

MULTIMEDIÁLNÍ A HYPERMEDIÁLNÍ SYSTEMY

11)
Distribuce multimédií

Petr Lobaz, 12. 5. 2015

HISTORICKÉ SOUVISLOSTI

PŮVODNÍ DISTRIBUCE OZVUČENÉHO FILMU

- filmový pás + zvuk na externím nosiči (např. gramofonová deska nebo CD)
 - nutnost synchronizace
 - snadná možnost špatného spárování
 - snadná možnost distribuce filmu s jazykovými mutacemi zvuku
- filmový pás se zvukovou stopou
 - souběžný synchronní přenos
 - už v 50. letech se objevují filmové pásy s dvojitou kvalitou zvukové stopy (základní kvalita – optický záznam, vícekanálový zvuk – magnetický záznam)

HISTORICKÉ SOUVISLOSTI

- situace navíc komplikovaná rozdělením filmu na několik cívek
 - nutnost dodržet předepsané pořadí promítání (playlist)
 - nutnost synchronního, nepozorovatelného napojení cívek při promítání
- prakticky stejné problémy se řeší ve světě digitálních multimédií
 - přenos mediálních dat
 - jejich synchronizace
 - jejich nepozorovatelné napojování

HISTORICKÉ SOUVISLOSTI

ŘEŠENÍ V ANALOGOVÉ TELEVIZI

- černobílá televize: souběžné televizní vysílání obrazové a zvukové stopy v odlišných frekvenčních pásmech
- barevná televize: zavedení dalšího frekvenčního pásma pro souběžné vysílání barvonosného (chrominančního) signálu
 - problém: jak zajistit, aby se tato pásma (tzv. *frekvenční multiplex*) navzájem nerušila
- řešení používané v satelitním přenosu: *časový multiplex* (soustavy MAC)
 - signály zvukové, jasové a barvonosné rozděleny na krátké úseky (pakety) odpovídající jedné řádce obrazu
 - rychlé střídavé vysílání postupně jeden paket za druhým
 - při dekódování složení paketů do potřebné podoby

DIGITÁLNÍ MULTIMÉDIA

ULOŽENÍ MULTIMEDIÁLNÍCH DAT

- obrazová a zvuková data zakódována do podoby elementárních proudů (*ES – elementary stream*)
 - kódování může (a nemusí) obsahovat techniky komprese dat
- požadavky na uložení/přenos při zajištění synchronizace ES jsou různé a protichůdné
 - obvykle se volí kompromis mezi univerzálností, snadností použití a hardwarové nezávislosti

ZÁKLADNÍ PODOBY ULOŽENÍ

ČASOVÝ MULTIPLEX

- obrazová a zvuková data z ES rozdělena do paketů
- střídavé ukládání do proudu multimedialních dat
- časově související data jsou uložena fyzicky blízko u sebe
- pakety opatřeny časovou známkou (timestamp)
 - zajištění synchronizace a odolnosti vůči výpadku paketu
- granularita (velikost paketů) závisí na velikosti vyrovnávacích pamětí vysílače/přijímače
- typická distribuční podoba – snadné přehrávání

ZÁKLADNÍ PODOBY ULOŽENÍ

PLAYLIST

- na datovém uložišti jsou uloženy jednotlivé ES
- každý ES opatřen obsahem (index)
 - informace, na jaké adrese začíná který snímek (případně struktura GOP atd.)
- multimedialní data: předpis přehrávání jednotlivých ES, případně jejich částí (EDL – edit decision list)
 - předpisem se mění pořadí a časování (synchronizace)
- výhodné pro střih a studiové zpracování
 - pracuje se jen se symbolickou informací
- výhodné pro skupinovou spolupráci na tvorbě videa
 - přenáší se jen symbolická informace (krátká)
 - samotné ES (objemné) jsou na společném uložišti

KONTEJNEROVÁ STRUKTURA

- konkrétní způsob propojení elementárních proudů
 - implementace závislá na paměťovém médiu a požadované funkcionalitě
- typický obsah:
 - samotná data elementárních proudů, případně informace o jejich umístění
 - technická charakteristika obrazu (rozlišení, poměr stran, barevný prostor, ... + typ kódování)
 - technická charakteristika zvuku (vzorkovací frekvence, počet kanálů, ... + typ kódování)
 - netechnické informace k elementárním proudům (název televizního programu, jazyk zvukové stopy, ...)
 - synchronizační značky
 - data k interaktivní funkcionalitě

