

2. přenáška

Hlas a jeho kódování

Osnova přednášky

1. Zpracování hlasu
2. Kodeky

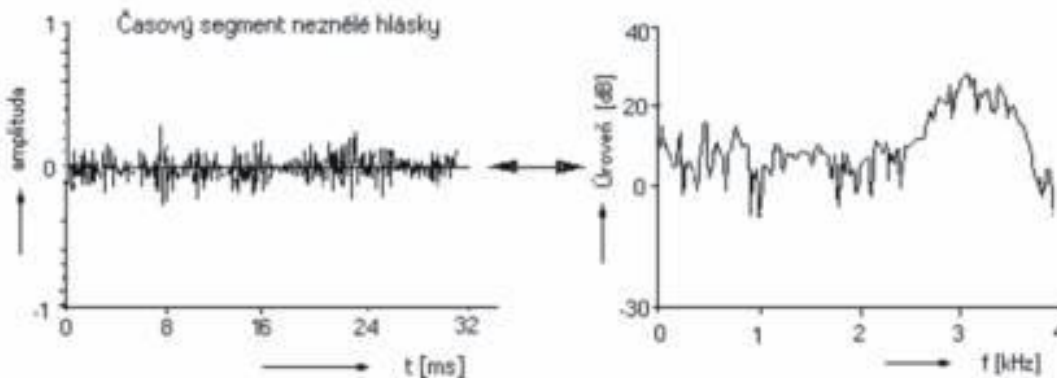
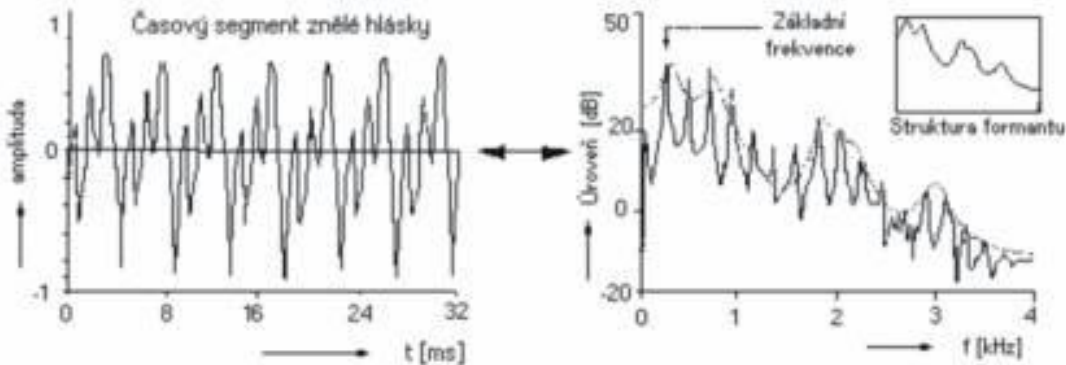
1. Hlas

Hlas

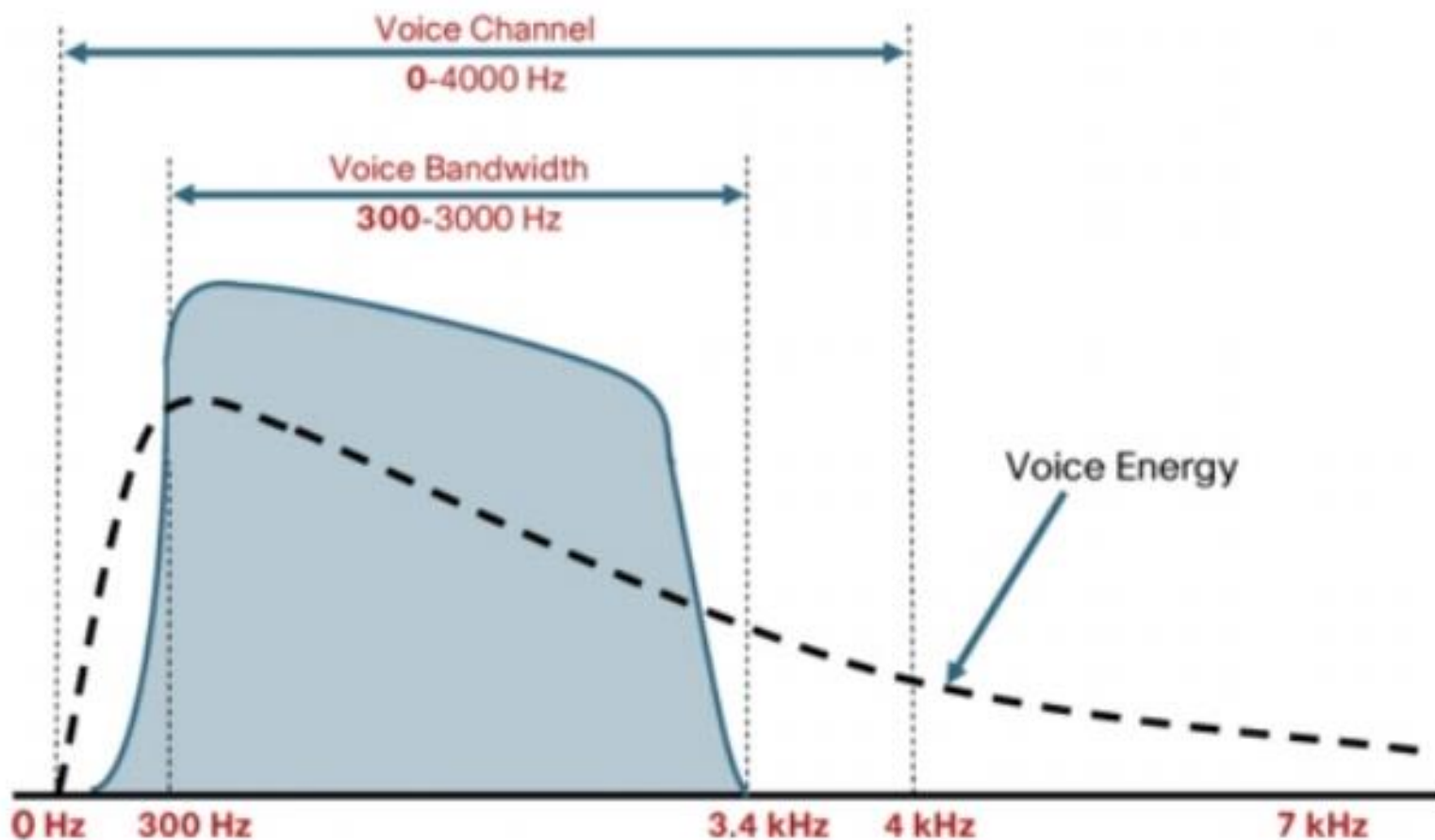
Kmitočet hlasivek je charakterizován základním tónem lidského hlasu (pitch periode) F_0 , který tvoří základ znělých zvuků (tj. samohlásek a znělých souhlásek). Kmitočet základního tónu je různý u dětí, dospělých, mužů i žen, pohybuje se většinou v rozmezí 150 až 4 000 Hz. Sdělení zprostředkované řečovým signálem je diskrétní, tzn. může být vyjádřeno ve tvaru posloupnosti konečného počtu symbolů. Každý jazyk má vlastní množinu těchto symbolů – fonemů, většinou 30 až 50.

Hlásky řeči dále můžeme rozdělit na znělé (n, e, ...), neznělé (š, č, ...) a jejich kombinace. Znělá hláska představuje kvaziperiodický průběh signálu, neznělá pak signál podobný šumu. Navíc energie znělých hlásek je větší než neznělých. Krátký časový úsek znělé hlásky můžeme charakterizovat její jemnou a formantovou strukturou. Formantem označujeme tón tvořící akustický základ hlásky. Ten vlastně představuje spektrální obal řečového signálu. Jemná harmonická struktura představuje chvění hlasivek.

Příklad časového a kmitočtového průběhu znělého a neznělého segmentu hovorového signálu



Pásmo potřebné pro telefonii



Pěvecké výkony jsou mimo oblast IP telefonie 😊



Deftones - [White Pony #07] Knife Party [foobar2000 v0.9.5.2]

File Edit View Playback Library Help

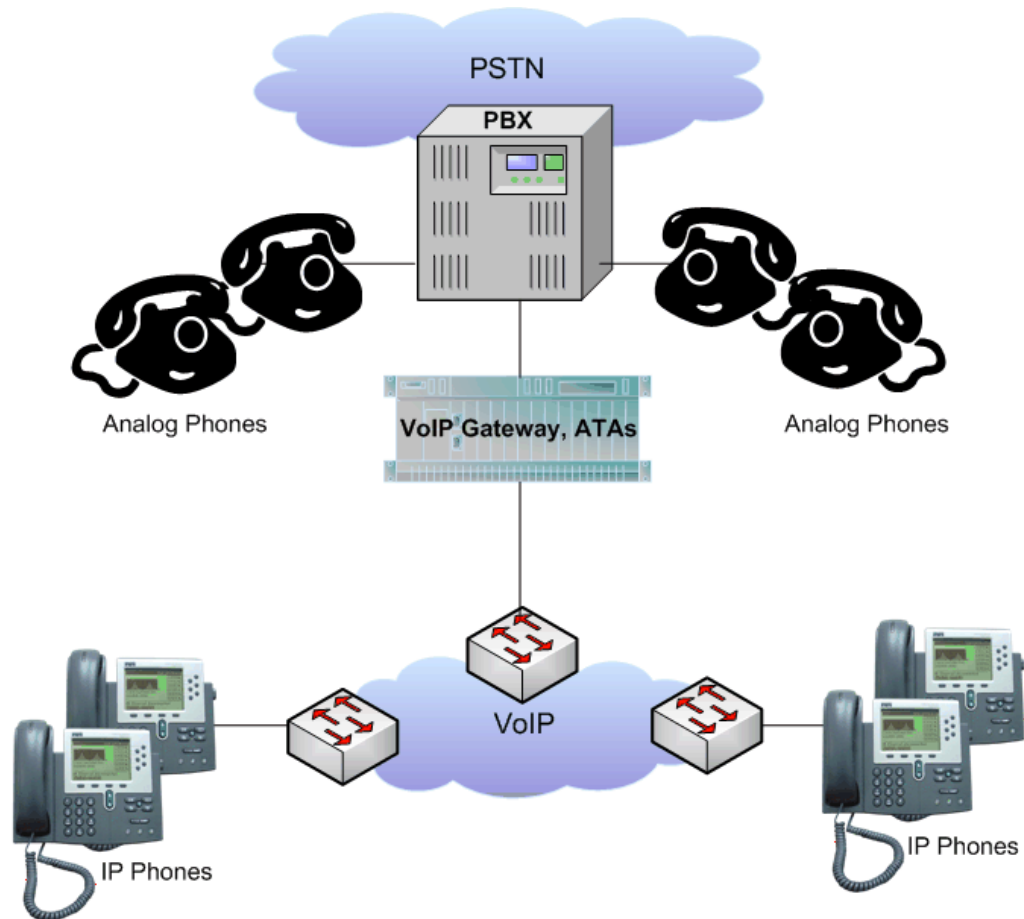
New Playlist (2)

Playi...	Artist/Album	Track No	Title / Track Artist	Dura...	cca. 16kHz
	Ear Drum Kru - Rádio_FM Vol.2 Pohoda_LIVE Exclusive [C...	04	Nizova Reda	4:19	
	TV On The Radio - Return To Cookie Mountain	02	Hours	3:55	
	Bloc Party - Intimacy	01.04	Biko	5:01	
	Bloc Party - Intimacy	01.09	Better Than Heaven	4:22	
	Bloc Party - Silent Alarm	01	Like Eating Glass	4:22	
	Bloc Party - Silent Alarm	12	Plans	4:10	
	Bloc Party - Silent Alarm	12	Plans	4:10	
	Bloc Party - Silent Alarm	12	Plans	4:10	
	Deftones - B-Sides & Rarities	14	Be Quiet And Drive (Far Away) (Acoustic)	4:32	
	Deftones - B-Sides & Rarities	01	Be Quiet And Drive (Far Away) (Acoustic)	4:32	

