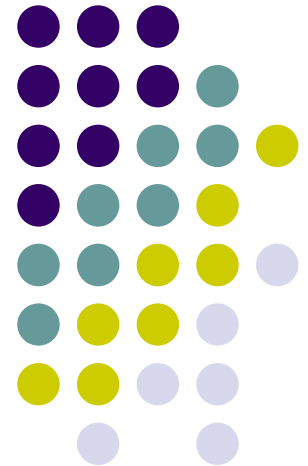
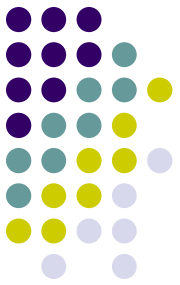


# Sběrnice

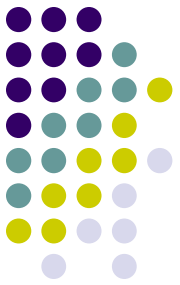
ISA, EISA, VL-BUS, PCI, PCI-E



# Sběrnice



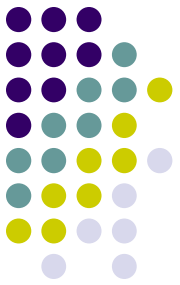
- Pod pojmem **sběrnice** obecně rozumíme soustavu vodičů, která umožňuje přenos signálů mezi jednotlivými částmi počítače.
- Pomocí těchto linek mezi sebou jednotlivé části počítače komunikují a přenášejí data
- Sběrnice představuje nejjednodušší způsob, jak propojit uvnitř počítače jednotlivé komponenty (naopak nejsložitějším způsobem by bylo vytváření dvoubodových spojů pro každou dvojici částí počítače, které budou chtít komunikovat)
- Sběrnice = nejjednodušší komunikační topologie
- Sběrnicí je vše propojeno do jednoho „uzlu“
- Na sběrnici smí být vždy **jen jedna** právě vysílající stanice
- Stanice, které nevysílají, musí mít vývody přepnuté do **stavu vysoké impedance „Z“**, aby na jejich vývodech nebyla ani jednička ani nula, která by mohla kolidovat s vysílanou informací
- Zařízení jako jsou procesor, koprocessor, cache paměť (pokud nebyla integrovaná přímo v procesoru), operační paměť, řadič cache paměti a operační paměti a některá další zařízení bývala propojena tzv. **systémovou sběrnicí**
- Osobní počítače musí být navrženy tak, aby bylo možné jejich snadné rozšiřování o další zařízení
- Takovéto rozšiřování je velmi často uskutečňováno pomocí tzv. **rozšiřující sběrnice** počítače (častěji označované pouze jako sběrnice), na kterou se jednotlivá zřízení zapojují



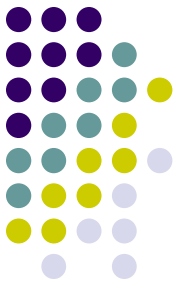
# Sběrnice

- Dle počtu linek, na kterých se současně posílají data, dělíme sběrnice na
  - **Sériové** - data se posílají bit za bitem po jedné lince
  - **Paralelní** - posílá se více bitů naráz po několika linkách vedených vedle sebe

# Sběrnice

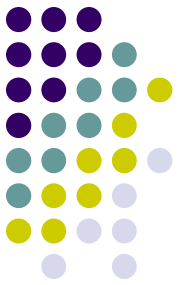


- Dle přítomnosti **hodinového signálu** dělíme sběrnice na
  - **Synchronní sběrnice** - platnost dat na sběrnici jednoznačně určuje hodinový signál.
  - Tímto způsobem dříve pracovala převážná většina všech sběrnic.
  - U synchronních sběrnic uvádíme taktovací **frekvenci** – ta vlastně udává počet přenosů za sekundu
- **Asynchronní sběrnice** – není zde žádný hodinový signál. Příjemce musí správně odhadnout okamžik, kdy má přijmout platná data



# Sběrnice

- Dle možnosti souběžného vysílání oběma směry zároveň dělíme sběrnice na
  - **Poloduplexní** – v jednu chvíli probíhá komunikace pouze jedním směrem
  - **Plně duplexní** – komunikující zařízení může zároveň vysílat i přijímat data současně



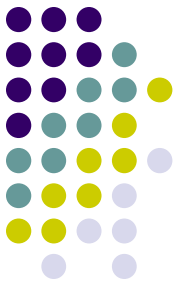
# Sběrnice

- Dělení dle lokalizace
  - **Vnitřní** – propojuje komponenty uvnitř počítače
  - **Vnější** – Slouží k připojování vnějších periferních zařízení (např. USB, SCSI)

# Sběrnice – další pojmy



- **Multimastering** - provoz sběrnice může řídit některá z rozšiřujících karet. Multimaster umožňuje, aby si mezi sebou vyměňovala data dvě periferie
- **Lokální sběrnice** - je připojena přímo na mikroprocesor, je jím řízena a propojuje mikroprocesor s obvody na základní desce
- **Burst režim** - přenáší se po sobě více následujících dat najednou (ve větší skupině), bez adresy u každého datového bloku. Původně se totiž data přenášela v krátkých blocích. Každý blok dat předcházela adresa určení. Burst režim sdruží data směřující na stejnou adresu a přenese je najednou - tím se přenos dat zrychlí



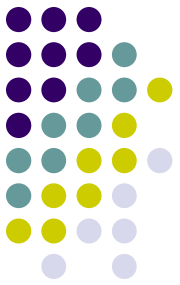
# Sběrnice

- Různé komponenty počítače vyžadují různé rychlosti přenosu dat.
- Z tohoto důvodu by použití jedné sběrnice pro vše bylo neefektivní a pro různé skupiny zařízení jsou proto vytvořeny různé sběrnice nebo komunikační linky

Zařízení	Přibližný datový tok
Klávesnice	20 B/s
Myš	50 B/s
Tiskárna	10 MB/s
Síťové rozhraní (Gigabitový Ethernet)	100 MB/s
Pevný disk	150 MB/s
SSD	500 MB/s
Grafický adaptér	2000 MB/s
Operační paměť (DDR4)	30 000 MB/s



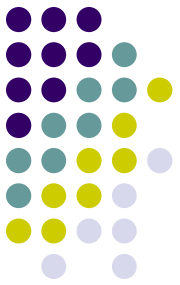
# Sběrnice



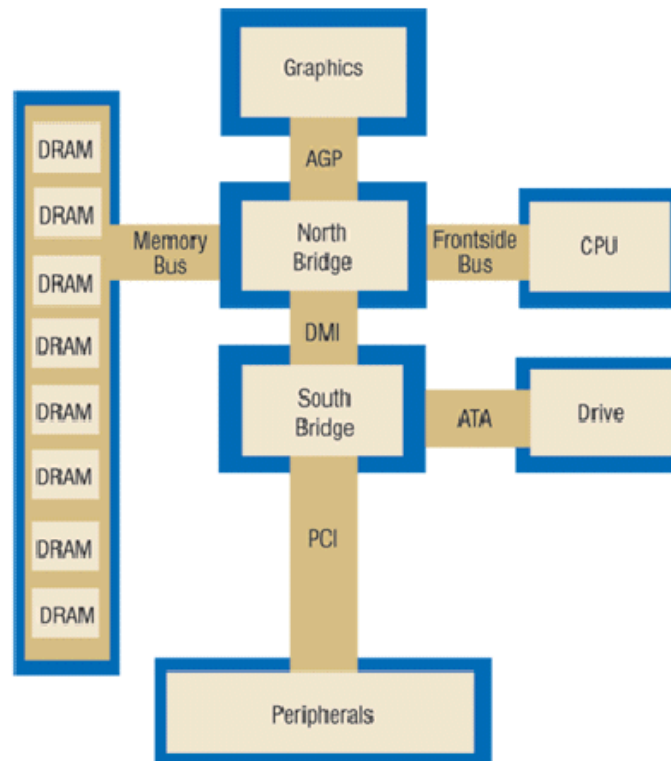
- Sběrnice má obvykle podobu soustavy linek (spojů) umístěných na základní desce PC.
- Aby bylo možno ke sběrnici připojovat různá zařízení, jsou kolmo k těmto spojům připájeny patice (sloty).
- Jednotlivé moduly (karty), které se ke sběrnici připojují, jsou pak umístěny na samostatných deskách a jejich vstupy a výstupy jsou vyvedeny na konektor. Zasunutím konektoru do slotu sběrnice je zařízení fyzicky připojeno ke sběrnici.



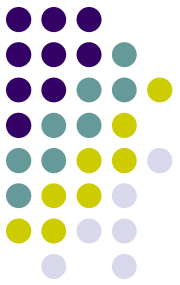
# Sběrnice



- Jednotlivé sběrnice jsou mezi sebou propojeny pomocí můstků (*bridge*) – Například dříve severní a jižní můstek



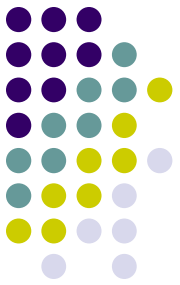
# Sběrnice



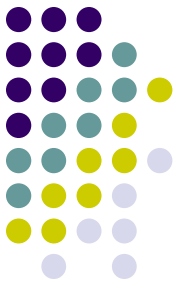
## Základní parametry sběrnice

- **Typ komunikace** – sériová/paralelní, synchronní/asynchronní, polo/plně duplexní
- **Šířka přenosu** – počet datových linek - počet bitů, které lze zároveň po sběrnici přenést
- **Počet vodičů** – je větší než šířka přenosu (nepřenáší se pouze data ale i řídící signály, napájení, hodinový signál...)
- **Propustnost** - Počet bajtů přenesených za sekundu. U synchronní sběrnice se dá vypočítat pokud vynásobíme frekvenci a šířku dat

# Sběrnice

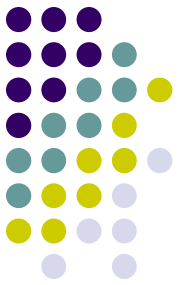


- Původně byly všechny sběrnice v počítačích PC **synchronní, paralelní a poloduplexní**
- **Paralelní sběrnice** je soustava mnoha vodičů, paralelně se generuje velké množství signálů a okamžik jejich platnosti je určen hranou hodinového signálu
- Paralelní sběrnice jsou **poloduplexní**. Přenos je možný v obou směrech, ale nikoliv oběma směry současně. Vysílání a příjem dat se musí v čase střídát. Plně duplexní paralelní sběrnice by vyžadovala dvojnásobný počet datových vodičů (oddělené datové vodiče pro vysílání dat a pro příjem dat) a byla by složitá.
- Signály na paralelní sběrnici lze rozdělit na
  - **Datové** – slouží k přenosu dat z/do připojených zařízení
  - **Adresační** – slouží k adresaci, tzn. k výběru zařízení, pro které jsou data určena. Každé zařízení má svou adresu (číslo brány). V případě grafických karet lze adresovat i jednotlivá místa v jejich videopaměti.
  - **Řídící** – synchronizace komunikace, určení směru komunikace, vyvolání přerušení, reset zařízení apod.
  - **Napájecí**



# I/O adresy

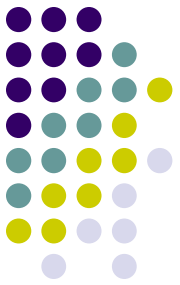
- Každé zařízení komunikující s procesorem má pro tuto komunikaci přidělenou skupinu adres, které nesmí používat žádné jiné zařízení
- například klávesnice má k dispozici adresy 60h a 64h, pevný disk adresy 3E8 až 3EFh
- Těmto adresám se říká adresy vstupu a výstupu - **I/O adresy** nebo **brány**
- **Stejně jako je třeba identifikovat jednotlivá místa v paměti, jsou jednotlivá zařízení identifikována svým číslem brány (I/O adresou)**
- Každá přídatná karta má tedy svoji I/O adresu (obvykle je jich více)
- Dříve bývala tato adresa na kartě napevno, dnes je možno ji v určitém rozsahu měnit



# PC bus

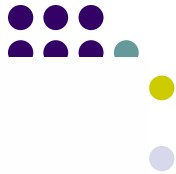
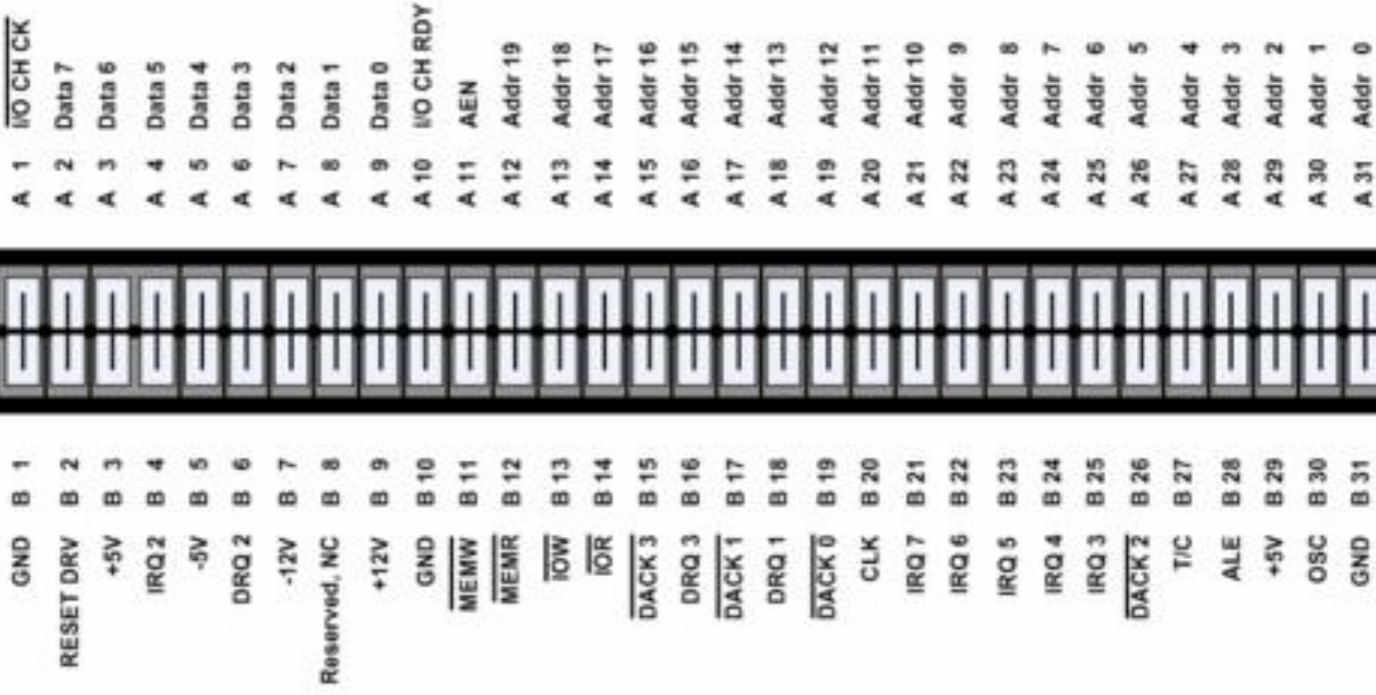
- Nejstarší sběrnice v osobních počítačích. Vychází z ní několik dalších modernějších sběrnice.
- Sběrnice **PC bus** byla navržena a vyrobena firmou IBM pro první počítače IBM PC a IBM **PC/XT**
- **Paralelní, synchronní, polo-duplexní**
- **62 signálů**, z nichž **8** slouží pro přenos dat
- Pro přenos **adresy** je na sběrnici PC bus **20 vodičů**, což odpovídá 20bitové adresové sběrnici procesoru 8086

# PC bus



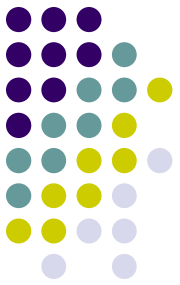
- Další signály:
  - Napájení (+5 V, -5 V, + 12V) a zem
  - Signál pro určení, zda přenášená adresa je adresou místa v paměti nebo adresou (branou) I/O zařízení
  - Signál pro určení, zda data na sběrnici budou čtena nebo mají být zapsána
  - Řídící signály (Reset, hodinové signály)
  - Přerušovací signály (IRQ) určené pro periférie, které potřebují vyžádat okamžité obsloužení procesorem (na sběrnici PC bus je jich 6 a jsou označeny jako IRQ2 - IRQ7)

## 8 Bit XT Bus – top view



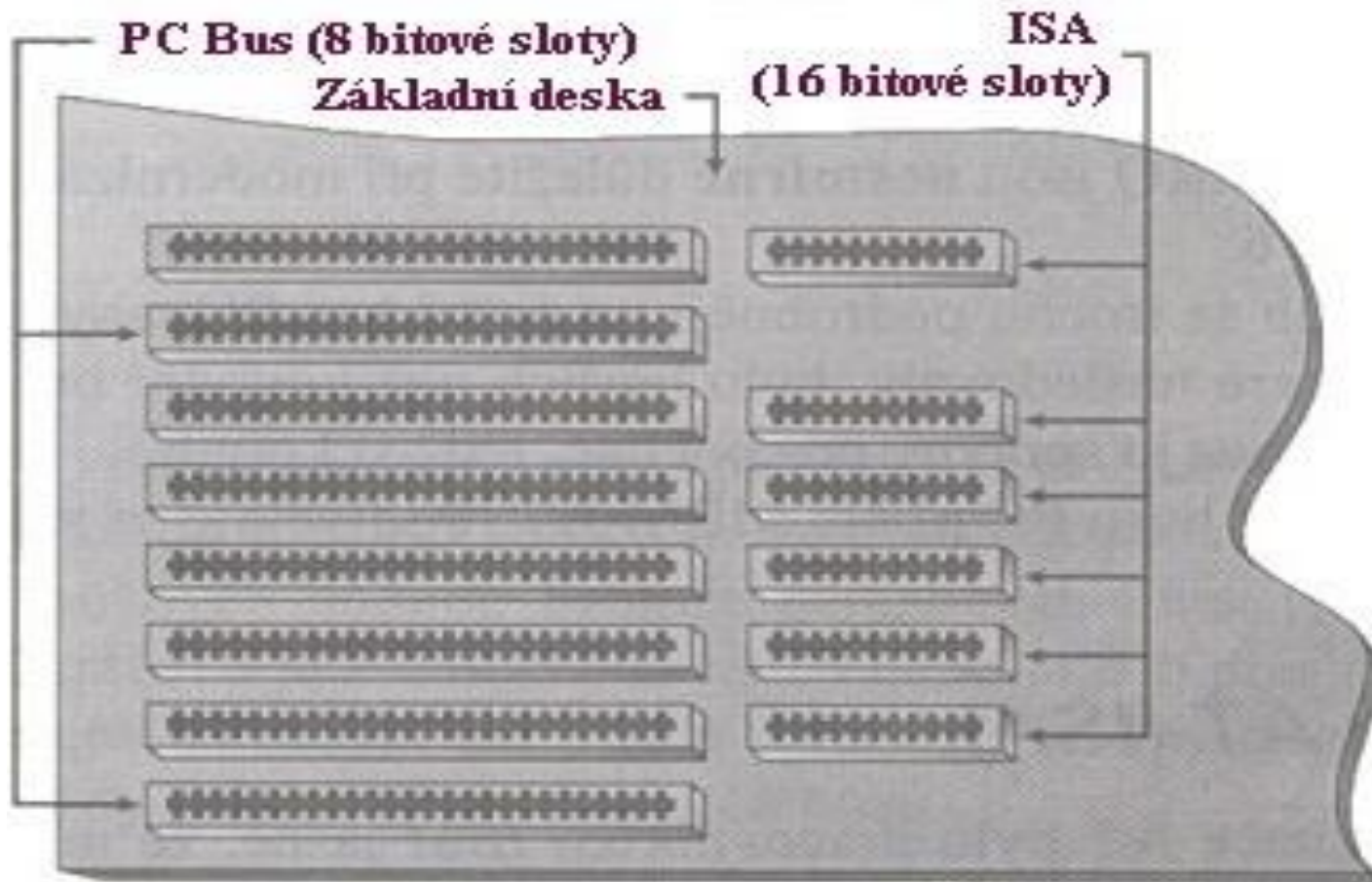
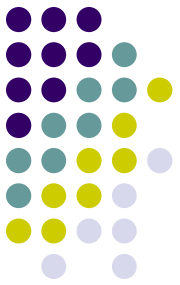


# Sběrnice ISA

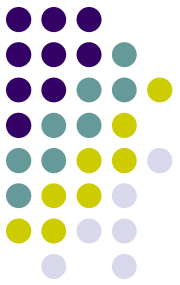


- Industry Standard Architecture
- Vyvinuta firmou IBM pro první počítače PC/AT – jako rozšířená PC bus
- stará 62-vodičová sběrnice se rozšířila o **dalších 36 vodičů** a odpovídající slot se rozšířil o **další konektor** umístěný v jedné řadě hned za starším 8bitovým slotem pro PC bus
- **16 datových** vodičů a **24 adresových** vodičů
- šířka přenosu **16 bitů**, tj. během jedné operace je možné po sběrnici přenášet nejvýše 16bitovou informaci
- Sběrnice je **synchronní** – okamžik platnosti dat je určen hranou **hodinového signálu sběrnice**
- Taktovací frekvence bývala shodná s frekvencí procesoru a postupně vzrostla z 4,7 MHz na max. **8 MHz**

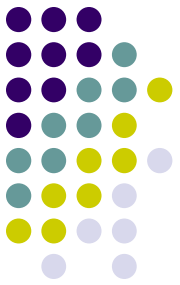
# ISA



# ISA

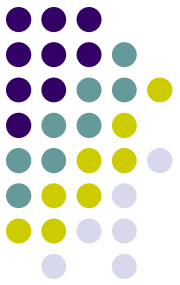


# ISA



- Signály:
  - vodič pro určení, zda přenášená adresa je adresou paměti nebo nějakého jiného vstup/výstupního zařízení (M/IO)
  - vodiče pro určení, zda data na sběrnici byla přečtena nebo mají být zapsána
  - vodiče pro napájení (+5 V, -5 V, + 12V) a elektrická zem
  - vodiče pro řídicí signály (Reset, hodinové signály, signály pro refresh paměti)
  - vodiče pro vyvolání přerušení (IRQ) určené pro periferie, které potřebují pro svou správnou činnost vyžádat okamžité obsloužení procesorem. Těchto linek je na sběrnici 11 a jsou označeny jako IRQ2-7, IRQ10-12 a IRQ14-15
  - Signály pro řízení přímého přístupu do paměti (DMA)

# Sběrnice ISA



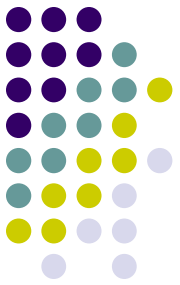
- Sběrnice ISA byla používána u většiny počítačů s procesorem **80286**, **80386** a u starších počítačů s procesorem **80486**
- Vzhledem k velkému množství přídavných karet, které jsou vyrobeny pro tuto sběrnici, byla ISA používána společně s jiným typem sběrnice i v modernějších počítačích
- Protože procesory 80286 a vyšší byly vyráběny s frekvencemi 16 MHz a vyšší, bylo nutné provádět vzájemné přizpůsobování rychlosti (Procesor je rychlejší než sběrnice)

