

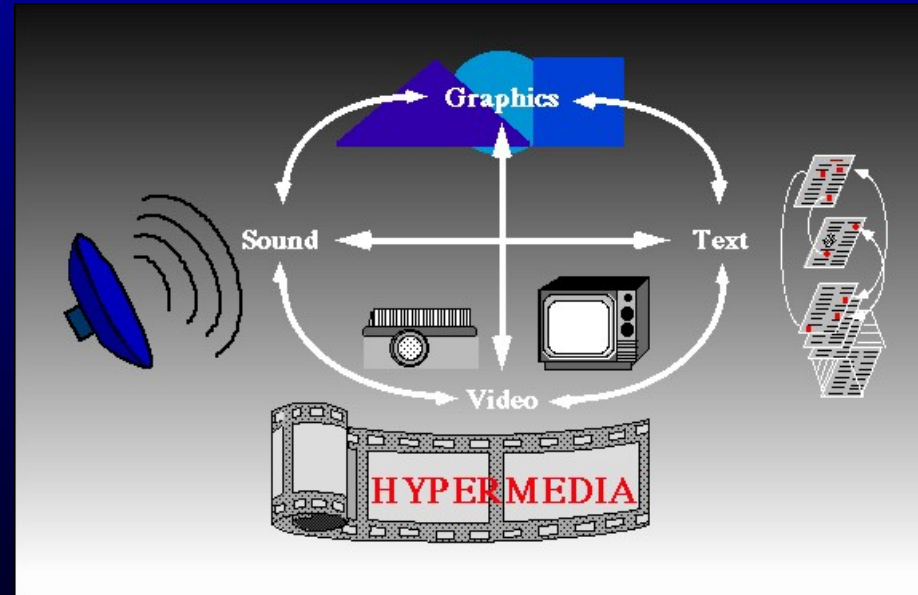
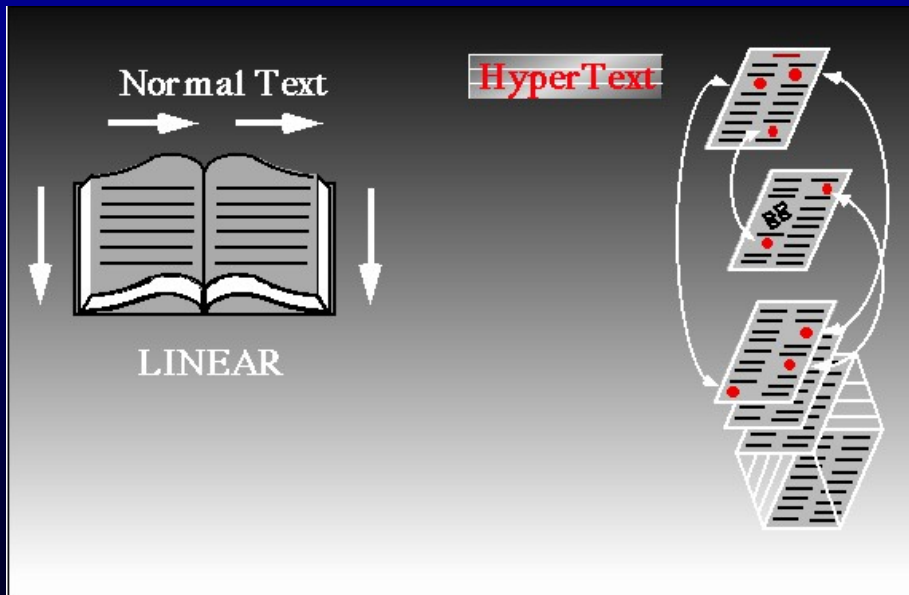
Multimediální technologie v sítích

eva@fi.muni.cz
hopet@ics.muni.cz

- Poznámka k barvě:
chcete-li v PowerPointu ČB verzi, jděte do menu View a zvolte položku Black and White

Přehled

- multimédia: přináší další způsoby interakce s počítačem – zvuk a video
- hypermédia – nelineární práce s médii – Ted Nelson (1965)



Interakce multimédií a sítí

- kódování do vhodného formátu a jeho dekódování u příjemce
- ošetření chyb přenosu
- parametry Quality of Service (QoS)
 - šířka pásma, zpoždění (latence, RTT), jitter
- synchronizace mezi médii
- klient – server vs. peer-to-peer (P2P) aplikace
- multicast

Aplikace

- WWW
- multimedialní produkty, databáze, hry
- videokonference
- streamování (aka video-on-demand, webcasting či netcasting), interaktivní TV
- virtuální realita
- zpracování videa, distribuované střížny

- ukázky

- WebMuseum, Paris

- <http://www.ibiblio.org/wm/>

- BBC, CNN

- <http://www.bbc.co.uk/>, <http://www.cnn.com>

- Index of Multimedia Information Sources

- <http://viswiz.gmd.de/MultimediaInfo/>

- archívy konferencí Internet2, Canarie

- <http://www.internet2.edu/>, <http://www.canarie.ca/>

- ResearchChannel

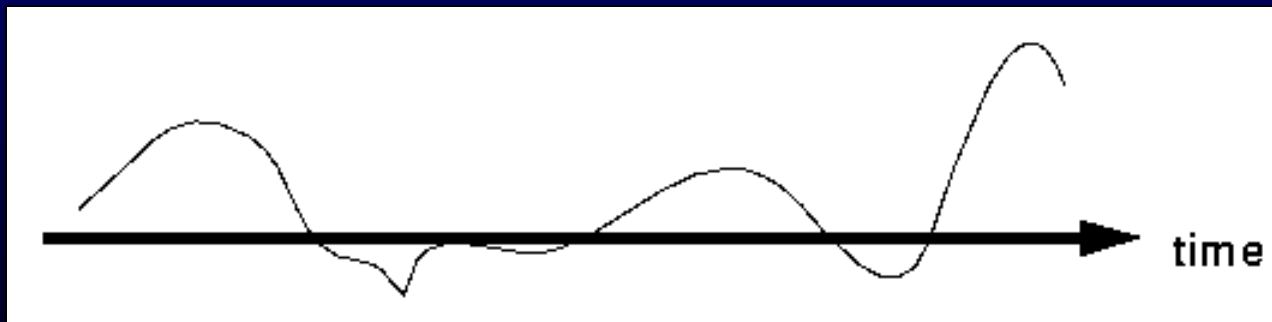
- <http://www.researchchannel.org/>

Audio

- nahrávání/přehrávání
- syntéza zvuku (příp. řeči)
- efekty nad zvukem
- posílání zvuku v digitální formě sítí

Audio – digitalizace

- zvuk
 - spojité tlakové vlny putující vzduchem
 - běžné fyzikální jevy vln: odraz, ohyb, difrakce
- mikrofon – převod zvuku na napěťový analogový signál



- PCM (Pulse Code Modulation)

- *vzorkování* – rozdělení časové osy

- většina lidí nedokáže rozlišit zvuky nad 20 kHz
 - vzorkovací frekvence musí být nejméně dvojnásobek frekvence zaznamenávané (Nyquistův teorém)
 - řeč: 11.025 kHz (telefon: 8 kHz)
zvuk nižší kvality: 22.05 kHz
CD: 44.1 kHz (frekvence >22.01 kHz se odfiltrují před zpracováním)
DAT: 48 kHz
 - vysoce kvalitní záznam: 96 kHz

– *kvantování (quantization)* – rozdělení svislé osy (intenzita – tlak/napětí)

- 1 bit odpovídá přibližně 6 dB (2x hlučnejší)
- CBR vs. VBR
- telefon: 8 bitů (256 úrovní, mono)
CD: 16 bitů (65536 úrovní, stereo)
- je-li V_{max} maximální pozitivní a negativní amplituda signálu a n je počet použitých bitů, pak kvantizační interval q je definován:

