VoIP QoS Factors

David Tipper
Associate Professor
Department of Information Science
Telecommunications
University of Pittsburgh
Slides 2

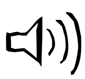
VoIP QoS

- Internet Telephone Quality of Service factors
  – Voice coder/decoder (codec) choice
  – Echo
  – Delay
  – Delay variation (delay jitter)
  – Packet loss
  – Reliability/Availability

Digital Speech Coding

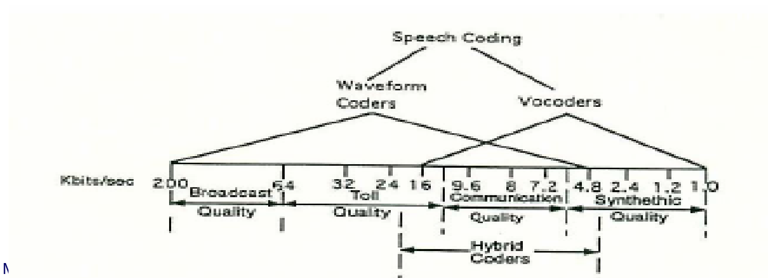
- Digital Speech
  - Convert analog speech to digital form and transmit digitally
- Applications
  - Telephony: (cellular, wired and Internet)
  - Speech Storage (Automated call-centers)
  - High-Fidelity recordings/voice
  - Text-to-speech (machine generated speech)

- Issues
  - Efficient use of bandwidth
    - Compress to lower bit rate per user => more users
  - Speech Quality
    - Want tollgrade or better quality in a specific transmission environment
    - Environment (BER, packet lost, out of order, delay, etc.)
  - Hardware complexity
    - Speed (coding/decoding delay), computation requirement and power consumption



Digital Speech

- Speech Coder: device that converts speech to digital
- Types of speech coders
  - Waveform coders
    - Convert any analog signal to digital form
  - Vocoders (Parametric coders)
    - Try to exploit special properties of speech signal to reduce bit rate
    - Build model of speech – transmit parameters of model
  - Hybrid Coders
    - Combine features of waveform and vocoders



Speech Quality of Various Coders

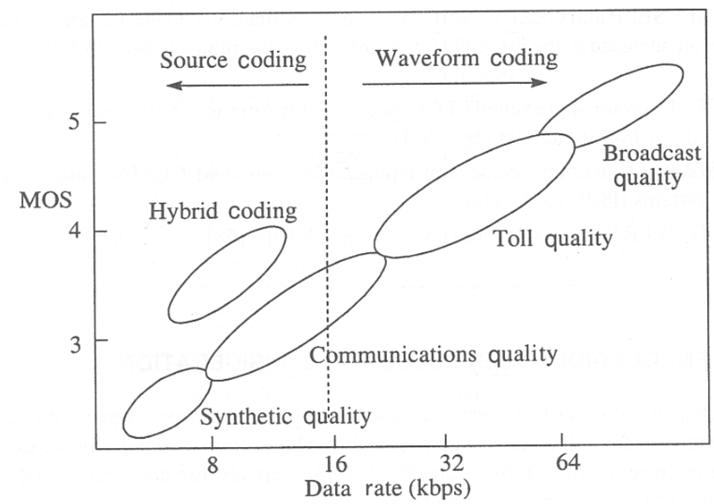
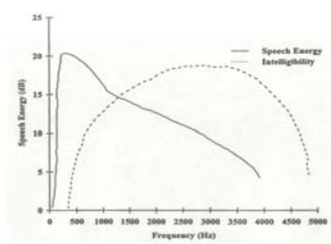
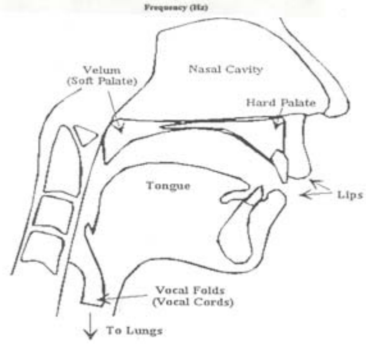


Figure 3.44 General speech quality versus transmission rate.

Characteristics of Speech

- Bandwidth
  - Most of energy between 20 Hz to about 7KHz ,
  - Human ear sensitive to energy between 50 Hz and 4KHz
- Time Signal
  - High correlation
  - Short term stationary
- Classified into four categories
  - Voiced : created by air passed through vocal cords (e.g., ah, v)
  - Unvoiced : created by air through mouth and lips (e.g., s, f)
  - Mixed or transitional
  - Silence

