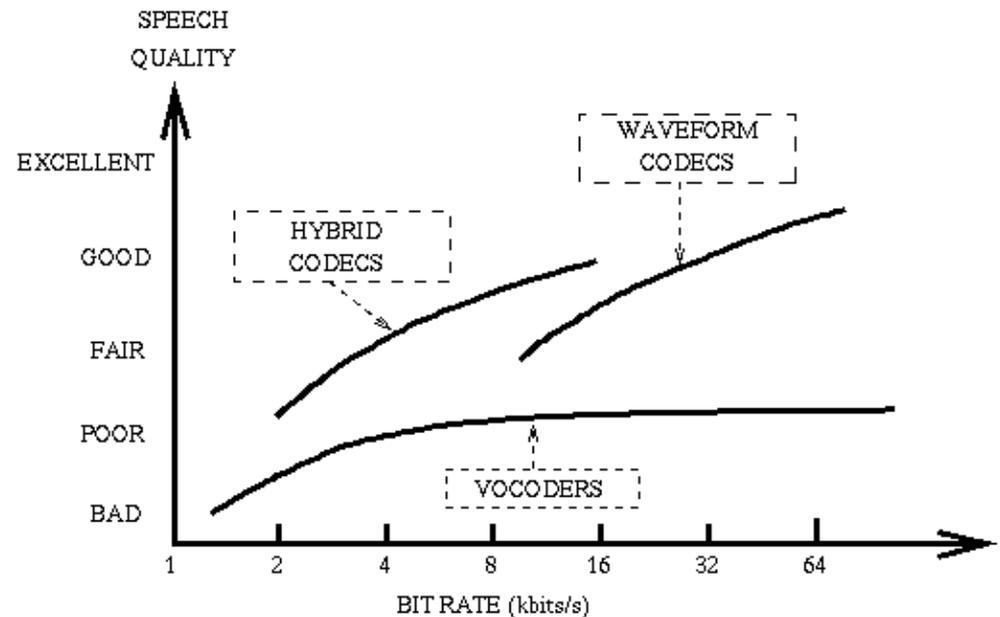
- International Telecommunication Union (ITU)
- Internet Engineering Task Force (IETF)
 - AVT: RTP
 - SIP: SIP per IP Telephony
 - IPTEL: CPL and GLP
 - SIPPING: estensioni SIP
 - SIMPLE: SIP per Instant Messagging
 - PINT (PIN): PSTN-Internet services
 - MEGACO: Media gateway control
 - SIGTRAN: Telephony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

VOIP

Codifica della voce

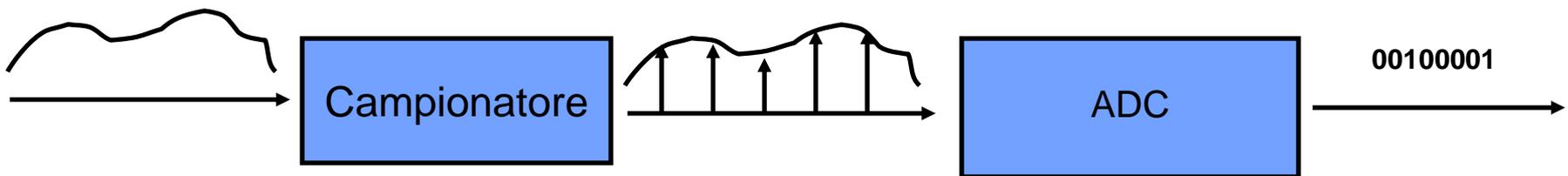
Codificatori Vocali

- Funzione: trasformano la voce in un flusso di bit
- Tipologie:
 - Waveform codecs
 - Source codecs (vocoders)
 - Hybrid codecs



Waveform Codecs

- Codificano il segnale di banda B senza perdite
 - Teorema di Nyquist: campionamento a $2B$
 - telefonia classica $B = 4 \text{ KHz}$
 - Quantizzazione (n. di livelli discreto e finito)
 - massimo rumore di quantizzazione tollerabile
 - Telefonia: $8 \text{ bit/campione} = 64 \text{ kb/s}$



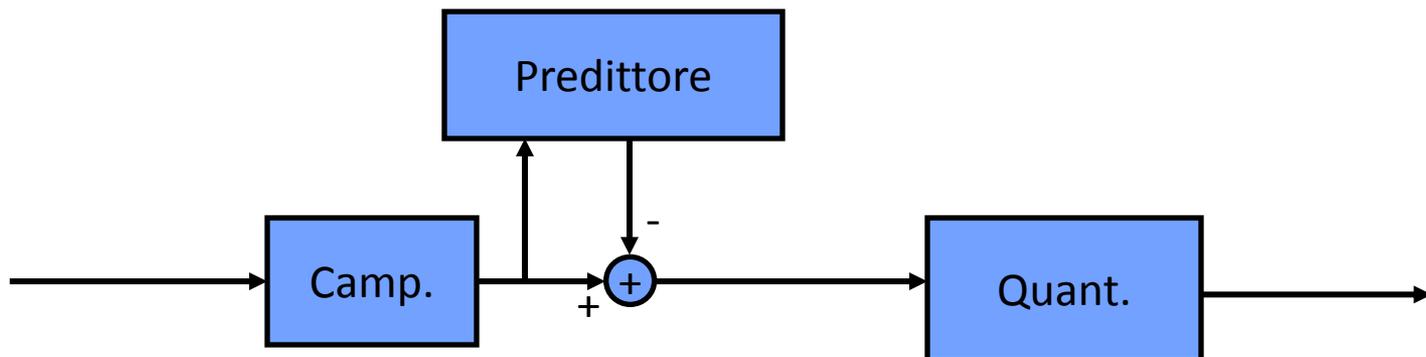
alta qualità, bassa complessità, basso ritardo (1 campione),
robustezza agli errori e al rumore di fondo

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)
 - per America (m-law) e
 - Europa (A-law)
 - Regole di conversione standard

