

Express jménem PCI

Budoucnost patří nové datové sběrnici

BERNHARD HALUSCHAK

Po PCI, PCI-X a také AGP by se PCI Express nyní měla stát jedinou, zato však univerzální datovou dálnicí pro PC. Popíšeme vám nový standard I/O a srovnáme jej s technologiemi od PCI až po PCI 3.0.

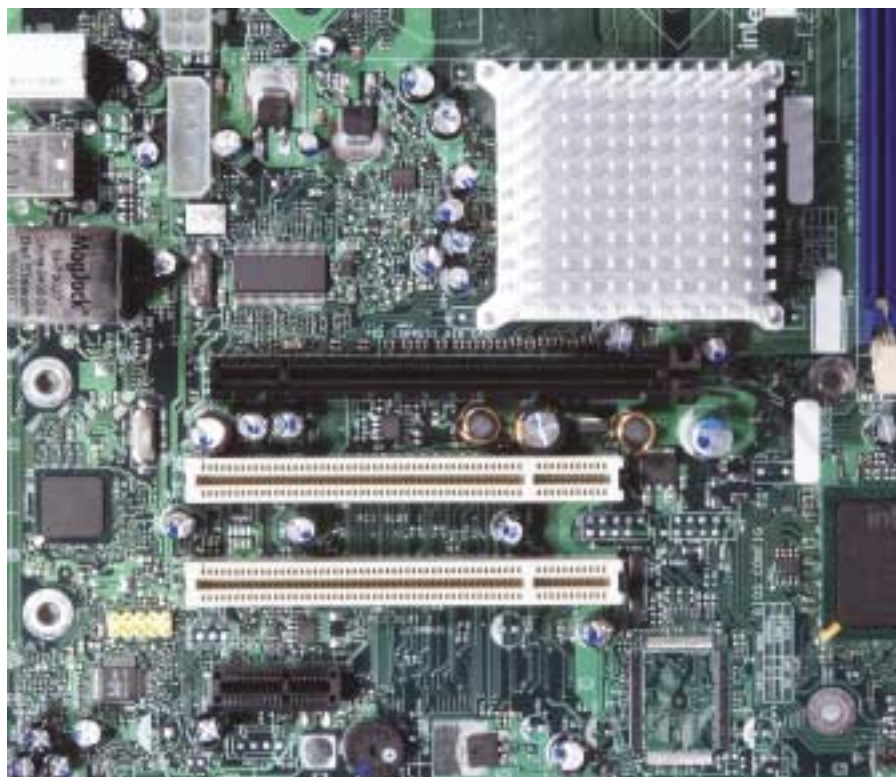
V oblasti serverů a pracovních stanic musí být na pevné disky systémů RAID a síťové karty přenášeno rychle a spolehlivě velké množství dat. Avšak nároky na plynulý datový přenos stouply i v segmentu stolních počítačů. Brzdou je zpravidla příliš pomalý systém sběrnice I/O. Ten musí mezi jednotlivými podsystémy řídit a zpracovávat proudy dat. Pokud přenosová rychlost architektury I/O nestačí, pak na datové dálnici nevyhnutelně dochází k zácpě. Právě toto manko by chtělo konsorcium PCI-SIG (http://www.pcisig.com/home), které tvoří firmy jako Intel, AMD, Microsoft, Dell a HP, dohnat pomocí nových technologií sběrnice.

Sběrnice PCI, představená již v roce 1993, zvládá maximálně 127,2 MB/s a je tím pádem pro moderní hardware I/O (jako 10Gigabitové ethernetové adaptéry) příliš pomalá. Tento deficit měla překonat PCI-X a především PCI Express. Tyto nové topologie sběrnice dosahují přenosové rychlosti až 7,95 GB/s (PCI-X 1066) a 9,53 GB/s v jednom směru (PCI Express s 32 páry vodičů). Systém sběrnice PCI-X se používá převážně v oblasti serverů a měl by být podle PCI-SIGu 2004 nahrazen právě pomocí PCI Express. Současně by PCI Express měla nahradit existující standardy i v ostatních oblastech. To se týká například rozhraní AGP, sběrnice Mini PCI a PC Card (dříve PCMCIA). První provozuschopné komponenty na bázi PCI Express existují již od konce roku 2003. Sériově vyráběné čipsety však známí výrobci jako Intel, VIA, či SiS představili teprve v polovině loňského roku.

Srovnání architektury sběrnice PCI

V následující tabulce naleznete souhrn všech verzí PCI. Zakládají se na paralelním přenosu dat a nejsou škálovatelné. Proti nim stojí sériový proces přenosu u PCI Express.

Podle specifikace má PCI Express šířku sběrnice od 1 do maximálně 32 linek. Při taktovací



▲ Shora dolů : PCI Express x16, dvojice PCI slotů a jeden PCI Express x1.

Přehled verzí PCI

Verze PCI	PCI 2.0	PCI 2.1	PCI 2.2 PCI 3.0	PCI 2.3/	PCI-X 1.0	PCI-X 2.0	PCI-X 3.0
Max. šířka sběrnice (bity)	64	64	64	64	64	64	64
Max. míra taktování (MHz)	33	66	66	66	133	533	1 066
Max. šířka pásma (GB/s)	0,12	0,5	0,5	0,5	0,99	3,97	7,95
Slots pro Bridge	4	2	2	2	1	1	1
Napětí (volty)	5	5/3,3	5/3,3	3,3	3,3	3,3/1,5	3,3/1,5
Zavedení (rok)	1993	1994	1999	2002/2004	1999	2002	2004

frekvenci 2,5 GHz dosahuje systém sběrnice maximální teoretické šířky pásma 19,1 GB/s. Zavedení nového systému sběrnice proběhlo v roce 2004. Ale ani výkony PCI-X ještě nejsou zcela na maximu. Konsorcium PCI-SIG plánuje vývoj PCI-X 2133 v nadcházejících letech.

Všechny standardy jsou stále přepracovávány a dále vyvíjeny. To je úkolem konsorcia PCI-SIG, které také zveřejňuje výsledné specifikace updatů.

Podrobnosti o PCI-Bus I

Současné systémy sběrnice PCI u stolních počítačů vycházejí ze specifikace PCI-Local-Bus 2.3. Sběrnice PCI disponuje kombinovanou datovou/adresovou sběrnici o šířce 32 bitů. Pracuje paralelně s taktovací frekvencí 33 MHz. Z toho

vychází výsledná šířka pásma 127,2 MB/s. Ale i varianta PCI se 64bitovou šířkou sběrnice a taktovací frekvencí 66 MHz používá stejnou specifikaci. Maximální přenosová kapacita pak dosahuje 508 MB/s. Tato konfigurace podporuje maximálně dvě místo čtyř PCI karet 3,3 V. Z toho vyplývá, že zastaralé 5 V karty nejsou s PCI sloty specifikace 2.3 kompatibilní. PCI 2.3 slot však ještě karty určené pro napájení 5 V mechanicky přijmout umí. Verze 3.0 uvedená v dubnu 2004 má jeden slot, do kterého pasují již jen univerzální grafické karty s napájením 3,3 V.

Pro vyřešení problému s počtem slotů využívá PCI takzvané PCI-to-PCI-Bridges. Ty nabízejí možnost provozovat na počítači sběrnice PCI paralelně nebo zapojené za sebou. PCI-to-PCI-Bridge navíc disponuje konfiguračním registrem k ukládání

inicializačních dat. Prefetch- a Posting-Buffer zařízení meziukládání čtených a zapisovaných dat. Navíc Bridge zvládá shrnutí jednotlivých přenosů do rychlých a efektivních dávkových přístupů.

