



# Signal Processing for Packet Voice Telephony

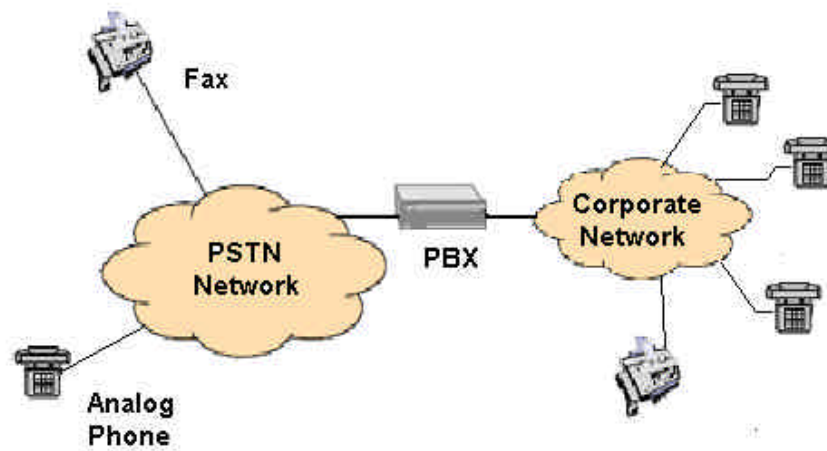
IEEE OEB ComSoc, San Ramon  
May 20, 2004

Krishna V V  
Principal Solutions Architect  
DSP Products Development  
LSI Logic

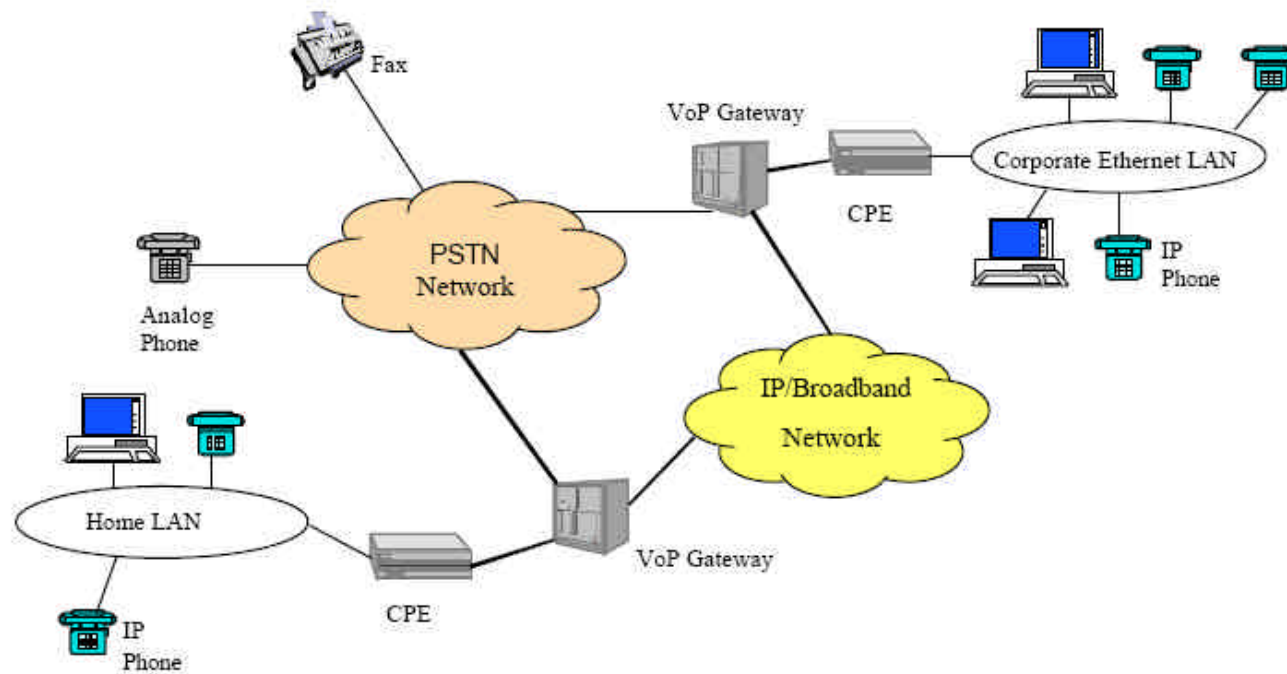
# Agenda

- About LSI Logic, briefly ...
- Packet Voice Networks Overview
- Signal Processing Layer Internals
  - Voice-specific functions
    - Speech codecs; packet loss concealment; echo cancellation; silence suppression; ...
  - Other related functions
- Processing needs for different contexts
- A few trends & issues

# Voice Networks: Past



# Voice Networks: Present



# Telling Quotes

“If you don’t do it, next year or the one after that you won’t be playing in the game”

Richard Notebaert (CEO, Qwest)

---

(while announcing Qwest’s plan to offer low-cost IP based phone service, 2003).

# Telling Quotes

“As of three weeks ago, all the long-distance (voice) traffic in Italy is carried over the IP network”

Stefano Pileri (Head, Domestic Network, Telecom Italia)

---

(announcement made at ITU Telecom Conference, Geneva October 15, 2003).

# Telling Quotes

“Telecom may be heading the way of DRAMs, where the price is set by the most idiotic competitor. ... It is a race to the bottom, and the bottom in this case is free service”

Robert Lucky

---

(Keynote Speech, Communications Design Conference, San Francisco, March 31, 2004).

# Not so SIPlе News

Singtel + SIPphone Press Release (April 5, 2004)  
(home.singtel.com/news\_centre/news\_releases/2004-04-05.asp)





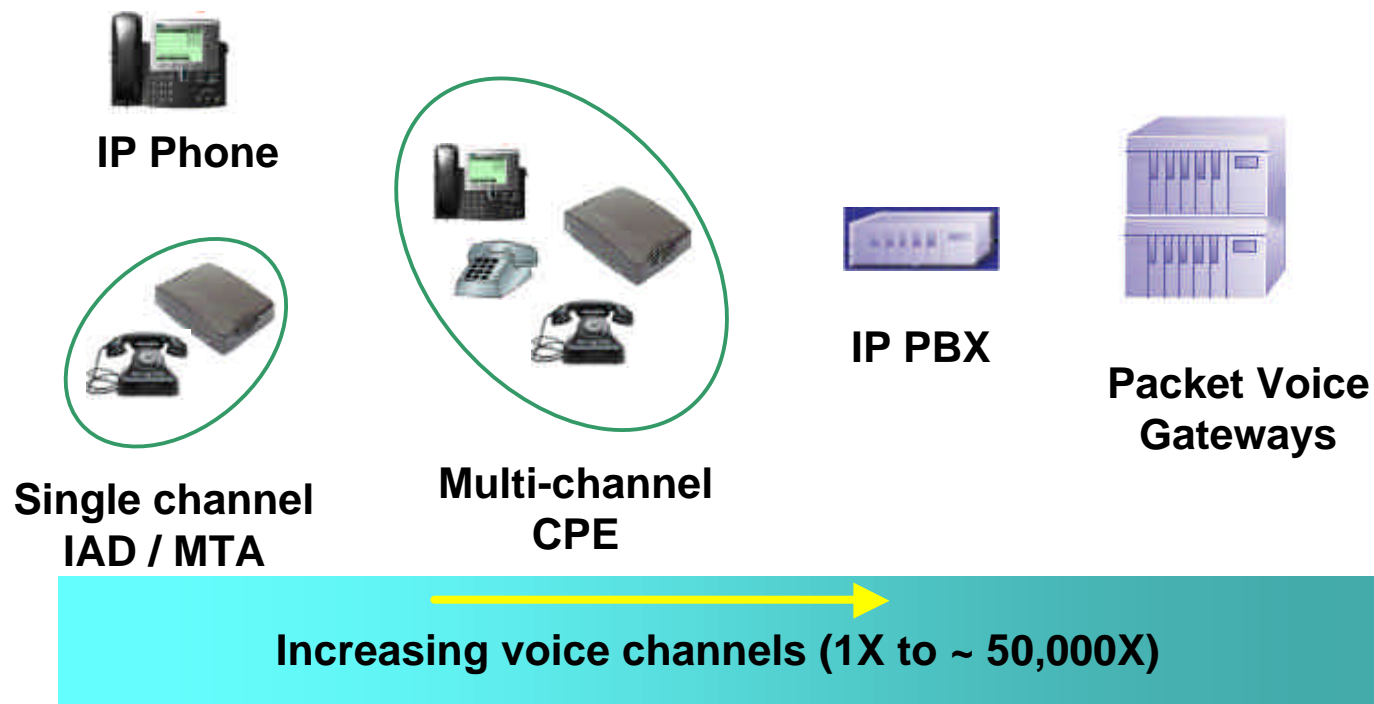
# Packet Pathways

- **Wired**
  - Cable Modems
  - DSL Networks
  - Corporate Ethernet LANS
  - Managed / Public Wide-area Networks (WANs)
- **Wireless**
  - WLANS (WiFi)
  - Satellite IP Networks (including VSATs)
  - BWA (WiMax) (?)