# Sběrnice MCA



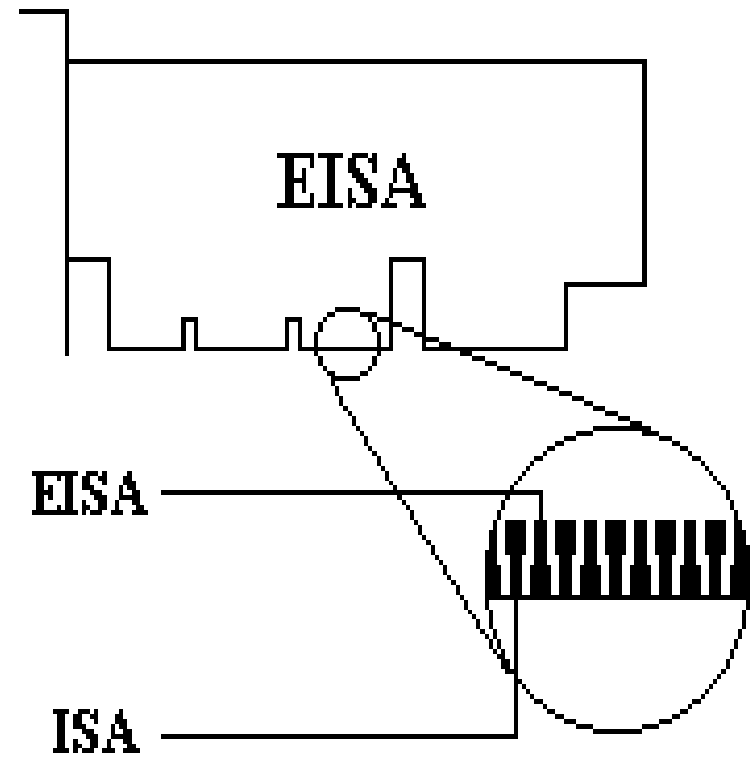
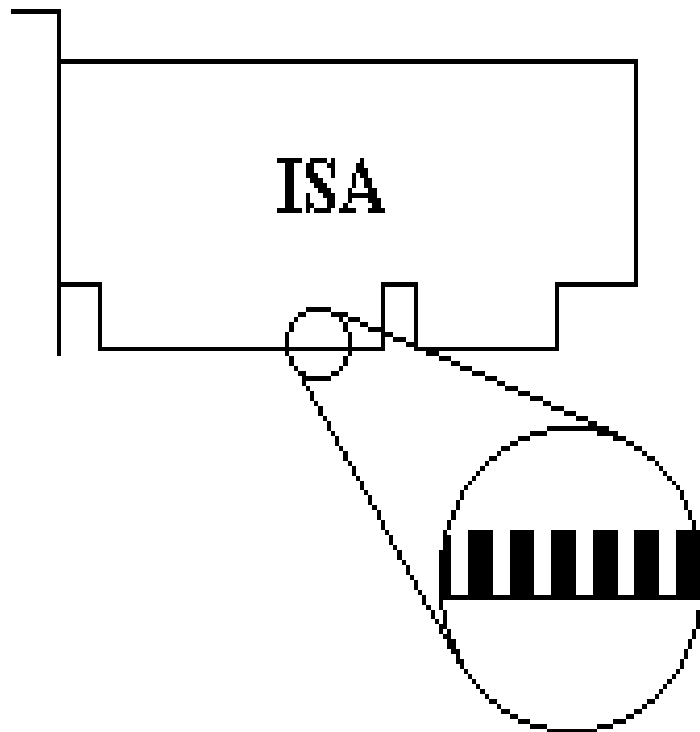
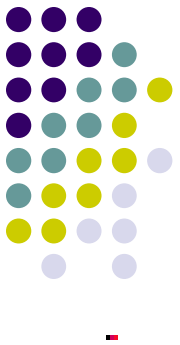
- Sběrnice **MCA** (MicroChannel Architecture) byla vyvinuta pro řadu počítačů firmy IBM s označením **IBM PS/2**
- Sběrnice MCA dovoluje běh s frekvencí **10 MHz** a přenáší data po **16 i 32 bitech**
- Sběrnice MCA nerozšířila kvůli **nekompatibilitě** s ISA a komplikovaným podmínkám licencování
- MCA má i tzv. **proudový režim**, ve kterém dokáže současně přenášet 64 bitů (po výběru adresy zařízení lze přenést 4x 16 bitů bez nutnosti opakovaně provádět další adresaci)
- Šířka adresové části je v závislosti na procesoru počítače 24 bitů (pro 80286) nebo 32 bitů (pro 80386)

# EISA



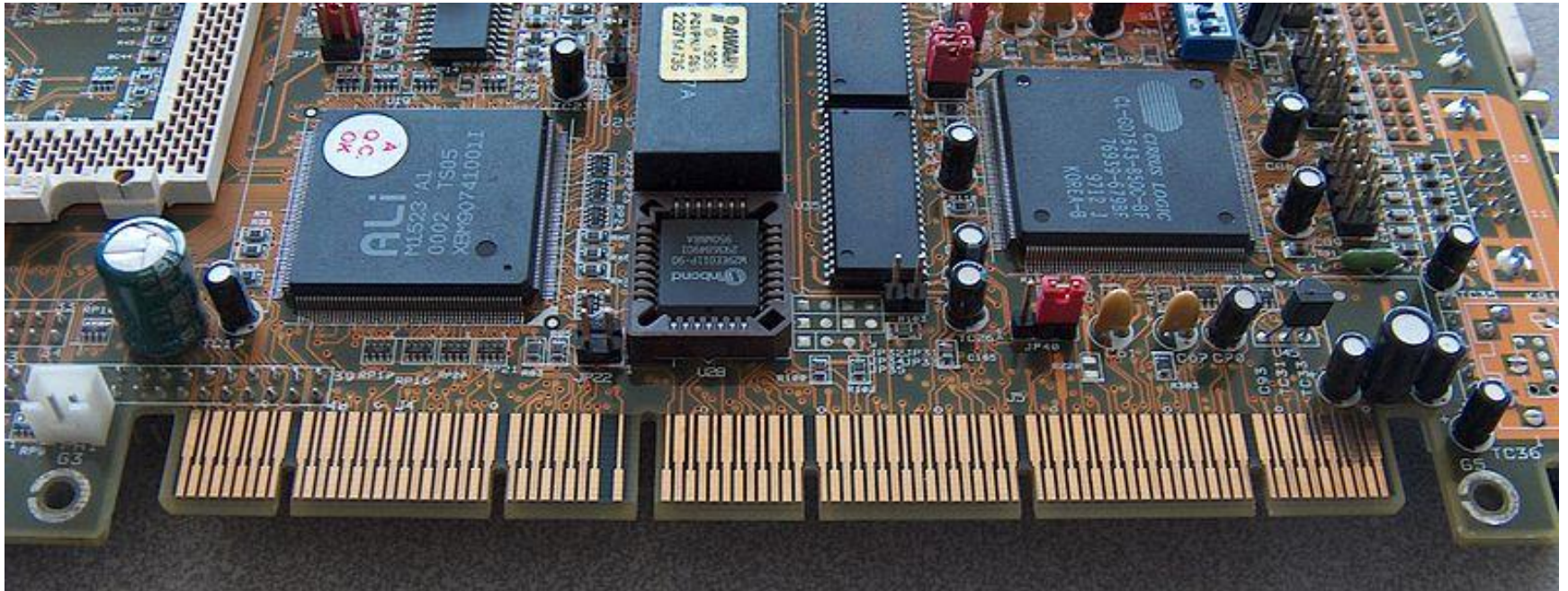
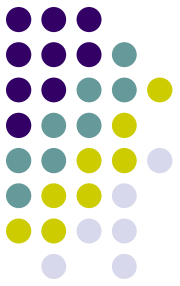
- Sběrnice **EISA** (Extended Industry Standard Architecture) byla vyvinuta 9 firmami (AST Research, Compaq, Epson, NEC, Olivetti, Tandy, Wyse a Zenith) v roce 1989
- šířka toku dat je **32 bitů**
- šířka adresy je **32 bitů**
- pracuje s frekvencí **8 MHz** (z důvodů kompatibility s ISA)
- dovoluje busmastering
- Slot sběrnice EISA má stejnou velikost jako slot ISA a obsahuje stejné vývody (62+36)
- Kromě těchto vývodů má ještě **59 nových vývodů** umístěných mezi starými vývody sběrnice ISA
- Tyto nové vývody zůstanou v případě zasunutí karty pro ISA sběrnici nezapojeny, čímž se dosahuje zpětné kompatibility EISA s ISA
- Sběrnice EISA byla používána zejména u počítačů s procesory **80386 a 80486**

# EISA

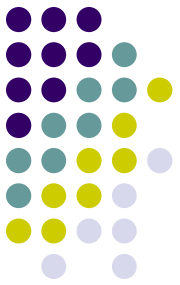




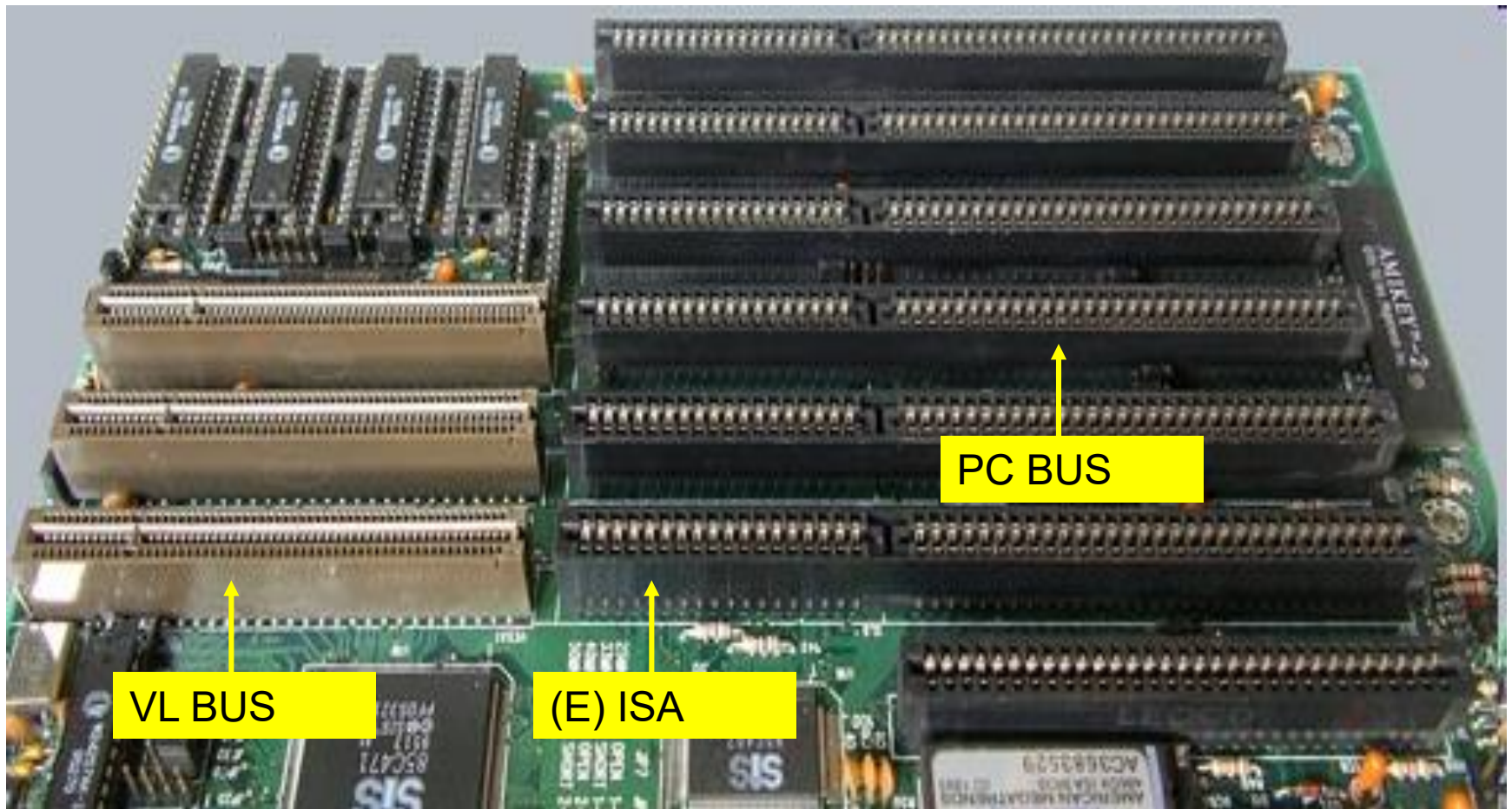
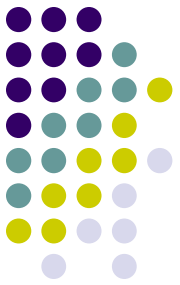
# EISA



# VL Bus

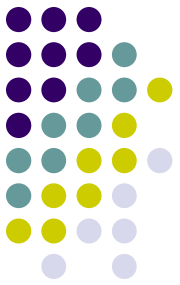


- Sběrnice **VL bus** (VESA Local Bus) byla navržena v roce 1992 konsorciem **VESA** (Video Electronic Standards Association)
- Byla navržena v době, kdy propustnost sběrnice EISA přestává stačit pro přenos dat do grafické karty
- Hlavní konstrukční vlastnost je, že **přebírá vývody** z procesoru 486 a přenáší je na konektory karty periferie – tedy je **lokální**
- **Silná procesorová závislost** způsobená přímým zapojením slotů VL busu na systémovou sběrnici procesoru 80486
- Tato závislost nedovoluje prakticky použít VL bus v jiném počítači, než je počítač s procesorem **Intel** nebo kompatibilním
- Další nevýhodou je práce s úrovněmi logiky TTL (5 V), zatímco nové procesory již pracují s napětím 3,3 V a nižším
- Konektor VL busu se nachází **v jedné řadě** za 16bitovým konektorem ISA a má **2x58 vývodů** (při své práci využívá některých signálů ISA sběrnice)
- Sběrnice VL bus byla používána zejména u počítačů s procesorem **80486** a prvních počítačů s procesorem **Pentium**
- Typicky se používala pro připojení grafické karty (umožňovala větší propustnost než EISA)

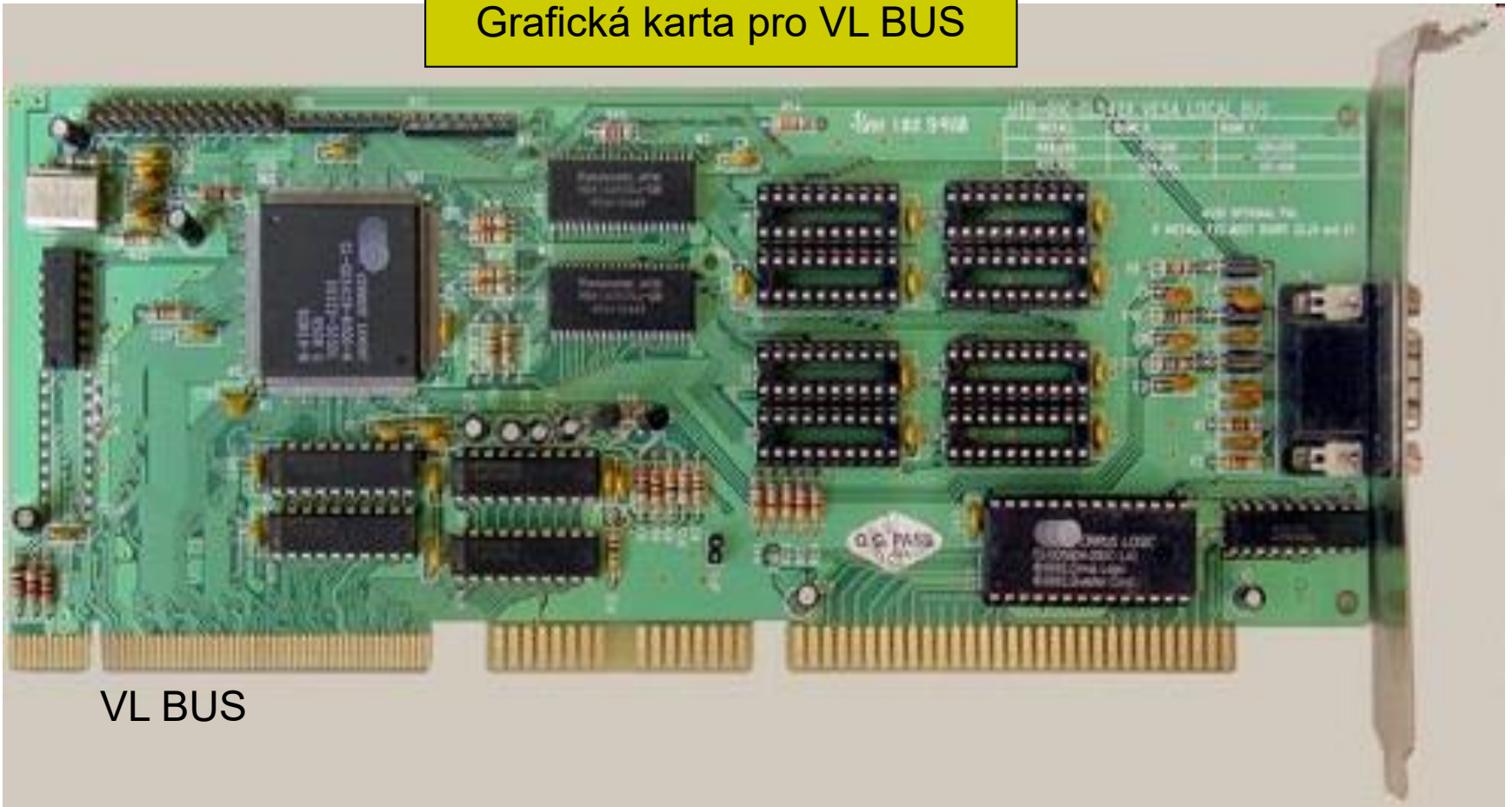




# VL BUS

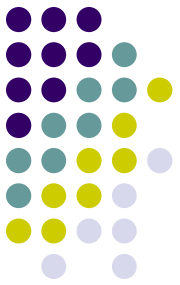


Grafická karta pro VL BUS



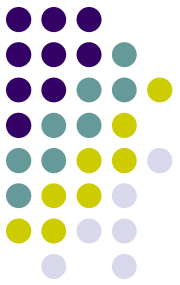
VL BUS

# PCI



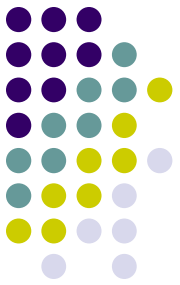
- **PCI** = Peripheral Component Interconnect
- navrhnutá firmou **Intel** pro počítače s procesory **Pentium**
- Používá se od roku 1994
- PCI již není klasickou **lokální sběrnici** jako VL bus, ale je k **systémové sběrnici** připojena přes tzv. **mezisběrníkový můstek**
- Toto řešení s sebou nese následující výhody
  - Možnost použití sběrnice PCI i v jiných počítačích než jsou PC (např. Macintosh, DEC, Alpha, PowerPC)
  - **Funguje v počítači s libovolným procesorem**
  - Můstek dovoluje provádět přizpůsobování napěťových úrovní (na rozdíl od VL BUS, kde záleželo na napájecím napětí procesoru)
- Šířka přenosu dat je **32 bitů** nebo **64 bitů**
- Při frekvenci **33 MHz** je propustnost sběrnice **132 MB/s** (32 bitů) nebo **264 MB/s** (šířka 64 bitů)

# PCI



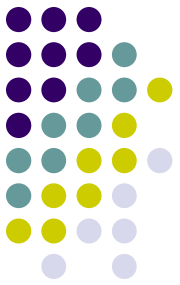
- PCI nemá oddělené adresní linky a datové linky - vodiče (32 nebo 64) slouží střídavě pro přenos dat i adres, adresa se posílá na začátku každé transakce
- časový multiplex – role vodičů se v čase střídá = **multiplexované adresové a datové vodiče** AD0, AD1, AD2....AD31 nebo AD63
- Přenosovou kapacitu sdílí všechna připojená zařízení (a např. jen síťová karta gigabitového Ethernetu může generovat datový tok 100 MB/s)

# PCI



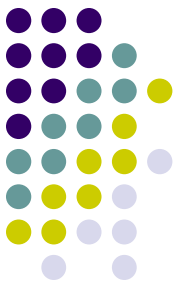


# PCI



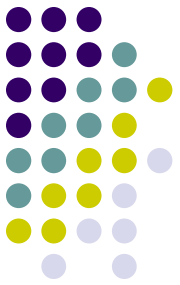
- PCI umožňuje **burst režim** a narozdíl od VL busu **není shora omezen** počtem 4 taktů
- podporuje **busmastering**, což dovoluje vyšší výkon počítače, protože přenosy řízené CPU vyžadovali spoustu procesorového času
- podporuje **Plug and Play** - dovoluje velmi snadnou konfiguraci desek pro PCI, popř. jejich automatickou konfiguraci bez zásahu uživatele





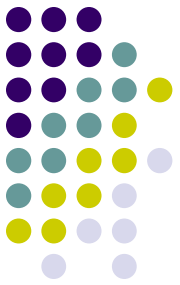
# PnP (Plug and Play)

- Na sběrnici **ISA**, bylo při přidání nové karty do systému nutné, aby její majitel nastavil nejprve kartu a pak stejně i její ovladač (nastavení IRQ, DMA, číslo brány ...)
- Zároveň bylo nutné, aby žádné dvě karty nebyly **v konfliktu** (nebyly nastaveny stejně nebo tak, že se jejich nastavení překrývalo)
- Přidat do počítače kartu tak dříve nebylo vůbec jednoduché
- **PnP – Plug and Play**
- Umožňuje automatickou konfiguraci karty
  - karta je vsunuta do slotu vypnutého počítače (připojování za běhu počítače není možné)
  - po zapnutí počítače je aktivována PnP část BIOSu
  - počítač rozpozná zasunutou kartu
  - postupně vyzve všechna zařízení připojená ke sběrnici k identifikaci
  - zařízení odesílají své identifikátory a požadavky
  - Počítač tak zjistí, které IRQ, DMA, I/O Adresy, RAM adresy potřebuje a které může použít
  - Automaticky se nastaví prostředky tak, aby nedošlo k žádnému konfliktu s již existujícími kartami,
  - OS vyhledá a nakonfiguruje potřebné ovladače, které by s nově nainstalovanou kartou spolupracovaly.