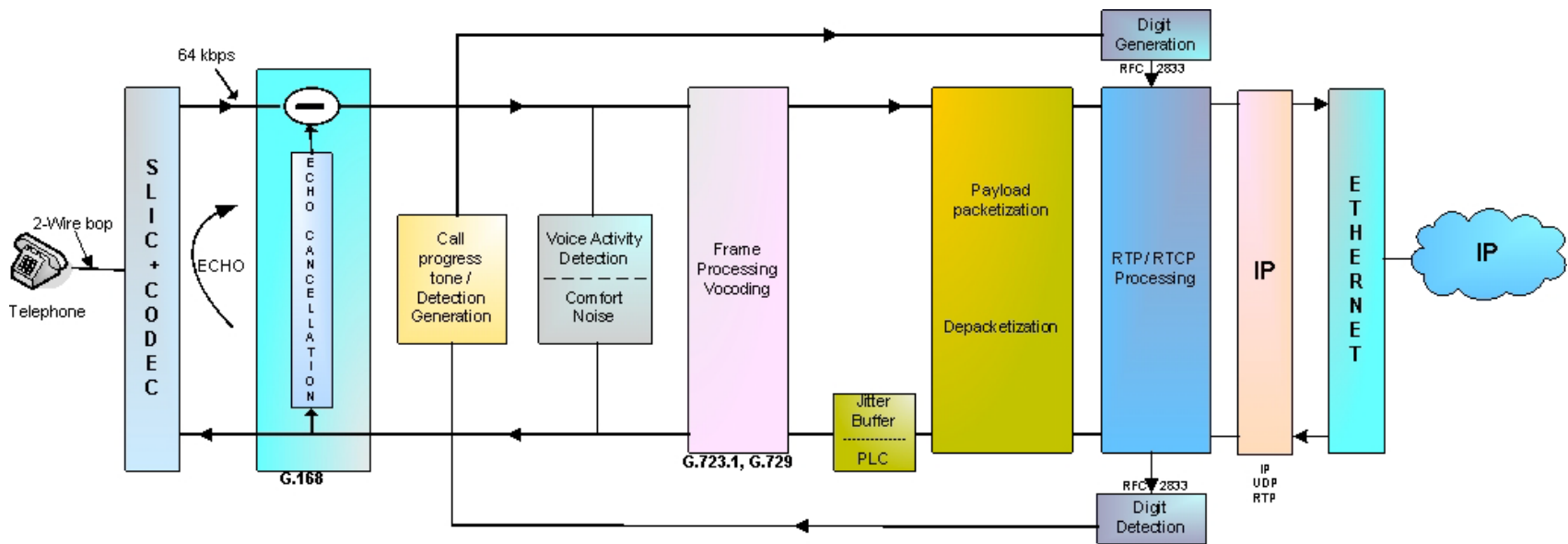
Blíže diskuze na <http://forum.avmania.e15.cz/viewtopic.php?style=2&f=1724&t=1062463&p=6027671&sid=>

2. Kodeky

Brána VoIP



Architektura VoIP brány



SLIC – Subscriber Line Interface Circuit

PLC – Packet Loss Concealment (odstranění důsledků ztrát rámců)

Echo canceliation – odstranění odezvy

Některé vlastnosti kodérů

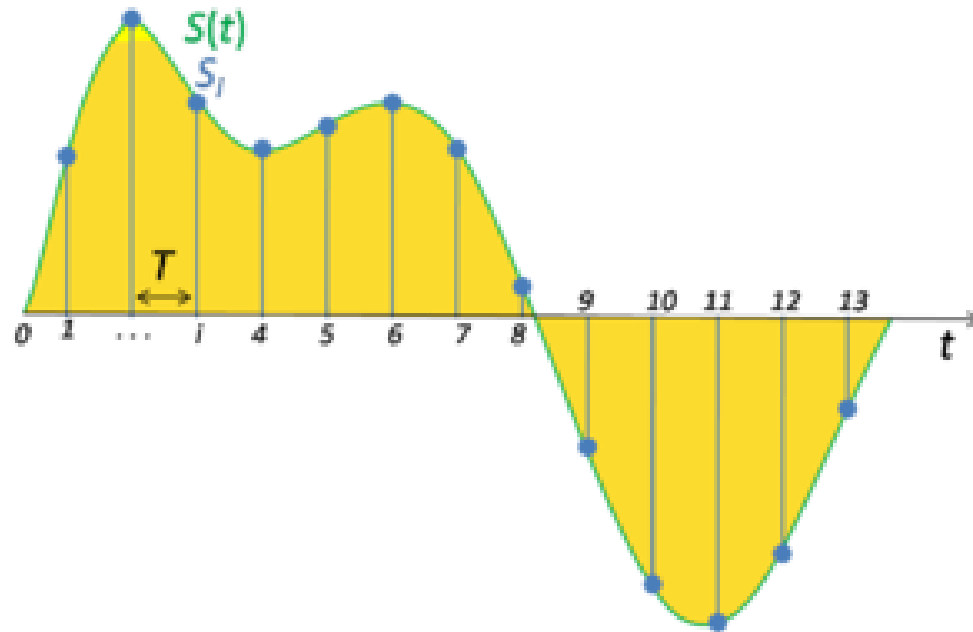
- **Voice Activity Detection (VAD)**
V pauze řeči je produkováno jen velmi malé množství bitů, které stačí na generování šumu.
- **Silence Suppression (DTX nebo SS) často VAD/SS**
Při delším odmlčení se úplně zastavuje přenos dat.
- **Comfort Noise Generation (CNG)**
Generování šumu pozadí
- **Packet Loss Concealment (PLC)**
Zahlazení ztrát paketů
- **Dynamic Jitter Buffer (Adaptive)**
Vyrovnávací paměť rozkmitu dob
- **Audio Codec Preferences**
Řešení preferencí kodeků
- **Perceptual enhancement**
Zeslabení šumu vytvořeného kódováním/dekódováním

Co je kodek?

Kodek (složenina z počátečních slabik slov „kodér a dekodér“, respektive komprese a dekomprese; převzato z anglického codec analogického původu) je zařízení nebo počítačový program, který dokáže transformovat datový proud (stream) nebo signál. Kodeky ukládají data do zakódované formy (většinou za účelem přenosu, uchování nebo šifrování), ale častěji se používají naopak pro obnovení přesně nebo přibližně původní formy dat vhodné pro zobrazování, případně jinou manipulaci.

(viz wikipedia)

Vzorkování a kvantování



Tři kroky zpracování hlasu:

1. Vzorkování (sampling) – diskretizace signálu v čase
2. Kvantování (quantizing) – diskretizace signálu v amplitudě
3. Kódování – diskrétní hodnoty vzorků jsou reprezentovány pomocí n -bitových čísel

Požadavky aplikací

Application	Application Availability	Delivery	Delay (One way)	Delay Variation
Real-Time	Actual/SLA	Actual/SLA		
Telephony	99.7/99.99%	>90%/99.0	40-80ms (reg) 100-250ms (global)	1-20ms
Data	99.7/99.99%	100%/99.0	50-100ms (reg) 150-350ms (global)	1-20ms
Video	99.7/99.99%	>95%/99.0	40-80ms (reg) 100-250ms (global)	1-20ms
Near Real-Time				
Audio	99.5/99.97%	>97%/99.5	50-100ms (reg) 150-350ms (global)	1-20ms
Data	99.5/99.97%	100%/99.5	60-150ms (reg) 150-350ms (global)	1-20ms
Video	99.5/99.97%	>95%/99.5	50-100ms (reg) 150-350ms (global)	1-20ms
Background/Batch				
Audio	99.0/99.95%	100%/99.5	80-200ms (reg) 200-500ms (global)	1-20ms
Data	99.0/99.95%	100%/99.5	80-200ms (reg) 200-500ms (global)	1-20ms
Video	99.0/99.95%	100%/99.5	80-200ms (reg) 200-500ms	1-20ms

Kodeky

název	modulace	bitový tok [kb/s]	MOS ⁴
<u>G.711</u>	PCM	64	4,4
<u>G.722</u>	SB-ADPCM	48/56/64	<i>neurčeno</i>
<u>G.723.1</u>	MPMLQ, ADPCM	5,3/6,4	3,98/3,5
<u>G.726</u>	ADPCM	16/24/32/40	3,85 pro 32 kb/s
<u>G.728</u>	LD-CELP	16	3,61
<u>G.729</u>	CS-ACELP	8	3,92/3,9 (G.729A)
GSM-FR	GMSK	13	3,7
iLBC	PCM	15,3/15,2	3,92/3,95
Speex	CELP	2,15–44,4	<i>neurčeno</i>

Mean Opinion Score – metoda založená na statistických šetřeních vybraného vzorku posuzovatelů s hodnocením od 1 (špatně) po 5 (výborně). Tím se liší od modernějších objektivních metod, např. PSQM (Perceptual Speech Quality Measurement) na níž je založen ITU standard P.861 (škála je od 0 po 6,5, ale některá testovací zařízení tyto hodnoty pak automaticky přepočítávají do škály MOS).

Zvláštnosti subjektivního hodnocení

- Starší dávají vyšší MOS než mladší, protože mají vytříbenější sluch.
- Muži dávají vyšší MOS než ženy, protože ženy sledují vyšší tóny a ty jsou uříznuty.
- Neznámý hlas vnímáme lépe a dáváme vyšší MOS.
- Cizí jazyk vede k nižšímu MOS.

Kodek G.711 (1965)

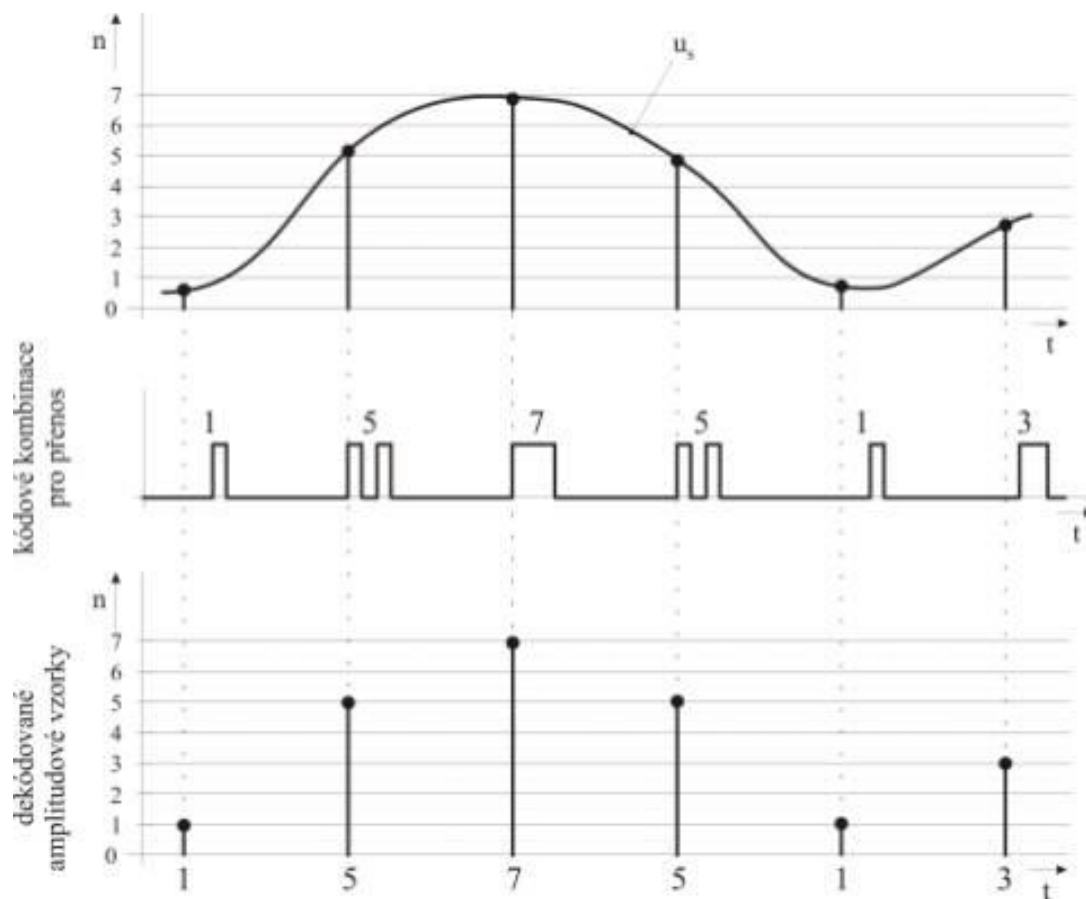
Kodek G.711 nepoužívá žádnou metodu komprese hlasu. Provádí pouze přímou PCM (Pulse Code Modulation), kde vzorek má velikost 8bitů a vzorkovací frekvence je 8 kHz, potřebná šířka pásma je 64 kb/s, zpoždění způsobené kódováním je pod 1 ms. Kvalita zvuku je při použití tohoto kodeku je obdobná kvalitě tradičního telefonního hovoru. Při logaritmické kompresi je 14b či 13b signálu převáděno na 8b signálu. G.711 je použit u v ISDN a na páteři sítí IP telefonie.