$$q = \frac{V_{max}}{2^n}$$

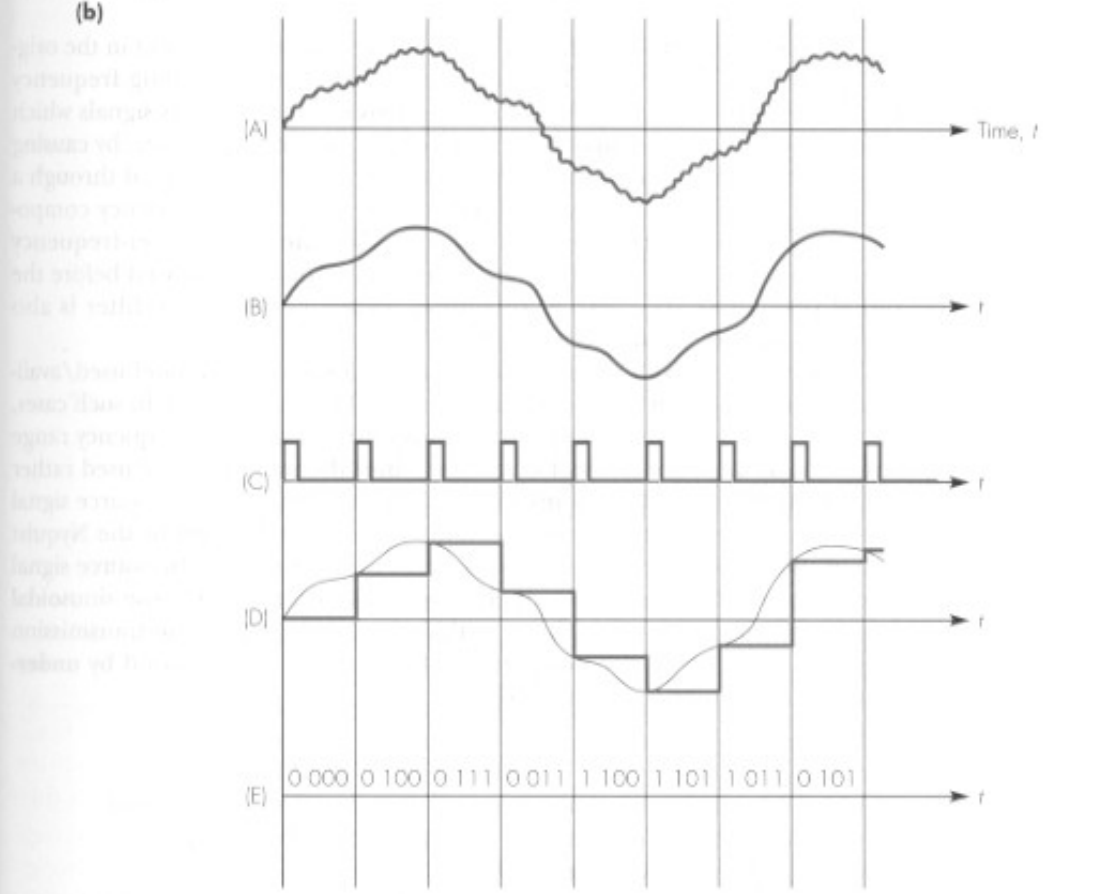
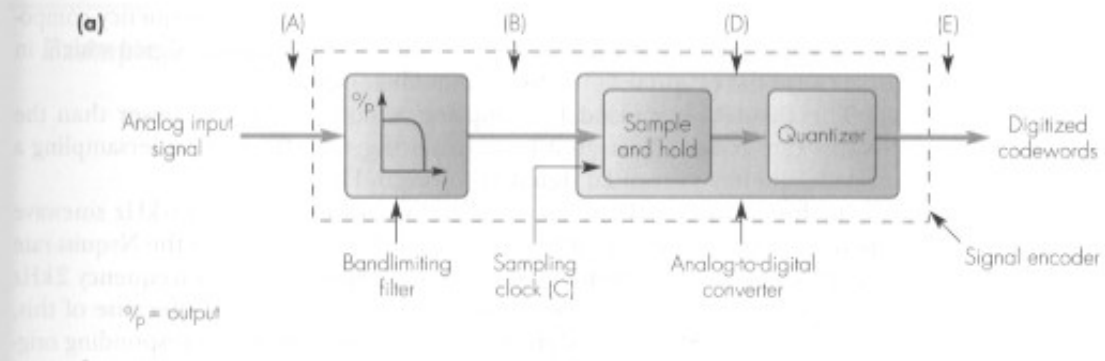
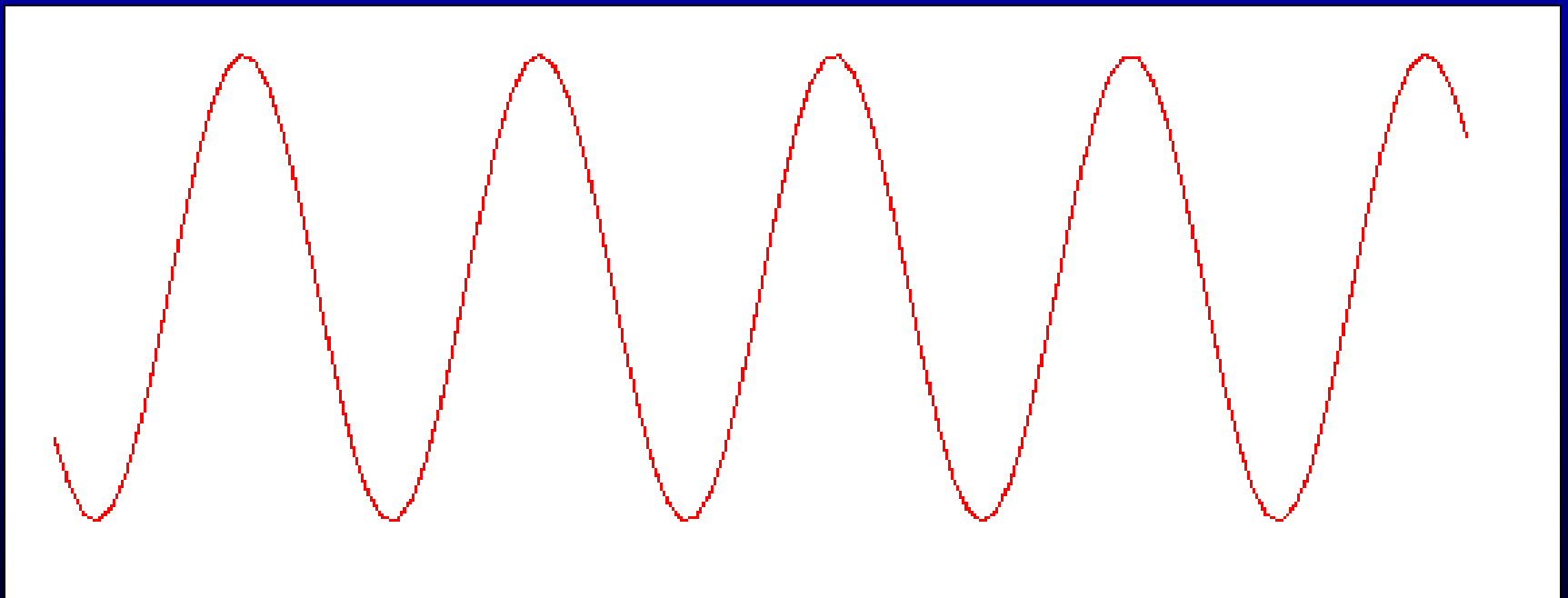
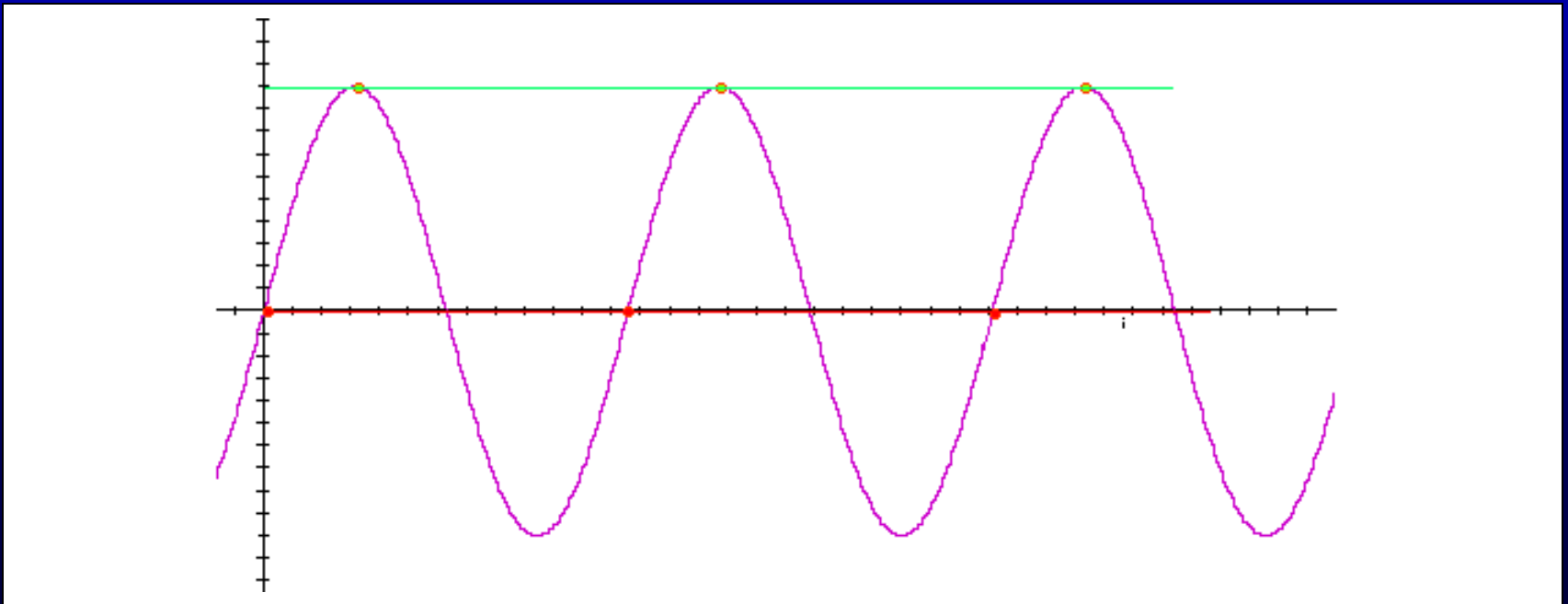


Figure 2.2 Signal encoder design: (a) circuit components; (b) associated waveform set.

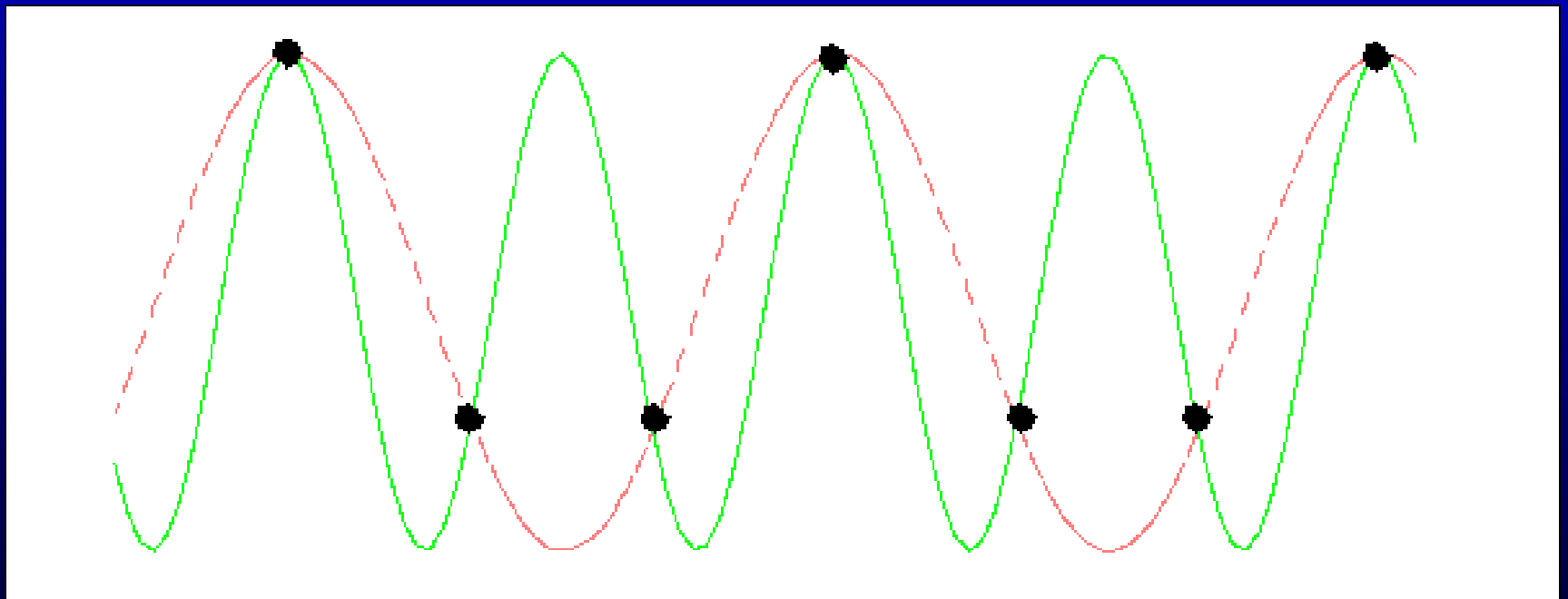
- odbočka k Nyquistovu teorému
– původní signál



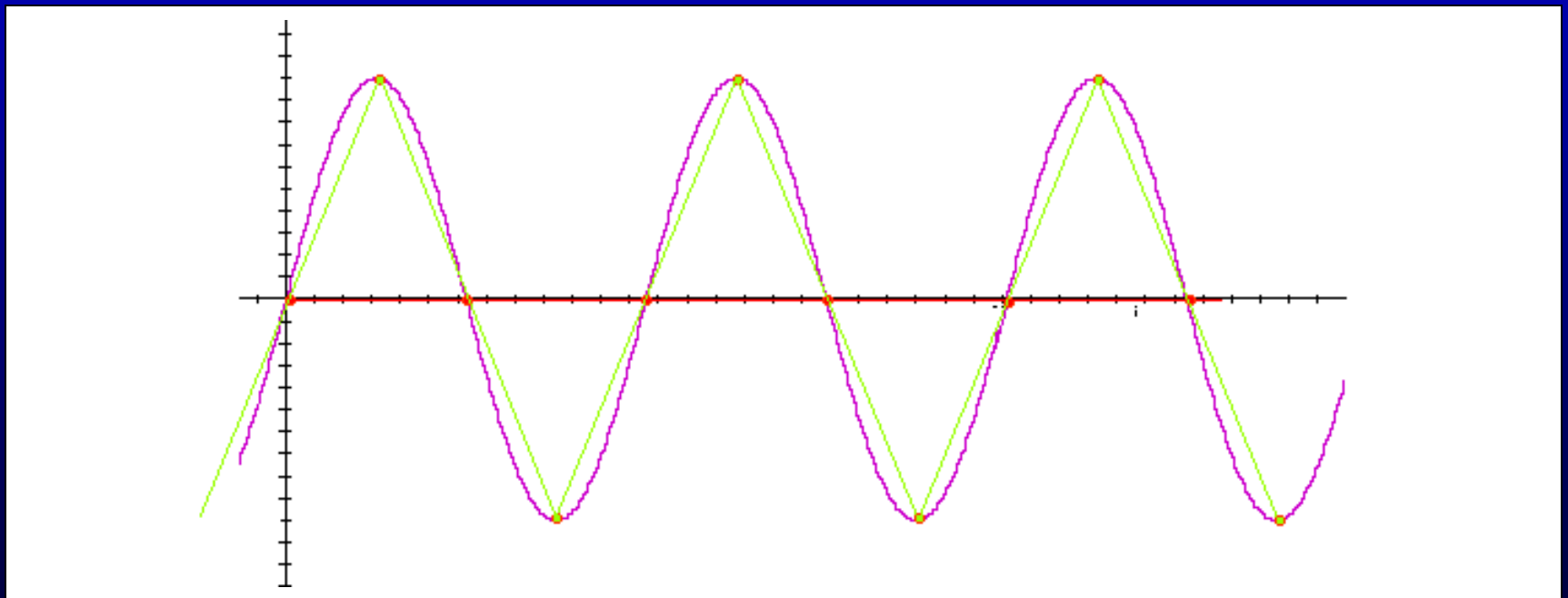
– vzorkování na stejné frekvenci



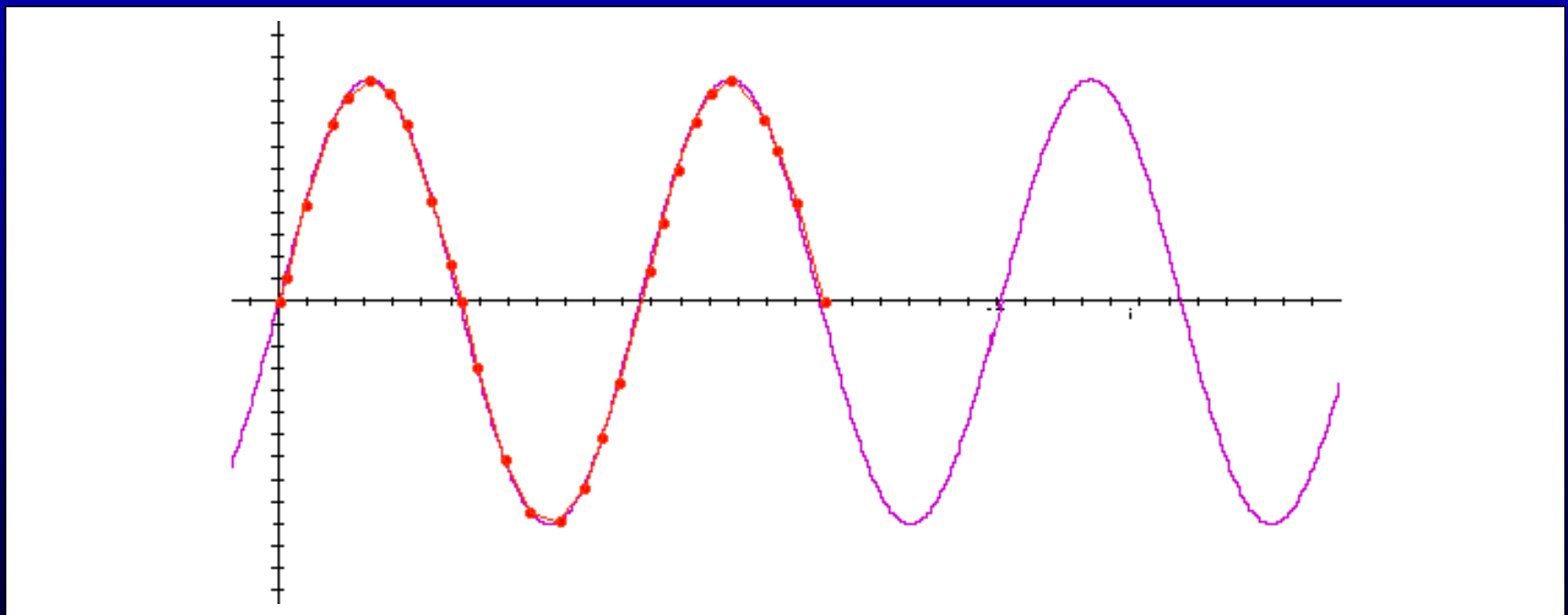
– vzorkování na 1.5-násobné frekvenci



– samplování na 2-násobné frekvenci



– vzorkování na ještě vyšší frekvenci



- podvzorkování (undersampling)
 - podvzorkovaná frekvence se „schová“ za jinou frekvenci

