

2. Síť a multimédia

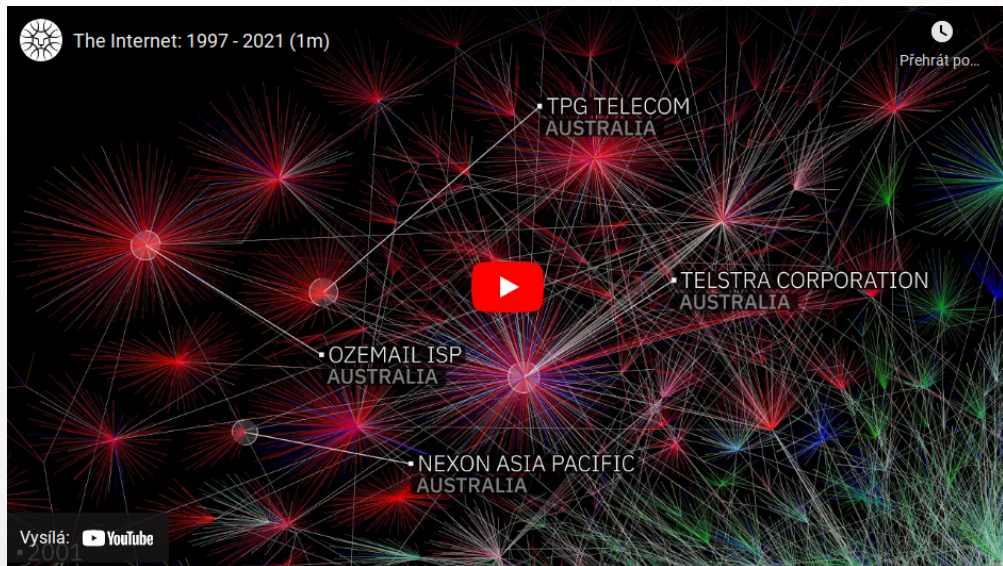
Miloš Liška
liska@fi.muni.cz

Ústav výpočetní techniky, MU

Druhý týden semestru

Internet

- Internet je celosvětový systém propojených počítačových sítí (sítí sítí), ve kterých mezi sebou komunikují počítače a další zařízení.

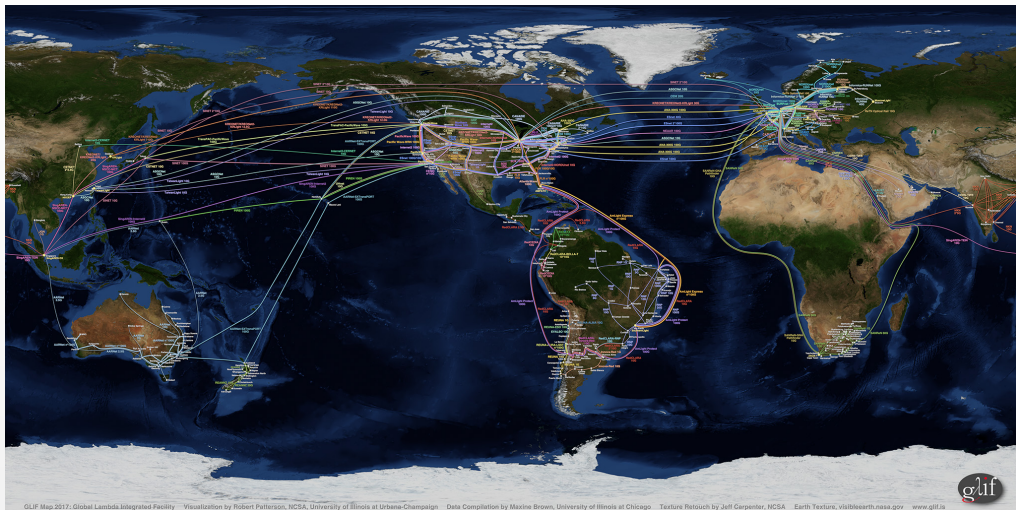


Počítačové sítě

Jak vzniká
digitální zvuk a
obraz

Přenos obrazu a
zvuku skrz
počítačovou síť

Aplikace
audiovizuálních
přenosů



Počítačové sítě

Jak vzniká
digitální zvuk a
obraz

Přenos obrazu a
zvuku skrz
počítačovou síť

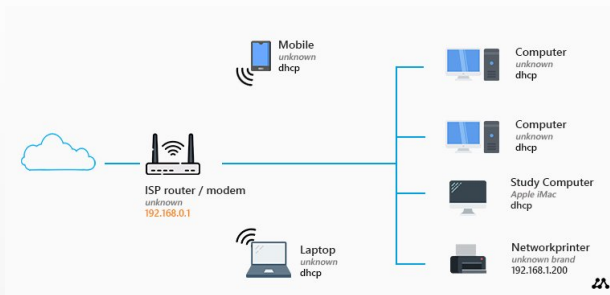
Aplikace
audiovizuálních
přenosů

Počítačové síť

Jak vzniká
digitální zvuk a
obraz

Přenos obrazu a
zvuku skrz
počítačovou síť

Aplikace
audiovizuálních
přenosů



■ Problémy domácích sítí

- Omezená šířka pásma směrem k poskytovateli (ISP)
- Nespolehlivá Wi-fi síť
- Zařízení uvnitř domácí sítě často nebývají dostupná z Internetu (NAT)

