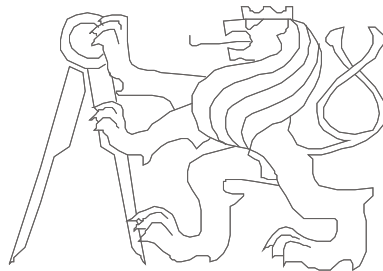


Pokročilé architektury počítačů

Architektura IO podsystému

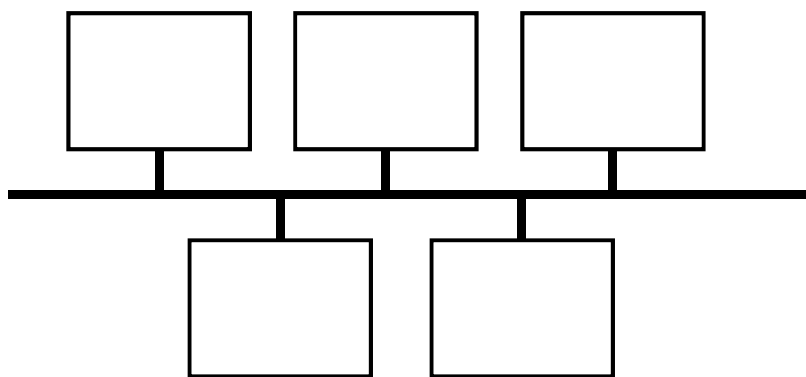


České vysoké učení technické, Fakulta elektrotechnická

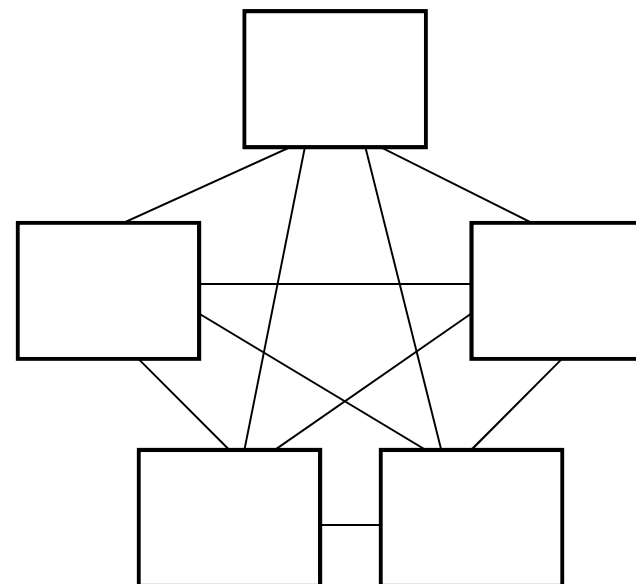
Co je úkolem?

- Propojit jednotlivé části výpočetního systému
- Požadavky:
- Vytvořit optimální datové cesty hlavně pro nejdůležitější periferie (vnější paměti).
- Možnosti řešení:
- S ohledem na závislost cena/výkon existuje hranice výkonnosti, kdy
 - datové cesty je možné sdílet, nebo
 - datové cesty je výhodné sdílet.


Připomenutí: sběrnice x dvoubodový spoj

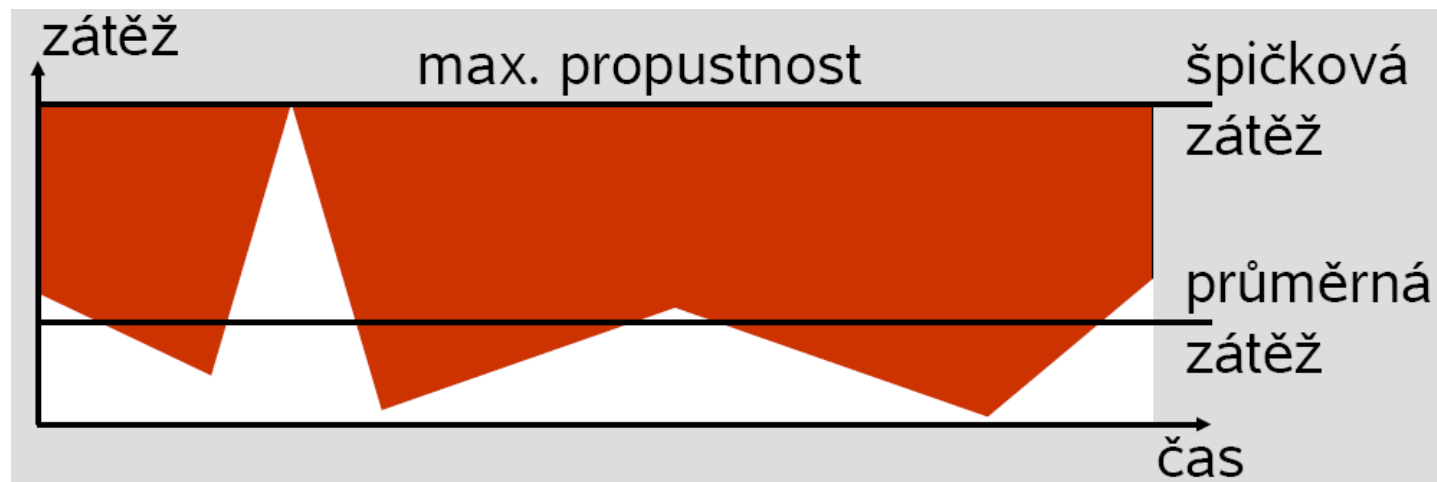


- Anglicky:
- Bus,
- point-to-point connection.

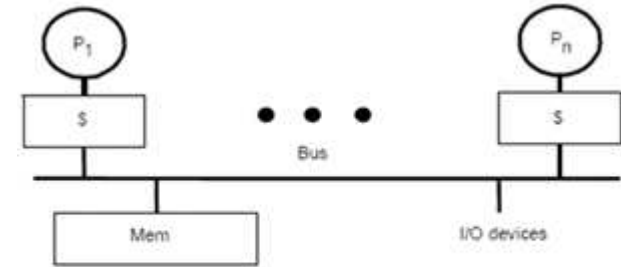


Průměrná a špičková zátěž

- zátěž: objem přenášených dat za čas
 - čas $\rightarrow 0$... špičková zátěž,
 - čas $\rightarrow \infty$... průměrná zátěž.
- Nevyužitá propustnost 

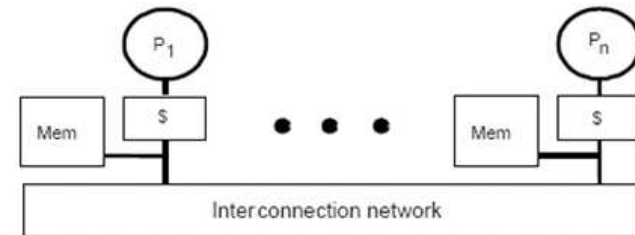
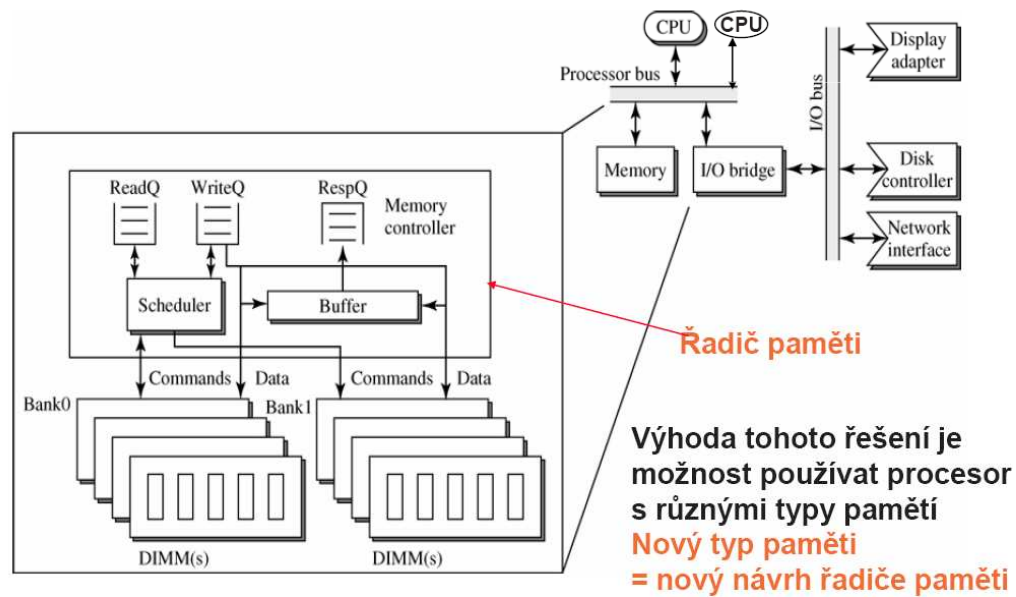