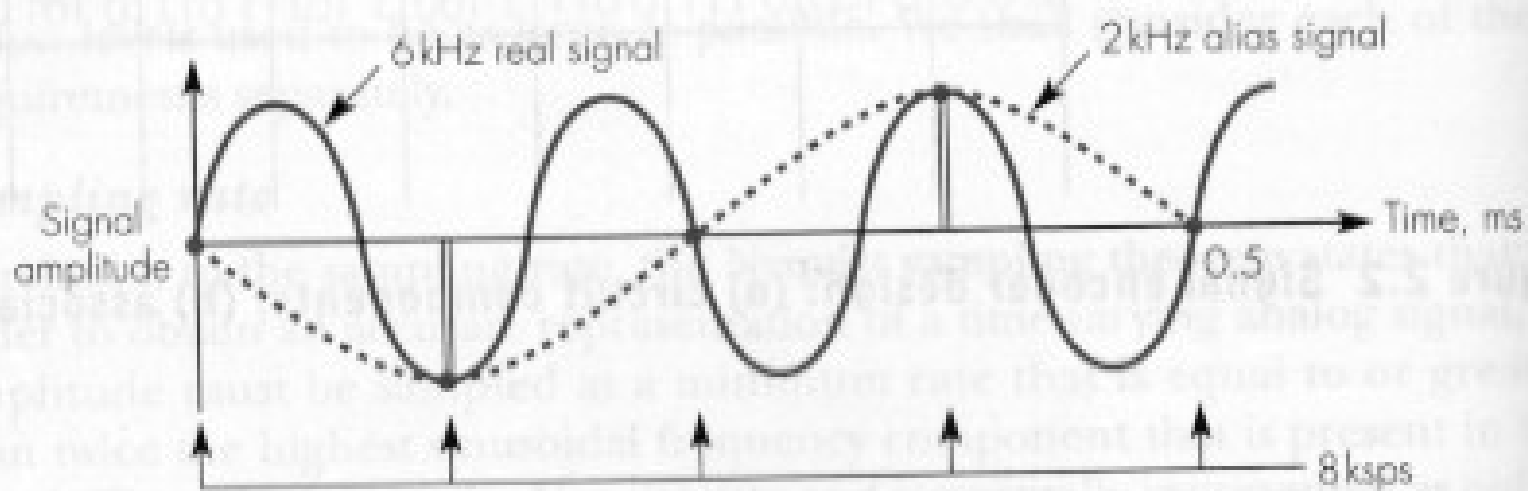


Figure 2.3 Alias signal generation due to undersampling.

- chyby v kvantování
 - chyba v záznamu může být $q/2$
 - několik úrovní intenzity se reprezentuje jednou hodnotou => kvantovací šum
 - vylepšit lze
 - ditheringem
 - nelineárním kvantováním
 - ucho je nelineární člen soustavy zvuk – mozek ;-)
(zhruba logaritmický)
 - zhuštění úrovní v oblasti, kde je to třeba; u zvuku je třeba mít jemněji rozlišenou oblast s nízkými kmitočty, protože ucho je v této oblasti citlivější na šum

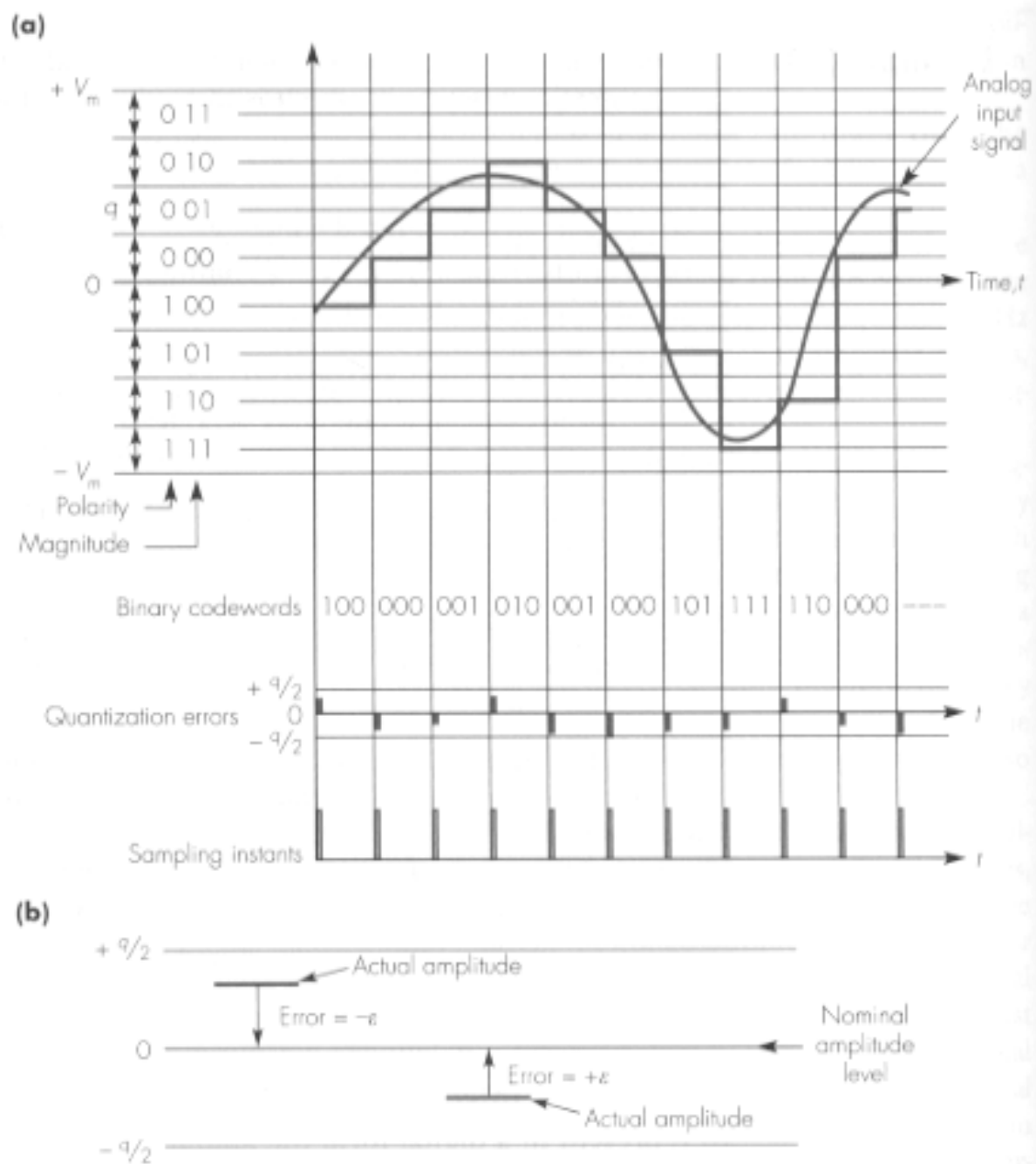


Figure 2.4 Quantization procedure: (a) source of errors; (b) noise polarity.

- šum
 - signal / noise ratio

$$SNR = 10 \log \frac{V_{signal}^2}{V_{noise}^2} = 20 \log \frac{V_{signal}}{V_{noise}}$$

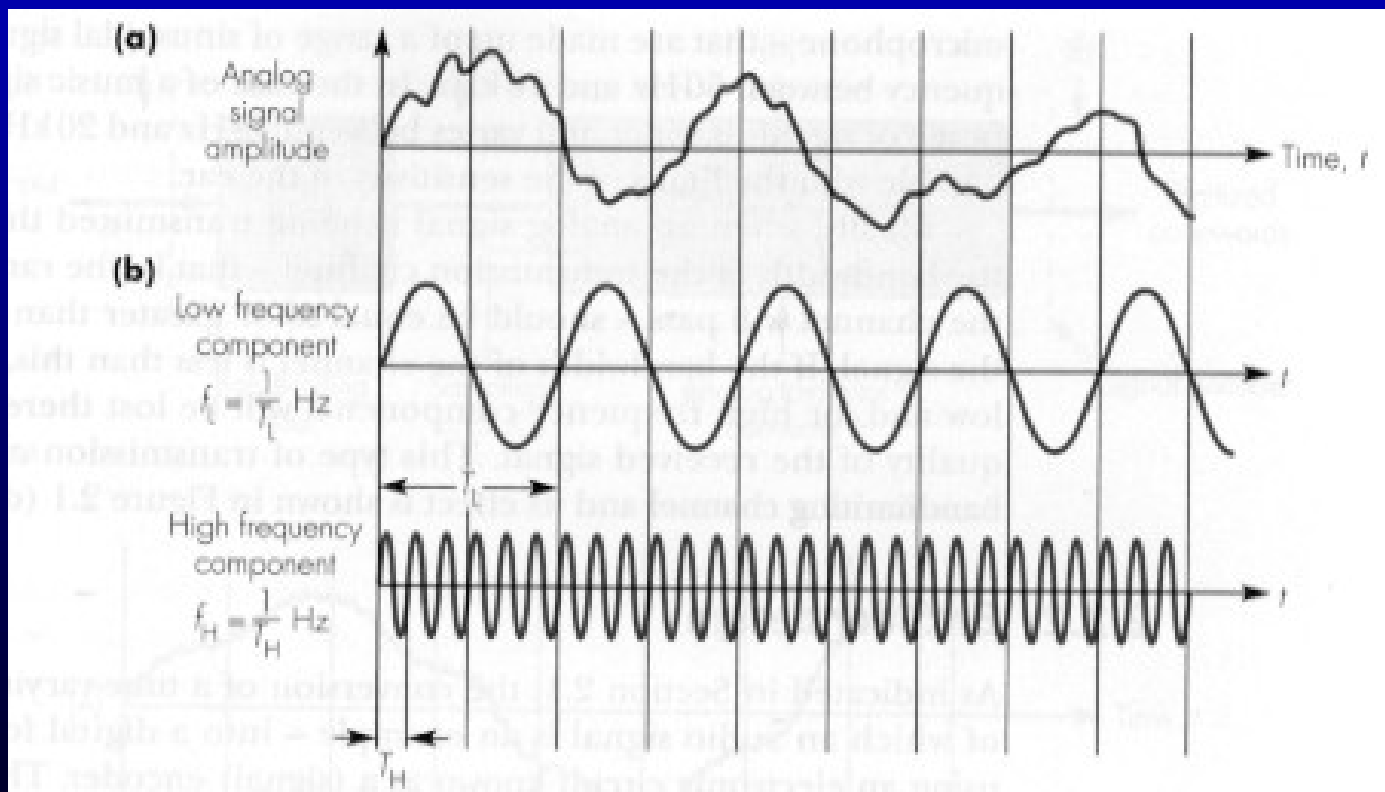
- 8 bitů – SNR 50 dB
- 16 bitů – SNR 98 dB

- příklady

Quality	Sample Rate (KHz)	Bits per Sample	Mono/ Stereo	Data Rate (Uncompressed)	Frequency Band
Telephone	8	8	Mono	8 KBytes/sec	200-3,400 Hz
AM Radio	11.025	8	Mono	11.0 KBytes/sec	
FM Radio	22.050	16	Stereo	88.2 KBytes/sec	
CD	44.1	16	Stereo	176.4 KBytes/sec	20-20,000 Hz
DAT	48	16	Stereo	192.0 KBytes/sec	20-20,000 Hz

