

Procesory Pentium

Pentium je procesor vyráběný od roku 1993 (původně 80586), od té doby se Pentium používá jako obchodní značka pro několik generací procesorů postavených na architektuře IA-32.

Kromě zvyšování výpočetního výkonu se výroba zaměřila také na hardwarovou podporu výpočetních systémů (OS, grafika, atp.), to vede k více jednotkám na čipu, větší spotřebě a nárokům na chlazení a to vede ke snižování napětí a tenčím litografickým čarám. Procesory Pentium se do jisté míry specializují (servery, notebooky, atp.).

Procesory Pentium jsou superskalární. Superskalární procesory využívají paralelní vydávání instrukcí a paralelní zpracování zároveň a obsahují dynamické plánování vydávání, což znamená, že je procesor schopen v průběhu vykonávání programu seskupovat instrukce do skupin, které je možno paralelně zpracovat.

Adresování

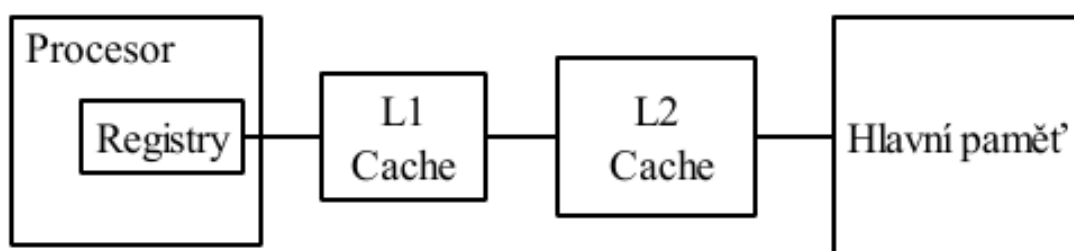
Všechny procesory od Intelu používají pro přístup k paměti adresu složenou ze dvou částí – segment a ofset, jejich spojením dostaneme binární číslo, které se používá na linkách mezi procesorem a pamětí. Ofset určuje vzdálenost od začátku bloku, segment pak samotný blok. Existují dva mechanismy převodu segmentu na binární číslo:

- reálný režim – segment se násobí 16 a fyzická adresa je pak toto číslo plus ofset
- chráněný – segment je pouze ukazatel do tabulky popisovačů, kde je uložena počáteční hodnota bloku

Segmentový registr má šířku 16 bitů, bity 0 a 1 určují úroveň oprávnění, bit 3 určuje tabulku popisovačů a bity 4-15 je index v této tabulce. Známe tři typy tabulek popisovačů:

- globální – unikátní a mohou zde být jakékoliv objekty se všeobecnou platností
- lokální – může jich být více a mohou obsahovat pouze objekty místní (task) platnosti
- pro přerušení – unikátní a obsahuje objekty obsluhující přerušení

Stránkovací mechanismus umožňuje programům užívat virtuální paměť, která je rozdělena na stránky konstantní délky, které jsou uloženy na externím zařízení. Fyzická paměť je také rozdělena na stránky této délky, OS pak zajišťuje aby potřebná stránka byla z externího média natažena do rámce fyzické paměti. Velikost rámce bývá 4KB, ale může být i 2 a 4MB.



CACHE

Kvůli malé rychlosti přístupu k operační paměti vznikla paměť CACHE, která je implementována uvnitř procesoru a slouží jako mezistupeň mezi operační pamětí a procesorem. Je rychlá, drahá a malá a většinou je rozdělena na CACHE pro data a CACHE pro instrukce. Úspěšné použití paměti předpokládá vyřešit tři problémy a to mapování, aktualizace a konzistence. Hledání v paměti probíhá tak, že se vstupní hodnota klíče porovnává s hodnotami klíčů v tabulce (všechny najednou). Možnosti organizace mapování:

- plně asociativní – libovolný blok hlavní paměti může být v libovolném bloku CACHE, nejlepší a nejdražší řešení
- přímé – určité bloky paměti jsou přiřazeny právě jednomu bloku CACHE
- částečně asociativní – kombinace výše uvedených metod, dnes nejrozšířenější