■ Optika

- Teoreticky nejrychlejší a nejspolehlivější
- V domácích podmínkách maximálně pro připojení k ISP

■ Metalické připojení

- Spolehlivé, v domácích podmínkách vhodné pro všechna zařízení na síti, která nejsou mobilní
- Dostatečná přenosová rychlost (viz dále)
- Staré telefonní linky s ADSL jsou také metalické připojení, o jeho spolehlivosti lze ale obvykle s úspěchem pochybovat

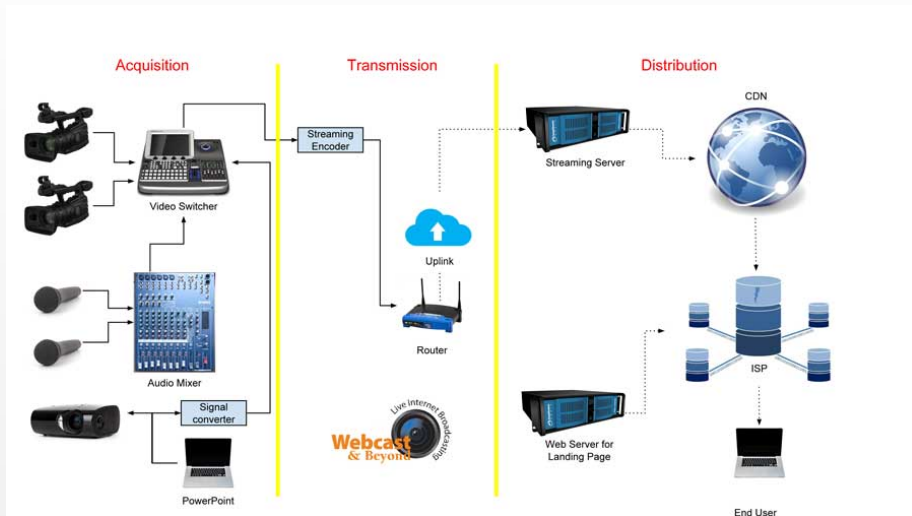
■ Wi-fi

- Principiálně nespolehlivá, sdílené pásmo, v jednu chvíli může vysílat/přijímat data pouze jedno zařízení
- Dnes běžně dvě frekvenční pásma na 2,4 GHz a 5 GHz
- Na 2,4 GHz je pouze 11–14 kanálů → typicky v místech s velkou koncentrací Wi-fi sítí (panelové domy) sdílí jeden kanál více sítí
- Na 5 GHz je 36–165 kanálů s větší šířkou pásma (propustností), ale s menším dosahem

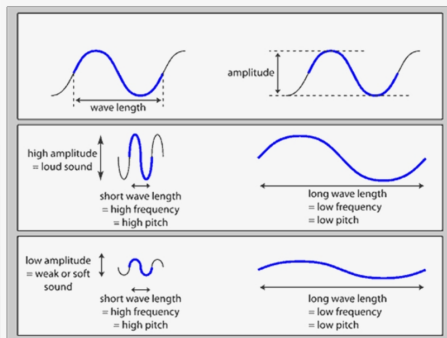
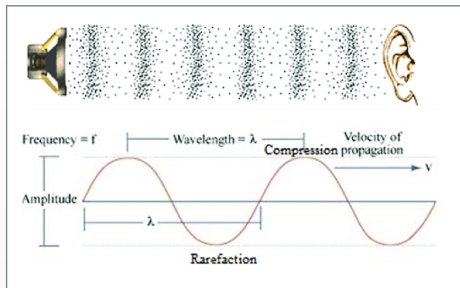
- Přenosová rychlost udává, jaký objem informace se přenesení za jednotku času
- Základní jednotkou přenosové rychlosti je bit za sekundu (bit/s, b/s, nebo anglicky bps = bits per second)
- Jednotka udává, kolik bitů informace je přeneseno za jednu sekundu.
 - Modem na analogové tel. lince → 14,4 až 52 kbps (Mimořadně text na tomto slidu má celkem 11,6 kb.)
 - 3G mobilní síť → cca 1 Mbps až jednotky Mbps
 - 4G mobilní síť → typicky maximálně desítky Mbps
 - 5G mobilní síť → ITU-T stanovuje 100/50 Mbps po dobu 95 % datového přenosu
 - ADSL připojení → jednotky Mbps až 250 Mbps
 - Připojení přes kabelovou TV → 38/10 Mbps až 444/122 Mbps
 - Optické síť → 1 Gbps až 400 Gbps
- Přenosová rychlost obvykle není symetrická
 - Výrazně ve prospěch stahování dat (download)
 - Odesílání dat (upload) je typicky omezené na jednotky až malé desítky Mbps
- Pozor na agregaci připojení

Co se děje když chci streamovat video na Youtube/Twitch...

- Magie ;)
- Trochu si jí rozebereme...