Používají se dva typy algoritmů, oba jsou semilogaritmické (companded PCM), ale liší se ve výběru konstant semilogaritmické křivky:

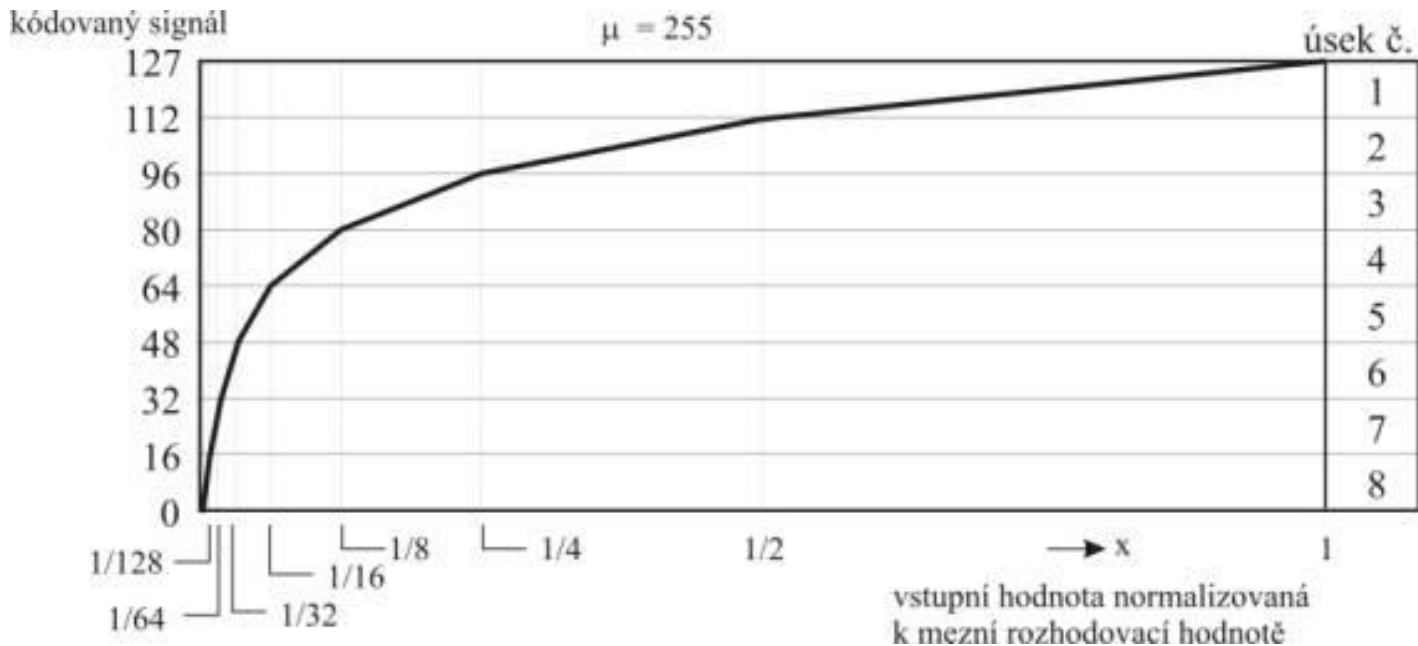
μ -law, který se používá v Severní Americe a Japonsku; Tam se používá vyšší komprese, protože v tamních telefonních sítích je používáno pouze 7 bitů pro přenos hlasu a osmý bit se používá na signalizaci.

a-law, který se používá ve zbytku světa a byl navržen s ohledem na jednodušší zpracování na počítači. V Evropě se používá pro přenos hlasu všech 8 bitů a signalizace je přenášena samostatným kanálem (Například pomocí SS7, CAS či jinou signalizací). V Evropě má tedy přenos hlasu v digitální telefonní síti vyšší kvalitu než stejný přenos pomocí telefonní sítě v Severní Americe či v Japonsku.

Princip kódování PCM



Kódování a-law



Pro a-law

1 bit určuje polaritu signálu

3 bity určují příslušný segment

4 bity kódují velikost kvantizační úrovně v daném segmentu

Při použití kódování čtyřmi bity je každý segment rozdělen do šestnácti dílčích hodnot, celkově pro kladné i záporné hodnoty signálu tak získáme rozlišení na 256 stupňů, což odpovídá kódování osmi bity.

Nové verze G.711

G.711.0 byl schválen v roce 2009. Poskytuje bezstrátovou kompresi G.711. Jeho plný anglický název je *Lossless compression of G.711 pulse code modulation*. Bývá také označován jako *G.711 LLC*.

G.711.1 byl schválen v roce 2008. Je to standard, který vede k rozšíření G.711 o možnost použít vzorkovací frekvenci 16 kHz a vyšší kvalitu zvuku s pomocí tří vrstev. Datový tok je 64, 80 nebo 96 kb/s. G.711.1 s datovým tokem 64 kb/s je kompatibilní s G.711. Vyšší datové toky nejsou kompatibilní s G.711.

Frekvenční pásmo je 50 až 7 000 Hz, tři módy, 64 kb/s, 80 kb/s and 96 kb/s. 2 bity navíc umožní 112 kb/s and 128 kb/s, tvoří modely R2B a R3.

Dolní pásmo – logaritmicky snímaná PCM + adaptivní vylepšování QoS.

Horní pásmo – MDCT (modified discrete cosine transformation), používá se i v MP3,
podstata: překrývání vzorků

Rozšíření roku 2010 (do 14 000 Hz, plánuje se 20 000 Hz) a bezstrátová bitová komprese.

RFC 5391

Network Working Group
Request for Comments: 5391
Category: Standards Track

A. Sollaud
France Telecom
November 2008

RTP Payload Format for ITU-T Recommendation G.711.1

Status of This Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Copyright Notice

Copyright (c) 2008 IETF Trust and the persons identified as the document authors. All rights reserved.

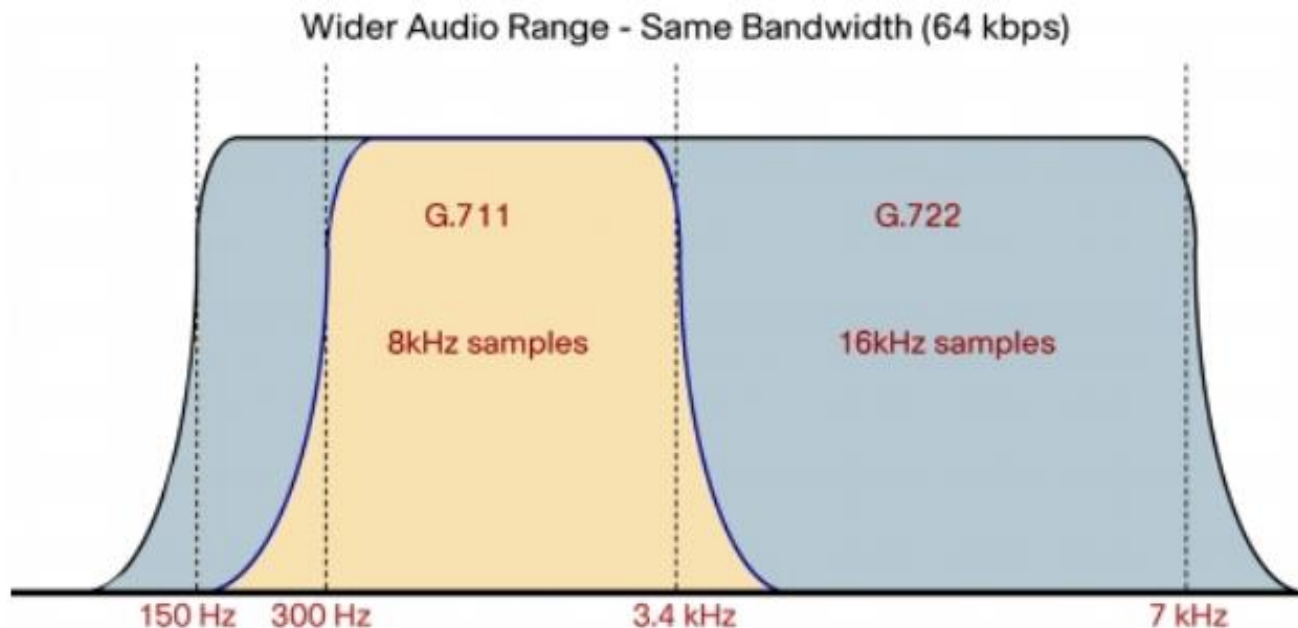
This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Abstract

This document specifies a Real-time Transport Protocol (RTP) payload format to be used for the ITU Telecommunication Standardization Sector (ITU-T) G.711.1 audio codec. Two media type registrations are also included.

G.722 (1988)

G.722 je nazýván širokopásmovým kodekem, protože používá dvojnásobný vzorkovací kmitočet (16 kHz), pásmo je rozšířeno na 7 kHz.



Tímto bylo dosaženo mnohem vyšší kvality hovoru, než u ostatních tehdejších kodeků. „Kvalita“ se projevuje především v tom, že lze jasně rozeznat, kdo hovoří, což je zvláště důležité u audiokonferencí a přináší rovněž výhodu z hlediska bezpečnosti (ochrana před falešným volajícím). Jde o 'embedded' kódér, což znamená, že se rychlost může volně přepínat mezi 48, 56 a 64 kb/s bez předchozího oznámení dekodéru.

Hodnocení G.722 (1988) a kodek G.722.1

Kodek G.722 je velmi vhodný pro všechny profesionální konverzační hlasové aplikace (algoritmické zpoždění je pouhých 1,5 ms) i pro HD. G.722 je podporován některými zařízeními pro videokonference a některými IP telefony.

Pro kódování je použit algoritmus SB-ADPCM (sub-band adaptive differential pulse code modulation – 2 kanály). Vyžaduje 30 % výkonu Pentia 100 resp. 10 MIPS DSP.

G.722.1 je nejnovější širokopásmový kodek pracující na 24 kb/s nebo 32 kb/s. Byl navržen společností PictureTel, později koupenou společností Polycom, která prodává 16 kb/s (verze kodeku Siréna). Kódovány jsou rámce 20 ms s predikcí na 20 ms. Verze 16 kb/s je podporována systémem Windows Messenger.

Rozdíl DPCM a ADPCM

Diferenciální pulsně kódová modulace **DPCM** (Differential Pulse Code Modulation) je modifikací PCM kódování a její princip byl publikován v roce 1952. Systém s DPCM kóduje rozdíly mezi okamžitou hodnotou vzorku signálu v daném vzorkovacím okamžiku a hodnotou predikovanou z předchozích vzorků. K zakódování rozdílu je zapotřebí nižšího počtu bitů. To vede k redukci přenosové rychlosti a tedy i množství přenášených dat se snižuje. DPCM se využívá především při kódování obrazu.

Adaptivní diferenciální pulsně kódová modulace **ADPCM** (Adaptive Differential Pulse Code Modulation) vylepšuje DPCM tak, že srovnávací průběh je vytvářen adaptivně a přizpůsobuje se konkrétním statistickým parametrům řeči. Výsledkem je ještě menší dynamický rozsah než v případě DPCM a tedy opět potřebujeme méně bitů k zakódování vzorku.

Kodek G.723.1 (1995)

V počátečních dobách VoIP VoIP Forum zvolilo tento kodek jako základní pro použití v protokolu H.323. Je rovněž používán mobilními videotelefony UMTS 99 (standard H.324M), kde Universal Mobile Telecommunication System je 3G systém standardů pro mobilní telefony. Spolu s G.729 jde o nejoblíbenější hlasový kodér pro VoIP.

Používá rámce délky 30 ms a odhaduje další průběh na 7,5 ms.

Má dva provozní režimy – 6,4 kb/s (r63 – 24bytové rámce) a 5,3 kb/s (r53 – 20bytové rámce). Mezi nimi je schopen přepínat při každém snímku. Norma dva režimy stanovuje jako povinné, bohužel řada implementací pracuje pouze v jednom ze dvou režimů.

Podporuje detekci hlasové aktivity (VAD – Voice activity detection), diskontinuální přenos (DTX – Discontinuous Transmission) a generování výplňového šumu (CNG – comfort noise generator). Ticho je kódováno ve velmi malých čtyřoktetových snímcích přenášených rychlostí 1,1 kb/s. Pokud informace o tichu nemusí být aktualizována, přenos se zastaví úplně.

Mimo systémy VoIP je G.723.1 použit v H.324 doporučeních a bude rovněž využit v nových standardech pro 3G-324 (organizace 3GPP, 3GPP2) pro 3G bezdrátové multimediální zařízení.

Je založen na 18 patentech, což může vést k lapáliím při licencování.

Způsob realizace kodeku G.723.1

Používá se ke kompresi zvuku, pracuje se zvukovými úseky délky 30 ms. G.723.1 se vyznačuje velmi nízkým bitovým tokem. Pro tento kodek jsou definovány dva algoritmy:

- MPC-MLQ (Multi-Pulse Maximum Likelihood Quantization) s bitovým tokem 6,3 kb/s;
- ACELP (Algebraic-Code Excited Linear Prediction) s bitovým tokem 5,3 kb/s. Například při ACELP a velikosti vzorku 160 bitů jsou nároky na pásmo jen 5,3 kb/s, za což se platí zpožděním způsobeným kódováním 37,5 ms a nižší kvalitou zvuku. PBX musí být nakonfigurovány na podporu ACELP.

Počet bitů velikosti vzorku (2, 3, 4, 5) má za následek 4 přenosové rychlosti. Není určen pro hudbu a nevysílá DTMF tóny spolehlivě (musí být předány out-of-band). Není určen ani pro modemy a faxové signály.

Výhoda ACELP

(Algebraic-Code Excited Linear Prediction)

Hlavní výhodou ACELP je, že použitá algebraická kódová kniha může být velmi velká (> 50 bitů) a to bez nepřiměřených nároků na paměť (RAM / ROM) či zpracování (CPU). Kodér algoritmu ACELP je založen na kódovacím schématu CELP, používá však specifickou algebraickou strukturu.

Všechny kódovací knihy mají problém s cizími jazyky a ty první.