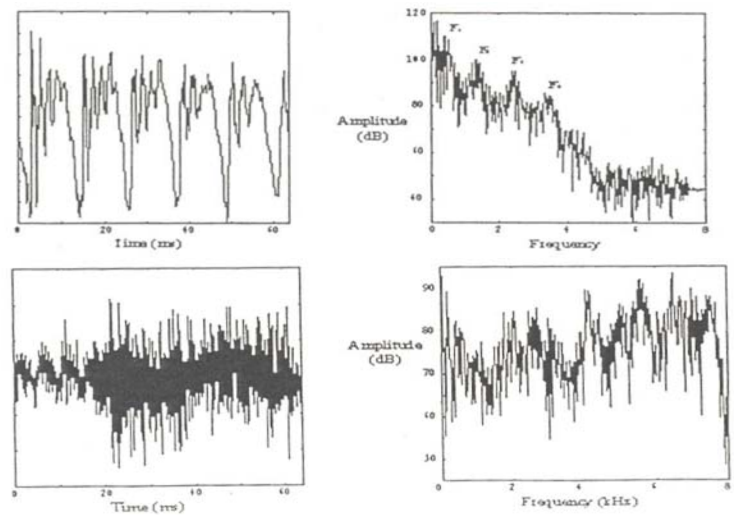




Characteristics of Speech

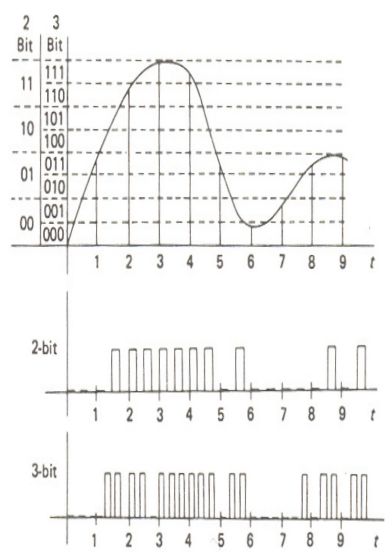
Typical
Voiced
speech

Typical
Unvoiced
speech

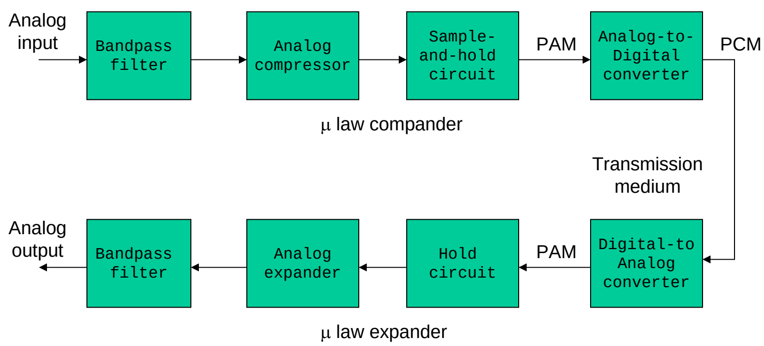


Waveform Coders (e.g.,PCM)

- **Waveform Coders**
  - Convert any analog signal to digital - basically A/D converter
  - Analog signal sampled > twice highest frequency- then quantized into ` *n* ‘ bit samples
  - Uniform quantization
  - Example Pulse Code Modulation
  - band limit speech < 4000 Hz
  - pass speech through μ–law compander
  - sample 8000 Hz, 8 bit samples
  - 64 Kbps DS0 rate
- **Characteristics**
  - Quality – High
  - Complexity – Low
  - Bit rate – High
  - Delay - Low
  - Robustness - High



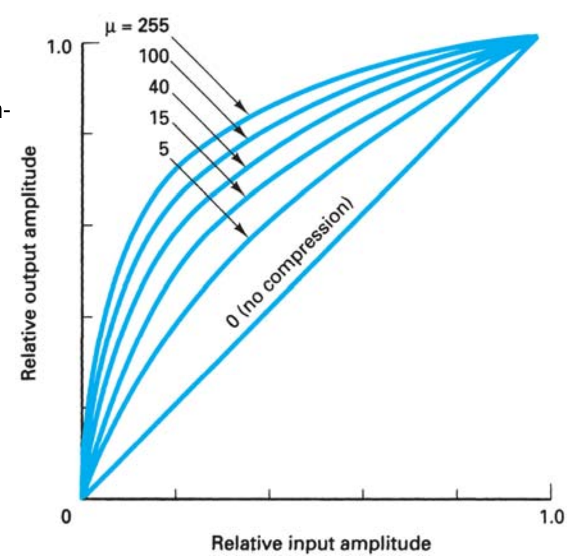
PCM Speech Coding



Pulse code modulation (PCM) system with analog companding – ITU G.700 standard basis for speech coding In PSTN in 60's

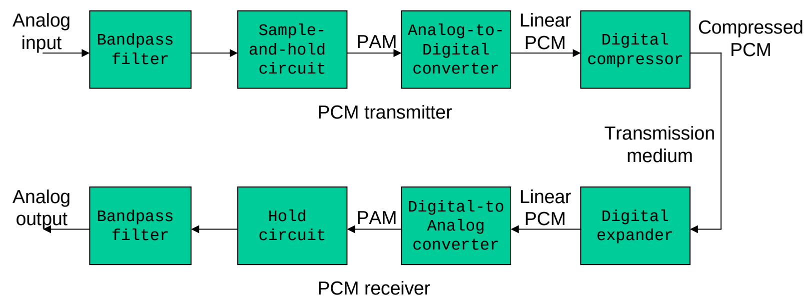
μ-law Compression/Companding

Companding
Nonlinear amplification in-order to equalize SNR across samples.



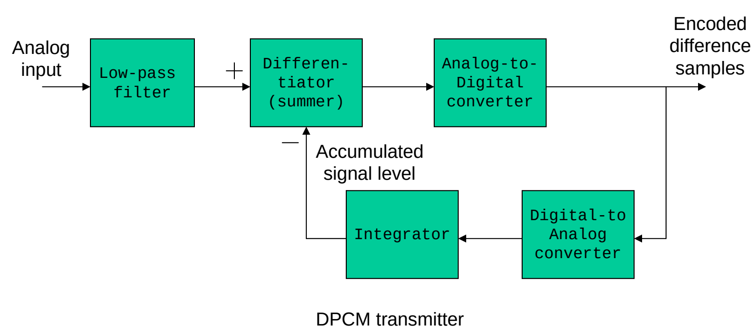
PCM Speech Coding

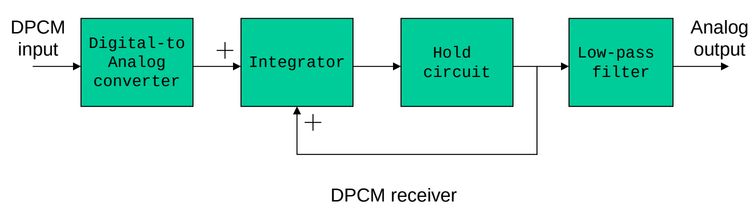
- Digitally companded PCM system – ITU G.711 standard
  - better quality speech than analog companding
  - baseline coder for PSTN 64 Kbps