- Podélné mechanické vlnění v látkovém prostředí (vzduch), které je schopno vyvolat v lidském uchu sluchový vjem
- Vlastnosti zvuku
 - Frekvence
 - Lidé slyší zvuky ve frekvenčním pásmu přibližně 16 Hz až 20 kHz
 - Intenzita vlnění
 - Intenzita vlnění (zvuku) je definována jako množství energie, které projde jednotkovou plochou kolmou na směr šíření na jednotku času.
 - Zavádíme prahová intenzitu jako nejnižší ještě slyšitelnou intenzitu čistého tónu o frekvenci 1 kHz
 - Hladina intenzity zvuku (dB) se zavádí jako logaritmická škála, která prahové intenzitě přiřadí nulovou hodnotu. Logaritmickou škálu používáme protože intenzita běžných zvuků kolísá v rozsahu několika dekád.
 - Hlasitost zvuku
 - Vztah mezi intenzitou vlnění a intenzitou vjemu závisí mimo jiné i na frekvenci
 - Proto definujeme ještě hladinu hlasitosti a hlasitost, které berou do úvahy fyziologické vnímání intenzity zvuku v závislosti na jeho frekvenci



Počítačové sítě

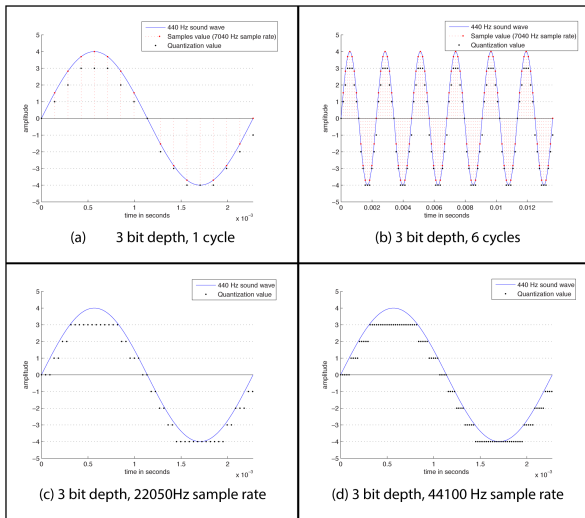
Jak vzniká
digitální zvuk a
obrazPřenos obrazu a
zvuku skrz
počítačovou síťAplikace
audiovizuálních
přenosů

- V pravidelných intervalech změříme intenzitu vstupního analogového signálu
- Interval mezi jednotlivými vzorky nazýváme vzorkovací interval
- Informaci o průběhu signálu mezi jednotlivými diskrétními měřeními ztrácíme
- Abychom byli schopni zaznamenat a rekonstruovat vstupní signál s nejvyšší frekvencí f , je třeba vzorkovací frekvence alespoň $2f$ (Nyquistův teorém)
 - Tj. pro nejvyšší slyšitelnou frekvenci 20 kHz potřebujeme vzorkovací frekvenci větší než 40 kHz
- Pro reprezentaci zvuk na CD používáme vzorkovací frekvenci 44,1 kHz
- Zvukové karty v PC pracují typicky se vzorkovací frekvencí 48 kHz (s touto hodnotou se lépe počítá)
- Pro vzorkování signálu, který odpovídá lidskému hlasu (např. pro digitální telefonii) stačí obvykle nižší vzorkovací frekvence 22 kHz, 11 kHz nebo i 8 kHz

- Hodnoty, naměřené při vzorkování, reprezentujeme jako čísla omezené délky (reprezentované konečným počtem bitů)
- Naměřené hodnoty zaokrouhlujeme na nejbližší diskrétní hodnotu
 - Vzniká kvantizační šum
- Počet bitů vyhrazených na reprezentaci intenzity vstupního signálu určuje kolik různých intenzit můžeme reprezentovat
 - Pro reprezentaci zvuku na CD pro reprezentaci jedné změřené hodnoty použijeme 16 b
 - Tj. můžeme reprezentovat $2^{16} = 65536$ různých intenzit signálu
 - Pro IP telefonii (kvantování hlasu) stačí opět nižší hodnota, například 8 b

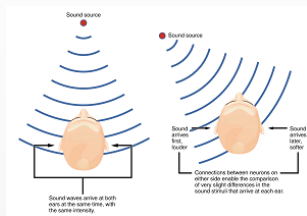
Příklad

- Vzorkujeme a kvantujeme analogový signál, jehož časový průběh je na diagramu
- Vzorkování je v diagramu naznačeno čárkovanými svislými čarami (okamžiky měření na časové ose)
- Pro reprezentaci intenzity signálu v každém vzorku používáme 3 bity



Mono, Stereo, Surround sound

- Člověk má dvě uši, mozek dostává dvě nezávislé informace o zvuku ze kterých na základě intenzity vnímaného zvuku vytváří zvukový prostorový vjem.



- Monofonní zvuk = 1 kanál
- Prostorový zvuk
 - Pro každý směrový zdroj zvuku nezávislý kanál
 - Klasický stereofonní zvuk přicházející „zleva a zprava“ jsou dva nezávislé audio kanály
 - 5.1 prostorový zvuk je 6 nezávislých audio kanálů
- Každý nezávislý zvukový kanál je vzorkovaný a kvantovaný tak, jak jsme si ukazovali na předchozích slidech

- $8 \text{ kHz} \times 8 \text{ b} \times 1 \text{ kanál} = 64 \text{ kbps}$
- $44,1 \text{ kHz} \times 16 \text{ b} \times 2 \text{ kanály} = 1411 \text{ kbps}$ (cca. 1,4 Mbps)
- $44,1 \text{ kHz} \times 16 \text{ b} \times 6 \text{ kanálů} = 4,233 \text{ kbps}$ (cca. 4,2 Mbps)
- $192 \text{ kHz} \times 24 \text{ b} \times 2 \text{ kanály} = 9000 \text{ kbps}$ (cca 9 Mbps)

A je to hodně dat nebo ne?