Obecně architektura počítače



tradiční SMP

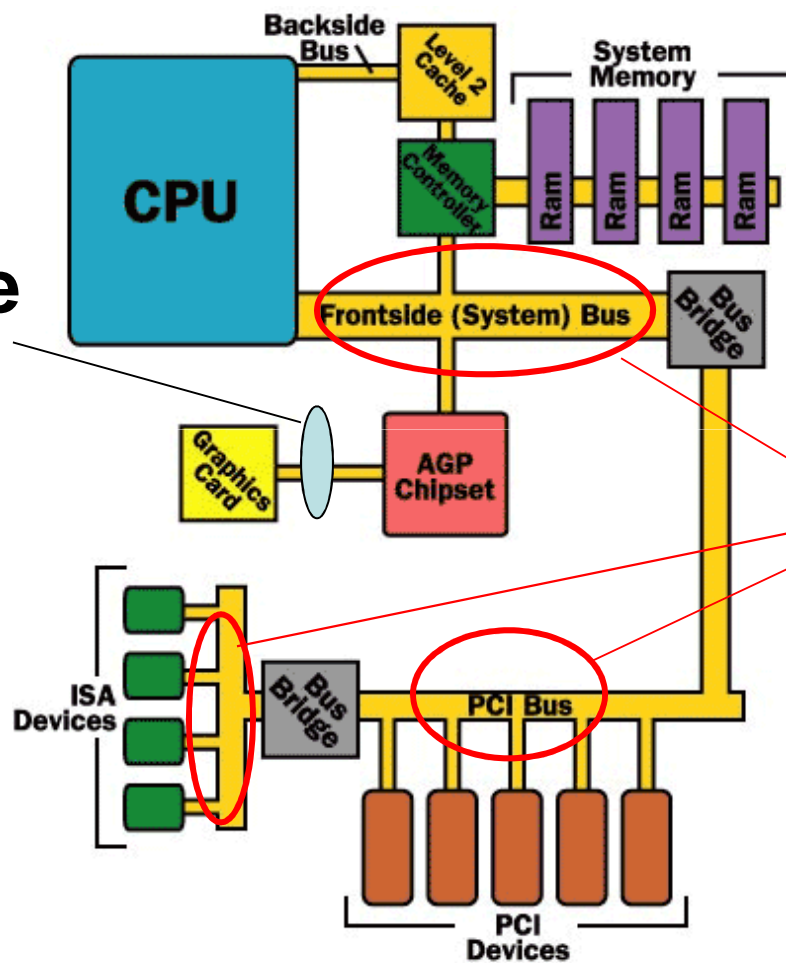
Organizace hlavní paměti – tradiční schema (Intel a další)



NUMA

Příklad z PC prostředí:

Pozor! Tohle je
dvoubodový
spoj!



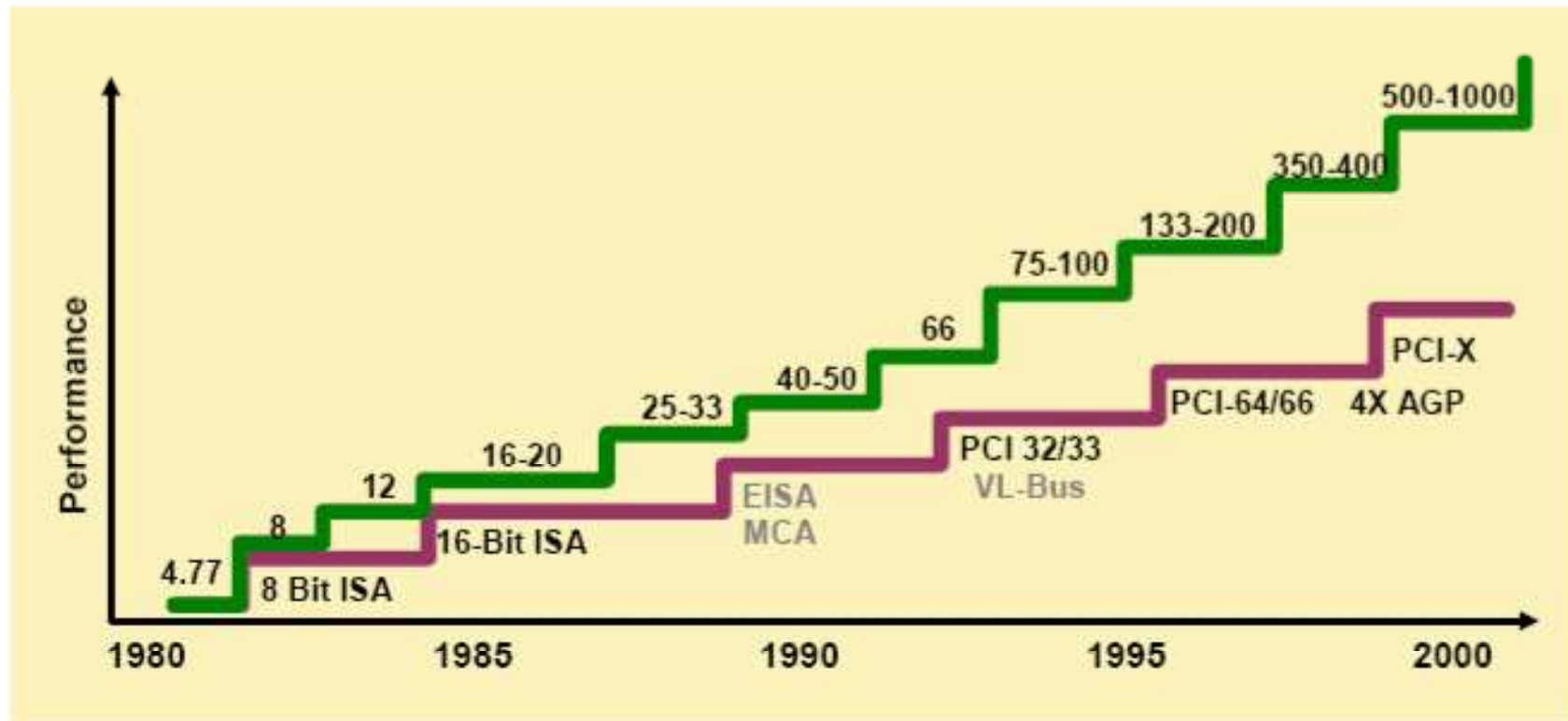
Sběrnice

©2001 HowStuffWorks

The I/O Bandwidth Problem

- While microprocessor performance continues to double every eighteen months, the performance of the I/O bus architecture has lagged, doubling in performance approximately every three years.
- Every time processor performance doubles, latency only increases by a factor of 1.2.

The I/O Bandwidth Problem



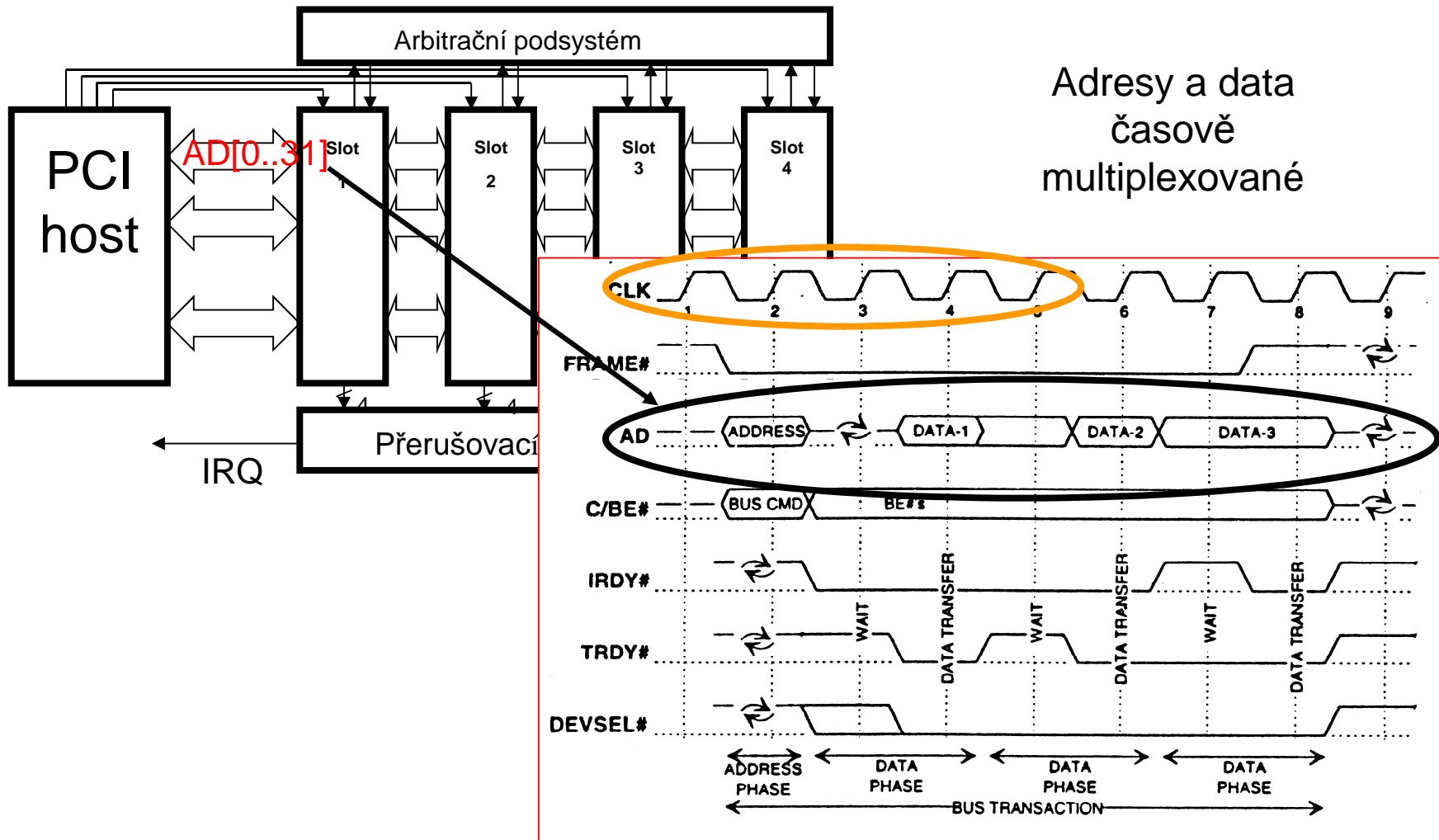
Terminologie interfejsingu I.

