

Sběrnice PCI, PCI-X, PCI Express

- Stručný přehled sběrnic v PC.
- Meze paralelních sběrnic.
- Důvody pro zavedení vysokorychlostních sériových protokolů do systémových sběrnic.
- Vlastnosti sběrnice PCI Express.
- Zobecnění získaných poznatků.

Historie sběrnice PCI a PCI-X

- V souvislosti s rozvojem systémových sběrnic se hovoří o generacích systémových sběrnic.

1. generace:

Sběrnice ISA – **paralelní 16 bitová, asynchronní sběrnice**, z pohledu pozdějších paralelních technologií pomalá sběrnice (synchronizace 8 - 10 MHz), šířka pásma 10 MB/s.

Zvyšování rychlosti sběrnice – **rozšiřování sběrnice** (extensivní rozvoj).

Jednoduchá komunikace – přenosy dat z prvků na straně procesoru (paměť, registry) do jiných prvků (registrů fyzicky přítomných v jiných komponentách, např. řadičích periferních zařízení).

Zabudovány mechanismy na generování **žádosti o přerušení** a **přímý přístup do paměti**.

Omezený počet prvků, které mohou (jsou schopny) řídit sběrnici:

Řadič sběrnice (po rozpoznání instrukce IN/OUT)

Řadič přímého přístupu do paměti (po generování žádosti o přímý přístup do paměti např. některým z řadičů PZ)

Struktura sběrnice ISA a její mechanické a elektrické vlastnosti neznamenal výrazný pokrok ve srovnání s technologiemi konstrukce sběrnic 60. a 70. let.

2. generace:

Sběrnice EISA, MCA, později PCI a PCI-X.

Paralelní synchronní sběrnice (PCI a PCI-X), přenosové rychlosti řádově jednotky GB/s.

Hierarchická struktura vytvořená pomocí mostů (bridge) – viz PCI, PCI-X. Vybavení sběrnice prostředky, které byly do té doby nestandardní – aplikace JTAG.

Sběrnice PCI

Standardizovaná, vysoce efektivní a spolehlivá sběrnice podporovaná několika počítačovými platformami (např. PC).

Pravidelná aktualizace **specifikací sběrnice PCI**.

3. generace:

Sběrnice PCI Express.

Sériová komunikace typu point-to-point. Možnost použití více spojů – zvýšení rychlosti přenosu a šířky pásma.

Do systémové sběrnice byla vložena sériová komunikace.

- Zobecnění trendů inovací sběrnice PCI:
Zvyšování rychlosti přenosu dat zdokonalováním elektrických a mechanických vlastností sběrnice PCI.
Příklad: snižování napájecího napětí – snižování logických hodnot přenášených přes sběrnici.
- Základní princip architektury na bázi sběrnice PCI:
Ke sběrnici procesoru je připojena přes výkonný **Northbridge** (severní most), přes nějž jsou připojena zařízení s vysokými požadavky na rychlost.
Southbridge – jižní most: připojení méně výkonných zařízení (standardních).
Bus Master – zařízení má schopnost realizovat přenosy přes sběrnici (řídí sběrnici) bez nutnosti použít prvek typu „řadič přímého přístupu do paměti“.

Komunikace mezi zařízeními ve sběrnici – přes paměť a registry – rozsah adres je konkrétnímu zařízení přidělen a je uveden v konfiguračním souboru.

Každé zařízení obsahuje sadu konfiguračních registrů – ty jsou adresovatelné a tudíž programovatelné:

⇒ každé zařízení má tři adresové prostory: paměťový, V/V registry a konfigurační registry.

Konfigurace může provádět pouze počítač, prvky bus master tuto možnost nemají.

Sběrnice PCI-X

- Rok 1999 – objevilo se rozšíření PCI-X založené na specifikaci PCI 2.3.
- Trend: další zvyšování rychlosti cestou zdokonalování elektrických a mechanických vlastností, snižování napájecích napětí – zkracování doby přepnutí mezi stavy aktivních prvků.

- Ve sběrnici PCI-X byla realizována strategie n-násobných přenosů vůči základní frekvenci 133 MHz.

PCI-X266 – dvojnásobná přenosová frekvence (za jednu periodu základního kmitočtu se přenos zrealizuje dvakrát).

PCI-X533 – čtyřnásobná přenosová frekvence (za jednu periodu základního kmitočtu se přenos zrealizuje čtyřikrát).

Vysoké kmitočty přenosu – nutnost nízkonapěťového rozhraní (1,5 V) a využití technik na opravu chyb – ECC (Error Correcting Codes – samoopravné kódy).

- Poznámka: AGP vychází ze sběrnice PCI.

