Freescale Technology Forum
Design Innovation.

5th Nov, 2008

The Emergence, Introduction and Challenges of Wideband Choice Codecs in the VoIP Market

PN101

Roger Chung

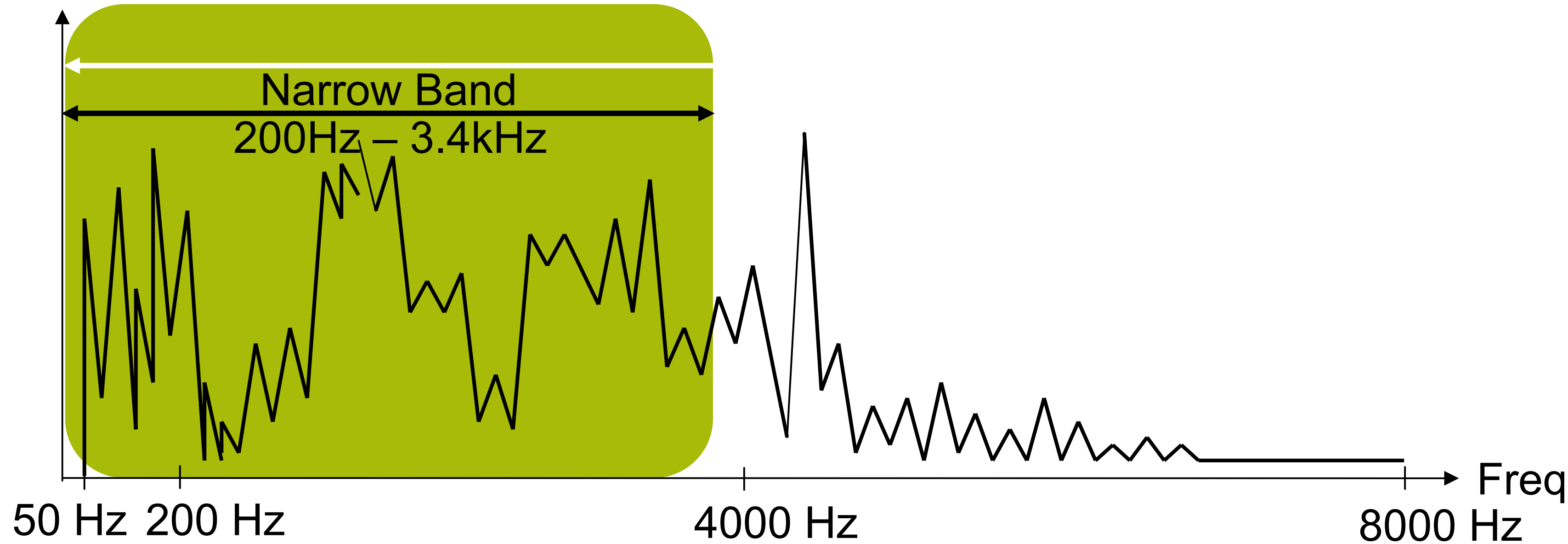
freescale™ semiconductor

Objectives

- To understand the benefits of wideband voice
- To understand the technical challenges of wideband voice
- To present the history of wideband and look at what the future holds

How narrow is narrow?

- Narrow band voice in a Telecommunications or Voice over IP context generally refers to an analog signal that is digitally sampled at 8kHz
- Narrow band refers to the frequency spectrum (or band) of the signal, which is a function of the sampling rate (Nyquist)
- Examples of narrow band codecs: G.711, G.726, G.723.1, G.728, G.729AB, iLBC, EVRC, AMR-NB, EFR, FR etc.