# Protocol Menu

- Protocols for VoIP Inter-operability:
  - MGCP (CableLabs / PacketCable)
  - H.323 (ITU, videoconferencing)
  - H.248 /Megaco (ITU)
  - SIP (IETF)
- Others key protocols (Quality, ...):
  - RTP, RTCP
  - UDP
  - RSVP, DiffServ, ...

# Packet Voice “Boxes”



# VOIP Market: DSP Slice

	2001	2006	
Annual Growth (voice channels):	12 Million	560 Million	
DSP Revenue:	US\$129M	US\$1400M	
ASP (DSP + S/W)	\$25	\$225	} Gateway Market Segment
Channels / DSP	< 5	~ 88	
Average cost / channel	\$5.70	\$2.60	

Source: VOIP & Packet Voice DSP Markets,  
Forward Concepts, AZ (April 2002)

# Packet Voice: Key Hurdles

- Delay
  - Typical end-to-end delays around 100-200ms
- Packet Jitter
  - Typical arrival time jitter around 20-50ms
- Packet Loss
  - Typical losses around 1-2%

# Delay: G.114 Guidelines

One-way Delay	ITU-T Classification (with echo "adequately controlled")
< 150ms	Mostly acceptable.
150-400ms	Acceptable (maybe).
> 400ms	Unacceptable (in general).

## TYPICAL DELAYS

**Terrestrial, national long distance PSTN: < 50ms**  
**Terrestrial, international PSTN: ~ 100ms**  
**Cellular: Mobile to PSTN: ~ 150ms**  
**Cellular: Mobile to Mobile: ~ 300 – 400ms**

# Delay in Packet Networks

Overall delay break-up:

Speech Codec	0.2 – 68ms	Low delay for PCM, ADPCM, G.728, ...
Packetization	5 – 30ms	
Interleaving	0 – 60ms	Optional
Transmission	25 – 150ms	
Jitter Buffer	50 – 100ms	
Total	~80 – 400ms	<b>Typical: 100 – 200ms</b>

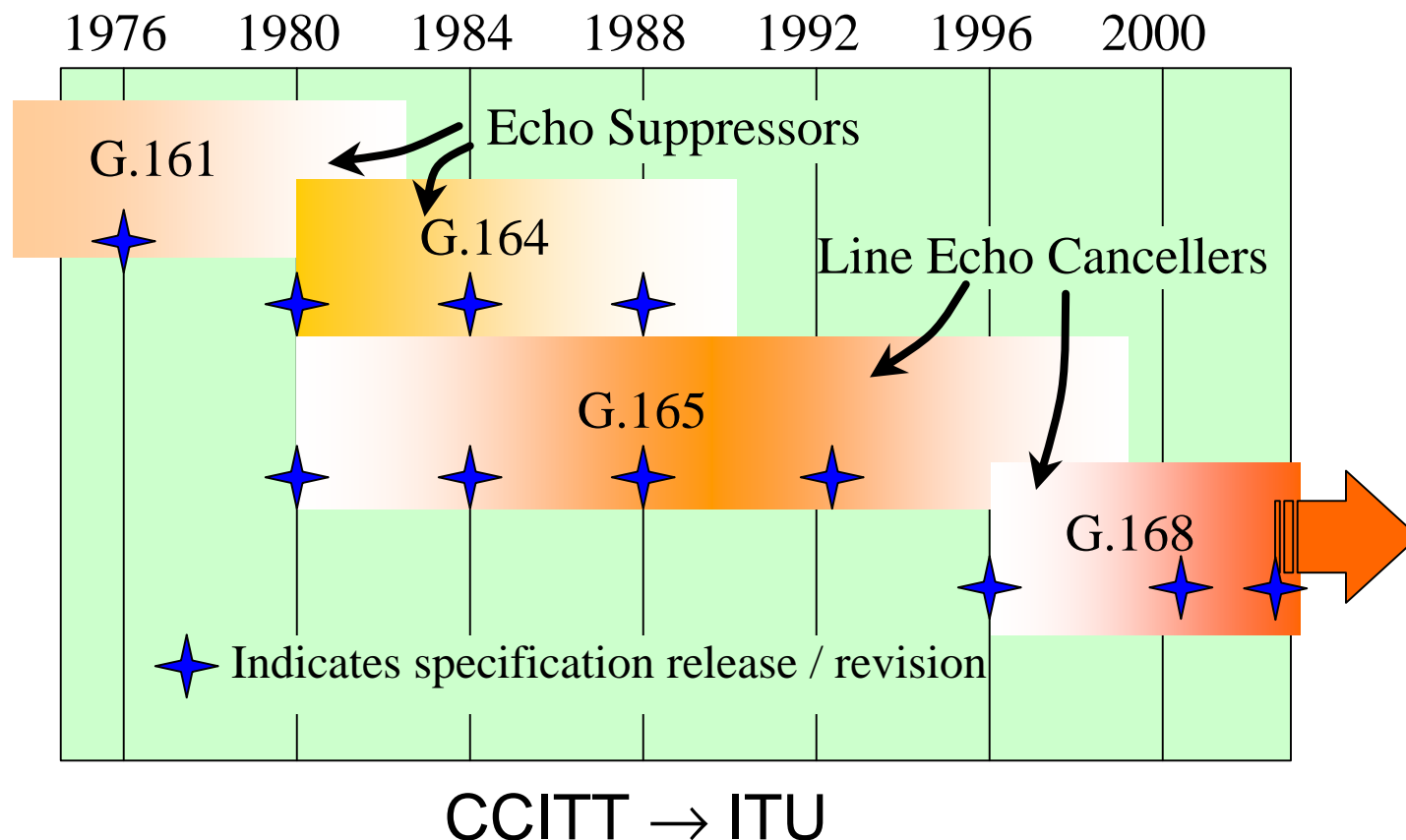
Revised from “Internet Telephony: Going like crazy”, by  
G. Thomsen, Y. Jani, IEEE Spectrum, May 2000.

# Handling Delay (Echo)

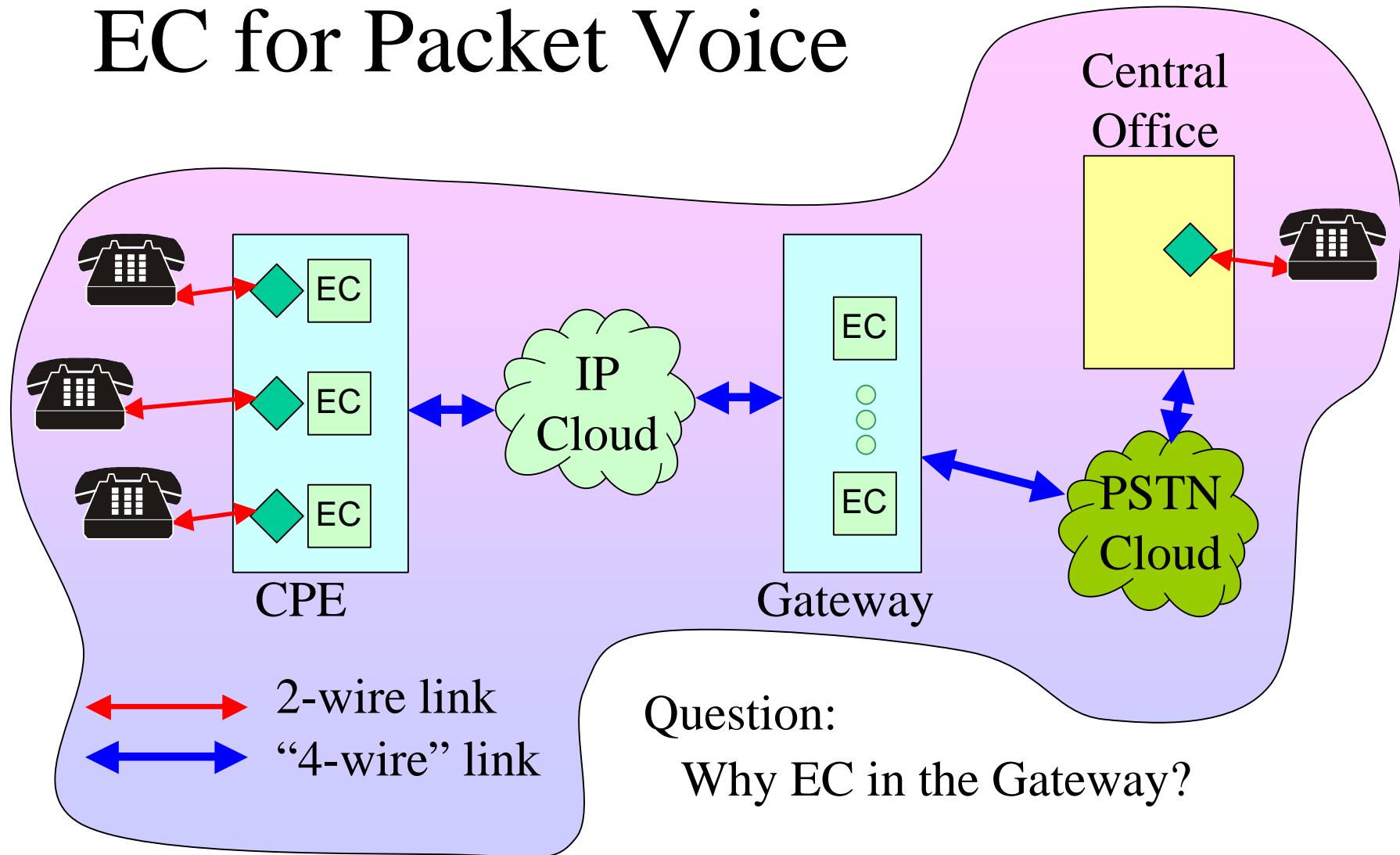
- One-way delays in packet voice networks > 100ms
- As recommended in ITU-T G.131, a network echo canceller (EC) is required.
- EC required only for:
  - PSTN interfaces on voice gateways
  - Analog phone (SLIC) interfaces on CPEs
- EC not required for digital IP phones (AEC is a different option)
- EC tail length – a much misused parameter
- ITU-T G.168 EC was initially developed for PSTN. Can it be applied as-is for packet voice networks?