Strategie zajištění konzistence dat mezi CACHE a hlavní pamětí:

- strategie současného zápisu (write-through) – data se modifikují jak v CACHE, tak v hlavní paměti, což je pomalé, protože hlavní paměť brzdí CACHE
- strategie zpožděného zápisu (write-back) – ukládá změny pouze do CACHE, blok je pak označen jako modifikovaný a v hlavní paměti se změny projeví až když je blok z CACHE uvolněn
- strategie předběžného zápisu – kombinace výše zmíněných

Proudové zpracování

Proudové zpracování znamená rozdělení činnosti na více posloupných kroků a v každém okamžiku je prováděno několik činností, každá v jiném stupni rozpracovanosti. Využívá se při provádění aritmetických operací a při vykonávání instrukcí. 3 předpoklady: soustavný přísun dat, činnost musí být rozdělitelná na více úkonů, trvání operací musí být přibližně stejně dlouhé. U proudového zpracování se musíme vyvarovat následujícím konfliktům:

- Datové - jsou způsobeny porušením závislostí mezi jednotlivými operacemi a jejich daty (jedna operace potřebuje ke svému vykonání data předešlé, která je ještě nevydala)
- Skokové - dochází k nim v případech kdy instrukce mění pořadí následujících instrukcí

Vyřešit tyto konflikty se může dvěma cestami, softwarově nebo hardwarově. Hardwarová řešení předpokládají vyřešení konfliktů během vykonávání programu aniž by došlo k výraznějšímu snížení výkonu. Používá se například predikce vícenásobného větvení. K dvěma předešlým závislostem je třeba dodat ještě jednu a tou je závislost prostředků. To znamená, že instrukce potřebuje prostředek, jež využívá instrukce předešlá (čtení/zápis operandů ve stejném taktu na stejném registru).

FPU

FPU slouží jako jednotka pro práci s čísly celými, v plovoucí řádové čárce nebo čísly BCD. Jednotka FPU vykonává příkazy z normálního proudu instrukcí procesoru. Může být umístěna buď vně nebo uvnitř CPU, dnes platí kvůli rychlosti zásadně druhý případ.

Obsahuje 8 80-ti bitových registrů, ve kterých jsou uložena čísla v přesnosti double extended precision. Pokud FPU dostane číslo jež není v tomto formátu automaticky jej na něj převede. Tyto registry jsou uspořádány jako zásobník a taky se k nim takto přistupuje.

FPU dále obsahuje 16 bitový stavový registr, který určuje stav FPU. Také má samostatnou hardwarovou sčítačku i násobičku a řídicí registr. Výroba se řídí dle standartu IEEE754.

Doplňující obrázky k FPU jsou na konci dokumentu.

Zobrazení reálných čísel

Existuje více zobrazení:

- přímé – mantisa i exponent jsou opatřeny znaménky zcela vlevo, řádové čárky vpravo
- exponent s posuvem
- se skrytým bitem – musí být normalizované

Reálné číslo je reprezentováno následujícím způsobem (ukázka na 20-ti bitové přesnosti):

Nejprve první bit určuje znaménko čísla $((-1)^n$, kde n je hodnota bitu), dále je 8 bitů exponent (od něj je odečteno 127) čísla 2 a nakonec 7 mantisa (1.mantisa kde mantisa je řetězec desetinných míst), všechno se následně vynásobí. Čili číslo $5 = 1 \times 4 \times 1.25$ je reprezentováno jako 0|100 0000 1|010 0000.