Zpracování signálu – FT

- Fourierova analýza



- Fourierova transformace

$$x(t) = \sum_{k=-\infty}^{\infty} x(t) \left(\cos\left(\frac{2\pi k t}{T}\right) + \sin\left(\frac{2\pi k t}{T}\right) \right)$$

- na počítači nemůžeme dělat nekonečné součty => diskrétní FT
- můžeme předstírat, že $x(t)$ je periodické se stejnou délkou, jako je délka záznamu
- základní frekvence = 1/délka záznamu
- window function – minimalizace nespojitosti fází na rozhraních segmentů

Zpracování signálu – DCT

- DCT používá reálná čísla
- obrázek o NxN pixelech:
- DCT je pole koeficientů:
kde:

$$DCT_{uv} = \frac{1}{2N} C_u C_v \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p_{xy} \cos \left[\frac{(2x+1)u}{2N} \pi \right] \cos \left[\frac{(2y+1)v}{2N} \pi \right]$$


kde:

$$C_u C_v = \begin{cases} 1 & \text{for } u, v = 0 \\ \sqrt{2} & \text{otherwise} \end{cases}$$

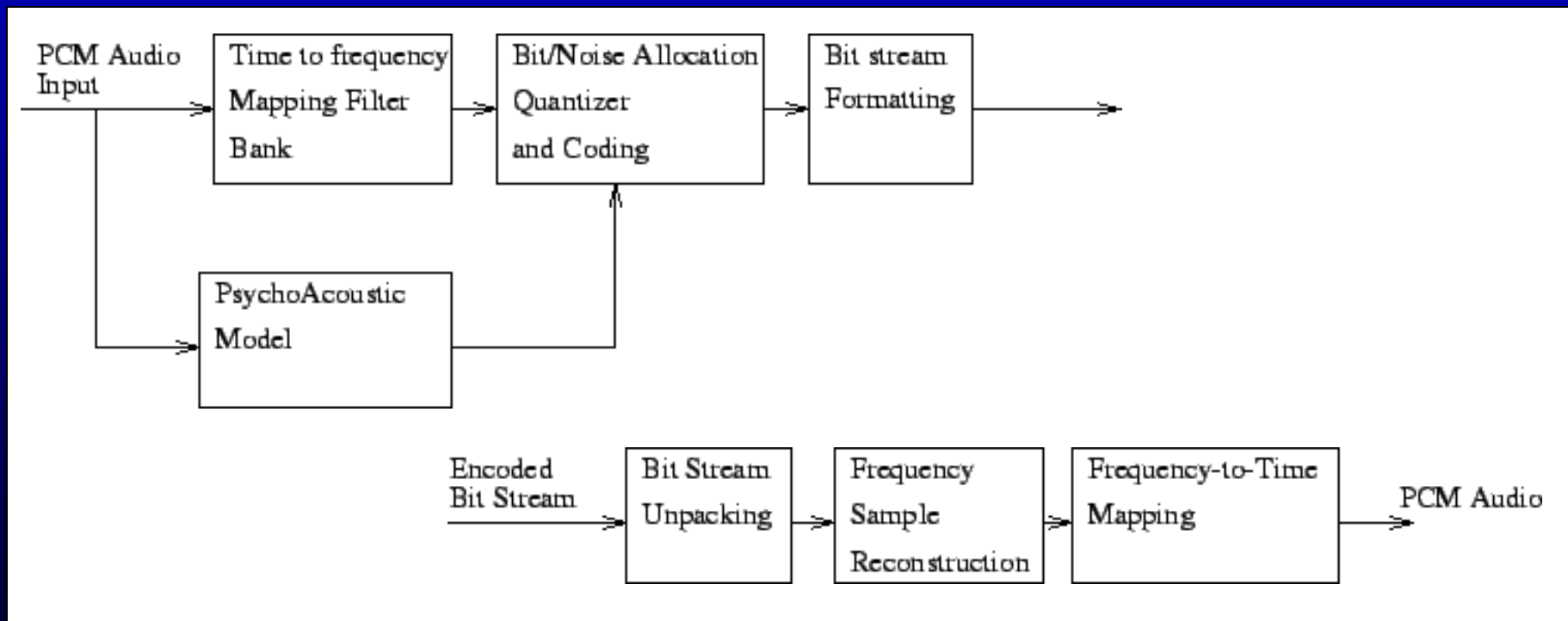
Poznámky k FT a DCT

- složitost:
 - DFT ... $O(n^2)$
 - FFT ... $O(n \cdot \log_2 n)$
 - DCT ... $O(n^4)$ resp. $O(n^2)$
- odkazy:
 - <http://www.dataq.com/applicat/articles/an11.htm>
 - <http://www.chipcenter.com/eexpert/bmasta/bmasta001.html>
 - http://bellatrix.isr.umd.edu/TechReports/ISR/1996/PhD_96-9/PhD_96-9.phtml

Formáty audia

- PCM
 - telefon: $1 \cdot 8000 \cdot 8 = 64$ kbps
 - CD kvalita: $2 \cdot 44100 \cdot 16 = 1.41$ Mbps
- DPCM (Differential Pulse Code Modulation)
 - ztrátová komprese, používá 4 bity na záznam rozdílů nezávisle na původním kvantování
- ADPCM (Adaptive Differential Pulse Code Modulation)
 - snaží se o prediktivní analýzu potřebného počtu bitů v závislosti na složitosti signálu
- aw (A-law v Evropě)
 - podobné ADPCM, použití v telefonech (kvantování pak odpovídá spíše 13 bitům než 8)

- LPC (Linear Predictive Coding), (Code Excited Linear Predictor) a GSM
 - optimalizace na hlas
- MP3

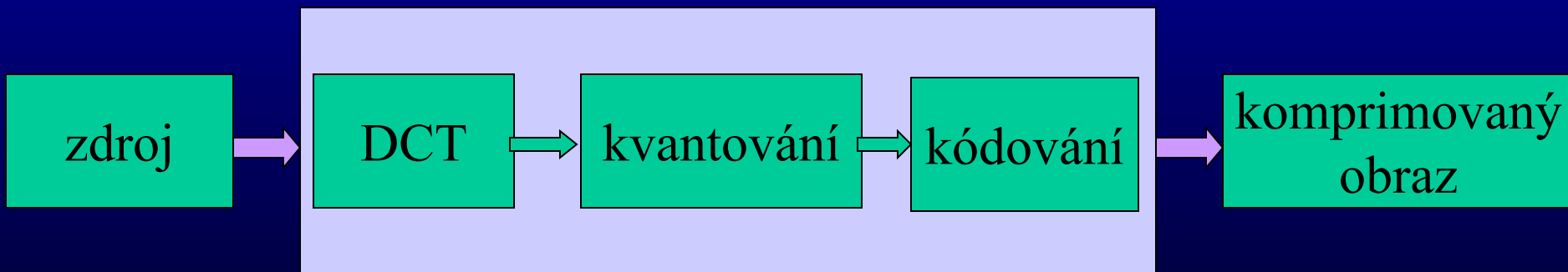


MP3

- frekvenční pásma – subbands
trik MP3 – výběr správných pásem
psychoakustický model
- DCT -> kvantování -> Huffmanovo
kódování
- Layer I – 384 kbps – 4:1
Layer II – 192 kbps – 8:1
Layer III – 128 kbps – 12:1

Video – kódování, komprese

- komprese obrázků
 - Lempel-Ziv, GIF (8 bitový index barev, slovníková LZ komprese; 10:1 až 2:1)
 - JPEG



- JPEG
 - DCT – bezeztrátová
 - DC, AC koeficienty
 - kvantování – tabulky kvantovacích koeficientů
 - kódování
 - RLE pro nulové koeficienty
 - Huffman pro nenulové
 - delta kódování pro DC koeficienty
 - barva (3 vrstvy)
 - možnost redukce 24 bitů na 8 bitů

Jak funguje MPEG

- intraframe vs. interframe komprese
- 3 typy snímků (frames)
 - I frames – intrapicture, referenční snímek
 - P frames – predicted, rozdíl oproti předchozímu snímku
 - B frames – bidirectional predicted, interpolace mezi předcházejícím a následujícím

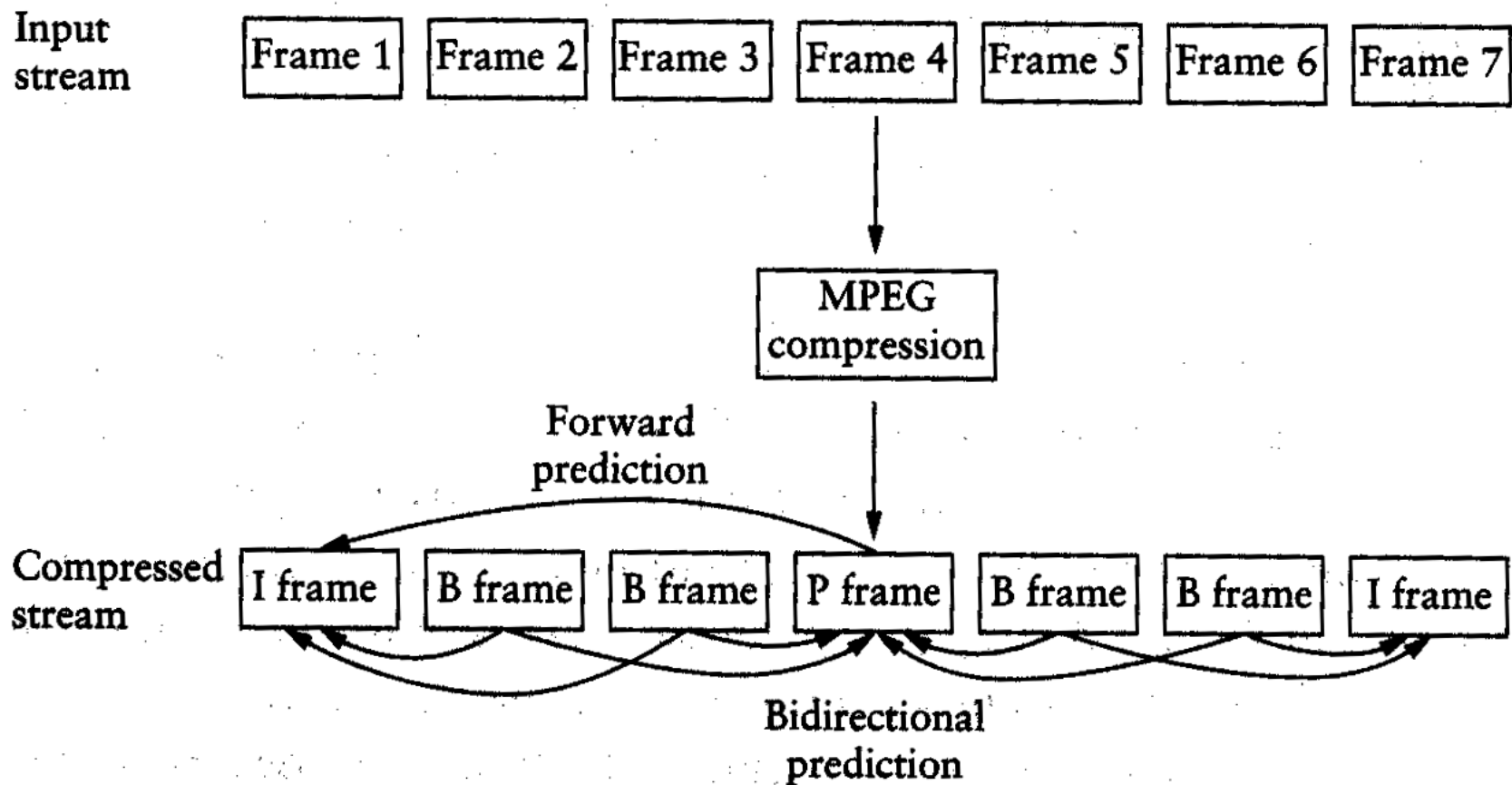
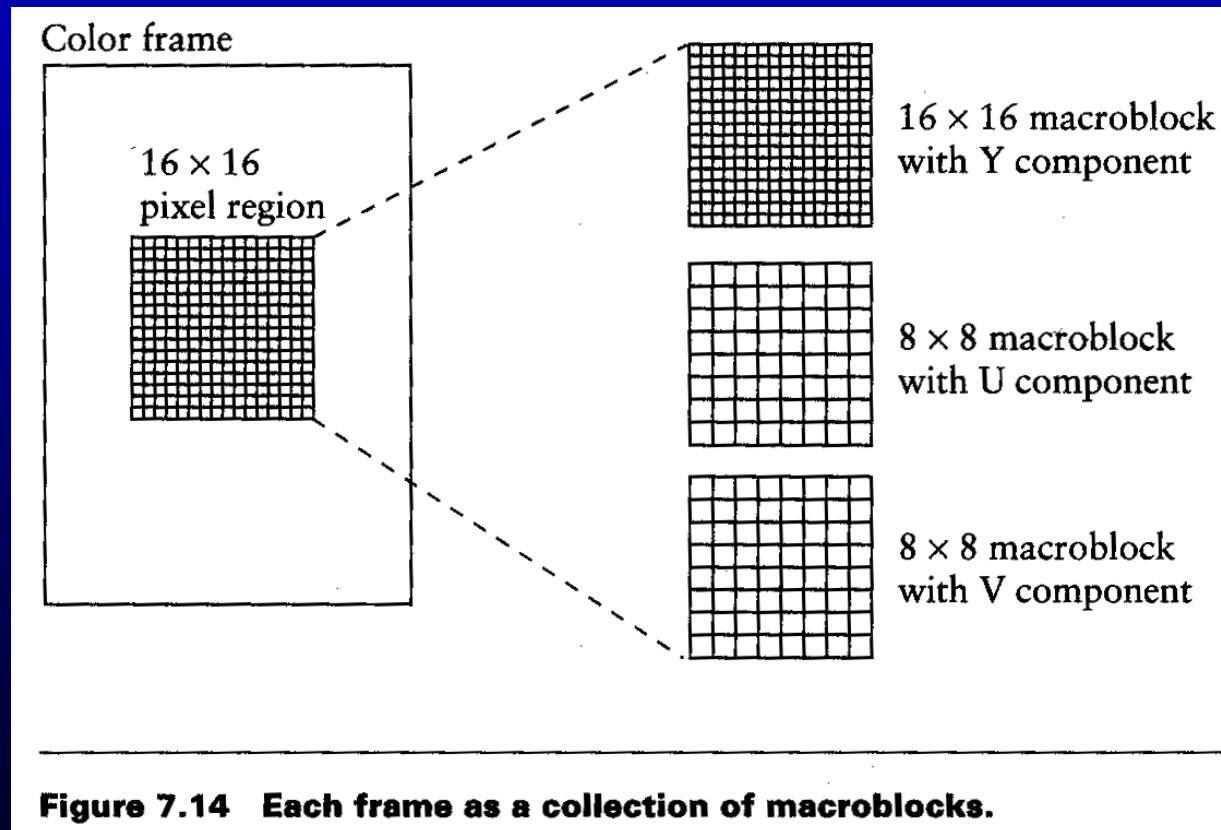