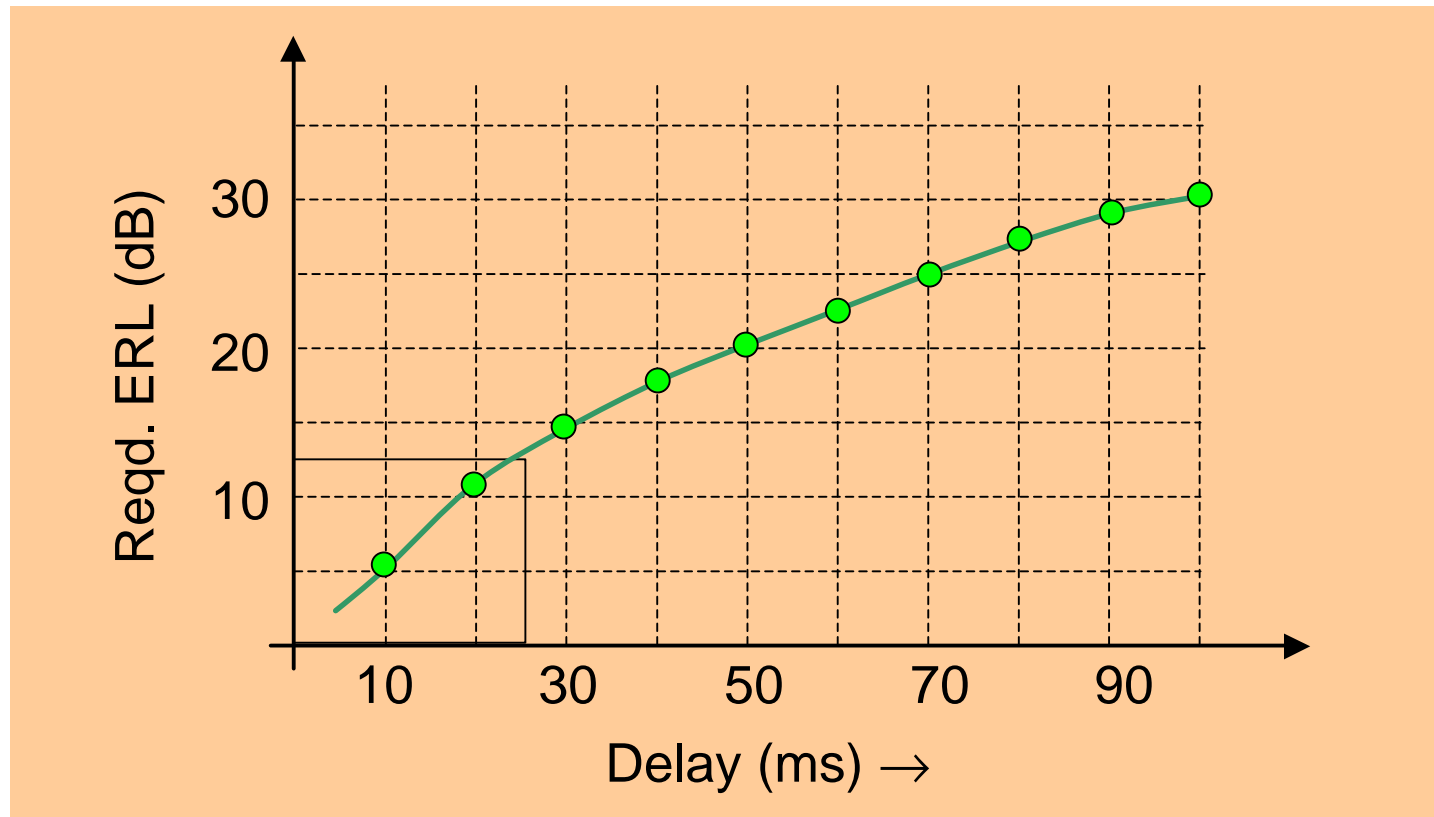
# Tackling Echo: ITU Standards



# EC for Packet Voice

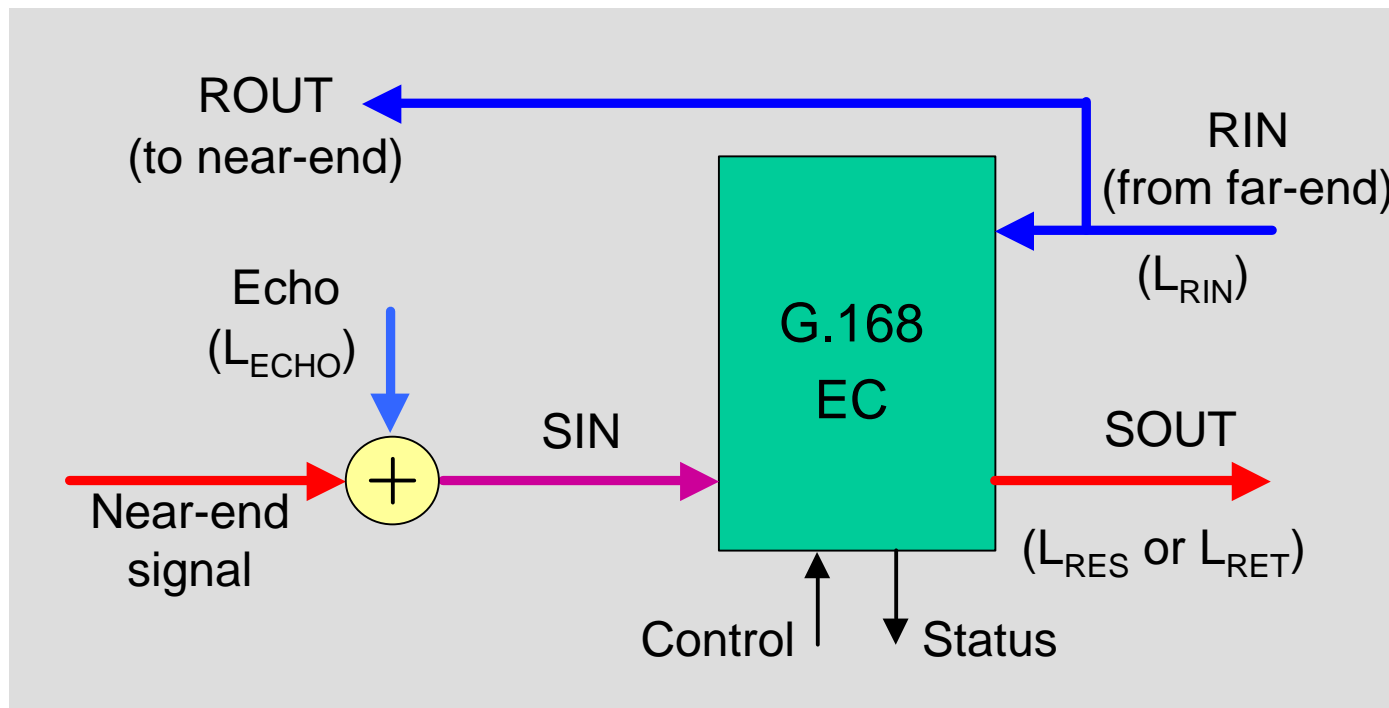


# Echo Level and Delay



ERL data from Table 1.1, "Acoustic Signal Processing for Telecommunication", S. L. Gay and J. Beneste (Ed.s), Kluwer Academic Publishers (2000)