MULTIMEDIÁLNÍ DATA

ZÁKLADNÍ OPERACE

- kódování
 - kódování (příp. komprese) zvukových, obrazových aj. dat do podoby ES
 - skládání ES do kontejnerové struktury
- dekódování
 - výběr ES z kontejnerové struktury
 - dekódování (příp. dekomprese) zvukových, obrazových aj. dat

MULTIMEDIÁLNÍ DATA

- transkódování
 - ideálně nevyžaduje dekódování a kódování ES (časově náročné, obvykle ztrátové)
 - změna charakteristik ES, např. redukce datového toku
 - přidání nebo ubrání ES v kontejnerové struktuře
 - převod na jinou kontejnerovou strukturu
 - plné dekódování a kódování ES se obvykle vyžaduje při změně obsahu ES
nebo změně kódování (např. typu komprese) ES

PAMĚŤOVÁ MÉDIA

- požadavky na kontejnerovou strukturu úzce souvisí s používaným paměťovým médiem

VYSÍLÁNÍ

- např. televizní nebo rozhlasové
- de facto nejde o paměťové médium
 - fyzické uložení vysílaných dat nemusí existovat
 - přijímač i vysílač mají pouze krátké vyrovnávací paměti pro aktuálně přenášený obsah
- vysílač ani přijímač nemají možnost regulovat proud dat
- přirozené požadavky: konstantní datový tok, časový multiplex ES
- např. MPEG TS (distribuční kontejner pro televizní vysílání)

PAMĚŤOVÁ MÉDIA

PÁSKOVÁ MÉDIA

- možnost zastavení a převíjení pásky, případně přehrávání jinou rychlostí
- během přehrávání se obsah chová podobně jako při vysílání – nemožnost ovlivnit tok dat
- doba převíjení pásky dlouhá
 - ⇒ nemožnost náhodného přístupu
 - ⇒ obsah musí být na pásce přesně připraven v prezentační podobě
- např. DV, HDV (digitální páskový videozáznam)
- přirozené požadavky: konstantní datový tok, časový multiplex ES, při převíjení musí být dekódovatelná data o pozici na pásce, resp. náhledový obraz

PAMĚŤOVÁ MÉDIA

„LINEÁRNÍ“ DISKOVÁ MÉDIA

- přesuny čtecí/záznamové hlavy na disku celkem rychlé
⇒ při dostatečně velké vyrovnávací paměti lze simulovat náhodný přístup k datům
- náhodný přístup využitý k interaktivitě
 - dělení pořadu na kapitoly
 - volitelné pořadí přehrávání
 - ...
- požadavky nejsou tak striktní jako u páskových médií
- obvykle časový multiplex ES + několik playlistů + mechanismus, jak mezi playlisty vybírat a přepínat
- např. Audio CD, DVD-Video

PAMĚŤOVÁ MÉDIA

SOUBOROVĚ ORIENTOVANÁ MÉDIA

- zajištěn rychlý náhodný přístup k datům
 - ⇒ data mohou být organizována podle smyslu do souborů a ne podle pořadí při záznamu/přehrávání
 - ⇒ data uvnitř souboru mohou být organizována implementačně nejpohodlnějším způsobem
- fyzická reprezentace může být různá (harddisk, paměťová karta, RAM, vzdálený přístup)
- požadavky se řídí konkrétní aplikací, omezeno jen rychlostí náhodného přístupu a velikostí vyrovnávacích pamětí
- např. MP4, MKV/WebM, MXF
- Blu-ray (distribuční i záznamová podoba) využívá prvky souborového i lineárního diskového média

PRÁCE S MM DATY

- požadavky závisí i na způsobu používání multimedálních dat

VYSÍLÁNÍ

- nepřerušovaný proud dat (*stream*)
- typicky televizní vysílání (*broadcast*), internetové vysílání
- musí umožňovat začátek dekódování „uprostřed proudu“ (pojem „začátek proudu“ de facto neexistuje)
- vyžadována odolnost vůči výpadku dat, chybě v přenosu
- v prostředí počítačové sítě vyžadováno řízení kvality, resp. datového toku
- často vyžadován přenos několika televizních programů v jednom proudu dat
- minimální požadavky na interaktivitu – přepínání mezi datovými proudy, zobrazení textových informací apod.