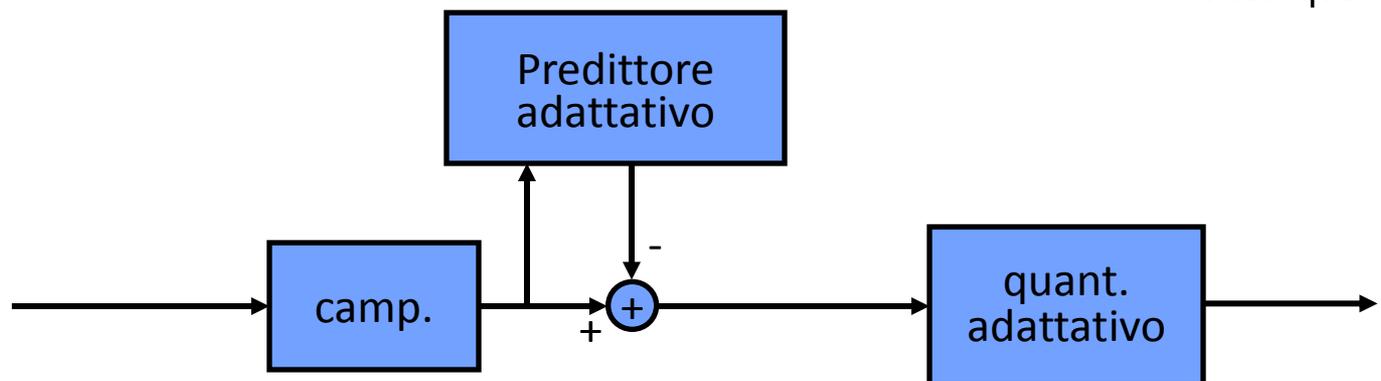
Differential PCM (DPCM)

- 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



Adaptive DPCM (ADPCM)

- Le prestazioni migliorano se predittore e quantizzatore sono adattativi
- 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



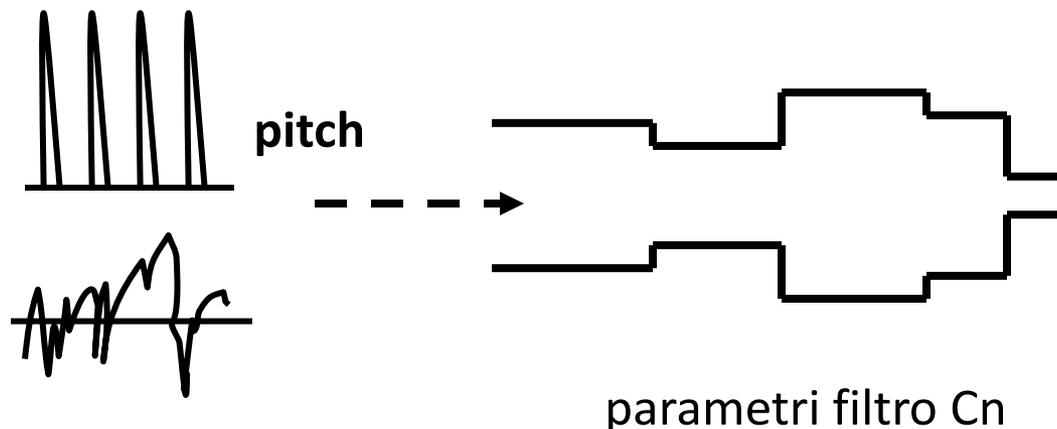
Bassa qualità

Source Codecs (vocoders)

- Si basano su modelli di generazione della voce umana
- I modelli permettono di “togliere la ridondanza” da segmenti vocali fino a un’informazione base 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 umani

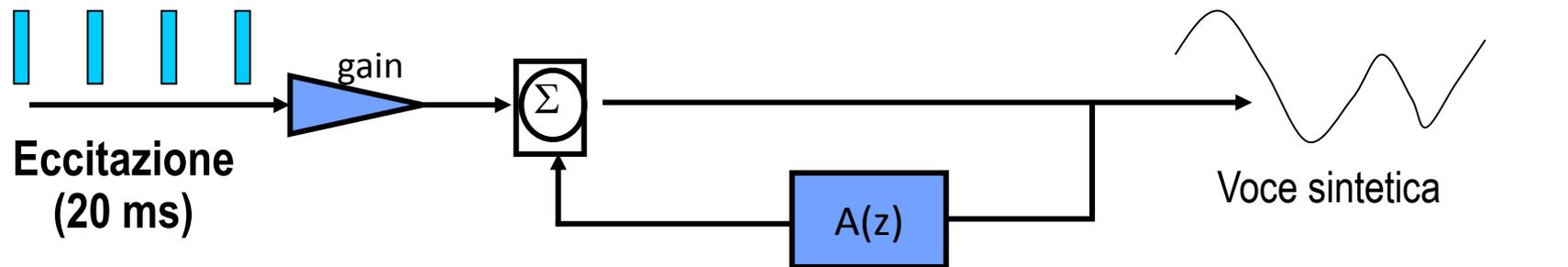
Modello a fonemi

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



Vocoder lineari (LPC)

- Modello della voce basato sul 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



La stima è effettuata minimizzando la varianza dell'errore

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

Vocoder lineari (LPC)