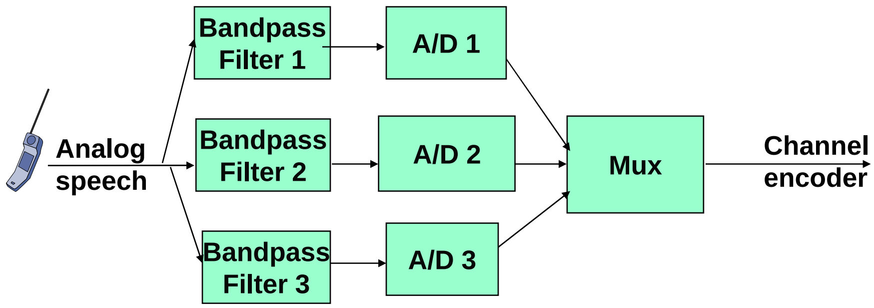
- Differential PCM (DPCM) : reduce bit rate from 64 Kbps to 32 Kbps)
  - since change is small between sample – transmit 1 sample
  - then on transmit difference between samples – use 4 bits to quantize
  - adaptively adjust range of quantizer – improves quality (ADPCM ITU G.726)

DPCM Speech Coding





Subband Speech Coding



Partition signal into non-overlapping frequency bands use different A/D quantizer for each band
Example: 3 subbands
5600+12000 + 13600 = 31.2 Kbps

band	Range	encoding
1	50- 700 Hz	4 bits
2	700-2000 Hz	3 bits
3	2000-3400Hz	2 bits

Vocoders

- Vocoders (Parametric Coders)
  - Models the vocalization of speech
  - Speech sampled and broken into frames (~25 msec)
  - Instead of transmitting digitized speech
    1. Build model of speech
    2. Transmit parameters of model
    3. Synthesize approximation of speech

- Linear Predictive Coders (LPC) basic Vocoder model
  - Models vocal tract as a filter
  - Filter excitation
    - periodic pulse (voiced speech) or noise (unvoiced speech)
  - Transmitted parameters:
    - gain, voiced/unvoiced decision, pitch (if voiced), LPC parameters

Vocoders

- Linear Predictive Coders (LPC)
  - Excitation
    - periodic pulse (voiced speech) or noise (unvoiced speech)
  - Transmitted parameters: gain, voiced/unvoiced decision, pitch (if voiced), LPC parameters

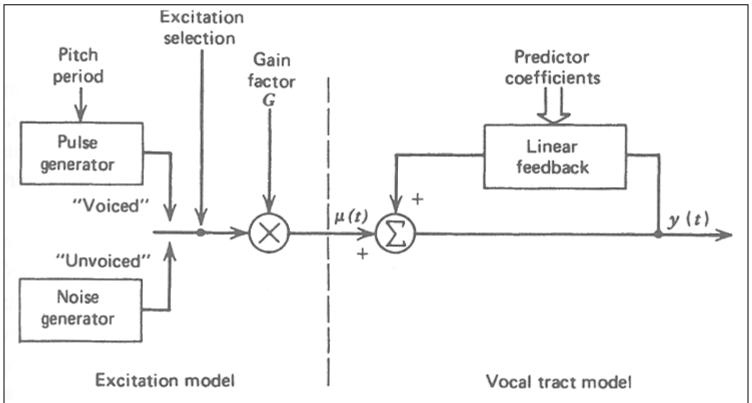


Figure 3.41 Speech generation model of linear predictive coding.

Vocoders

- Example Tenth Order Linear Predictive Coder
  - Samples Voice at 8000 Hz – buffer 240 samples => 30 msec
  - Filter Model
    - (M=10 is order, G is gain, z^{-1} unit delay, b_k are filter coefficients)

$$H(z) = \frac{G}{1 + \sum_{k=1}^{M} b_k z^{-k}}$$

- G = 5 bits, b_k = 8 bits each, voiced/unvoiced decision = 1 bit, pitch = 6 bits => 92 bits/30 msec = 3067 bps

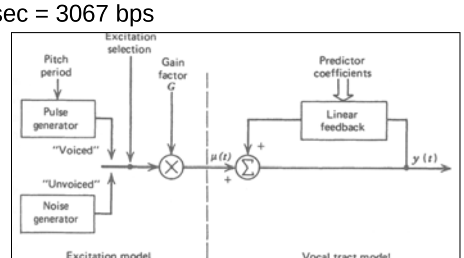


Figure 3.41 Speech generation model of linear predictive coding.

Vocoders

- LPC coders can achieve low bit rates 1.2 – 4.8 Kbps
- Characteristics of LPC
  - Quality – Low
  - Complexity – Moderate
  - Bit Rate – Low
  - Delay – Moderate
  - Robustness – Low

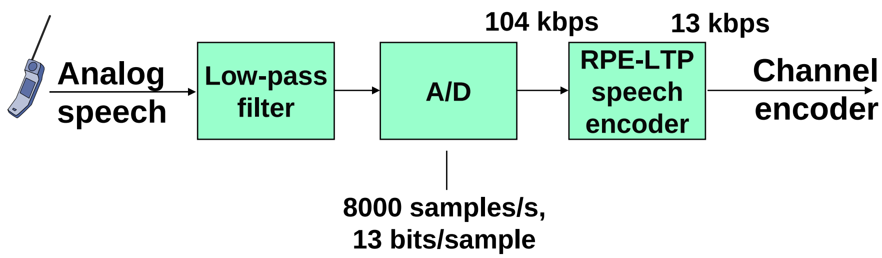
- Quality of pure LPC vocoder too low for internet telephony - try to improve quality by using hybrid coders

- Hybrid Coders
  - Combine Vocoder and Waveform Coder concept
    - Residual LPC (RELP)
    - Codebook excited LPC (CELP)