Podrobnosti o PCI-Bus II

Sběrnice PCI pracuje se dvěma druhy sběrnice-ových zařízení – iniciátorem a cílovým zařízením. Iniciátor nastartuje přístup na sběrnici tak, že převezme kontrolu nad řídicím vodičem a pak generuje cílovou adresu. Navíc určí délku a začátek sběrnice-ového přenosu.

Cílové zařízení signalizuje připravenost ke čtení pomocí potvrzujícího signálu, nebo vytvoří Wait-States. Datové přenosy však nemůže inicializovat sám, je odkázán na zařízení Bus-Master nebo na procesor. Komponenty Bus-Master mají přístup ke všem zdrojům systému a mohou spouštět funkce jako Burst-Transfers. Pokud požaduje více Bus-Masterů přístup na sběrnici, pak o tom, který z nich přístup získá, rozhoduje předem stanovená arbitrážní logika.

S funkcemi jako Plug-and-Play přebírá PCI automatickou konfiguraci zařízení PCI. Během bootování systému jsou určeny zdroje jako přerušování, kanály DMA, adresy portů a oblasti paměti.

Podrobnosti o PCI-X-Bus I

Pro pokrytí nároků stoupajících na sběrnici PCI byla v září roku 1999 představena specifikace pro PCI-X 1.0, která byla v květnu roku 2002 nahrazena revizí 2.0. Výkonnostní schopnosti sběrnice PCI-X byly dále zvyšovány. Již od listopadu 2002 se konsorcium PCI-SIG zabývá vývojem standardu PCI-X-3.0. Všechny systémy sběrnice o šířce 64 bitů PCI-X 1.0/2.0/3.0 garantují plnou hardwarovou a softwarovou kompatibilitu se současnými standardy PCI.

Sběrnice PCI-X 1.0 pracuje v normálním módu s taktovací frekvencí 133 MHz a slotem s napájením 3,3 V. V této konfiguraci dosahuje datové propustnosti 0,99 GB/s. Podporuje však i frekvence 100 a 66 MHz se dvěma, potažmo čtyřmi sloty. U sběrnice PCI-X 2.0 se o zdvojnásobení, respektive o zčtyřnásobení šířky pásma při zachování taktovací frekvence 133 MHz starší dva vodiče Address-Strobe AD_STB a AD_STB#. V módu Double-Data-Rate (2 × 133 MHz) dosahuje sběrnice datové průchodnosti 1,98 GB/s, v provozu Quad-Data-Rate (4 × 133 MHz) se dostane až na 3,97 GB/s. Proti PCI-X 1.0 pracuje PCI-X 2.0 vedle 3,3 V i s úrovní napětí 1,5 V. Rozšířený standard PCI-X navíc nabízí rozpoznávání chyb, korekturu datového proudu přes proces ECC a Device-ID-Messaging. Standard PCI-X 3.0, který je stále ještě vyvíjen, by měl disponovat i vylepšeným Power-Managementem.

Dále by měl zvládat isochronní datový přenos, stejně jako redundantní datové cesty při přenosu dat na sběrnice systému. Od označení PCI-X 1066 můžeme odvodit, že datová sběrnice zvládá 1 066 megatransferů za sekundu. Tím dosahuje teoretické šířky pásma 7,95 GB/s. I když PCI-X 1066 ještě uvedena nebyla, připra-

vuje konsorcium PCI-SIG také standard PCI-X 2133. Podrobnosti o tomto standardu zatím k dispozici nejsou. Podle PCI-SIG by měl konkrétní podobu získat až v průběhu několika let. Pokud však bude PCI-X 2133 i přes konkurenci PCI-Expressu realizováno, pak dosáhne výpočetní šířky pásma 15,9 GB/s. Pro srovnání můžeme uvést, že současné technologie jako 10GB ethernet nebo Fibre-Channel pracují s teoretickými hodnotami přenosu 1,16 GB/s. Budoucí přenosové procesy pro 40/100GB ethernet a Fibre Channel by měly dosahovat až desetinásobku hodnoty dnešních datových přenosů.

Podrobnosti o PCI-X Bus II

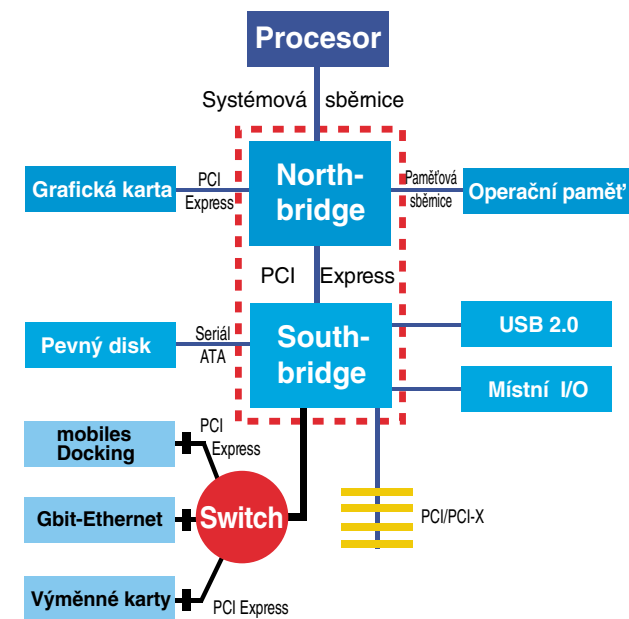
Přenosové procesy u PCI-X 1.0 až 3.0 fungují téměř identicky jako standard PCI, navíc jsou směrem dolů kompatibilní. Kromě toho byly přidány některé nové registry a funkce pro zlepšení kontroly datového toku mezi odesílatelem a příjemcem. Mezi ně patří například Split-Transactions. Ty odesílateli dovolují provádět další přenosy dat na ostatní zařízení, aniž by nejprve čekaly na odpověď od adresáta. Tento Polling, který je okrádá o čas, je jinak u PCI nutný. Již zavedený sy-

stém sběrnice PCI-X 1.0 se používá především v oblasti serverů. Přídavné karty jako 10GB ethernet adapter a SCSI-RAID Controller požadují od systémů sběrnice I/O značnou šířku pásma. PCI-X 2.0 a především nový standard 3.0 by měly výkonostní požadavky karet s vysokými nároky (jako 40Gigabitový ethernet adapter) pokrýt. PCI-SIG navíc již vypracovává kritéria pro PCI-X 2133.

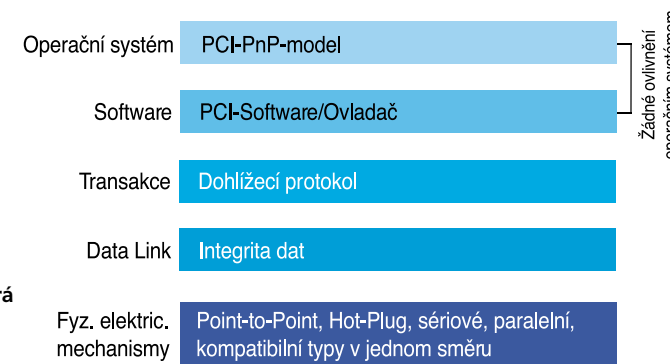
Architektura PCI Express

Flexibilní architektura PCI Express – známá též pod označením 3GIO a Arapahoe – slibuje nahrazení dnešních systémů sběrnice PCI/PCI-X a portu AGP. PCI Express se hodí pro univerzální rychlé vzájemné propojení interní komponent jako USB řadičů, karet s vysokým průtokem dat jako SCSI-RAID, nebo 10GB ethernetové adaptéry. Může být využit i při propojení dokovacích stanic a k nim připojeným externím přístrojů, například přenosných počítačů. Nový standard I/O je schopný nahradit obvyklá spojení Chip-to-Chip jako V-Link od VIA, MuTIOL od SiS či Hub Link od Intelu. Rychlost zde může být proti zatím nejrychlejšímu procesu MuTIOL (0,99 GB/s) zvýšena až devětkrát.