PRÁCE S MM DATY

VÝMĚNA DAT

- např. pro uložení hotového pořadu
- výsledek se využije pro tvorbu vysílací podoby, podoby pro Blu-ray, ...
- záznam by měl obsahovat jen data + informace, jak je dekodovat
- na strukturu kontejneru nejsou zvláštní požadavky

PRACOVNÍ STUDIOVÝ FORMÁT

- založeno na playlistu, resp. EDL (edit decision list)
- obrazová a zvuková data mohou být uložena v samostatných ES, nebo v jiných pracovních studiových formátech

PRÁCE S MM DATY

PRACOVNÍ FORMÁT K PŘÍPRAVĚ VYSÍLACÍ PODOBY

- obsahuje jen data z hotového pořadu v časovém multiplexu „rozumné“ granularity
- formát připraven k snadné tvorbě různých distribučních podob (pro pozemní vysílání, internetové vysílání apod.)
- formát ideálně obsahuje jen návod, jak distribuční podobu připravit, aniž by transkodér musel rozumět detailům komprese obrazu a zvuku

PŘEHRÁVÁNÍ Z LOKÁLNÍHO ULOŽIŠTĚ

- typicky vyžadován náhodný přístup, strukturovaný přístup (dělení na kapitoly), interaktivita

UKÁZKA KONTEJNERU

- moderní kontejnerové struktury jsou často navrženy tak, aby umožňovaly různá využití
 - jejich vnitřní organizace se ale podle způsobu využití liší

ZÁKLADNÍ NAIVNÍ KONTEJNER

- příklad vychází z kontejnerové struktury AVI
- příklad slouží k ukázce různých myšlenek a pastí v návrhu kontejnerové struktury
- data organizována v hierarchické struktuře „krabiček“
- „krabička“: (délka v bytech), (typ krabičky), (data krabičky)
- podle typu „krabičky“ mohou data obsahovat další „krabičky“ nebo neorganizovanou sekvenci bytů

UKÁZKA KONTEJNERU

```
[ záhlaví
  délka videa
  počet ES
  [ informace o ES1: obraz
    barevný prostor, poměr stran, snímková frekvence
    typ kódování (komprese) ]
  [ informace o ES2: zvuk
    počet kanálů, vzorkovací frekvence, počet bitů na vzorek
    typ kódování ]
  [ informace o ES3: alternativní zvuk ... ]
  ... ]
[ obsah
  [ data ES1 (obraz) snímku 1 ]
  [ data ES2 (zvuk) snímku 1 ]
  [ data ES3 (alternativní zvuk) snímku 1 ]
  ...
  [ data ES1 (obraz) snímku 2 ]
  [ data ES2 (zvuk) snímku 2 ]
  ...]
[index pozic v souboru
  pozice začátku snímku 1
  pozice začátku snímku 2
  ...]
```

UKÁZKA KONTEJNERU

VÝHODNÉ VLASTNOSTI

- jednoduchost
- hodí se pro lokální přehrávání videa s jednoduše kódovanými ES
- pro jednoduché jednoúčelové použití je návrh v pořádku, problémy začne klást při používání nad rámec původního záměru

UKÁZKA KONTEJNERU

NEVÝHODNÉ VLASTNOSTI

- pro dekódování je nutné znát záhlaví
⇒ nehodí se pro vysílání
- při poškození záhlaví je zbytek souboru nedekódovatelný
- jednotlivé ES musí mít stejnou délku
⇒ nehodí se pro pracovní formát
- data obrazového ES musí být uložena v pořadí přehrávání
⇒ nehodí se pro komprese se složitější strukturou GOP
- data zvukových ES musí být dělitelná na časové úseky odpovídající jednomu snímku ⇒ nehodí se pro komprese zvuku pracující s většími bloky vzorků
- umožňuje pouze pevnou snímkovou frekvenci ⇒ nehodí se např. pro záznamy přednášek („komentovaný powerpoint“), kodéry reagující na opakující se snímky ve videu apod.