# PCI - přerušení

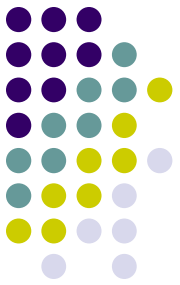
- K dispozici jsou pouze čtyři přerušovací signály označené **INTA#**, **INTB#**, **INTC#** a **INTD#**, které jsou v PCI mostu asociovány s vnějšími přerušeními procesoru
- Fyzicky jsou jednotlivé linky přerušení na sběrnici zapojeny **kaskádovitě**, takže INT#A je v prvním slotu na první pozici, ve druhém slotu na druhé pozici atd...
- všechny z nich jsou dostupné každému zařízení a při inicializaci sběrnice mohou jednomu zařízení být přiřazena až čtyři přerušení, což však není běžné
- Pozdější revize PCI specifikací přidávají podporu pro **přerušení signalizované zprávou** (ne signálem)



# PCI - revize

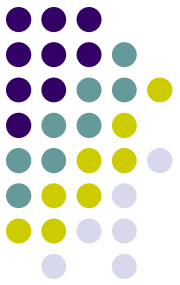
- Následné revize PCI přidaly nové vlastnosti a vylepšení výkonu - frekvence vzrostla **66 MHz**, používané napětí kleslo na **3,3 V**
- Běžné PCI karty mají buď jeden nebo dva **klíčovací zářezy** podle používané napěťové úrovně
- Karty vyžadující 3,3 V mají zářez vedle přední strany karty (kde se nacházejí externí konektory) zatímco ty, vyžadující 5 V mají zářez na druhé straně. Takzvané univerzální karty mají oba zářezy a mohou přijímat oba typy signálů
- **PCI 2.2** umožňuje 66 MHz signalizaci (vyžaduje 3,3 V signalizaci - nejvyšší možná rychlost přenosu 533 MB/s)
- **PCI 2.3** dovoluje použít 3,3V a univerzální klíčování, ale nedovoluje klíčování pro 5V
- **PCI 3.0** je konečný oficiální standard pro PCI sběrnici, byla úplně odstraněna možnost 5V

# PCI - X

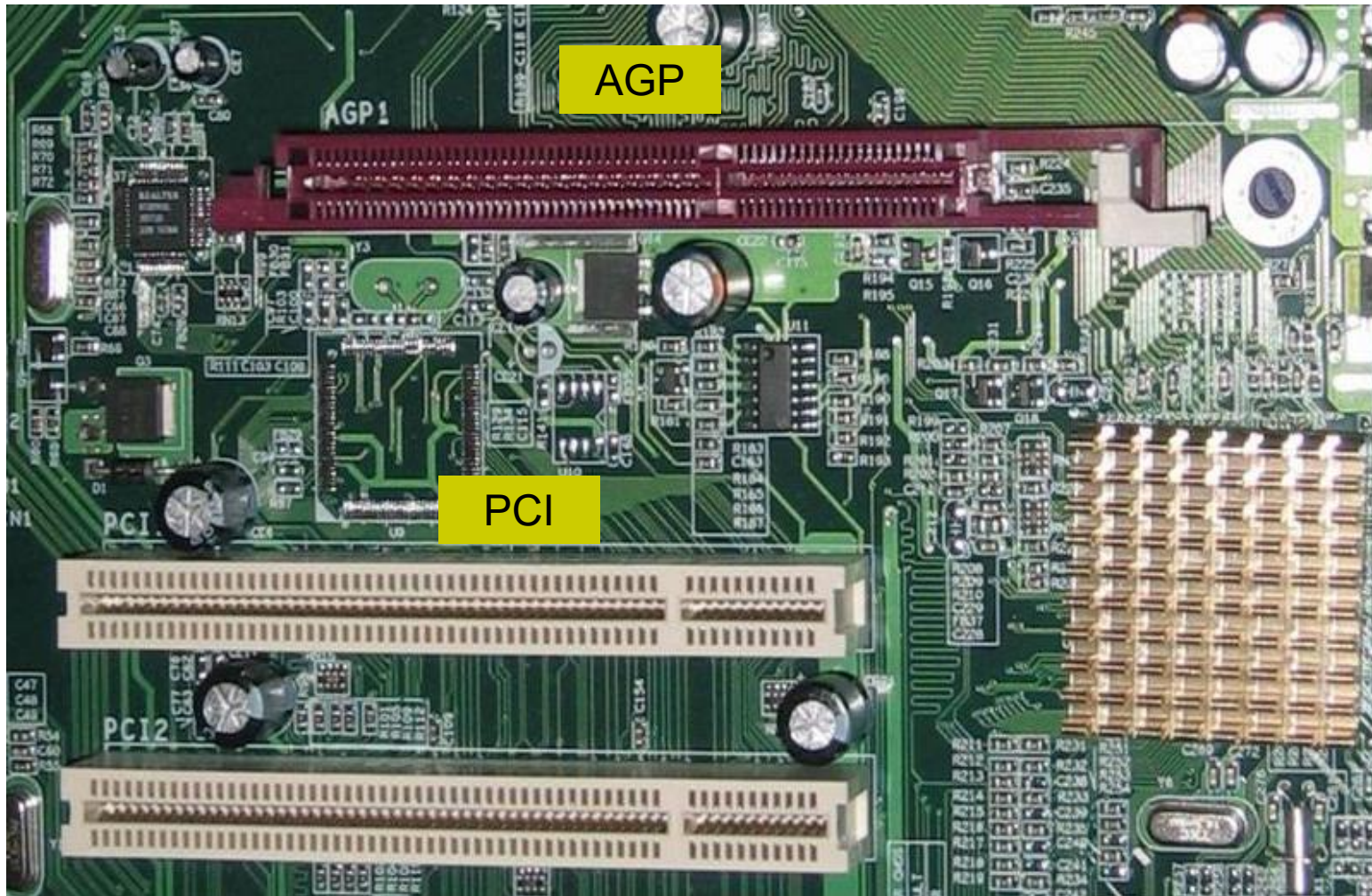
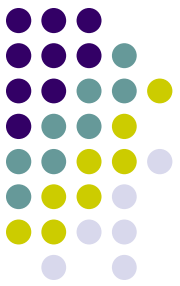


- V roce **1998** navržena jako nástupce sběrnice **PCI 2.3** (firmy HP, Compaq, IBM)
- Šířka datové sběrnice **64b**
- zvýšená maximální frekvenci na **133 MHz** (z původní 66 MHz)
- V dalších revizích frekvence rostla až na **533 MHz** (to umožňuje datový tok až **4,3 GB/s**)
- Paralelní přenos na takto vysoké frekvencí je již extrémně problematický z hlediska **rušení, přeslechů a synchronizace**
- Proto u posledních revizí této sběrnice musel být zaveden **ECC - Error Correction Code**, který automaticky opravuje chyby vzniklé při přenosu dat
- Další zrychlení paralelní komunikace se ukázalo jako nemožné – nelze zvýšit frekvenci ani přidat další paralelní datové linky
- Sběrnice **PCI-X** nenahradila sběrnici **PCI**, protože se dostatečně nerozšířila
- Nástupcem sběrnice **PCI** tak byla až sběrnice **PCI-Express**
- Karty **PCI-X** je možné použít v **PCI** slotu, jen na základních deskách, které podporují napájení **3,3V**
- Ve sběrnici **PCI-X** mohou pracovat **3,3V PCI** karty
- Karty **PCI-X** a **PCI** mohou být používány na **PCI-X** sběrnici současně, ale rychlost je pak limitována nejpomalejší kartou

# AGP

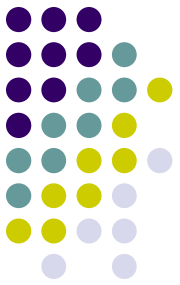


- **Accelerated Graphics Port**
- sběrnice **PCI** přestaly stačit nárokům **3D grafiky**
- V roce 1998 Intel vyvinul novou vysokorychlostní sběrnici věnovanou výhradně grafice a nazval ji **AGP**, ale nejedná se v pravém slova smyslu o sběrnici, neboť k AGP lze připojit pouze **jedno zařízení**
- propojuje dvoubodově základní desku a **grafickou kartu**
- Základní pracovní frekvencí sběrnice AGP je **66 MHz**
- Technologie **AGP** vznikla úpravou sběrnice **PCI**, některé řídicí signály jsou podobné (adresová a datová část AGP používá stejné vodiče pro přenos jako PCI)
- Odstraněn je **arbitrážní obvod** (ten se stará a přidělování sběrnice a bezkolizní provoz) - proto je možné připojit pouze jedno zařízení



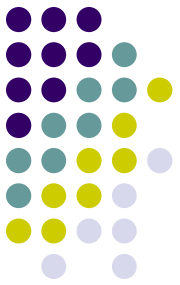


# AGP



- **AGP 1.0**
  - 1x (266 MB/s) (8 B přeneseno za dva takty)
  - 2x (533 MB/s) (8 B přeneseno během jednoho taktu)
- **AGP 2.0**
  - 1x (266 MB/s) (8 B přeneseno za dva takty)
  - 2x (533 MB/s) (8 B přeneseno během jednoho taktu)
  - 4x (1.07 GB/s) (16 B přeneseno během jednoho taktu)
- **AGP 3.0**
  - 4x (1.07 GB/s) (16 B přeneseno během jednoho taktu)
  - 8x (2.1 GB/s) (32 B přeneseno během jednoho taktu)

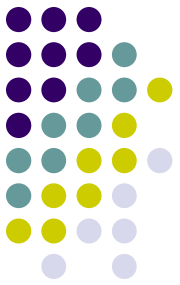
# PCI - Express



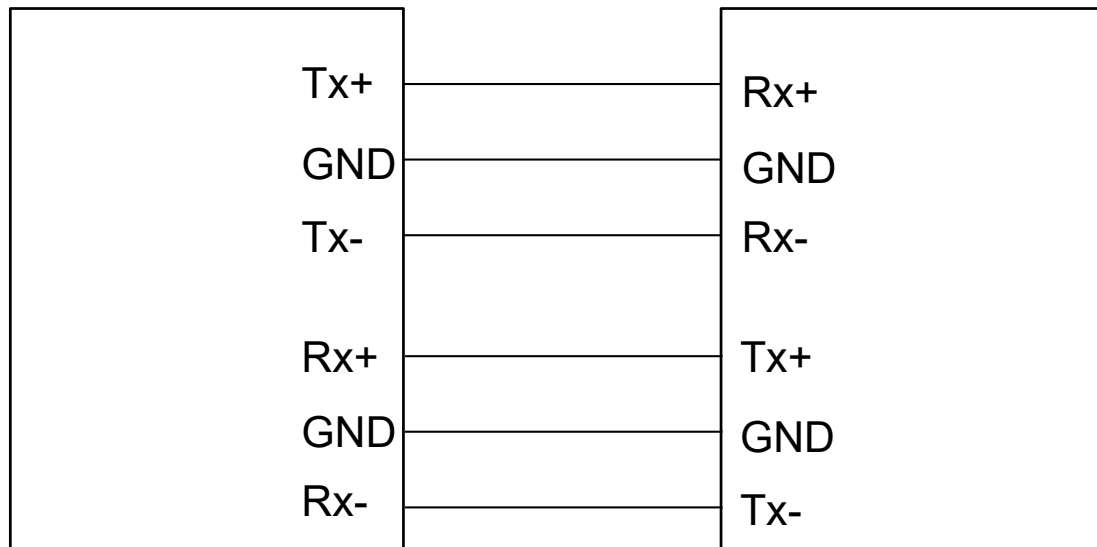
- Klasická sběrnice PCI byla paralelní a poloduplexní - všechny vodiče slouží pro přenos dat oběma směry, ovšem nikoli oběma směry zároveň, přenáší se několik bitů naráz po několika paralelní vodičích
- Další zrychlování paralelního přenosu se ukázalo je velmi obtížné a neefektivní – na PCI-X se dospělo do bodu, kdy propustnost nešla dále zlepšit
- V roce 2004 se **PCI-Express** stává **novým standardem** pro osobní počítače
- **PCI Express** přechází na **sériovou komunikaci**
- **Sériová komunikace** umožňuje dosažení velmi vysokých frekvencí (odpadají potíže s rušením přeslechy a synchronizací mnoha vodičů, kde hodiny tikají pomalu, až když se ustálí signál na všech linkách)
- Sběrnice **PCI-E** je **sériová, plně duplexní a asynchronní**
- Sběrnice PCI Express existuje v několika variantách, které se liší počtem sériových linek. Každá jednotlivá sériová linka je tvořena několika vodiči, aby bylo možné přenášet data současně oběma směry (plně duplexní provoz) a aby se zamezilo rušení
- Signál se vysílá s nízkým napětím 0,8 V



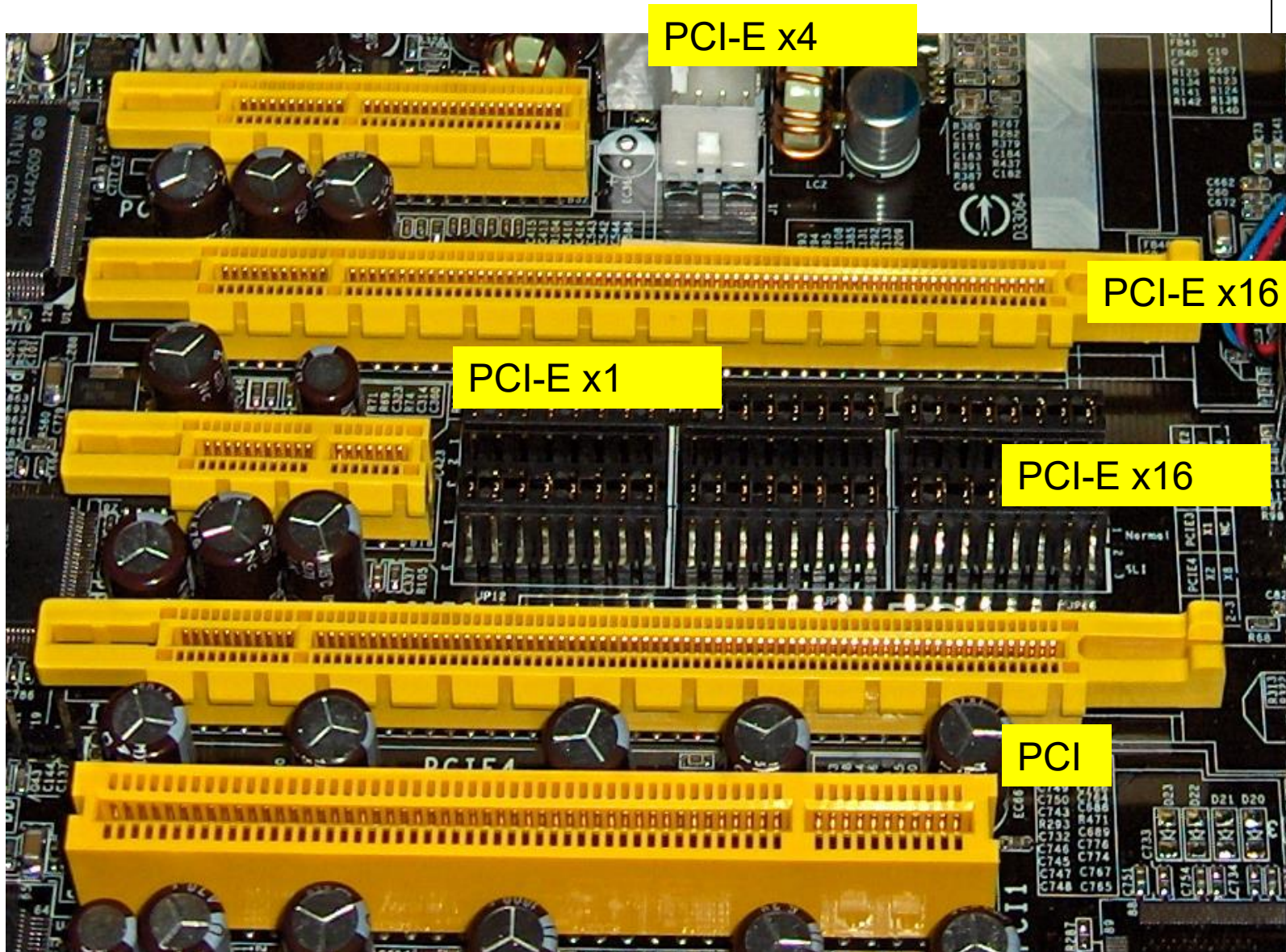
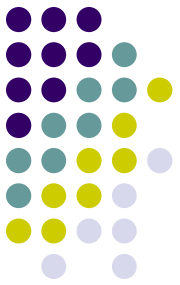
# PCI-E x1



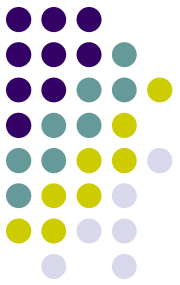
- Základní modul ("x1") sběrnice PCI-Express má **jediný sériový fullduplexní kanál**
- Základní modul sběrnice tvoří **jeden symetrický pár vodičů** pro TX (transmit) a druhý pro RX (receive)
- Vyšší násobky (x2..x32) vznikají paralelním vedením několika těchto základních jednobitových kanálů



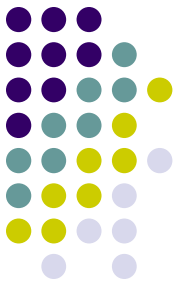
Konektory PCI-E pro různé "násobky" jsou různě dlouhé, nicméně mají společnou rozteč vývodů a "ideově jednotnou" modulární stavbu



# PCI - E



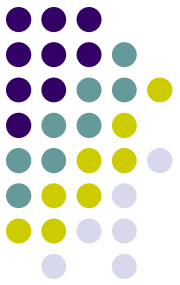
- Data na jednotlivém kanálu jsou kódována systémem **8b10b**, takže jeden užitečný bajt je transformován na desetibitovou posloupnost – zaručen dostatečný počet změn signálu a vyvážený poměr nul a jedniček
- Blok několika bajtů odeslaných za sebou tvoří **rámec**
- Každý **rámec** je opatřen **adresou odesilatele, adresou příjemce, sekvenčním číslem a kontrolním součtem** – komunikace se podobá počítačové síti
- PCI-E každou **transakci** rozdělí do dvou rámců, tj. na požadavek a odpověď
- Konektor sběrnice PCI-E úplně **postrádá** vodiče pro tradiční **signály přerušení**
- Sběrnice PCI-E uplatňuje nový způsob doručování přerušení - doručování pomocí zpráv (tzv. message-signaled interrupts)



# PCI - Express

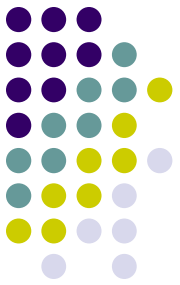
- **Propustnost**
- Datové toky uváděné v případě PCI-E obvykle znamenají propustnost jedním směrem. V případě efektivního využití plného duplexu je pak celkové množství současně přenesených dat dvojnásobné
- Takt = **2,5 G taktů /s.**
  - PCI-E x1 = 250 MB/s (1B = 10 bitů kvůli 8b10b)
  - PCI-E x4 = 1 GB/s
  - PCI-E x8 = 2 GB/s
  - PCI-E x16 = 4 GB/s

# PCI - E

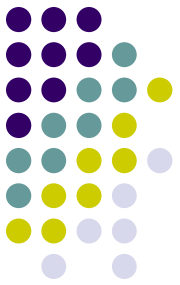


- Komunikace je **vždy sériová**
- Jednotlivé bity bajtu jsou transformovány kódem 8b10b na nových 10 bitů a tyto bity jsou odeslány postupně za sebou **sériově po jedné lince**
- Odesílaný bajt je součástí **rámce**, který má hlavičku, datový obsah a kontrolní součet
- Celý rámec se vysílá **sériově po jedné jediné lince**
- PCI-E x2 používá stále sériovou komunikaci – dvě nezávislé sériové komunikace probíhající **vedle sebe**. Na každé z linek probíhá vysílání jiného rámce
- PCI-E x8 je stále **sériovou komunikací**. O paralelní komunikaci by se jednalo v případě, že by jednotlivé bity bajtu byly odeslány současně přes osm paralelních linek, ale tak tomu zde není.
- PCI-E x8 umožňuje 8 zcela nezávislých sériových komunikací. Ve stejný okamžik může být po osmi linkách odesíláno osm různých rámců. Komunikace na jednotlivých linkách je časově zcela na nezávislá. Zahájení odesílání rámců i jednotlivých bitů probíhá na jednotlivých linkách v různých okamžicích.
- PCI-E x16 = 16 nezávislých sériových linek

# PCI – E



- **PtP (point-to-point)** funkce dovoluje zařízením na sběrnici komunikovat jednomu s druhým, aniž by při tom zatěžovaly procesor
- Grafické karty hlavních výrobců byly postupně předělávány z **AGP** na **PCI-Express**
- většina grafických karet pro PCI-E je ve variantě pro **PCI-E 16×**
- **PCI-E 1× a 4×** se používají pouze pro zvukové karty a další zařízení, která nepotřebují velký datový tok

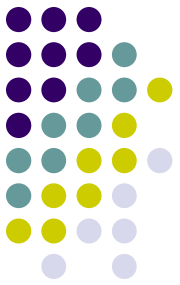


# PCI-E 2.0

- Dvojnásobná frekvence – **5 G taktů /s**
- přenosová rychlost jedné linky zvýšena z **250MB/s** na **500 MB/s**
- Tato verze je zpětně i dopředně **kompatibilní**, můžete tedy karty s podporou PCI-Express 2 zapojit do základní desky, která obsahuje pouze podporu verze 1.1 a naopak
- Propustnost PCI-E (2.0) je
  - 1× - 500 MB/s (obousměrný součtový provoz 1 GB/s)
  - 4× - 2 GB/s
  - 8× - 4 GB/s
  - 16× - 8 GB/s



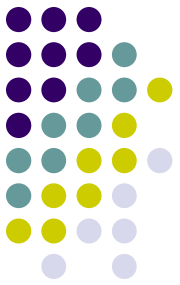
# PCI-E 3.0



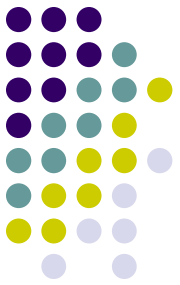
- Nová verze sběrnice již nepoužívá jednoduché kódování 8b10b, ale mnohem složitější a přitom efektivnější metodu pro zajištění dostatečného počtu změn signálu – **scrambling**
- Původních 128 bitů se převádí na 130 bitů (převod není definován kódovací tabulkou, ale matematickým algoritmem, který pro stejných 128 bitů může dát pokaždé jiný výsledek v závislosti na předchozím postupu vysílání)
- *Pro zajímavost případný kód 128b130b by byl definován hypotetickou tabulkou, které by musela mít  $2^{128}$  řádků (tj. 340 000 triliard)*
- Propustnost je přibližně dvojnásobná oproti PCI-E 2.0
- Jedna linka v jednom směru přenese téměř **1 GB/s** při rychlosti 8GT/s



# PCI-E 4.0 a 5.0



- V dalších generacích sběrnice PCI-E se opět vždy **zdvojnásobila** propustnost díky zvýšení rychlosti odesílaných bitů
- Na sběrnici **PCI-E 4.0** jsou data na jedné lince odesílána sériově rychlostí **16 Gb/s**. Propustnost jedné linky v jednom směru je pak **2 GB/s**
- Na sběrnici **PCI-E 5.0** jsou data na jedné lince odesílána sériově rychlostí **32 Gb/s**. Propustnost jedné linky v jednom směru je pak **4 GB/s**.
- Kdyby linka byla synchronní, hodinový signál by zde měl frekvenci 32 GHz. Tedy frekvence odesílání jednotlivých bitů je několikanásobně vyšší než frekvence mikroprocesoru. (Linka je ale samozřejmě asynchronní, bez hodin)
- Je zde vidět, že moderní technologie již umožňují výrobu chipů s velmi vysokou taktovací frekvencí. Bohužel mikroprocesor je příliš složitý na to, aby jako celek (miliardy tranzistorů) mohl běžet na takové frekvenci a při tom se dal uchladiť.