► **Hudba budoucnosti: takhle vypadá topologie PCI Express u stolního počítače. Sběrnice PCI Express řídí datový provoz k zařízením I/O.**



© tecChannel.de



© tecChannel.de

Datové přenosy PCI Express orientované na pakety se však kvůli layer-modelu nehodí pro vzájemné propojení procesoru nebo paměti (na rozdíl třeba od HyperTransportu). Kódování 8B/10B a s ním spojené dlouhé latenční doby, stejně jako komplikovaný Routing 32 linek nepřinášejí žádné výhody. Přesvědčivá topologie sběrnic však není bez přídavných modulů I/O-Bridge možná.

PCI-Express Bridges

Dlouhodobým cílem PCI Express je nahradit rozdílné standardy I/O jako PCI, PCI-X či AGP. V přechodném období však musí být zaručeno, že nová technologie bude z důvodu nákladů nadále

podporovat běžné karty. Zároveň musí PCI Express bezproblémově spolupracovat s aktuálními rozhraními jako USB 2.0, SCSI, IEEE1394b, Ethernet a Infiniband. Řešení této „kompatibilizační šňůry“ nabízejí moduly Bridges.

Pomocí modulů Bridge je možné všechny komponenty nekonformní s PCI Express zapojit do architektury PCI Express. Ztráty výkonu by se přitom objevovat neměly. Například výrobce grafických čipů NVIDIA používá modul AGP-to-PCI Express, aby do vyvinutí „domácího“ řešení mohl současně grafické čipy nasadit do architektury PCI Express. Tento modul Bridge se dá provozovat i v opačném směru. Podobný koncept sleduje výrobce grafických čipů ATI. Spo-

lečnost používá vlastní grafické čipy PCI Express a kombinuje je s PCI-Express-to-AGP Bridge, aby byly kompatibilní se stále aktuálním rozhraním AGP.

I v oblasti serverů, kde jsou například komponenty PCI-X velmi rozšířené, by v budoucnu měly moduly Bridge usnadnit zapojení této technologie do architektury PCI Express. Nasazením technologie Bridge lze v tomto segmentu ušetřit náklady na obnovenou validaci karet, které jsou přestavěny na PCI Express.

PCI Express, architektura Switch

Vedle modulů Bridge má technologie PCI Express k dispozici i nový funkční modul Switch. Ten spravuje různá koncová zařízení a disponuje inteligentní kontrolou toku. Ta podporuje isochronní datové přenosy v rámci Switchu z jednoho koncového zařízení na druhé pomocí takzvaných virtuálních tunelů, tedy bez oklady přes host. Isochronní datový tok je možný i směrem z hostu ke koncovému zařízení.

Oba způsoby přenosu urychlují datový provoz a zabraňují během přenosu dat problémům s „úzkými“ místy. PCI a PCI-X (kromě PCI-X 3.0) takové možnosti nenabízejí. Aby byla zaručena softwarová kompatibilita se standardy PCI-2.3/3.0, má každý vstup a výstup Switchu interní virtuální PCI-to-PCI Bridge. U datových přístupů PCI Express by však bez ohledu na kompatibilitu s PCI mohly v rámci nových řešení odpadnout.

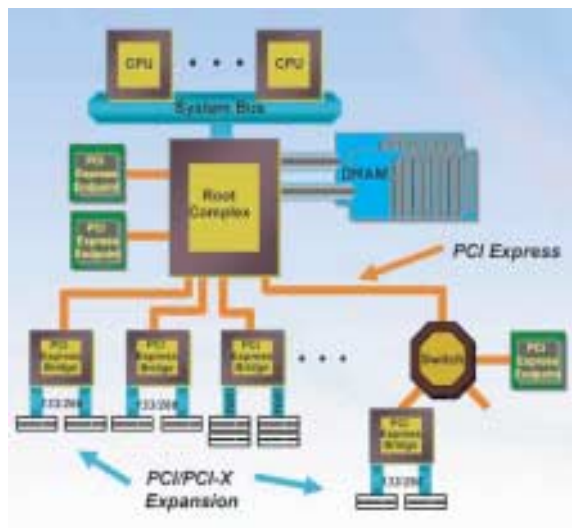
PCI Express Switch může například rozdělit datový přenos Host Bridge na více zařízení. Ty pak již „nevisí“, podobně jako u běžné architektury sběrnic PCI, v řetězci za sebou, nýbrž vedle sebe. Každé zařízení je pak pro Host dosažitelné přímo, s nižší latenční dobou. Struktura Switchu navíc zvyšuje jistotu při výpadku, neboť defekt jedné komponenty již nechromuje celý řetězec.

PCI Express Advanced Switching

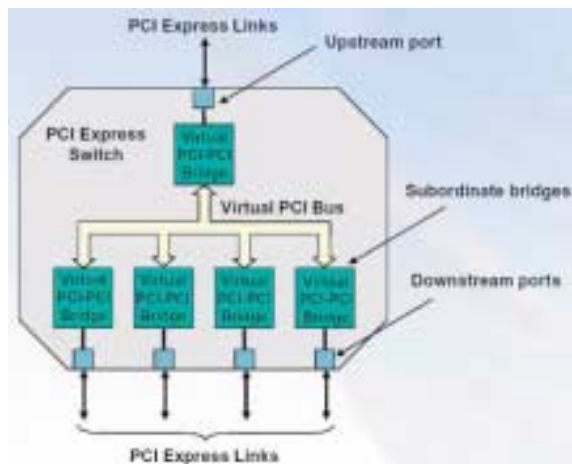
Pod vedením Intelu doplnila Advanced Switching Interconnect Special Interest Group (<http://www.asi-sig.org/home>, ASI SIG) v prvním návrhu 1.0 technologii PCI Express o standard Advanced Switching, zkráceně PCI AS. V lednu 2004 se konalo představení tohoto standardu. Ten by měl učinit z PCI Express spojovací techniku, využitelnou v oblasti sítí a telekomunikací. Navíc je použití PCI AS v serverech plánováno jako standard Backplane.

PCI AS by mělo sloužit jako spojovací technika například v síťových procesorech routerů a urychlovat a zjednodušovat výměnu dat v oblasti serverů, založených na telekomunikacích. S technologií Advanced Switching je možná i realizace takzvaných Fabrics. V této topologii zapojení je možné převádět data mezi libovolnými body v závislosti na adrese a odesílateli. Pro technologii PCI AS zůstává hardware architektury PCI neměnný, pouze software musí odpovídat zvláštním nárokům výměny dat v síti.