Odolnost kodeku G.723.1

G.723.1 dosahuje MOS skóre 3,7 s režimu 5,3 kb/s a MOS 3,9 v režimu 6,4 kb/s. V prvním režimu je používán algoritmus MPC-MLQ (Multipulse LPC with Maximum Likelihood Quantization), v druhém. ACELP (adaptivní s dynamickou rychlostí při kompresi).

Tabulka porovnává výkon G.723.1 (6,3 kb /s) a ADPCM v Bell Labs v roce 1994. Cílem výmazů rámců bylo odsynchronizovat kodér a dekodér (aby bylo třeba více snímků na synchronizaci). V praxi bývá chybovost pod 3 % (v ideálním případě až pod 1%).

Reakce tandemu kodeků na ztrátu rámce

podmínky	G.721.1 6,3 kb/s	MOS	32 kb/s ADPCM
Čistý kanál	3.901		3.781
3% výpadků rámců	3.432		—
Tandem dvou kodeků	3.409		3.491

Kodeky G.726

Používá techniku ADPCM k zakódování G.711 bitových řetězců slova o délce 2, 3 resp. 4 bity, což vede k datové tokům 16, 24, 32 (dnes nejpoužívanější) či 40 kb/s. Vzorkovací frekvence je 8 kHz (typically 80 vzorků u rámce 10 ms). Dosahuje MOS 4,3 a je často brána jako měřítko kvality. Jde o kodek s malým zpožděním. Velikost snímků je 125 ms, není použita predikce. Zabere 30 % výkonu Pentia 100 resp. 10 MIPS DSP. K dispozici je rovněž verze známá jako G.727.

Je primárně používán na mezinárodní trunky v telefonní síti. Rovněž je používán v bezdrátových telefonních systémech a je používán na některých fotoaparátech Canon.

Kodek G.728 (1992–1994)

G.728 používá algoritmus LD-CELP (low-delay code excited linear prediction). Při pouhých 16 kb/s má MOS skóre srovnatelné s G.726 na 32 kb/s. V porovnání s PCM nebo ADPCM, které ignorují charakter signálu, je CELP kódér optimalizovaný pro hlas (vocoder). Modeluje konkrétní hlas za použití knihy lineárních prediktivních kódů). Předáváno je pak to, co nejlépe odpovídá charakteru daného hlasu. Pro hudbu je tento kodek nevhodný.

V rámci protokolu H.320 je používán pro videokonference a některé H.323 videokonferenční systémy, využívá se zde jeho nízkého zpoždění (625 μ s až 2,5 ms). Spotřebuje zato plný výkon Pentia 100 s 2 kB RAM.

Variantou tohoto standardu s použitím 15,2 kb/s (nižší rychlost) je RealAudio 28_8. Vyvinuté RealNetworks v roce 1995.

Rozdíl CELP a LD-CELP

CELP (Code - Excited Linear Prediction) jsou lineárně predikční kodeky s vektorovým kvantováním budícího signálu. Jejich přenosová rychlost se pohybuje od 4 do 8 kb/s. V pamětech kodéru a dekodéru jsou uloženy soubory možných posloupností vstupního signálu. Je-li v kódové knize nalezena posloupnost odpovídající budícímu signálu, je příslušná adresa budící posloupnosti (tj. vektoru uloženého v kódové knize) binárně přenesena do přijímače. Přijímač pak podle přijaté adresy generuje na základě shodné kódové knihy příslušnou budící posloupnost pro LP (linear prediction) hlasový syntezátor. Optimální vektor je určen užitím kritéria minima střední kvadratické chyby.

Jedna z nevýhod původního CELP algoritmu je velká výpočetní náročnost prohledávání kódových knih. Mnoho CELP algoritmů vyžaduje procesory schopné zpracovat více než 20 MIPS a prostor pro kódovou knihu kolem 40 kbytu. Kódovací zpoždění je až 35 ms a omezuje použití CELP kodeků pro přenosy na delší vzdálenosti a jejich vícenásobné užití. Z toho všeho vyplývá snaha o vývoj vhodných prohledávacích algoritmů.

Způsob kódování s krátkou dobou zpoždění **LD-CELP** (LD – Low Delay) je založen na zpětnovazebním řízení adaptace s kódovacím zpožděním do 2ms a přenosovou rychlostí 16 kb/s. Koncepte LD-CELP se od kodeků CELP liší v řešení vlastního hlasového syntezátoru a způsobu jeho řízení. U kodeků CELP se přenáší v časovém multiplexu lineárně predikční koeficienty vyčíslené v kodéru krátkodobou (STP) a dlouhodobou analýzou (LTP), které slouží v přijímači k periodickému nastavování LPC filtrů hlasového syntezátoru. Naproti tomu mezi kodérem a dekodérem LD-CELP se tato data nepřenášejí a jak kodér, tak dekodér jsou řízeny zpětnovazebně.

Kodek G.729 (1995–1996)

Jde o velmi populární kodek pro přenos hlasu přes Frame Relay a pro V.70 hlasové a datové modemy. Spolu s G.723 jde o nejoblíbenější hlasový kodér pro VoIP, ale stále ještě není nativně podporován na platformě Windows.

Jedná se o velmi používaný zvukový kodek pracující se zvukovými vzorky délky 10 ms, který používá kompresi zvuku CS-ACELP (Conjugate-Structured Algebraic-Code Excited Linear Prediction). Predikce je na 5 ms. Dosahovaný bitový tok 80bitových rámců má hodnotu 8 kb/s, existuje však i rozšíření pro 6,4 kb/s a 11,8 kb/s, které zajišťuje větší kvalitu přenášeného zvuku. Tento kodek není určen pro přenos hudby, nepřenáší ani spolehlivě DTMF tóny. Neumožňuje přenos faxových a modemových signálů.

Přílohy A a B G.729 definují VAD, CNG, a DTX schémata pro G.729.

Rámce poslané k aktualizaci pozadí hluku jsou 15 bitů dlouhé a zasílány pouze v případě změny pozadí.

Je zde 20 patentů, což opět pro projektanty přináší problémy.

Většinu patentů mají v rukou SiproLabs (www.sipro.com), platí to i pro G.723.1. ³²

Jak pracuje kodek G.729

- CS-ACELP kodér pracuje s rámcí řeči o délce **10 ms**, což při vzorkovací frekvenci 8000 Hz odpovídá 80 vzorkům za sekundu. Každý 10 ms je rámec řečového signálu analyzován pro získání parametrů CELP modelu. Tyto parametry jsou kódovány a přenášeny komunikačním kanálem.
- V dekodéru jsou tyto parametry použity k obnovení excitačního signálu a koeficientů syntetického filtru. Řečový signál je pak rekonstruován filtrací excitačního vektoru přes syntetický filtr. Kvalita řečového signálu je pak ještě zvýšena postfiltrem.
- Celkové zpoždění nutné pro výpočet algoritmu je **15 ms**, další přídatná zpoždění mohou vzniknout během přenosu v komunikačním kanálu nebo při multiplexování dat.
- **Velkou nevýhodou** je zde (stejně jako u G.728 LD-CELP) velký výpočetní výkon potřebný k prohledání celé kódové knihy a nalezení nejlepšího excitačního vektoru.

G.729A

G.729 (1996) vyžaduje cca 20 MIPS pro kódování a 3 MIPS pro dekódování (asi o 30% méně než G.723.1).

G.729A (1995): Je zpětně kompatibilní s G.729 a má výhodu v menší náročnosti výpočtu.

G.729 je často používán v VoFR, je podporován výrobci mnoha zařízení Frame Relay.

Kodek GSM

(Global System for Mobile communications)

Testování kódovacích algoritmů se provádělo pro těchto sedm jazyků: angličtina, němčina, francouzština, finština, japonština, španělština a hindština. Výběr probíhal z těchto typů kódování:

Typ kódování	MOS	Přenosová rychlost(kb/s)	MIP S	Výrobce
RPE - LPC	3,54	14,77	1,5	Philips (Německo)
MPE - LTP	3,27	13,20	4,9	IBM (Francie)
SBC - APCM	3,14	13,00	4,9	Ellemtel (Švédsko)
SBC - ADPCM	2,92	15,00	4,9	Telecom Res. (Anglie)

Výsledkem byla nakonec kombinace řešení od firmy Philips a IBM. Vznikl tak kodek s označením RPE - LTP (Regular Pulse Excitation – Long Term Prediction). Obvody kodeku jsou realizovány signálovými procesory a lze je podle funkce rozdělit na tyto hlavní bloky:

- předzpracování signálu,
- analýza LPC,
- krátkodobá analýza a filtrace,
- kódování RPE a analýza LTP.

Celková bitová rychlost signálu bitový 13 x 8 kHz = 104 kb/s se po analýze RPE–LTP změní z původní rychlosti 104 kb/s na 13 kb/s.

Kodeky GSM-FR, GSM-HR, GSM-EFR

GSM full-rate Používá techniku RPE-LTP, kóduje rámce 20 ms rychlostí 13 kb/s. Nepoužívá predikci. Není extrémně náročný, požaduje 4,5 MIPS a méně než 1kB RAM. Je patentován Philipsem.

Primární určení tohoto prvního GSM digitálního kodeku bylo zařízení DCME (Digital Circuit Multiplex Equipment), pro satelitní telefonii a pro bezdrátové standardy jako je DECT (Digital Enhanced Cordless Telecommunications).

Dnes už kvalita zvuku nevyhovuje a tak byl kodek GSM-FR nahrazen modernějšími variantami FR kodeku, které poskytují vyšší kvalitu zvuku při nižší přenosové rychlosti.

GSM half-rate (1994) má menší nároky na pásmo (5,6 kb/s) Rámce jsou dlouhé 20 ms a predikce je 4,4 ms. Vyžaduje cca 30 MIPS a 4 Kb RAM. Kodek byl citlivý na hluk pozadí, a proto nebyl úspěšný.

GSM enhanced full rate (1995) používá techniku CD-ACELP. Rámce jsou dlouhé 20 ms a rychlost 12,2 kb/s. Standardizován TIA a je základním pro US služby celulární telefonie.

MOS GSM kodeků v různých prostředích

kodek	klid	auto	ulice
GSM FR	3.71	3.83	3.92
GSM HR	3.85	3.45	3.56
GSM EFR	4.43	4.25	4.18
Reference bez kódování	4.61	4.42	4.35

iLBC (2004)

iLBC (Internet Low Bit Rate Codec) je další kodek založený na lineární predikci. Jde o otevřený úzkopásmový zvukový kodek specifikovaný v RFC 3951. Zásadní rozdíl oproti starším standardům založených na PCM (jako je G.711) je , že pro lineární predikci používá velmi kvalitní blokově nezávislý algoritmus.

Vzorkovací frekvence je 8 kHz/16 bit (160 vzorků pro 20 ms rámce, 240 vzorků pro 30 ms rámce). Má fixní bitovou rychlost 15,2 kb/s pro 20 ms rámce, 13,33 kb/s pro 30 ms rámce). Má fixní velikost rámce (304 bitů na blok 20 ms rámce, 400 bitů na blok 30 ms rámce). Zátěž procesoru je srovnatelná s G.729A, má ale vyšší odolnost vůči ztrátě paketů.

Kodek je určen pro aplikace reálného času jako je telefonie a videokonference.

Vyvinuly ho Global IP Solutions a je popsán v RFC 3951. Používají ho Gizmo5, QuteCom, Google Talk, Yahoo! Messenger, Polycom IP Phone, Maemo Recorder (pro Nokie) a další.

iLBC v RFC 3951

Network Working Group
Request for Comments: 3951
Category: Experimental

S. Andersen
Aalborg University
A. Duric
Telio
H. Astrom
R. Hagen
W. Kleijn
J. Linden
Global IP Sound
December 2004

Internet Low Bit Rate Codec (iLBC)

Status of this Memo

This memo defines an Experimental Protocol for the Internet community. It does not specify an Internet standard of any kind. Discussion and suggestions for improvement are requested. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (2004).

Abstract

This document specifies a speech codec suitable for robust voice communication over IP. The codec is developed by Global IP Sound

iLBC freeware

<http://www.iLBCfreeware.org>
iLBCfreeware



[News](#)

[Software](#)

[Documentation](#)

[Support](#)

What is iLBC?

iLBC (internet Low Bitrate Codec) is a FREE speech codec suitable for robust voice communication over IP. The codec is designed for narrow band speech and results in a payload bit rate of 13.33 kbit/s with an encoding frame length of 30 ms and 15.20 kbps with an encoding length of 20 ms. The iLBC codec enables graceful speech quality degradation in the case of lost frames, which occurs in connection with lost or delayed IP packets.