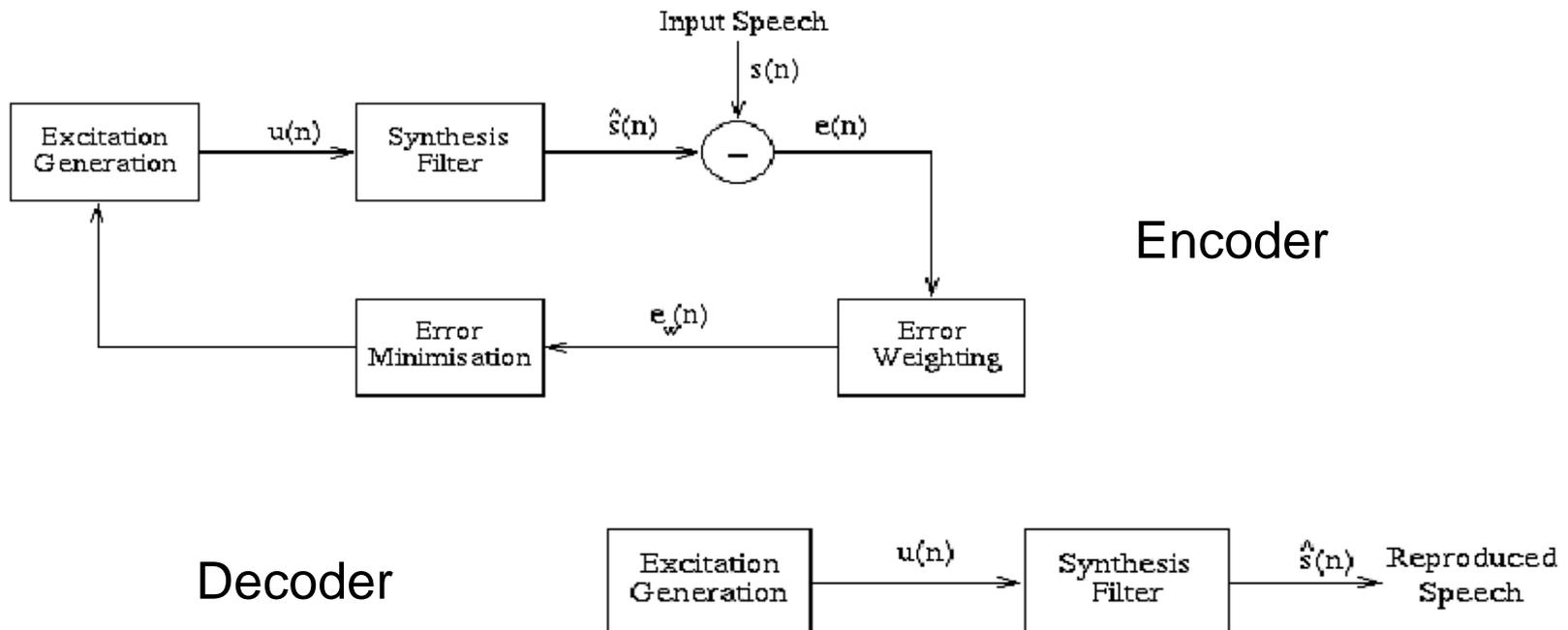
- In decodifica un sintetizzatore utilizza i parametri ricevuti per riprodurre il segnale
- Ritardi elevati: segmentazione, analisi, sintesi
- Qualità: intelligibile ma non naturale (limiti modello + problemi con rumori di fondo)
- Bit rate basso: < 2.4 kbit/s

Modelli di Eccitazione

- Eccitazione classica LPC ad due stati (es. LPC-10)
 - gain, pitch, flag voice/unvoiced
- Eccitazione mista
 - eccitazione periodica e pseudo-random insieme
 - 2 filtri di sintesi
 - frequenze basse → eccitazione periodica
 - frequenze alte → eccitazione pseudo casuale
- Eccitazione residuale
 - Eccitazione ideale: $e(n)$
 - Codifica a basso bit-rate di $e(n)$
 - Non e' più un vero vocoder ... Verso codificatori ibridi

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 (MPLP), 1982
 - Impulsi non uniformi di ampiezze variabili
 - Posizione e ampiezza di ciascun impulso sono determinate con procedura iterativa che minimizza una funzione dell'errore
 - es. MPLP 9.6 kbit/s di BTI per servizio Skyphone
- Regular Pulse Excitation (RPE)
 - Sequenza di impulsi equispaziati
 - Parametri: posiz. Primo impulso e periodo
 - es. LPT-RPE GSM 13 kbit/s

Hybrid Codecs

- Code Excited Linear Prediction (CELP)
 - La sequenza di eccitazione viene scelta tra un insieme di sequenze (code-book)
 - Le sequenze nel code-book sono realizzazioni di processi gaussiani
- 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