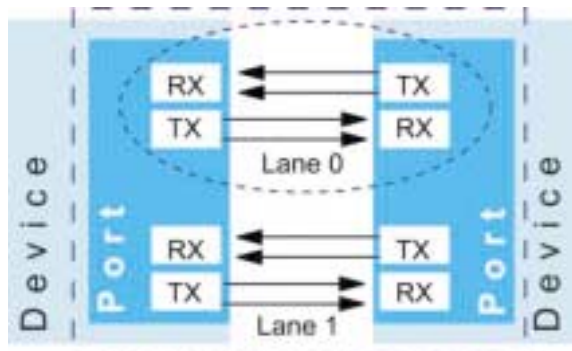
◀ **PCI Express Bridges: pomocí modulů Bridges lze běžné standardy I/O jako PCI nebo PCI-X bezproblémově propojit s architekturou PCI Express (zdroj: Intel).**



◀ **PCI Express Switch: umožňuje isochronní datové přenosy přes virtuální tunely. Aby byla zaručena kompatibilita s PCI, má každý vstup a výstup PCI-to-PCI Bridge (zdroj: Intel).**



◀ **PCI Express Link: vždy dva diferencní vodiče k příjemci a odesílateli tvoří jednu Lane.**



Slot a Bus I

PCI Express Slot by měl být finančně výhodný. Proto se u PCI Express Slotu ve srovnání se standardním PCI slotem snížením počtu pinů zmenšují i rozměry. PCI Express vychází ze sériového spojení Point-to-Point. Nejjednodušší spojení mezi příjemcem a odesílatelem se skládá ze dvou jednosměrných diferencně zapojených párů vodičů s nižším a usměrněnějším napětím signálu, zvanými též Lane. To zabraňuje křížovému modulaci se sousedním signálním vedením.

Při dvou datových vodičích (odpovídá x1 Link se dvěma diferencními páry vodičů) má včetně vodičů pro data, adresy, řízení a zásobování proudem minimálně 36 pinů. U sběrnic Link o šířce x4 to je 64, u x8 98 a u x16 maximálně 164 pinů. Naproti tomu 32bitový PCI slot potřebuje 84 jednotlivých připojení. U AGP 8x to je 108 a u slotu PCI-X 150 pinů.

Podle počtu potřebných Lanes může být PCI Express slot bez mechanických omezení vsazen před nebo vedle existujících PCI slotů. I systémová řešení pro PC karty, Cartridge nebo kabelová spojení k ostatním počítačům, včetně podpory Hot Plugu, jsou díky sběrnicí PCI Express možná.

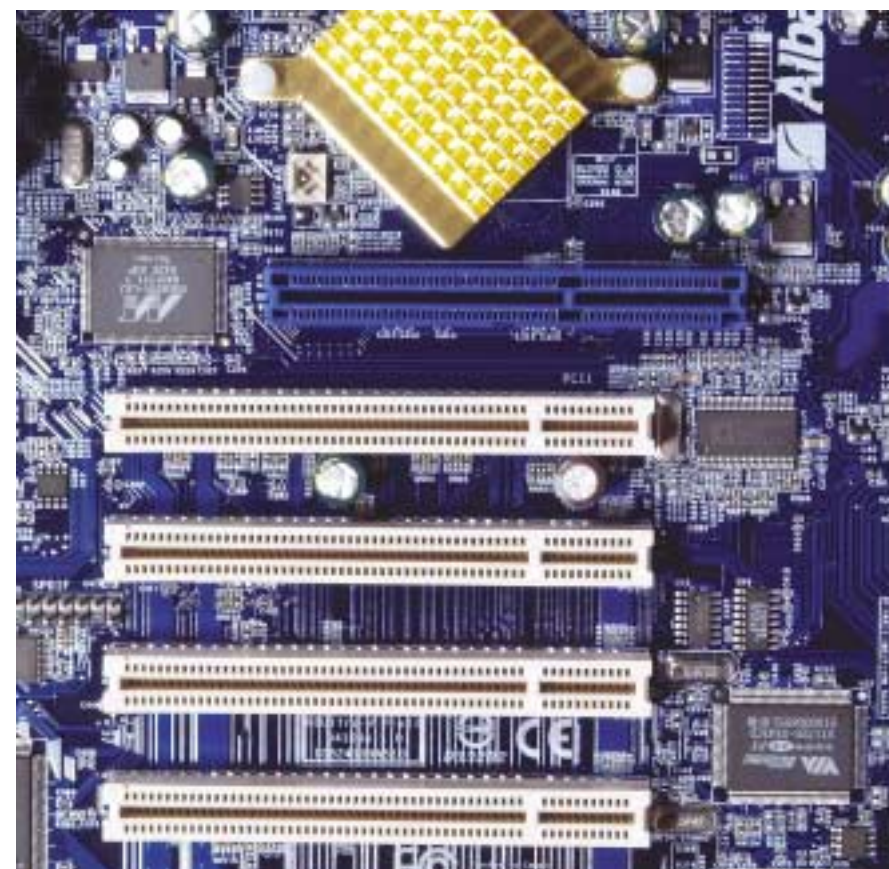
Slot a Bus II

Sběrnice PCI Express se dá výhodně implementovat na čtyřvrstvé základní desky. Pokud chceme zabránit rozdílným dobám chodu signálů linek na základní desce, pak musí být měděné dráhy vedení stejně dlouhé. Specifikace předepisuje maximální délku vedení u 16linkového spojení 50,8 cm. Navíc musí být vodiče rozvržené na impedanci 100 ohmů. Maximální napětí Peak-to-Peak mezi dvěma diferencními vodiči signálu nesmí překročit 1,2 V. Zároveň nesmí překročit 0,8 V na vstupu relé a 0,175 V na výstupu přijímače. Exaktní údaje definovaných úrovní napětí jsou u diferencně sériových přenosů dat obzvláště důležité. Přijímač totiž nezískává informace z diferencí amplitud, nýbrž ze změn toku v datovém proudu.

PCI Express Slot volitelně nabízí i integrovaný SMBus. Sériový interface se skládá ze dvou vodičů, přes které jsou vyměňovány informace mezi komponenty systému se stejným rozhraním. SMBus je schopna řídit nezávisle na dalších kontrolních vodičích sběrnicí PCI Express úkoly systému a power-managementu. Navíc předává i údaje výrobce jako jména modelů a sériová čísla připojených komponent, okolnosti události Suspend a protokoly o statutu při chybách zařízení.

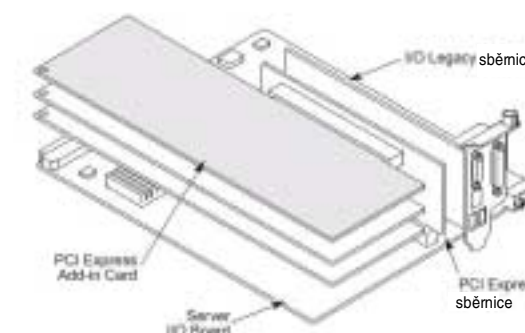
Kompatibilita

Důležitou vlastností standardu PCI Express je softwarová kompatibilita s běžným standardem PCI. Ani systém, ani software aplikací nemusí být na systém sběrnic nastaven. Zvláštní nároky na kompatibilitu mají obě kritické fáze: inicializace a provoz. Během inicializační fáze systém zprostředkovává hardwarovou konfiguraci a rezervuje odpovídající zdroje jako paměť,

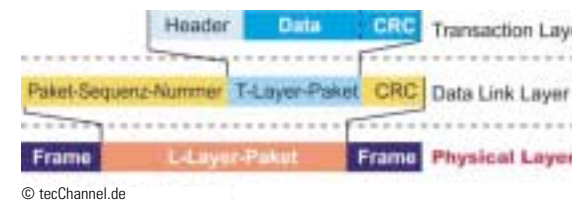


▲ Pozice AGP (na obrázku modrou barvou) je postupně nahrazována modernějším řešením PCI Express x16

▶ **Server Express: s odpovídající kartou Riser je možné osadit PCI Express i na desky serverů, kde je málo místa.**



▶ **Základ: Physical Layer tvoří nejnižší úroveň modelu vrstev. Je přímým pojítkem ke slotu PCI Express.**



oblast I/O a interrupty. Koncept PCI Express přijímá beze změn dosavadní rutinu konfigurace PCI a programovatelnost připojených zařízení I/O. Tím je zaručeno, že všechny systémy budou s PCI Express spolupracovat bez nutnosti speciálních úprav.