Tabulka 1: Přehled konfigurace jednotlivých typů PCI sběrnice

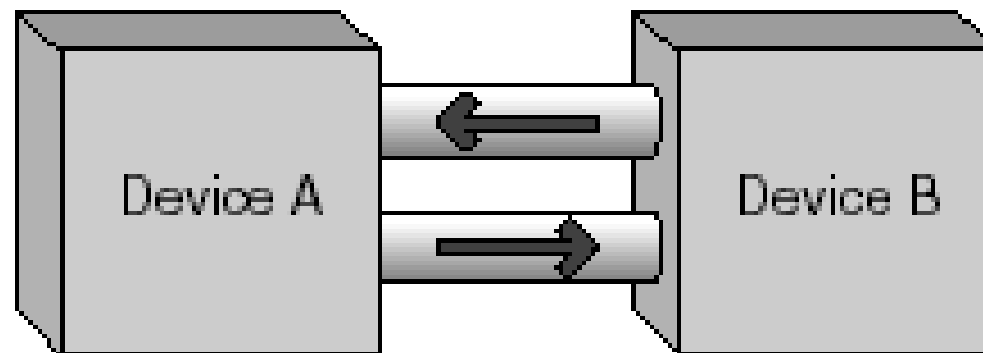
Typ sběrnice	PCI - 33		PCI - 66		PCI-X 66		PCI-X 133		PCI-X 266			PCI-X 533		
Počet datových bitů	32	64	32	64	32	64	32	64	16	32	64	16	32	64
Počet pinů ¹	49	81	49	81	50	82	50	82	36	50	82	36	50	82
Přenosová rychlost MB/s ²	133	266	266	533	266	533	533	1066	533	1066	2133	1066	2133	4266
Napájecí napětí	5/3,3V		3,3V		3,3V		3,3V		1,5V a 3,3V			1,5V a 3,3V		

- Trend: zvyšování základní frekvence, více přenosů realizovaných v rámci jedné periody základního kmitočtu (v této tabulce nezachyceno), snižování napájecího napětí.

Sběrnice PCI Express (PCI E)

- O sběrnici PCI Express se hovoří jako o stavu, který nastal, když křivka reflektující vývoj systémových sběrnic ustrnula na inflexním bodě, poprvé se v sestavách počítačů objevila v r. 2004.
- Požadavky na PCI E (definované v době, kdy se rozhodlo o vytvoření této sběrnice):
 - rozšiřitelnost (scalability) základního spoje s cílem dosáhnout vyššího výkonu),
 - efektivnější využívání spoje (např. vývodů konektoru) – průměrná rychlost realizovatelná přes jeden vývod konektoru,
 - cílová skupina výpočetních prostředků: personální počítače (stolní i přenosné, servery, komunikační prostředky),
 - perspektiva využití v budoucích sestavách vyžadujících vysokou rychlost přenosu dat.

- Výrazný požadavek: stabilita architektury v mnoha dalších aplikacích v budoucnu (vyšší rychlost – do PCI E se zařadí další spoj, architektura se výrazně nemění).
- Charakteristika:
PCI Express je **vysokorychlostní, sériový, nízkonapěťový, diferenciální spoj** pro komunikaci mezi dvěma zařízeními.
Spoj je realizován jako dvojité jednosměrné propojení (dva jednosměrné spoje).



- Spojení **Full duplex** – současně je možné přenášet data v obou směrech maximální možnou frekvencí.
- Jiný princip: **Half duplex** – je možné přenášet data pouze v jednom směru.

- Proč vysokorychlostní?

Rychlost přenosu: 2,500,000,000 bitů/s (2,5 Gb/s) - Standard PCI Express 1.1 (250 MB/s) – v jednom spoji (hodnota Gb/s se dělí 10 – používá se kódování 8b/10b).

Spoj PCI Express x 16 – rychlost 4 GB/s (250 MB/s v jednom spoji).

Standard PCI Express 2.0 – 5 Gb/s (500 MB/s v jednom spoji).

Standard PCI Express 3.0 – 10 Gb/s (1 GB/s v jednom spoji).

Standard PCI Express 4.0 – 20 Gb/s (2 GB/s v jednom spoji).

Standard PCI Express 5.0 – 40 Gb/s (4 GB/s v jednom spoji).

- Proč nízkonapěťový?

Pracuje se s napětími 0,8 – 1,2 V (technika LVDS – Low Voltage Differential Signalling).

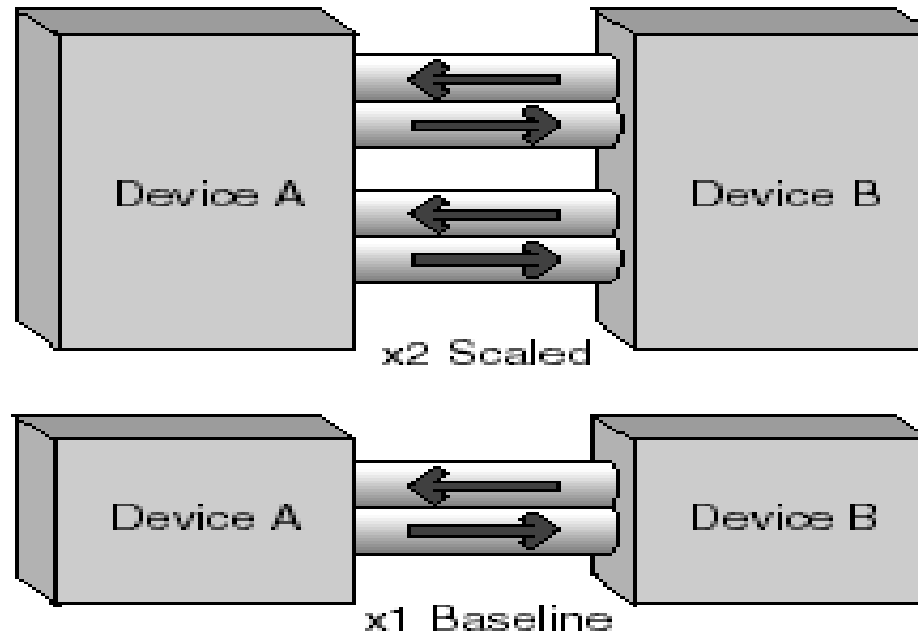
Vyšší šířka pásma sběrnice PCI Express než šířka pásma sběrnice PCI, počet vývodů konektoru je přibližně 10x menší (relace „rychlost“ versus „počet vývodů konektoru“ – jeden ze způsobů posuzování rychlosti přenosu – vychází v PCI E výrazně lépe).