# PCI-E x16

- Moderní grafické karty komunikují přes slot PCI-E x16
- Celková propustnost všech šestnácti linek v jednom směru je pak
- PCI-E 1.0      4 GB/s
- PCI-E 2.0      8 GB/s
- PCI-E 3.0      16 GB/s
- PCI-E 4.0      32 GB/s
- PCI-E 5.0      63 GB/s (tady již započítáváme i vliv kódování – scramblingu 128b130b)
- Kromě grafických karet se v současné době sběrnice PCI-E používá také k připojování rychlých SSD.
- SSD se připojují k PCI-E **x4**
- Pomalé SSD se připojují přes SATA (rozhraní pro klasické pevné disky)

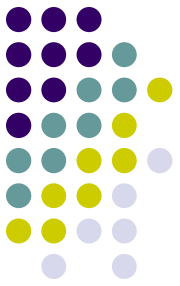
# Sběrnice



Nejvýznamnější vnitřní sběrnice osobních počítačů

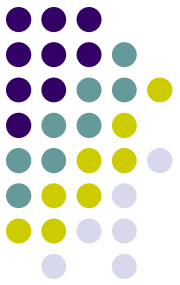
Označení	Bitová šířka	Frekvence (MHz)	Procesory
PC BUS	8	4,77	8088, 8086, 80186
ISA (AT BUS)	16	8	80286 až 80486
MCA	32	8	80286, 80386
EISA	32	8	80386, 80486
VL-BUS	32	33–50	80486 až Pentium
PCI	32/64	33–66	Pentium a vyšší
PCI-X	64	až 533	Pentia
PCI Express 1.0	1 (probíhá více spojení současně)	Asynchronní (2500)	Pentium4, AMD Athlon a vyšší

# Komunikační cesty na základní desce

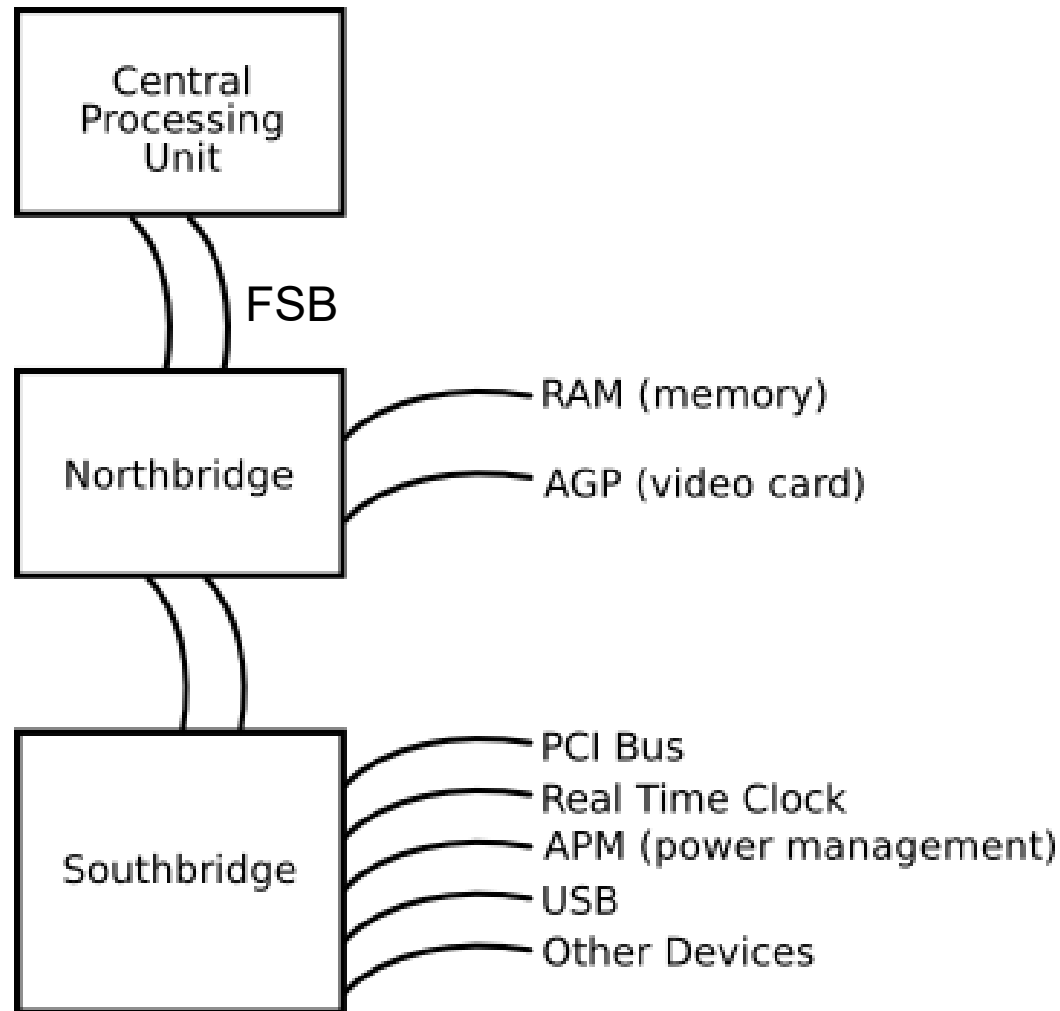
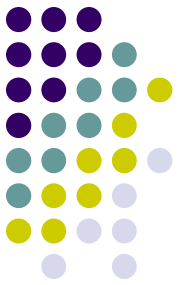


- **Chipset** – sada podpůrných obvodů okolo mikoprocessoru
- **Northbridge** - Severní můstek- je také znám jako systémový řadič
- **Southbridge**- jižní můstek - je také znám jako vstupně-výstupní řadič (I/O Controller Hub)

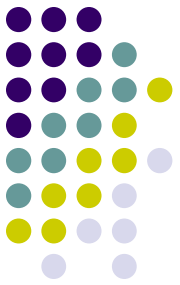
# FSB



- **Front Side Bus**
- Soustava vodičů vedoucích přímo z **vývodů mikroprocesoru**, kterými komunikuje s nejbližším okolím (severním můstkem)
- Jde v podstatě o synchronní paralelní poloduplexní lokální systémovou sběrnici
- Již od procesoru 80486 je běžné, že frekvence, na které zde probíhá komunikace, je **nižší** než vnitřní taktovací frekvence mikroprocesoru (viz *overdrive*)
- Přenos mnoha bitů paralelně vysokou rychlostí je možný pouze v miniaturním prostředí uvnitř mikroprocesoru (přesuny mezi registry a cache) a navenek není možné tuto rychlost udržet i pro komunikaci mikroprocesoru s okolím – To je také mimo jiné důvodem, proč moderní architektura SandyBridge obsahuje integrovaný grafický čip a severní můstek přímo v mikroprocesoru
- Taktovací frekvence mikroprocesoru je odvozena od frekvence FSB
- Moderní BIOS umožňuje nastavit **frekvenci FSB** a hodnotu **násobiče** taktovací frekvence pro mikroprocesor
  - například: FSB 200 MHz, násobič 9x → Mikroprocesor poběží na 1800 MHz
- FSB tvoří komunikační cestu mezi mikroprocesorem a **chipsetem**
- **Chipset** a jeho severní a jižní můstek je spojovacím bodem pro všechny ostatní sběrnice (komunikační cesty) v počítači

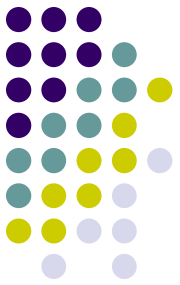


# Northbridge



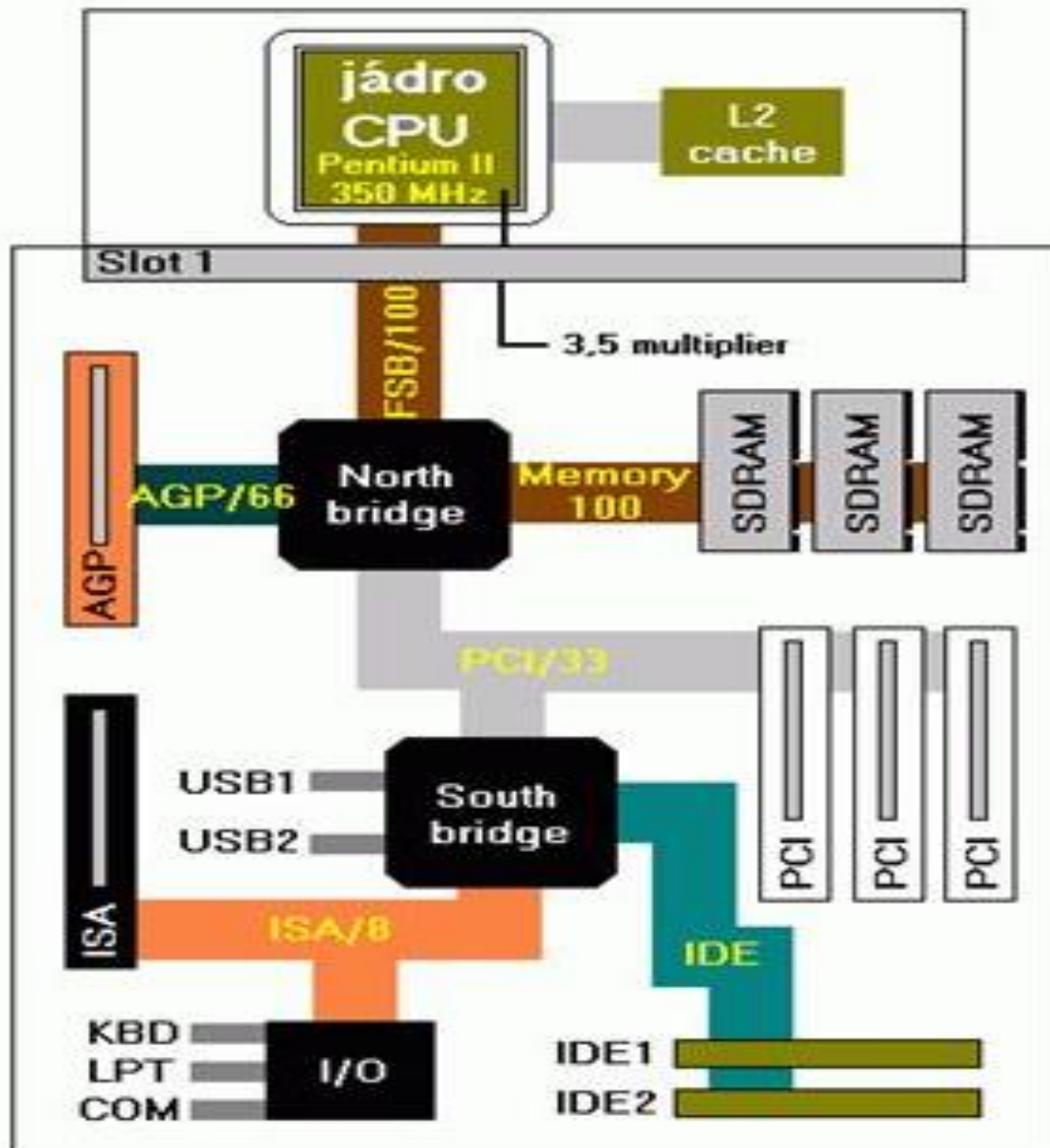
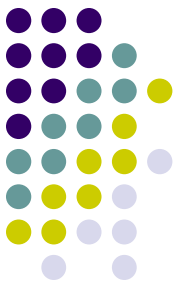
- **Severní most** zajišťoval komunikaci mezi CPU, pamětí RAM (řadič paměti), AGP portem nebo PCI Express sběrnici a také zajišťuje spojení s jižním mostem
- Některé severní můstky obsahují integrované grafické karty
- Protože různé procesory a paměti používají rozdílné signály, pracuje severní můstek pouze s jedním nebo se dvěma typy procesorů a zpravidla pouze s jedním typem paměti RAM
- Severní můstek tedy byl na základních deskách základním prvkem, který určuje rychlost, typ procesorů a typ paměti RAM, který lze na daném motherboardu použít
- Severní můstek hraje důležitou roli v tom jak dalece mohou být počítače (pře)taktovány
- Jeho frekvence je stanovena pro procesor jako základní, od níž se jeho vlastní frekvence procesoru odvíjí pomocí násobiče

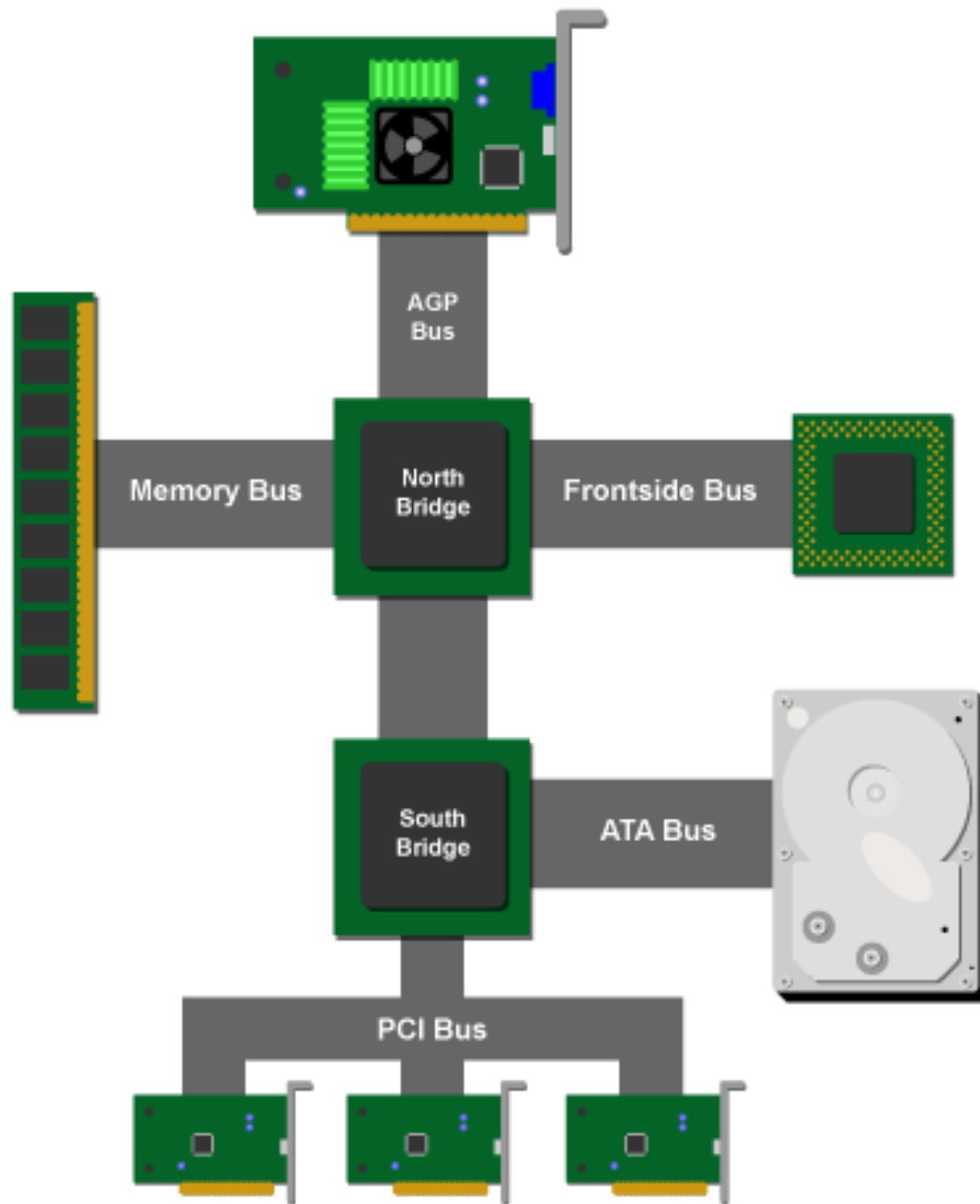
# Southbridge

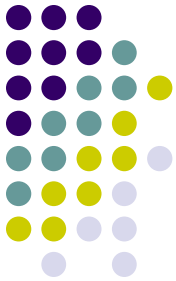
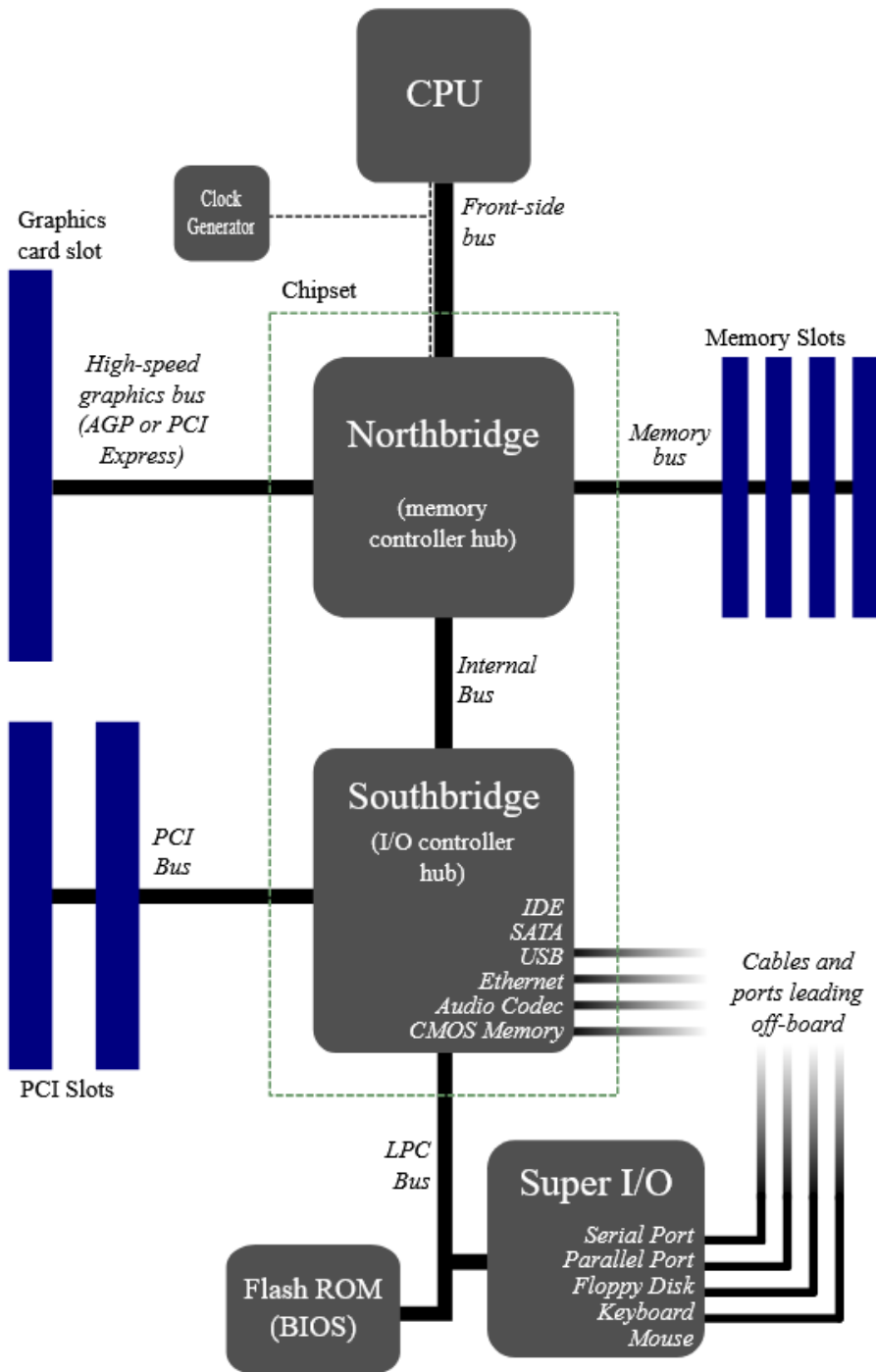


- Čip realizuje **pomalejší I/O** funkce základní desky
- Protože jižní můstek je více **vzdálen od procesoru**, má v typickém počítači na starosti obsluhu pomalejších zařízení
- Dříve byla pro komunikaci mezi severním a jižním můstkem využívána PCI sběrnice
- Toto řešení vytvářelo úzké místo, proto většina moderních chipsetů využívá pro vzájemnou komunikaci vlastní nestandardní rozhraní s vysokou propustností (Direct media interface)



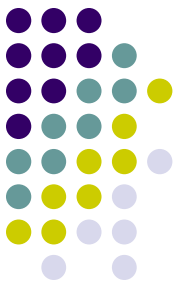






### Super I/O - SIO

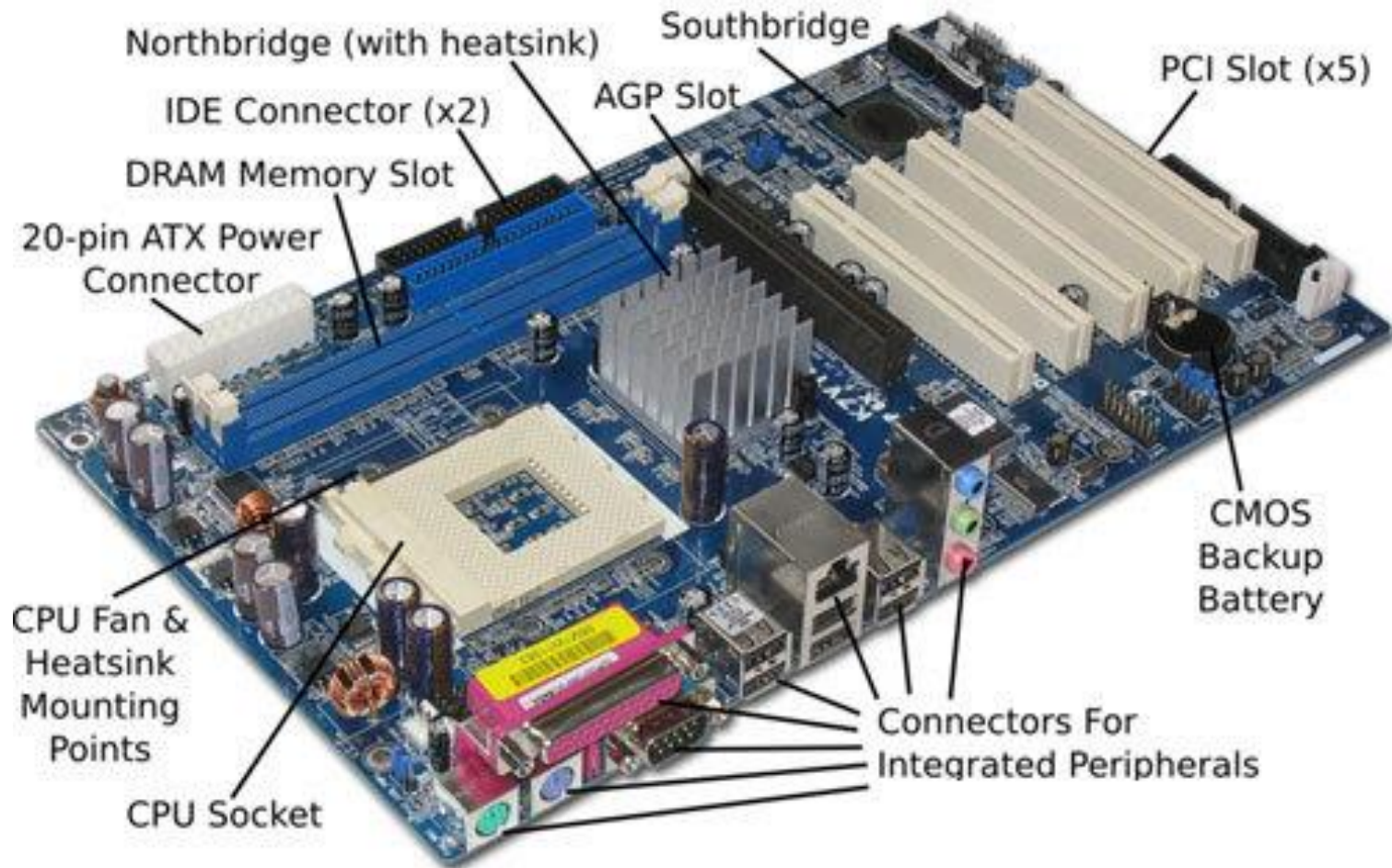
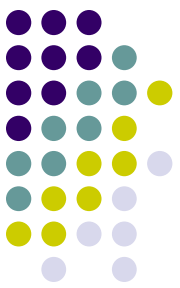
Chip nebo sada chipů pro řízení komunikace s pomalými vstupy a výstupy