PCI Express rozšiřuje inicializační a konfigurační data, takzvaný Configuration Space, na 4 096 bytů na jeden funkční blok. Naproti tomu stojí 256 bytů, které jsou dány specifikací PCI

v revizi 2.3. Configuration Space obsahuje informace o tom, která zařízení jsou připojena na sběrnicí a jak jsou řízena. Tato data jsou důležitá pro proces bootování. PCI Express proto rozděluje Configuration Space na oblast kompatibilní s PCI 2.3 s maximem 256 bytů a na rozšířenou oblast PCI Express s 4 096 minus 256 bytů. Přístup do oblasti může probíhat jak podle pravidel PCI 2.3, tak i podle PCI Express. Přístup do rozšířeného prostoru PCI Express přes PCI 2.3

možný není. Aby se zabránilo chybám, musí být konfigurační nároky kompatibilní s PCI v rozšířené oblasti PCI Express vyplněny nulami. Pro implementaci PCI Express do systému PC nejsou potřebné nějaké nové nebo upravené ovladače. Všechny modifikované procesy sběrnice PCI Express ovlivňují výlučně Physical Layer, Data Link Layer a Transaction Layer. Způsob fungování a stavba jednotlivých vrstev PCI Express podrobně vysvětlují, jak bude budoucí standard I/O fungovat.

Physical Layer

Physical Layer představuje nejnižší úroveň modelu vrstev a reguluje datový tok přes jednotlivé sériové vodiče sběrnice.

Před každým přenosem kóduje data odesílatel a příjemce podle procesu 8B/10B. Tato sériová kódovací technika převádí 8 bitů dlouhé informace do 10 bitů dlouhých datových bloků. V tomto datovém toku jsou pomocí za sebou následujících bitových změn obsaženy taktovací informace, s nimiž se synchronizují přenosové

stanice. Další signály Sideband nejsou pro kontrolu datového toku nutné. Nevýhodou kódování 8B/10B je o dvacet procent nižší efektivita Dateratu. Sběrnice PCI Express dosahuje při základní frekvenci 2,5 GHz maximální míry přenosu dat 2,5 GB/s na jeden směr a pár vodičů, což odpovídá 2,0 GB/s uživatelských dat. Plánováno je zvýšení frekvence až na 10 GHz, čímž bychom dosáhli maximální fyzické rychlosti přenosu v měděných vodičích 10 GB/s.

Sběrnice PCI Express dovoluje pomocí vložené linky flexibilní lineární indexaci v potřebné šířce pásma. Physical Layer architektury PCI Express podporuje linky o šířce $\times 1$, $\times 2$, $\times 4$, $\times 8$, $\times 16$ a $\times 32$. Data určená k přenosu jsou rozdělena na jednotlivé linky a zakódována procesem 8B/10B.

Sběrnice široká 32 linek pak dosáhne maximální rychlosti přenosu 9,31 GB/s v jednom směru. Během inicializační fáze synchronizují odesílatel a příjemce šířku linky a přenosovou frekvenci. Tento proces neovlivňuje systémové nebo softwarové vrstvy, a tak garantuje plnou kompatibilitu s běžným standardem PCI. Update specifikace ve formě nových technologií encodingu nebo řízení rychlosti přenosu jsou pomocí struktury vrstev snadno proveditelné. Nová zadání a nároky mění výlučně funkční parametry ve Physical Layeru.

Data Link Layer

Hlavní úkol vrstvy Data Link Layer spočívá v tom, připravit pomocí Physical Layeru datové pakety Transaction Layeru (TLPs) a obráceně. Při tom dává k dispozici servisní funkce jako rozpoznávání chyb a s tím spojené opakování přenosu dat. Pakety Data Link Layeru navíc přebírají určité funkce Link Managementu. Jsou to potvrzení o přijetí paketů Transaction Layeru, inicializační status a Power status, stejně jako předávání informací o kontrole datového toku. Bloky Data Link Layeru jsou mezi dvěma komponentami přenášeny přes Point-to-Point Transfer. Naproti tomu se TLPs dají rozvést přes více napojených zařízení.

Aby byla zaručena integrita dat TLPs, připojí Link Layer na začátek bloku datového paketu jednoznačné identifikační číslo, takzvané Paket-Sequencenumber. Tím je přenosový proces v rámci oblasti PCI Express jednoznačně identifikován. Link Layer navíc na konec bloku zařadí kontrolní součet pro rozpoznávání chyb (CRC). Pokud příjemce signalizuje chybné doručení dat, potom Data Link Layer inicializuje obnovený přenos bloku. Když Data Link Layer tento datový blok vyšle, pak dorazí k příjemci s jistým zpožděním. Tato latentní doba závisí na faktorech jako frekvence a šířka pásma sběrnice, zpoždění signálu kvůli drahým vodičům a opakování přenosu při rozpoznání chyby.

Struktura Data Link Layeru

Pakety Data Link Layeru (DLLPs) se skládají z pěti bytů počínaje bytem nula. Osm bitů bytu nula specifikuje typ DLLPs. Bity nula až pět bytu jedna a bity šest a sedm bytu dvě definují například

Header-String pro inicializaci kontroly toku. K tomu patří i bity nula až sedm bytu tři a bity nula až tři bytu dvě, které tvoří Date-String funkce. Poslední String může při potvrzení příjmu obsahovat i sekvenční číslo chybně přenesených TLPs. Pokud Data Layer přenáší funkce Powermanagementu, pak zůstanou byty neobsazené. Zakončení DLLPs tvoří byte čtyři s kontrolním součtem CRC.

Pokud musí Data Link Layer přenášet blok Transaction Layeru, pak na začátek TLPs přidá sekvenční číslo, které se skládá ze dvou bytů. První čtyři bity Date-Stringu zůstanou neobsazené a následujících dvanáct bitů tvoří sekvenční číslo. Na konec datového paketu pak Data Link Layer navěsí čtyři byty dlouhý kontrolní součet k rozpoznávání chyb (LCRC).

Transaction Layer

Transaction Layer zpracovává nároky na zapisování a čtení, stejně jako na speciální události. Je pojítkem mezi vrstvou softwaru a Data Link Layerem. Aby se zabránilo nepotřebným datovým přenosům, podporuje Transaction Layer protokol podle takzvané kontroly Credit-Based-Flow. Tento proces předává datové pakety pouze tehdy, pokud je v cíli volná vyrovnávací paměť. To šetří opakovaně předávání dat a tím pádem i přebytečné zatěžování šířky pásma sběrnice. Pokud odesílatel potřebuje potvrzení o přijetí, pak přenos probíhá v modu Split-Transfer. Ten dovoluje odesílateli až do odpovědi adresáta provádět další přenosy na ostatní příjemce.