- Možnosti PCI Express jsou rozšiřitelné (scalable) v tom smyslu, že přidáním dalšího spoje se zvýší šířka pásma.
- PCI Express nahrazuje AGP, PCI, PCI-X, její současné schopnosti nejsou konečné, náhrada AGP v konstrukci grafických adaptérů.

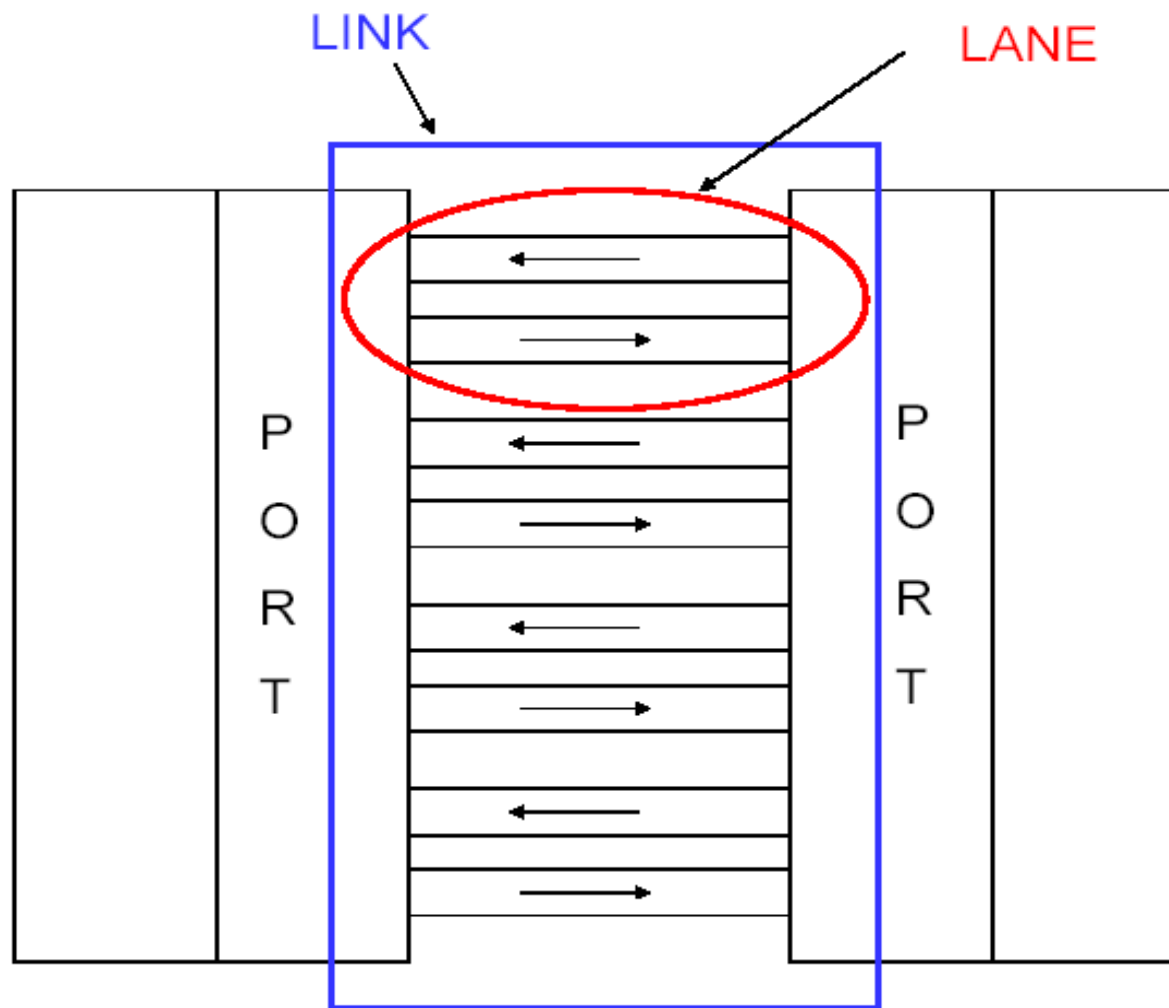
- Výhoda: point-to-point spoje se neovlivňují, pokud potřebujeme na jednom spoji realizovat větší šířku pásma, ostatní spoje to neovlivňuje.
- PCI Express Base Specification:
Jsou definovány tyto konfigurace sériových spojů: x1, x2, x4, x8, x12, x16, and x32

Příklad: konfigurace **x1** představuje jeden spoj oběma směry.