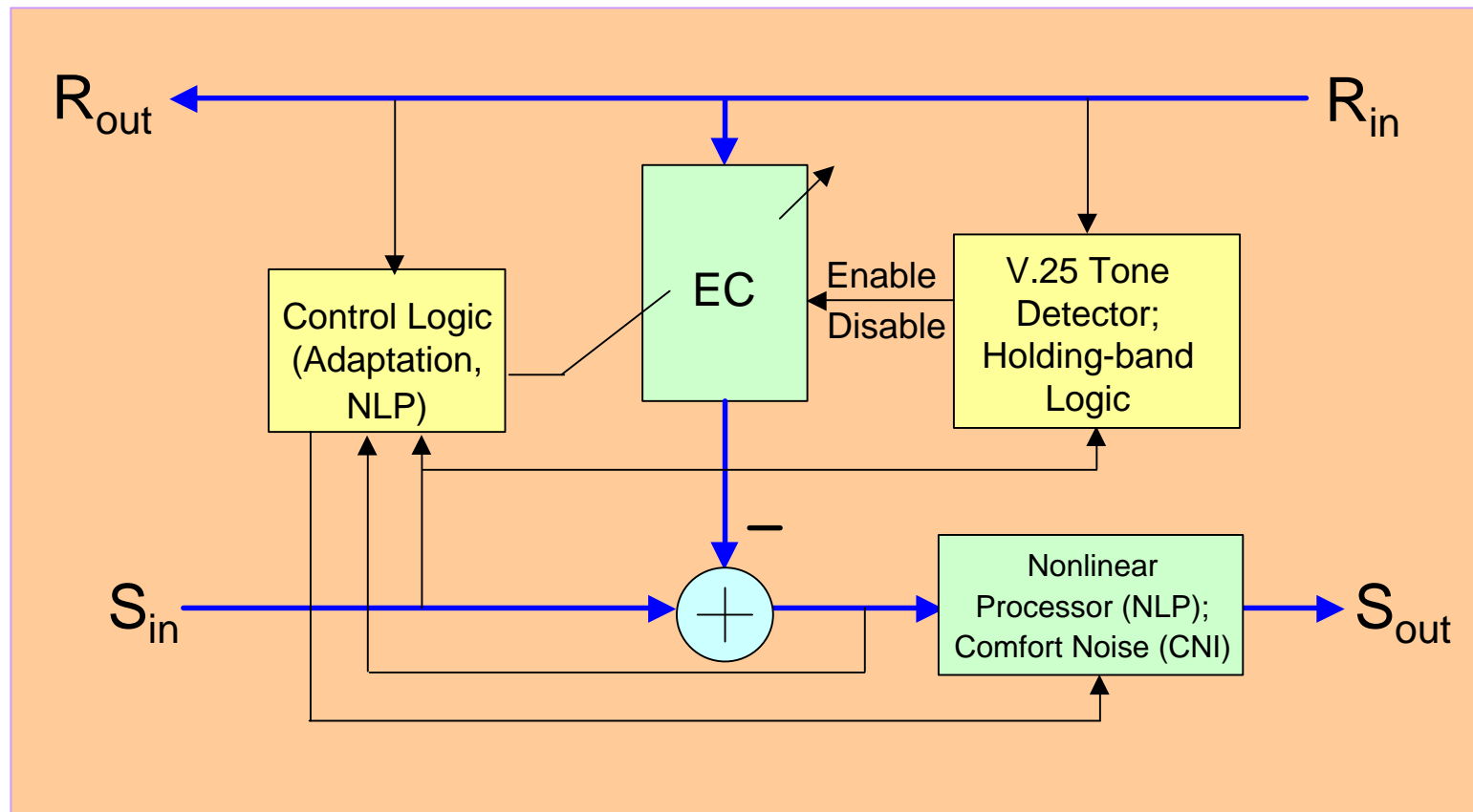
# EC – A Black-box View



$$ERL = L_{RIN} - L_{ECHO}$$

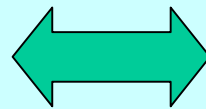
$$A_{COM} = L_{RIN} - L_{SOUT} \text{ (near-end signal absent)}$$

# G.168 EC Internals



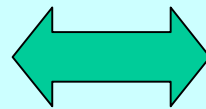
# Some EC Design Options

“Full tail”



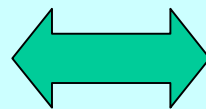
“Tail independent”  
or “Floating window”

Single filter with  
robust control

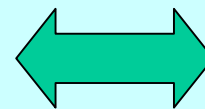


Double filter with  
simpler controls

Time  
domain

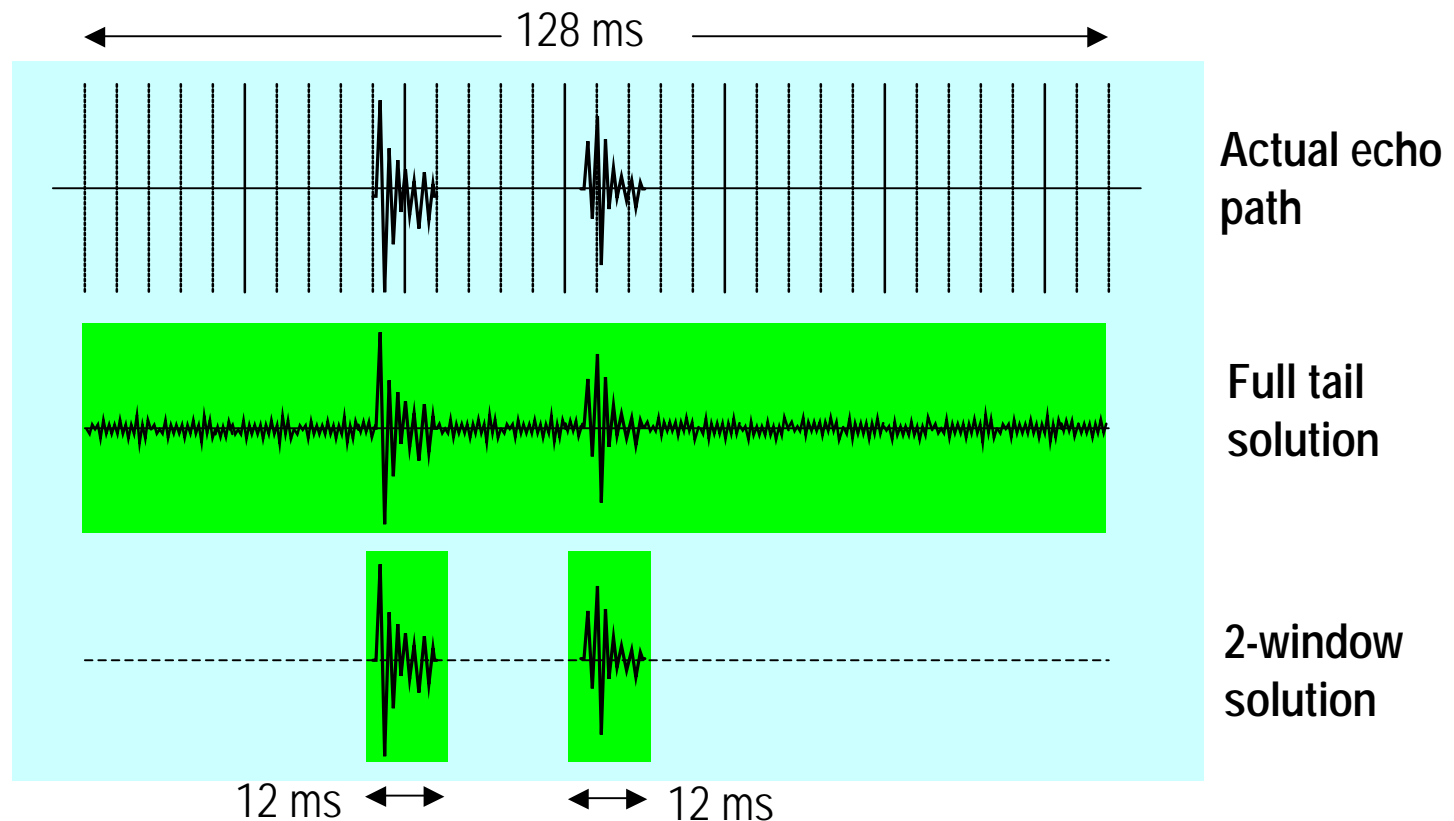


Transform  
domain



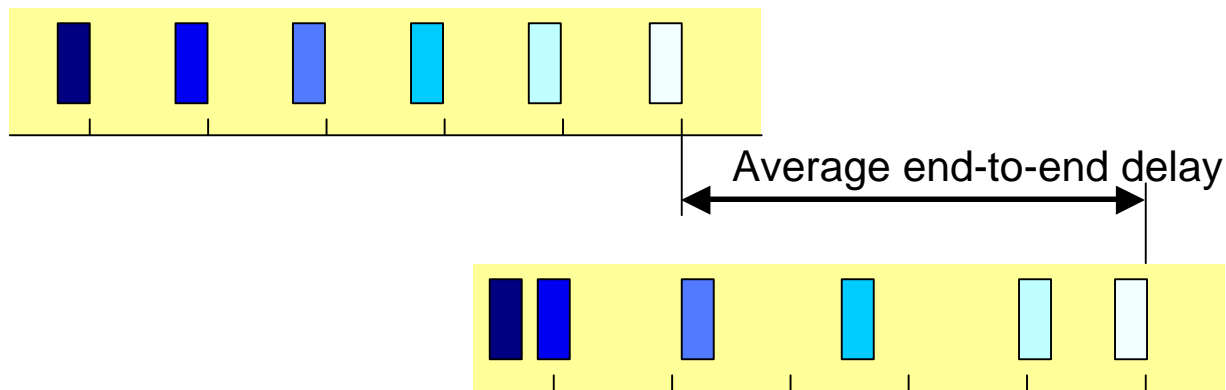
Subband  
structure

# Full Tail / Floating Window

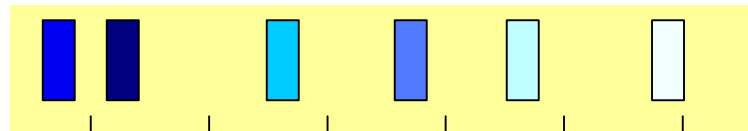


# Handling Packets

Packets assembled and ready to go ...