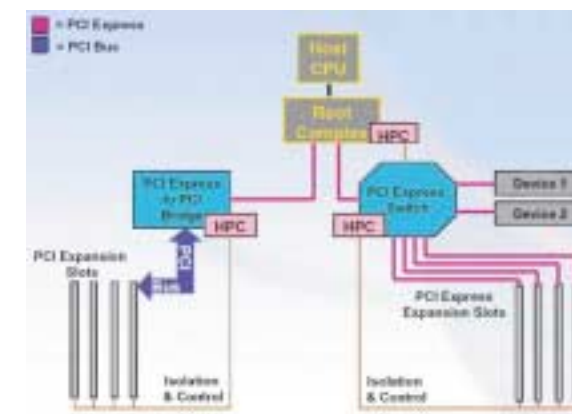
Transaction Layer navíc stanovuje pravidla pro přenos a management TLPs, jako třeba pro kompatibilitu PCI/PCI X. Rozdělení datových paketů podle určitých skupin Traffic Class dovoluje inteligentní kontrolu toku. Ve switchích jsou tyto jednotky před datovým přenosem podle použití rozděleny nebo shrnuty do virtuálních kanálů. Tento proces složí k efektivnímu využití existující šířky pásma sběrnice, protože jsou přetvářeny pouze ty datové pakety, které patří k sobě. Specifikace PCI Express pro výměnu dat pomocí Transaction Layeru definuje čtyři druhy adresování. Jsou to Memory, I/O, Configuration a Message, s různými mody přenosu jako Reads a Writes.

Struktura Transaction Layeru I

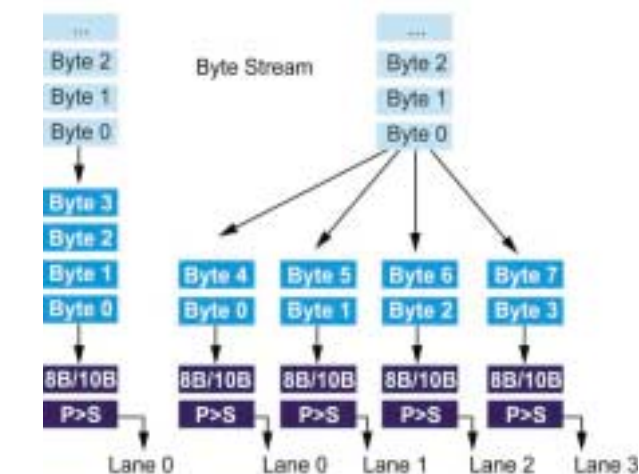
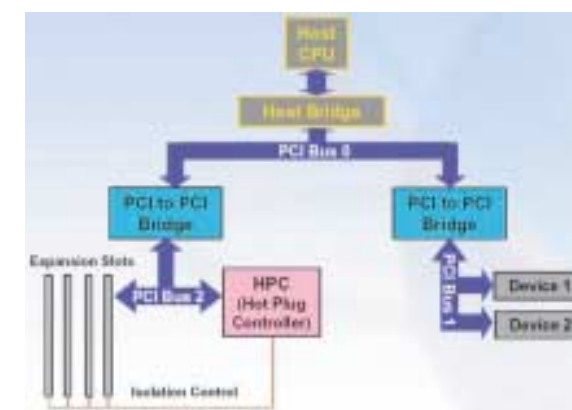
Transaction Layer se skládá z Headeru, datové oblasti a případně i přílohy ke zjišťování chyb. Obecný význam prvních čtyř bytů (dvojslovo) Transaction Headeru je u všech paketů Transaction Layeru stejný. Začíná bitem nejvyššího řádu bytu nula. Bity pět a šest bytu nula udávají délku Headeru (Fmt). Ten může být dlouhý tři až čtyři dvojslova a případně může obsahovat i datovou přílohu. Bity nula až čtyři bytu jedna (Type) a dva bity délky Headeru definují mód přenosu, jako například 32/64-Bit-Memory či I/O-Reads/Writes. Ty také stanovují další skladbu Transaction-Headeru.

Bity Traffic-Classu (TC) čtyři až šest bytu jedna kódují osm různých obslužných tříd, které pře-

► **PCI Express Native Hot Plug: funkce definovaná ve standardu PCI Express pracuje nezávisle na předlohách výrobce (zdroj: Intel).**

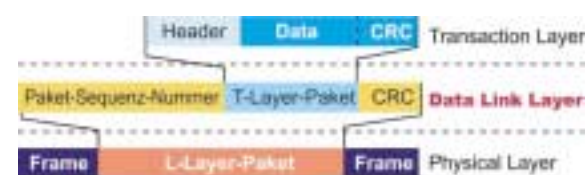


► **PCI Hot Plug: běžný PCI Hot Plug potřebuje nastavení od výrobce pro BIOS a pro ovladače systému a zařízení (zdroj: Intel).**



◀ **Rozdělení a kódování: pokud nějaký port pošle data přes několik linek, pak Physical Layer tyto byty v určitém pořadí rozdělí na jednotlivé linky a zakóduje je.**

© tecChannel.de



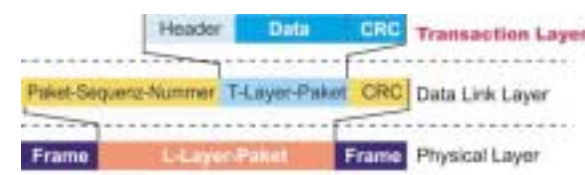
◀ **Prostřední vrstva: Data Link Layer usazený mezi Physical Layer a Transaction Layer vykonává služby jako předávání servisních funkcí a hlášení o stavu.**

© tecChannel.de



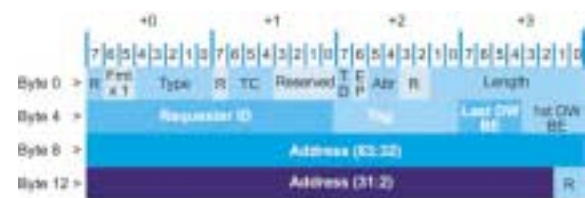
◀ **Odolný proti chybám: zakončení paketu Data Link se skládá z 16bitového kontrolního součtu CRC.**

© tecChannel.de



◀ **Transaction Layer: skládá se z oblasti Headeru a dat. Podle potřeby může následovat 32bitový kontrolní součet.**

© tecChannel.de



◀ **Vše v jednom: v Headeru paketu Transaction Layeru jsou všechna data pro příjemce relevantní. Tak může být obsah bloku správně dekódován.**

© tecChannel.de

nos dat ovlivňují. Pouze systémový software v prostředí PCI Express může definovat obslužné třídy, jako například přenos dat podle priority. Defaultní nastavení (sled bitů: 000) podporuje veškeré komponenty PCI Express. Sedmý bit bytu dvě v Transaction-Headeru (TD) signalizuje, že za paketem Transaction Layeru následuje zvláštní příloha, které se skládá z jednoho dvojslova. To obsahuje speciální kontrolní součet k rozpoznávání chyb End-to-End-CRC (ECRC) a má tak zaručovat vyšší integritu dat. Usazením šestého bitu bytu dvě (EP) příjemce rozpozná chybný TLP. Zpráva o statutu odesílatele navíc informuje o vzniklé chybě a přiměje ho k opakování datového přenosu. Rozpoznávání chyb je prováděno pouze u zapisovacích a čtecích operací.