Konfigurace **x2** představuje dva sériové spoje (každý pro oba směry).



- V počítačích se sběrnici PCI Express není zapotřebí arbitr pro rozhodnutí o přidělení sběrnice pro přenos dat – spoj point-to-point.
- Zařízení na spoji PCI Express může kdykoliv do sériového spoje poslat data, stejně tak musí být připraveno ze vstupního spoje data přijmout. Každý spoj je dedikovaný.



- Terminologie:

Link - spoj mezi dvěma zařízeními na PCI Express.

Lane – dvě dvojice vodičů, přes něž se přenáší diferenciální signál (dva směry) - vysílání a příjem.

Žádné další adresové, řídicí signály a datové signály (typické např. pro sběrnici PCI, komunikace s využitím paketů).

Architektura spoje PCI Express je modulární – architektura je rozšiřitelná z hlediska počtu spojů – **důležité z hlediska perspektivy sběrnice** (je rozšiřitelná).

Poznámka: pokud je v architektuře počítače prvek, který není tzv. rozšiřitelný (scalable), pak „nemá v architektuře perspektivu“ – snaha o změnu.

- Pracuje se s kódováním 8b/10b.

- Závěr: sběrnice PCI Express může být „nastavena“ pro řadu aplikací, protože má nastavitelnou různou rychlost přenosu. Proto má sběrnice PCI Express široké použití v řadě aplikací.

Jak se počítá výkon sběrnice pro potřeby srovnání sběrnic?

- Vyjadřuje se v průměrné rychlosti – celková rychlost přepočtená na vývod konektoru (pin) nebo spoj.
- V tomto parametru je PCI E nejefektivnější ze všech doposud používaných systémových sběrnic (pin-efficient, link-usage-efficient).
- Parametr Bandwidth-per-Pin (šířka pásma přepočtená na jeden vývod).

Konvenční sběrnice PCI

- *Kmitočet: 33 MHz * 4 slabiky = 133 MB/s*
133 MB/s : 74 vývodů = 1,8 MB/vývod

Sběrnice PCI-X

- *PCI-X 2.0: 533 MHz, datová část systémové sběrnice 64 bitů obousměrných (8 slabik), 150 vývodů.*

$$533 \text{ MHz} * 8 \text{ slabik} = 4264 \text{ MB/s}$$

$$4264 \text{ MB/s} : 150 \text{ vývodů} = 28,4 \text{ MB/vývod}$$

Sběrnice PCI E

- *PCI E sestává ze dvou spojů, má k dispozici dva jednosměrné nezávislé spoje, k 8 bitům přidává další 2 bity (kódování 8b/10b)*