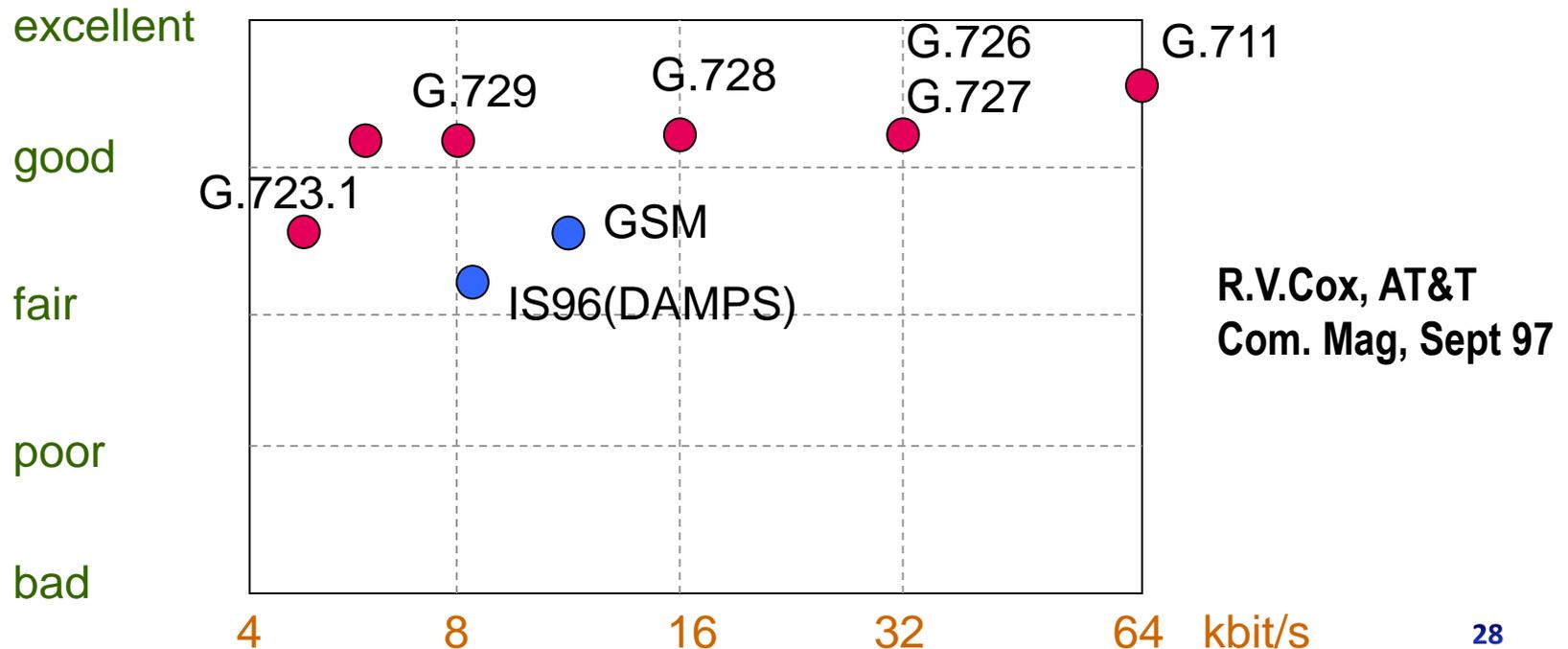
Principali Codifiche

Compressione	Year	Bit rate (kbit/s)	Frame size (ms)	Look ahead (ms)
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 MP-MLQ	1995	6.3	30	7.5
G.723.1 ACELP	1996	5.3	30	5
RPE-LTP (GSM)	1987	13	20	0

Prestazioni dei Codec

■ Mean Opinion Score

- 12-64 ascoltatori
- 5 livelli: excellent, good, fair, poor, bad



Qualità della voce ricostruita

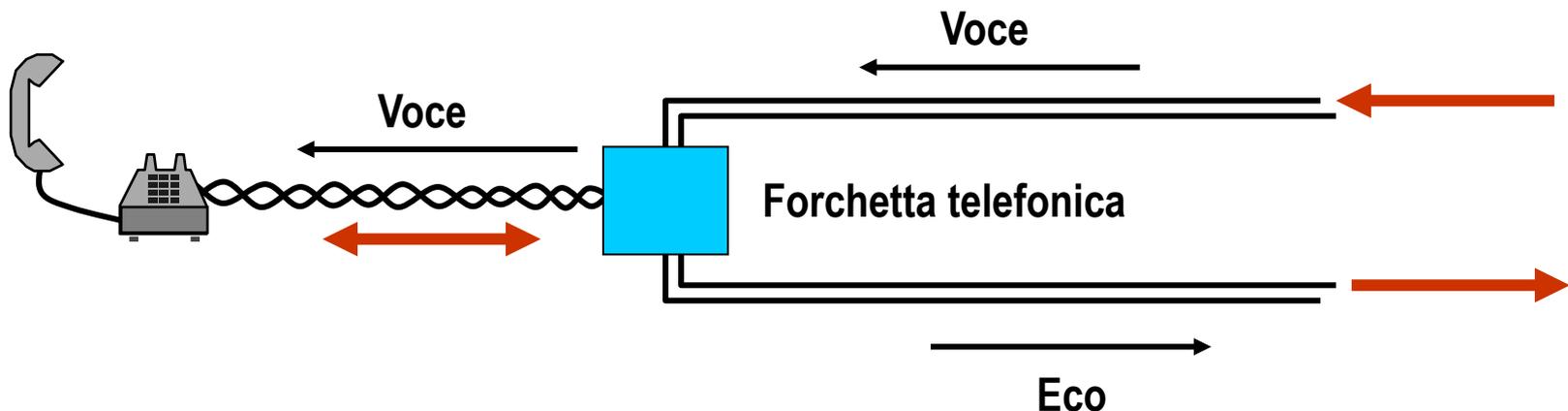
- Oltre alle caratteristiche intrinseche della codifica, nelle reti trasmissive la qualità della riproduzione è influenzata anche da
 - Ritardi di riproduzione eccessivi
 - Eco
 - Perdita totale o parziale di segmenti codificati
 - Codifica-decodifica multipla

Soppressione dei silenzi

- Nelle conversazioni duplex il canale è usato 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
- Esistono molti metodi di silence-detection
 - Decisione segmento per segmento basata su livello di energia
 - Decisione con memoria (es. almeno due segmenti consecutivi di segmenti sotto soglia per l'attivazione)
 - Soglie adattative
 - Meccanismi legati al codec

Cancellazione dell'eco

- L'eco può essere anche dovuto all'accoppiamento fra speaker e microfono
- Con ritardi maggiori di 25 ms l'eco va cancellato
- I cancellatori d'eco aggiungono ritardi e possono essere inglobati nella stessa scheda del codificatore