- **Interfejs** = interface = rozhraní = propojení = styk = mezixicht.
 - Společná komunikační část sdílená dvěma systémy, zařízeními nebo programy.
 - Zahrnuje i prvky této hranice a doplňkové řídicí obvody určené k jejich propojení.
- Sběrnice x dvoubodový spoj.
- Adresová, datová, řídicí sběrnice.
- Brána.
- Multiplexovaná/oddělená sběrnice.
- Procesorová, systémová, lokální, V/V sběrnice.

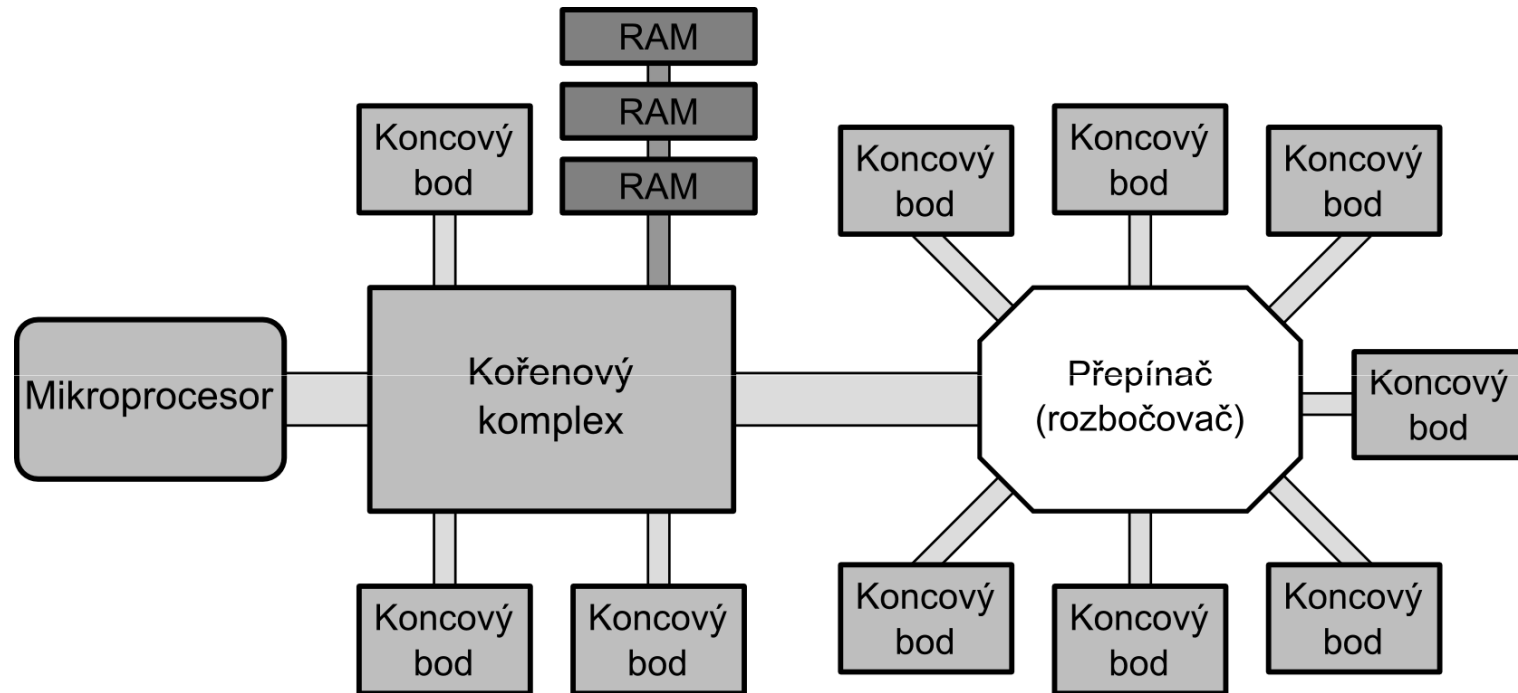
Ale pozor!

- Point-to-point
 - není totéž co
- peer-to-peer!
 - Peer-to-peer arch. = rovný s rovným
- a k tomu protějšek je
 - architektura klient-server!

Multiplexovaná sběrnice?



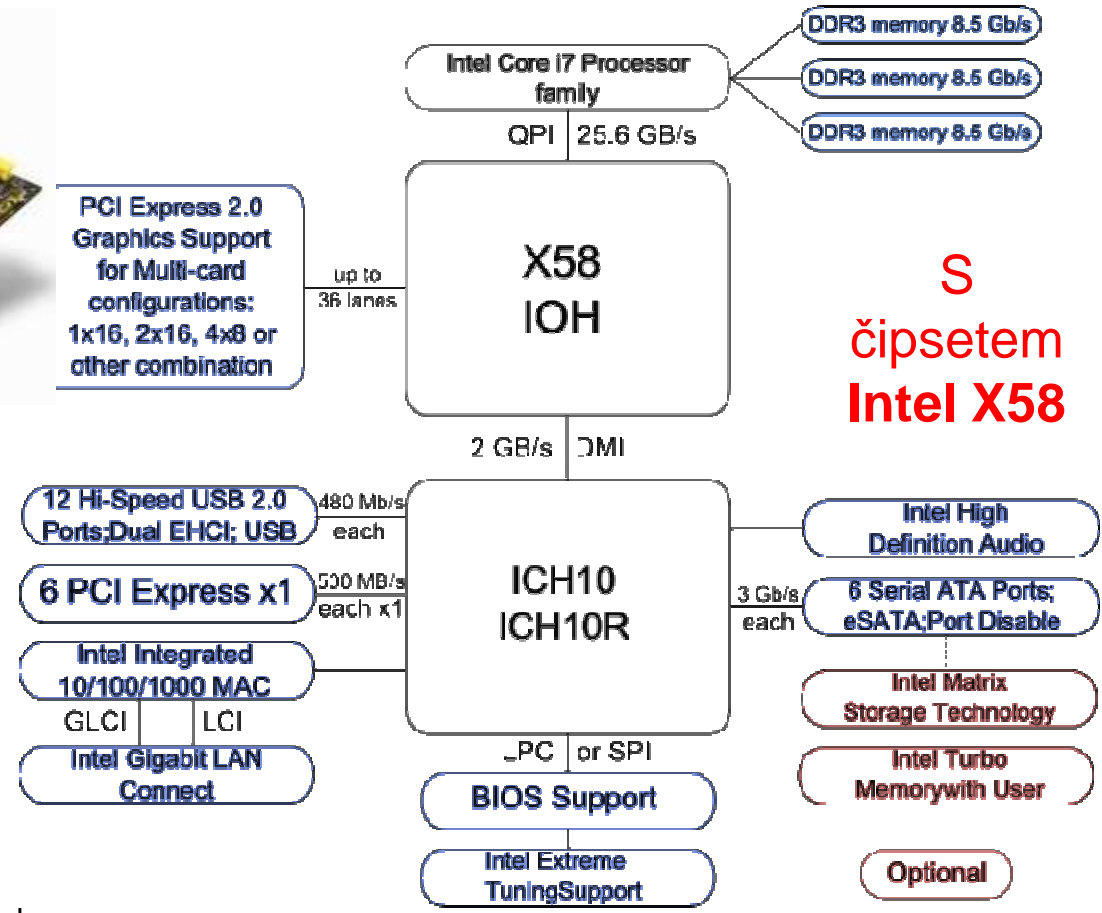
Architektura dnešního PC



Aktuální systémová deska (02/09)



PCI Express 2.0 Graphics Support for Multi-card configurations: 1x16, 2x16, 4x8 or other combination