Možnosti zapsání čísel:

Binární čísla bez znaménka:

- bajt – 8 bitů
- slovo – 16 bitů
- dvojslovo – 32 bitů
- čtyřslovo – 64 bitů
- dvojitě čtyřslovo – 128 bitů

Celá čísla kladná nebo záporná:

- bajt – 128 kladné až 127 záporné
- slovo – 32.768 kladné až 32767 záporné
- dvojslovo – 2.147.483.648 kladné až 2.147.483.647 záporné
- čtyřslovo – 1.844×10^{19} z toho polovina kladných a polovina záporných

Reálná čísla:

- jednoduchá přesnost – 32 bitů, z toho mantisa 24 a exponent 8
- dvojitá přesnost – 64 bitů, z toho mantisa 52 bitů a exponent 11
- rozšířená dvojitá přesnost – 80 bitů, z toho mantisa 63 a exponent 15

Pracovní režimy:

- Chráněný – standardní režim 32-bitových instrukcí. Chrání data jedné aplikace v paměti proti přepisu od druhé. Víceúlohový (multitasking) – umožňuje zastavit pouze chybný program.
- Reálný – pro zajištění kompatibility se staršími procesory (80286, 8086 a 8088). Mohou v něm běžet 16-bitové aplikace.
- Virtuální reálný režim – může zde běžet 16-bitová aplikace, ale je jí přidělen pouze 1MB paměti. Plně emuluje prostředí 8088.

Přepínání úloh

Chráněný režim umožňuje zavedení více úkolů do paměti z nichž právě jeden je vykonáván.

Úkolem se rozumí program, obsluha přerušení nebo výjimečná situace. Přepínání mezi těmito úkoly obstarává hardware. Jsou složeny ze dvou částí:

- výkonný prostor – data a kód
- stav úlohy – TSS (Task-State Segment), slouží k uchování kontextu v momentě přepnutí. Kontext obsahuje data statická (nastavují se při inicializaci procesu) a dynamická (mění se při přepnutí).

Instrukční sada

Instrukční soubor procesorů Intel prošel velkým vývojem. Základní instrukční soubor u I80386 měl 128 instrukcí, I80486 přidalo dalších 7 pro řízení CACHE a konverzi velký/malý indíán. Počáteční verze Pentia přidaly 5 instrukcí (CUID a RDTSC pro práci s registrem časového razítka). Pentium Pro rozšířilo sadu o dalších 5 instrukcí (CMOV a FCMOV pro řídicí závislosti), Pentium II přidalo 4 (SYSENTER a SYSEXIT pro rychlý přístup ke službám OS), Pentium III dalších 9 pro práci s FPU a nakonec Pentium IV o dalších 8 instrukcí (zajímavé jsou prefixy pro řízení predikce skoků).

Dalšími podstatnějšími rozšířeními jsou:

- MMX – 47 instrukcí pro práci s multimédií
- SSE – 79 instrukcí pro práci se 128 – bitovými položkami
- SSE2 a SSE3 – přidáno 144 instrukcí respektive 13 instrukcí na Pentiu IV

P5

Pentium (1993)

Pentium byl jako první mikroprocesor navržen kombinací CISC a RISC. Umožňuje symetrický multiprocessing (dva procesory na jedné desce).

První verze Pentia měla tloušťku čáry 0,8 μm a 3,1 milionu tranzistorů na čipu. Frekvence byla 66 MHz, spotřeba 16W, napájecí napětí 5V, šířka datové sběrnice 64 bitů a adresové 32 bitů (může adresovat 4GB paměti a 64TB virtuální). Registry mají délku 32 bitů, obsahuje dvě CACHE 8KB. Pouzdro bylo nazýváno PGA a mělo 273 vývodů.

Další verze běhali na frekvence až 266 MHz, měli napájení 3,3V a CACHE se dočkala zvětšení na 16KB.

Pentium MMX (1997)

Frekvence byla 166 MHz, instrukční sadu mělo rozšířeno o 57 nových instrukcí pro zpracování grafiky, zvuku a videa. Výpočetní výkon byl 120 MIPS, tloušťka čáry 0.35 μm a na čipu bylo 4,5 milionu tranzistorů.