Packets arriving at destination with jitter.



Packets arriving at destination with jitter and out of order (likely on public networks).



# Jitter Buffer

- Evidently, jitter buffer is a crucial module in the receiver.
- Out-of-order packet arrivals can be sorted based on RTP time stamp.
- Trade-off of voice quality versus latency.
  - A small buffer helps minimize the extra latency, but drops packets that arrive too late.
- Adaptive jitter buffer that grows or shrinks as needed, is one solution.

# Packet Losses

- In addition to packet drops by jitter buffer, packet losses are likely due to
  - UDP (does not offer guarantee of delivery)
  - Network congestion (bandwidth)
  - Router overload (packet throughput)
- Up to 5% (or more?) packet losses considered likely
  - Even 1% packet loss degrades voice quality significantly
  - Packet Loss Concealment (PLC) is yet another essential module in the receiver
  - PLC is not sufficient to handle certain tone signals (DTMF digits, V.25 tone for EC disabling, etc.)

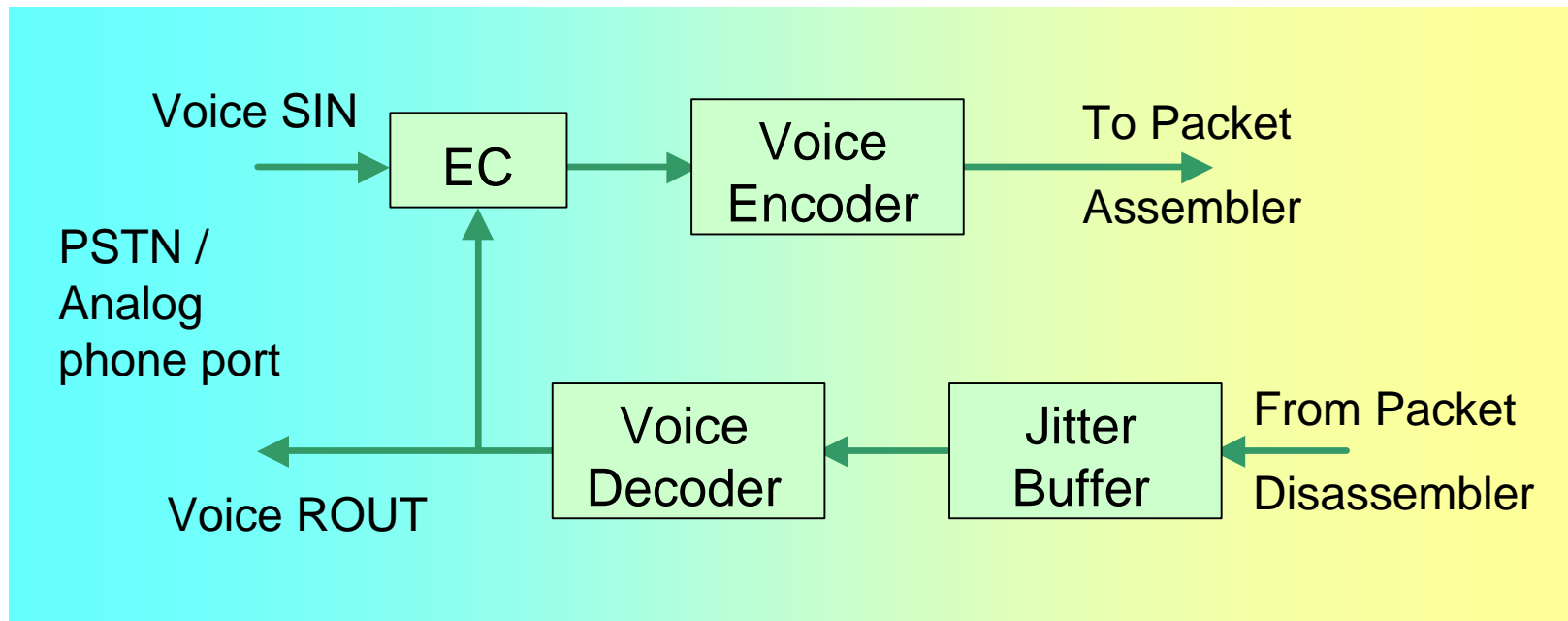
# Tone Relay

- Helps in reliable transfer of DTMF digits and other signaling tones (packet losses)
- Fast DTMF detection also avoids possible leakage problems
  - Fast detection particularly important with low bit rate voice codecs such as G.723.1 or G.729.
- Q: Does tone relay use UDP or TCP?

# Dealing With Packet Loss

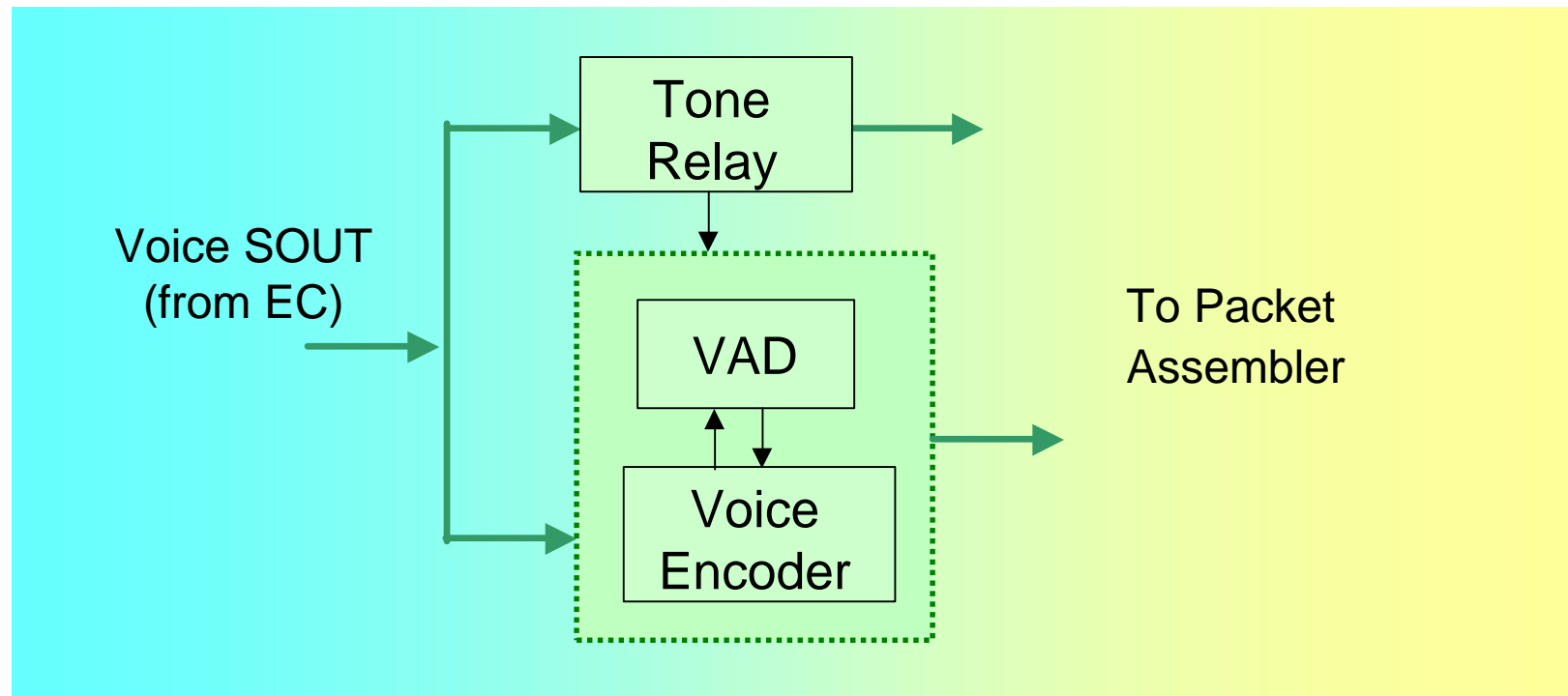
- Network Level (transparent to DSP)
    - QoS protocols
    - Call Admission Control
  - Other Non-transparent Means
    - Adaptive Jitter Buffer
    - Interleaving
    - Transmit Redundant Packets
    - Silence Suppression
- Quality gained at the cost of extra latency
- ← Indirect approach – reduce network congestion