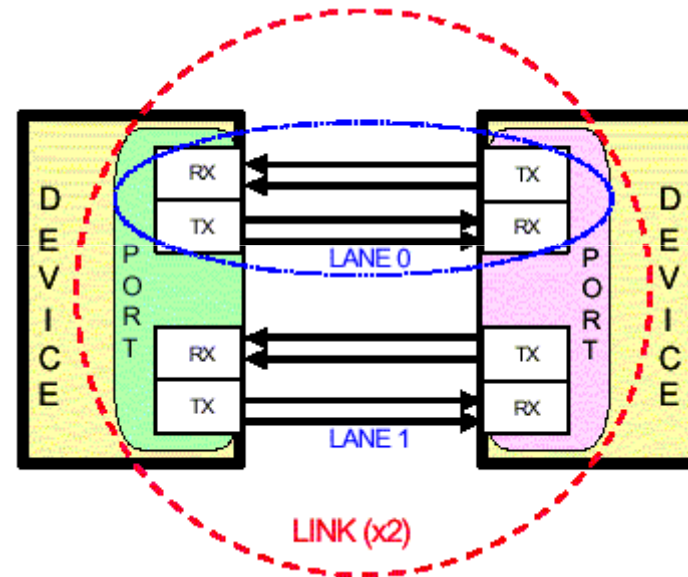


S
čipsetem
Intel X58

Zdroj původního odkazu:
http://www.svethardware.cz/art_doc-2E98BCAFAE7771A1C1257551005B666F.html

PCI Express

- Proud (lane) tvoří 4 vodiče.
- Podle počtu proudů se určuje link (x2).
- Jedná se o jednosměrný dvoubodový spoj!



Electrical Configuration

- The signaling technology used in HyperTransport technology is a type of low voltage differential signaling (LVDS).
- LVDS has been widely used in these types of applications because it requires fewer pins and wires.
- Cost and power requirements are reduced because the transceivers are built into the controller chips.

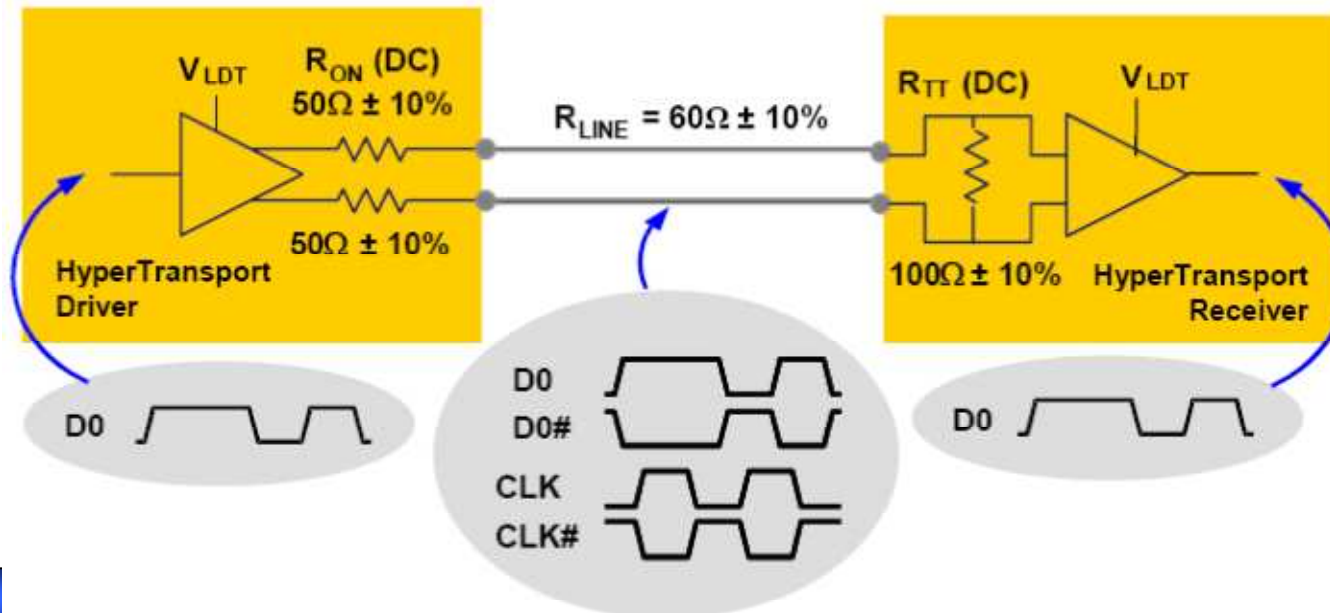
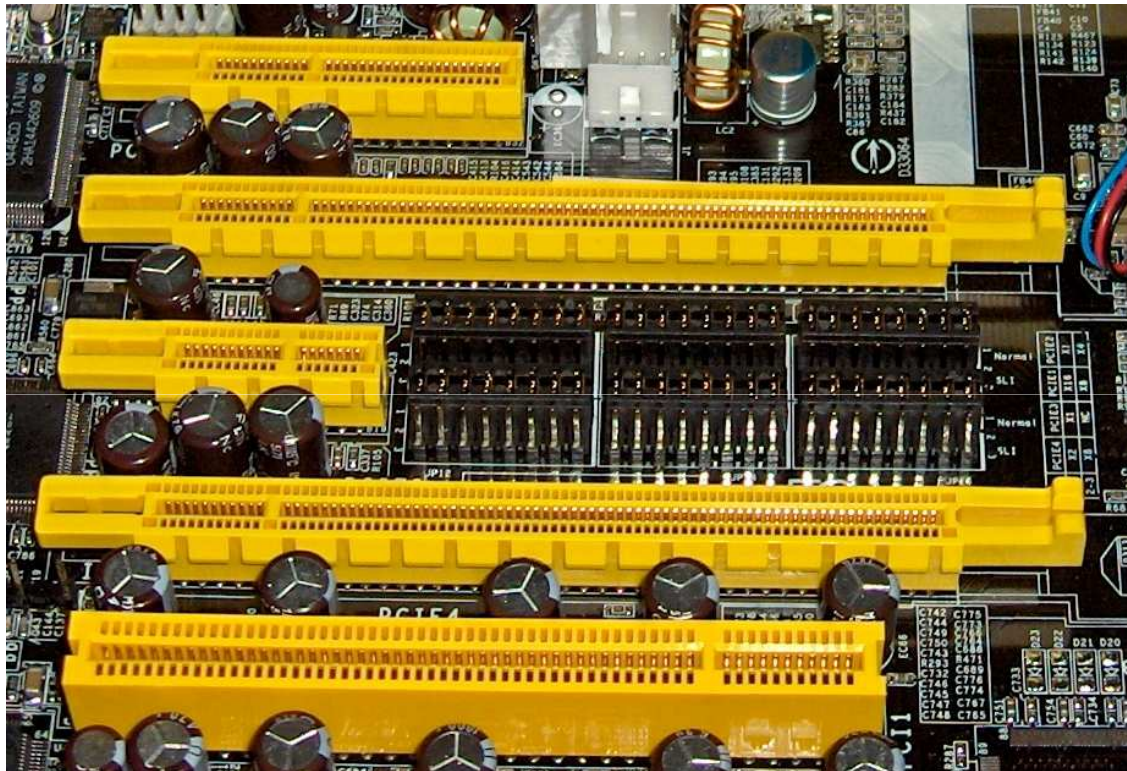


Figure 4. Enhanced Low-Voltage Differential Signaling (LVDS)

PCIe



PCI Express x4,

x16,

x1 a

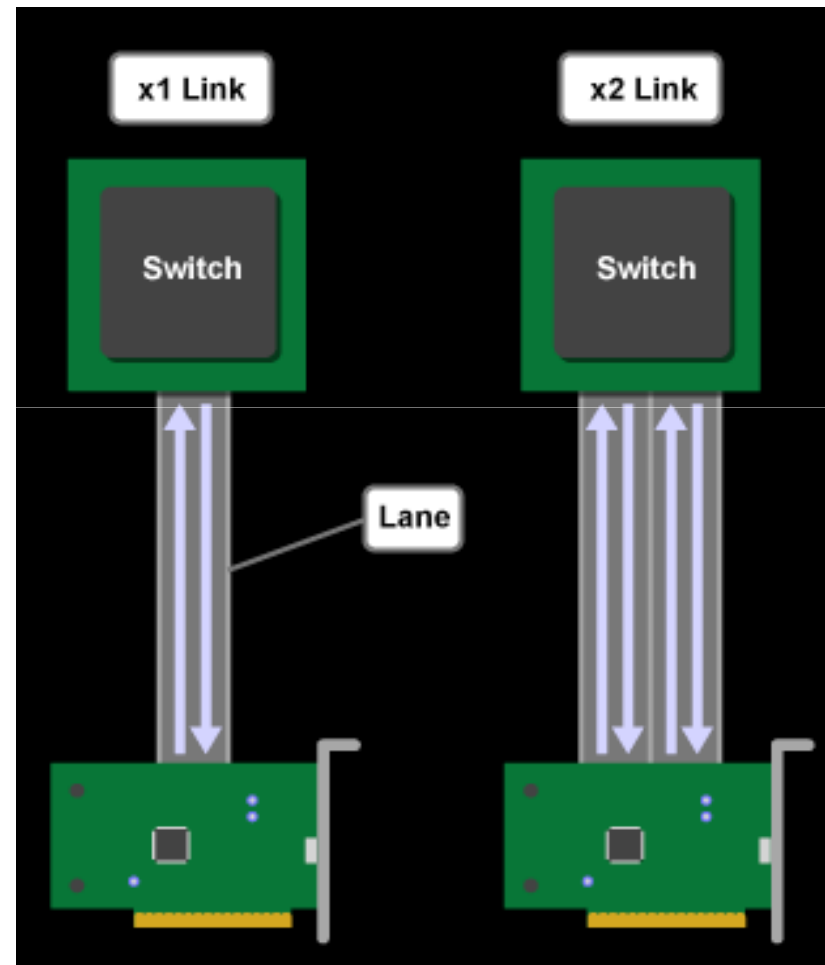
x16 sloty,

spolu se standardním 32b PCI slotem

na desce LanParty nF4 Ultra-D jednoho výrobce (DFI).