Figure 7.13 Sequence of I, P, and B frames generated by MPEG.

MPEG - komentář

- pro přenos po síti je důležité:
 - snímek P nemá smysl bez snímku I
 - snímek B nemá smysl bez snímků I a P
 - v případě snímků B se snímky neposílají v sekvenčním pořadí
- MPEG nedefinuje poměr I/P/B snímků
 - pouze I snímky ~ MJPEG

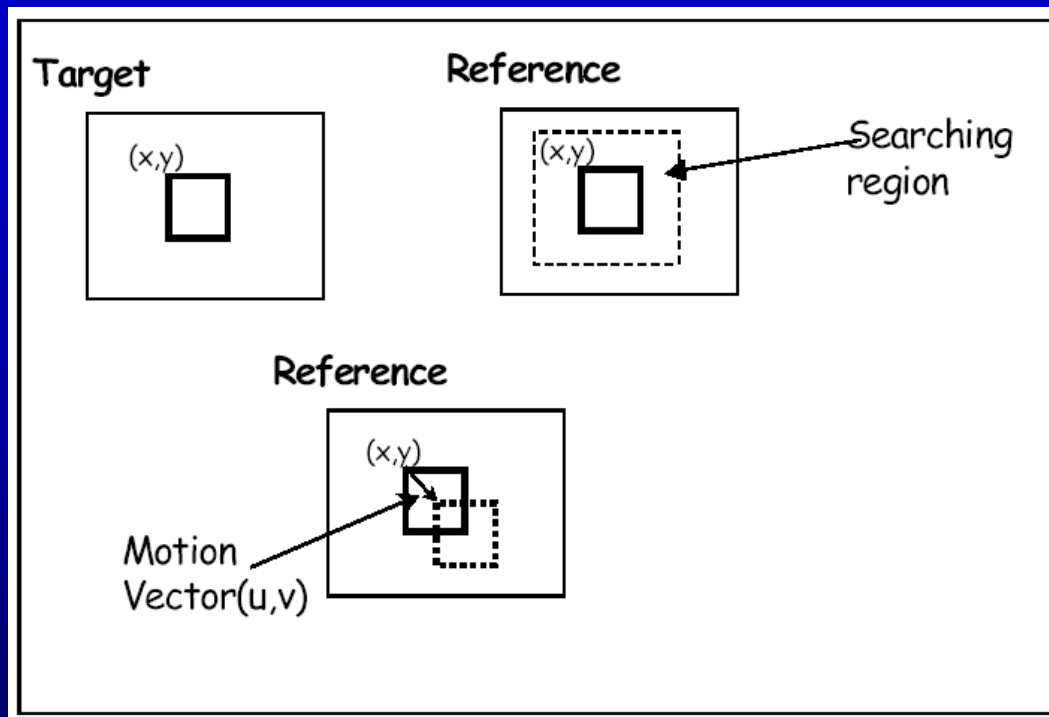
- snímky I – makrobloky 16x16



- snímky B a P jsou zpracovány také po makroblocích
- snímky B
 - výjimečně může být makroblok stejný jako v případě I snímku (v případě velké změny)
 - čtveřice
 - souřadnice makrobloku ve snímku
 - motion vector vzhledem k předchozímu referenčnímu snímku
 - motion vector vzhledem k následujícímu referenčnímu snímku
 - ■ pro každý pixel – změna pixelu vzhledem k referenčním snímkům (komprese DCT)

$$F_c(x,y) = (F_p(x+x_p,y+y_p) + F_f(x+x_f,y+y_f))/2 + \blacksquare(x,y)$$

- motion vector & motion estimation



- MPEG v reálném čase? S podporou HW ano, bez ní to trochu bolí...

MPEG formáty

- MPEG-1
 - rok 1993
 - 1 – 1.5 Mbps
 - náhodný přístup – 0.5 s
- MPEG-2
 - rok 1995
 - digital TV, DV, DVD
 - až 100 Mbps
 - od 4 Mbps výrazně lepší kvalita než MPEG-1

– MPEG-2 main profile

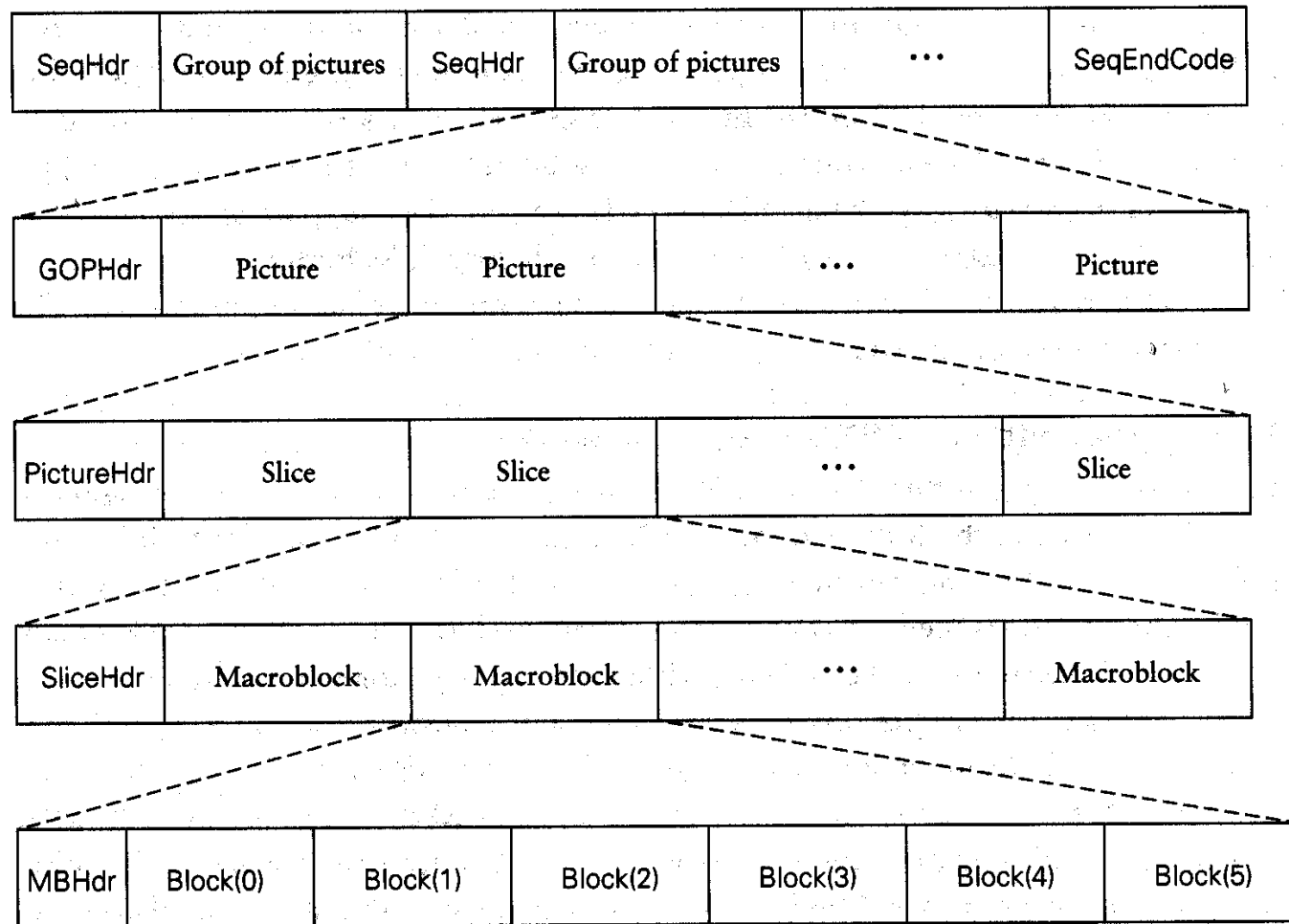


Figure 7.15 Format of an MPEG-compressed video stream.

- CBR vs. VBR (změny BR pro jednotlivé GOP – např. změnou kvantování)
- co je optimální BR v případě multicastu??
=> odpověď: vrstvení dat
- problémy s retransmisí v TCP – nepřijatelně velká latence v případě výpadku
- přenos po UDP – má smysl pečlivě volit rozhraní paketů, aby ztráta paketu „moc nebolela“ => Application Level Framing
- problémy se ztrátami I snímků => diffserv
- problém s latencí:
IBBBBBPBBBBBI
člověk vnímá hranici 100 ms, ale 5x 67ms (15 fps)
= 335 ms!

- kombinace s audiem – interleaving
- optimalizace v případě ucpání sítě
 - vynechání B snímků a přenos audia místo nich, protože člověk je mnohem citlivější na nekvalitní audio než na nekvalitní video

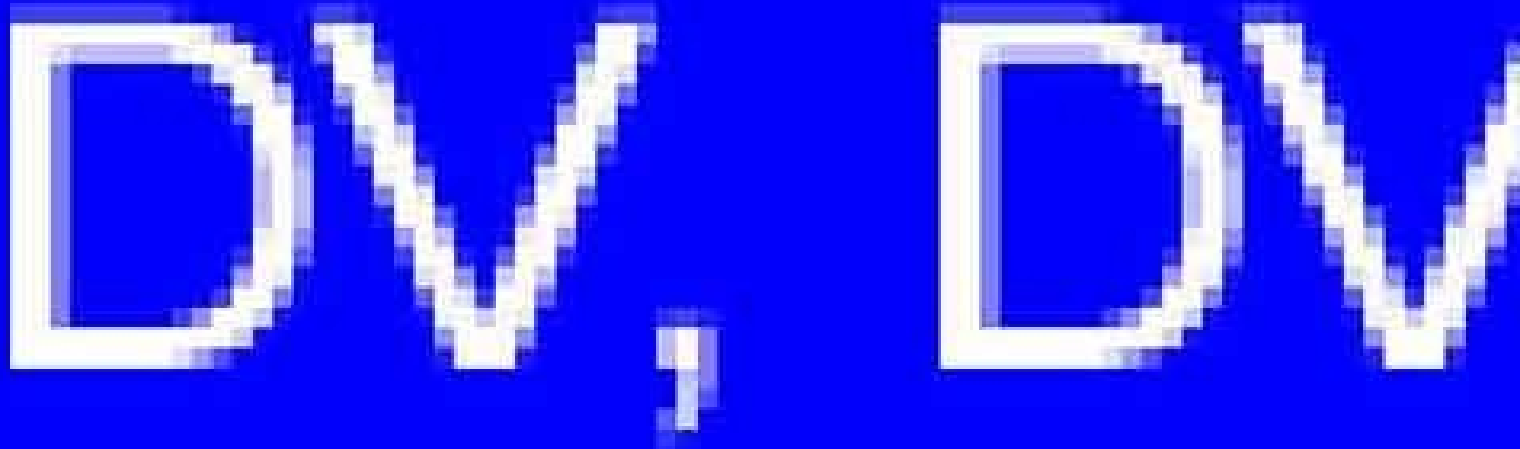
- MPEG-4

- určeno pro malou šířku pásma 10 kbps až 1 Mbps
- důležitý bezchybný přenos
- organizace do objektů, k nimž lze přistupovat samostatně
- MS (Streaming) Media, RealMedia, DivX