# Packet Voice: Key Blocks



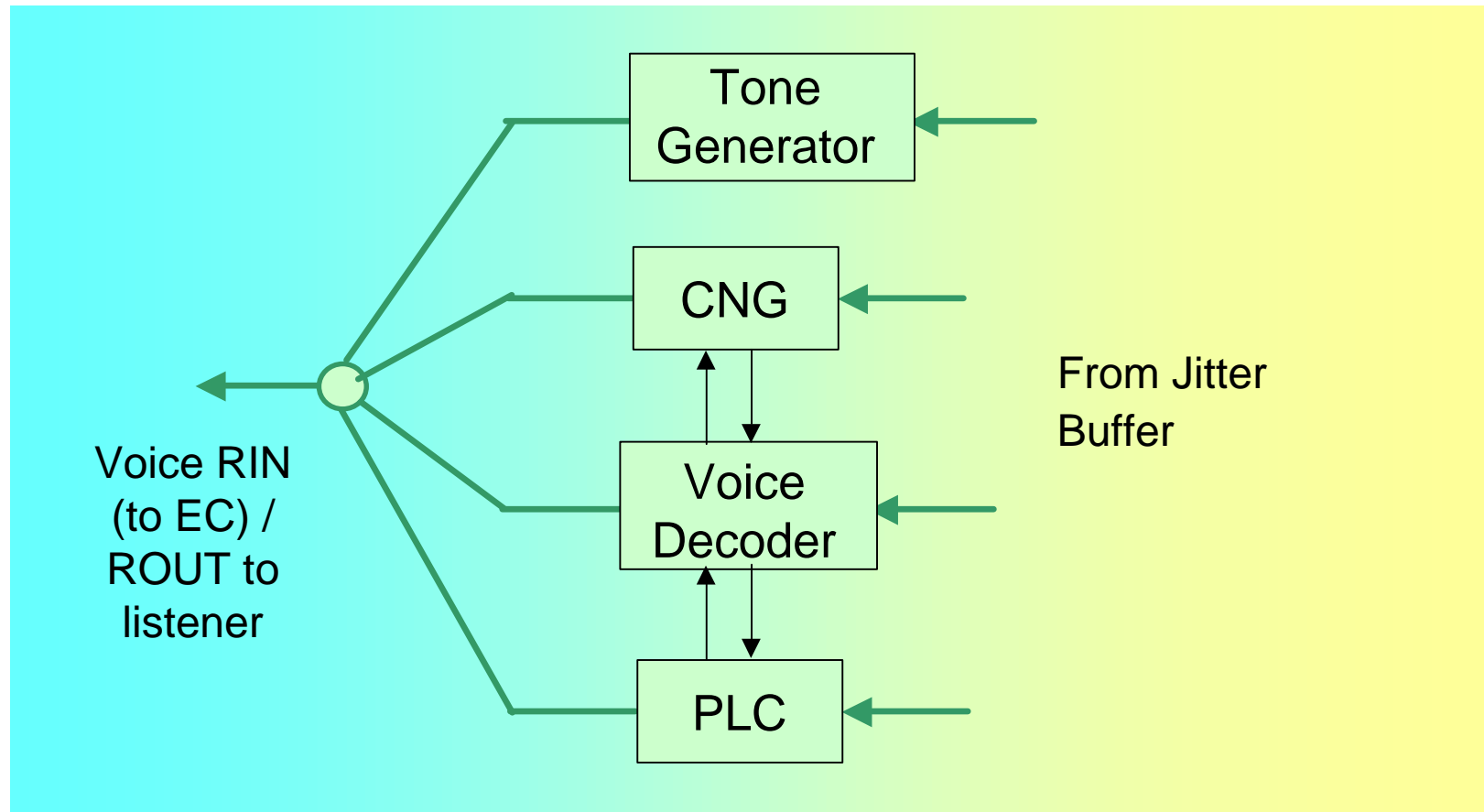
- EC required only if echo can occur in ROUT / SIN loop.
- EC (if deployed) should be properly aligned with any I/O buffering delays in the loop.

# Packet Voice: SND Path Detail

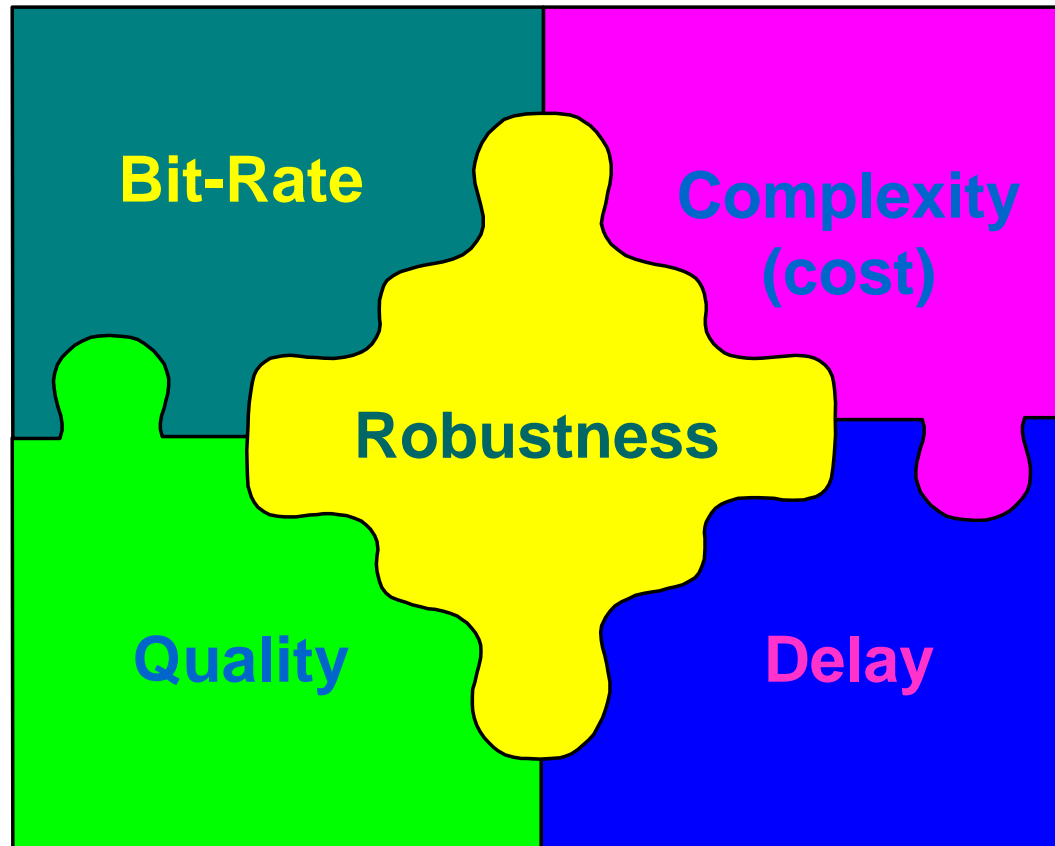


- If an EC is present, how should its NLP + CNF module be used?  
Preferable to integrate this function with the VAD.