Bandwidth, quality, cost and bit rate

- G.711 is the de-facto standard of today's digital telephone network (T1 and E1 digital trunks all use G.711 as a standard for voice)
- G.711 takes a linearly quantized signal (13 or 14 bits @ 8k samples/s) and applies logarithmic quantization (not all bits get equal weighting) to reduce to 8 bits per sample @ 8k samples/s thus 64 Kbps.
- Waveform Codecs assume no a-priori knowledge of how the signal was generated thus are signal independent
- G.711 & G.726 are examples of waveform codecs
- Source or Hybrid coding uses a psychoacoustic model of the vocal tract to model speech and achieve lower bit rates but generally at lower voice quality and higher complexity compared to G.711.
- G.729, G.723.1, G.728 are examples of CELP based source/hybrid codecs

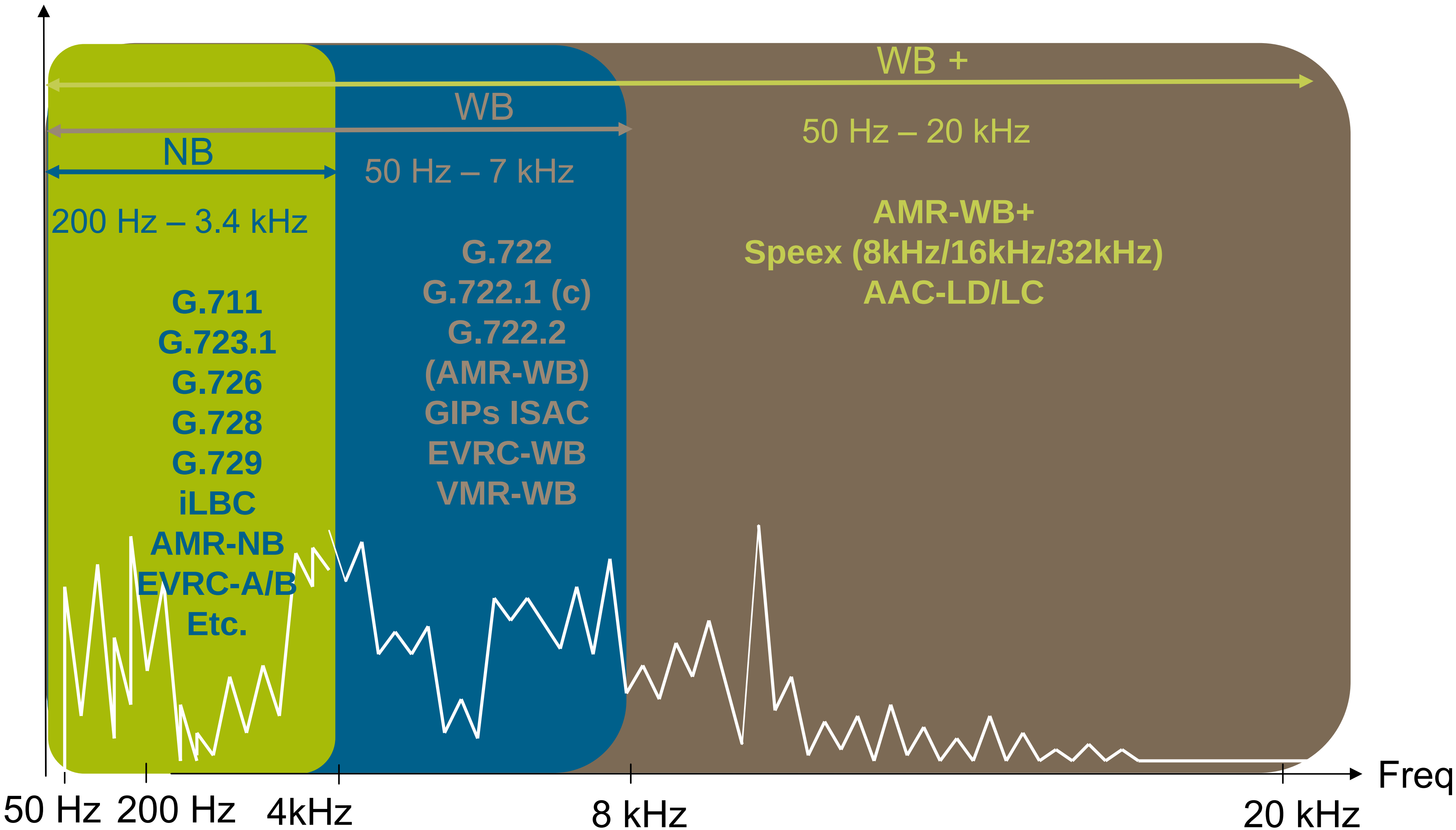
Bandwidth, quality, cost and bit rate

Algorithm	kbit/s	StarCore® DSP Performance (approx mcps)	Typ PESQ (no packet loss)
G.711(with VAD/CNG & PLC)	64	1.9	4.1 – 4.3
G.726 (24kbps)	16 - 40	4.5	3.6
G.728	16	11-12	3.8
G.729 A/B	8	4	3.6
G.723.1	5.3 / 6.3	6.5	3.5
GSM-AMR	4.75 - 12.2	8.5	3.9
iLBC	13.3 - 15.2	10	3.8

Who needs wide?