Features

- Bitrate 13.33 kbps (399 bits, packetized in 50 bytes) for the frame size of 30 ms and 15.2 kbps (303 bits, packetized in 38 bytes) for the frame size of 20 ms
- Basic quality higher than G.729A, high robustness to packet loss
- Computational complexity in a range of G.729A
- Royalty Free Codec

Current Release

- [Announcements](#)
- IETF RFCs:
 - [iLBC codec spec](#)
 - [RTP Payload Format for iLBC](#)

Project News

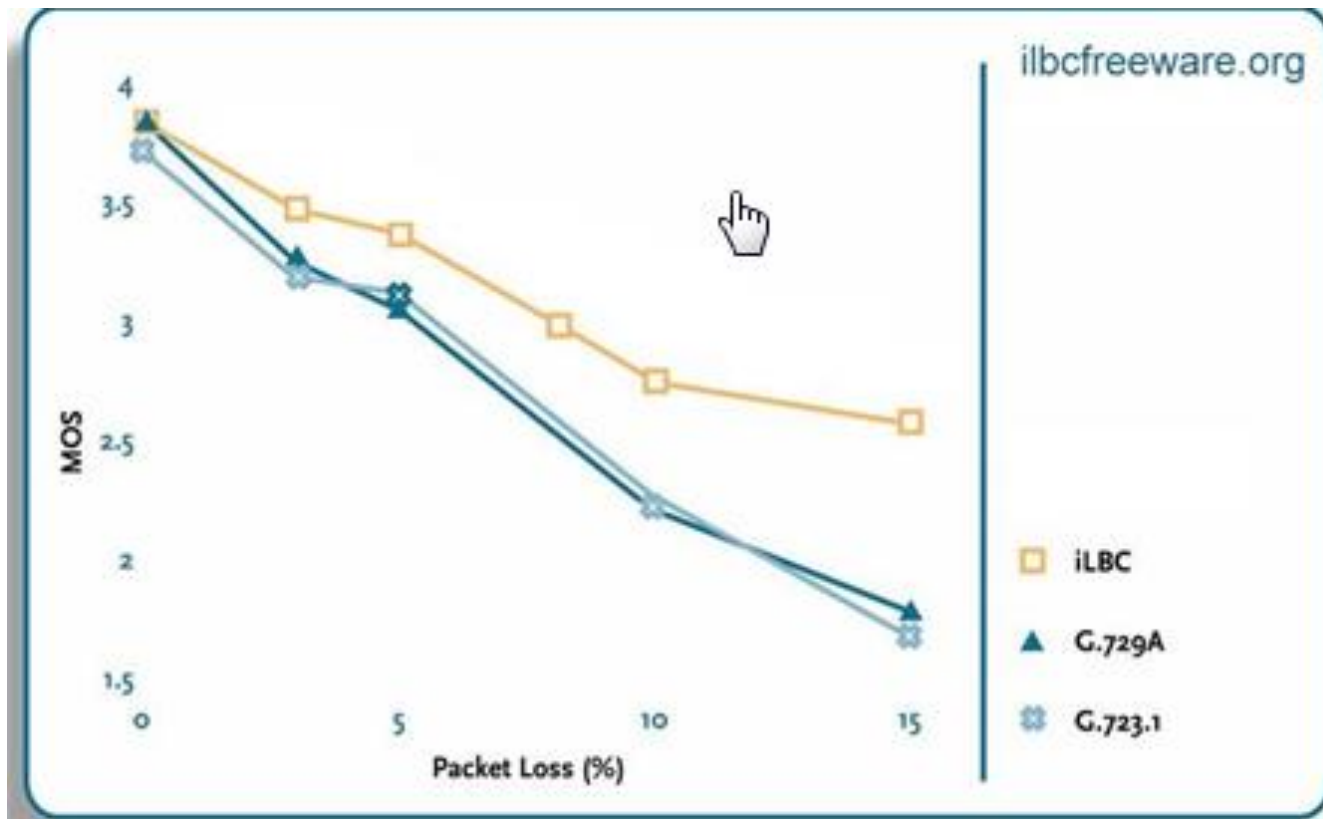
Dec 6th, 2004. The final IETF RFC and the [iLBC RTP Profile draft](#) is

Aug 9th, 2004. GIPS iLBC has been the recently released [CableLabs Codec](#)

[More...](#)

Implementations

Přednost iLBC: vyšší imunita vůči výpadkům rámců



iLBC ob stojí i bez výpadků

Condition: Clean channel				
ILBC is:	> G.723.1	>= G.729	>= G729.A	<= G.729.E

Condition: Packet loss				
ILBC is:	>> G.723.1	>> G.729	>> G729.A	> G.729.E

Bliže: <http://www.packetizer.com/codecs/ilbc/iLBC.WP.pdf>

Podpora iLBC Cisco

Cisco IOS Software Releases 12.4 T

iLBC Codec Support

[HOME](#)

[SUPPORT](#)

[PRODUCT SUPPORT](#)

[CISCO IOS AND NX-OS SOFTWARE](#)

[CISCO IOS SOFTWARE RELEASES 12.4 T](#)

[CONFIGURE](#)

[FEATURE GUIDES](#)


iLBC Codec Support



[Table Of Contents](#)

[iLBC Support](#)

Downloads

[iLBC Codec Support](#) 

iLBC Support

The internet Low Bitrate Codec (iLBC) is a standard, high-complexity speech codec that is suitable for robust voice communication over IP. iLBC has built-in error correction functionality that helps the codec perform in networks with a high-packet loss.

Configuration Information

Configuration information is included in the "[Dial Peer Overview](#)" chapter and "[Dial Peer Features and Configuration](#)" chapter in [Dial Peer Configuration on Voice Gateway Routers](#) at the following URL:

http://www.cisco.com/univercd/cc/td/doc/product/software/ios123/123ccqr/vvfax_c/int_c/dpeer_c/index.htm

Command Reference Information

Command reference information is included in the Cisco IOS Voice Command Reference, Release 12.4T, at the following URL:

http://www.cisco.com/univercd/cc/td/doc/product/software/ios124/124cr/hvr_r/index.htm

New or Modified Commands

The following commands are new or modified for this feature:

- `codec ilbc`
- `codec preference`
- `rtp payload-type`

List of All Release 12.4T Commands

An alphabetical list of all Cisco IOS Release 12.4T commands is in the [Cisco IOS Master Commands List Release 12.4T](#) at the

Feedback: Help us help you

Please rate this document.

- Excellent
- Good
- Average
- Fair
- Poor

This document solved my problem.

Speex (2002)

Jedná se opět o otevřený zvukový kodek, který používá pro kompresi zvuku metodu CELP. Používá šířku pásma od 2 do 44 kb/s. Vzorkovací frekvence lze zvolit 8, 16 a 32 kHz. Jde o open a free software, neboli odpadají problémy s patenty a licencemi. Je součástí projektu GNU a je dostupný v rámci revidovaných BSD licencí. Jde tedy o alternativu k drahým proprietárním kodekům. Jeho specifikace mu umožňuje dosahovat velmi dobrých kompresních poměrů při zachování srozumitelnosti. Není primárně určen (jako ostatní podobné kodeky) na kompresi telefonních hovorů, ale na VoIP.

API Speexu ke stažení



- Home
- News
- Comparison
- Samples
- Downloads
- Documentation
- Roadmap
- Plugins & Software
- Contact

Downloads

The latest development version of Speex is always available through Git by at <http://git.xiph.org/speex.git>. Note that packaging information for Debian can be found on the [packages page](#). In addition, [RareWares](#) maintains useful [Windows utilities](#) such as SpeexDrop.

Current *Unstable Release (recommended)*

1.2rc1

This release adds support for acoustic echo cancellation with multiple microphones and multiple loudspeakers. It also adds an API to decorrelate loudspeaker signals to improve multi-channel performance. In the bugfix department, there are fixes for a few bugs in the echo canceller, jitter buffer and preprocessor. At this point, the API for 1.2 should be stable and only a few very minor additions are planned.

Downloads:

- [Source Code](#)

1.2beta3

The most obvious change in this release is that all the non-codec components (preprocessor, echo cancellation, jitter buffer) have been moved to a new *libspeexdsp* library. Other changes include a new jitter buffer algorithm and resampler improvements/fixes. This is also the first release where libspeex can be built without any floating point support. To do this, the float compatibility API must be disabled (`--disable-float-api` or `DISABLE_FLOAT_API`) and the VBR feature must be disabled (`--disable-vbr` or `DISABLE_VBR`).

Downloads:

- [Source Code](#)
- [Windows binaries](#)

Další práce ITU-T SG16 Speech/Audio

G.718 Vůči chybám odolné úzkopásmové i širokopásmové kódování řeči a audio 8-32 kb/s + Annex B superpásmové rozšíření (28–48 kb/s). Je zde možnost stereo rozšíření.

G.719 nekomplexní, pro kódování audia ve vysoké kvalitě, aplikace jsou konverzačního charakteru (32–128 kb/s)

Další nové kodeky

SILK je u Skype nástupce původního SVOPC. Byl vyvinut pro desktop s matematickým koprocesorem, pokud není koncové zařízení PC, nevyhovuje. Má jedinou vzorkovací rychlost 16kHz, bitovou rychlost 8 kb/s. Založen je na lineární predikci 5,2 ms.

SVOPC měl audio a řečový mód, SILK tento rozdíl zmenšuje.

SILK má vzorkovací frekvenci 8, 12, 16 či 24 kHz (tři navíc) a bitovou rychlost od 6 do 40 kb/s (proměnlivou dle podmínek přenosu).

Může použít nízké algoritmické zpoždění 25,2 ms (20 ms rámeček + 5,2 ms prognóza).

BroadVoice je založen na dvoufázovém kódování TSNFC (Two-Stage Noise Feedback Coding) nahrazujícím populární CELP (Code-Excited Linear Prediction). Má velmi nízké zpoždění, algoritmická práce s vyrovnávacími paměti zpožďuje o pouhých 5 ms (ve srovnání s 15 až 40 ms řady dalších kodeků).

CELT (Constrained-Energy Lapped Transform) je kodek s ultranízkým zpožděním (typicky od 5 do 22.5 ms) a plným audio pásmem (≥ 20 kHz; vzorkovací rychlost je 8 kHz až 48 kHz). Podporuje hlas i hudbu včetně sterea. Bitová rychlost je 32 kb/s až 128 kb/s i vyšší.

Tento kodek je zamýšlen jako most, který vyplňuje mezeru mezi kodeky Vorbis a Speex pro aplikace, které vyžadují jak vysoce kvalitní zvuk, tak nízké zpoždění potřebné pro dialog (Vorbis je svobodný software zamýšlený jako náhrada MP3).

Skype Silk

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 9, 2010

K. Vos
S. Jensen
K. Soerensen
Skype Technologies S.A.
March 8, 2010

SILK Speech Codec
draft-vos-silk-01

Abstract

This document describes SILK, a speech codec for real-time, packet-based voice communications. Targeting a diverse range of operating environments, SILK provides scalability in several dimensions. Four different sampling frequencies are supported for encoding the audio input signal. Adaptation to network characteristics is provided through control of bitrate, packet rate, packet loss resilience and use of discontinuous transmission (DTX). And several different complexity levels let SILK take advantage of available processing power without relying on it. Each of these properties can be adjusted during operation of the codec on a frame-by-frame basis.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-

BroadVoice

The screenshot shows the Broadcom website's 'Downloads & Support' page for the 'BroadVoice® Speech Codec Open Source C Code'. The page features a red header with the Broadcom logo and navigation links. The main content area includes an 'Overview' section with a text box describing the BroadVoice family of algorithms, a 'Technology' section, and a 'LINKS' sidebar with a list of related resources.

BROADCOM. 日本語 한국어 简体中文 繁体中文 Search

About Us Investors Broadband Communications Mobile & Wireless Network Infrastructure Downloads

Downloads & Support

Home > Support > BroadVoice® Speech Codec Open Source C Code adjust font size: A

BroadVoice® Speech Codec Open Source C Code

Overview

BroadVoice® is a family of speech coding algorithms created by Broadcom and standardized by CableLabs®, SCTE®, and ANSI for Voice over IP applications in cable telephony. BroadVoice is also part of the ITU-T Recommendations J.161 and J.361. To encourage wide-spread use of BroadVoice in diverse speech compression applications, Broadcom is providing both the floating-point and fixed-point C source code of BroadVoice on a royalty-free basis under the GNU Lesser General Public License (LGPL), version 2.1, as published by the Free Software Foundation.

Technology

LINKS

- + Bluetooth
- BroadVoice®
 - ▾ Codec Comparison
 - ▾ Demo Samples
 - ▾ Downloads
 - ▾ Documentation
 - ▾ Published Papers
 - ▾ Contact
- + Communications Processors
 - ▾ Crystal HD
 - ▾ Dune Network Fabrics

CELT

AVT Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 14, 2010

J-M. Valin
Octasic Semiconductor
T. Terriberry
Xiph.Org Foundation
G. Maxwell
Juniper Networks
C. Montgomery
Xiph.Org Foundation
July 13, 2009

Constrained-Energy Lapped Transform (CELT) Codec
draft-valin-celt-codec-01

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

Na dokumentaci pro CELT se ještě čeká



- Home
- News
- Comparison
- Samples
- Downloads
- Documentation
- Presentations
- Software
- Contact

Documentation

Manual

Unfortunately, there is no formal manual for CELT yet. However, you can find some information about it:

- The [README file](#) distributed with libcelt
- On [Jean-Marc Valin's blog](#)
- The [celt-codec.org presentations page](#)

API Reference

The CELT API Reference Manual provides information on programming with libcelt.