$$(2500 \text{ MHz} * 1 \text{ bit}) / 10 \text{ b} = 250 \text{ MB/s},$$

$$\text{Výsledný výkon: } 500 \text{ MB/s}$$

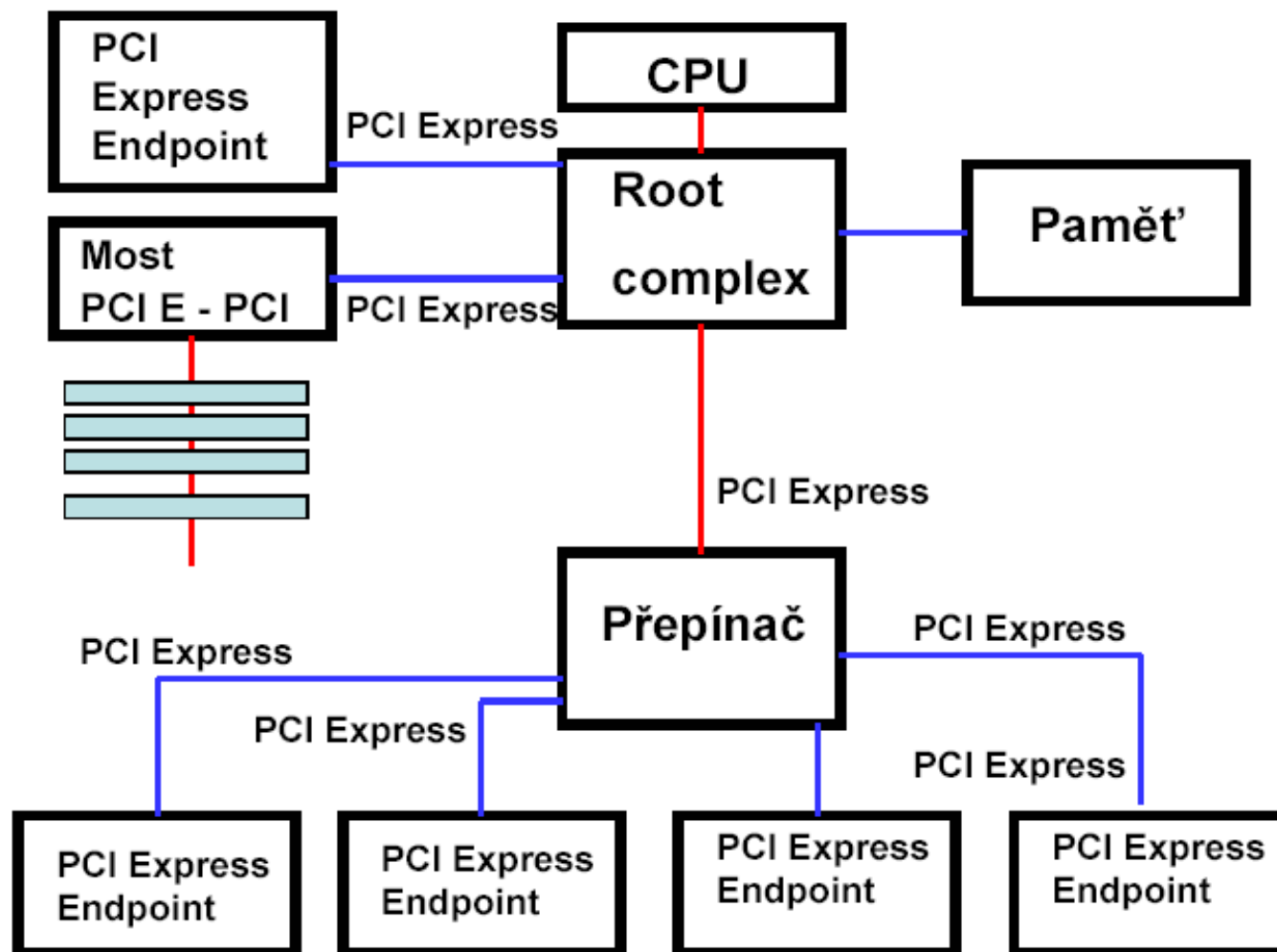
Vývody: 4 pro data (2 pro Rx), (2 pro Tx), 2 pro napájení, 2 pro „zem“).

$$500 \text{ MB/s} : 8 \text{ vývodů} = 62,5 \text{ MB/vývod}$$

Nárůst počtu „lane“ – větší efektivita využití napájecích spojů, (nenarůstá) s počtem „lane“ – dosáhne se až 100 MB/vývod.

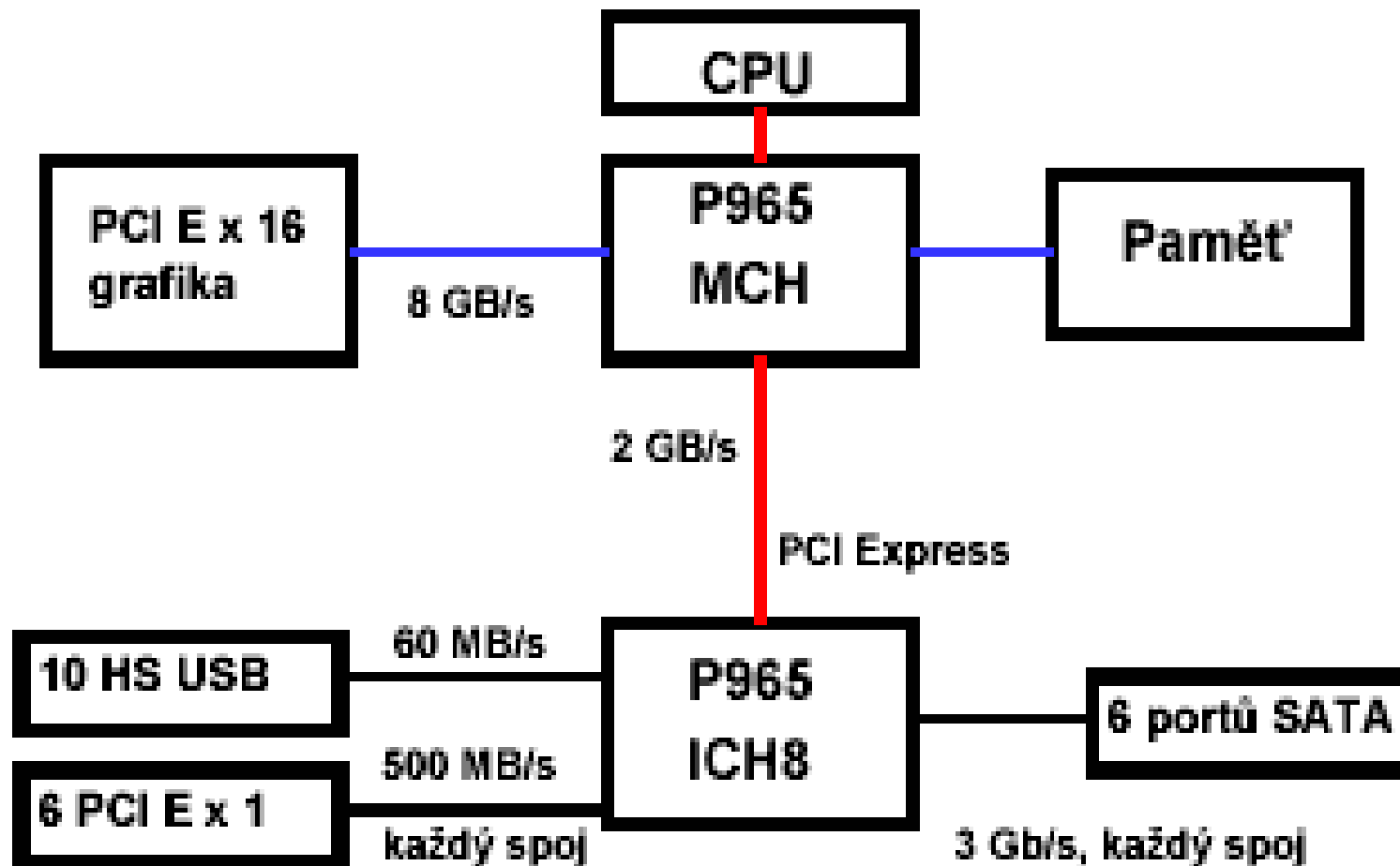
- Závěr: v parametru výkon/vývod je sběrnice PCI E zatím nejefektivnější.