SYNCHRONIZAČNÍ STRUKTURA

- základ všech kontejnerů
- elementární proud rozdělen na bloky (pakety), každý má:
 - prezentační časovou známku (PTS – presentation time stamp): kdy se má blok začít zobrazovat
 - dekódovací časovou známku (DTS – decoding time stamp): kdy má dekodér začít s dekódováním bloku
- umožňuje, aby dekodér nemusel o ničem rozhodovat, ale jen vykonával předepsanou posloupnost činností (např. rozhodování o přehazování snímků v GOP může být komplikované)
- struktura bloků obecně nesouvisí se strukturou snímků (blok může obsahovat část dat pro snímek, data jednoho nebo více snímků)

SYNCHRONIZAČNÍ STRUKTURA

- časové známky se nemusí objevovat u každého bloku
 - snižuje množství nadbytečně přenášených dat
 - vyžaduje složitější dekodér
 - pokud blok obsahuje více snímků, vztahují se časové známky k prvnímu
 - pokud blok obsahuje jen část dat snímku, časové známky chybí
 - chybějící časové známky se odvozují např. z konstantní snímkové frekvence
- v klasickém vysílání (pozemním, satelitním) je nutné zařídit stejnou rychlost hodin vysílače a přijímače
 - v proudu dat další synchronizační značky, přijímač podle nich doladuje rychlost interních hodin

SYNCHRONIZAČNÍ STRUKTURA

- časové známky musí být kvůli přesnosti založeny na celých číslech
 - typicky se specifikuje časová báze (např. 27 MHz)
 - časové známky se udávají vzhledem k časové bázi
 - každý elementární proud má typicky svou vlastní časovou bázi
 - pro omezené pořady stačí reprezentace celým číslem
 - musí se vyřešit, jak časové známky reprezentovat při nepřetržitém vysílání (obvyklá zkratka: 24/7)
- bloky se označují identifikačním číslem, v každém bloku vyšší o 1
 - umožňuje kontrolu výpadku bloku
- takto upravený elementární proud: **PES** (packetized ES)

MPEG-2 PS (PROGRAM STREAM)

- hodí se pro bezporuchová prostředí
- typicky kódování jednoho programu (obrazové + zvukové stopy)
- není zamýšlený pro přímou distribuci
- rozdělen do svazků (pack): záhlaví + několik PES paketů
 - PES pakety ve svazku mají typicky podobné časové známky ⇒ časový multiplex
- kromě svazků se průběžně objevují:
 - záhlaví proudu: definuje parametry časování
 - deskriptory programu: např. typ komprese
 - tyto průběžné informace umožňují začít dekódovat proud dat víceméně odkudkoliv

DVD-VIDEO

- založeno na MPEG-2 Program stream
 - multimedialní data uložena v souborech *.VOB (video object) – MPEG-2 PS
 - zásadní informace o jejich obsahu (např. snímková frekvence, poměr stran apod.) uloženy v souborech *.IFO
- VOB obsahuje
 - obrazové stopy
 - související zvukové stopy
 - statické obrázky s časovými známkami – titulky
 - příkazy jednoduchého programovacího jazyka

DVD-VIDEO

- příkazy specifikují, jak má probíhat přehrávání
 - nejčastějším příkazem je skok na jistou část VOB
 - skoky mohou být podmíněné obsahem vnitřních proměnných
 - programovací jazyk dále definuje, jak vnitřní proměnné nastavovat

⇒ de facto algoritmicky generovaný playlist – program chain (PGC)
- pomocí různých PGC se definuje
 - interaktivní chování menu disku
 - uživatelský výběr zvukové stopy, titulkové sady apod.
 - přehrávání alternativních verzí filmu
 - ...

MPEG-2 TS (TRANSPORT STREAM)

- navržen pro nespolehlivá prostředí
- primární účel: vysílání televizního signálu
- používán i jako kontejner v kamerách nebo součást složitějších systémů (Blu-ray)
- rozdělen do krátkých paketů stejné délky (188 bytů)
- do paketů se v časovém multiplexu ukládají
 - samostatné PES
 - hotové MPEG-2 PS
 - hotové MPEG-2 TS (na jejich složení platí stejná pravidla)
- v záhlaví paketu uložen identifikátor, ke kterému proudu paket patří
 - pakety se speciálním identifikátorem 0 přenášejí sružovací tabulku (PAT – program association table)
 - sružovací tabulka určuje, které proudy patří k sobě

BLU-RAY

- struktura původně navržena pro záznam obrazu
⇒ založeno na MPEG-2 TS
 - samotná AV data uložena v souborech *.m2ts
- kvůli snadné editaci založena na playlistech
 - playlisty uloženy v samostatných souborech *.mpls
 - informace o TS v samostatných souborech *.clpi
- playlisty je možné (podobně jako na PGC na DVD)
„vytvářet“ algoritmicky
- díky oddělené struktuře playlist / AV obsah je možné kombinovat AV data, která nejsou multiplexovaná
 - titulky mohou být v samostatném souboru
 - rychlost obrazu a zvuku se mohou lišit (např. interaktivní prohlížení obrazové galerie může být s asynchronním zvukovým doprovodem)