- [libcelt API](#)
- [internal API](#) (only useful for developing CELT itself)

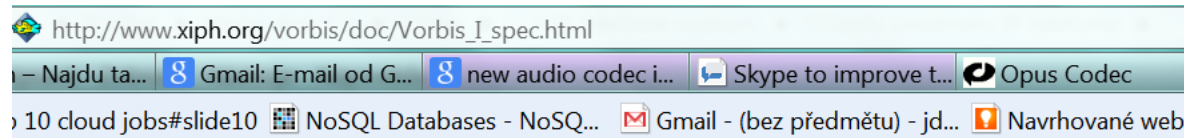
Papers

The first paper describes version 0.3.2 of the codec, while the second describe the low-complexity mode of version 0.5.1. While there are minor differences with the current version (mostly with the bit allocation), mc of the paper is still valid.

- J.-M. Valin, T. B. Terriberry, C. Montgomery, G. Maxwell, [A High-Quality Speech and Audio Codec With Less Than 10 ms delay](#), To appear in *IEEE Transactions on Audio, Speech and Language Processing*, 200
- J.-M. Valin, T. B. Terriberry, G. Maxwell, [A Full-Bandwidth Audio Codec with Low Complexity and Very](#)

Vorbis

- Vorbis je ztrátový audio formát, který se měl stát náhradou formátu MP3. Jeho vývoj je zastřešen organizací Xiph.Org Foundation. Vorbis audio bývá nejčastěji uloženo v kontejneru Ogg a v tomto spojení je nazýváno Ogg Vorbis.
- Softwarová implementace formátu Vorbis vyvíjená nadací Xiph.Org, kodek libvorbis, je svobodný a open source software uvolněný pod licencí BSD.



Vorbis I specification

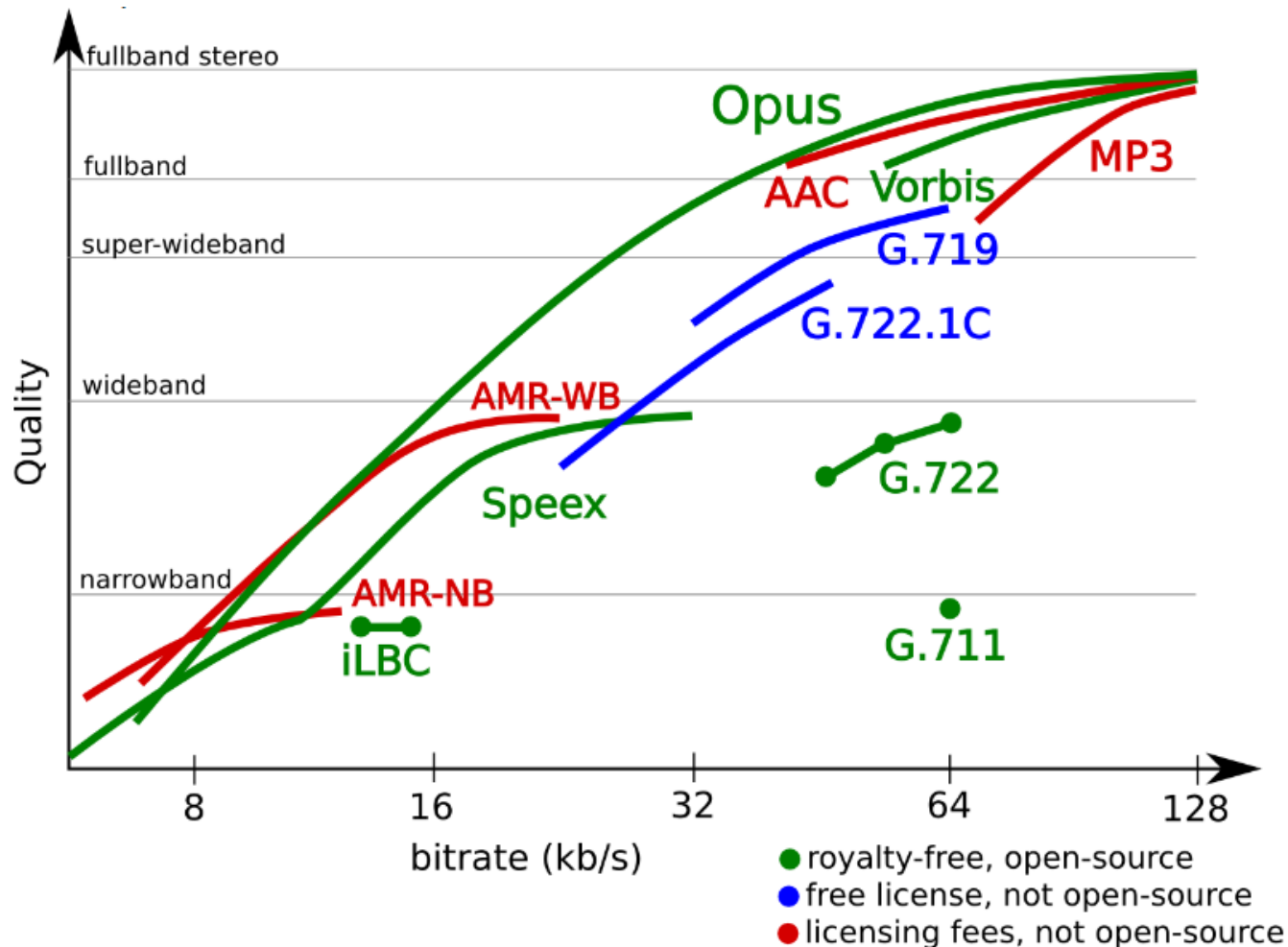
Xiph.Org Foundation

February 3, 2012

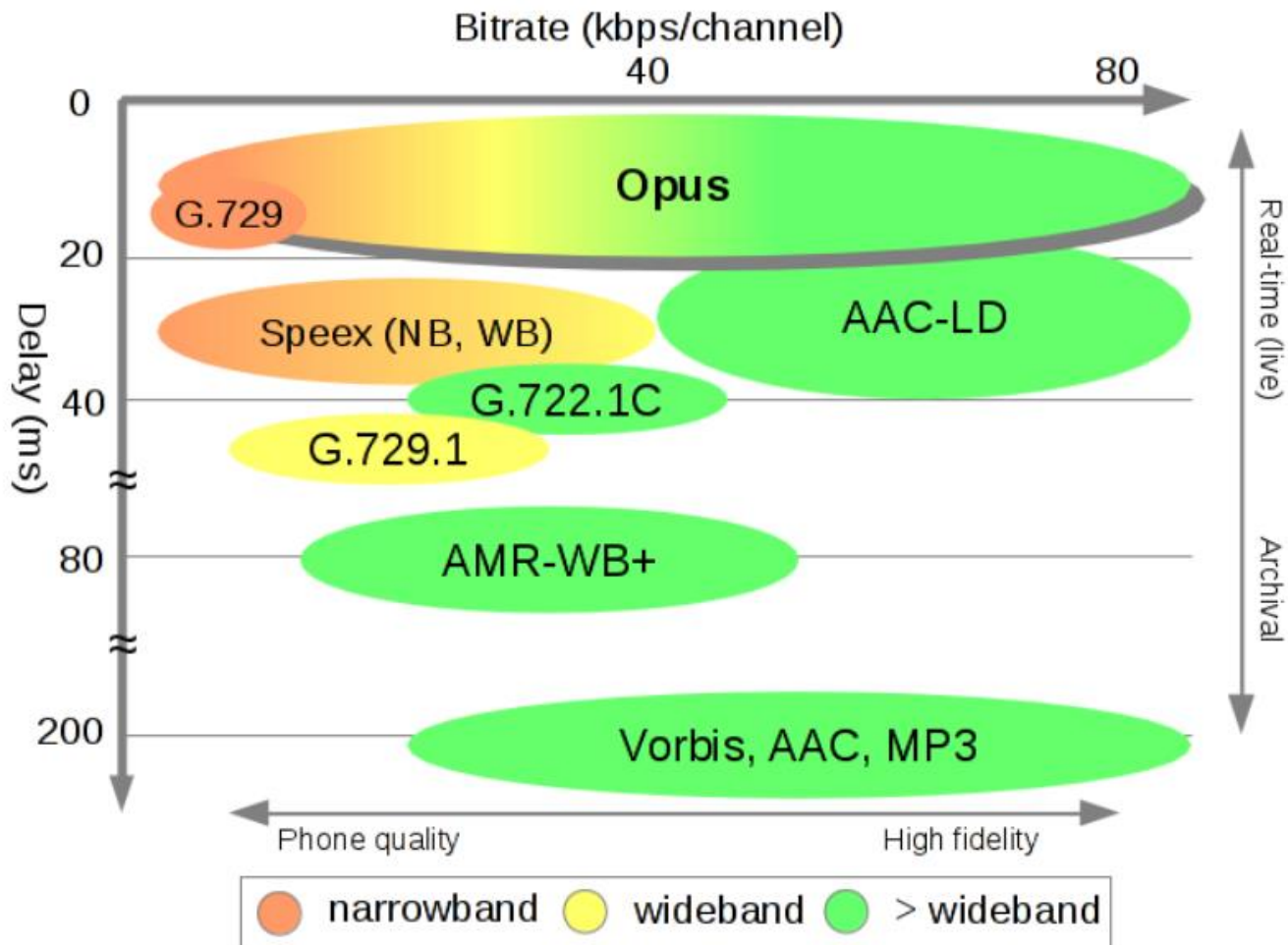
Hybridní Opus - 2011

- Mix technologií - Octasic, Skype, Xiph.Org and Broadcom.
- V hybridním Opusu je každý rámeček nejdříve kódován na LP (linear prediction) vrstvě (do 8 kHz) a pak na MDCT vrstvě (nad 8 kHz). Vrstva SILK je VBR (Variable bit rate) a CELT je CBR (Constant bit rate). SILK potřebuje prognózu na 5,2 ms, CELT na 2,5 ms. Nesoulad je řešen zpožděním CELT o 2,7 ms.
- Modifikovaná diskrétní kosinová transformace (MDCT – modified discrete cosine transform) je transformace blízká Fourierově. Je založená na typu IV diskrétní kosinové transformace (DCT-IV), Má delší časová okna a vlastnost překrývání: druhá polovina bloku se překrývá s první polovinou bloku dalšího. Tato finta vyhlazuje přechody mezi bloky, proto je použita ve ztrátových formátech typu MP3 (MPEG-1 Audio Layer 3) či Media Audio.

Kvalita versus bitová rychlost



Bitová rychlost kontra zpoždění



Opus standardizován v roce 2012

[\[Docs\]](#) [\[txt|pdf\]](#) [\[draft-ietf-codec-...\]](#) [\[Diff1\]](#) [\[Diff2\]](#) [\[IPR\]](#)

PROPOSED STANDARD

Internet Engineering Task Force (IETF)
Request for Comments: 6716
Category: Standards Track
ISSN: 2070-1721

JM. Valin
Mozilla Corporation
K. Vos
Skype Technologies S.A.
T. Terriberry
Mozilla Corporation
September 2012

Definition of the Opus Audio Codec

Abstract

This document defines the Opus interactive speech and audio codec. Opus is designed to handle a wide range of interactive audio applications, including Voice over IP, videoconferencing, in-game chat, and even live, distributed music performances. It scales from low bitrate narrowband speech at 6 kbit/s to very high quality stereo music at 510 kbit/s. Opus uses both Linear Prediction (LP) and the Modified Discrete Cosine Transform (MDCT) to achieve good compression of both speech and music.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in [Section 2 of RFC 5741](#).

Most mezi přenosem hudby a telefoníí

AAC Low Delay (AAC-LD) je jiný název pro MPEG-4 Low Delay Audio Coder, což je audio kompresní formát kombinující výhody audio kódování s výhodami dvousměrné komunikace s nízkým zpožděním. Tento formát je odvozen a formátu MPEG-2 (Advanced Audio Coding – AAC). Maximální zpoždění je 20 ms. Tento formát je schopen zajistit stereo signál v pásmu 7 Hz na jedné ISDN lince či v pásmu 15 kHz na dvou ISDN linkách.

Kvalitní firemní kodeky, např. L16

Document Number: TR-41.3.3/02-05-008
TR-41.4.4/02-05-017

STANDARDS PROJECT: PN-4705, PN-3-4462-URV

TITLE: L16-256 Wide Band Codec Definition

SOURCE: Polycom
1000 West 14th Street
North Vancouver, B.C. V7P 3P3
Canada

CONTACTS: Peter Baker
Phone: 604 697 9315, 604 990 5415 x105
Internet: peter.baker@polycom.com

DATE: May 1, 2002

DISTRIBUTION TO: TR-41.3.3, TR41.4.4

Cisco telefon SPA525G 3 vyžaduje výběr tří kodeků z těchto:
G711u, G711a, G726-32, G729a, G722 a L16.
Moderní L16 tam nechybí.

Pro vojenské aplikace: MELP (1997)

MELP (Mixed Excitation Linear Prediction) je určen pouze pro vojenské aplikace, satelitní komunikaci, utajená rádiové zařízení.

Parametry: bloky 22,5 ms zahrnují 180 vzorků, vzorkovací rychlost 8 000 kHz. Pokrývá pásmo 100 Hz - 3800 Hz.