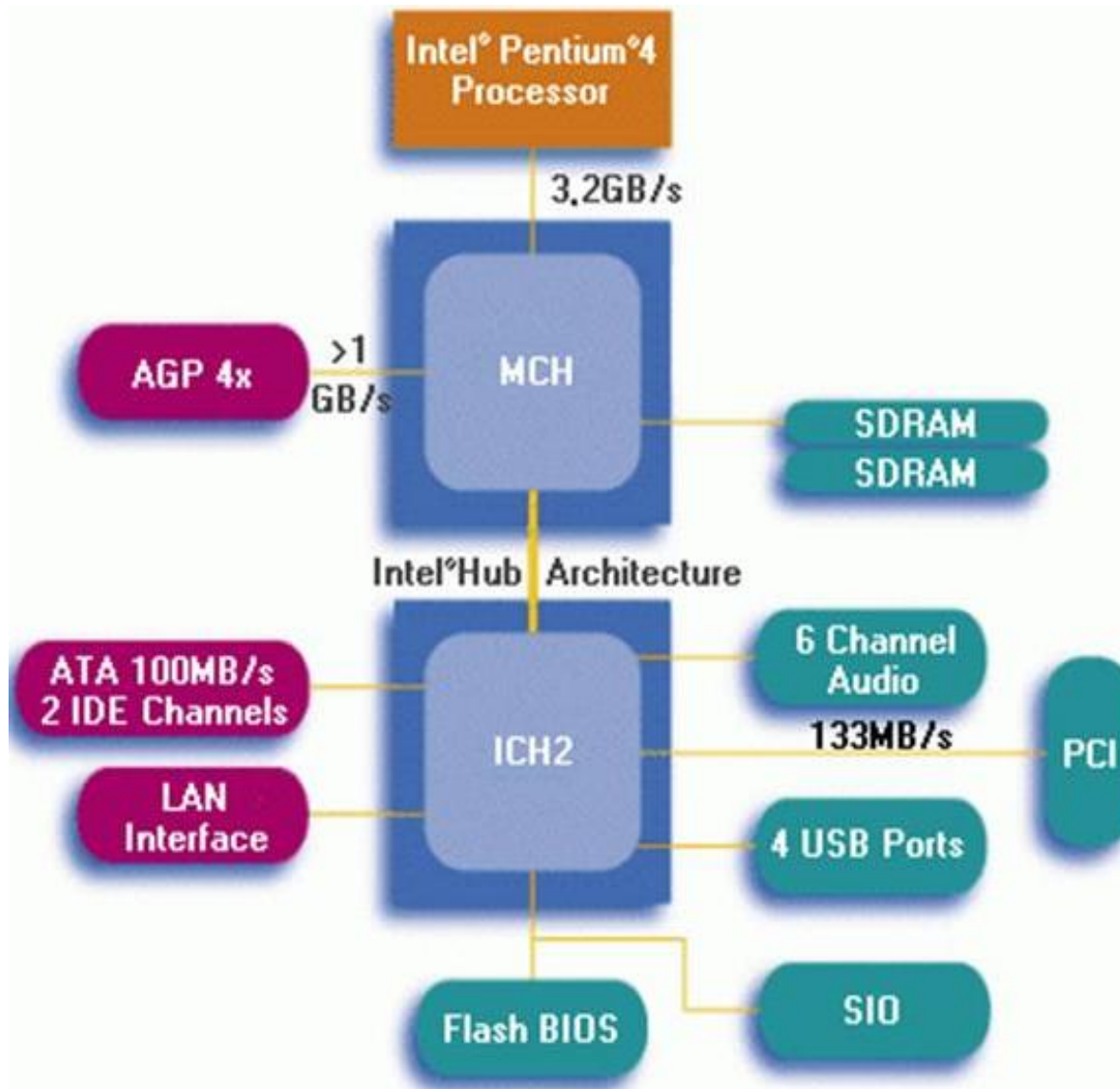
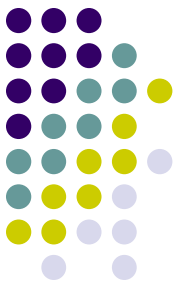


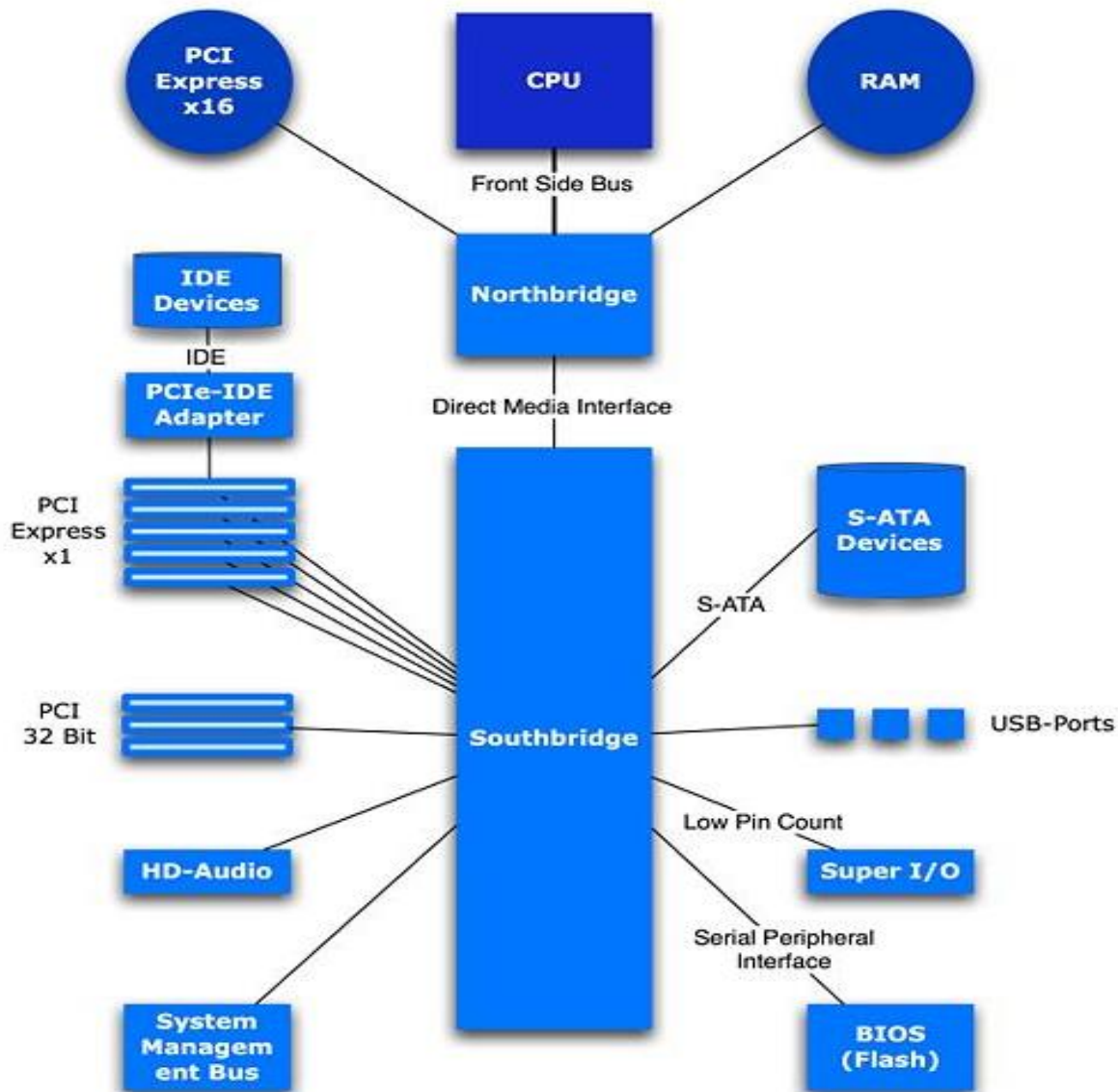
**SOUTHBRIDGE**

**NORTHBRIDGE**

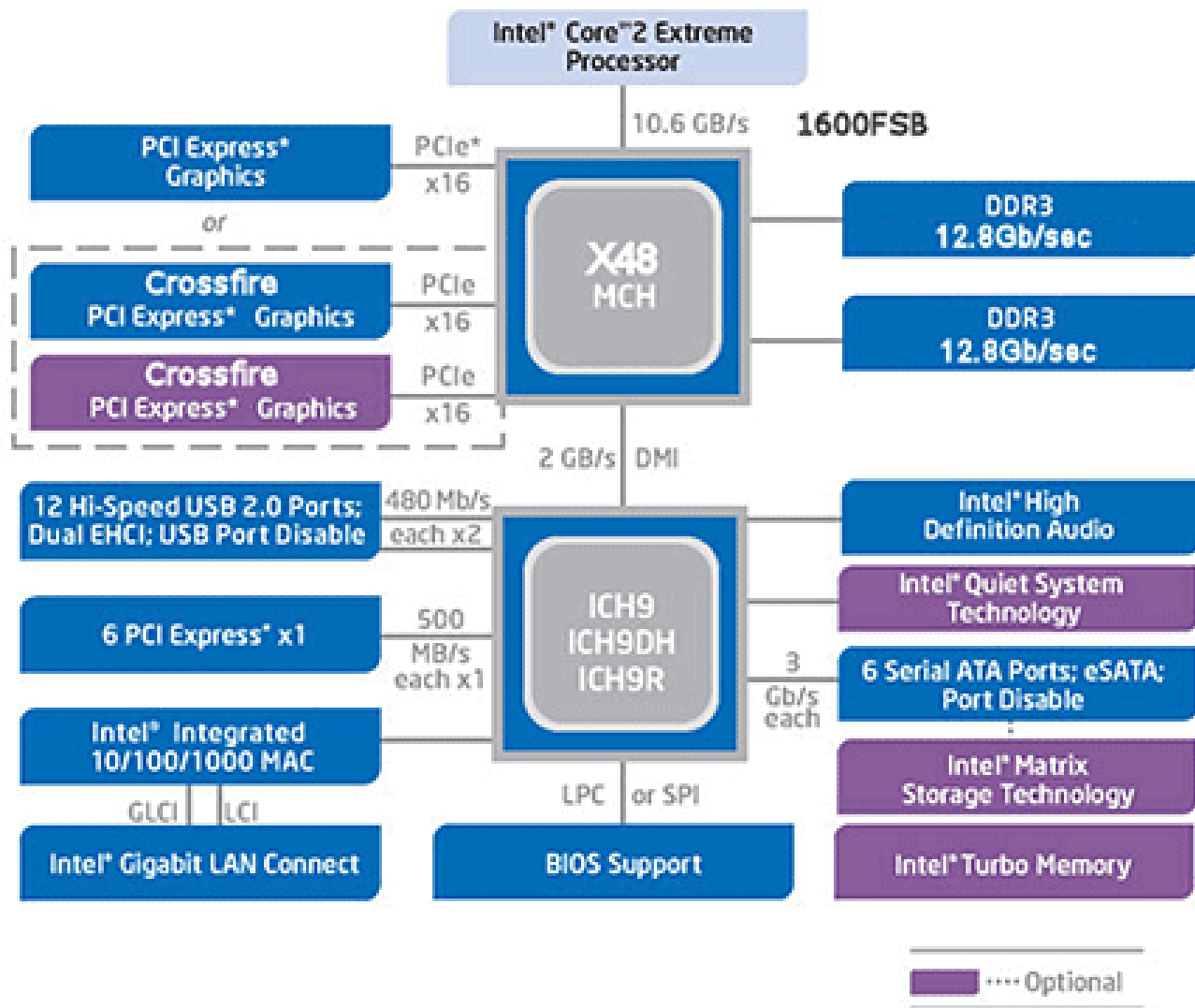






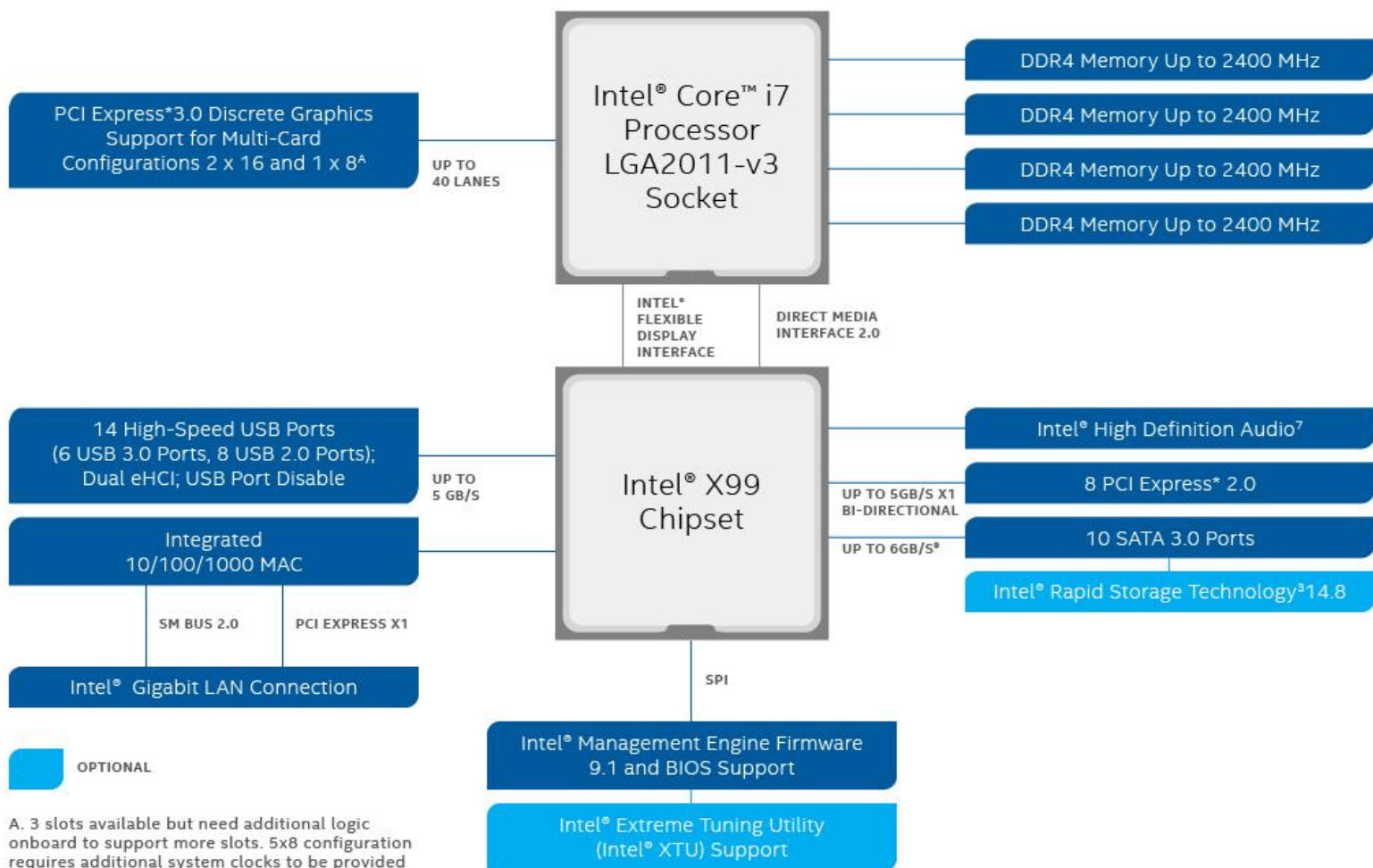






Intel® X48 Express Chipset Block Diagram

## INTEL® X99 CHIPSET BLOCK DIAGRAM



A. 3 slots available but need additional logic onboard to support more slots. 5x8 configuration requires additional system clocks to be provided by third party components.

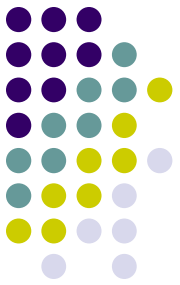
B. All SATA ports capable of 6 GB/s.

# Komunikace s pamětí, SDRAM



- Synchronní DRAM = **SDRAM**
- SDRAM má vstup pro **hodinový signál** a veškeré události jsou synchronizovány a odvozeny od tohoto hodinového signálu
- Mikroprocesor dříve **nekomunikoval** s pamětí **přímo** – paměť a FSB mikroprocesoru mohou běžet na zcela jiných frekvencích a navíc mikroprocesor nemá čas starat se o správné načasování výběru sloupce a řádku
- O adresaci a časování paměti se tedy staral **severní můstek** – systémový řadič, řadič paměti
- Všechny komponenty systému (CPU, severní můstek, paměti) jsou spolu spojeny přes sběrnice o různých frekvencích a s různými napěťovými úrovněmi signálů
- **Základní deska** musí umět sjednotit a synchronizovat tyto různé sběrnice mezi sebou, aby data mohla proudit plynule a bez potíží

# Adresový dekodér



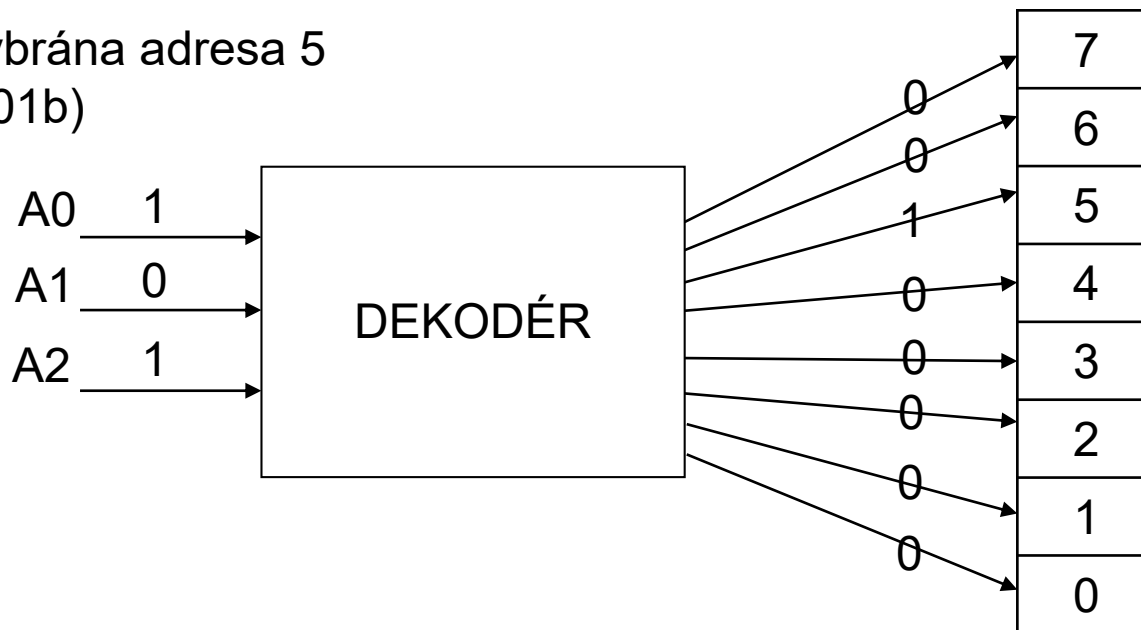
- Signály z **adresačních vstupů** jsou přivedeny do **dekodéru**
- Jeho úkolem je **binárně zakódovanou adresu** převést na signál, kterým bude **aktivována** jedna správná paměťová buňka
- Například dekodér v paměti s kapacitou 256 Bajtů bude mít *8 vstupů (A0..A7)* a *256 výstupních signálů*
- Bude-li na vstupu takového dekodéru kombinace 00010111, objeví se na výstupu dekodéru jednička na signálu, který vede do paměťové buňky s adresou 23 a na všech ostatních výstupech bude 0
- Bude-li na vstupu takového dekodéru kombinace 11111111, objeví se na výstupu dekodéru jednička na signálu, který vede do paměťové buňky s adresou 255 a na všech ostatních výstupech bude 0
- Bude-li na vstupu takového dekodéru kombinace 01010101, objeví se na výstupu dekodéru jednička na signálu, který vede do paměťové buňky s adresou 85 a na všech ostatních výstupech bude 0

# Adresový dekodér



Paměť s kapacitou 8B

Je vybrána adresa 5  
(5=101b)