DV

- MPEG-2
- 25 Mbps (DVCPRO-50 50 Mbps)
- kompresní poměr 5:1
- intraframe a interfield komprese
 - frames (snímky) vs. fields (půlsnímky)

DV artefakty - fathering



DV, DVCAM, DVCPRO

DV artefakty - quilting



DV artefakty - motion blocking



H.261 a H.263

- H.261
 - vznikl primárně pro konference nad ISDN
 - variabilní šířka pásma
 - 40 Kbits/s až 2 Mbits/s (typicky 64 či 128 kbps)
 - obdoba P snímků

Picture Formats Supported

Picture format	Luminance pixels	Luminance lines	H.261 support	Uncompressed bitrate (Mbit/s)			
				10 frames/s		30 frames/s	
				Grey	Colour	Grey	Colour
QCIF	176	144	Yes	2.0	3.0	6.1	9.1
CIF	352	288	Optional	8.1	12.2	24.3	36.5

- H.263

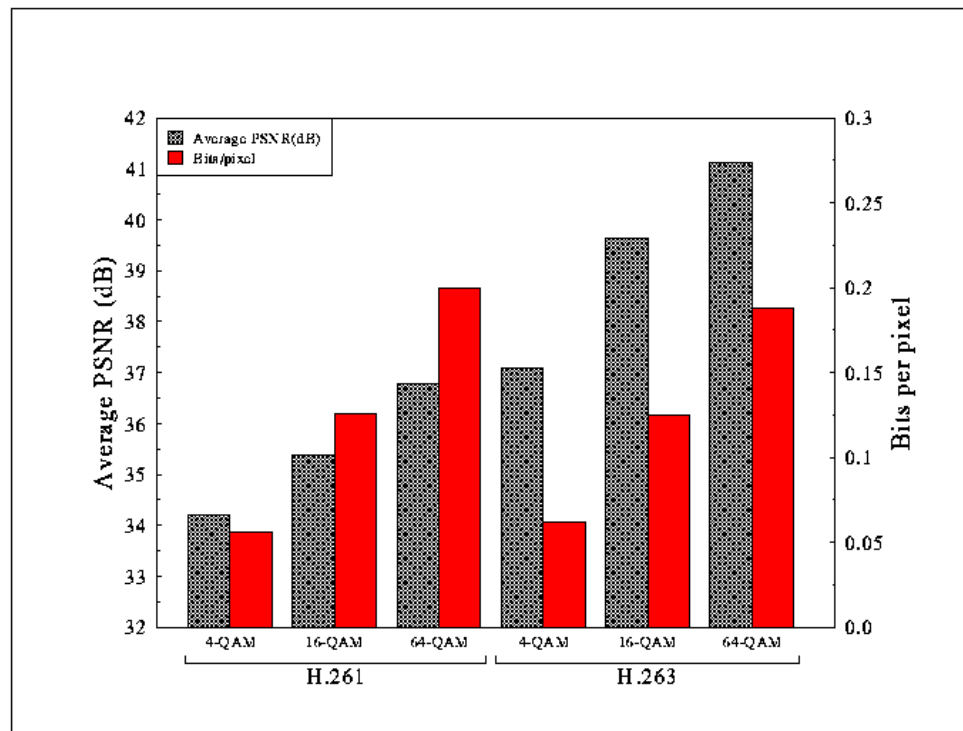
- lepší kompenzace pohybu (půlpixelová přesnost)
- hierarchická struktura proudu
- obdoba P a B snímků (proč to nemusí být vždy dobrý nápad?)

Picture Formats Supported

Picture format	Luminance pixels	Luminance lines	H.261 support	H.263 support	Uncompressed bitrate (Mbit/s)			
					10 frames/s		30 frames/s	
					Grey	Colour	Grey	Colour
SQCIF	128	96		Yes	1.0	1.5	3.0	4.4
QCIF	176	144	Yes	Yes	2.0	3.0	6.1	9.1
CIF	352	288	Optional	Optional	8.1	12.2	24.3	36.5
4CIF	704	576		Optional	32.4	48.7	97.3	146.0
16CIF	1408	1152		Optional	129.8	194.6	389.3	583.9

- porovnání H.261 a H263

- PSNR = Peak Signal Noise Ratio
- 15 kbps – 4-QAM
- 30 kbps – 16-QAM
- 45 kbps – 64-QAM



Další formáty

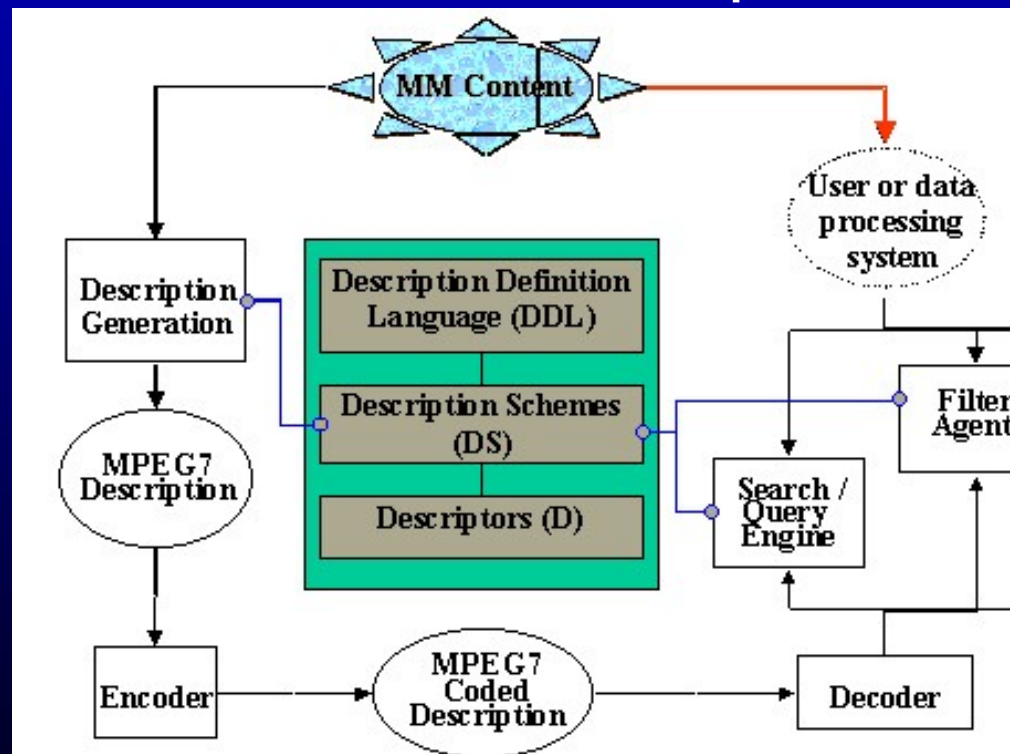
- Motion JPEG
 - série obrázku komprimovaná pomocí JPEGu
 - často s podporou HW
- AVI
 - formát MS
 - obálkový formát
- MOV
 - formát pro QuickTime
 - různé komprese

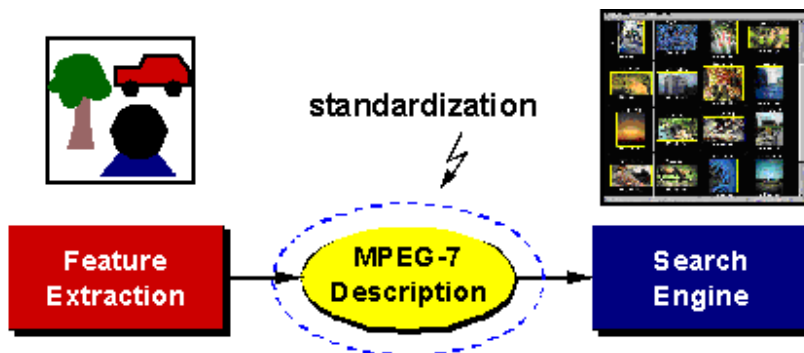
Nekomprimované video

- $768 * 576 * 24 * 25 = 265,420,800$
- k čemu to může být dobré?
 - vysoká kvalita
 - nízká latence

Metadata o multimédiích

- MPEG-7
 - Multimedia Content Description Interface

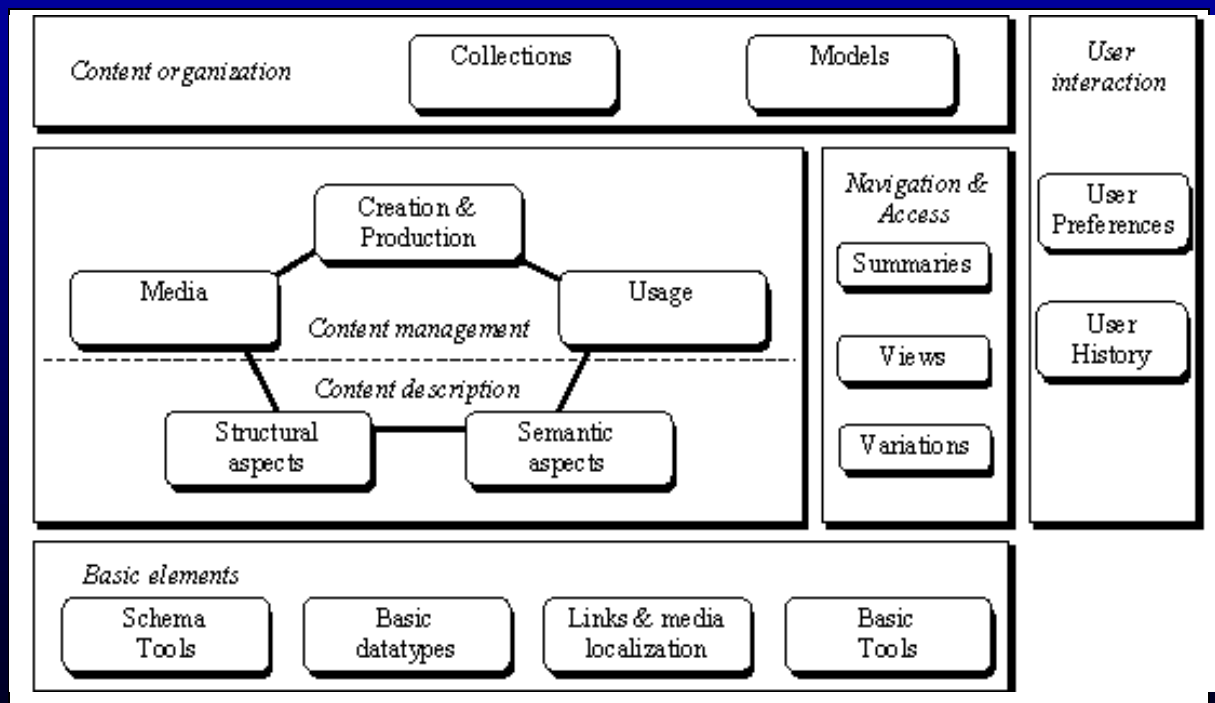




Feature Extraction:
 Content analysis (D, DS)
 Feature extraction (D, DS)
 Annotation tools (DS)
 Authoring (DS)

MPEG-7 Scope:
 Description Schemes (DSs)
 Descriptors (Ds)
 Language (DDL)
 Ref: MPEG-7 Concepts

Search Engine:
 Searching & filtering
 Classification
 Manipulation
 Summarization
 Indexing



- popis obrazu a zvuku:
 - obraz: tvar, velikost, povrch, barva, pohyb...
 - zvuk: klíč, nálada, tempo...
 - na nejvyšší úrovni: „This is a scene with a barking brown dog on the left and a blue ball that falls down on the right, with the sound of passing cars in the background.“
- ukazuje na různé části jednoho materiálu

- další informace
 - formát
 - podmínky přístupu k materiálu
 - klasifikace (např. předdefinované kategorie, např. adult)
 - odkazy na další relevantní materiály
 - kontext pořízení materiálu
- MPEG-21