Typy zařízení na sběrnici PCI Express



- Root Complex:
Propojení V/V podsystemu na procesor a paměť.
Může to být např. prvek (G)MCH (Graphics & Memory Controller Hub) nebo jeho kombinace s ICH (I/O Controller Hub).
- Přepínače.
- Most PCI E na PCI.
Možnost koexistence PCI Express s technologií PCI/PCI-X.
- Každý spoj je realizovatelný na různých rychlostech.
- Kompatibilita zdola – existuje možnost instalace zařízení vybavených sběrnici PCI.

Reflexe obecné architektury v čipové sadě fy Intel

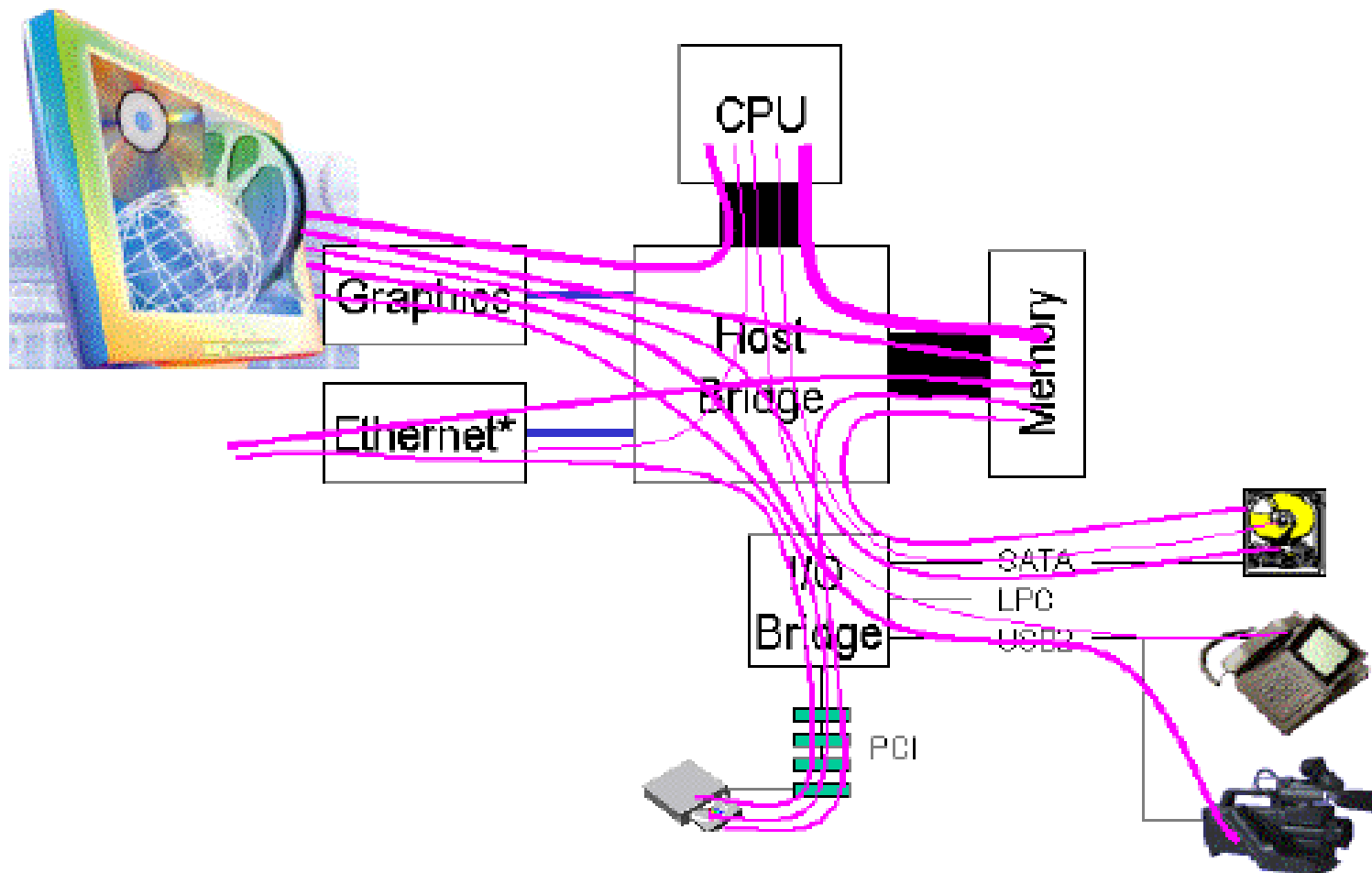


- Poznámky:

Grafický adaptér připojen přes PCI Express x 16, každý spoj 500 MB/s (5 000 Gb/s).

Konkrétní architektura typická pro konkrétní sběrnici si vynutí existenci **čipové sady**.

Toky dat v počítači se sběrnicí PCI Express



Shrnutí:

- PCIe, je na rozdíl od předcházejících typů sběrnic koncipována jako několik point-to-point seriových spojů (lanes) – nikoliv jako nededikovaný paralelní spoj. **Sběrnice PCIe je sítí point-to-point spojů.**
- Tyto spoje jsou směrovány v rozbočovači – ten funguje jako přepínač. Tam jsou tyto spoje ustaveny.
- Tato koncepce dovoluje, aby v konkrétním okamžiku spolu komunikovalo více dvojic zařízení.
- Rozdíl oproti dřívějším sestavám se sběrnicí PCI – **na sběrnici mohlo vkládat informaci pouze jedno zařízení.**
- **PCI E - možnost násobných existencí point-to-point spojů a přenosů mezi více zařízeními.**
- Počet spojů (lanes) je „dohodnut“ při zapínání nebo až v průběhu činnosti – může se odehrávat dynamicky (tzn. až v okamžiku potřeby).

- Principy práce se sběrnici PCIe umožňují **nepoužívat vůbec mechanismy přidělování sběrnice, nemusíme řešit problémy s čekáním na uvolnění sběrnice** (tzn. klasické problémy známé z paralelních sběrnic).