- Wide Band is sampled at 16 kHz (approx 7 kHz usable bandwidth)
- Wideband speech coding delivers major improvements in speech quality
  - **Extended LF contributes to increased naturalness, presence and comfort**
  - **Extended HF provides better fricative differentiation (differentiation between certain unvoiced or plosive utterances, such as “s” and “f” or “p” and “t”)**
  - **Improves the intelligibility and naturalness of speech**
  - **Adds a sense of transparent communication, eases speaker recognition & reduces listener fatigue**
- Siemens – “…wideband transmissions can reduce speech ambiguities by as much as 90 percent, increasing conversational intelligibility and reducing listener fatigue.” (2003 press release)
- Polycom – “For single syllables, 3.3 kHz bandwidth yields an accuracy of only 75 percent, as opposed to over 95 percent with 7 kHz bandwidth.” (2003 white paper)
- Clearer transmission of voice for a wider demographic
  - **Original narrowband codecs were designed by Western speakers for use in communications between Western dialects**
  - **Wideband codecs better support global IP communications**
- Ability for OEMs & network operators to deliver a service with differentiation other than price
- A voice communication experience in line with the expectations of today's more HD-aware consumers

How wide is wide?



Wideband demo—can you hear the difference??

"Seed, feed, seed" at different bandwidths and additive noise levels.