Example Hybrid Vocoder: GSM

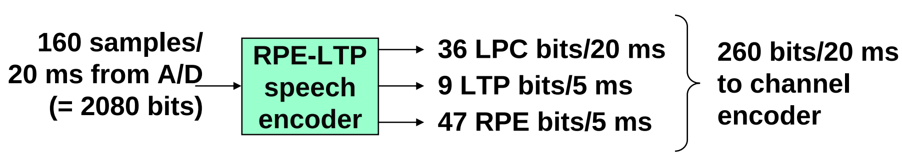
- PCM: 64kbps too wasteful for cell phones
- GSM uses Regular Pulse Excited -- Linear Predictive Coder (RPE--LPC) for speech
  - Basically combine DPCM concept with LPC
  - Information from previous samples used to predict the current sample.
  - The LPC coefficients, plus an encoded form of the residual (predicted - actual sample = error), represent the signal.
  - 20 millisecond samples: each encoded as 260 bits =>13 kbps (Full-Rate coding).

GSM Speech Coding



GSM Speech Coding (cont)

Regular pulse excited - long term prediction (RPE-LRP) speech encoder (RELP)speech coder

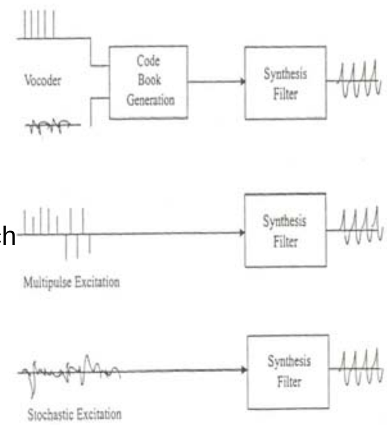


LPC: linear prediction coding filter
LTP: long term prediction – pitch + input
RPE: Residual Prediction Error:

Hybrid Vocoders

- **Codebook Excited LPC (CELP)**
  - Problem with simple LPC is U/V decision and pitch estimation doesn't model transitional speech well, and not always accurate

  - Codebook approach – pass speech through an analyzer to find closest match to a set of possible excitations (codebook)
  - Transmit codebook pointer + LPC parameters
  - *ITU G.729* standard 8kbps
    - Popular for VoIP
  - *ITU G.723.1* standard 6.3 Kbps and 5.2 Kbps versions
    - Popular for VoIP on PCs



CELP Encoder

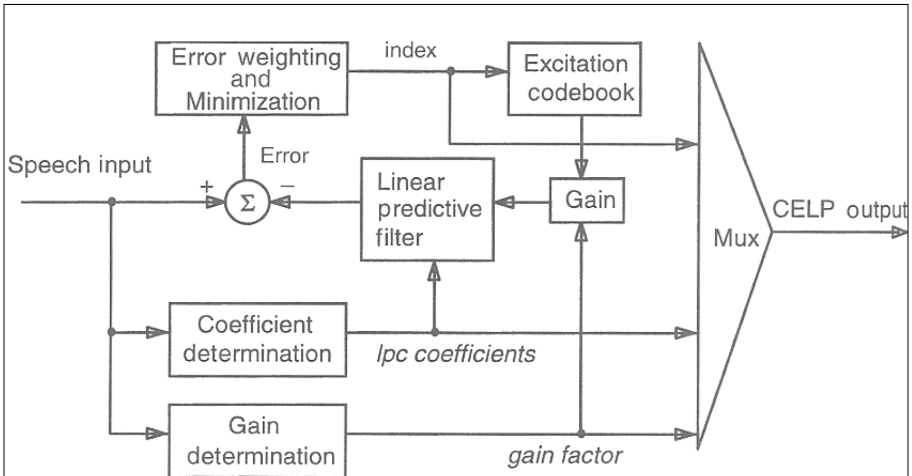


Figure 3.43 CELP encoder block diagram.

Evaluating Speech Coders

- **Qualitative Comparison**
  - based on subjective procedures in ITU-T Rec. P. 830
- **Major Procedures**
- ***Absolute Category Rating***
  - Subjects listen to samples and rank them on an absolute scale - result is a mean opinion score (MOS)
- ***Comparison Category Rating***
  - Subjects listen to coded samples and original un-coded sample (PCM or analog), the two are compared on a relative scale – result is a comparison mean opinion score (CMOS)

Mean Opinion Score (MOS)

Excellent	5
Good	4
Fair	3
Poor	2
Bad	1

Comparison MOS (CMOS)

Much Better	3
Better	2
Slightly Better	1
About the Same	0
Slightly Worse	-1
Worse	-2
Much Worse	-3

Evaluating Speech Coders

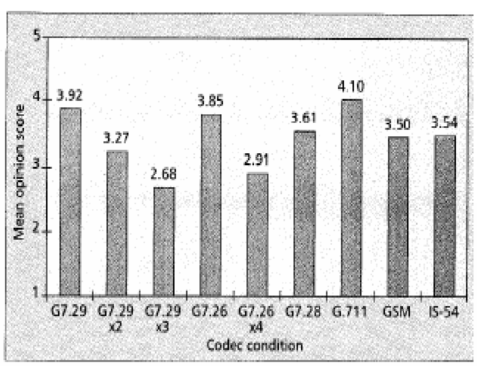
MOS for clear channel environment – no errors
Result vary a little with language and speaker gender

Standard	Speech coder	Bit rate	MOS
PCM	Waveform	64 Kbps	4.3
DECT	ADPCM	32 Kbps	4.1
GSM	Hybrid RELPC	13.3 kbps	3.54
QCELP	Hybrid CELPC	14.4 Kbps	3.4
QCELP	Hybrid CELPC	9.6 Kbps	3.4
LPC	Vocoder	2.4 Kbps	2.5
ITU G.729	Hybrid CELP	8.Kbps	3.9
ITU G.723.1	Adaptive CELP	6.3Kbps	3.9
ITU G.723.1	Adaptive CELP	5.2Kbps	3.4