Funkční jednotky:

- FPU – jednotka pro práci s čísly v plovoucí řádové čárce
- Celočíselná jednotka – obsahuje ALU1, ALU2 a registry
- Segmentační jednotka – připravuje a vypočítává adresy
- Stránkovací jednotka – stará se o virtuální správu paměti
- Řadič – obsahuje frontu instrukcí, dekodér instrukcí a řadič
- CACHE instrukcí – vybírá instrukce a má v sobě frontu instrukcí

P6

Pentium PRO (1995)

Obsahuje v jednom pouzdře dvě jednotky – vlastní procesor a CACHE L2 o velikosti 256KB, 512KB nebo 1MB. L1 CACHE je umístěna přímo v procesoru a má velikost 16KB - 8KB je dvoucestná asociativní pro instrukce a 8KB je čtyřcestná asociativní pro data.

Používá architekturu DIB, což znamená dvě nezávislé sběrnice a L2 CACHE, která je k procesoru připojena samostatnou sběrnicí na plné rychlosti procesoru (umožňuje i multiprocessorový systém až ze 4 procesorů). Umí dynamické vykonávání instrukcí (instrukce zpracovány na základě logiky a ne podle seznamu), existují 3 metody tohoto vykonávání: predikce vícenásobného větvení, analýza dat a spekulativní vykonávání instrukcí.

Datová sběrnice má šířku 64 bitů, adresová 36 bitů. Napájeno je 3,3V, později 3,1V. Příkon je 25W a maximální frekvence 200MHz. Nemá MMX.

Pentium II (1997)

Pentium II je v zásadě vylepšené Pentium PRO. L1 CACHE má velikost 32KB (2x16), tloušťka čáry je 0,35μm později 0,25μm. Je napájeno 2,8V, později 2V. Frekvence dosahuje 233 až 450 MHz, spotřeba 35W (při 233MHz), 27W(při 450MHz – menší čára a menší napětí). Jsou zde implementovány instrukce MMX a podporuje symetrický multiprocessing.

Pentium III (1999)

Instrukční sada se dočkala rozšíření o SSE. Obsahuje ochranu proti přetaktování a teplotní čidlo. Tloušťka čáry je 0,25μm, později 0,18μm. Původně L2 CACHE měla 512 KB mimo jádro a pracovala na poloviční rychlosti jádra, později bylo 256KB L2 CACHE integrováno na čip při plné rychlosti jádra. Frekvence maximálně 1GHz, podpora symetrického multiprocessingu.

P7

Pentium IV (2000)

Inovace oproti předešlým verzím:

- Dvoustupňová jednotka proudového zpracování instrukcí
- Nová architektura CACHE
- ALU pro celá čísla pracující na dvojnásobné rychlosti
- Vylepšená FPU
- Rozšiřující sada instrukcí (SSE – 2) podporující 128 bitovou aritmetiku
- Zrychlená systémová sběrnice

Postaveno na architektuře NetBurst – spoléhá na vysoké frekvence, obsahuje SSE2, Hyper-Threading (simuluje 2 virtuální procesory pro OS). Čára 0,18, 0,13 a 0,09μm. 42 miliónů tranzistorů.

Prvně běželi na frekvencích 1,4GHz a 1,5GHz, později od 1,5 – 4 GHz. Spotřeba při základní taktovací frekvenci 55W.

Existuje více verzí – Celeron (low-endová verze) a Xeon (high-endová verze).

4 základní funkční části:

- vstupní – zpracovává instrukce v předepsaném pořadí, zajišťuje přísun instrukcí z L2 a upravuje instrukce na zpracovatelné v prováděcí jednotce. Obsahuje CACHE L2, jednotku predikce skoků, dekodovací jednotku a paměť ROM mikrokódu.
- blok řízení zpracování mikroinstrukcí mimo pořadí – obsahuje jednotku alokace vyrovnávacích pamětí, jednotku přejmenování registrů a jednotku plánování provádění mikroinstrukcí mimo pořadí
- blok výkonných jednotek – obsahuje dvě ALU pro celá čísla, dvě FPU a jednotky generace adres operandů a datovou CACHE L1.
- paměťový subsystém – stará se aby v CACHE bylo to co být má a nebylo to, co se potřebovat nebude.