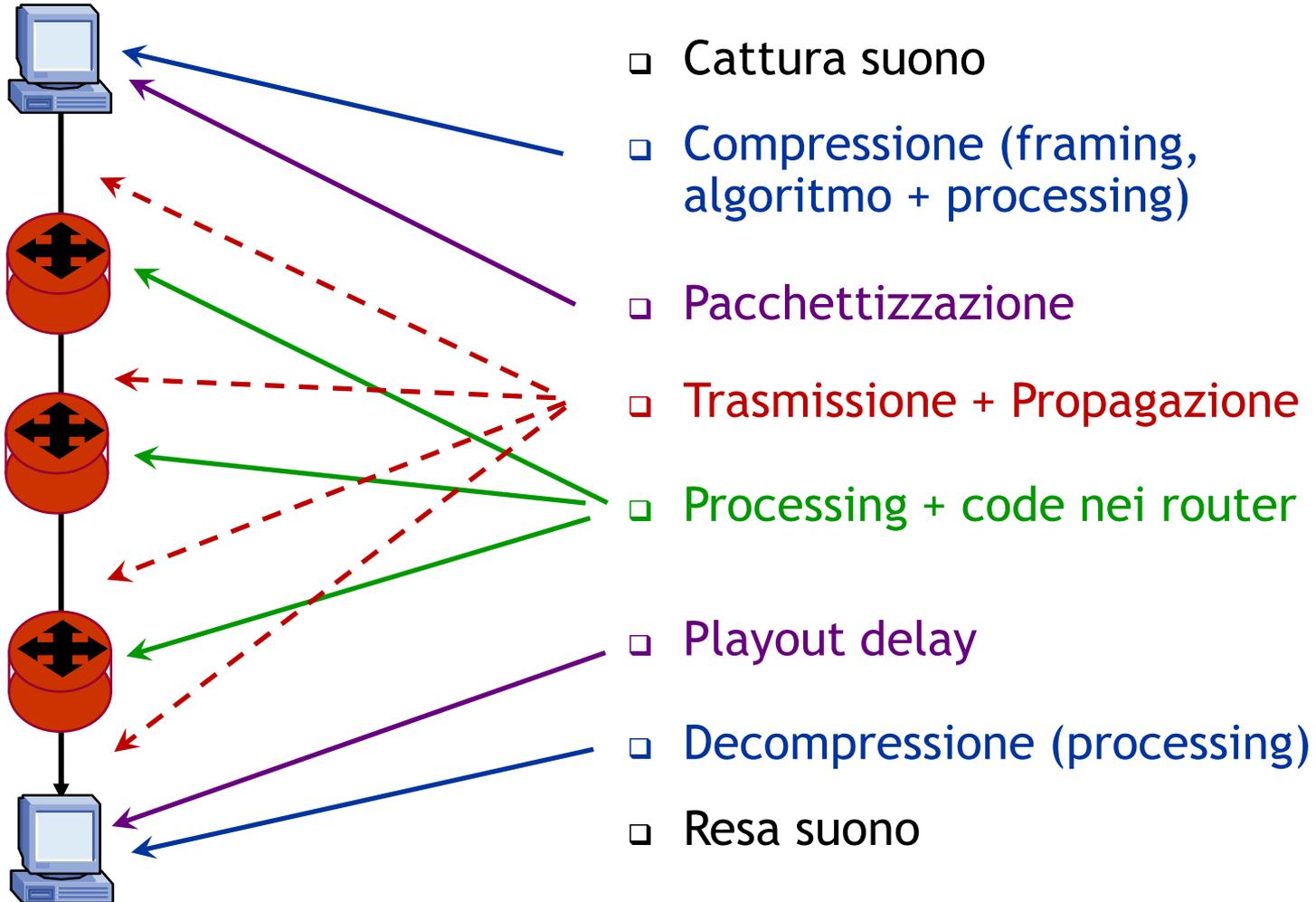


End-to-End QoS (ETSI Typhoon)

- NOTA: Sopra 25 ms di ritardo occorre buona compensazione dell'eco

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

Cause di ritardi in VoIP

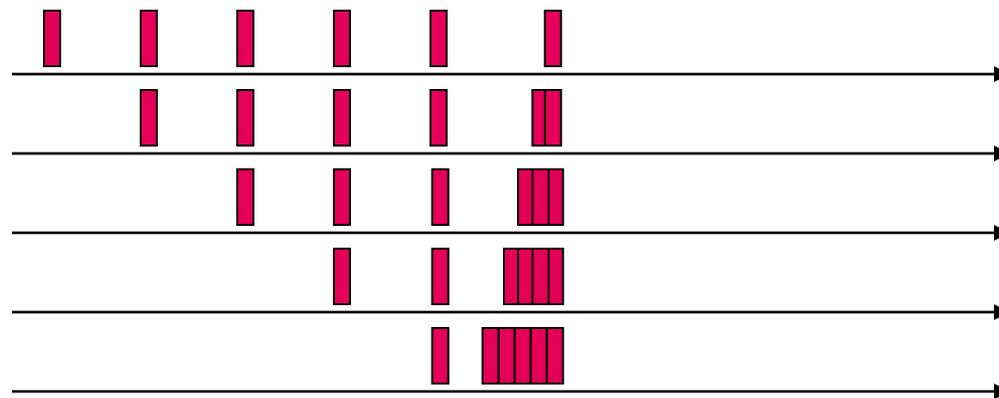


Ritardo di compressione

Compressione	Bit rate (kbit/s)	Frame size (ms)	Look ahead (ms)	Processing (ms)
G.711 PCM	64	0.125	0	0
G.726 ADPCM	32	1	0	0
G.728 LD-CELP	16	0.625	0	0
G.729 CS-ACELP	8	10	5	0-10
G.723.1 MP-MLQ	6.3	30	7.5	0-30
G.723.1 ACELP	5.3	30	5	0-30
RPE-LTP (GSM)	13	20	0	0-30

Pacchettizzazione

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



Ritardo di
pacchettizzazione

Ritardo di pacchettizzazione

- Dipende dalla lunghezza del pacchetto (ms di voce)
- Si compone in modo più o meno additivo al ritardo di compressione
- In generale: buona compressione + pacchettizzazione prendono 20-30 ms + processing