Preview



Detect
temporal / spatial keys

Compose
a logical structure

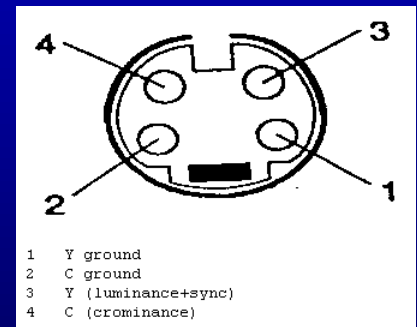
Annotate
MPEG-7

```
<Mpeg7 xmlns:="...">  
  <ContentDescription xsi:type="ContentDescription">  
    <AudioVisualContent id="" xsi:type="AudioVisualContent">  
      <SegmentDecomposition id="10" type="temporal" criteria="structure">  
        <Segment id="" xsi:type="AudioVisualSegmentType">  
          <MediaTime>  
            <MediaRefTimePoint timeBase="">  
              PT0S30F  
            </MediaRefTimePoint>  
            <MediaIncrDuration timeUnit="PT1N30F">  
              270  
            </MediaIncrDuration>  
          </MediaTime>  
        </Segment>  
      </SegmentDecomposition>  
    </AudioVisualContent>  
  </ContentDescription>  
</Mpeg7>
```

Hardware

- audio:
 - zvukové karty (D/A, A/D převodníky; syntéza zvuku)
 - vstupy: mikrofonní, linkové vstupy, digitální vstupy (S-PIDF)
 - výstupy: sluchátkové, linkové výstupy (zesilovač - > reproduktory), digitální výstupy (S-PIDF), podpora více než 2 reproduktorů
 - hw podpora: např. DSP procesory, podpora pro syntézu zvuků
- mikrofony
- zesilovače a repro vs. aktivní reproduktory

- video:
 - grafické karty
 - hw podpora kódování
 - grabovací karty (capture cards)
 - kompozitní signál vs. S-Video
 - firewire (IEEE-1394)



- scannery (bubnové, ploché – flatbed, filmové)
- videokamery
 - analogové (square pixel):
 - PAL (768x576, 25fps)
 - NTSC (640x480, 30/29.97 fps)
 - digitální (DV-based)
 - PAL (720x576, 25fps)
 - NTSC (720x480, 30/29.97 fps)
 - HDTV (mnoho možností)
 - interlaced vs. progressive scan

- HDTV

Common Name	Resolution	Frame Rate*	High-Definition†	Worth Saving?	
1080p	1920×1080	30 p	✓	✓	
	1920×1080	29.97 p	✓		
	1920×1080	24 p	✓	✓	
	1920×1080	23.976 p	✓		
1080i	1920×1080	30 i			
	1920×1080	29.97 i			
720p	1280×720	60 p	✓	✓	
	1280×720	59.94 p	✓		
	1280×720	30 p	✓	✓	
	1280×720	29.97 p	✓		
	1280×720	24 p	✓		
	1280×720	23.976 p	✓		
	704×480	60 p			
	704×480	59.94 p			
	704×480	30 p			
	704×480	29.97 p			
480p	640×480	60 p			
	640×480	59.94 p			
	640×480	30 p			
	640×480	29.97 p			
	640×480	24 p			
	640×480	23.976 p			
	480i	640×480	30 i		
		640×480	29.97 i		

- barevné prostory

- RGB

- YUV

- Y – luminance | luma

- U, V – chrominance

- $$Y = 0.299R + 0.587G + 0.114B$$

- $$U = 0.147R + 0.289G + 0.436B = 0.492(B - Y)$$

- $$V = 0.615R + 0.515G + 0.100B = 0.877(R - Y)$$

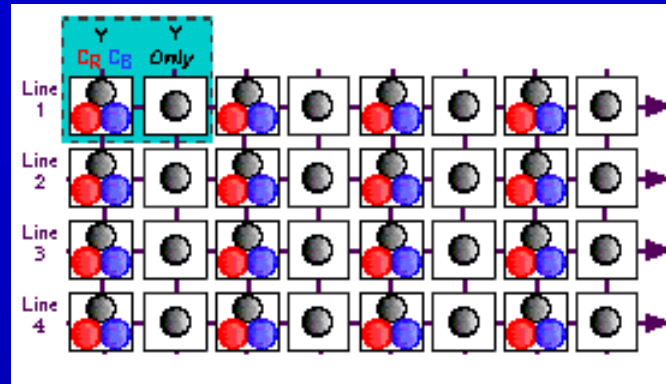
- YCrCb (YUV + scale + offset)

- $$Y = (77/256)R + (150/256)G + (29/256)B$$

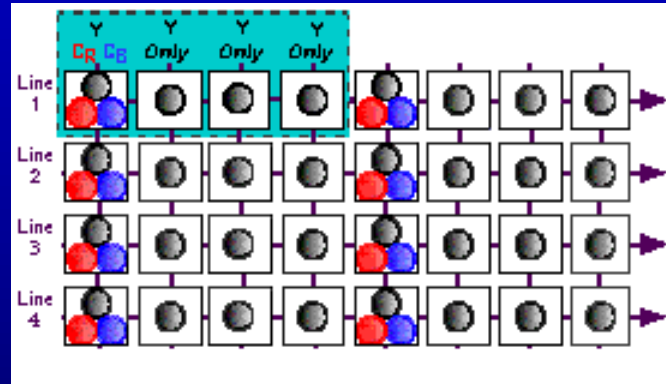
- $$Cb = - (44/256)R + (87/256)G + (131/256)B + 128$$

- $$Cr = (131/256)R + (110/256)G - (21/256)B + 128$$

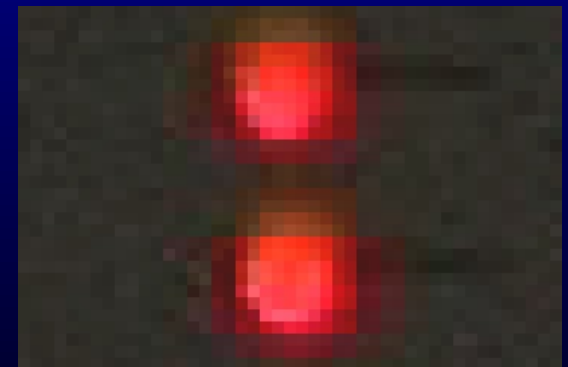
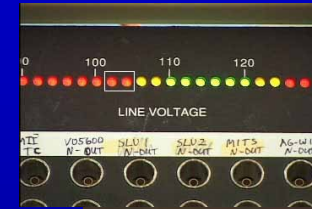
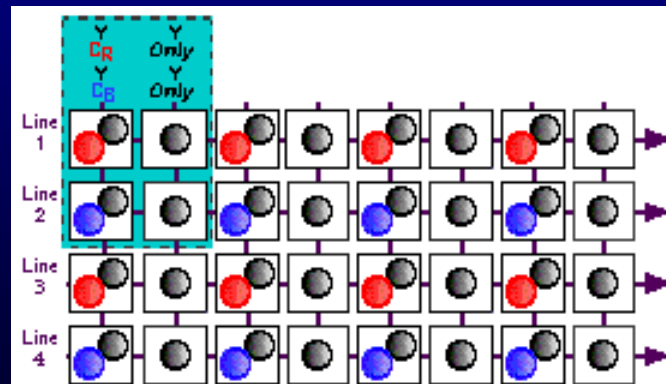
- vzorkování
– 4:2:2



- 4:1:1



- 4:2:0



Projekce

- zobrazování
 - CRT monitory, LCD displeje, plasmové obrazovky
 - projekce
- 3D
 - 3D projekce, helmy
 - polohování v prostoru
 - snímání objektů ve 3D
- haptika

Nároky multimedialních přenosů

- video
 - velká šířka pásma (negarantovaná), nepříliš velké nároky na QoS (výpadky nejsou příliš rušivé)
- audio
 - střední šířka pásma, důležité nároky na QoS
- haptika
 - extrémní nároky na QoS – problém s latencí (částečně limitované i rychlostí světla), nutnost existence lokálního modelu

Streamování

- video a zvuk
- díky jednosměrnosti můžeme data dostat na cílový počítač ještě před tím, než jsou použita
- bufferování (na straně klienta, příp. i serveru)
- potřeba rozumně spolehlivého síťového připojení

Streamovací produkty

- obálkové protokoly
 - RealMedia
 - Microsoft Media
 - QuickTime
 - mohou v sobě obsahovat různé formáty komprimovaného a nekomprimovaného videa a audia (např. MPEG-4 video a MP3 audio)
- pokusy se streamováním DivX
- synchronizační a režijní jazyky: Lingo (Macromedia), SMIL (W3C standard)

SMIL

- synchronizace
- SMIL (Synchronized Multimedia Integration Language)
 - <http://WDVL.com/Authoring/Languages/XML/SMIL/Intro/smil.html>
 - <http://www.w3.org/TR/REC-smil>
 - XML
 - např. RealPlayer G2, Quicktime >=4.0
 - MS neSMILní ;-)

- synchronizační tagy
 - dur (trvání)
 - begin (zpoždění)
 - seq (sekvenční přehrávání)
 - par (paralelní přehrávání)
- události
 - posílání události
 - čekání na událost
- switch tag

```
<smil>
<head>
  <layout>
    <root-layout height="400" width="600" background-color="#000000" title="Dreaming out Loud"/>
    <region id="satfam" width="564" height="400" top="0" left="0" background-color="#000000" z-index="2" />
    <region id="jupfam" width="349" height="400" top="0" left="251" background-color="#000000" z-index="2" />
    ...
    <region id="title" width="125" height="25" top="40" left="237" background-color="#ffffff" z-index="2" />
  </layout>
</head>

<body>
  <par>
    <audio src="media/dreamworldb.auz" dur="61.90s" begin="3.00s" system-bitrate="14000" />
    <seq>
      
      ...
      
      <par>
        
        
        <text src="media/music.txt" region="music" begin="3.00s" end="50.00s" />
        <text src="media/dreamland.txt" region="dreamland" begin="4.00s" end="50.00s" />
        ...
        <text src="media/me.txt" region="me" begin="20.00s" dur="3.00s" />
        <text src="media/jose.txt" region="jose" begin="23.00s" end="50.00s" />
      </par>
      <text src="media/title.txt" region="title" begin="3.00s" end="25.00s" />
    </seq>
  </par>
</body>
</smil>
```

Videokonference

- díky požadavku interaktivity není možné bufferování
- potřeba využívat kodeků s minimální latencí

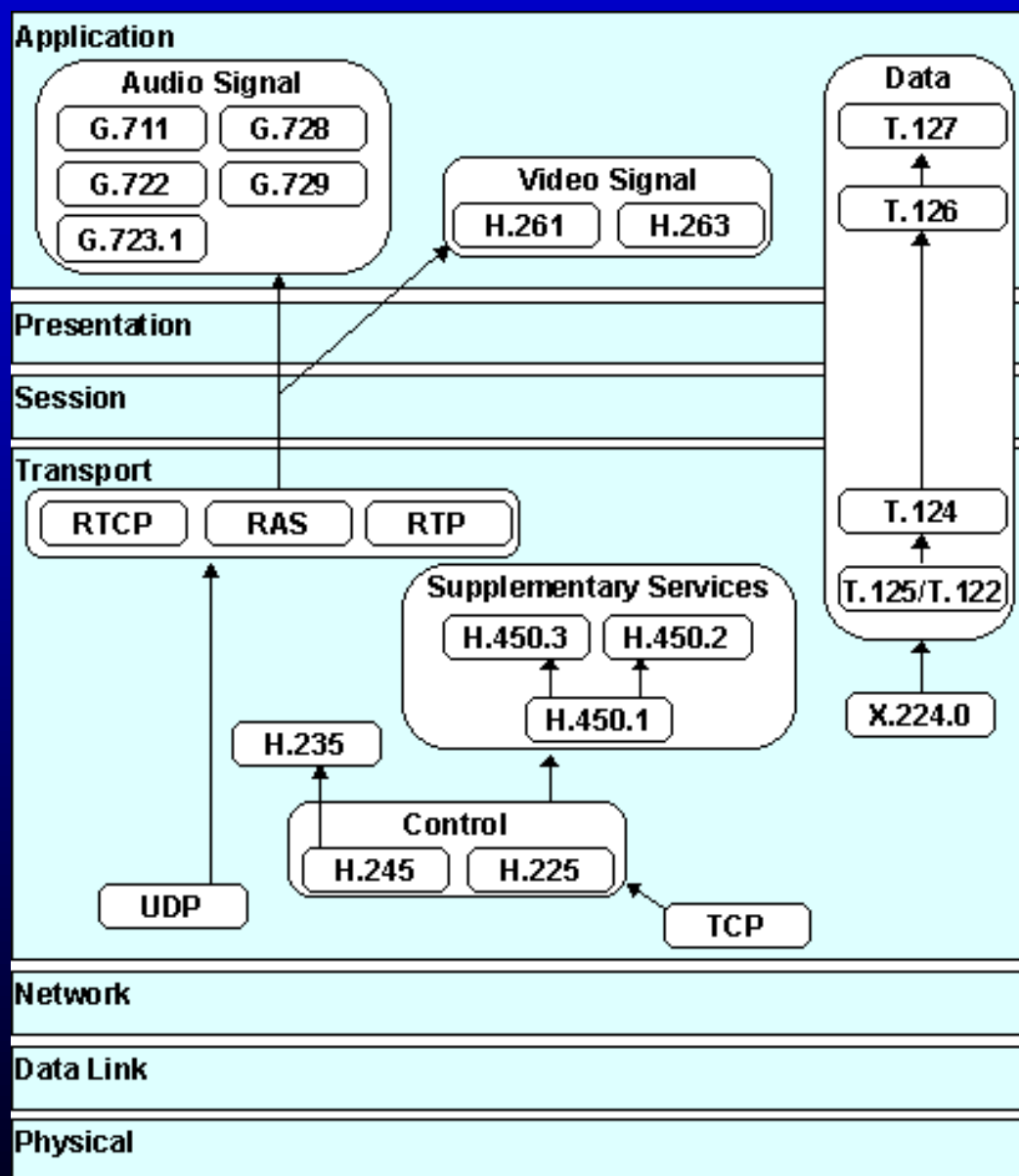
Standardy ve videokonferencích

- MBone tools
 - multicastový a unicastový režim
- nástroje vyvíjené nad H.323 (příp. SIPem)
 - většinou komerční s podporou HW
 - MS Netmeeting
 - CUSeeMe
 - OpenH323