MPEG-4 PART 12

- ISO Base Media File Format
- odstraňuje nevýhodu předchozích přístupů:
ES musel být převeden na pakety (PES)
- vychází z kontejneru QuickTime,
od něj je odvozen např. MP4, Flash Video (F4V)
- struktura podobná „základnímu naivnímu kontejneru“,
ale s podstatnými rozdíly
- soubor složen z „krabiček“ (box)
 - jejich pořadí není pevně dané (!)
- box může obsahovat neorganizovaná data a/nebo další boxy
- data ES mohou být uložena kdekoliv (!)
(v rámci jednoho souboru v boxech typu „mdat“,
v externích souborech, na síti, ...)

MPEG-4 PART 12

- box typu „trak“ rozděluje data ES pomocí indexů na „vzorky“ (několik snímků, část zvuku apod.)
 - každému „vzorku“ specifikuje časové známky PTS a DTS
 - ES tím pádem mohou zůstat v původní podobě
 - pokud je to vhodné, mohou být ES libovolně multiplexovány do sebe
- box typu „hint“ obsahuje podrobný návod, jak vytvořit podobu AV dat vhodnou k vysílání (streamování)
 - v jednom souboru může být víc hint boxů (např. jeden pro tvorbu MPEG TS, další pro MPEG PS atd.)

MATROSKA (MKV)

- kontejner především určený pro lokální přehrávání
- díky vnitřnímu uspořádání snadný převod do vysílací (streamovací) podoby
 - ⇒ z MKV odvozen formát WebM pro internetové video
- hierarchická struktura založená na XML
 - ⇒ velmi volné konkrétní pořadí elementů

INTERNETOVÁ DISTRIBUCE MM

- příprava kontejneru vhodného k lokálnímu přehrávání + stažení souboru
 - např. iTunes, knihovny filmů apod.
- příprava kontejneru vhodného k streamování + použití speciálního síťového protokolu
 - často použity protokoly RTP (přenos dat) + RTCP (řízení toku dat)
 - upřednostňuje se co nejrychlejší doručení síťového paketu před spolehlivostí doručení
 - vhodné pro přímé přenosy i přehrání z archivu (media on demand)

INTERNETOVÁ DISTRIBUCE MM

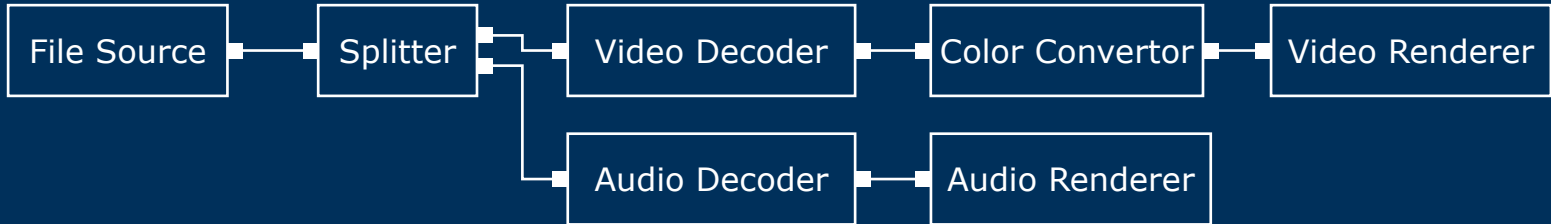
- progresivní stahování
 - rozdělení obsahu na krátké kousky (např. 5 s)
 - kousky uloženy do kontejnerů vhodných k lokálnímu přehrávání
 - příprava playlistu složeného z kousků
 - stažení playlistu, průběžné stahování kousků
 - jednotlivé kousky mohou být připraveny v různé kvalitě, při stahování dalšího kousku automatický výběr kvality podle aktuálního stavu síťového připojení
 - dá se použít i pro přímý přenos: ke stažení k dispozici několik kousků, playlist se průběžně obnovuje
 - pro stahování se dá použít i http protokol (viz např. standard MPEG-DASH)