Evaluating Speech Coders

- Types of environments recommended for testing coder quality
  - Clean Channel no background noise
  - Vehicle : emulate car background noise
  - Street : emulate pedestrian environment
  - Hoth : emulate background noise in office environment (voice band interference)
- Consider environments above for cases of
  - Perfect Channel – no transmission errors
  - Random channel errors
  - Bursty channel errors
- May consider repeated encoding/decoding (e.g.,VoIP to Mobile call)

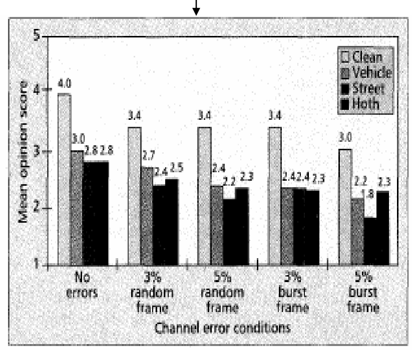
Evaluating Speech Coders



■ **Figure 1.** *Mean opinion scores for the basic coded conditions, including multiple incodings by a single codec.*

Repeated coding degrades quality

Background noise and errors degrade quality



■ **Figure 7.** *Mean opinion scores for G.729 conditions with background noise and random and burst frame erasure.*

Speech Quality

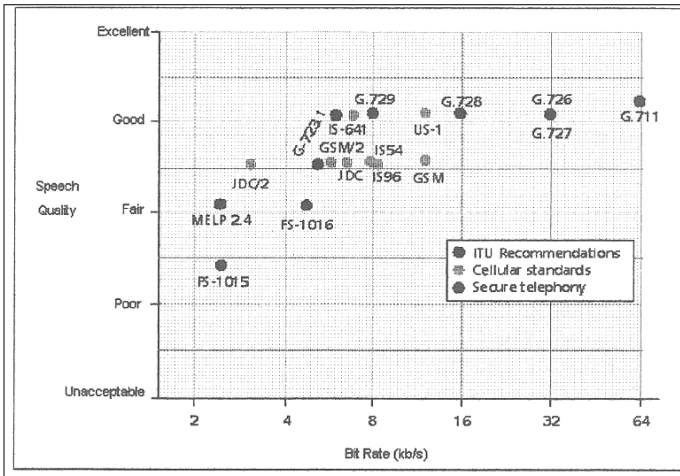


Figure 3.45 Speech quality of standard encoding algorithms. (From R. V. Cox, "Three New Speech Coders from the ITU Cover a Range of Applications," *IEEE Communications Magazine*, September 1997.)

Codec Selection

- For VoIP need to consider Quality, Complexity, Delay, Compression Rate

Coder	Bit Rate	Coding Delay	Decoding Delay	Complexity
G.711	64 Kbps	0	0	Low
G.729	8 Kbps	15 ms	7.5 ms	Medium
G.723.a,b	6.4/5.3 Kbps	35.5 ms	18.75 ms	High

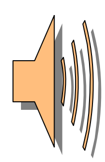
Silence Compression

- Much of a conversation is Silence (~40%)
  no need to transmit
- Voice Activity Detector (VAD)
  - Hardware to detect silence period quickly

- Discontinuous transmission (DTX)
  - Stop transmitting frames
    - Send minimal # of frames to keep connection up
- Comfort Noise Generator (CNG)
  - Synthesize background noise avoids: “Did you hang up?”
    - Random noise or reproduce speaker’s ambient background
- Popular VoIP G.723.1 codec has VAD/DTX/CNG

Echo

- Talker echo
  - Talker hears his own voice, delayed
  - Analogous to near-end cross-talk



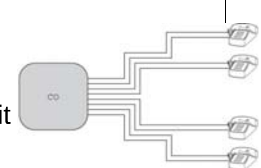
- Listener echo
  - Listener hears talker, and attenuated echo
  - Analogous to video “ghost”
- Two causes: Acoustic & Electrical
- Problem in all forms of Telephony worsened by delay of VoIP systems

Acoustic Echo

- Speaker-to-mike audio feedback
  – Hands-free "speaker-phones" (not head sets)
  – Analogous to microphone "squeal"
  – A's echo 10-15 dB below B's voice
  – Annoying to A ("barrel") if delay > 40 ms
- Easy Fix - use a head set
- Echo cancellers slightly effective
  – They work best for small, constant delays

Echo Compensation

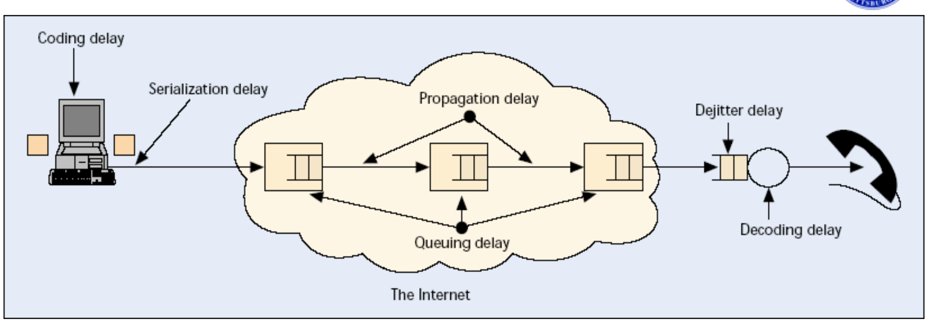
- Electronic echo occurs at any hybrid 4-wire to 2-wire analog interface
- Echo suppressor – old analog technology
  – Incoming speech-activated match impedance loss
- Echo canceller
  – Build estimate of echo digitally and remove it
    - Cancels talker echo
  – Operation
    - Record last T seconds of speech
    - Find pattern match in ear signal, subtract it
  – Long delays → expensive EEC
    - Put EEC close to the hybrid – VoIP Gateway or in terminal
    - Software versions of EC in many PC VoIP implementations
  – Not necessary in CATV or 4-wire systems



Delay

- Voice is Real Time Communication
  - large delay in transport is a problem
  - Unnatural conversation pattern
    - Example call over a GEO satellite incurs .25 sec delay leads to noticeable disruption in normal conversation pattern – users start to step on each others speech
- Some Typical Delay Values
  - Satellite Hop: 250 ms
  - Coast to Coast in North America: 24-60 ms
  - US to Japan (fiber) propagation delay: 80-100 ms
- End-to-End Delay Guidelines
  - Delays less than 150 ms can largely be ignored
  - Delays 150 - 400 ms acceptable over large distances NYC - Bangkok
  - Delays > 400ms unacceptable as users perception of voice quality very poor
- VoIP end-to-end delay can easily exceed 400 ms!