	3.3 kHz LP	7 kHZ LP
CLEAN		
24 dB SNR		
12 dB SNR		
0 dB SNR		

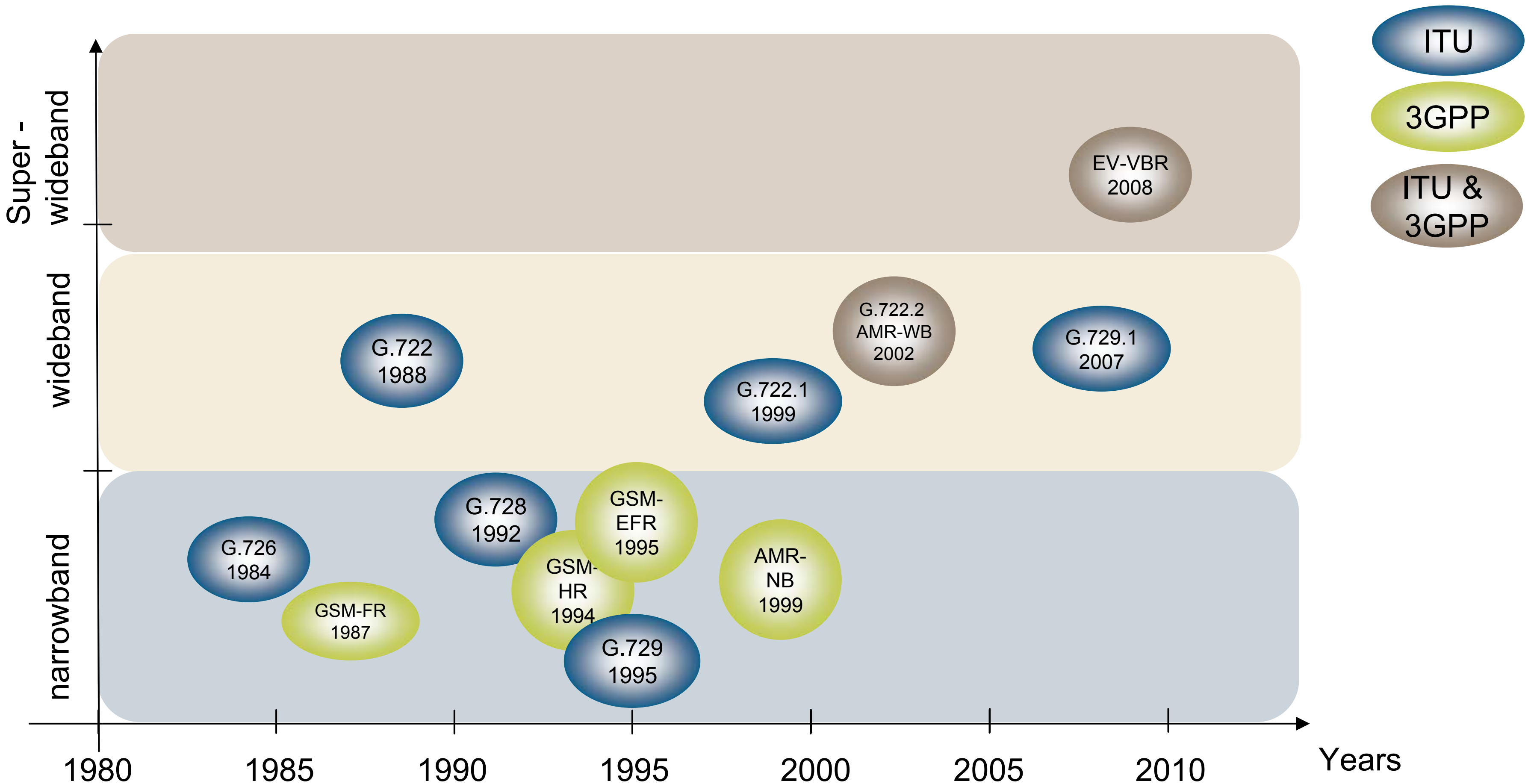
ETSI

Some Aspects of Wideband Speech in Enterprise Telephony

2nd Workshop on Wideband Speech, June 2005

freescale™ semiconductor

ITU and 3GPP codec roadmap



DSP algorithm complexity – can you afford it?

Algorithm	kbit/s	Freescale StarCore DSP Performance (approx mcps)
G.722	48,56,64	6.4
G.722.1	24,32	4.0
G.722.1c (Siren14)	24,32,48	9.24
G.722.2 (AMR-WB)	6.6 - 23.85	15-16
G.729.1	14 - 32	17.5 (est)
EVRC-C (Wide Band)	0.8, 2,4,8.55	45
GIPS ISAC	10 – 32	33 (sub opt)

Wideband voice enhancement devices

Wideband complexity extends beyond voice codecs e.g.

- WB VAD/CNG
- WB Noise Reduction: will require enhancements in the FFT/IFFT and other critical areas
- WB ALC: anticipated enhancements in the areas of (a) energy estimation (b) regulation speed control
- WB (Acoustic) ECAN: it will require development in several areas, due to increased expectations in voice quality.
  - Up-scaling (memory/CPU requirement increase)
  - Upgrade the adaptive filter (to work on fixed-point issues, increase the convergence speed/depth), double-talk detector (to enhance the double-talk detection quality; reduce temporal clipping).

Migration of AEC (32ms/192ms) to the wideband domain

Functions/functionalities affected (if no substantial design change implemented and no precision upgrade needed):

