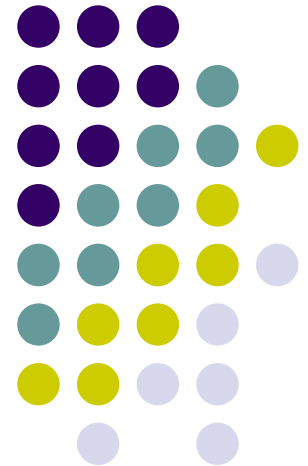


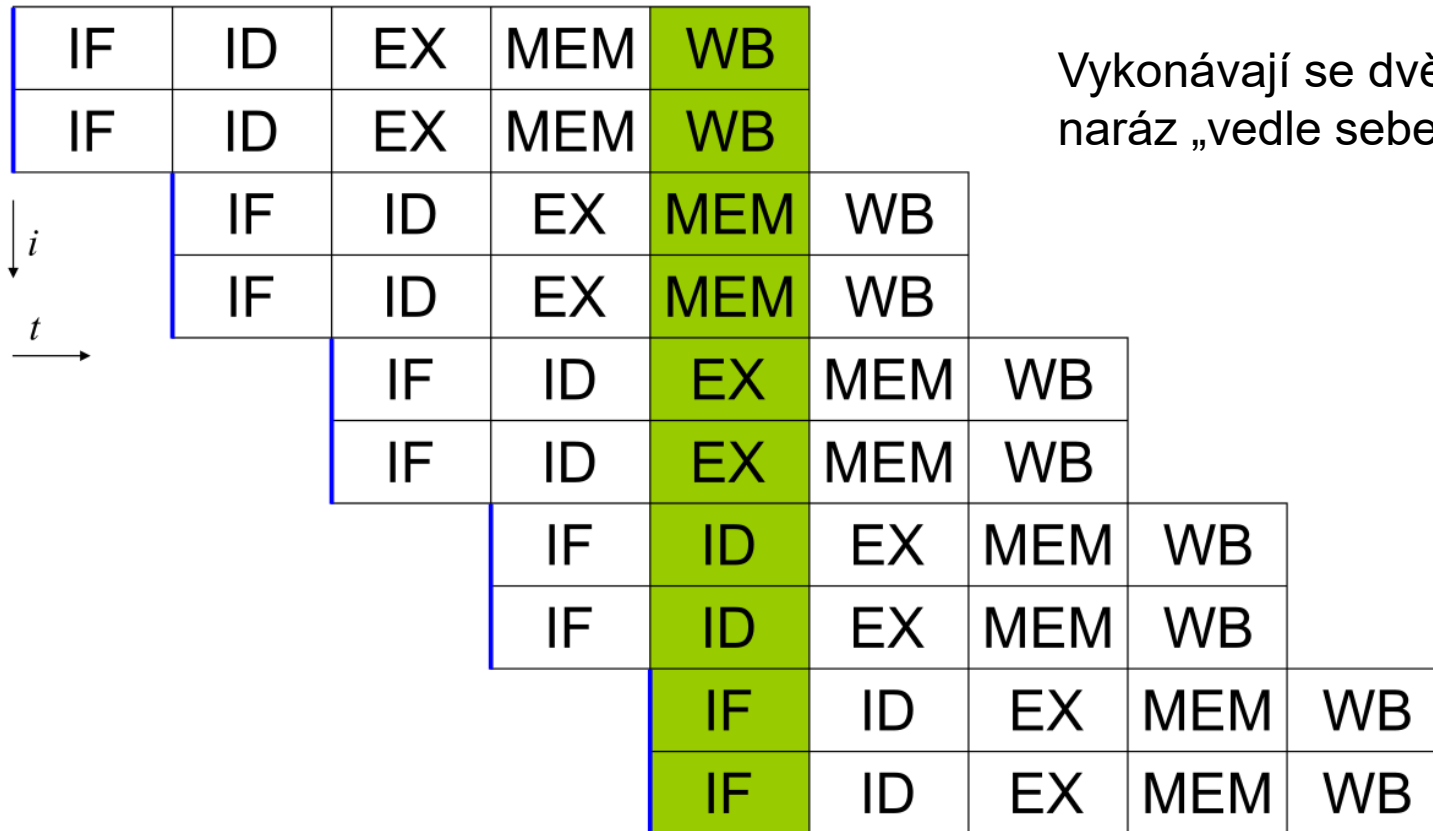
První superskalární mikroprocesory

Pentium, Pentium PRO, Pentium MMX





Superskalární mikroprocesor



Vykonávají se dvě instrukce
naráz „vedle sebe“

Pentium

- Uveden na trh v roce **1992**
- V prvních zprávách označován jako 80586 nebo jako **P5**, ale nakonec se ujalo označení **Pentium**
- Procesor je navržen kombinací technik **RISC** a **CISC**
- Pentium je prvním **superskalárním** procesorem firmy Intel
- Je dodáván v pouzdře PGA s **273** vývody
- na čipu o rozměrech 12,8 x 12,8 mm integruje asi **3,1 milionu** tranzistorů
- Šířka adresové sběrnice **32 bitů** (adresuje 4 GB paměti)
- Šířka datové sběrnice **64 bitů** (**dvojnásobná** oproti předchůdci 486)
- Registry a ALU jsou **32-bitové**
- FPU pracuje s **64b** nebo **80b** daty (reálnými čísly)
- Procesor považujeme stále ještě za **32-bitový**
- Na chipu jsou integrované dvě **oddělené cache pro program a data**, z nichž každá má kapacitu **8 kB**
- V **datové cache** je uloženo 8 kB nejčastěji používaných dat
- V **instrukční cache** jsou uloženy naposledy provedené části programu (program má tendenci se vracet, urychlí se provádění cyklů) Tato cache je read-only bez funkce write-back



© www.cpu-world.com

Superskalární procesor



- **Superskalární** procesor - procesor, který má **více než jednu frontu** pro zřetězené zpracování instrukcí a více ALU, ve kterých se jednotlivé stejné fáze různých instrukcí současně provádějí
- Konkrétně Pentium má dvě fronty označované jako **U, V**
- Během jednoho taktu může **dokončit až dvě instrukce** zároveň
- Ve frontách U a V je realizován **6-stupňový pipeling**
- **V procesoru jedou dva 6-stupňové pipeliningy vedle sebe**
- Po spuštění programu tedy mohou být v 6. taktu dokončeny zároveň první a druhá instrukce, v 7. taktu třetí a čtvrtá instrukce současně, v 8. taktu pátá a šestá instrukce.....
- Díky tomu Pentium dosahuje při stejné frekvenci vyššího výkonu než procesory 80486

Superskalární procesor



- Pentium kombinuje filosofii RISC i CISC
- Složité CISC instrukce jsou rozloženy na **mikroprogram** popsany mikrokódem
- Každý mikropovel tohoto mikroprogramu pak lze vykonat v **6 krocích**
- Jednoduché instrukce se vykonávají bez překladu na mikroprogram – přímo hardwarovými prostředky
- Při superskalárním zpracování však není možné zpracovávat vždy dvě instrukce naráz