PCIe Links and Lanes

- Each link consists of one or more lanes
 - Each lane is 1-bit wide (4 wires, each 2-wire pair can transmit 2.5Gb/s in one direction)
 - Upstream and downstream now simultaneous and symmetric
 - Each Link can combine 1, 2, 4, 8, 12, 16 lanes- x1, x2, etc.
 - Each byte data is **8b/10b** encoded into 10 bits with equal number of 1's and 0's; net data rate 2 Gb/s per lane each way.
 - Thus, the net data rates are 250 MB/s (x1) 500 MB/s (x2), 1GB/s (x4), 2 GB/s (x8), 4 GB/s (x16), each way



Perspective

- What about PCI Express?



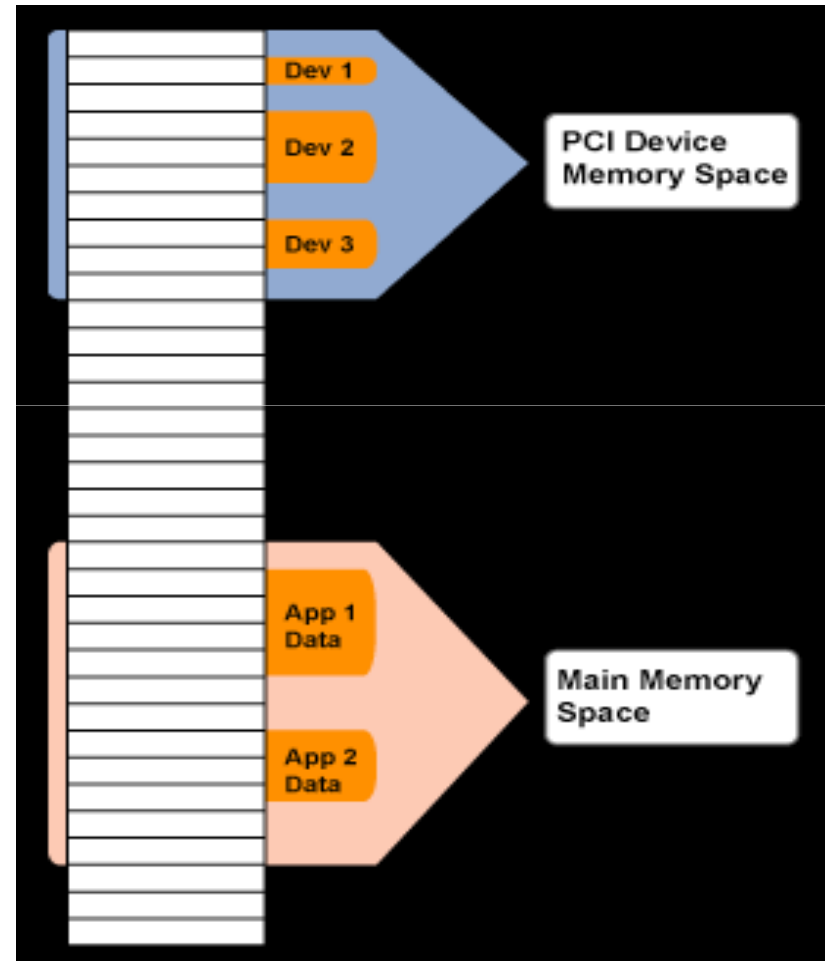
LVDS

Each data lane of a PCI Express card transmits 250 Mbytes/s in each direction.

As shown here an 8 lane PCI express connector has 49 total pins

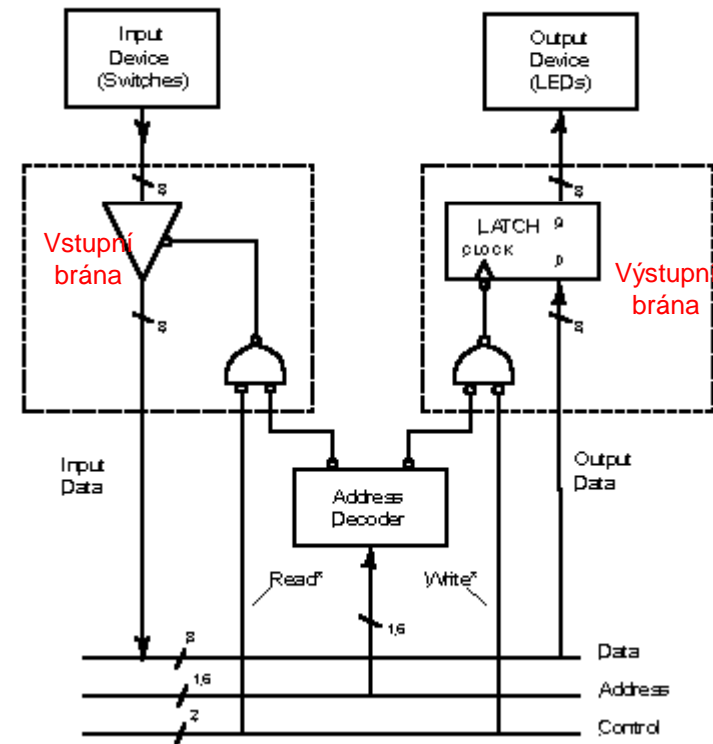
PCI as Memory Mapped I/O

- PCI device registers are mapped into the CPU's physical address space
 - Accessed through loads/ stores (kernel mode)
- Addresses assigned to the PCI devices at boot time
 - All devices listen for their addresses



Brána, anglicky port

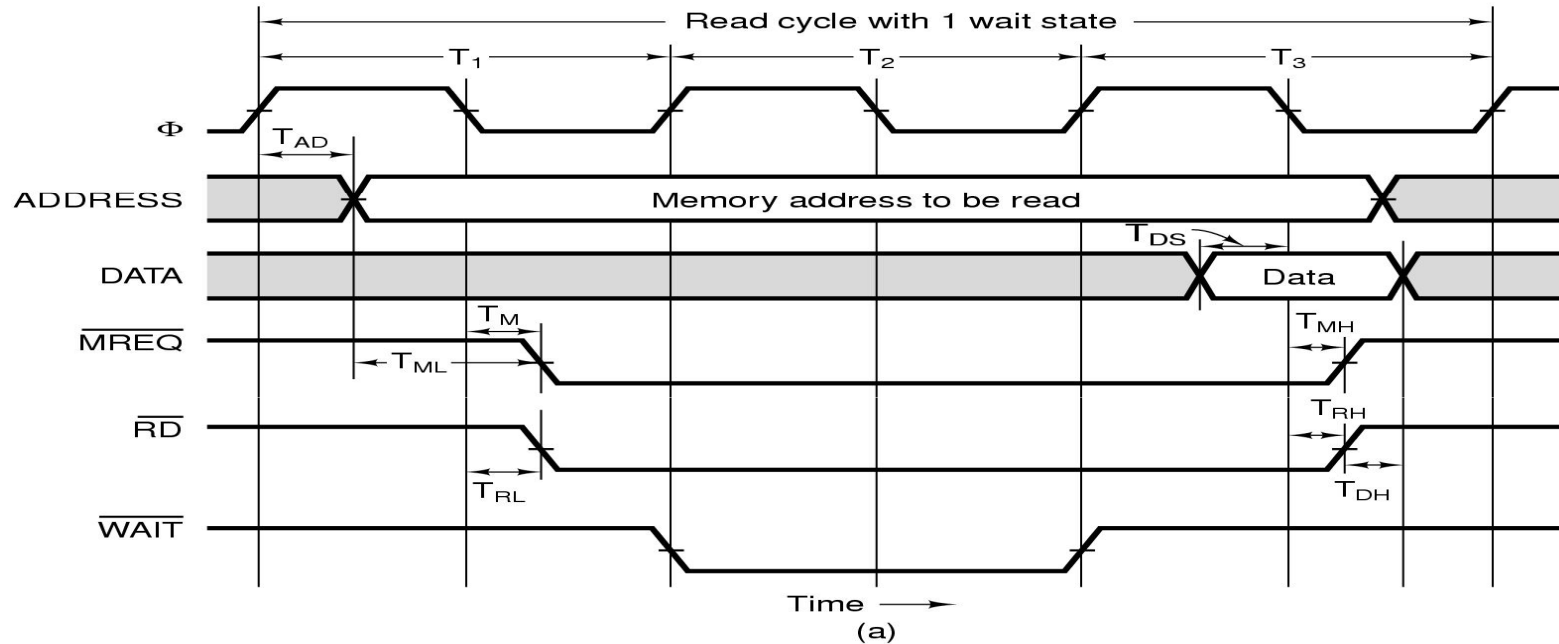
- Buňka v adresovatelném prostoru V/V zařízení nebo paměti.
- Stavební prvek interfejsingu.
- Obecně je to 8/16/32 b buňka, registr, ale ...
- uvidíme, že někdy se bez klopných obvodů obejde. Podrobnosti ve cvičeních.