Procesory – vývoj do budoucna

Tzv. Moorův zákon praví, že každých 18 měsíců se počet jednotek na čipu zdvojnásobí. I když to do této doby platilo, je jisté že jednou se musíme dostat na úroveň, kterou už nebudeme schopni překročit a to z těchto důvodů:

- Technologický limit – procesory se vyrábějí pomocí vrstvy citlivé na světlo (maska), dnes se však pracuje s tak malou vlnovou délkou, že menší už by znamenalo použít rentgenové záření, na to však není známo z čeho vyrobit masku.
- Fyzikální limit – velikosti tranzistorů na čipu jsou již tak malé, že se dostaneme k bodu, kdy už menší vyrobit nelze.

Kvůli těmto limitům se již dále zvyšovat výkon „hrubou silou“ (tzn. přidáváním jednotek na čip) nedá, proto se musí přejít na jinou strategii. Z toho důvodu se v dnešní době už s architekturou Pentia IV do budoucna nepočítá a přešlo se zpět na architekturu Pentia III a architekturu Core, jež znamená menší odběr elektrického proudu, menší zahřívání a zaměřuje se na větší kapacitu paměti CACHE.

Upřesnění

CISC

CISC architektura má v instrukční sadě implementované komplexní instrukce. To jsou instrukce, které předepisují provedení více operací. Architektura CISC se vyznačuje:

- proměnlivou délkou kódování instrukcí
- ALU instrukce mají obvykle dva operandy a operandem může být i adresa paměťového místa
- instrukce mohou implementovat více než jednu operaci
- komplikovanou realizací proudového zpracování instrukcí

U novodobých procesorů pro PC, které jsou označovány jako „spíše“ CISC, však provádění mikroinstrukcí zajišťuje RISC – ový procesor , který je označován jako jádro procesoru. Využívá se tak rychlosti, se kterou umí procesory typu RISC proudově zpracovávat instrukce. Pro aplikace procesor poskytuje instrukční sadu, která obsahuje i komplexní instrukce (CISC architektura).

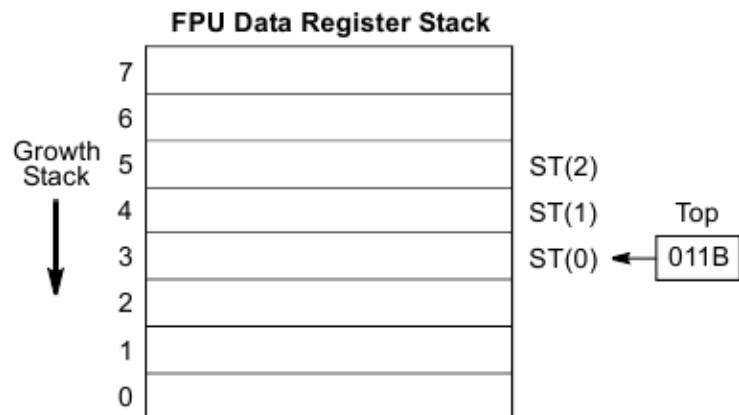
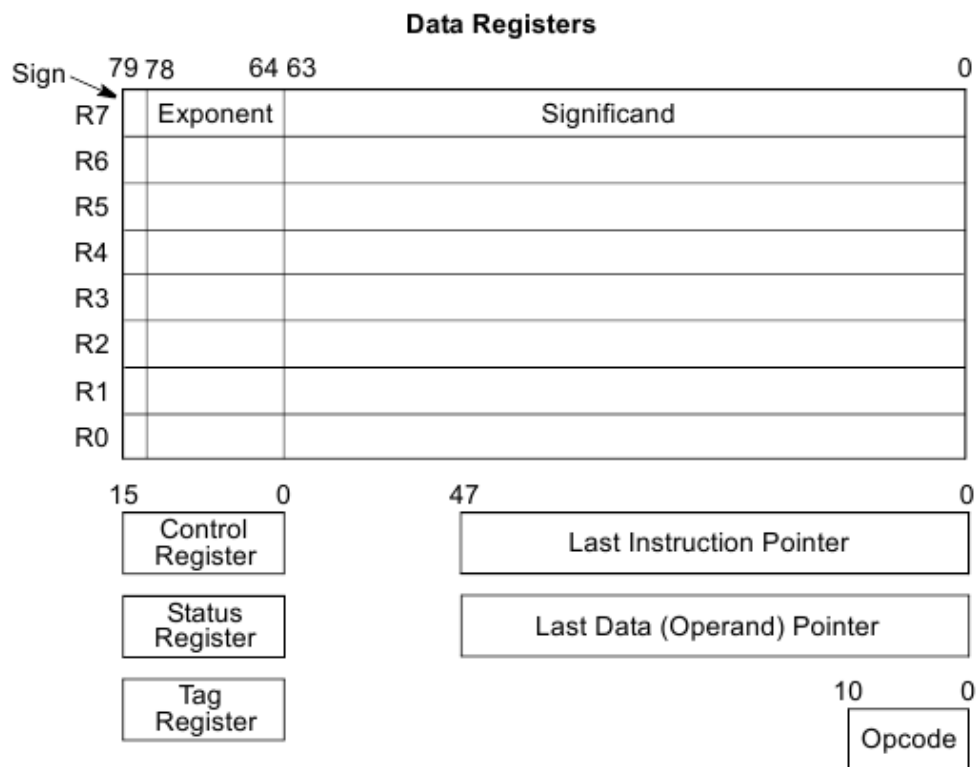
Hyper Threading