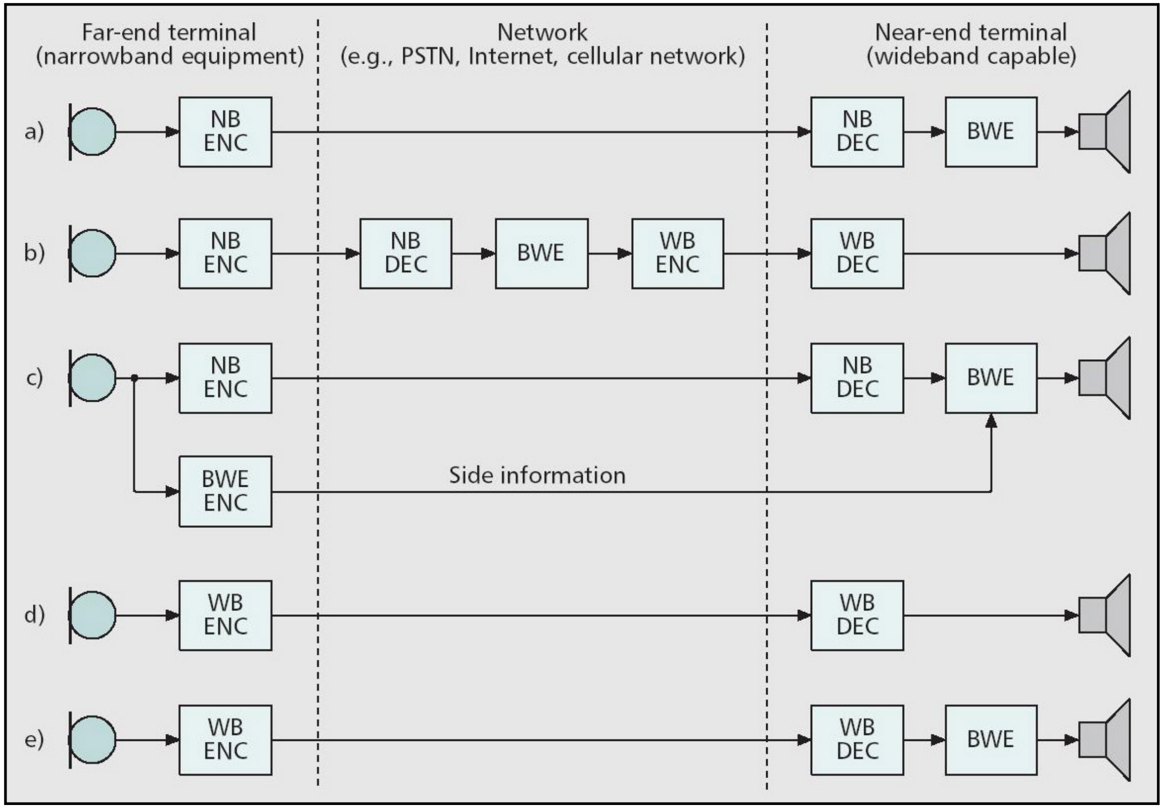
- update_fir (assuming same coverage, 24ms): cycles x 4
- apply_fir: cycles x 4
- update_sub_fir: cycles x 4
- apply_sub_fir: cycles x 4
- energies: cycles x 4 (assuming same windows in ms)
- decimation_filter: cycles x 2 (Q: why only 2x? A: because IIR type, not FIR type)
- dc_notch: cycles x 2
- sm_analys_filt_bank, sm_synthe_filt_ban: cycles x 2 (for the time being excluded)
- misc decisions: cycles x 1.5-2 (more investigation required)
- other misc code: cycles x 1.5-2 (more investigation required)

Summary: 7.5 MCPS (approx) NB increases to 21 MCPS (approx); memory – x3

The projected MCPS numbers are optimistic & do not take into account the following:

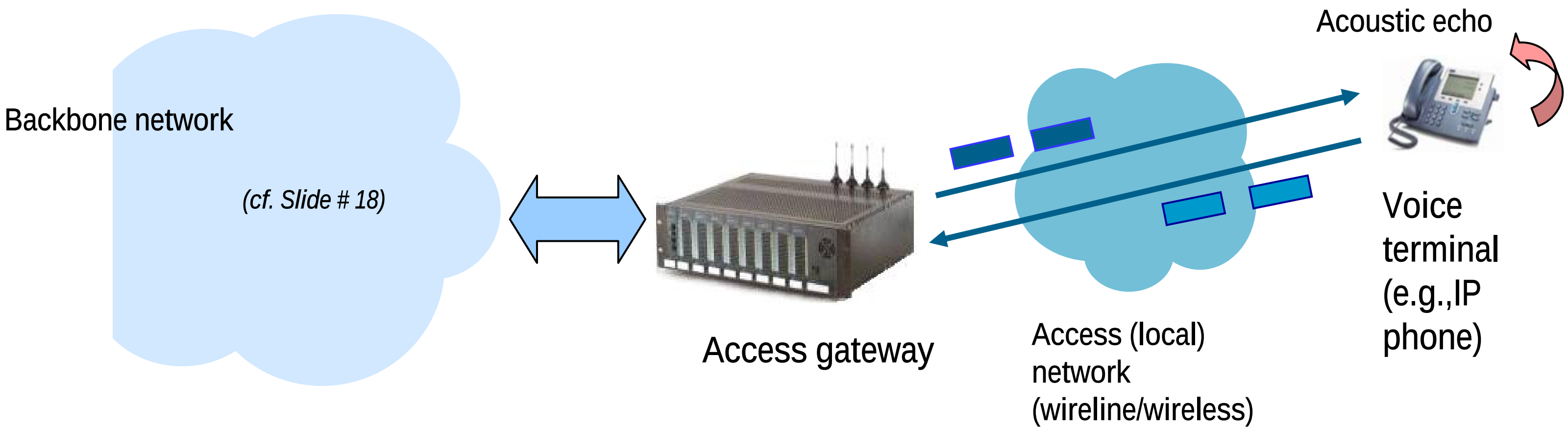
- **need for increasing the dispersion coverage to 64ms/128ms range**
- **need for increasing the fixed-point precision; if the dispersion coverage is 64ms the current precision used for ECAN and AEC solution is no longer adequate**
- **need for upgrading the NLP/CN functionalities to spectrum matching**
- **need for upgrading other essential functions to meet new demands associated with WB**