IP Telephony Delays



•Consider VoIP only network (no gateways or PSTN)
•Major Delays in IP Telephony Systems
- •Coding
- •Packetization/Serialization
- •Queueing at Routers
- •Propagation
- •Dejitter
- •Decoding

IP Telephony Delays

- Coding Delay
  - Time to gather speech sample compute vocoder model values for transmission
  - Value depends on vocoder utilized (0-50ms)
- Packetization and Serialization
  - Packetization: Time to gather data from coder for packet payload, attach headers
  - Remember the protocol stack for VoIP

    Output of Vocoder packed in Real Time Protocol (RTP) packets

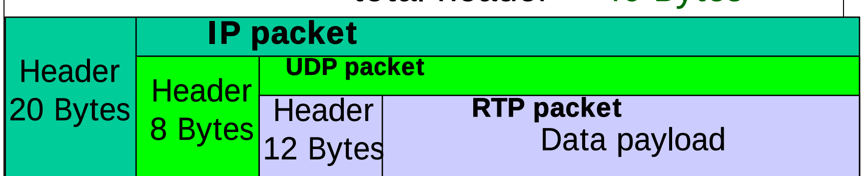
    Which are payload for User Datagram Protocol (UDP) packets

    Which are payload for Internet Protocol packets (IP)

Packetization Delay

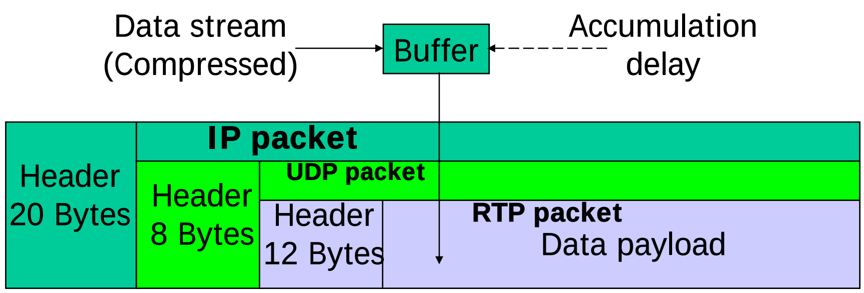
- VoIP packet (RTP/UDP/IP)

total header = 40 Bytes



- Assume
  - Delay: N voice samples → T ms -> payload P
  - Payload efficiency: P/(P+Header) %
  - Net data rate: (P+Header)/T = R Kbps

Packetization and Delay



- For example: 10Byte payload from 4-to-1 compression rate vocoder
  - Delay: 10Byte → 40 samples → 40×125μs = 5ms
  - Packet efficiency: 10/(40+10) = 20%
  - Net data rate: 50B/5ms = 80 Kbps (>64 kbps)

ITU Standard Codecs

- Built-in
  - Compression algorithm
  - Framing for the IP packet

Codec	BW	CR	Data Rate	Pkt Dly	IP Rate	MOS
G.711	4 kHz	1:1	64 kbps	1 ms	384 kbps	4.2
G.728	4 kHz	4:1	16 kbps	2.5ms	144 kbps	4.3
G.729	4 kHz	8:1	8 kbps	10 ms	40 kbps	4.0
G.723.a	4 kHz	10:1	6.4 kbps	30 ms	17 kbps	3.9
G.723.b	4 kHz	12:1	5.3 kbps	30 ms	11 kbps	3.7
G.722	7 kHz	2:1	56 kbps	10 ms*	88 kbps*	4.3+
CD	22 kHz	1:2	700 kbps	10 ms*	200 kbps*	5.0

Note payloads are usually small => high header overhead

Serialization and Transmission

- Serialization Delay: time to transmit on access line both from caller to network also have this at the other end of network to called party
  - 1 byte on 64kbps line => 125 μsec
  - G.723a codec over modem: 64byte packet /56kbps=11ms
  - 1byte on OC-3 fiber to home line (155Mbps) => 0.05 μsec
  - Insignificant on high-speed links
- Propagation Delay
  - Time to propagate packet down link - depends on distance of link and medium
    - Satellite Hop wireless link 250 ms
    - Coast to Coast in North America fiber optic propagation 24 ms
    - For example fiber optic cable propagates at roughly 2/3 speed of light (3×10^8) meter/sec - so 200km link has propagation delay of $200/(3 \times 10^8) = 0.66$ ms
  - Small enough on short fiber links to ignore

Network Delays

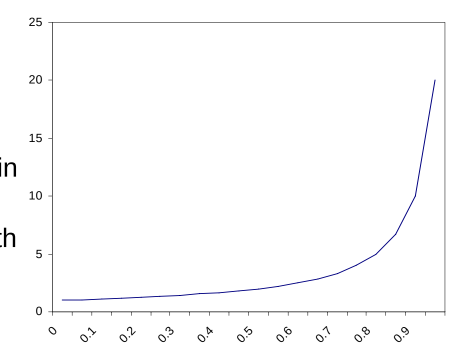
- Router delay
  - Time for router to process/transmit packet + delay in router queues
  - Time to process/transmit packet depends on router switch speed and link speed – for high bandwidth links and core network routers small amount of time 10 – 20 μsecs