V roce 2001 MELPe (Enhanced MELP) US standardem MIL-STD-3005, v roce 2002 NATO standardem STANAG-4591.

Používá periodické i neperiodické vzorkování a adaptivní filtr. Neperiodické vzorkování odstraňuje choulostivé přechodné úseky signálu mezi hlasem a dalšími zvuky.

Návrh zpracoval student Georgia Tech, dnes má pro komerční použití licenci Polycomm.

FLEXI-232 a FLEXI-422

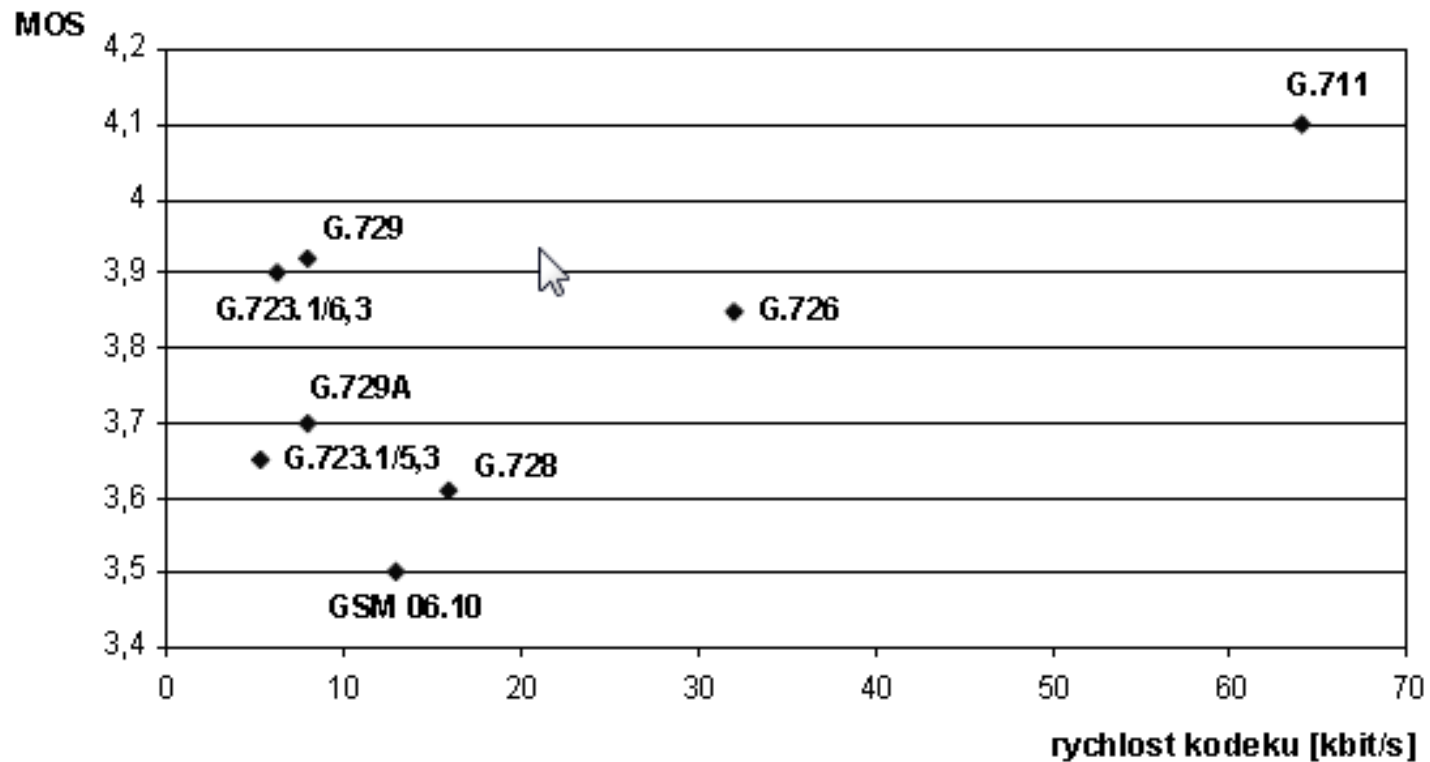


Karta Flexi-232 pro modemy



Karta v DTE doplněná PTT zařízením
(PTT – Push to Talk resp. Press-to-Transit, určeno pro HDX)

Porovnání MOS



Porovnání nároků na výkon

Kodek	Algoritmus	MIPS	CBR (Kbps)	NEB (Kbps)	MOS
G.711	PCM	---	64	87.2	4.1
G.726	ADPCM	1	32	55.2	3.85
G.728	LD-CELP	30	16	31.5	3.61
GSM 06.10	RPE-LP	10	13	---	3.5
G.729A	CS-ACELP	11	8	---	3,7
G.729	CS-ACELP	20	8	31.2	3.92
G.723.1	MP-MLQ	16	6.3	21.9	3.9
G.723.1	ACELP	20	5.3	20.8	3,65

Informace převzaty ze stránek fy. MATTES AD, spol. s r.o.

CBR – Codec Bit Rate, přenosová rychlost kodeku (v b/s)

NEB – Nominal Ethernet Bandwidth, nominální ethernetová šířka pásma (v b/s)

Který kodek je nejlepší?

Nelze jednoznačně určit, který kodek je ten nejlepší, protože některé se vyznačují nízkým bitovým tokem, jiné nízkou náročností na výpočet a samozřejmě různou kvalitou hovoru. Volba kodeku by proto měla vždy záležet na konfiguraci a možnostech nainstalovaného VoIP zařízení a síťové propustnosti.

Problémy s kodeky na PC

- Kodeky mohou být napsány pro 32bitové nebo 64bitové operační systémy. Pokud používáte 64bitovou verzi systému Windows, musíte nainstalovat kodeky pro 64bitovou verzi. Pokud například v 64bitovém operačním systému nainstalujete 32bitový kodek, přehrávač možná nebude moci přehrát žádné soubory, které tento kodek vyžadují.
- Mnoho starších kodeků je k dispozici pouze ve 32bitové verzi. Pokud poskytovatel kodeku neurčí, zda jde o 32bitovou nebo 64bitovou verzi kodeku, pravděpodobně se jedná o 32bitovou verzi.

Media Player

Zvukové kodeky

Typ	Název	Formát	Typ Binary	Verze
ACM	Kodek Microsoft IMA ADPCM	0011	imaadp32.acm	6.1.7600.16385
ACM	Kodek Microsoft CCITT G.711 A-Law and u-Law	0007	msg711.acm	6.1.7600.16385
ACM	Zvukový kodek Microsoft GSM 6.10	0031	msgsm32.acm	6.1.7600.16385
ACM	Kodek ADPCM	0002	msadp32.acm	6.1.7600.16385
ACM	Fraunhofer IIS MPEG Layer-3 Codec (decode only)	0055	l3codeca.acm	1.9.0.401
ACM	BANDI_MPEG-1 Audio Decoder	0050		
ACM	Převaděč PCM	0001		
DMO	WMAudio Decoder DMO	0160, 0161, 0162, 0163	WMADMOD.DLL	6.1.7601.17514
DMO	WMAPro over S/PDIF DMO	0162	WMADMOD.DLL	6.1.7601.17514
DMO	WMSpeech Decoder DMO	000A, 000B	WMSPDMOD.DLL	6.1.7601.17514
DMO	MP3 Decoder DMO	0055	mp3dmod.dll	6.1.7600.16385

Videokodeky

Typ	Název	Formát	Typ Binary	Verze
ICM	Microsoft RLE	MRLE	msrle32.dll	6.1.7601.17514
ICM	Microsoft Video 1	MSVC	msvidc32.dll	6.1.7601.17514
ICM	Microsoft YUV	UYVY	msyuv.dll	6.1.7601.17514
ICM	Kodek Intel IYUV	IYUV	iyuv_32.dll	6.1.7601.17514
ICM	Toshiba YUV Codec	Y411	tsbyuv.dll	6.1.7601.17514
ICM	Kodek Cinepak společnosti Radius	cvid	iccvid.dll	1.10.0.13
ICM	Bandi MJPEG Video Decoder	MJPG		
ICM	Bandi MPEG-1 Video Decoder	MPEG		
DMO	Mpeg4s Decoder DMO	mp4s, MP4S, m4s2, M4S2, MP4V, mp4v, XVID, xvid, DIVX, DX50	mp4sdecdec.dll	6.1.7600.16385
DMO	WMV Screen decoder DMO	MSS1, MSS2	wmvsdecdec.dll	6.1.7601.17514
DMO	WMVideo Decoder DMO	WMV1, WMV2, WMV3, WMVA, WVC1, WMVP, WVP2	wmvdecod.dll	6.1.7601.17514
DMO	Mpeg43 Decoder DMO	mp43, MP43	mp43decdec.dll	6.1.7600.16385
DMO	Mpeg4 Decoder DMO	MPG4, mpg4, mp42, MP42	mpg4decdec.dll	6.1.7600.16385

Kodeky, které podporuje Asterisk

Asterisk Codecs

Asterisk supports the following narrow-band and wideband (HD audio) codecs:

- [G.711](#) ulaw (as used in US)
- [G.711](#) alaw (as used in Europe)
- [G.722](#) - 16 kHz wideband codec; passthrough, playback and recording in Asterisk 1.4; full support incl. transcoding in Asterisk 1.6, a [backport for 1.4](#) is available, or use this [more up-to-date patch](#)
- [G.723.1](#) - [pass-thru](#) for people who need a [license](#), [free](#) for other people
- [G.726](#) - 32kbps only (16/24/32/40kbps supported in `format_g726` for files); flawed until Asterisk 1.4 which corrected the implementation and introduced codec **g726aal2** and setting **g726nonstandard** for backwards compatibility with Asterisk 1.2 installations
- [G.729](#) - may require a [license](#) unless using [pass-thru](#), [free](#) version available for use in countries without patents or for educational use only
- [GSM](#)
- [iLBC](#)
- LPC10 (not recommended!)
- [Speex](#) - configurable 4-48kbps, VBR, ABR, etc. [see bug 2536](#). For Asterisk 1.4. there is [patch 10519](#) available that adds wideband support for the OpenWengo software client

Dnes preferované kodeky u Cisco

Audio Codec Preference Order for Cisco Unified Communications Manager 8.0(1)

If Low Loss Is Configured for Link Loss Type	If Lossy Is Configured for Link Loss Type
AAC-LD—256 kb/s	AAC-LD—256 kb/s
L16—256 kb/s	L16—256 kb/s
G.722 64k—64 kb/s	iSAC—32 kb/s
iSAC—32 kb/s	G.722 64k—64 kb/s
G.722.1 32k—32 kb/s	G.722.1 32k—32 kb/s
G.722 —56 kb/s	G.722 —56 kb/s
G.722.1—24 kb/s	G.722.1—24 kb/s
G.722—48 kb/s	G.722—48 kb/s
G.711 mu-law 64 k—64 kb/s	G.711 mu-law 64 k—64 kb/s
G.711 A-law 64k—64 kb/s	G.711 A-law 64k—64 kb/s
G.711 mu-law 56k—56 kb/s	G.711 mu-law 56k—56 kb/s
G.711 A-law 56k—56kb/s	G.711 A-law 56k—56kb/s
iLBC—16 kb/s	iLBC—16 kb/s
G.728—16 kb/s	G.728—16 kb/s
GSM Enhanced Full Rate—13 kb/s	GSM Enhanced Full Rate—13 kb/s
GSM Full Rate—13 kb/s	GSM Full Rate—13 kb/s
G.729b—8 kb/s	G.729b—8 kb/s
G.729ab—8 kb/s	G.729ab—8 kb/s
G.729—8 kb/s	G.729—8 kb/s
G.729a—8 kb/s	G.729a—8 kb/s
GSM Half Rate—7 kb/s	GSM Half Rate—7 kb/s
G.723.1—7 kb/s	G.723.1—7 kb/s

Dříve preferované kodeky u Cisco

Table 2-4 Audio Codec Preference Order for previous versions of Cisco Unified Communications Manager

If Low Loss Is Configured for Link Loss Type	If Lossy Is Configured for Link Loss Type
---	iLBC—16 kb/s
AAC-LD—256 kb/s	AAC-LD—256 kb/s
L16—256 kb/s	L16—256 kb/s
G.722.1 24k—24 kb/s	G.722.1 24k—24 kb/s
G.722.1 32k—32 kb/s	G.722.1 32k—32 kb/s
G.722 64k—64 kb/s	G.722 64k—64 kb/s
G.711 mu-law 64k—64 kb/s	G.711 mu-law 64k—64 kb/s
G.711 A-law 64k—64 kb/s	G.711 A-law 64k—64 kb/s
G.722 56k—56 kb/s	G.722 56k—56 kb/s
G.711 mu-law 56k—56 kb/s	G.711 mu-law 56k—56 kb/s
G.711 A-law 56k—56 kb/s	G.711 A-law 56k—56 kb/s
G.722 48k—48 kb/s	G.722 48k—48 kb/s
iLBC—16 kb/s	---
G.728—16 kb/s	G.728—16 kb/s
GSM Enhanced Full Rate—13 kb/s	GSM Enhanced Full Rate—13 kb/s
GSM Full Rate—13 kb/s	GSM Full Rate—13 kb/s
G.729b—8 kb/s	G.729b—8 kb/s
G.729ab—8kb/s	G.729ab—8kb/s
G.729—8 kb/s	G.729—8 kb/s
G.729a—8 kb/s	G.729a—8 kb/s
GSM Half Rate—7 kb/s	GSM Half Rate—7 kb/s
G.723.1—7 kb/s	G.723.1—7 kb/s