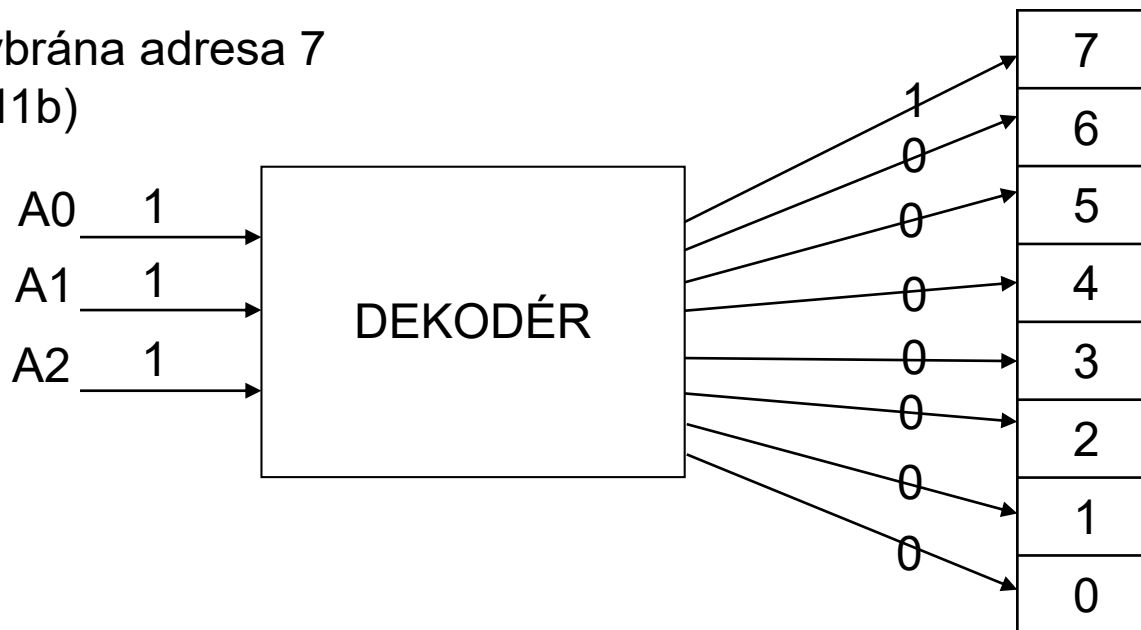
Na všech výstupech dekodéru jsou nuly, kromě signálu, který vede do paměťové buňky s adresou 5

# Adresový dekodér



Paměť s kapacitou 8B

Je vybrána adresa 7  
(7=111b)



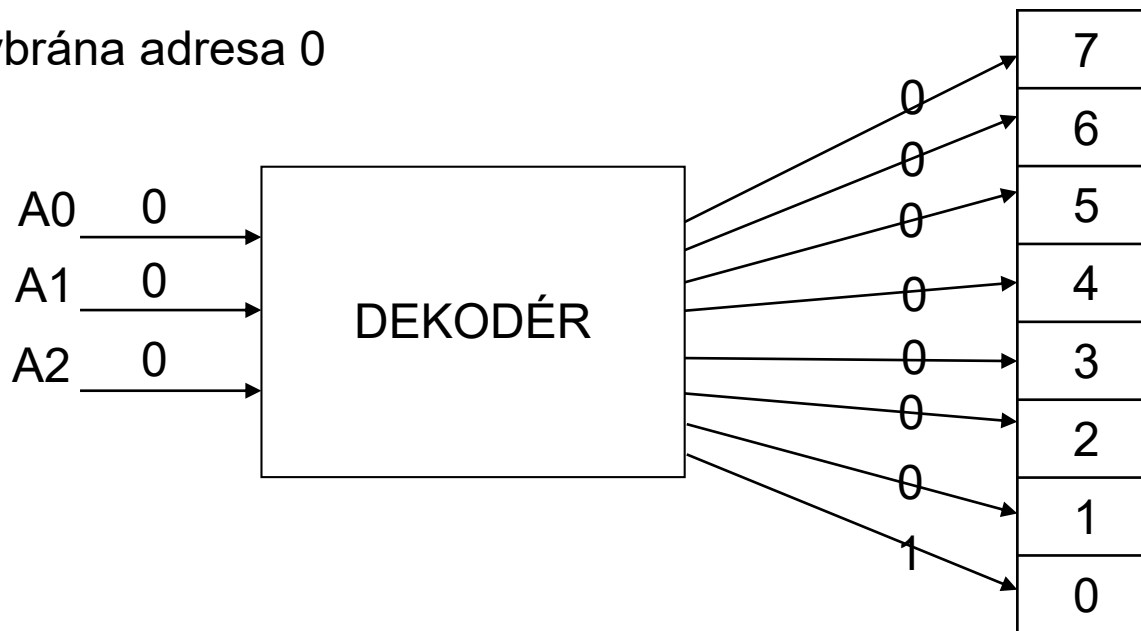
Na všech výstupech dekodéru jsou nuly, kromě signálu, který vede do paměťové buňky s adresou 7

# Adresový dekodér



Paměť s kapacitou 8B

Je vybrána adresa 0



Na všech výstupech dekodéru jsou nuly, kromě signálu, který vede do paměťové buňky s adresou 0

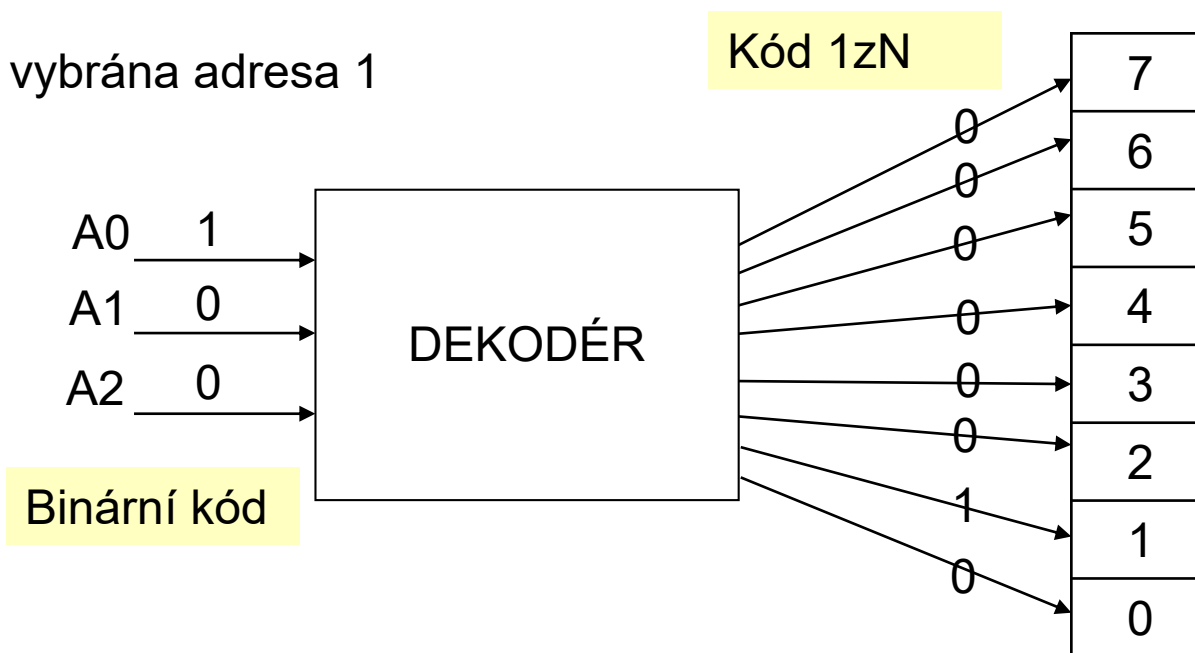


# Adresový dekodér

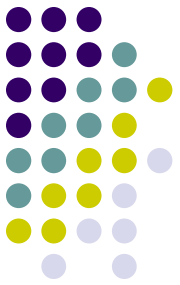


Paměť s kapacitou 8B

Je vybrána adresa 1

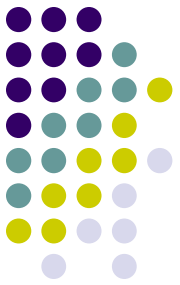


Na všech výstupech dekodéru jsou nuly, kromě signálu, který vede do paměťové buňky s adresou 1



# Dekodér

- Čím má paměť **větší kapacitu**, tím je její adresový dekodér **složitější**
- Například v paměti s kapacitou 1 MB musí mít dekodér 20 vstupů (A0..A19) a 1048576 výstupů
- Pokud by na vstupu takového dekodéru byla kombinace 10110111110000010000, objeví na výstupu jednička na vývodu číslo 752656 a na všech ostatních výstupech bude nula
- Paměťový chip s kapacitou 1 GB by již musel mít **extrémně složitý dekodér** s 30 adresačními vstupy a 1073741824 výstupy – to by bylo nerealizovatelné
- Proto u paměti s větší kapacitou dochází k rozložení dat do „**řádků a sloupců**“ a použití dvou jednodušších dekodérů

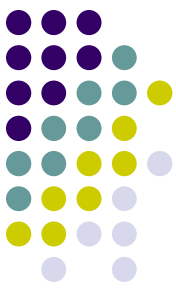


# Dekódování adresy

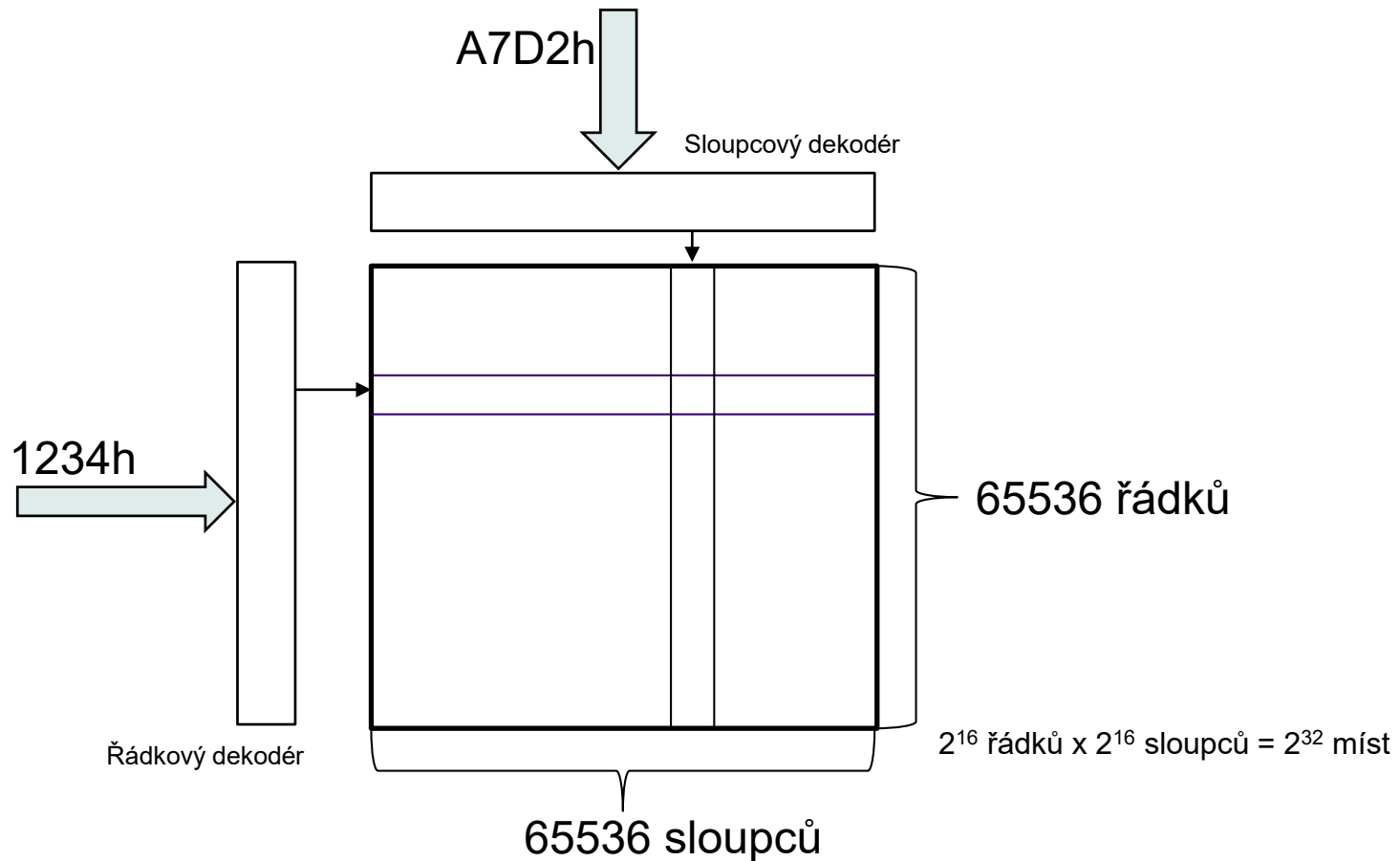
- Kapacita moderních velkých pamětí je příliš veliká na to, aby v nich mohl být jeden jediný adresový dekodér
- Například paměť s kapacitou 4 GB by musela mít dekodér na jehož vstupu by bylo 32 signálů A0-A31, ale na jeho výstupu by bylo  $2^{32}$  (tj. 4 294 967 296 ) signálů, kterými by byl vybrán jeden konkrétní řádek paměti obsahující požadovaný bajt (8 bitů)
- Moderní paměť je organizována podle adres do řádků a sloupců – používá dekodér pro sloupce a dekodér pro řádky
- Je jednodušší mít v paměti dva 16-bitové dekodéry (s 65536 výstupy) než jeden supersložitý 32-bitový (s 4 294 967 296 výstupy)
- **Horní bity adresy vybírají řádek**, na kterém v paměti nalezneme požadovaná data
- **Spodní bity adresy, pak vybírají sloupec** (tzn. kde na daném řádku) data nalezneme

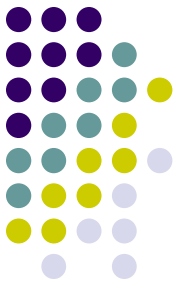
Při **sekvenčním přístupu** k následujícím adresám (1234A7D3h, 1234A7D4h, 1234A7D5h...) zůstáváme na řádku 1234 h (horních 16-bitů adresy se nemění)

Při běžné práci s daty, která jsou uložena v paměti postupně za sebou, by tedy stačilo vybrat **trvale řádek** a pak adresovat pouze pomocí spodních 16 bitů sloupec



- Příklad
- Výběr adresy 1234A72D h
- Horních 16 bitů použijeme k výběru řádku – vybere se řádek číslo 1234 h
- Spodních 16 bitů použijeme k výběru sloupce – vybere se sloupec číslo A7D2 h

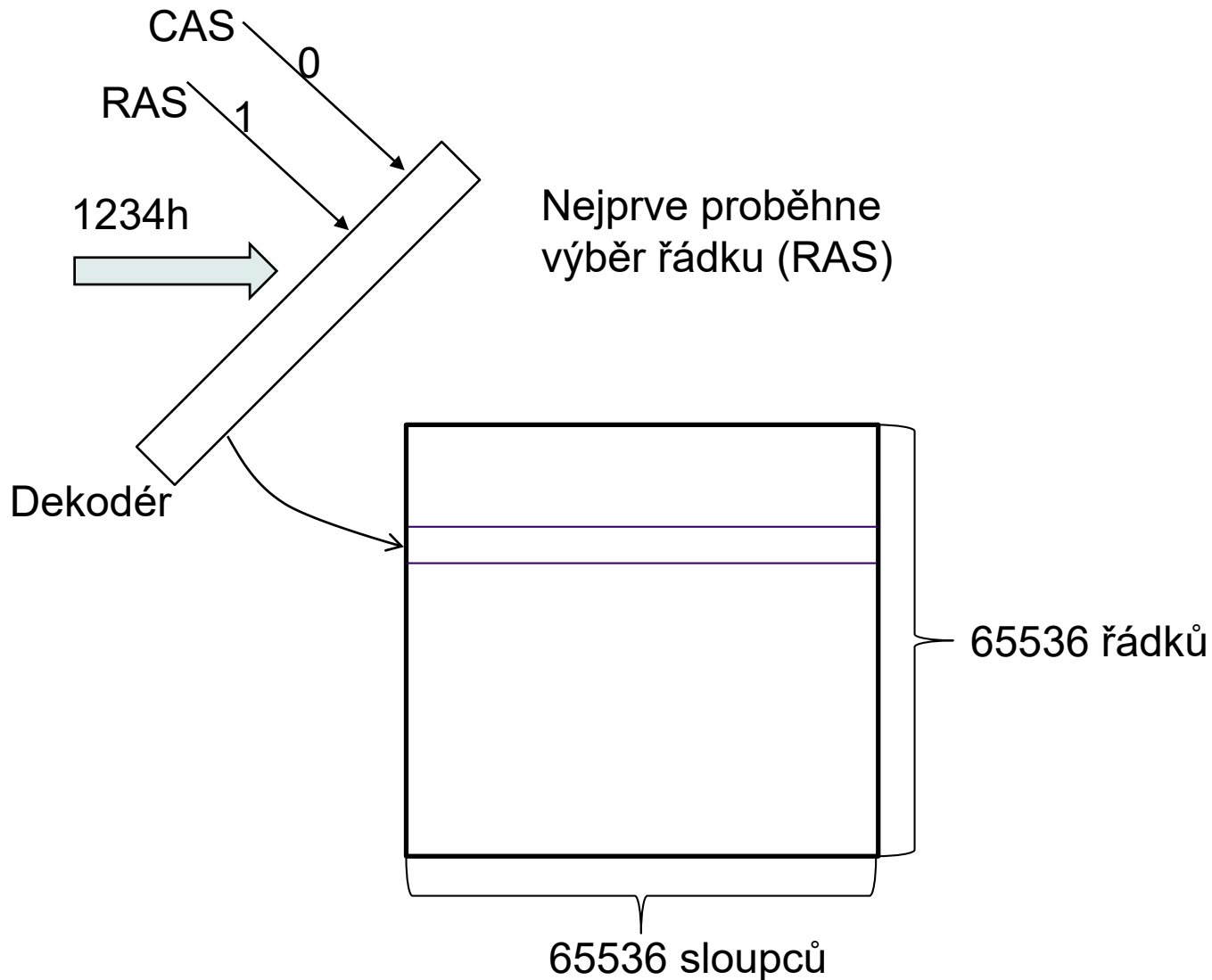
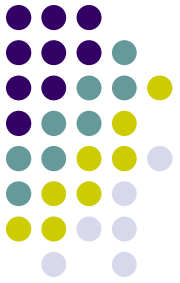




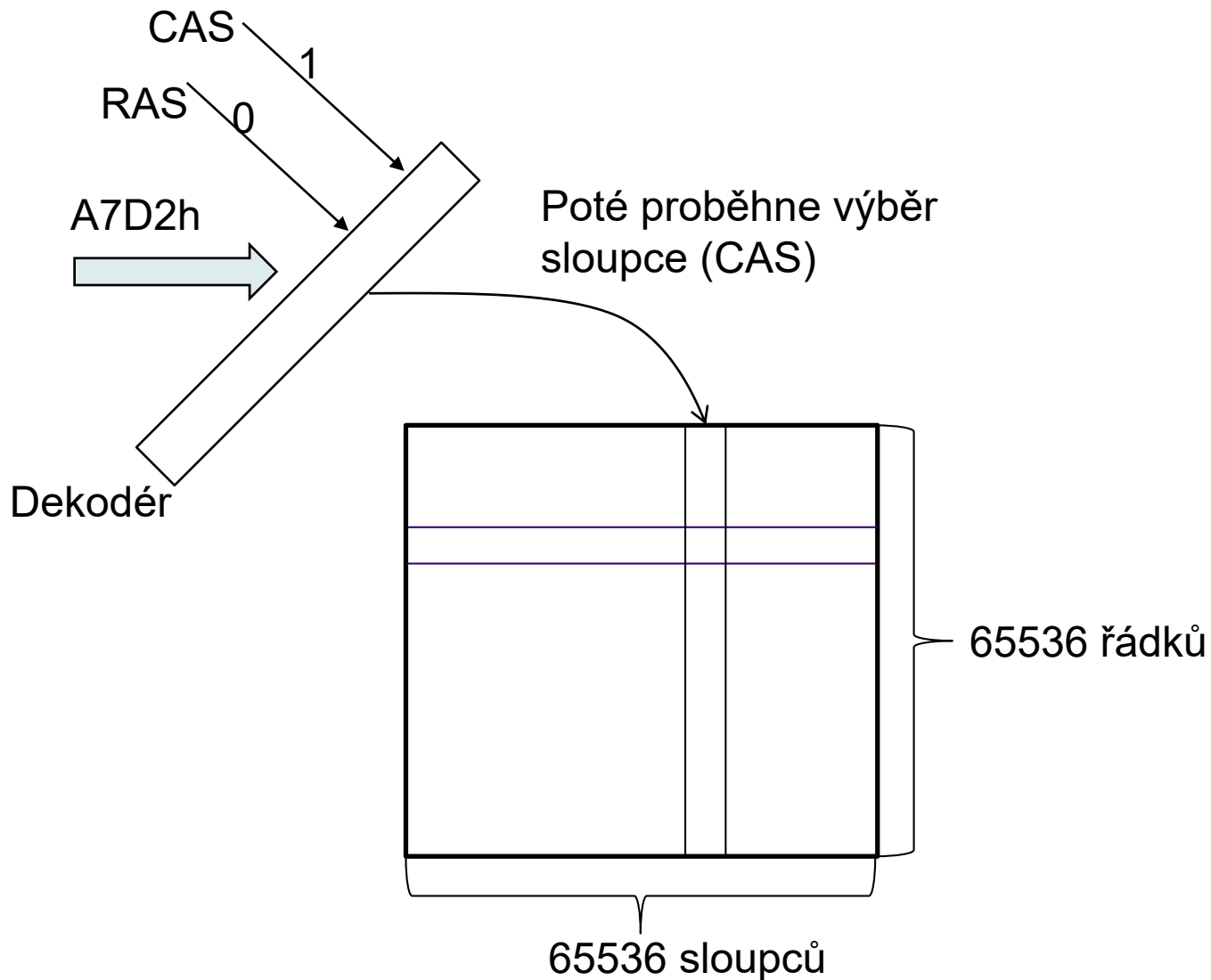
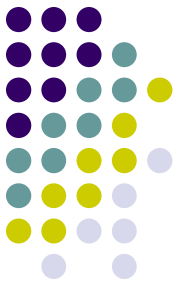
# Dekódování adresy (RAS a CAS)

- Při přístupu do paměti vybíráme nejprve řádek (přivádíme pouze horní bity adresy a pracuje pouze řádkový dekodér) – **Row Access**
- Výběr řádku trvá určitou dobu – Row Access Strobe (RAS)
- Poté pomocí spodních bitů adresy a sloupcového dekodéru vybíráme konkrétní adresu – **Column Access**
- Data na výstupu paměti se poté objeví se zpožděním několika taktů. Toto zpoždění se označuje jako Column Access Strobe – CAS
- RAS to CAS je další důležitý časový parametr, který udává, kolik taktů je potřeba počkat mezi výběrem řádku a sloupce
- Vzhledem k tomu, že výběr řádku a sloupce neprobíhá současně, bylo by možné použít k výběru řádku i sloupce jeden stejný dekodér
- Výstupy takového dekodéru by vedly k řádkům i ke sloupcům a pomoc řídicích signálů by se přepínalo, jestli se má dekodér právě použít k výběru řádku nebo sloupce