Queueing Delay

– Time waiting in router buffers for processing and transmission

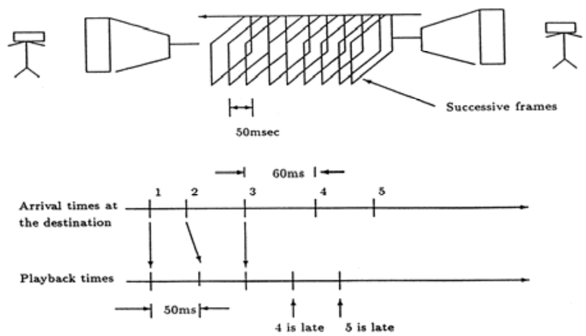
– Value highly dependent on load and QoS mechanisms deployed in router 10's msec to 10's secs

• Queueing Delay nonlinear with increases of network load



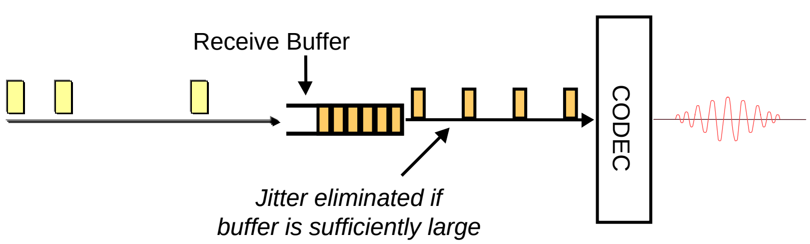
Network Delays

- Delay Jitter defined as the variation of the delay for two consecutive packets
  - Due to variation of
    - Routes of packets
    - Router delay (processing time + queueing time)



Network Delays

- Jitter buffer
  - Jitter buffer to smooth out playout of packets to destination
    - Allows packet delivery times to vary
    - Allows packets to arrive out of order
  - Note 30 ms holds one G.723 packet, typical values 30-100 msec



Measurement Delay Example

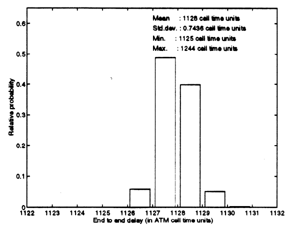


Figure 1: PSC-NOR Series 1: Histogram of end-to-end delay at 4125 Kbps

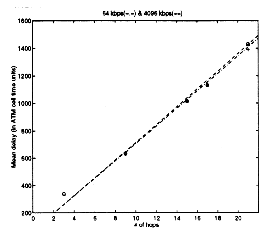


Figure 3: Mean delay versus the number of hops

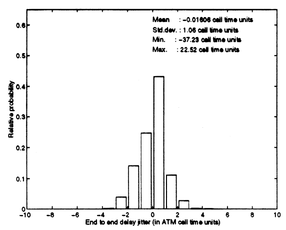


Figure 2: PSC-NOR Series 1: Histogram of end-to-end delay jitter at 4125 Kbps

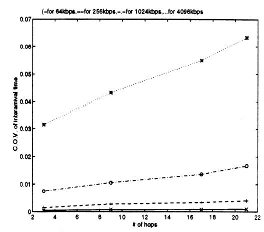
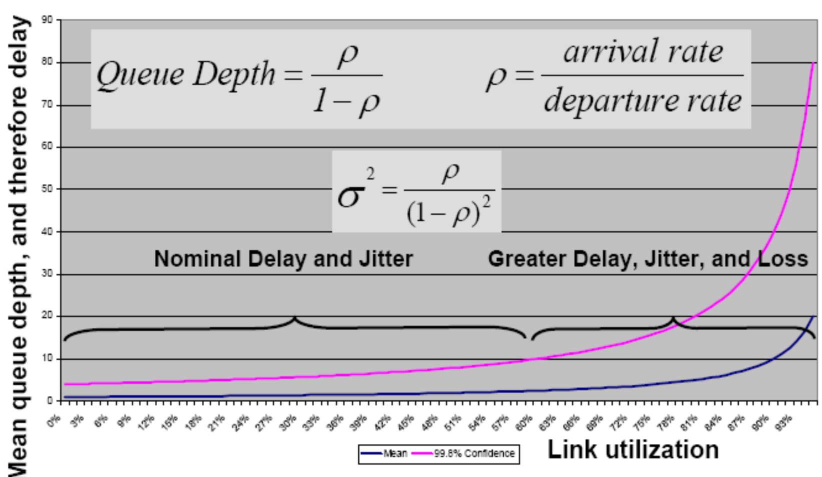


Figure 4: Coefficient of Variation for cell inter-arrival times, Series 1

Queueing Theory

Queueing Models Provide an Analytical Model of Router Delay and Jitter vs. Load- example M/M/1 model of router port



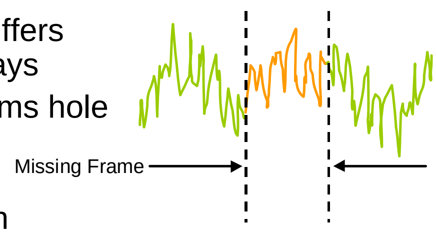
Note: curves like this (M/M/1) only possible when arrival process is known

Example of End-to-End Delay Budget

- Often design on basis of a Target Delay Budget
- Sender
  - Coding Delay 5
  - Packetization delay 30
  - Serialization delay 11
- Network
  - Routers 5 @ 7ms each 35 If no congestion.
  - Propagation 25
- Receiver
  - Jitter buffer 30
  - Receive, de-packet, decode <u>46</u>
- Total 182 ms
- Well below 400ms but above ideal of 150ms