# Packet Voice: RCV Path Detail



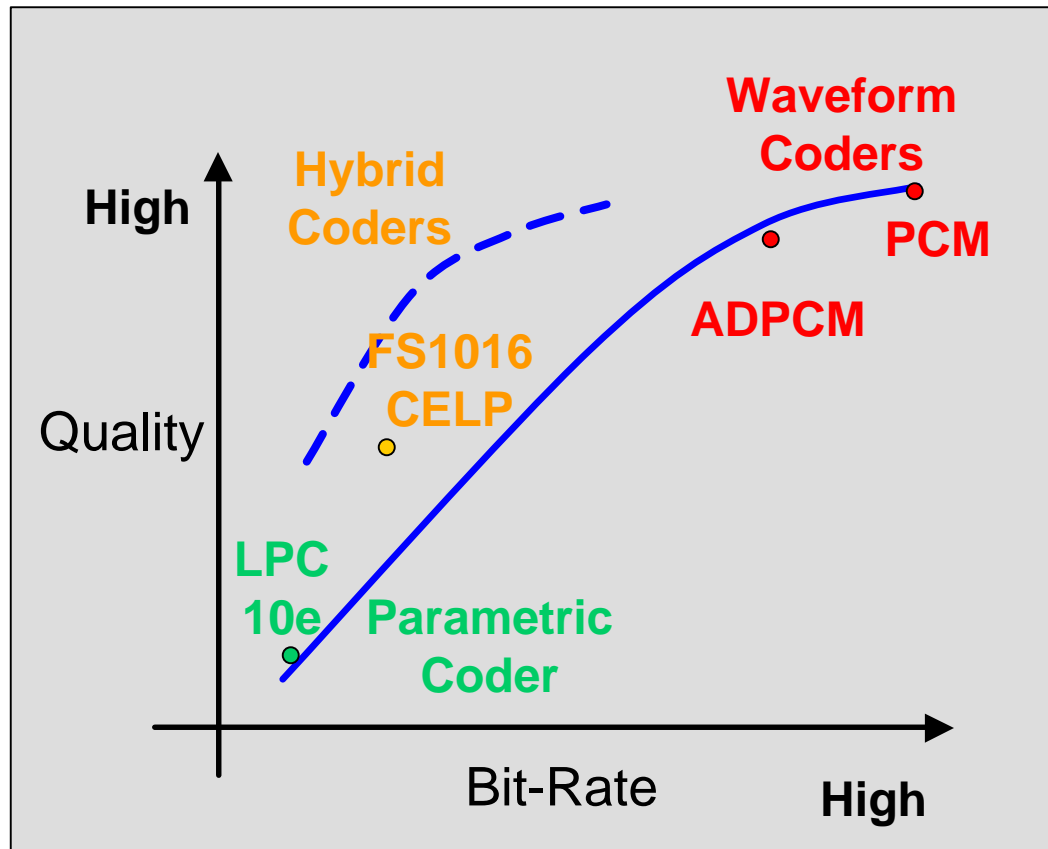
# Speech Coding Criteria



Depending on the end-use, each criterion requires a different weightage.



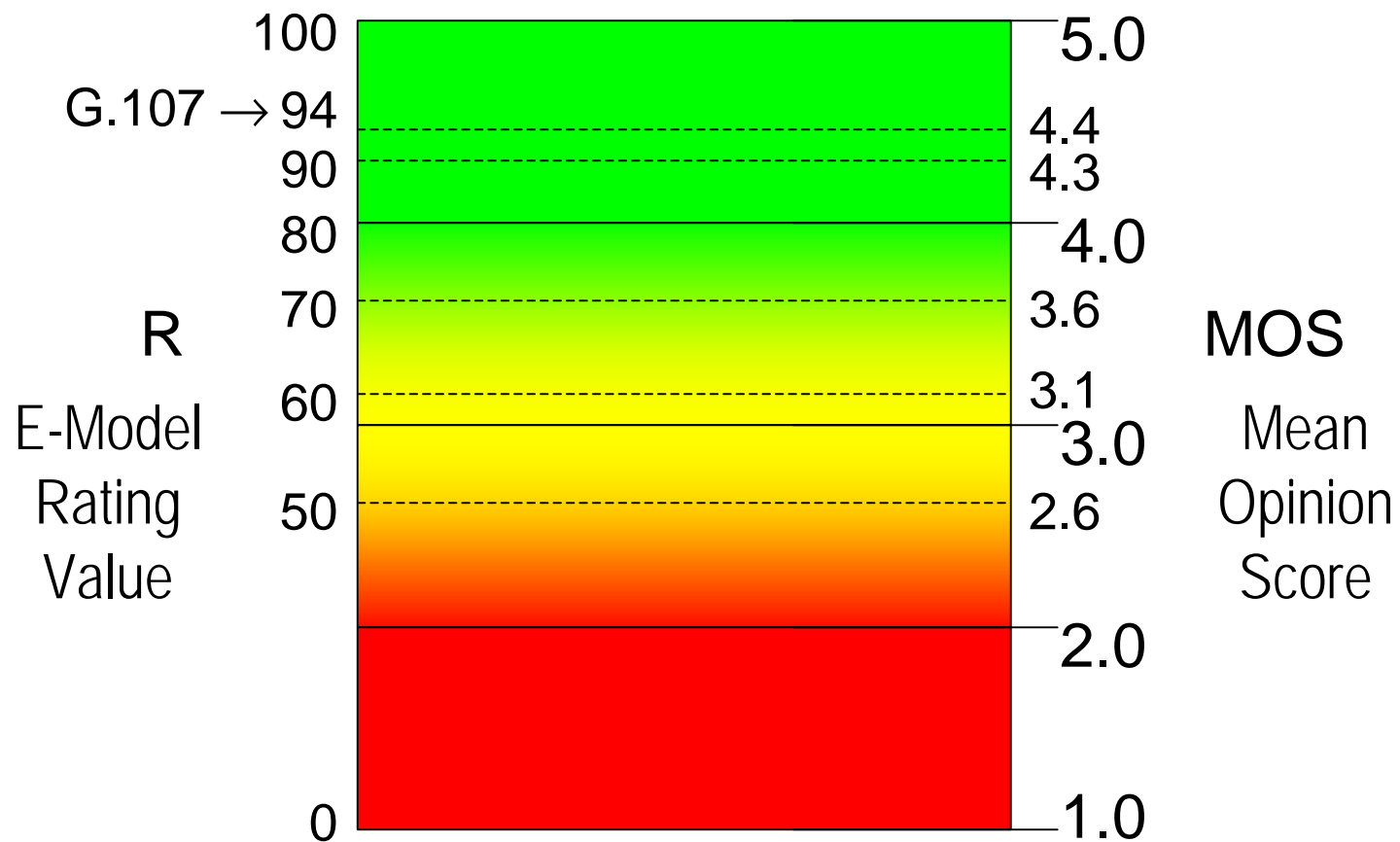
# Speech Coding Approaches



Hybrid Codecs employ tools such as vector quantization to gain quality at low bit rates.

Hybrid codecs are more costly to implement.

# Voice Quality



# Some Voice Codec Options

Codec	Bit Rate (kbits/s)	Delay	Notes
G.711 (PCM)	64	0.125ms	Toll quality
G.726 (ADPCM)	32	0.25ms	Toll Quality
G.728	16	0.625ms	Licensing required
G.723.1	5.3 / 6.3	30ms (7.5ms)	Licensing required
G.729 / G.729A	8	10ms (5ms)	Licensing required
GSM (RPE-LTP)	13	20ms	Patents? Low Quality
iLBC	13.4 / 15.2	30ms / 20ms	Royalty-free / IETF
Speex	2.2 - 44	30ms	Open source

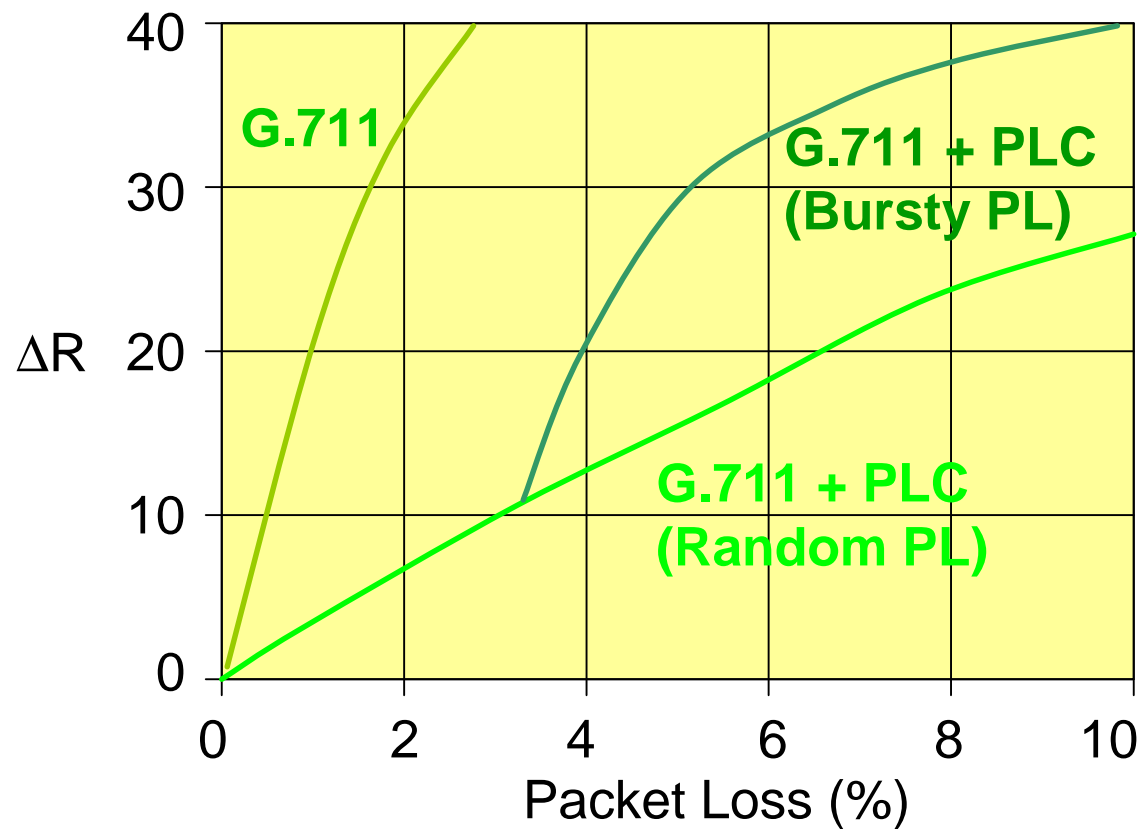
NOTE: Delay column shows frame size  $T_F$  and look-ahead buffer duration  $T_{LA}$ , if any.  
Total codec processing delay is  $2T_F + T_{LA}$

# Voice Codec Costs

Codec	Bit Rate (kbits/s)	DSP Processing Budget
G.711 (PCM)	64	< 1 MHz (with PLC)
G.726 (ADPCM)	32	4 – 9 MHz (with PLC)
G.728	16	20 – 30 MHz
G.723.1	5.3 / 6.3	12 – 20 MHz
G.729 / G.729A	8	12 – 20 / 6 – 10 MHz
GSM (RPE-LTP)	13	3 – 8 MHz
iLBC	13.4 / 15.2	8 – 12 MHz
Speex	2.2 - 44	?