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

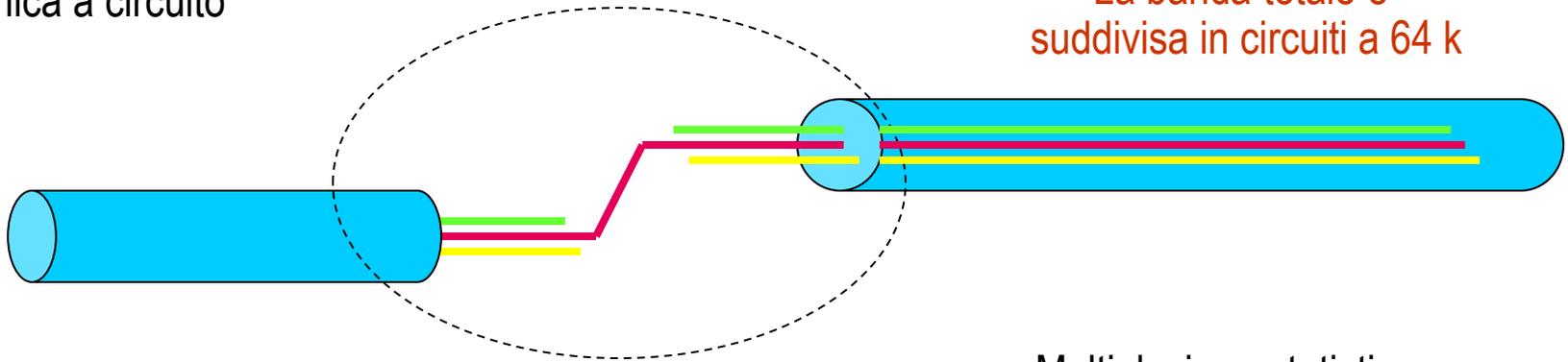
**X n. di link
attraversati**



Ritardo di propagazione: 5 ms / 1000 km

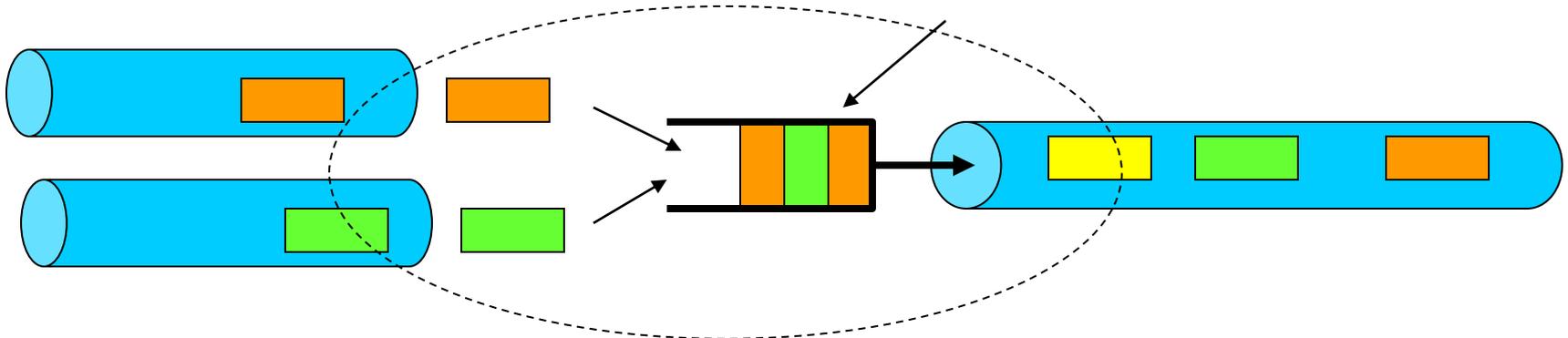
Ritardi in rete

Tecnica a circuito



La banda totale è
suddivisa in circuiti a 64 k

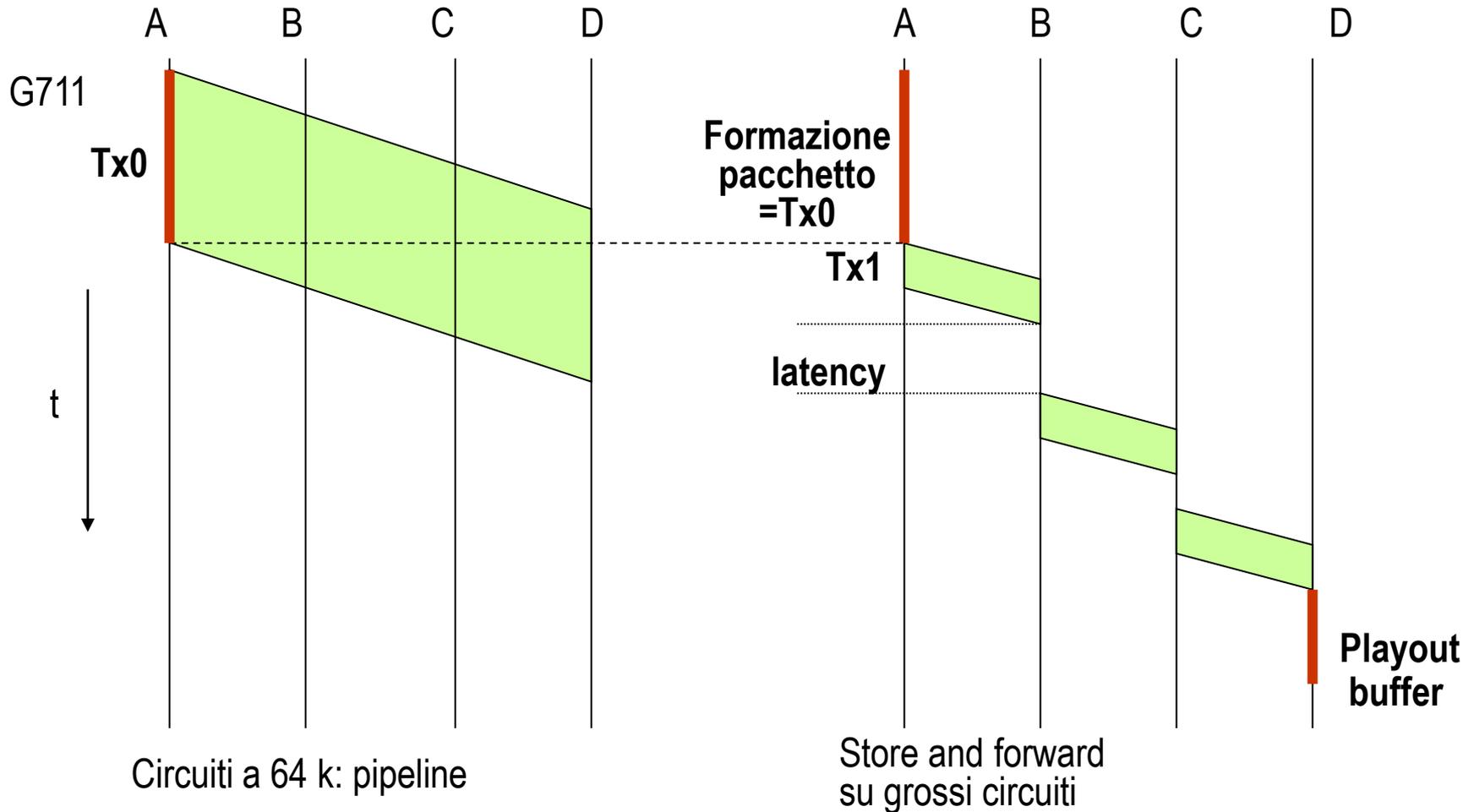
Tecnica a pacchetto