Připomenutí - terminologie interfejsingu

- Synchronizace přenosu údajů
 - asynchronní,
 - synchronní,
 - pseudosynchronní,
 - izochronní.
- Pozn.:
 - následujícím časovým diagramům, signálovým sledům, se obvykle říká **protokol** sběrnice.
- Uvidíte jeden sběrnicový cyklus.

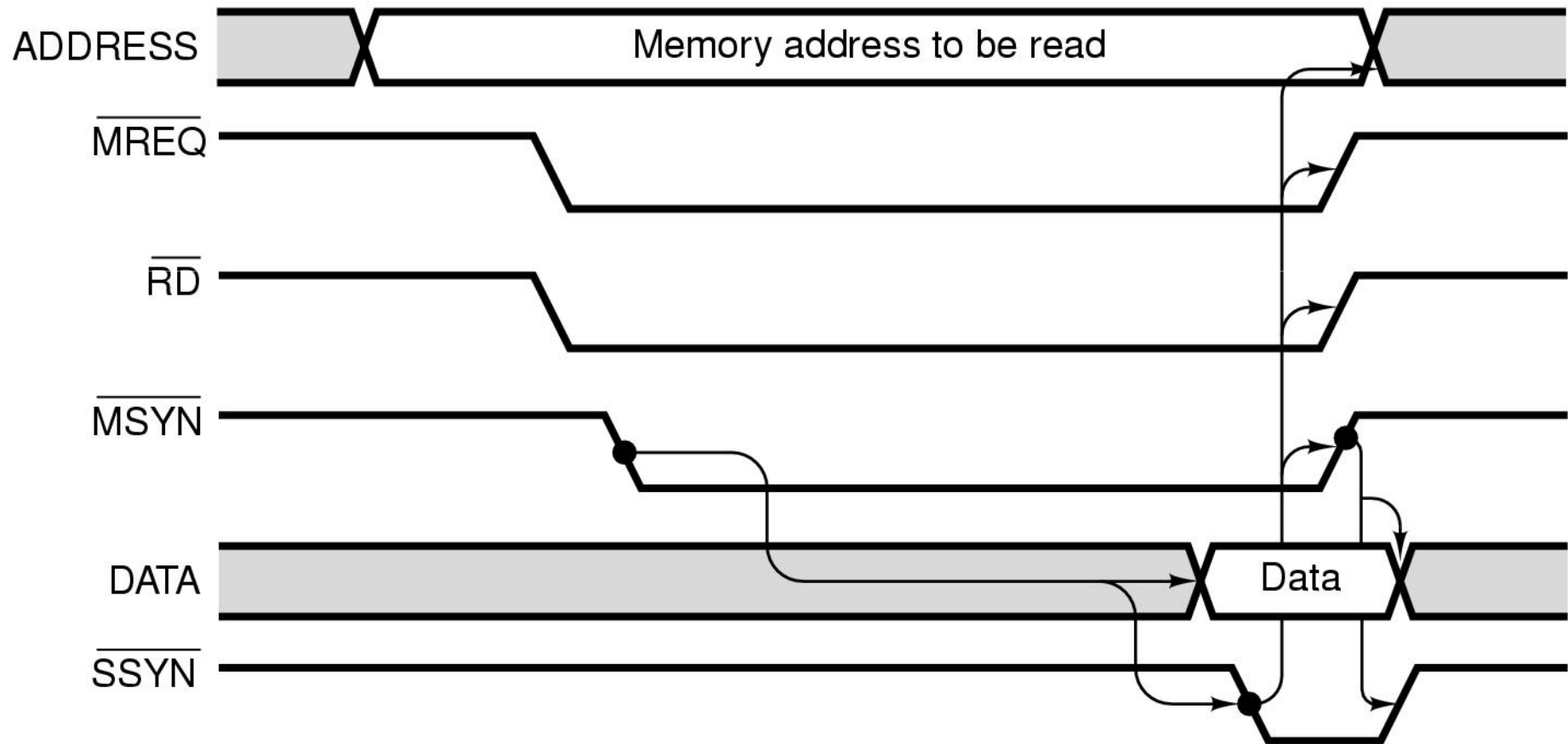
Příklad synchronního přenosu



Symbol	Parameter	Min	Max	Unit
T_{AD}	Address output delay		11	nsec
T_{ML}	Address stable prior to \overline{MREQ}	6		nsec
T_M	\overline{MREQ} delay from falling edge of Φ in T_1		8	nsec
T_{RL}	\overline{RD} delay from falling edge of Φ in T_1		8	nsec
T_{DS}	Data setup time prior to falling edge of Φ	5		nsec
T_{MH}	\overline{MREQ} delay from falling edge of Φ in T_3		8	nsec
T_{RH}	\overline{RD} delay from falling edge of Φ in T_3		8	nsec
T_{DH}	Data hold time from negation of \overline{RD}	0		nsec

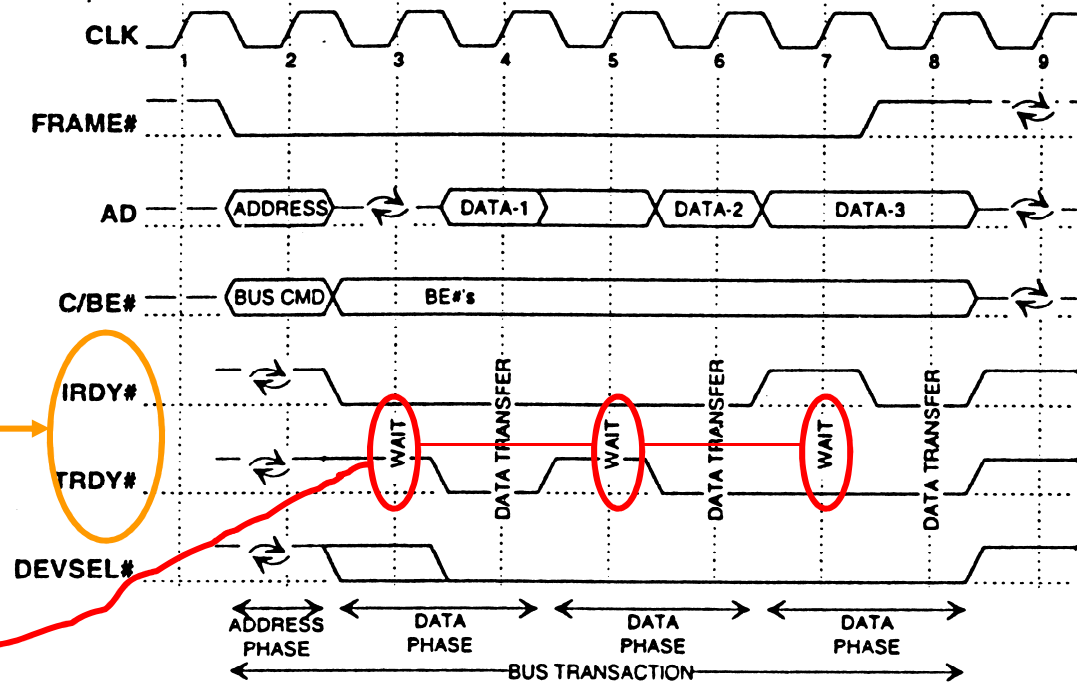
(b)

Příklad asynchronního přenosu



Jde to i jinak?