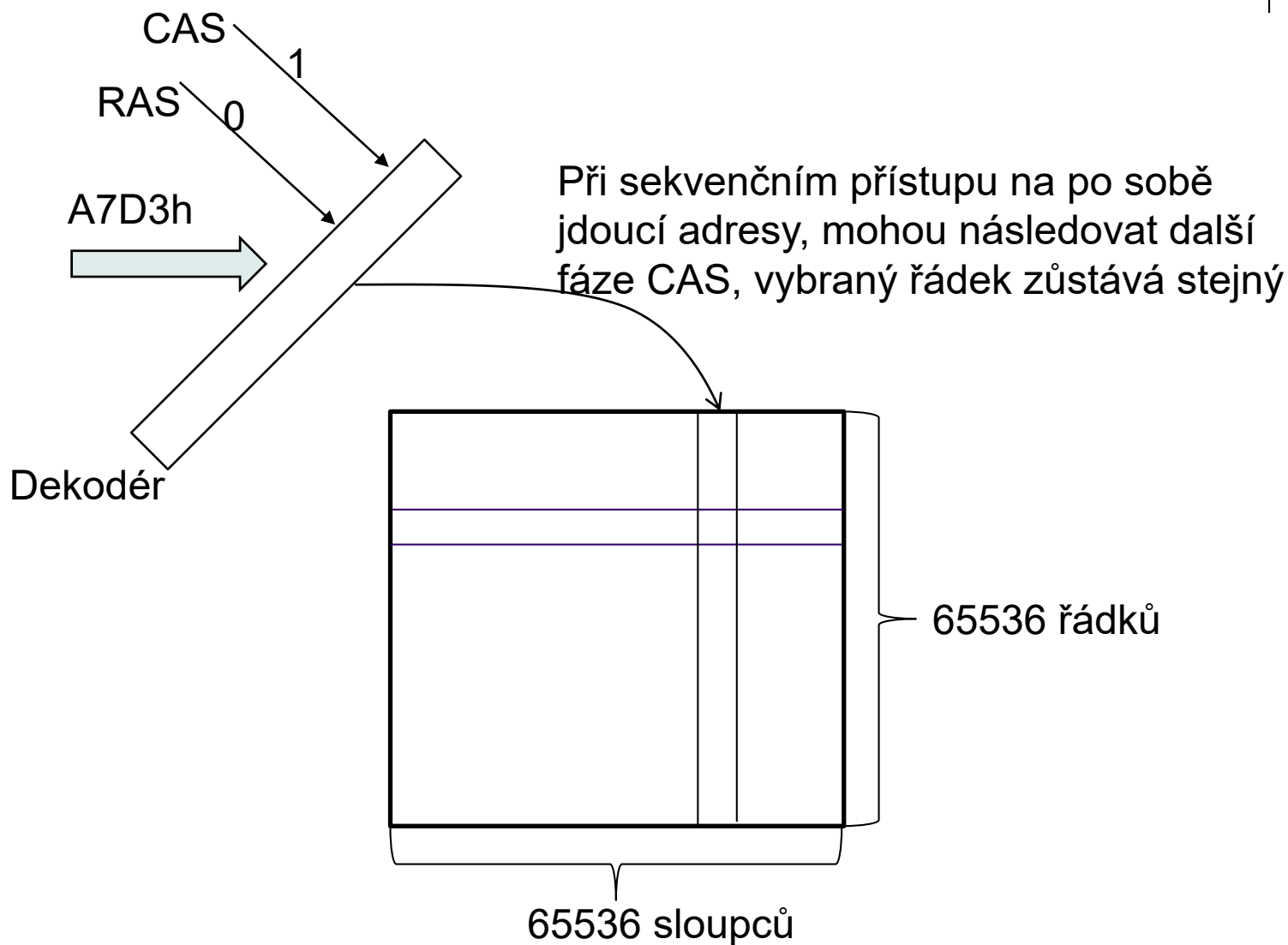
# Dekódování adresy



# Dekódování adresy

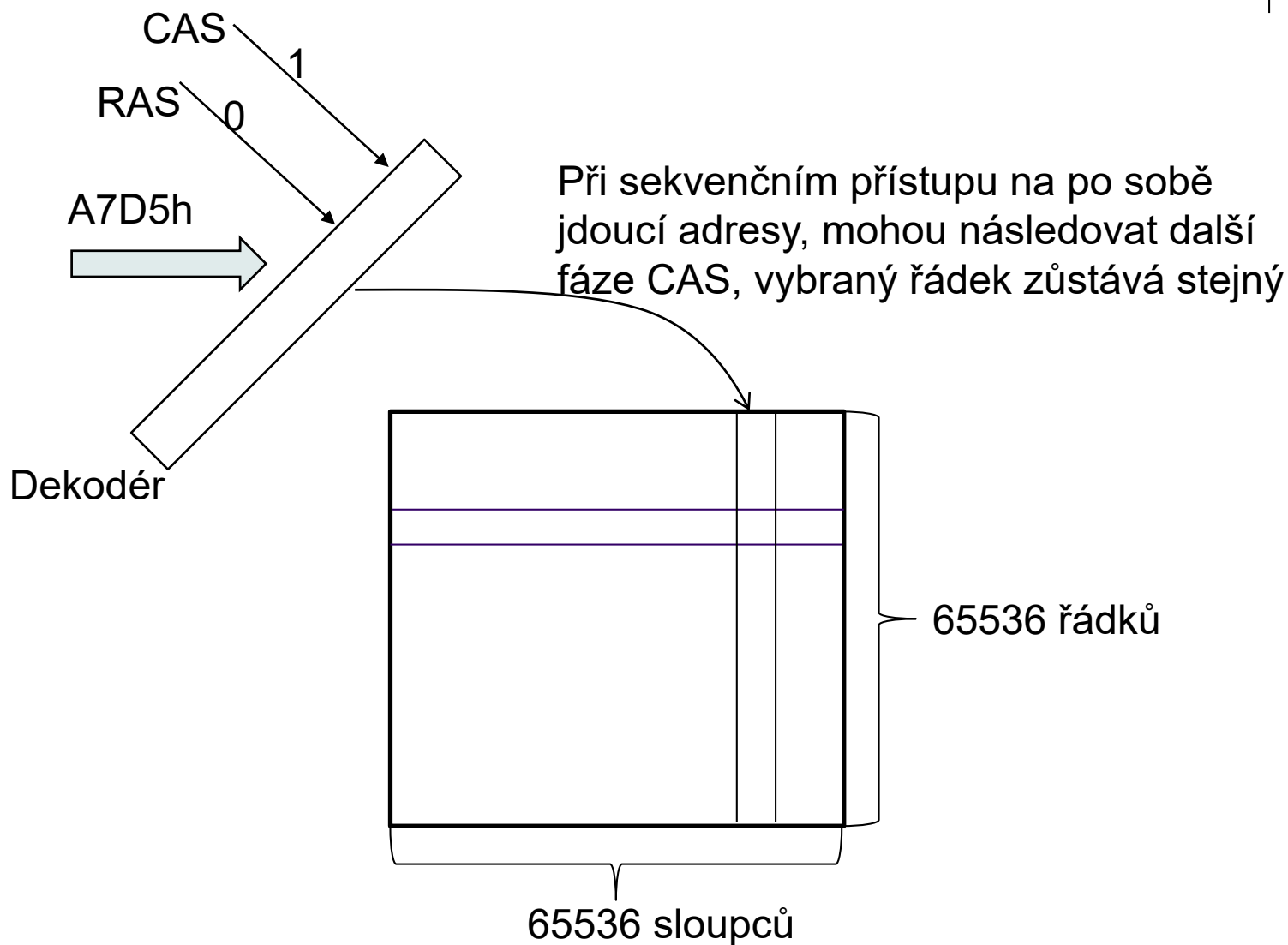


# Dekódování adresy

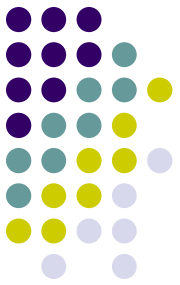




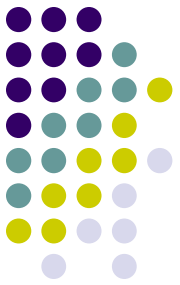
# Dekódování adresy



# Časování paměti



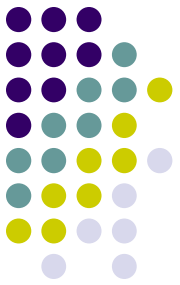
- Vhodným nastavením časování paměti lze zvýšit výkon celého počítače, ale také snížit stabilitu
- Správné nastavení časování paměti patří skoro do oblasti „počítačové alchymie“. Problematika je velmi komplikovaná a přesahuje rámec středoškolských možností
- Základní parametry časování paměti pak bývají obvykle udávány jako kombinace čtyř čísel CAS - RAS to CAS - RAS precharge - Active to precharge
- Například: 2-3-3-6
- **CAS Latency** – Označováno také jako **CL** – Kolik taktů je třeba počkat na data, po tom, co byl vybrán sloupec.
- Čím je CL nižší, tím lepší
- V podstatě jde o nejdůležitější parametr, protože obvykle čteme data sekvenčně
- Přístup do paměti bude nejrychlejší, pokud program pracuje s adresami, které jsou postupně za sebou, protože ty budou uloženy ve stejném řádku – vybírá se tedy pouze sloupec
- Naopak náhodný přístup do paměti na různá od sebe vzdálená místa je vždy zdlouhavější
- **RAS to CAS** – počet taktů, které proběhnou mezi výběrem řádku a výběrem sloupce
- **RAS precharge** – počet taktů potřebný ke změně řádku. Pokud chcete přestat používat vybraný řádek, je třeba provést akci „precharge“
- **Activate to Precharge** – Nový řádek je vybrán akcí „activate“. Další „precharge“ (tedy deaktivace tohoto řádku) smí proběhnou až po určitém minimálním počtu taktů.



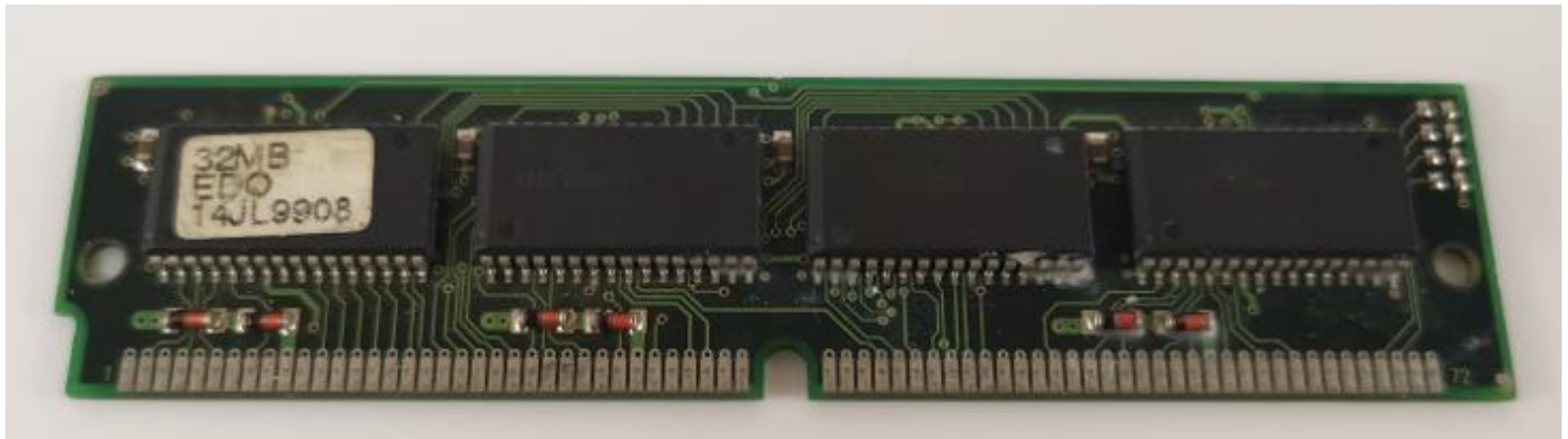
# Paměťové moduly

- V osobních počítačích není paměť tvořena jedním samostatným chipem, ale několik chipů je zapojeno dohromady a tvoří **paměťový modul**
- Chipy na paměťových modulech pracují **paralelně** – Paměti jsou pomalé, takže když už trvá přístup do paměti dlouho, čteme při takovém přístupu co nejvíce dat naráz z několika chipů paralelně (když jedete na nákup jen jednou za týden, nakoupíte si toho hodně do zásoby)
- První paměťové moduly byly složeny z osmi paralelně zapojených jednobitových pamětí.
- Každý bit bajtu byl uložen v jiném paměťovém obvodu.
- Vybraná adresa je přivedena na vstup všech osmi paměťových obvodů
- Tyto paměťové moduly se nazývají DIP (Dual-in-package)

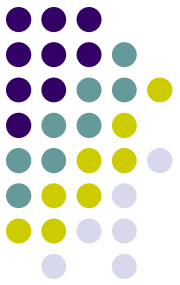
# SIMM



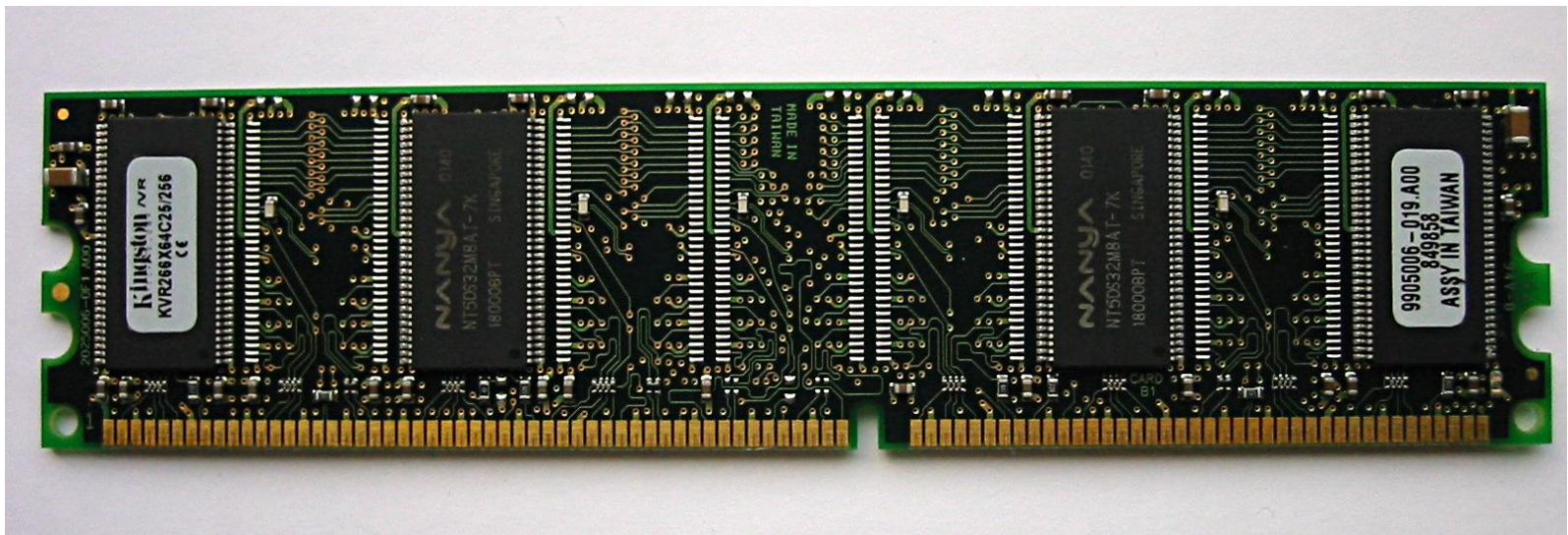
- Později byli paměťové obvody umísťovány na malou desku zvanou SIMM (Single In-line Memory Module)
- Typická varianta SIMM má 72 vývodů a obsahuje 4 paměti s šířkou slova 8 bitů, které jsou zapojené paralelně, takže celková šířka dat je 32 bitů
- Tato varianta paměťových modulů se používala ve 32-bitových počítačích a umožňovala zápis/čtení 32 bitů naráz (4 bajty ležící na sousedících adresách)



# DIMM

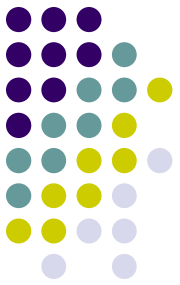


- Pro mikroprocesory Pentium vznikly paměťové moduly DIMM
- Dual In-line Memory Module
- V podstatě jde o dva paralelně zapojené SIMM moduly
- Šířka dat je **64 bitů (čte se osm po sobě jdoucích adres současně)**
- Modul má **168 vývodů**
- Vývody jsou po obou stranách modulu
- kontakty na starších SIMM byly zdvojené (tzn. po obou stranách stejné)
- Pokud byl procesor Pentium používán s moduly SIMM, musely být osazovány v párech, aby dohromady daly 64-bitovou šířku dat

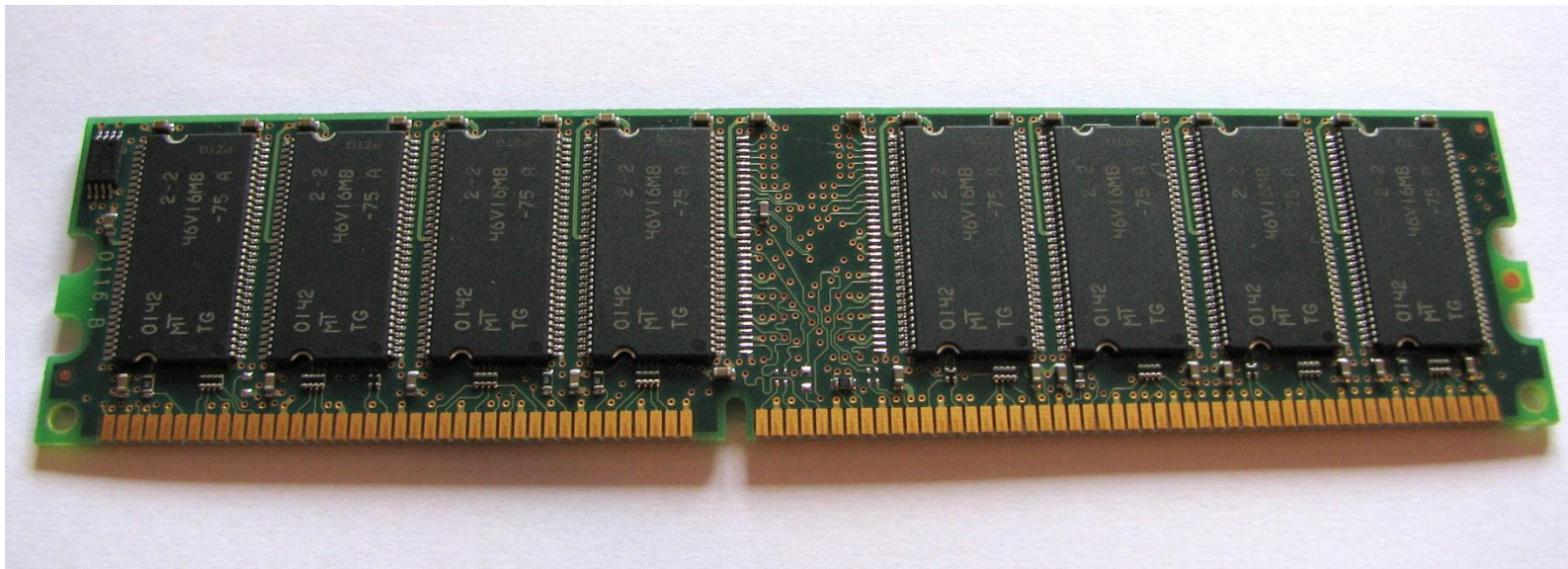




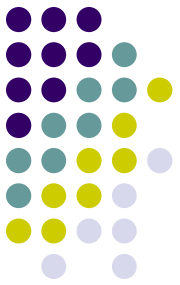
# DDR DRAM



- Double Data Rate DRAM
- Typ paměti používaný v současných počítačích PC
- Paměťové moduly DDR mají 184 pinů a zámek, který zaručuje správnou instalaci do paměťového slotu na základní desce

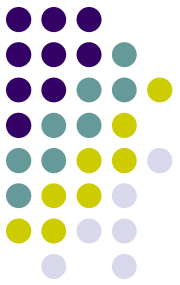


# DDR



- Synchronní DRAM = **SDRAM**
- SDRAM má vstup pro **hodinový signál** a veškeré události jsou synchronizovány a odvozeny od tohoto hodinového signálu
- **SDR** – *Single* data rate
- Do této doby běžně používaná synchronní DRAM paměť
- Tyto paměti mají 168 pinů a napájecí napětí typicky 3.3 V
- Data se z/do paměti přenášejí při každé **vzestupné hraně** hodinového signálu
- **DDR** – *Double* data rate
- Data se z/do paměti přenášejí při **vzestupné i sestupné** hraně hodinového signálu (tedy dvakrát během jedné periody)
- nejpomalejší DDR-200 pracují na hodinové frekvenci 100 MHz, ale ke čtení/zápisu může docházet 200 000 000 krát za sekundu
- Mezi procesorem a pamětí probíhá výměna **64 bitů** naráz (8 Bajtů), takže **DDR-200** běžící na frekvenci **100MHz** má teoretickou datovou propustnost **1600 MB/s** (8x200) a používá se i označení **PC1600**

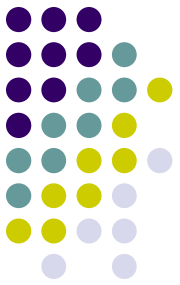
# DDR



- Dvojnásobné propustnosti je dosaženo sledování **sestupné i vzestupné hrany** hodinového signálu
- Během **jedné periody** hodinového signálu dochází ke **dvěma změnám** (hranám)
- Paměť DDR vykoná během jednoho hodinového cyklu dva pracovní cykly odstartované oběma hranami
- DDR paměti pracují s napájecím napětím 2,5 – 2,6 V
- Běžně snesou i vyšší napájecí napětí, které se používá při jejich přetaktování na vyšší frekvenci (stoupne spotřeba el. energie a zahřívání, ale zvýší se stabilita)
- Nejdůležitější parametrem je
  - CL – CAS latence (zpoždění při čtení dat po výběru sloupce), Obvykle CL=2
  - Taktovací frekvence



# DDR



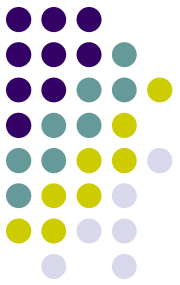
- Příklad:
    - **Paměť DDR-200** je taktována frekvencí 100 MHz a vykoná 200 milionů pracovních cyklů za sekundu
    - Šířka toku dat je 64 bitů (8B). Při souvislém toku dat lze přenést 1600 MB/s
    - Tento modul bude označen jako PC1600
  - **Paměť DDR-600** je taktována frekvencí 300 MHz, přenést lze  $600\,000\,000 \times 8\text{ B} = 4800\text{ MB/s}$
  - Modul bude označen jako PC4800
- 
- Zásadní otázka:
  - Jak je možné u DDR-600 přenést data 600 000 000 krát za sekundu, pokud CL=2? Neměl by počet přenosu být jen 150 000 000 krát za sekundu? DDR umožňuje dvě přenosy během jednoho taktu, ale CL=2 nám říká, že po adresaci trvá dva takty, než budou data k dispozici. To přeci vůbec nedává smysl.....