The co-existence of NB and WB



- (a) NB @ TX and BWE @ RX end
- (b) NB @ TX and BWE in the network
- (c) NB @ TX with BWE sent through the network
- (d) True WB encoding and decoding
- (e) WB @ TX and super-BWE @ RX end

Wideband VED's positioning: terminal vs. gateway



- The following advantages can be identified for the case of WB VED (DSP) functionality residing at the terminal :
  - No bulk delay of significance (Echo Cancellation)
  - No packet loss affecting VED operations
  - DSP dedicated to single channel (thus, typically there is lesser need for functional compromises)

- The following advantage can be identified for the case of WB VED (DSP) functionality residing at the access gateway:
  - more efficient utilization of DSP resources, thus, overall, more competitive solution (if cost is a predominant factor)
  - Maintenance/upgrade per user more straightforward and less costly

Source (graphic elements only): QiiQ Communications Inc., Cisco (Cisco IP Phone 7940G)

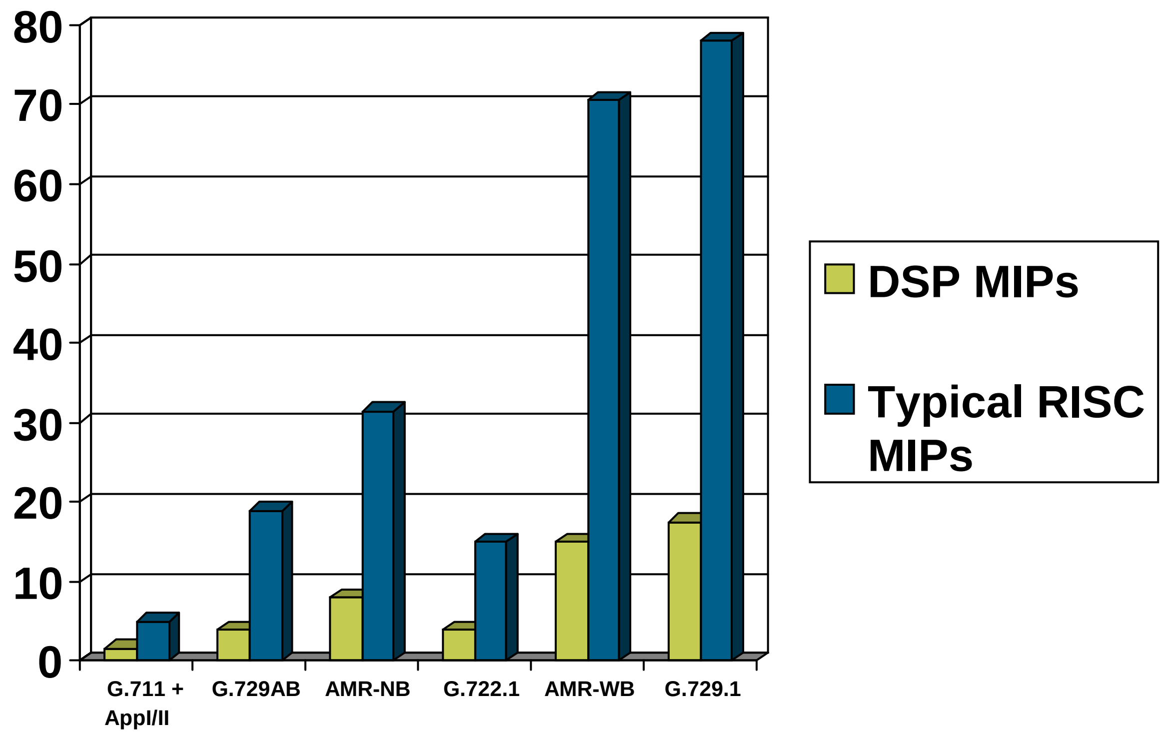
Wideband VED's positioning: terminal vs. gateway