Stavové signály,
kterými zařízení
(iniciátor nebo
target) žádá o
vlození **WAIT-**
taktů

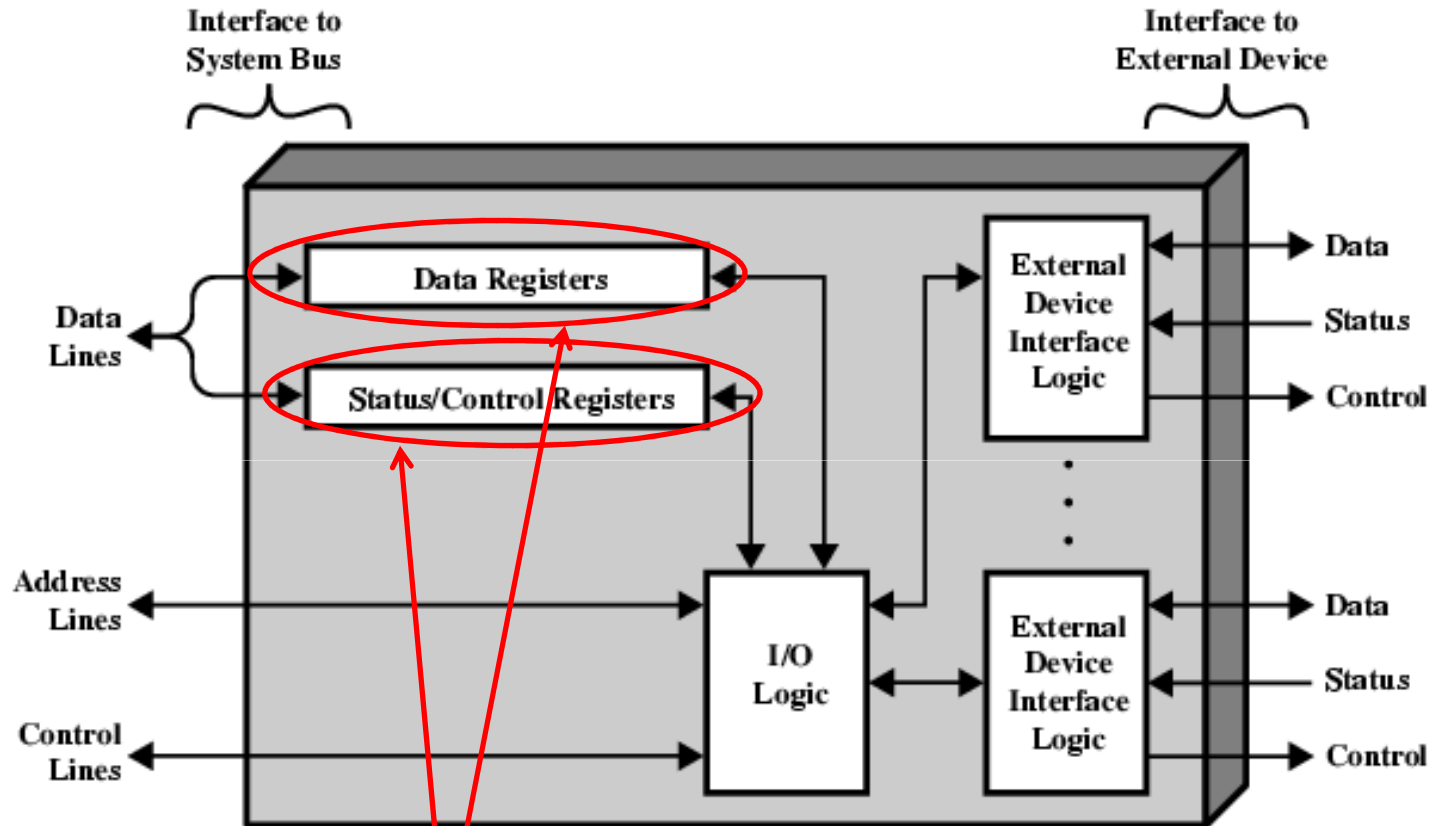


Ano, pseudosynchronně!

A kdy je synchronizace **izochronní**?

- V izochronním přenosu se přenáší údaje s **konstantní průměrnou rychlostí!**
- Jinak: za určitý časový interval se přenesou vždy stejný objem dat, ale okamžitá rychlost přenosu nemusí být stále stejná.
- Typické pro moderní multimediální zařízení.
- Řešení: používá se několik souběžných přenosových kanálů, jedním kanálem se obsluhuje několik periférií, atd.
- Připomenutí:
 - synchronní přenos – přenos konstantní okamžitou rychlostí,
 - asynchronní přenos – zařízení jsou z hlediska časování nezávislá.

Blokové schéma adaptéru



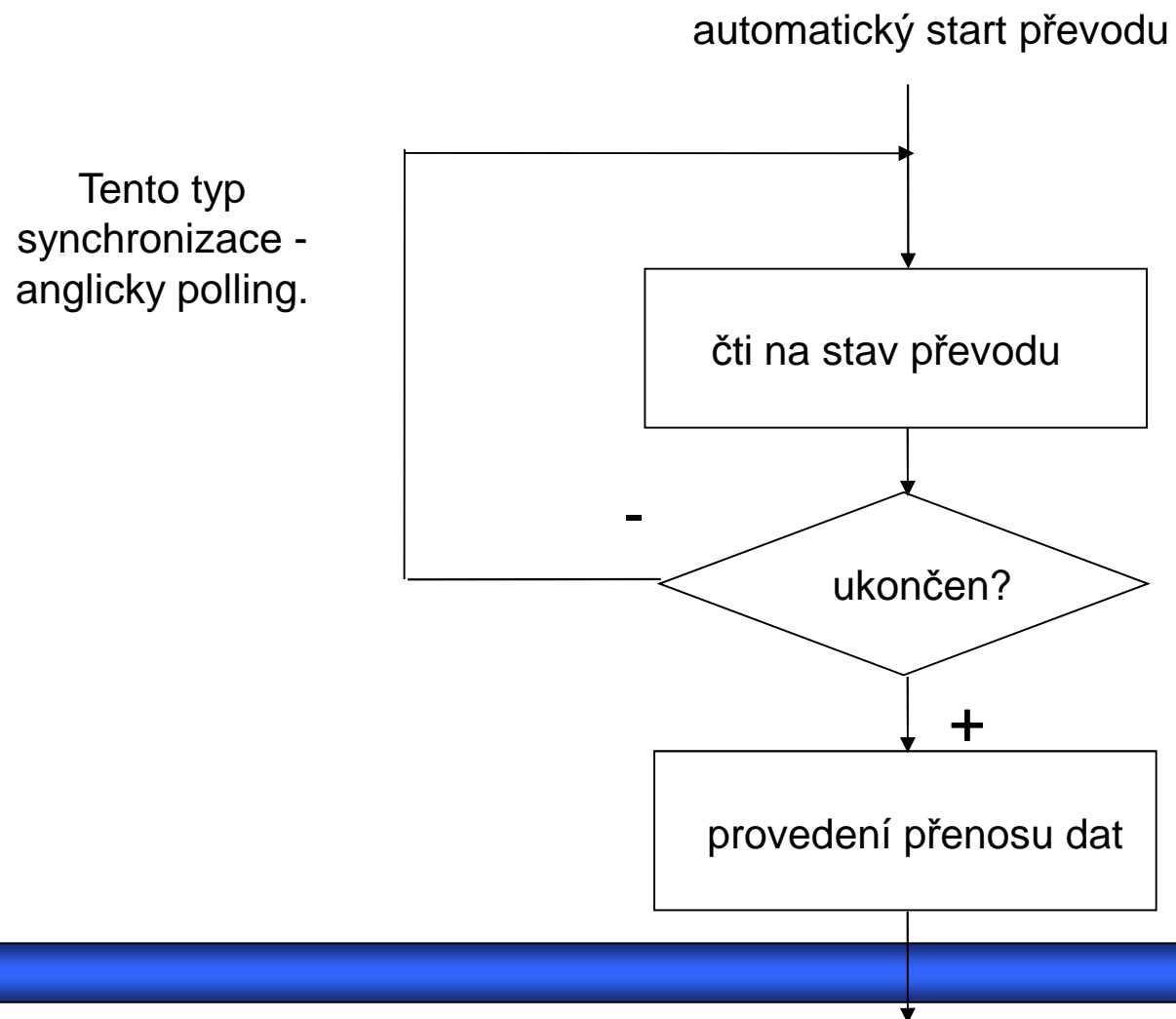
A kde v něm jsou ty brány?

Synchronizace součinnosti procesoru a periferie

Vycházejme z toho, že

- Procesor neumí nic jiného, než
 - pracovat podle předem připraveného programu, nebo
 - reagovat na (vnější) přerušení (startem obslužného programu).

Vývojový diagram synchronizace přenosu



Jsou i jiné možnosti synchronizace?

- Ano! Kromě programovaného V/V dat můžeme ještě použít:
- V/V s přerušením, angl. Interrupt driven I/O,
- V/V pomocí DMA, DMA driven I/O,
- V/V pomocí specializovaného procesoru.

Podrobnější diskuze k DMA

- Záleží na tom, jak je DMA implementován.
- Možnosti jsou:
 - zastavovaným procesorem,
 - kradením cyklů,
 - transparentní DMA.