Struktura Transaction Layeru II

Ordering-Attribute-Field (Attr) TLP se skládá z obou bitů čtyři a pět bytu dvě. Pomocí čtvrtého bitu je možné adresovat pole Cache-Coherency-Managementu. Pokud je bit deaktivován, přenosová operace nutí hardware provádět koherentní přístupy na cache. Jeden bit signalizuje upuštění od této formy přístupu (No Snoop). Pátý bit rozhoduje, zda budou datové přístupy probíhat přísně podle pravidel PCI nebo podle určitých předpokladů specifikace PCI X. To garantuje zpětnou kompatibilitu se standardy PCI/PCI X.

Délku uživatelské datové oblasti (Length) reprezentují byty tři a dva, a třetího jsou vyhodnocovány pouze první dva bity. Celkový počet de-

seti bitů představuje maximální datovou množinu 1 024 dvojslov, což v přepočtu znamená 4 KB uživatelských dat.

Header obsahuje v bitových polích Fmt, Type a TD všechny nutné informace, aby mohl určit velikost celého Transaction Layer paketu.

Powermanagement

PCI Express Powermanagement podporuje všechny funkce, které jsou dány v Bus Powermanagement Interface Specification Revision 1.1 a v Advanced Configuration and Power Interface Specification Revision 2.0. Na rozdíl od PCI s Device-Powermanagement-States D0, D1, D2, D3hot a D3cold definuje PCI Express takzvané Link-Powermanagement-States. Jednotlivé Link-States L0, L0s, L1, L2 a L3 jsou odvozeny z Device-States a mají podobné funkce. V módu L0/D0 zásobuje napáječ proudem veškerý hardware. Stav L0s/D0/D1 je stejný jako mód Standby. L1/D1/D2 rozšiřuje provoz Standby tak, že odpojí speciální funkční jednotky karet. Ve stavu L2/D3hot je aktivováno pouze pomocné napětí napáječe. Zde se pak systém nachází v takzvaném Low-Power-Sleep-State. Naproti tomu mód L3/D3cold všechna napětí vypne.

Schopnost Hot Plug

Funkce Hot Plug umožňuje zapojovat nebo odebírat hardwarové komponenty počítače v běžícím stavu. Přerušení provozu přes reset nebo Power-Down není nutné. První oficiální specifikace Hot-Plugu pro PCI byla uvedena v roce 1997. V roce 2001 následovala přepracovaná



▲ Grafická karta PCI Express: výrobci grafických čipů již pracují na prvních prototypch grafických karet s rozhraním PCI Express (zdroj: Intel).

verze se zlepšenou funkcionalitou Standard-Hot-Plug-Controlleru (SHPC). Vývoj specifikace pro PCI Express Native Hot Plug pro využití u sběrnic PCI Express začal již v roce 2002.

K výhodám PCI Express Native Hot Plugu patří i plná podpora Legacy PCI Hot Plugu, která je založena na modelu SHPC. Navíc využívá celou řadu standardních programů, které jsou již implementovány v současných operačních systémech jako Windows 2003 Server a Windows XP. Native Hot Plug přitom upouští od verzí BIOSu, ovladačů a modifikovaných řadičů zařízení, které jsou nastaveny výrobcem. Každý komponent PCI Expressu, jako Root Complex, Bridge nebo Switch, disponuje vlastním Hot Plug Controllerem, který pracuje nezávisle na

nastavení od výrobce, tedy podle specifikací PCI Express. Dalším plusem PCI Express Hot Plugu jsou takzvané funkce Power Budgeting. Pokud je příkon PCI Express Slotu omezen operačním systémem na určitou hodnotu nebo pokud se jednomu slotu v systému nedostává dostatečného elektrického výkonu, pak PCI Express Native Hot Plug tento stav rozpozná. Doručí tyto informace administrátorovi systému, který může přijmout odpovídající opatření.

Naproti tomu stojí PCI Hot Plug se svojí nestandardizovanou a komplexní funkcionalitou. Centrální ovládání PCI slotů přebírá externí Hot Plug Controller. Společně se speciálně nastaveným ACPI-BIOsem výrobce jsou Controller a BIOS schopny rozpoznat určitá zařízení Hot

Plugu v PCI slotech a spustit odpovídající průběhy funkcí podle údajů výrobce.

Na rozdíl od PCI Hot Plugu nabízí standardizovaný PCI Express Native Hot Plug administrátorovi při správě systému rozhodující výhody. Pokud použité komponenty odpovídají standardu PCI Express, pak není potřeba brát při používání funkcí ohled na výrobní vlastnosti. Pokud například vložíme do PCI Express slotu zařízení typu Native Hot Plug, je tento stav přímo signalizován operačnímu systému. To nahrazuje definované Service-Interrupty ovladačů a nastartuje odpovídající funkce Hot Plugu, které operační systém již podporuje.

PCI Express Graphic Interface

Dny konektorů AGP jsou sečteny. S prvními základními deskami s podporou PCI Express jsme se setkali již v loňském roce. Výrobci grafických čipů a karet jako ATI, NVIDIA a 3DLabs současně nabízejí produkty, které jsou pro novou sběrnicí určeny.

Grafický port pro PCI Express je již od počátku k dispozici v provedení $\times 16$ Link. Tato konfigurace umožňuje teoretickou šířku pásma 3,73 GB/s v jednom směru. Proti tomu stojí interface AGP $8\times$ s 1,99 GB/s. Počet přípojek stoupne z 132 pinů u Universal AGP na 164 pinů u zástrčky PCI Express.

Jednou z výhod nové technologie sběrnic PCI Express je bezkonfliktní routing vodičů. Pouze jediný pár vodičů musí být absolutně symetrický a musí mít stejně dlouhý layout. U délek různých párů vodičů ponechává tato specifikace více volnosti, protože informace o Clocku a datech jsou přenášeny přes každý pár vodičů, proto není Timing kritický. Architektura AGP se skládá z oddělených vodičů Data a Strobe, které spolu korelují. Kvůli kompenzování signálních dob jednotlivých vodičů je potřebné komplexní a exaktně definované vedení vodičů (Serpentine Routing). Zásobování napětím pro rozhraní je omezeno na 3,3 a 12 V, běžné rozhraní AGP požaduje navíc i zásobování napětím 5 a 1,5 V.

PCI Express slot může dávat až 75 W elektrického výkonu. Naproti tomu u standardu AGP je podle specifikace možné dosáhnout maximálně 25 W. V provedení AGP Pro je v závislosti na verzi k dispozici 50 a 110 W. Verze PCI Express Pro se stále nachází ve vývojové fázi konsorcia PCI-SIG (<http://www.pcisig.com/home>).

Mini PCI Express

Mini PCI Express je založena na standardu PCI Express. Toto rozhraní, které není koncipováno pro konečného uživatele, by mělo zvládat speciální rozšíření Onboard jako WLAN, modem v notebooku nebo PC.

K hlavním změnám oproti běžné zásuvce Mini PCI Type III patří poloviční součinitel tvaru karet Mini PCI Express. Teoretická datová výkonost se u nového rozhraní navíc zvýšila ze 127,2 MB/s na 596 MB/s (jedna PCI Express linka - viz

– s taktovací frekvencí 2,5 GHz). Tato vysoká šířka pásma umožňuje při datovém přenosu přestup od paralelní na sériovou technologii PCI Express.