Architektura H.323 videokonferencí

- klienti
 - HW a SW (SW povětšinou nejsou příliš kompatibilní)
- gatekeepery
 - konvertování dat pro různé sítě
 - AAA
- MCUs
 - H.323 je v podstatě point-to-point protokol
 - MCU přidává schopnosti point-to-multipoint

Architektura H.323 protokolu



VoIP

- integrace sítí s telefonními službami
- protokoly H.323 a SIP
- telefonní přístroje a ústředen s podporou VoIP
- někdy i s přenosem obrazu

Prostředky na sdílení prezentací

- sdílená pracovní plocha
 - wb/wbd/sdílené malování v NetMeetingu
 - VNC a sdílení aplikací pomocí NetMeetingu
- sdílení prezentací
 - distributed PowerPoint
 - VNC

QoS v operačních systémech

- *QoS není jen o sítích!*
 - QoS je třeba zajišťovat od uživatele k uživateli, včetně všech mezičlánků
- QoS na úrovni operačních systémů
 - přístup k procesoru
 - přístup k filesystému
 - přístup k síťovým rozhraním
 - podpora v aplikacích

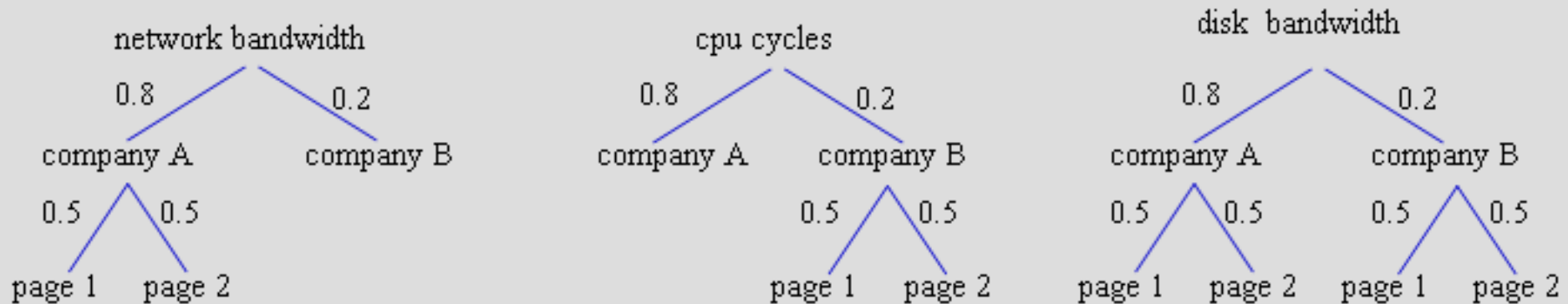
Eclipse/BSD

- začlenění podpory QoS do FreeBSD 3.4
- motivace
 - podpora aplikací (web servery, video servery)
 - QoS parametry na základě
 - klienta
 - serveru
 - přenášeného materiálu

- cíle

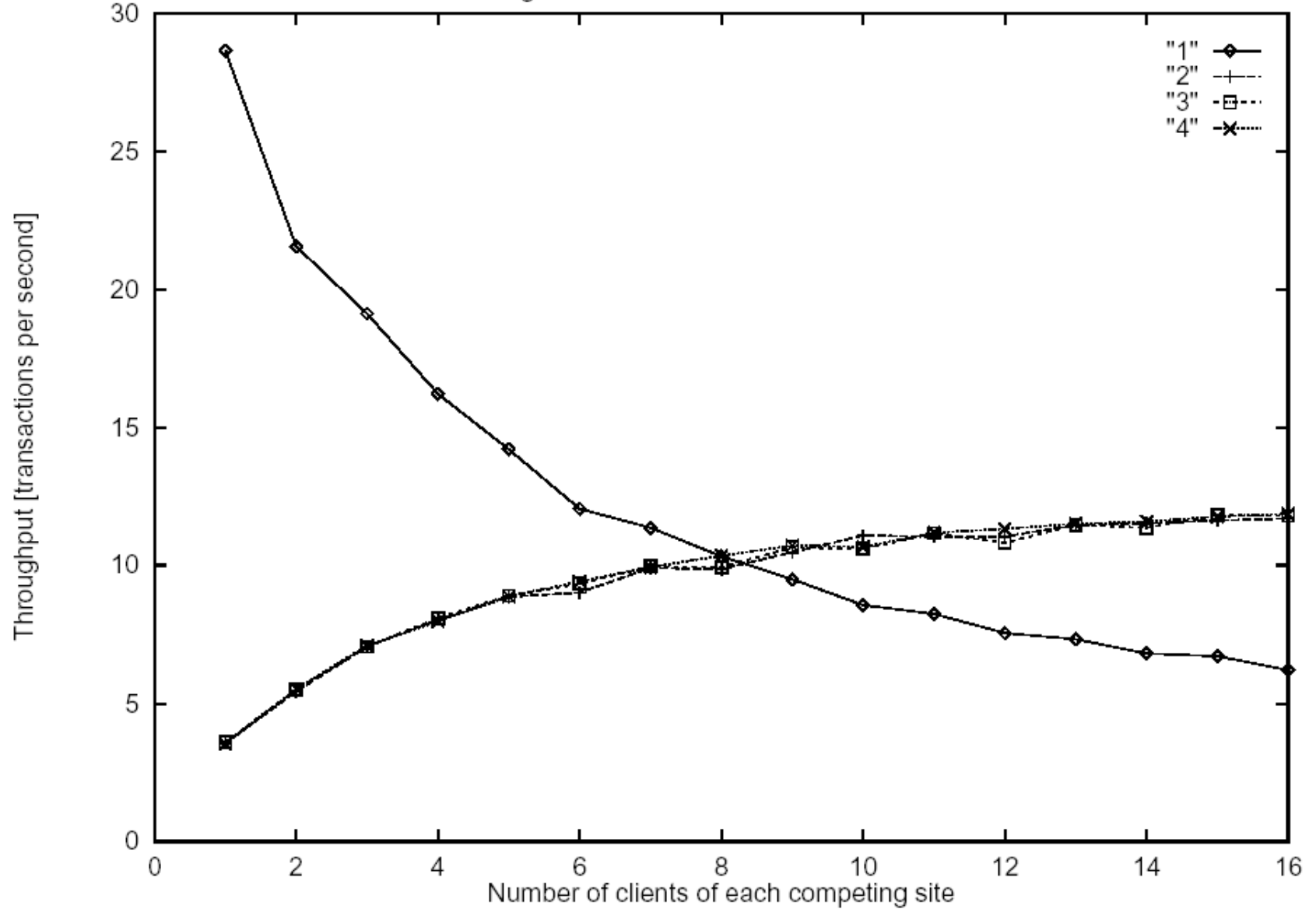
- isolation
- differentiation
- fairness
- (cumulative) throughput
- flexibilní management zdrojů
- podpora aplikací bez nutnosti významnějších změn jejich architektury
- hierarchičnost v plánování

- příklad – dvě společnosti mají hostována svá webová přípojná místa – každá společnost má dvě stránky

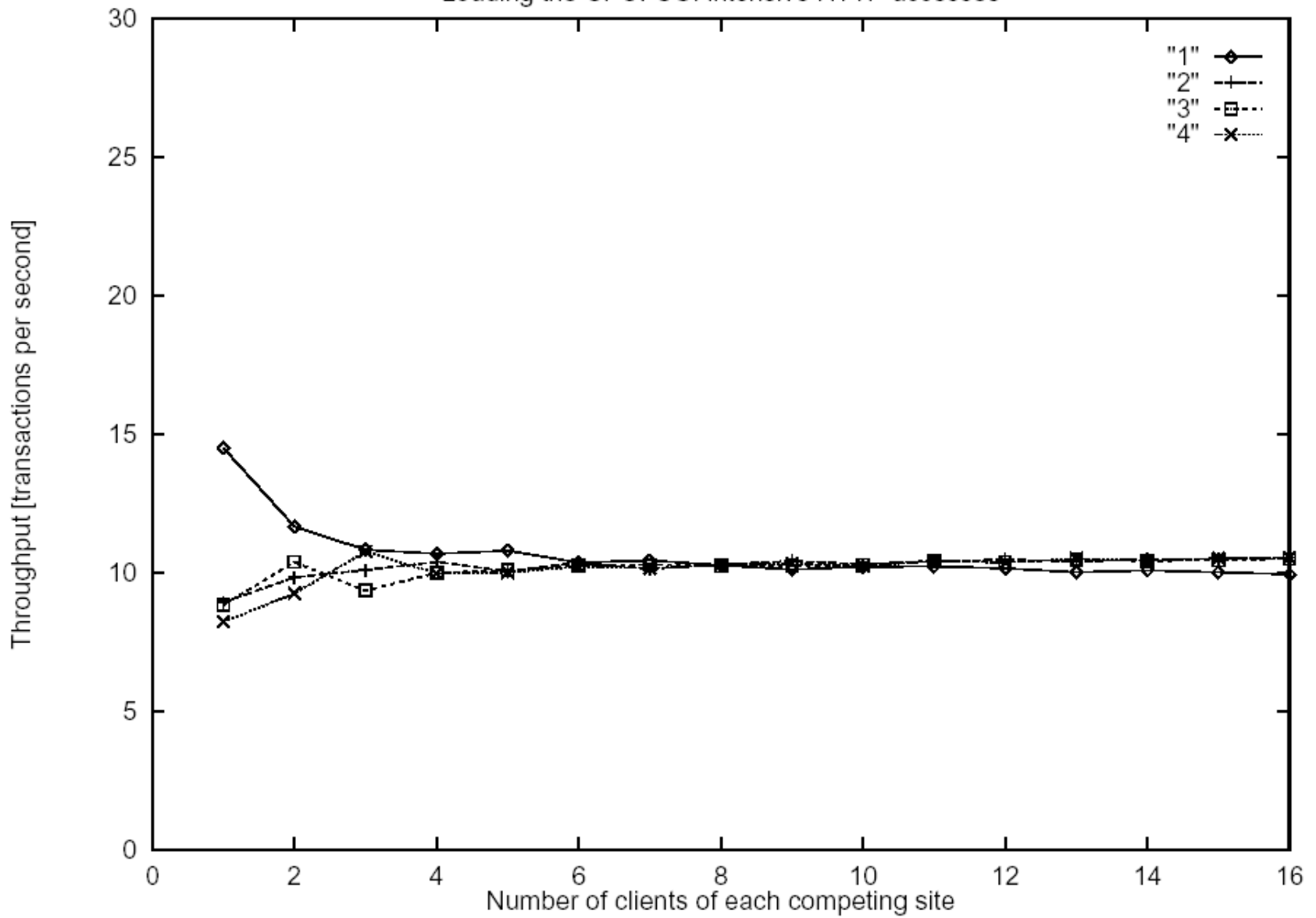


- ukázka chování pod zátěží – 1 Apache se 4 virtuálními hosty a modulem `mod_eclipse`
 - fixovaný počet klientů k místu 1
 - mění se počet klientů k místům 2, 3, 4
- obrázek 1 – FreeBSD
- obrázek 2 – Eclipse/BSD

Loading the CPU: CGI intensive HTTP accesses



Loading the CPU: CGI intensive HTTP accesses



Odkazy

- [http://www.cs.cf.ac.uk/Dave/Multimedia/BSC_MM_C
ALLER.html](http://www.cs.cf.ac.uk/Dave/Multimedia/BSC_MM_C<u>ALLER.html</u>)
- <http://www.bell-labs.com/project/eclipse/release/>
- <http://www.mpeg.org/MPEG/>
- <http://ipsi.fhg.de/delite/Projects/MPEG7/>
- <http://www.protocols.com/pbook/h323.htm>
- <http://www.protocols.com/pbook>
- <http://www.zvon.org>
- <http://www.dv.cz>

RUDOLPH YOU IDIOT...
I SAID THE **SCHMIDT** HOUSE!