Nastavení preferencí kodeků na Cisco směrovači

```
voice class codec 100
codec preference 1 g729r8
codec preference 2 g723r63
codec preference 3 g711alaw
!
```

Dohoda o kodeku u SIP

(odchyceno Wiresharkem)

```
INVITE sip:novak@sip.domainB.cz SIP/2.0
Via: SIP/2.0/UDP 195.122.198.236:5065;rport;branch=z9hG4bK1441F37B74154C09BD150D68AD8B83F7
From: markl <sip:markl@sip.domainA.cz:5065>;tag=1603324369
To: <sip:novak@sip.domainB.cz>
Contact: <sip:markl@195.122.198.236:5065>
Call-ID: 650ADF5C-0EB4-499A-9744-2B2561B30C94@192.168.2.111
CSeq: 9126 INVITE
Max-Forwards: 70
Content-Type: application/sdp
User-Agent: SIPphone Lite release 1104v
Content-Length: 321
v=0
o=markl 212548077 212548116 IN IP4 195.122.198.236
s=SIPphone Lite
c=IN IP4 195.122.198.236
t=0 0
m=audio 8000 RTP/AVP 0 8 3 98 97 101
a=rtpmap:0 pcmu/8000
a=rtpmap:8 pcma/8000
a=rtpmap:3 gsm/8000
a=rtpmap:98 iLBC/8000
a=rtpmap:97 speex/8000
a=rtpmap:101 telephone-event/8000
a=fmtp:101 0-15
a=sendrecv
```

Výsledků testů MOS

Test Set 1								
Ideal Conditions: No Network Load with Artificial Voice								
	G.711u	G.711a	G.726 32kbps	G.729a	G.723.1	iLBC	GSM AMR 12.2kbps	Speex
MOS AVG (PSQM)	4.45	4.44	4.3	4.04	4.08	4.1	4.14	3.84
Test Set 2								
Ideal Conditions: No Network Load with Both Gender Voice								
	G.711u	G.711a	G.726 32kbps	G.729a	G.723.1	iLBC	GSM AMR 12.2kbps	Speex
MOS AVG (PSQM)	4.3	4.29	4.21	3.76	3.74	3.81	3.96	3.78
Test Set 3								
Normal Conditions: 1% Packet Loss with Both Gender Voice								
	G.711u	G.711a	G.726 32kbps	G.729a	G.723.1	iLBC	GSM AMR 12.2kbps	Speex
MOS AVG (PSQM)	4.24	4.25	4.14	3.74	3.75	3.82	3.94	3.79
Test Set 4								
Duress: 5% Packet Loss with Both Gender Voice								
	G.711u	G.711a	G.726 32kbps	G.729a	G.723.1	iLBC	GSM AMR 12.2kbps	Speex
MOS AVG (PSQM)	4.13	4.11	3.79	3.51	3.57	3.74	3.79	3.59

Čipy OPL-2, OPL-3

(Operator Type-L)

OPL-2 – zvukový čip Yamaha YM3812

OPL-3 – zvukový čip Yamaha YMF262
(populární ve zvukové kartě Soundblaster 16)

OPLL – zvukový čip Yamaha YM2413

244 registrů pouze pro zápis, 9 kanálů zvuku, každý produkovaný dvěma oscilátory.
Frekvence (10 bitů) oktáva (3 bity), tvar vlny (celá, půlka, čtvrtka...), tremolo, vibráto, nosná...

Moduly složené z procesorů

(Packet Voice Digital Signal Processor Modules)

PVDM2-8	8-Channel Packet Fax/Voice DSP Module	1 DSP (TI 2505)
PVDM2-16	16-Channel Packet Fax/Voice DSP Module	1 DSP (TI 2510)
PVDM2-32	32-Channel Packet Fax/Voice DSP Module	2 DSPs (TI 2510)
PVDM2-48	48-Channel Packet Fax/Voice DSP Module	3 DSPs (TI 2510)
PVDM2-64	64-Channel Packet Fax/Voice DSP Module	4 DSPs (TI2510)

Kodeky s vysokou komplexností:

G.723.1, G.728, G.729, G.729b, GSM-EFR and Modem Relay

Kodeky se střední komplexností:

G.711, G.729a, G.729ab, G.726, GSM-FR and Fax Relay

Výpis přehledu DSP

HQ-1#show voice dsp

```
DSP  DSP                DSPWARE CURR  BOOT                PAK    TX/RX
TYPE NUM CH CODEC      VERSION STATE STATE  RST AI VOICEPORT TS ABORT  PACK COUNT
=====
```

```
-----FLEX VOICE CARD 0 -----
          *DSP VOICE CHANNELS*
```

```
CURR STATE : (busy)inuse (b-out)busy out (bpend)busyout pending
LEGEND      : (bad)bad   (shut)shutdown (dpend)download pending
```

```
DSP  DSP                DSPWARE CURR  BOOT                PAK    TX/RX
TYPE NUM CH CODEC      VERSION STATE STATE  RST AI VOICEPORT TS ABRT PACK COUNT
=====
```

```
          *DSP SIGNALING CHANNELS*
```

```
DSP  DSP                DSPWARE CURR  BOOT                PAK    TX/RX
TYPE NUM CH CODEC      VERSION STATE STATE  RST AI VOICEPORT TS ABRT PACK COUNT
=====
```

C5510	002	01	{flex}	8.2.0	alloc	idle	0	0	0/2/0	02	0	0/0
C5510	002	02	{flex}	8.2.0	alloc	idle	0	0	0/2/1	02	0	0/0

```
-----END OF FLEX VOICE CARD 0 -----
```

Vytváření DSP profilů

```
R1(config)#voice-card 0
R1(config-voicecard)#dsp services dspfarm
R1(config)#dspfarm profile 1 transcode
R1(config-dspfarm-profile)#codec g711ulaw
R1(config-dspfarm-profile)#codec g711alaw
R1(config-dspfarm-profile)#codec g729ar8
R1(config-dspfarm-profile)#codec g729abr8
R1(config-dspfarm-profile)#codec g729r8
R1(config-dspfarm-profile)#maximum sessions 6
R1(config-dspfarm-profile)#associate application
SCCP
R1(config-dspfarm-profile)#no shutdown
```

```
R2(config)#voice-card 0
R2(config-voicecard)#dsp services dspfarm
R2(config)#dspfarm profile 1 conference
R2(config-dspfarm-profile)#codec g711ulaw
R2(config-dspfarm-profile)#codec g711alaw
R2(config-dspfarm-profile)#codec g729ar8
R2(config-dspfarm-profile)#codec g729abr8
R2(config-dspfarm-profile)#codec g729br8
R2(config-dspfarm-profile)#maximum sessions 2
R2(config-dspfarm-profile)#associate application
SCCP
R2(config-dspfarm-profile)#no shutdown
```

Vztah velikosti paketu a rychlosti přenosu paketů

	Kodek a paketizační interval			
	<u>G.711</u> 20 ms	<u>G.711</u> 30 ms	<u>G.729</u> 20 ms	<u>G.729</u> 40 ms
Pásmo [kb/s]	64	64	8	8
Paketizační interval [byte]	160	240	20	40
Záhlaví protokolů IP, UDP a RTP [byte]	40	40	40	40
Velikost VoIP části paketu [byte]	200	280	60	80
Rychlost [pakety/s]	50	33,33	50	25

Závislost zpoždění způsobené kódováním na velikosti vzorku

ITU-T Standard	Coding Scheme	Bit Rate	Sample Size (Bits)	Encoding Delay (Time)	Mean Opinion Score (MOS)
G.711	PCM	64K CBR	8 bits	<1 ms	4.4
G.726	ADPCM	32/24/16K CBR	4/3/2 bits	1 ms	4.2, 3.8, 3.2
G.728	LC-CELP	16K VBR	40 bits	2 ms	4.2
G.729	CS-ACELP	8K VBR	80 bits	15 ms	4.2
G.723.1	ACELP	5.3K VBR	160 bits	37.5 ms	3.5

Požadavky na pásmo pro jednotlivé kodeky

Bandwidth Consumption	802.1Q Ethernet	PPP	MLP	Frame-Relay	ATM
G.711 at 50pps	93kbps	84kbps	86kbps	84kbps	106kbps
G.711 at 33pps	83kbps	77kbps	78kbps	77kbps	84kbps
G.729A at 50pps	37kbps	28kbps	30kbps	28kbps	43kbps
G.729A at 33pps	27kbps	21kbps	22kbps	21kbps	28kbps

Závislost požadavků na pásmo na paketizačním intervalu

Voice Type	over Link Type	Full/Half Duplex	Link Speed (kbps)	10 ms Packetization, VAD off			10 ms Packetization, VAD on		
				G.711	G.729	G.723	G.711	G.729	G.723
				Bandwidth demand (bps)	Bandwidth demand (bps)	Bandwidth demand (bps)	Bandwidth demand (bps)	Bandwidth demand (bps)	Bandwidth demand (bps)
Digital	TDM	Full	n/a	64,000	8,000	6,300	64,000	8,000	6,300
VoIP	802.3	Half	10,000	252,800	140,800	n/a	176,960	98,560	n/a
VoIP	802.3	Half	100,000	252,800	140,800	n/a	176,960	98,560	n/a
VoIP	802.3	Full	100,000	126,400	70,400	n/a	88,480	49,280	n/a
VoIP	Frame Relay	Full	any	100,800	44,800	n/a	70,560	31,360	n/a
VoIP	PPP	Full	any	102,400	46,400	n/a	71,680	32,480	n/a
VoIP	ATM (AAL-5)	Full	any	127,200	84,800	n/a	89,040	59,360	n/a
VoIP/hc	Frame Relay	Full	any	72,800	16,800	n/a	50,960	11,760	n/a
VoIP/hc	PPP	Full	any	74,400	18,400	n/a	52,080	12,880	n/a
VoFR	Frame Relay	Full	any	71,200	15,200	n/a	49,840	10,640	n/a
VcATM	ATM (AAL-5)	Full	any	84,800	42,400	n/a	59,360	29,680	n/a
Voice Type	over Link Type	Full/Half Duplex	Link Speed (kbps)	30 ms Packetization, VAD off			30 ms Packetization, VAD on		
				G.711	G.729	G.723	G.711	G.729	G.723
				Bandwidth demand (bps)	Bandwidth demand (bps)	Bandwidth demand (bps)	Bandwidth demand (bps)	Bandwidth demand (bps)	Bandwidth demand (bps)
Digital	TDM	Full	n/a	64,000	8,000	6,300	64,000	8,000	6,300
VoIP	802.3	Half	10,000	169,600	57,600	54,400	118,720	40,320	38,080
VoIP	802.3	Half	100,000	169,600	57,600	54,400	118,720	40,320	38,080
VoIP	802.3	Full	100,000	84,800	28,800	27,200	59,360	20,160	19,040
VoIP	Frame Relay	Full	any	76,267	20,267	18,667	53,387	14,187	13,067
VoIP	PPP	Full	any	76,800	20,800	19,200	53,760	14,560	13,440
VoIP	ATM (AAL-5)	Full	any	84,800	28,267	28,267	59,360	19,787	19,787
VoIP/hc	Frame Relay	Full	any	66,933	10,933	9,333	46,853	7,653	6,533
VoIP/hc	PPP	Full	any	67,467	11,467	9,867	47,227	8,027	6,907
VoFR	Frame Relay	Full	any	66,400	10,400	8,800	46,480	7,280	6,160
VcATM	ATM (AAL-5)	Full	any	84,800	14,133	14,133	59,360	9,893	9,893

Typický příklad použití v SLA

85 % volání MOS \geq 4,2

95 % volání MOS \geq 4,0

4 % volání MOS $<$ 3,8

Trendy vývoje v kodecích

- Vylepšování kvality, vývoj směrem k širšímu pásmu:
Narrowband (-3.4 kHz) --> Wideband (-7 kHz) --> Superwideband (-14 kHz) --> Fullband (-20 kHz)
- Zvyšování odstupu od vnějšího šumu (auto, letadlo, vrtulník...)
- Multimódové kodeky určené pro řeč i hudbu
- Kolaborace ve standardizaci, dříve, speciálně pro ITU platilo pravidlo „vítěz bere vše“.
- Řada proprietárních řešení (Skype...)
- Open Source, volný software
- Co bude následovat?

Domácí úkol

Na http://www.vocal.com/audio_examples/itu_coders.html

ITU Speech Coder Audio Examples

Speech coder audio examples are provided for each ITU speech coder.

- Reference PCM sample file

- G.711 A-law encoded

- G.711 μ -law encoded

- G.723.1 5.3k bps

- G.723.1 6.4k bps

- G.726 16k bps

- G.726 24k bps

- G.726 32k bps

- G.726 40k bps

- G.729 8k bps

- G.729A 8k bps