- Red Book audio standard (klasické CD) od Philipsu a Sony z roku 1980
 - Právě $44,1 \text{ kHz} \times 16 \text{ b} \times 2 \text{ kanály} = 1411 \text{ kbps}$ (cca. 1,4 Mbps)
- Rychlost čtení z 3,5" diskety byla cca 50 kbps
- Celá ČR byla v roce 1993 připojena k Internetu rychlostí 9,6 kbps
- 1,4 Mbps také odpovídá 31,5 MB za 3 minuty (relativně běžná délka písničky)
- Kapacita 3,5" diskety je až 1,44 MB
- V roce 1995 byla obvyklá kapacita pevného disku 500 MB
- Kapacita úložiště „MP3“ přehrávače v roce 2005 byla např. 1000 MB (1 GB)
- Potřebujeme kompresi!

- Kompresní mechanismy MPEG-1 Layer 3 (mp3), MPEG-2 Layer 3 (aac) apod.
- Kompresní poměr 1:10 a více
 - Tj. na reprezentaci komprimovaného audia (typicky 2 kanály, 48 kHz vzorkovací frekvence a 16 bitů na reprezentaci vzorku) potřebujeme 10x méně dat
- Zavádíme pojem **bitrate** (b/s, bps)
 - Průměrný počet bitů za sekundu, kterými kódujeme zvuk
 - Typicky daný kapacitou média (CD, kapacita „MP3“ přehrávače) nebo dostupnou šířkou pásma.
 - Obvykle 8–324 kbps pro ztrátovou kompresi
 - Např. bitrate u služby Spotify se pohybuje mezi 96 a 160 kbps (stereofonní zvuk)
 - 384 kbps pro prostorový zvuk na DVD
 - Srovnej např. s dostupnou šířkou pásma v 2G mobilní síti (50 kbps)

- Pohyblivé obrázky
- Video je jen sekvence statických snímků zachycených a přehrávaných s konkrétní, stabilní frekvencí
- Temporální (časová) dimenze videa
 - Video zachycuje scénu a v ní se pohybující objekty (ev. pohybující se celou scénou)
 - Pohyb ve scéně je spojitý a je třeba jej diskretizovat
 - Vzorkování dělením spojitého pohybu na sekvenci statických snímků
- Spaciální (prostorová) dimenze videa
 - Jednotlivé statické snímky videa zachycují dvourozměrný, v čase statický obraz konkrétní scény a objektů v ní
 - Statický obraz opět nese spojitou informaci
 - Statický obraz dále diskretizujeme opět vzorkováním a kvantováním

- Video (film) je sekvence statických snímků zachycených a zpět přehraných s danou snímkovou frekvencí, která má za úkol vyvolat dojem plynule pohybujících se objektů
- Vzorkování videa v temporální časové (dimenzi)
- Psychologická hranice pro udržení dojmu přirozeně se pohybujících objektů je 12 snímků za sekundu (frames per second, fps)
- Časování snímání u nejstarší techniky (a němých filmů) bylo zcela manuální (kameraman točil klikou a ručně posouval filmový pás, na kterém se exponovala jednotlivá políčka), průměrně 16 fps
- Tradičních 24 fps u filmu vychází jako kompromis mezi cenou filmového materiálu a nemožností synchronizovat zvuk s obrazem při nízkých snímkových frekvencích
- S nástupem televize se snímková frekvence začala řídit frekvencí střídavého proudu (50 Hz resp 59.94 Hz)
 - Odtud jsou odvozeny snímkové frekvence 25 fps resp. 29,97 (30) fps
- Dnes se běžně používá i 60 fps – věrnější vzorkování rychle se pohybujících objektů ve scéně

Počítačové sítě

Jak vzniká
digitální zvuk a
obraz

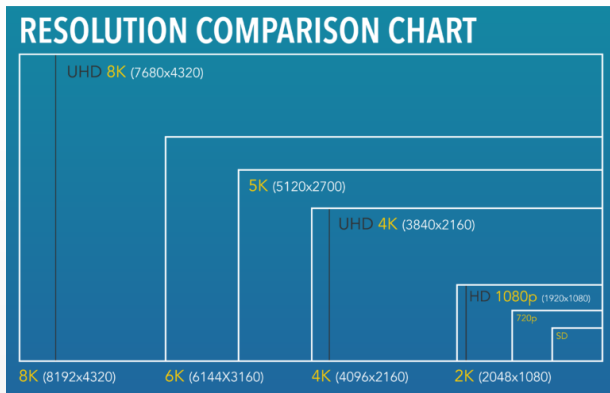
Přenos obrazu a
zvuku skrz
počítačovou síť

Aplikace
audiovizuálních
přenosů

- Elektromagnetické vlnění s velmi úzkou šířkou spektra (viditelné světlo) odražené od objektů v okolí a dopadající na světlocitlivé buňky sítnice oka.
- Vzruchy z oka jsou dále zpracovávány a interpretovány mozkiem.
- Obraz interpretujeme na základě zkušeností.
- V porovnání s jakoukoliv digitální technikou je v jistých oblastech oko spolu s mozkiem velice dokonalá soustava (dynamický rozsah vidění, schopnost vyvážení bílé, rychlost ostření, noční vidění atd.)