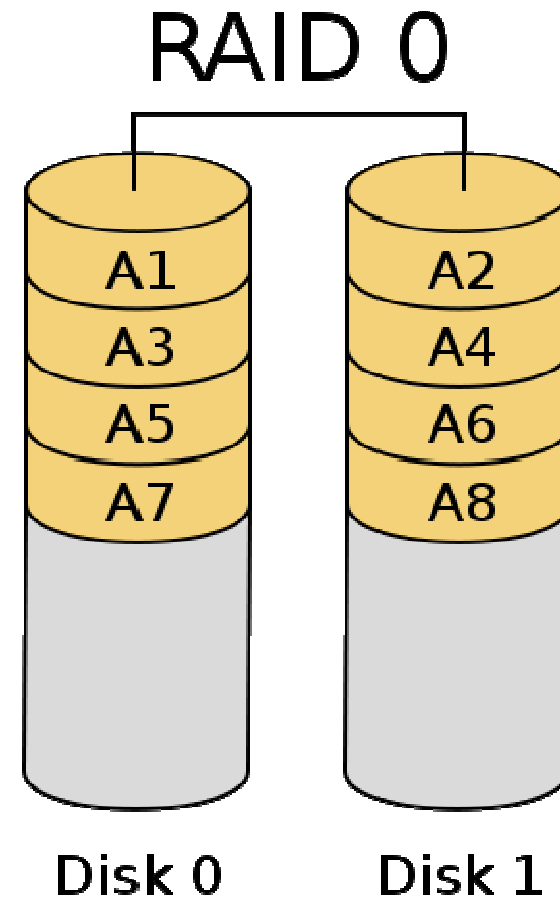
Nejdůležitější periferií je disk. Tak ho musíme zrychlit...

RAID N

- Redundant Array of Inexpensive/Independent Disks.
- Výhoda? Zvýšení chybové odolnosti vnějších pamětí.
- Poznámka: *ne všechny architektury skutečně obsahují redundantní disky. Některé jsou optimalizovány na rychlost.*

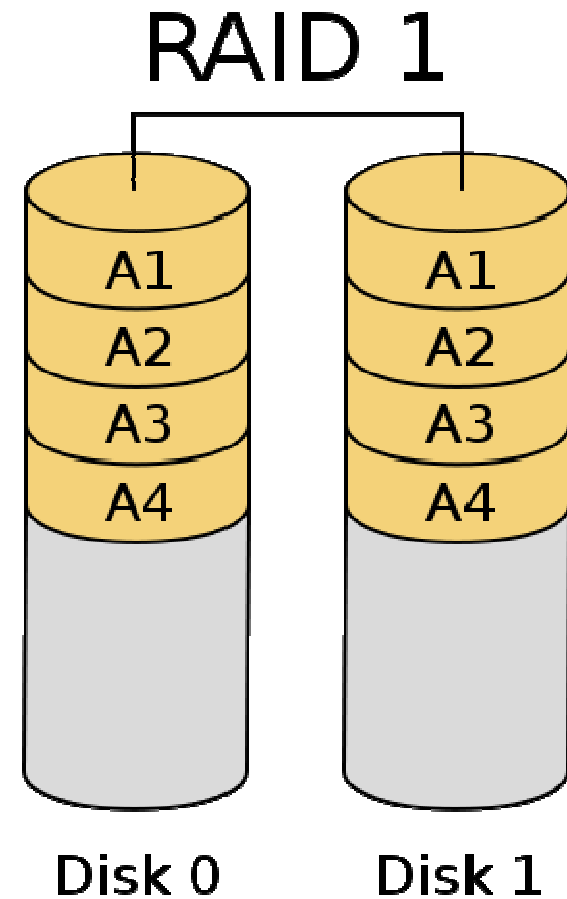
RAID 0

- Pro zvýšení výkonu systému pevných disků.
- tzv. “stripping” (proužkování)



RAID 1

- Pro zvýšení spolehlivosti uložených dat.
- Označuje se jako “Mirroring”.
- Nezrychluje, ale zvyšuje spolehlivost.

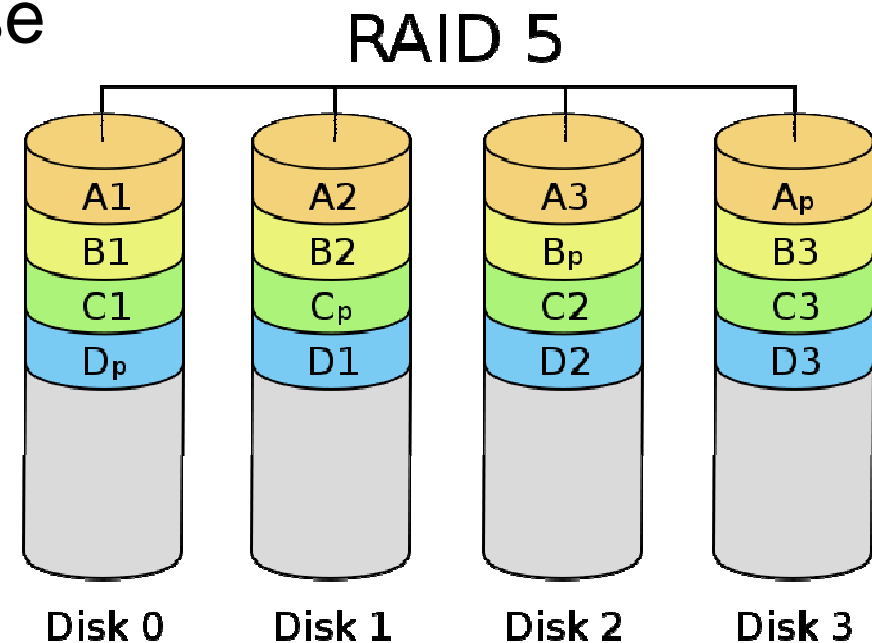


RAID 10

- Kombinace obou výše popsaných.
- Vytvoří se RAID 0 a ten se pak zrcadlí na RAID 1. Výsledkem jsou vlastně dva RAID 0 obsahující identická data.
- RAID 10 zvyšuje jak výkon, tak spolehlivost, musíte ovšem použít nejméně čtyři disky, nejlépe se stejnými parametry.

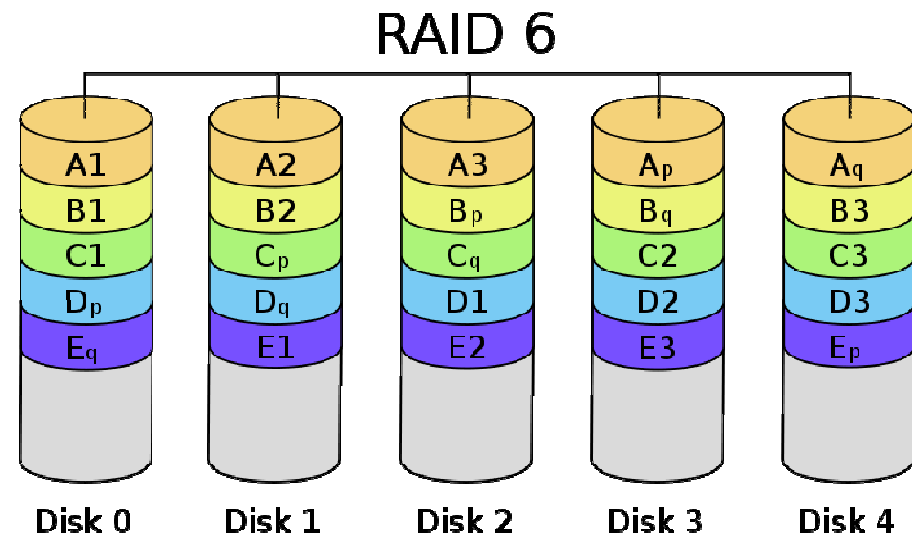
RAID 5

- Ukládá paritní informace, nikoli však na jeden vyhrazený disk.
- V degradovaném režimu se musejí data uložená na vadném disku odvodit z dat zbývajících disků a parity.
- Zrychluje čtení, zpomaluje zápis.



RAID 6

- Obdoba RAID 5, používá dva paritní disky s různě vypočtenou paritou.
- Odolný proti výpadkům 2 disků.
- Rychlost čtení jako RAID 5, zápis ještě pomalejší.



Realizace datové cesty

Vyrovnávací paměti v IO podsystemu

- Optimálním řešením disproporce mezi
 - požadavkem vysokého (ale drahého) HW dimenzovaného na špičkovou propustnost,
 - výhodnou průměrnou propustnost
- jsou vyrovnávací paměti.



Kolísající zátěž.
HW dimenzovaný na
špičkovou zátěž

Stálá zátěž.
HW dimenzovaný na
průměrnou zátěž

Realizace?

- Frontou.
- Kruhovou frontou,
- Paměť,
- Skrytou paměť (cache).
- Bohužel všechny zvyšují latenci.

The I/O Bandwidth Problem

A number of new technologies are responsible for the increasing demand for additional bandwidth.

- High-resolution, texture-mapped 3D graphics and high-definition streaming video are escalating bandwidth needs between CPUs and graphics processors.
- Technologies like high-speed networking (Gigabit Ethernet, InfiniBand, etc.) and wireless communications (Bluetooth) are allowing more devices to exchange growing amounts of data at rapidly increasing speeds.
- Software technologies are evolving, resulting in breakthrough methods of utilizing multiple system processors. As processor speeds rise, so will the need for very fast, high-volume inter-processor data traffic.