Multiplicazione statistica

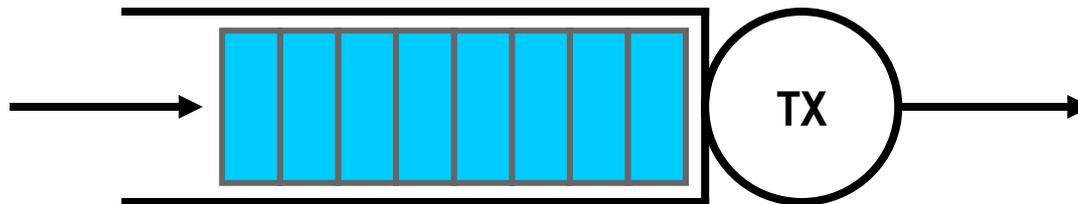
Ritardi

Effetti della velocità dei link

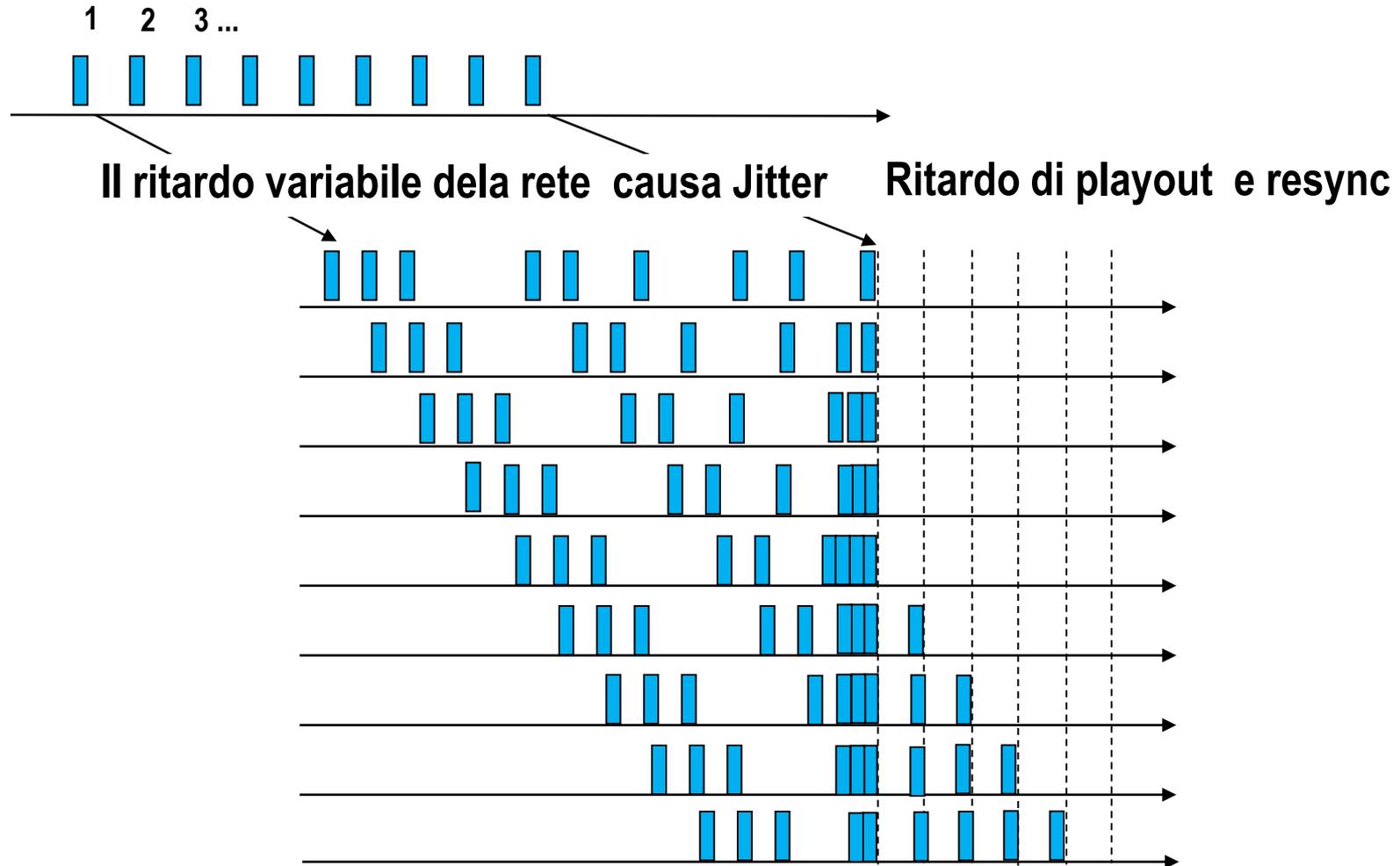


Ritardo in coda

- La teoria prova che i ritardi medio e percentile nella coda di trasmissione sono proporzionali al tempo di trasmissione T
- Un carico accettabile produce un ritardo pari a circa 1-2 volte T