- Obraz vzorkujeme ve dvou rozměrech konečně velkou maticí $M \times N$ bodů (pixelů) a obrazovou informaci (barvu) v každém bodě kvantujeme do K úrovní
- Vzorkování
 - Vzorkování obrazu se řídí Shanonovou větou
 - Nejmenší detail, který chceme zachytit, v digitálním obrazu musí být minimálně dvakrát větší než velikost vzorku
 - Velikost vzorku je daná fyzickými rozměry zachycené scény a počtem pixelů v obou dimenzích





Počítačové sítě

Jak vzniká
digitální zvuk a
obraz

Přenos obrazu a
zvuku skrz
počítačovou síť

Aplikace
audiovizuálních
přenosů

- Při daných rozměrech scény a zmenšujícím se rozlišení přicházíme o detaily v zachyceném obrazu
- Příklad: snímáme scénu o rozměrech $8 \times 4,5$ m
 - 8K rozlišení \rightarrow 1 pixel zabírá plochu $1,04 \times 1,04$ mm
 - HD rozlišení \rightarrow 1 pixel zabírá plochu $4,17 \text{ mm} \times 4,17 \text{ mm}$
 - SD rozlišení (analogová TV) \rightarrow 1 pixel zabírá plochu $1,04 \times 0,78$ cm

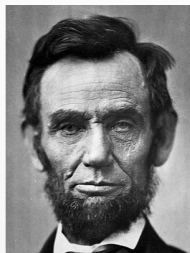
Kvantování obrazu

- Pro každý diskretní vzorek obrazu (pixel) potřebujeme reprezentovat jeho barvu
- Barev ve viditelném spektru je opět spojitě mnoho
- Kvantujeme je do K úrovní. Otázka je kolik úrovní potřebujeme pro věrné zachycení barev.
- Abychom dokázali reprezentaci barev uchopit, používáme barevné modely
 - Model RGB - barva se skládá ze tří složek Červená, Zelená, Modrá
 - Model YUV - barva se skládá z jasové složky a dvou barevných složek
- Každou složku reprezentujeme typicky 8, 10 (HDR) nebo 12bity (HDR)
- Při 8 bitech na barevnou složku máme k dispozici $(2^8)^3 = 16.8mil.$ barev (může se to zdát hodně, ale nemusí to stačit)
- Při 10 bitech na barevnou složku máme k dispozici $(2^{10})^3 = 1mld.$ barev



Vzorkování a kvantování obrazu

- Příklad speciálního vzorkování obrazu 12×16 vzorky (pixely) a kvantování 8 bity (256 úrovněmi) v jednom barevném kanálu
- $12 \times 16 \times 8b = 1536b = 192B$



187	183	174	168	160	162	126	183	172	183	166	186
165	182	163	74	76	62	88	17	110	216	180	184
180	180	80	14	34	6	10	39	48	106	159	181
206	186	6	124	131	111	120	204	166	18	86	180
194	68	137	251	237	236	236	228	227	87	71	201
172	186	207	233	233	214	220	236	228	98	74	206
188	88	179	208	186	215	211	158	130	76	20	169
189	87	165	84	16	168	134	91	91	62	22	148
199	168	191	189	162	227	178	142	182	116	96	200
208	174	165	252	236	231	149	178	228	43	90	234
180	216	116	149	236	187	86	150	79	38	218	241
190	224	147	106	227	170	127	143	94	111	265	224
180	214	173	66	133	143	96	80	2	109	249	219
187	186	236	76	1	81	47	0	6	217	256	211
183	202	237	145	0	0	12	106	190	136	243	236
186	206	123	387	177	121	123	200	176	13	96	218

187	183	174	168	160	162	126	183	172	183	166	186
165	182	163	74	76	62	88	17	110	216	180	184
180	180	80	14	34	6	10	39	48	106	159	181
206	186	6	124	131	111	120	204	166	18	86	180
194	68	137	251	237	236	236	228	227	87	71	201
172	186	207	233	233	214	220	236	228	98	74	206
188	88	179	208	186	215	211	158	130	76	20	169
189	87	165	84	16	168	134	91	91	62	22	148
199	168	191	189	162	227	178	142	182	116	96	200
208	174	165	252	236	231	149	178	228	43	90	234
180	216	116	149	236	187	86	150	79	38	218	241
190	224	147	106	227	170	127	143	94	111	265	224
180	214	173	66	133	143	96	80	2	109	249	219
187	186	236	76	1	81	47	0	6	217	256	211
183	202	237	145	0	0	12	106	190	136	243	236
186	206	123	387	177	121	123	200	176	13	96	218

- 192 B je velmi málo dat na reprezentaci obrazu
- Při vzorkování 12×16 má nejmenší detail, který jsme schopni v takovém obrazu zachytit jedním pixelem rozměry cca 4×4 cm (viz Shannonova věta)
- Kvantování v jednom barevném kanálu 8 bity umožňuje reprezentovat pouze 256 úrovní jasu/odstínů šedé

Počítačové sítě

Jak vzniká
digitální zvuk a
obraz

Přenos obrazu a
zvuku skrz
počítačovou síť

Aplikace
audiovizuálních
přenosů

Kolik dat tedy potřebujeme pro reprezentaci nekomprimovaného videa?