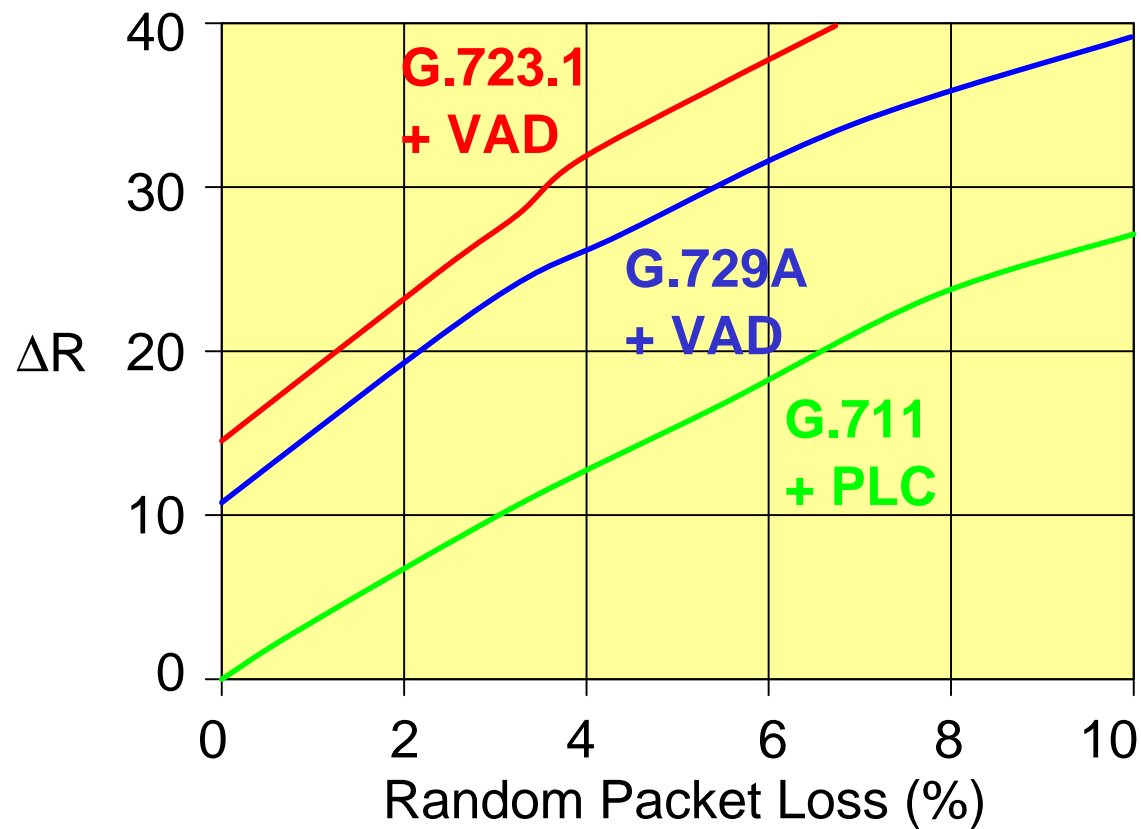
NOTE: Program / data memory requirements of each codec should also be considered.

# R-degradation: Packet Loss



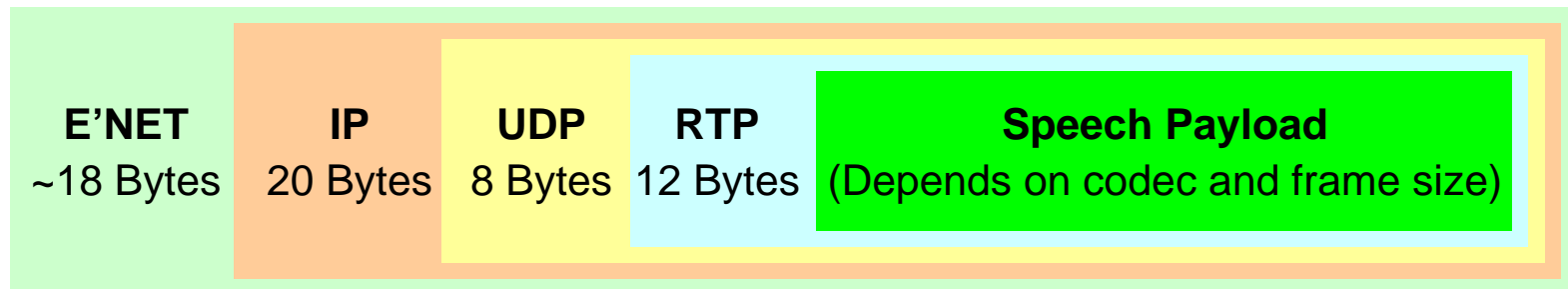
Source: Luis F Ortiz (Brooktrout Technology), RTC Magazine, July 2001

# R-degradation: Packet Loss



Source: Luis F Ortiz (Brooktrout Technology), RTC Magazine, July 2001

# Packet Header Overheads



- Packetization overheads can be significant.
- Header compression (cRTP) is possible for IP/UDP/RTP. Use of cRTP needs low round-trip delays (for header repair requests). This is primarily useful for (low-speed) local links.
- Large packets amortize overheads at the cost of extra latency.
- Other overheads (such as RTCP) are not counted here.

# Bit Rates With Overheads

Voice Payload (Bytes) :

Codec:	Frame Size		
	10ms	20ms	30ms
G.711	80	160	240
G.726	40	80	120
G.729A	10	20	30

Channel Bit Rate (kbits/s) :

Codec:	Frame Size		
	10ms	20ms	30ms
G.711	110.4	87.2	79.5
G.726	78.4	55.2	47.5
G.729A	54.4	31.2	23.5

Unless packet header overheads are reduced, benefits of low bit-rate codecs are not fully utilized.



# Signal Processing Options

- Low-channel CPEs
  - One General purpose  $\mu$ P alone
  - One DSP alone
  - Single-chip IP Phone SoC ( $\mu$ P + DSP + I/O)
  - One uP (host) + one or more DSP(s)
- Large CPEs, Gateways
  - Multiple hosts + DSP Farms
  - Hosts + DSPs with HW Accelerators
  - Few hosts + SoC (Processors + HW + I/O)

# Signal Processing Costs

- DSP MHz numbers below typically scale up by 1.5X to 3.0X on general purpose processors.
- Only functions that contribute to peak processor load are listed.
- Memory usage (not presented) is often a critical factor.

Functions	DSP MHz / voice channel
<b>Codec (VAD-CNG, PLC)</b>	<b>3 – 30 (depends on codec choice)</b>
<b>G.168 EC</b>	<b>3 – 10 (depends on EC design)</b>
<b>DTMF Tx + Rx</b>	<b>1 – 4</b>
<b>Caller ID Tx</b>	<b>1 – 2</b>
<b>Other Functions</b>	<b>4 – 14 (Jitter buffer, packet processing, I/O handling, task scheduling)</b>
<b>Total:</b>	<b>12 – 60 MHz</b>

# Trends & Issues

- Wide band (7 kHz) voice codecs
- Stereo audio (conferencing) (?!?)
- Improved multi-party conferencing support
  - Conference bridges with multi-casting?

- Improved QoS
- Improved security
- Lower power consumption

In particular, the emerging VoWLAN (or VoWiFi) market needs this support.

# Packet Voice References

- Books
  - F. Ohrtman, "Voice over 802.11", Artech (2004)
  - A. Sulkin, "PBX Systems for IP Telephony", McGraw-Hill (2002)
  - D. J. Wright, "Voice over Packet Networks", Wiley (2001)
- IETF
- Cable Labs (PacketCable)
- ATM Forum
- Frame Relay + MPLS Forum
- Historical papers (packet voice problem not new):
  - W. A. Montgomery: "Techniques for packet voice synchronization" IEEE JSACS, SAC-1, 6, Dec 1983
  - J. Gruber, L. Strawczynski, "Subjective effects of variable delay and speech loss in dynamically managed systems", IEEE Globecom (1982).