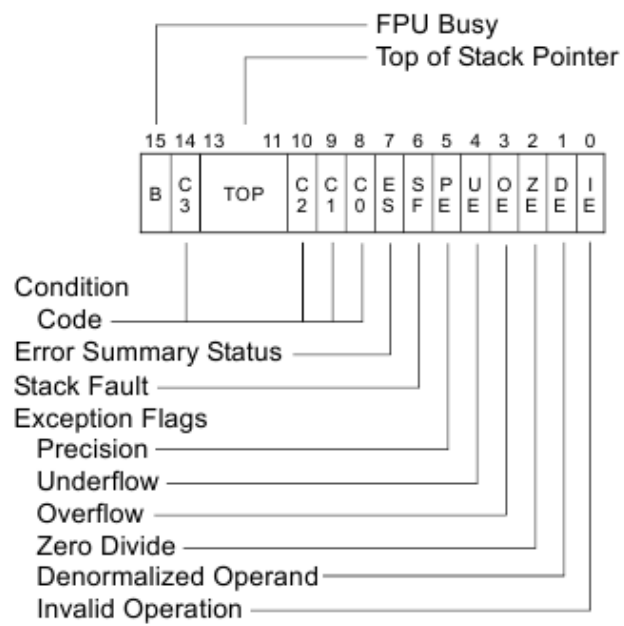
HT znamená zpracovávání více úloh najednou. Toho je docíleno díky replikaci funkčních bloků, které využívají souběžné vykonávání několika instrukcí. Někdy se jim říká logické procesory a každý z nich má své, dělené zdroje mezi jednotlivé logické procesory a sdílené zdroje, které využívá každý. Každý procesor s touto technologií má všechny fyzické registry replikované.

IEEE754

Standart pro práci s čísly v plovoucí řádové čárce. Defínuje 4 formáty pro práci s těmito čísly (32 bit single precision – podle standartu je jediná vyžadována, 64 bit double precision, ≥ 43 bit a nakonec ≥ 79 bit většinou 80 bit extended precision). Všechny formáty jsou ukládány stejným způsobem – znaménko(1 bit), pak exponent a mantisa (u extended precision – 15 bitů exponent a 64 bitů mantisa). V registrech je možno ukládat tyto typy čísel: normalizovaná čísla, denormalizovaná čísla (nulový exponent), nula, nekonečno a nečíselná hodnota (maximální exponent a nenulová mantisa).

Obrázky k FPU





Struktura stavového registru