- Analogová TV (PAL):

$$768 \times 576 \times 24[b/\text{pixel}(RGB)] \times 25[\text{fps}] = 265,420 \text{ kbps} = 259 \text{ Mbps}$$

- Full HD video:

$$1920 \times 1080 \times 24[b/\text{pixel}(YUV)] \times 2/3[4 : 2 : 2\text{sampling}] \times 30[\text{fps}] = 995,328 \text{ kbps} = 972 \text{ Mbps}$$

- 4K video:

$$3840 \times 2160 \times 30[b/\text{pixel}(YUV)][4 : 4 : 4\text{sampling}] \times 60[\text{fps}] = 14,929,920 \text{ kbps} = 14 \text{ Gbps}$$

- 8K video:

$$7680 \times 4320 \times 24[b/\text{pixel}(YUV)] \times 2/3[4 : 2 : 2\text{sampling}] \times 60[\text{fps}] = 31,850,496 \text{ kbps} = 31 \text{ Gbps}$$

- Nekomprimované video typicky znamená obrovské datové objemy
 - Nejen vzhledem k propustnosti počítačových sítí ale i např. sběrnic v PC
- Beztrátová komprese videa přináší kompresní poměry 1:5 až 1:12 – stále málo
- Ztrátová komprese (typicky H.264 nebo H.265/HEVC) přináší kompresní poměry 1:20 až 1:200 – to je již zajímavé
- Doporučený bitrate při kompresi videa pomocí kompresního mechanismu H.264 na Youtube (bitrate může být reálně i nižší pokud se pohybujeme stále ve stejných řádech).

Počítačové sítě

Jak vzniká
digitální zvuk a
obraz

Přenos obrazu a
zvuku skrz
počítačovou síť

Aplikace
audiovizuálních
přenosů

Type	Video Bitrate, Standard Frame Rate (24, 25, 30)	Video Bitrate, High Frame Rate (48, 50, 60)
2160p (4k)	35-45 Mbps	53-68 Mbps
1440p (2k)	16 Mbps	24 Mbps
1080p	8 Mbps	12 Mbps
720p	5 Mbps	7.5 Mbps
480p	2.5 Mbps	4 Mbps
360p	1 Mbps	1.5 Mbps

Recommended video bitrates for HDR uploads

Type	Video Bitrate, Standard Frame Rate (24, 25, 30)	Video Bitrate, High Frame Rate (48, 50, 60)
2160p (4k)	44-56 Mbps	66-85 Mbps
1440p (2k)	20 Mbps	30 Mbps
1080p	10 Mbps	15 Mbps
720p	6.5 Mbps	9.5 Mbps

Zpoždění při zpracování zvuku

- Zvuk zpracováváme v řetězci operací (pipeline)
- Digitalizace zvukového signálu zvukovou kartou → komprese zvuku (→ síťový přenos) → dekomprese zvuku → převod zvukového signálu zpět do analogové podoby na zvukové kartě
- Zpoždění zpracování audia v celé pipeline je tzv. end-to-end zpoždění
- Zvukový signál nemůžeme zpracovávat po jednotlivých vzorcích
 - Při vzorkovací frekvenci 48 kHz je vzorkovací interval dlouhý 20 μ s
 - Časování zpracování vzorků zvuku musí být velmi přesné, každá nepřesnost je slyšitelná
 - Vzorky, které nestihneme zpracovat v daném čase, můžeme už pouze zahodit, protože už potřebujeme zpracovávat další vzorky zvuku
 - Tak přesného časování nejsme schopni na procesoru s běžným operačním systémem dosáhnout
- Proto každá operace v pipeline bufferuje (ukládá do mezipaměti) a zpracovává více vzorků naráz
- Obvykle 64, 128 nebo 265 vzorků audia naráz → zpoždění každé operace je 1,3 ms až 5,3 ms
- Zpoždění pipeline pro zpracování zvuku tedy může tedy být i cca. 20 ms (bez zpoždění síťového přenosu)

- Zpoždění nabíráme v celém řetězci počínaje kamerou (zdrojem videa), přes akvizici obrazu a zvuku až po přehrávání obrazu a jeho zobrazování na displeji
 - Kamera čas daný vzorkovacím intervalem mezi jednotlivými snímky na expozici statického obrazu
 - Při 60 fps je vzorkovací interval $1/60 \text{ s} = 16,6 \text{ ms}$. Při 25 fps je vzorkovací interval $1/25 \text{ s} = 40 \text{ ms}$
 - Pokud není scéna dostatečně nasvícena, jsou to vlastně relativně krátké expoziční časy
 - Video dále zpracováváme typicky po celých snímcích v rámci posloupnosti operací (pipeline)
 - Typicky akvizice snímku → komprese snímku (→ síťový přenos) → dekomprese snímku → zobrazení snímku na grafické kartě (výstupní kartě pro zpracování videa → zobrazení snímku na displeji)
 - Doba maximálního trvání každé operace je daná délkou vzorkovacího intervalu mezi jednotlivými statickými snímky