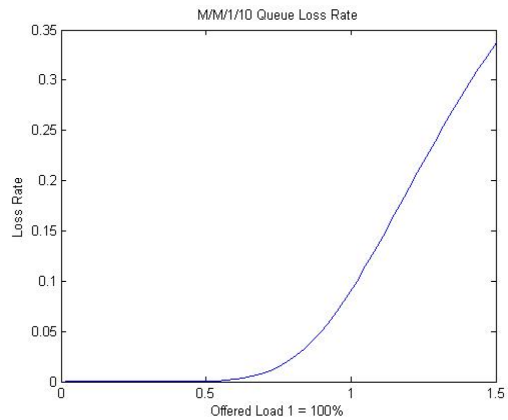
Packet Loss

- Packet Loss occurs when buffers overflow at routers or gateways
- A lost packet leaves a 40-80ms hole (depending on codec) in the conversation
- Retransmission not an option
- Some low level of packet loss can be made up by human brain from context
- MOS drops quickly with increasing packet loss rate
- For quality comparable to PSTN need very low loss rate < 0.5%



Packet Loss

- Packet Loss increase is highly nonlinear with load increase
- Can be studied with Queueing Theory for M/M/1/K queue



$$P_b = \frac{(1-\rho)\rho^K}{1-\rho^{K+1}} \qquad \rho \neq 1$$

$$P_b = \frac{1}{K+1} \qquad \rho = 1$$

½ % loss at 60% load

Reliability

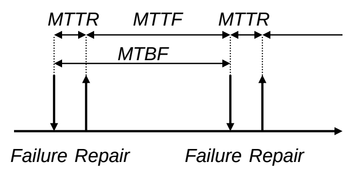
- Reliability
  - *Ability of an item to perform a required function for a stated period of time*
    - *The "item" could be a component, subsystem, or system.*
  - A common unit to measure reliability
    - Mean Time To Failure (MTTF)
  - Another related term: Mean Time Between Failure (MTBF)
    - Time between consecutive failures of the item
    - Usually, MTTR $\ll$ MBTF
    - To be exact, MBTF = MTTF + MTTR,
      - MTTR is Mean Time To Repair
    - Carrier Grade Telco equipment has large MBTF
      - For example, Bidirectional Optical Amplifier on a Fiber MBTF = 5*10^5 hours

Availability

- Availability (*A*)
  - *Ability of an item to perform stated function at over time or*
  - *Fraction of the time that an item can be used when needed*
  - Value in the 0.0 to 1.0 range
  - Mean Time To Repair (MTTR)
    - An average time to restore a full functionality to an item
      - This may include time to diagnose, isolate, remove and replace the failed part

$$A = 1 - \frac{MTTR}{MTBF}$$

Carrier class PSTN equipment requires availability in the range five nines (0.99999)



Availability Goals

- Availability Goals depend on application and user requirements
- PSTN designed for five 9's availability
- Most data networks have much lower availability values => VoIP will not be as reliable
- Also VoIP can have congestion outages where equipment is highly loaded resulting in such poor performance difficult to compete/maintain usable call

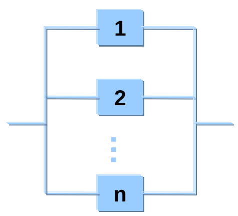
Availability level	Downtime per year
99.999%	5.25 min
99.97%	157.68 min
99.9%	8 hours 46 min
99%	87 hours 4 min

System Availability

- System availability calculated from component availability A_i *will always be less than component availability*
- If devices in series

$$A_{series} = \prod_{i=1}^{n} A_i$$

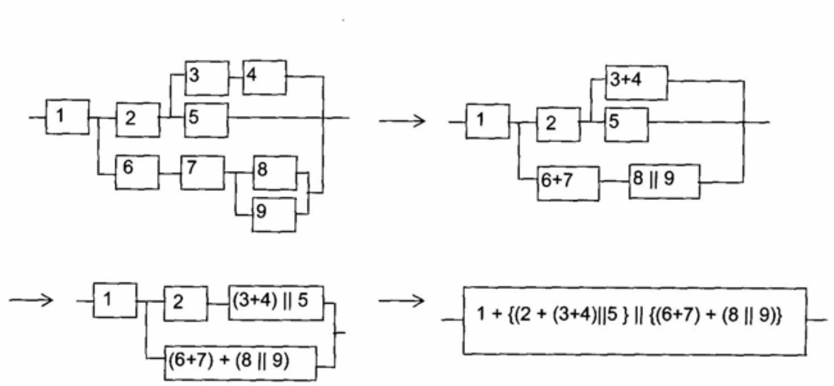
- If devices in parallel



$$U_{parallel} = \prod_{i=1}^{n} U_i$$

$$A_{parallel} = 1 - \prod_{i=1}^{n} (1 - A_i)$$

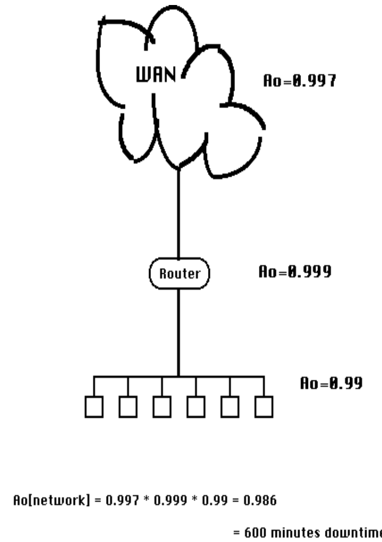
Series-Parallel Reduction



+ series
|| parallel

Availability

- High availability requires increased cost $$
- Need
  1. increased component availability
  2. system redundancy via parallel routes
- Avoid single point of failure
- Note routers have much more software than PSTN switches –> more prone to failure



VoIP QoS

- Internet Telephone Quality of Service factors
  – Voice coder/decoder (codec) choice
  – Echo
  – Delay
  – Delay variation (delay jitter)
  – Packet loss
  – Reliability/Availability
- VoIP is envisioned as one of many services on shared infrastructure - how can QoS be provided VoIP?