- Principy PCIe jsou vyvíjeny a udržovány skupinou PCI-SIG (PCI Special Interest Group) – skupina více jak 900 společností, které udržují také specifikaci klasické sběrnice PCI.

Transakce na sběrnici PCI Express

- Tzv. „split-transaction“ protokol (rozdělený protokol).
Transakce sestává ze dvou částí: **požadavek** (request) a **provedení** (completion).
- Mechanismus:
Iniciátor transakce (žadatel – **requester**) vyšle paket „žádost“.
Cíl (**completer**) paket přijme, provede požadované činnosti, pak odpoví paketem (tzv. **completion packet**).
Paměťové transakce – 32 bitů / 64 bitů adresy
- Dvě zařízení spojená spojem nemusí být nutně ve vztahu **requester – completer**.
- Příklad: přepínač dostane transakci z root complex (bude dále poslána do Endpoint), přepínač není completer.
Endpoint dostane transakci z přepínače, přepínač není requester.

Typy transakcí

Paměťové transakce

- Operace (zápis/čtení) – prvky adresované v adresovém prostoru paměti – *Memory Read Request*, *Memory Read Completion*, *Memory Write Request*.

V/V transakce

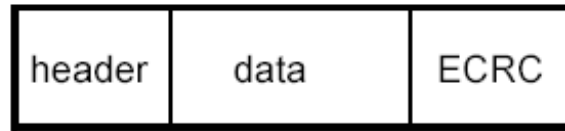
- Operace (zápis/čtení) – prvky adresované ve V/V adresovém prostoru – *I/O Read Request*, *I/O Read Completion*, *I/O Write Request*, *I/O Write Completion*.

Konfigurační transakce

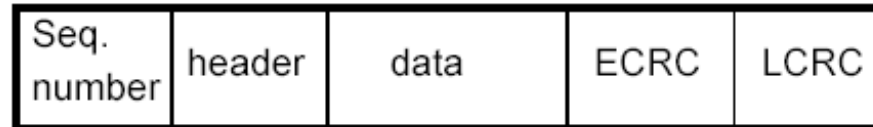
- Operace (zápis/čtení) – prvky adresované v konfiguračním adresovém prostoru – *Configuration Read Request*, *Configuration Read Completion*, *Configuration Write Request*, *Configuration Write Completion*.

Transakce typu „zpráva“

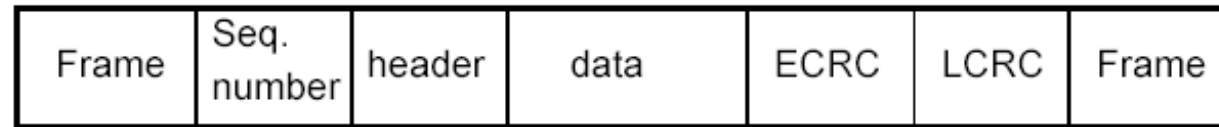
- Využívá se např. při chybových hlášeních (nahrazuje signály PERR – Parity Error, SERR – System Error, které jsou k dispozici v klasické sběrnici PCI), power management.