Zpoždění při zpracování obrazu

- Akvizice jednoho statického, digitálního snímku z kamery trvá celý vzorkovací interval
- Kompresi a dekomprese je typicky výrazně rychlejší, k síťovému přenosu se ještě dostaneme
- Zobrazení snímku na grafické kartě nebo kartě pro výstup videa trvá opět celý vzorkovací interval
- Zpoždění displeje může být řádově jednotky milisekund u herních displejů až po desítky milisekund
- Tj. při zpracování videa po celých snímcích bude zpoždění větší než $2 \times 1/fps$
- Pokud je vzorkovací interval 40 ms (25 fps), pak bude zpoždění při zpracování obrazu minimálně 80 ms, neobvyklé není ani zpoždění 250 ms (1/4 s)
 - I komprese a dekomprese každého snímku může trvat až 40 ms, zobrazování bývá také pomalejší

Vadí nám takové zpoždění?

- Typicky očekáváme synchronní obraz a zvuk
- Pokud obraz a zvuk explicitně nesynchronizujeme, bude se jejich zpoždění rozcházet
- Rozdíl zpoždění > 100 ms je již viditelný in na desynchronizaci řeči a pohybu rtů mluvčího ve videu (lip-sync)
- V případě hudby a například záběrů prstů na klaviatuře klavíru vadí i rozdíl zpoždění v řádu malých desítek ms
- Mezinárodní telefonní unie ITU-T stanovuje, že maximální end-to-end zpoždění v telefonii je 150 ms
- Při zpoždění nad 300 ms už může být nepříjemné vést delší dobu normální rozhovor

- Přenos obrazu a zvuku je z hlediska využití šířky pásma v porovnání s aplikacemi jako email, přístup na webové stránky, Facebook apod. řádově náročnější
- Přenos obrazu a zvuku skrz počítačovou síť navíc časový rozměr
- Streaming videa a audia například z Youtube nemusí probíhat v reálném čase, uživatel může interagovat typicky pouze pomocí textového chatu a zpoždění přenosu v desítek sekund nebo i jednotek minut lze tolerovat
 - Můžeme řešit i krátkodobé výpadky dostupné šířky pásma
- Videokonference (Skype, Whatsapp, Google Meet, Zoom apod.) v reálném čase probíhají, ale lze tolerovat zpoždění až v řádu malých stovek ms
 - I malé výpadky v dostupné šířce pásma typicky znamenají ztrátu obrazových a zvukových dat
- Čím nižší zpoždění přenosu obrazu a zvuku vyžadujeme, tím větší klademe na síť nároky

Zpoždění přenosu v počítačových sítích

- Zpoždění je způsobeno jednak fyzikálně při přenosu signálu (bezdrátově, v metalických kabelech i v optice) a zpožděním zpracování paketů v síťových prvcích
- Rychlost šíření signálu je shora omezená rychlostí (pomalostí světla)
 - Cca 300000 km/s
 - Ale např. nejkratší vzdálenost mezi Brnem a San Diegem je 9825 km tj. zpoždění přenosu signálu (i po optice) bude cca $0,033 \text{ s} = 33 \text{ ms}$
 - Reálná délka všech optických kabelů na trase je mnohem větší (délky transatlantických optických kabelů se pohybují mezi 6000 a 13000 km)
- Zpoždění jednotlivých síťových prvků je v řádu desítek mikrosekund
- Reálné zpoždění vysokorychlostních sítí při přenosech v rámci ČR je okolo 2 ms
- Zpoždění kabelového připojení v síti Vodafone (UPC) je při přenosech v rámci ČR jen okolo 3 ms
- Ale při použití telefonních linek (ADSL v síti O2) a domácí Wi-fi se budeme pohybovat už kolem 20 ms
- A reálné zpoždění vysokorychlostí sítě při přenosu z Brna do San Diega bude okolo 85 ms
- Zpoždění síťového přenosu je potřeba přičíst ke zpoždění zpracování obrazu a zvuku!

- Streaming videa z Youtube a videokonference přes Skype, Whatsapp, Facebook, Google Hangouts Meet, Zoom apod. všichni známe.
- Jaké nové aplikace založené na audiovizuálních přenosech současné sítě umožňují?
 - Hraní her prostřednictvím audiovizuálních přenosů včetně VR/AR.
 - Je možné operovat prostřednictvím videa na dálku?
 - Může produkční tým pracovat na nejnovějším hollywoodském filmu z obýváku?
 - Mohou spolu hrát hudebníci po síti?