# Konstrukce DDR SDRAM paměti

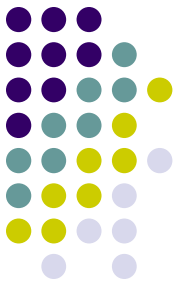


- Příklad - **512 MB SDRAM**
- Paměťový modul bude složen z **osmi chipů** s kapacitou **64 MB** ( $8 \times 64 \text{ MB} = 512 \text{ MB}$ )
- Každý z osmi chipů přispívá svými **osmi bity** k celkové šířce **64 bitů** dat, které lze naráz přečíst nebo zapsat
- Poslední tři bity fyzické adresy určují, v kterém z osmi chipů leží vybraný bajt, ale k vlastní adresaci paměťového obvodu se nepoužijí
- Každý 64 MB chip obsahuje **čtyři 16 MB banky**
- Každá banka bude organizována jako matice, které má **4096 řádků** a **4096 sloupců** ( $4096 \times 4096 = 16 \text{ Mega}$  //  $2^{12} \times 2^{12} = 2^{24}$ )
- K výběru řádku je potřeba 12 bitů (A0 – A11)
- K výběru sloupce je potřeba 12 bitů (A0 – A11)
- Adresační vstupy použité k výběru řádku a sloupce jsou totožné (A0-A11)
- K výběru banky jsou potřeba 2 bity (B0, B1)
- Řádek je obvykle vybrán na delší dobu
- Ke změně sloupců dochází po každém čtení nebo zápisu (pokud nechceme používat stále stejnou adresu pro čtení/zápis)
- Během čekání (CL) mohou být vybrány další banky a zasílány jim adresační pokyny – časování paměti funguje prokládaně, každá banka pracuje nezávisle

# DDR2



- Při stejné taktovací frekvenci nabízí dvakrát vyšší výkon než DDR, protože se pracuje dvojnásobnou šířkou dat – 16 B (128 b)
- Pracují s nízkým napájecím napětím 1,8 V
- Jsou dodávány v modulech DIMM s 240 vývody
- Díky snížení napájecího napětí a možnosti dosáhnout stejné propustnosti při poloviční frekvenci než DDR první generace, mají DDR2 výrazné nižší spotřebu energie
- Příklad:
  - **DDR2-400**
    - Taktovací frekvence 100 MHz, při tom vykoná 200 milionů pracovních cyklů
    - Při jednom pracovním cyklu lze přenést 2 x 64 bitů (16 B)
    - Propustnost je tedy 3200 MB/s a paměť se bude značit jako PC2-3200
  - **DDR2-1066**
    - Taktovací frekvence 266 MHz, při tom vykoná 533 milionů pracovních cyklů
    - Propustnost je tedy 8533 MB/s a paměť se bude značit jako PC2-8500



# Srovnání DDR a DDR2

- **DDR-400**

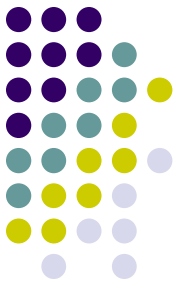
- Proveďte se 400 000 000 přístupů za sekundu
- Taktovací frekvence 200 000 000 Hz
- Přenáší se 8 B naráz
- Propustnost je  $400\,000\,000 \times 8\text{B} = \mathbf{3200\,MB/s}$
- Paměť se označí také jako **PC3200**

Číslo **400** uvedené u paměti DDR nám udává počet přístupů za sekundu (v milionech)

- **DDR2-400**

- Proveďte se 200 000 000 přístupů za sekundu
- Taktovací frekvence 100 000 000 Hz
- Přenáší se **16 B naráz**
- Propustnost je  $200\,000\,000 \times 16\text{B} = \mathbf{3200\,MB/s}$
- Paměť se označí také jako **PC2-3200**

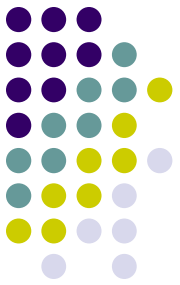
Číslo **400** uvedené u paměti DDR2 nám neudává ani taktovací frekvenci, ani počet přístupů za sekundu. Paměti DDR400 a DDR2-400 mají stejnou propustnost a proto mají v označení i stejná čísla



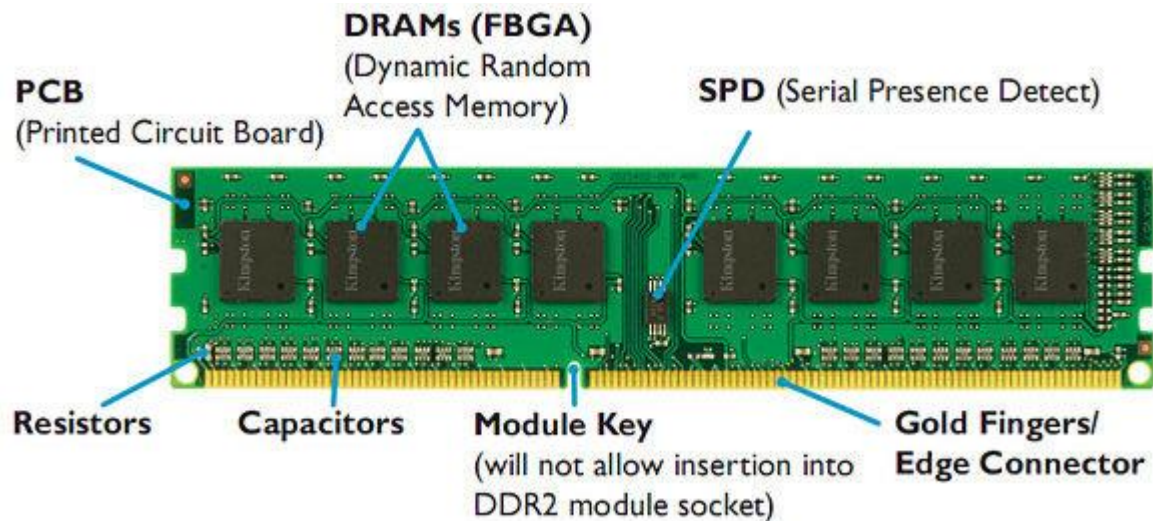
# Srovnání DDR a DDR2

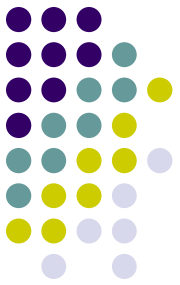
- **DDR-600**
- Provede se 600 000 000 přístupů za sekundu
- Taktovací frekvence 300 000 000 Hz
- Přenáší se 8 B naráz
- Propustnost je  $600\,000\,000 \times 8\text{B} = \mathbf{4800\,MB/s}$
- Paměť se označí také jako **PC4800**
  
- **DDR2-600**
- Provede se 300 000 000 přístupů za sekundu
- Taktovací frekvence 150 000 000 Hz
- Přenáší se 16 B naráz
- Propustnost je  $300\,000\,000 \times 16\text{B} = \mathbf{4800\,MB/s}$
- Paměť se označí také jako **PC2-4800**

# DDR3

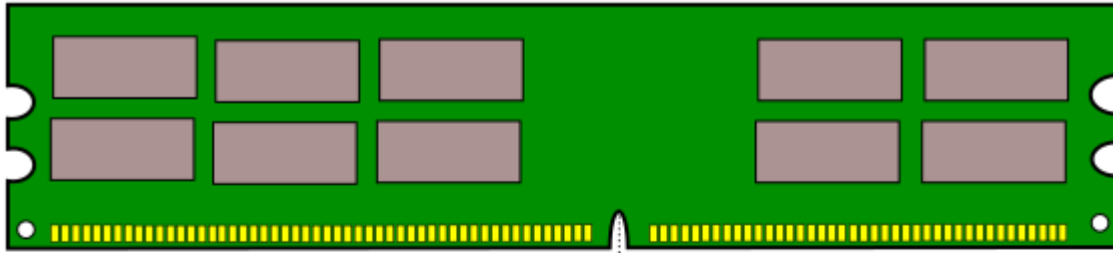


- Napájecí napětí 1,5 – 1,35 V
- Díky nízkému napájecímu napětí mají ještě nižší spotřebu energie než předchůdce DDR2
- Zpočátku byly velmi drahé a nemuseli být vždy výkonnější než DDR2, protože mají obvykle vyšší latenci (více taktů, ale takty jsou kratší = CAS latence cca 10 ns)
- 240 vývodů - jsou elektricky nekompatibilní s předchůdci DDR a DDR2
- Jejich éra začala s příchodem architektury Nehalem (tento procesor a všechny pozdější mají integrovaný řadič paměti, který již podporuje pouze DDR3)
- Udáváné teoretické propustnosti (např 17 GB/s) jsou jen velmi těžko v praxi dosažitelné

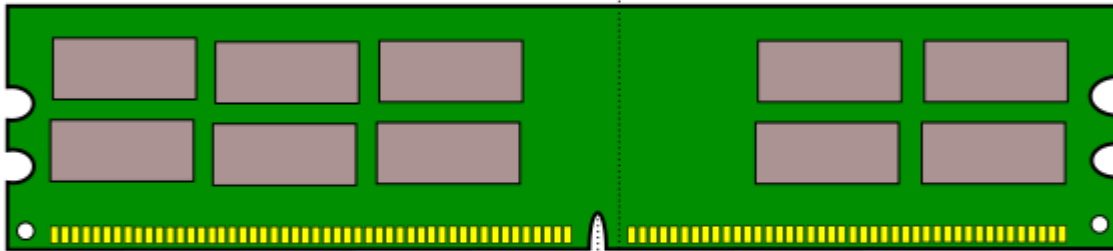




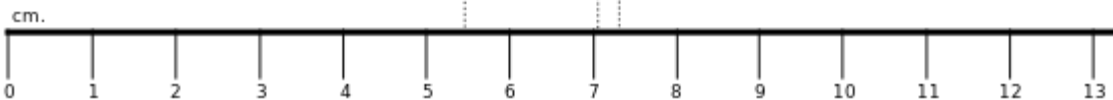
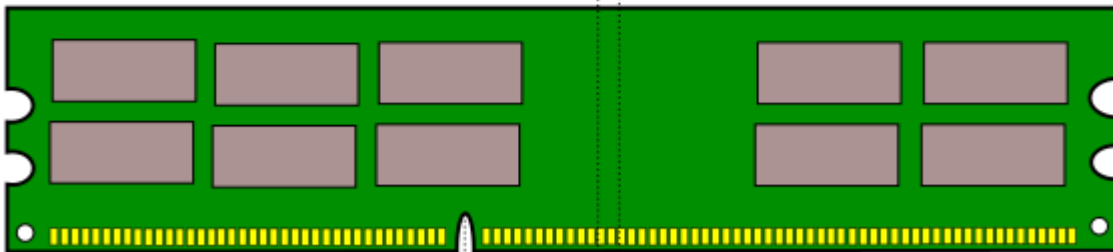
# DDR



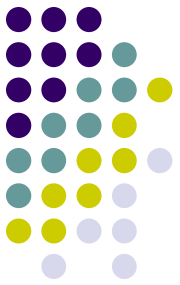
# DDR 2



# DDR 3



# DDR4

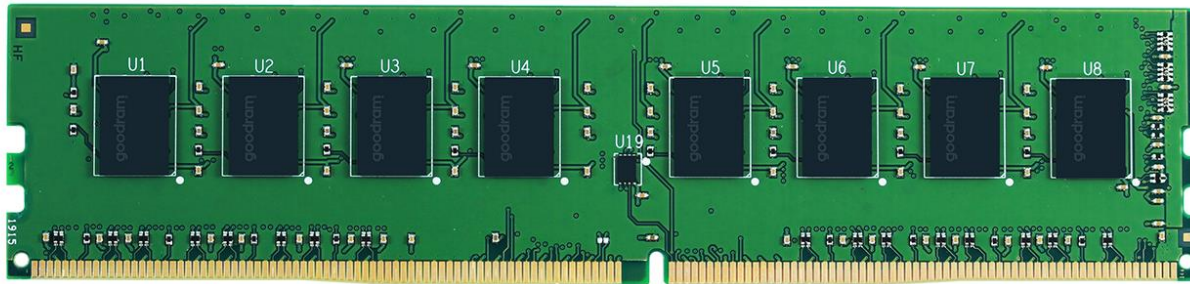


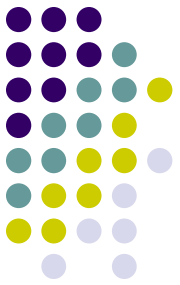
- Používají se od roku 2014
  - U DDR4 pamětí se mění výrobní proces (z 30nm se dále přechází na 20nm)
  - Provozní napětí sníženo na 1,05 – 1,2 V
  - Zvýšily se pracovní frekvence (nad 2000 MHz) a tím i přenosové rychlosti
  - Negativní změnou je zvýšení latence na CL12
  - Díky těmto změnám se snížila spotřeba energie až o 40 % při stejné propustnosti oproti předchozí generaci
- 
- U těchto pamětí se obvykle udává taktovací frekvence a nikoliv počet přístupů za sekundu.
  - Příklad DDR4 2400 MHz – PC4-19200
  - Paměť má taktovací frekvenci 2400 MHz
  - Propustnost je 19 200 MB/s (19,2 GB/s)
- 
- Pozor, u paměti DDR (první generace) by při taktovací frekvence 2400 MHz, bylo možné provést 4 800 000 000 přístupů, takže by se přeneslo 4 800 000 000 x 8B za sekundu a propustnost by byla 38 400 MB/s
  - Všimněte si, že u pamětí DDR4 již tato úvaha neplatí. Kvůli vysoké latenci, je zde třeba s propustností zacházet jiným způsobem





# DDR4

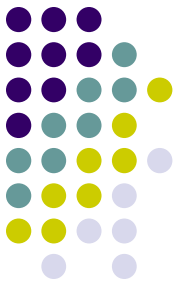
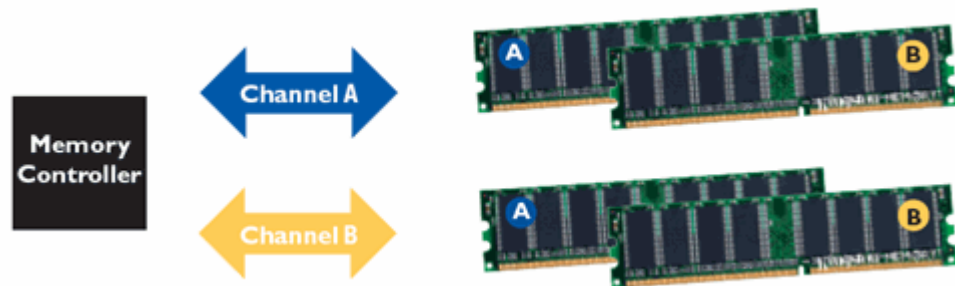
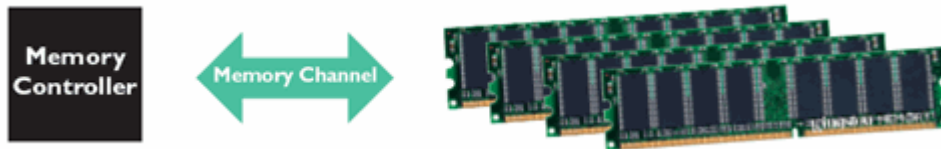




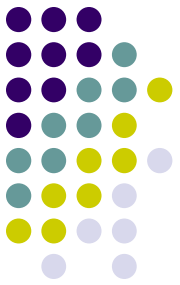
- S vývojem pamětí se CL (CAS latence) postupně zvyšuje.
- U DDR byla CL mezi 2 a 3 u DDR2 se zvyšuje na hodnoty mezi 3 a 6 a DDR3 to je obvykle někde mezi 5 až 10
- U DDR4 pak přes 12, dnes obvykle přes 15
- Taktovací frekvence pamětí se zvyšuje
- Se zvyšujícími se frekvencemi se ale zhoršuje odezva

# Multichannel

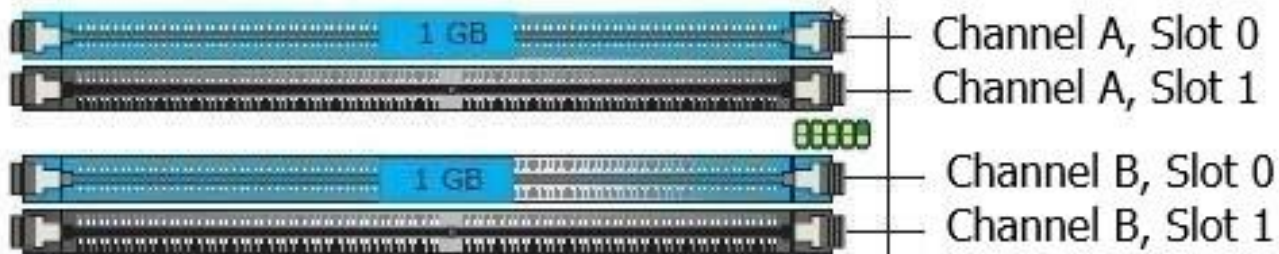
- Single-channel – řadič paměti obsluhuje paměťové moduly přes jednu sdílenou komunikační sběrnici
- Dual-channel – řadič paměti dokáže komunikovat dvěma oddělenými paměťovými sběrnici najednou. To pak zvyšuje propustnost v případě, že se čtou naráz data ležící ve dvou různých paměťových modulech a každý z nich je připojen k jiné paměťové sběrnici

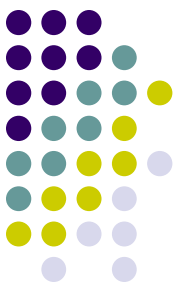


# Multichannel



- Pokud máte na základní desce 4 paměťové sloty, pak bývá obvykle barevně nebo jinak odlišeno, které dva sloty jsou na stejné paměťové sběrnici (kanálu)
- Paměťové moduly se pak vyplatí umístit do slotů tak, aby každý byl na jiné sběrnici





# Kontrolní otázky

- Proč vznikla kroucená dvoulinka? – **je lepší jak klasická dvojlinka, menší rušení**
- Vysvětlete význam vodiče GND – **slouží k měření oproti napětí**
- Vysvětlete kolik vodičů GND bude vhodné použít při paralelním přenosu 8 bitů - **8**
- Proč všechna moderní komunikační rozhraní (PCI-E, S-ATA, USB) používají sériovou komunikaci – **je rychlejší než paralelní**
- Vysvětlete rozdíl mezi synchronním a asynchronním přenosem dat – **synchronní má hodinový signál**
- Vysvětlete rozdíl mezi paralelním a sériovým přenosem dat – **paralelní přenáší více dat**
- Vysvětlete rozdíl mezi plně duplexním a poloduplexním přenosem dat – **plně duplexní může komunikovat zároveň oběma směry**
- Vysvětlete význam vodičů Tx a Rx – **odesílající pin a přijímající (Transmit, receive)**
- Vysvětlete význam vodičů Tx+ a Tx- - **Tx+ posílá standardní signál, Tx- v záporném napětí**
- Co je to přeslech? – **vzájemné rušení několika paralelních signálů**
- Vyjmenujte všechny paralelní sběrnice – **PCI BUS, ISA, EISA, MCA, PCI, AGP, PCI-X**
- Jak funguje kódování 8b10b – **z 8 bitů udělá 10 bitů, pravidelné střídání 1 a 0**
- Bývá sériová komunikace obvykle synchronní nebo asynchronní a proč? – **asynchronní**
- Bývá paralelní komunikace obvykle synchronní nebo asynchronní a proč? - **synchronní**
- Kódování 8b10b má smysl použít při synchronním nebo asynchronním přenosu a proč? - **asynchronním**
- Vysvětlete pojem „sběrnice“ – **nejjednodušší komunikační topologie, slouží k zapojení komponentu do počítače a komunikaci mezi nimi**
- Popište roli severního můstku – **část chipsetu, komunikace mezi RAM a CPU, mezi GPU a CPU**
- Proč architektura mikroprocesorů SandyBridge přesunula řadič paměti přímo do mikroprocesoru? – **uvnitř CPU může probíhat paralelní komunikace na velmi vysoké rychlosti**
- Popište roli jižního můstku – **I/O zařízení, které nepožadují takovou rychlost - klávesnice, myš, pevné disky**
- Musí být frekvence FSB a taktovací frekvence paměti stejná? – **ne, protože se dokáže provádět na jiné frekvenci komunikace s CPU a na jiné frekvenci s RAM - proto tam vlastně je**
- Vysvětlete zkratku SDRAM – **synchronní DRAM**
- Bude se nějak lišit výkon mikroprocesoru při nastavení – **frekvence CPU se odvozuje na základě FSB \* násobič (FSB – propojení CPU se severním můstkem)**
  - FSB 200 MHz, násobič 9x
  - FSB 300 MHz, násobič 6x – **bude rychlejší, protože FSB bude přenášet rychleji data**



# Kontrolní otázky

- Vysvětlete pojem „systémová sběrnice“ – **sběrnice řízená přímo CPU**
- Jaké různé varianty sběrnice PCI znáte (počet bitů, frekvence, napájecí napětí) – **32 vs 64 bitů, 33MHz vs 64 MHz, 5V vs 3.3V**
- Co je to Plug and play ? Jak fungovala instalace nové karty na starých sběrnicích bez Plug and play ? – **počítač automaticky vyhledá vhodný ovladač pro dané zařízení a sběrnici nastaví**
- Co udává u paměti hodnota CL2? – **kolik taktů musíme čekat na poskytnutí dat z paměti**
- Jaká je taktovací frekvence paměti DDR-400 – **200 MHz**
- Jaká je taktovací frekvence paměti DDR2-400 – **100 MHz**
- Kolik bitů naráz se zapisuje do paměti DDR – **64 bitů**
- Co je to RAS a CAS ? – **row access, column access – nastaví, zda vybírám sloupec / řádek**
- Proč se v paměťových obvodech nepoužívá jeden velký adresový dekodér? – **byl by velmi složitý**
- Která část adresy se použije k výběru řádku? – **první část (část zleva), vyšší bity**
- Na které sběrnici se musí používat ECC a proč? – **PCI-X (error connection code), přenáší se zde 64bitů velmi vysokou frekvencí a dochází zde k přeslechům a rušení**
- Jakou propustnost má sběrnice PCI-E 2.0 x4 **2 GB/s**
- Jakou propustnost má sběrnice PCI-E x16 **4 GB/s**
- Jakým způsobem žádá zařízení o přerušení na sběrnici PCI-E – **vyšle speciální signál (rámec)**
- Která z probíraných sběrnic má největší počet paralelně vedených signálů? – **PCI, PCI-X, EISA (pokud včetně adresových)**
- Jakou propustnost má sběrnice PCI, je-li frekvence 66 MHz, šířka dat 64 bitů. **4 GB/s**
- Proč neběží FSB na stejné frekvenci jako jádro mikroprocesor? – **CPU běží na velmi vysokých frekvencích, FSB je paralelní a proto by to nešlo**
- Podaří se přes PCI-E x8 odeslat rámec s daty rychleji, než by se stejný rámec vysílal přes PCI-E x1? – **ne, protože je to furt jen rámec**