Transaction Layer



Data Link Layer



Physical Layer

ECRC end-to-end CRC

LCRC link CRC

- **Transaction Layer (transakční vrstva)** – konvertuje požadavky na nějakou činnost do paketů (TLP - Transaction Layer Packet)
Header (hlavička) - typ transakce (I/O, paměť, konf. prostor, zpráva)
 - Požadavek na přenos z paměti (délka dat a adresa paměti):
Na úrovni Transaction Layer se vytvoří paket a doplní se hlavičkou.
Např. hlavička pro paměťové transakce obsahuje adresu, s níž se bude pracovat (a další informace).
Typ požadavku:
MRd – Memory Read Request
MWr – Memory Write Request
IORd – I/O Read Request
IOWr – I/O Write Request
- Důležité:** v paketu (její hlavičce) je k dispozici struktura, která obsahuje informace, které byly v paralelní sběrnici přenášeny paralelně.

- **Data Link Layer (spojovací vrstva)**

Paket je doplněn o **sequence number** (pořadové číslo paketu), pro následnou kontrolu pořadí paketů na straně příjemce – spojovací vrstva má v sobě integrovány kontrolní funkce.

Další funkce na této úrovni na vysílací straně: **doplnění LCRC (Link CRC)**.

Kontrola LCRC na straně příjemce.

- **Physical Layer (fyzická vrstva)**

Je reprezentovaná obvody, které jsou potřebné pro realizaci protokolu: **vyrovnávací paměti, obvody pro transformaci paralelní informace na sériovou a naopak, obvody fázového závěsu (generování synchronizace), impedanční přizpůsobení sériových linek.**

Konverze 8b na 10 b.

Konverze paralelních dat na sériová.

Doplnění „frame“ – rámeček, vymezení začátku a konce paketu.

- Na přijímací straně je postup opačný ve všech vrstvách. Na přijímací straně se kontroluje to, co bylo na vysílací straně na příslušné úrovni do paketu vloženo.

Co je dále řešeno ve sběrnici PCI Express:

- Data Scrambling.
- Kódování 8b/10b.
- Synchronizace generovaná na přijímací straně s využitím parametru RL – nutnost jeho dodržení.

Data Scrambling

- Význam slova „Scramble“ – zakódovat (kromě jiného).
- Cíl: odstranění jevu, který je označován jako „elektromagnetická interference (EMI)“.
- Elektromagnetická interference (vyzařování): vznikne jako výsledek opakovaných přenosů stejných vzorů dat přes spoj.
- Přenášená data jsou přenášena přes LZPR (Lineární Zpětnovazební Posuvný Registr – Linear Feedback Shift Register - LFSR).
- LZPR realizuje polynom $G(x) = X^{16} + X^5 + X^4 + X^3 + 1$, tzn. vysílaná data jsou dělena tímto polynomem.
- Scrambling (vysílací strana) – descrambling (přijímací strana).

Kódování 8b/10b

- Účel: „vlození“ informace o synchronizaci (parametr RL) do přenášených dat.
- Poznámka: uvědomme si problém „clock skew“ při paralelních přenosech – posunutí některého z bitů proti hraně vzorkovacího signálu.

Tentýž problém pokud sériově přenášená data jsou na přijímací straně synchronizována synchronizací přenášenou samostatným spojem.

Řešení: „vlození“ informace o synchronizaci (dodržení parametru RL) do přenášeného signálu.

Vložení informace o synchronizaci – uspořádání posloupnosti datových bitů tak, aby se v ní pravidelně vyskytovaly změny signálu. Od těchto změn signálu je možné na přijímací straně zesynchronizovat obvody fázového závěsu, které generují synchronizační pulsy.

- Výsledek překódování: „00h“ v bitové podobě „00000000“ (8 bitů) se překóduje na posloupnost: „1101000110“ (10 bitů).

Výsledek:

Ve vysílaných posloupnostech bitů reprezentujících jednotlivé znaky se vyskytují tyto alternativy (totéž jako v SATA):

pět „1“ a pět „0“

čtyři „1“ a šest „0“

čtyři „0“ a šest „1“

Změna vysílaného signálu tak, že je splněn požadovaný parametr RL (Run Length Limited) – na přijímací straně je možné tyto změny využít na zesynchronizování PLL (phased-locked loop – fázový závěs).

Průběžná disparita

- Počet „1“ a počet „0“ přenášený přes kabel je shodný.
- Využití pro detekci chyb na přijímací straně.

Klidová úroveň přenášeného napětí

- Sleduje se průběžná disparita na vodiči, přes nějž se signál přenáší – musí se přenášet stejný počet „0“ a „1“.

Závěr: technika kódování dat využitá v PCI E je shodná s technikou využitou v SATA.