Kvůli zaručení vyšší flexibility rozhraní Mini PCI Express integrovali vývojáři kromě sběrnic PCI Express i sběrnic SM a USB 2.0. Rozhraní Mini PCI disponuje celkem 52 zásuvkovými kontakty. Odstup jednotlivých pinů je 0,8 mm. Interface dodává 750 mA přes kontakty 3,3 V a 375 mA přes kontakty 1,5 V. Konečnou specifikaci pro interface Mini PCI Express konsorciem PCI SIG () předložilo v roce 2003.

ExpressCard

ExpressCard je další evolucí standardu PC Card na bázi PCI Express. PCMCIA Trade Association představila kartu ExpressCard specifikace 1.0 v září roku 2003 pod označením NEWCARD. Standard ExpressCard díky většímu výkonu a menšímu součiniteli tvaru nahradí dosavadní PC Cards v notebookech a měl by být využit i v oblasti PC.

Vyšší šířku pásma u standardu ExpressCard garantuje technologie PCI Express. Volitelně lze použít i USB 2.0. Univerzální zástrčky ExpressCardu disponují dostatečným množstvím kontaktů pro obě varianty a podporují oba dva systémy sběrnic. Budoucí notebooky a PC nebudou k ovládání ExpressCardu potřebovat žádné přídatné Controllery: PCI Express a USB 2.0 pak budou patřit ke standardním vlastnostem systémových čipsetů. Notebooky se zástrčkami PC karet ještě potřebují vlastní PC Card Controller, který tak zvyšuje náklady.

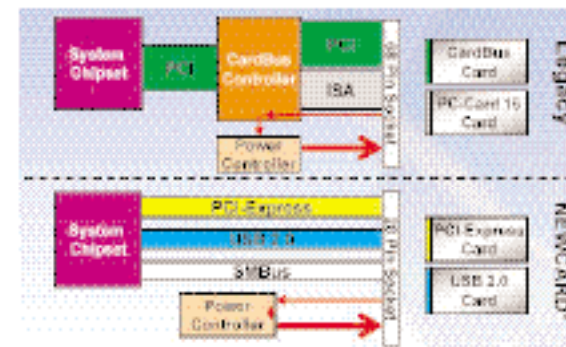
ExpressCard specifikace 1.0 definuje dva součinitele tvaru: ExpressCard/34 s šířkou 34 mm a ExpressCard/54 s šířkou 54 mm. Druhá odpovídá šířce běžné PC Card a přijímá například odpovídající pevné disky. Co se zástrček týče, jsou plánovány také oba součinitele tvaru. Univerzálně použitelný slot ExpressCard/54 může každopádně přijímat i ExpressCard/34. Výška ExpressCardu je standardně 5 mm, při shodné délce 75 mm.

PCMCIA poprvé nabízí u ExpressCard specifikace 1.0 společný standard pro rozšiřující karty, který platí pro notebooky i PC. Současné PC Cardy se prosadily výlučně v mobilních počítačích. Podobně jako USB 1.1/2.0 má i ExpressCard v budoucnosti spojovat oba druhy, a tím nabízet vyšší funkcionalitu. Vedle PCMCIA Trade Organisation podporují vývoj ExpressCard i Intel, Microsoft, Dell, Lexar Media a SCM Microsystems. Průmyslové skupiny USB-IF (<http://www.usb.org/home>) a PCI-SIG spolupracují na základě svých rozhraní i na rozšířeném standardu. Měla by existovat i široká infrastruktura systémů s čipsety PCI Express, které by přímo podporovaly komponenty ExpressCardu.

Shrnutí

Zásadní výhoda PCI Express spočívá ve škálovatelnosti, softwarové kompatibilitě s PCI a v univerzálním použití například v oblasti ser-

► **DualBus: zástrčky ExpressCard (označení NEWCARD) nabízejí jak systém sběrnic USB 2.0, tak PCI Express. Karty si mohou svobodně vybrat. Nahoře na obrázku: běžné řešení PC Cardu se zvláštním čipem Controlleru (zdroj: Intel).**



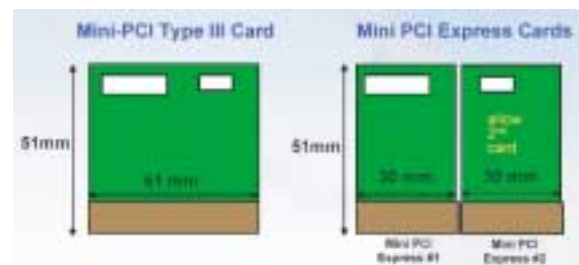
▲ ExpressCard budou používány majiteli notebooků i běžných stolních PC ►

verů a pracovních stanic. Podle potřebné šířky pásma je totiž možné paralelně zapojit na jednom portu více linek. Minimální konfigurace s jednou linkou dosahuje 2,5 GB/s v jednom směru. Vývojáři považují za možné dokonce i zečtyřnásobení přenosového výkonu. Vysoký výkon přenosu však relativizuje o dvacet procent snížený podíl na uživatelských datech, což je podmíněno kódovacím procesem 8B/10B. PCI Express se navíc dá použít nejen jako spojení I/O pro karty, ale rovněž jako vzájemné propojení Chip-to-Chip. „Sériovým“ systémem sběrnic a především nákladným protokolem přenosu s kódováním 8B/10B vznikají při přenosu dat dlouhé latenční doby, takže PCI Express nepřichází v úvahu jako sběrnic pro procesory nebo paměti. V tom má výhodu HyperTransport, neboť zde tato zátěž neexistuje. Nová technologie sběrnic bude použita i jako základ pro používání rozhraní AGP a Mini PCI, stejně jako u rozhraní PC Card. Podle plánů Intelu by další vývoj PCI Express Advanced Switchingu pod názvem PCI AS měl přinést i jeho využití jako datové sběrnic v síťových procesorech. Dalším plusem PCI Expressu by měly být i nízké náklady, vzhledem k malému počtu vodivých drah. Na druhé straně PCI Express vyžaduje novou spojovací architekturu na systémové desce a nové karty, ja-



ko například grafické karty. Je třeba brát zřetel na to, že PCI-X je jako varianta se zpětnou kompatibilitou s PCI stále k dispozici a v provedení PCI-X 1066 nabízí i do budoucna dostatečný výkonostní potenciál pro systémová řešení. Předběžné odhady hovořily o tom, že PCI Express bude mít v zaváděcí fázi těžkosti, aby se prosadila proti etablovaným architekturom sběrnic a zavedeným základům karet. Praxe však ukazuje, že výrobci grafických karet se již chopili příležitosti a architekturu PCI Express plně podporují.

5 0074/VAC □



◀ **Mini PCI Express: taktó vypadají první prototypy karty ve formátu Mini PCI Express (zdroj: Intel).**

Přehled maximálního příkonu PCI Express slotů

Součinitel tvaru PCI Express	Příkon – standardní karta	Příkon – nízkoprofilová karta
PCI Express $\times 1$	10 W (1) / 25 W	10 W
PCI Express $\times 4/\times 8$	25 W	10 W
PCI Express $\times 16$	25 W (1) / 75 W	25 W

Hodnoty označené (1) používá PCI Express během inicializační fáze. Po konfiguraci jako highpower zařízení pracuje PCI Express s vyšší hodnotou, nutnou pro komponentu.