INTERNETOVÁ DISTRIBUCE MM

- obvykle je distribuce doplněna generováním náhledových obrázků (thumbnails) pro usnadnění náhodného přístupu do videa
- distribuce obvykle probíhá přes Content Delivery Network (CDN)
 - data duplikována na mnoha serverech
 - při požadavku na sledování multimediálního obsahu se automaticky vybere nejdostupnější server
- technická podoba se volí s ohledem na typ distribuce
 - 1:1 (jeden vysílač, jeden přijímač) – např. telefonie
 - 1:N – např. internetová televize
 - M:N – např. videokonference
 - technická podoba musí zejména řešit minimalizaci síťového provozu a synchronizaci mezi účastníky

POČÍTAČOVÉ ZPRACOVÁNÍ MM

- obvykle založeno na představě modulů předávajících si obsah
- moduly se skládají do grafů



POČÍTAČOVÉ ZPRACOVÁNÍ MM

- modul poskytuje standardní funkce typu přijmiData(), odešliData() a interní zpracujData()
- propojení filtrů $A \rightarrow B$
 - de facto funkce A.odešliData() volá B.přijmiData() (metoda push)
 - nebo funkce B.přijmiData() volá A.odešliData() (metoda pull)
 - proces propojení filtrů: A, resp. B se dozví adresu funkce, kterou má volat
 - proces předávání dat je mírně složitější zařazením vyrovnávacích pamětí, ale princip zůstává

POČÍTAČOVÉ ZPRACOVÁNÍ MM

- jednotlivé moduly a mechanismus skládání grafů poskytuje operační systém nebo externí knihovny
 - Windows: aktuálně Microsoft Media Foundation, stále se lze setkat s DirectShow, Video for Windows
 - MacOS: QuickTime
 - externí knihovny: např. libavcodec (typicky na Linuxu)

POČÍTAČOVÉ ZPRACOVÁNÍ MM

TYPICKÉ MODULY

- hlavní charakteristikou je počet vstupů a výstupů do grafu
- datové vstupy: ze souboru, ze síťového připojení, z externího zařízení, ...
- datové výstupy
- splitter (demuxery, demultiplexory): rozdělení kontejnerové struktury na elementární proudy
- muxery (multiplexory): skládání kontejnerové struktury z elementárních proudů
- dekodéry: dekodování (dekomprese) elementárních proudů
- kodéry: kódování (komprese) elementárních proudů

POČÍTAČOVÉ ZPRACOVÁNÍ MM

- konvertory: např. konverze barevného prostoru, počtu zvukových kanálů, ...
- mixážní filtry: kombinace několika proudů do jednoho
- replikátory: umožňují rozvětvit proud dat (připojení několika modulů na jeden výstup jiného modulu)
- renderery: koncové moduly zobrazující dekodovaný obsah (vykreslení obrazu, přehrání vzorků zvuku, ...)

POČÍTAČOVÉ ZPRACOVÁNÍ MM

- data se postupně umísťují do bloků paměti
 - blok se samostatným významem: událost
 - událost opatřena časovými známkami
 - je-li to možné, předávají si moduly pouze symbolické odkazy na události (např. předání dat mezi splitterem a dekodérem)
 - je-li třeba data události výrazně změnit (např. před dekódováním a po něm), vytváří modul novou událost a tu posílá symbolicky dál

POČÍTAČOVÉ ZPRACOVÁNÍ MM

- celý graf musí být synchronizován jedněmi hodinami
 - generování časové známky pro aktuální stav všech modulů grafu
 - často se odvozují z audio rendereru
- datový tok obvykle řízen výstupními moduly (metoda pull)
 - renderer chce v čase T zobrazit snímek
 - pokud nemá k dispozici data, požádá o ně předchozí modul
 - ...
- moduly typicky disponují vyrovnávací pamětí
 - kromě aktuálně zpracovávané události mají ve vyrovnávací paměti jednu nebo několik dalších
 - tok dat grafem se musí řídit tak, aby nedocházelo k přetečení ani podtečení vyrovnávacích pamětí

POČÍTAČOVÉ ZPRACOVÁNÍ MM

TVORBA GRAFU MODULŮ

- ručně: podle programátorových požadavků
- poloautomaticky: programátor pouze specifikuje počáteční a koncový modul
- automaticky: podle připravených šablon, např.
„vstup ze souboru, výstup na obrazovku a zvukovou kartu“
- automatické propojování modulů A, B:
 - vstupy a výstupy modulů jsou označeny typem dat (např. nekomprimovaná obrazová data ve formátu YUY2)
 - pokud je typ výstupu A shodný se vstupem B, lze moduly propojit
 - pokud je vstup → výstup nekompatibilní, hledá se sekvence konverzních modulů mezi dostupnými moduly