Compensazione del jitter



Compensazione del jitter

- Il ritardo di playout del primo pacchetto deve essere settato per compensare il jitter massimo
- T_{out} costante: settato a priori al valore massimo tollerato
- T_{out} adattativo: variato durante i periodi di silenzio sulla base dei ritardi osservati nel periodo di attività precedente

Valutazioni di ritardo

- Nella rete telefonica, il cammino massimo USA da costa a costa offre un ritardo misurato (propagazione e apparati) di 90 ms
- Nel backbone le velocità sono elevate e i ritardi di trasmissione a pacchetto, e nelle relative code, sono frazioni di ms
- Ritardo di codifica, pacchettizzazione, decodifica, e processing 50 ms
- Assumendo un ritardo di buona qualità (200 ms max) restano 60 ms per ritardo nelle code e nel playout buffer

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

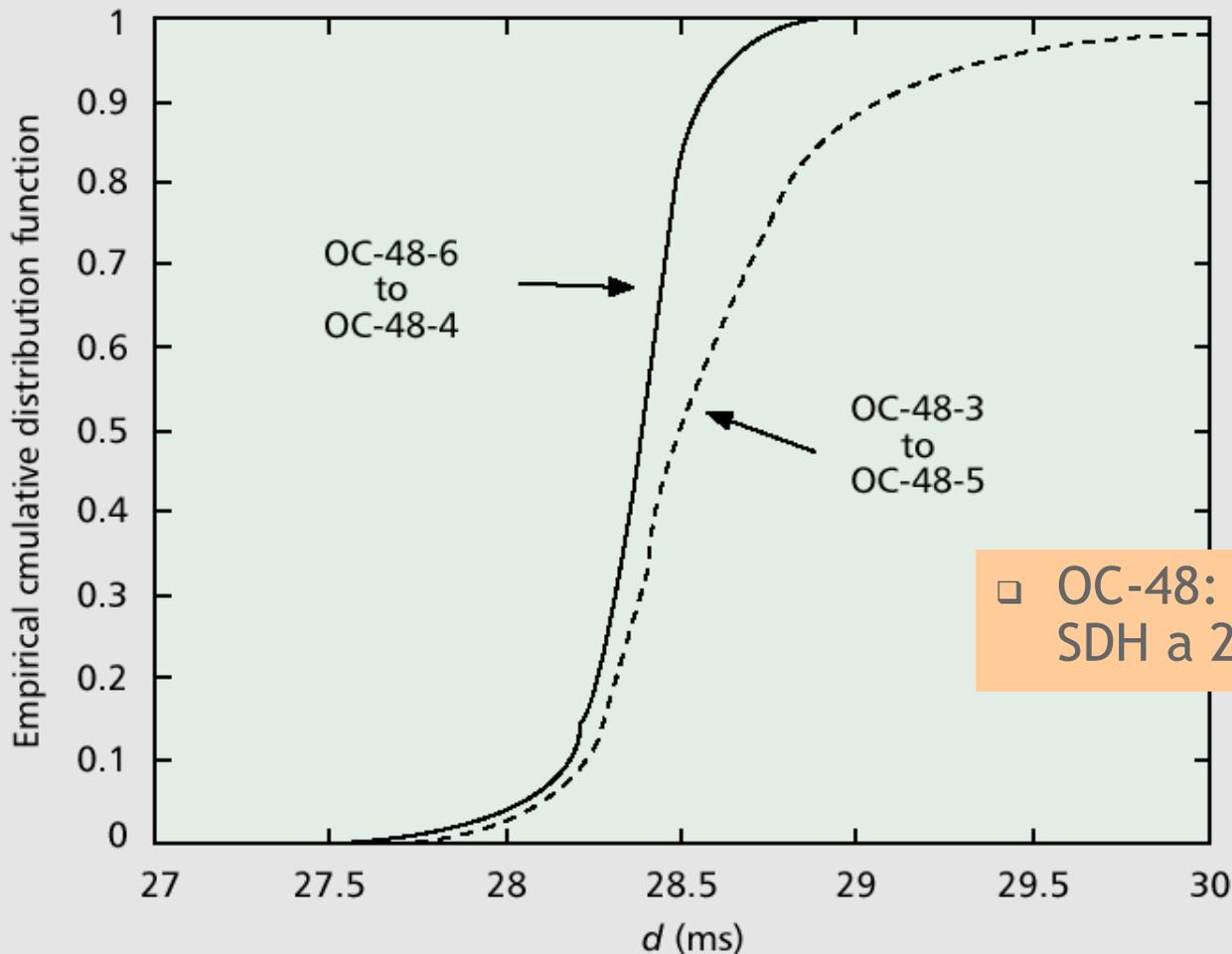
Packet-Level Traffic Measurements from the Sprint IP Backbone

IEEE Network Nov. Dec. 2003

- U.S. transcontinental delay distributions in Fig. 11 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.

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

Valutazioni di Qualità

- Mean Opinion Score (MOS)
 - Misura la chiarezza della voce
 - Basato sull'opinione di un gran numero di ascoltatori
 - Standardizzato nella raccomandazione P.800

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

- Perceptual Speech Quality Measurement (PSQM)
 - Raccomandazione P.861 del 1996
 - Utilizza un confronto matematico fra segnale elaborato e segnale degradato operato attraverso operazioni che tengono conto delle modalità percettive
 - I risultati sono tarati in modo da predire il MOS

Difetti:

- pensato per i CODEC
- non riscontra i difetti tipici delle reti (errori del canale, perdita di pacchetti, clipping..)

Valutazioni di Qualità

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

Valutazioni di Qualità

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