Summary

- Network terminal appears to be the natural location for hosting DSP functions related to WB audio support;
- Network terminals such as IP phones are typically equipped with low to moderate DSP functionality or RISC performance
- The physical design of Wide Band terminals must change to accurately render wideband speech
  - Loudspeaker Enclosures (Frequency Response, THD @ LF)
  - Earpiece (Frequency Response @ LF)
- Network terminals with "bare bone" functionality may have reduced DSP functions; thus, the necessary DSP functions (VED: such as AEC, NR and alike) have to be located at the access gateway
- Positioning of DSP functions (VED) at the gateway contributes to:
  - more efficient utilization of DSP resources,
  - challenges related to inherent dilemma 'quality vs. channel density'
  - signal degradation due to 'tandem' effects
  - signal degradation due to packet loss

Quality, but at what cost?

- Processing power!!!
- High-compression wideband codecs ~2 to 4x complexity of narrowband counterpart
- Wideband Acoustic Echo Canceller ~2x the complexity of a narrowband AEC
- The good news – Moore's law and highly efficient coding techniques mean hands free wideband VoIP is a reality on today's RISC processors without DSPs or co-processors



Media framework roadmap

IP to IP (Voice/Fax/Data)

Phase #1
• Custom Media Channel Template

•Phase #2 Standard
• G.711 (5ms)
• Clear Channel
• G.726A (5ms)
• G.729AB
• G.723.1
•AMR-NB (IF1 & IF2)
• EVRC-A & EVRC-B (IF1 & IF2)
• Tones (inband & relay)
• Conferencing (3-Way)
• T.38 Fax Relay
• VBD

•Phase #3 Standard (TBD)
• G.728
• iLBC
• G.168 Ecan
• ALC

IP to TDM (Voice/Fax/Data)

Phase #1 Standard (Phase#2 IP to IP) plus:
• G.168 LEC
• NR
• ALC
• CID Type I (FSK & DTMF)
• CW Type II
• V.152
• Record & Playback Announcments

Phase #2 Standard

• iLBC
• G.728

Wide Band (Voice/Fax/Data) IP to IP & IP to TDM

Phase #1 Standard WB (IP to IP)

(Phase#2 IP to IP) plus:
• G.722 / G.722.1
•AAC-LC/LD/HE/++
• G.722.1C
• AMR-WB
• EVRC-WB

Phase #2 Standard WB (IP to IP)

• G.729.1
• WMV9
•iPCM-WB
• ISAC

Conferencing (Voice/Fax/Data & Video) IP to IP & IP to TDM

Phase #1 Standard Conf

(Phase#1 Wide Band) plus:
• Conferencing (N Way)
• Video Conferencing (Separate DSP Load)

IMS

Phase #1 Standard IMS

• TBD

freescale™ semiconductor

Conclusions & summary

- ► Wideband voice codecs offer significant advantages over narrow band
- ► Wideband processing impinges on many processing elements in the voice channel
- ► Computational complexity increases significantly at terminal and/or within infrastructure (factor of 2x to 4x or more)
- ► Today, DSP and RISC performance is up to the task
- ► Freescale has a rich and evolving wideband offering for infrastructure DSP & RISC based terminals (i.MX)

Related Session Resources

Session Location – Online Literature Library

http://www.freescale.com/webapp/sps/site/homepage.jsp?nodeId=052577903644CB

Sessions

Session ID	*Title*

Demos

Pedestal ID	*Demo Title*

freescale™
semiconductor