- Dnes relativně běžné technologie (Google Stadia, NVidia GeForceNOW)
- Obraz a zvuk her je generovaný v cloudu a přenášený v podstatě obdobně jako streaming z Youtube, zpět se přenáší události z klávesnice/herního ovladače
- Větší nároky na kvalitu a šířku pásma (cca 35 Mbps pro 4K video)
- Obvykle implementováno pomocí webových technologií, které nejsou nutně stavěné na co nejnižší zpoždění
- Jednosměrné zpoždění herní cloud → obrazovka uživatele je běžně cca 100 ms
 - K tomu je potřeba přičíst zpoždění ovládání a zpoždění sítě mezi herním zařízením a cloudovou platformou
- Při hraní na lokálním PC je zpoždění obvykle kolem 40 ms (60 fps video)
- I tak je to docela použitelné a dá se na hraní prostřednictvím přenosů obrazu a zvuku zvyknout

Je možné operovat prostřednictvím videa?

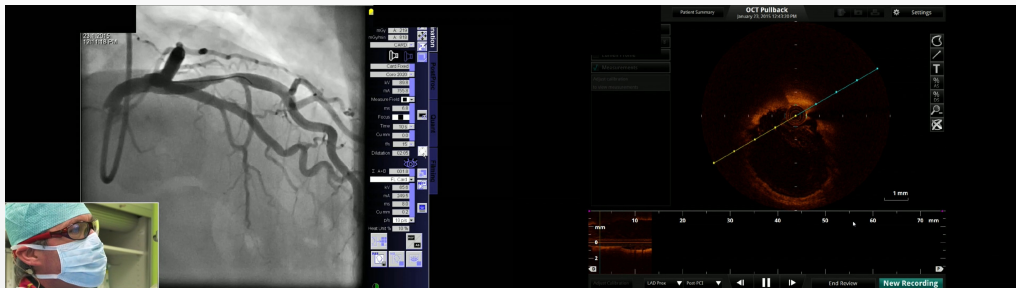
- Operace, kdy operatér provádí zákrok na základě obrazů z kamer a dalších medicínských zobrazovacích metod jsou dnes zcela běžné
- Šetrnější pro pacienta, malé operační rány
- Endoskopické/laparoskopické zákroky, kardiologie, dentální chirurgie apod.

Počítačové sítě

Jak vzniká
digitální zvuk a
obraz

Přenos obrazu a
zvuku skrz
počítačovou síť

Aplikace
audiovizuálních
přenosů



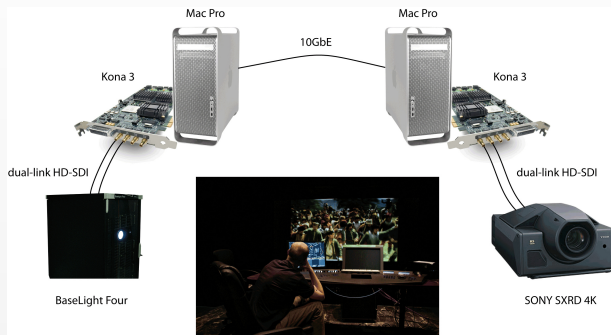
Je možné operovat prostřednictvím videa na dálku?

- A proč ne? Z pohledu operátora je to stejná aplikace jako hraní her prostřednictvím audiovizuálních přenosů
- da Vinci Surgical System
- Menší nároky na interaktivitu než u akčních her
- Operátor je schopen se naučit, že ovládání nástrojů je opožděno až cca o 0,5 s (500 ms)
- Vyšší nároky na kvalitu obrazu než u běžných videokonferencí
- Potřebná šířka pásma v řádu menších desítek Mbps



Produkce filmů z obýváku

- Vzdálená postprodukce filmového materiálu (např. color grading) se přímo nabízí
 - HDR video s reprezentací barev až 16 bity na barevný kanál je nutná
 - Dnes velká rozlišení (4K–8K)
 - Datové toky v řádech stovek Mbps až jednotek Gbps



- Filmová studia by ráda tento proces nasadila tak, aby například koloristi mohli pracovat z domu (zejména v souvislosti s COVID-19)
 - Studio má typicky desítky koloristů
 - I při sofistikované kompresi je potřeba 100 Mbps na stream
 - Problém s domácími sítěmi

Mohou spolu hrát hudebníci po síti?

- Rychlost zvuku ve vzduchu je 343 ms^{-1} , při šířce orchestřiště 10 m je zpoždění mezi hudebníky do cca 35 ms, na takové zpoždění jsou hudebníci cvičeni
- End-to-end zpoždění 35ms je pro tuto aplikaci svatý grál i při síťových přenosech
 - Ideálně pro zvuk i obraz
- Takové zpoždění je extrémně obtížně dosažitelné
 - End-to-end zpoždění při běžném zpracování zvuku na PC je v řádu malých desítek ms (bez síťového přenosu)
 - Pokud zpracováváme obraz po jednotlivých snímcích, pak je zpoždění přímo závislé na snímkové frekvenci a dosáhnout end-to-end zpoždění $\leq 35 \text{ ms}$ je možné jen při velmi vysokých snímkových frekvencích
 - Síťové zpoždění při přenosech v rámci ČR může být 10 ms v jednom směru, mezi Evropou a Amerikou/Asií ale 100 a více ms

UltraGrid

High-quality and Low Latency Multimedia Transmissions
Using Commodity Hardware

www.ultragrid.cz

cesnet
■■■■■■■■■■



Počítačové sítě

Jak vzniká
digitální zvuk a
obraz

Přenos obrazu a
zvuku skrz
počítačovou síť

Aplikace
audiovizuálních
přenosů