- Aby se tak mohlo dít, musí být splněny jisté předpoklady:
 - následující instrukce nesmí být **závislá** na instrukci předcházející (následující instrukce nesmí potřebovat výsledek instrukce předcházející)
 - obě instrukce musí být „**jednoduché**“, tj. nejsou prováděny *mikroprogramově*, ale *hardwarově* (nerozkládají se na jednodušší)



Superskalární procesor

- Program:

A=5

B=17

C=A+B

D=A-B

E=C/2

F=E+4

G=A*B

Datová závislost znemožňuje současné provádění těchto dvou instrukcí

Superskalární zpracování

A=5 B=17

C=A+B D=A-B

E=C/2

F=E+4 G=A*B

Tyto dvě instrukce mohou být provedeny současně

Tyto dvě instrukce mohou být provedeny současně

Tato instrukce se dokončí samostatně, protože následující instrukce je datově závislá na jejím výsledku (pracuje s hodnotou E)

Tyto dvě instrukce mohou být provedeny současně

Superskalární procesor



Program

A=17

B= 2 *A

C=B + 7

D= \sqrt{C}

E=D+9

F=E+A

- Každá následující instrukce tohoto programu je datově závislá na výsledku předchozí
- Tento program na superskalárním procesoru bude bohužel vykonán pouze stejně rychle jako na nesuperskalárním
- V programu nejsou žádné dvě instrukce, které by procesor mohl dokončit současně



Datové závislosti

- Program

A=17

B=9

C=A+3

D=B*2

E=C/2

C=A+B

Pozor! Tyto dvě instrukce **nejsou datově závislé**.
Následující instrukce nepracuje s výsledkem
předcházející

- V tomto programu nejsou po sobě jdoucí instrukce datově závislé
- Všechny instrukce tohoto programu mohou být prováděny po dvou naráz „vedle sebe“

Superskalární procesor



- Díky plynulému pipeliningu lze v každém taktu dokončit jednu rozpracovanou instrukci
- Superskalární procesor umožňuje dokončení více instrukcí naráz ve stejném taktu, ale to neznamena, že se to opravdu v každém taktu povede
- Na superskalárním procesoru lze v některých taktech dokončit dokonce dvě instrukce naráz
- Díky tomu superskalární procesor může za sekundu dokončit více instrukcí, než jaká je jeho taktovací frekvence
- Příklad
- Mikroprocesor Pentium měl taktovací frekvencí 60 MHz
- Díky pipeliningu a superskalární architektuře může za sekundu dokončit 60 milionů až 120 milionů instrukcí (záleží na konkrétním programu a datových závislostech v něm a také na případných problémech plynulosti pipeliningu kvůli podmíněným skokům)

Zřetězené provádění instrukcí



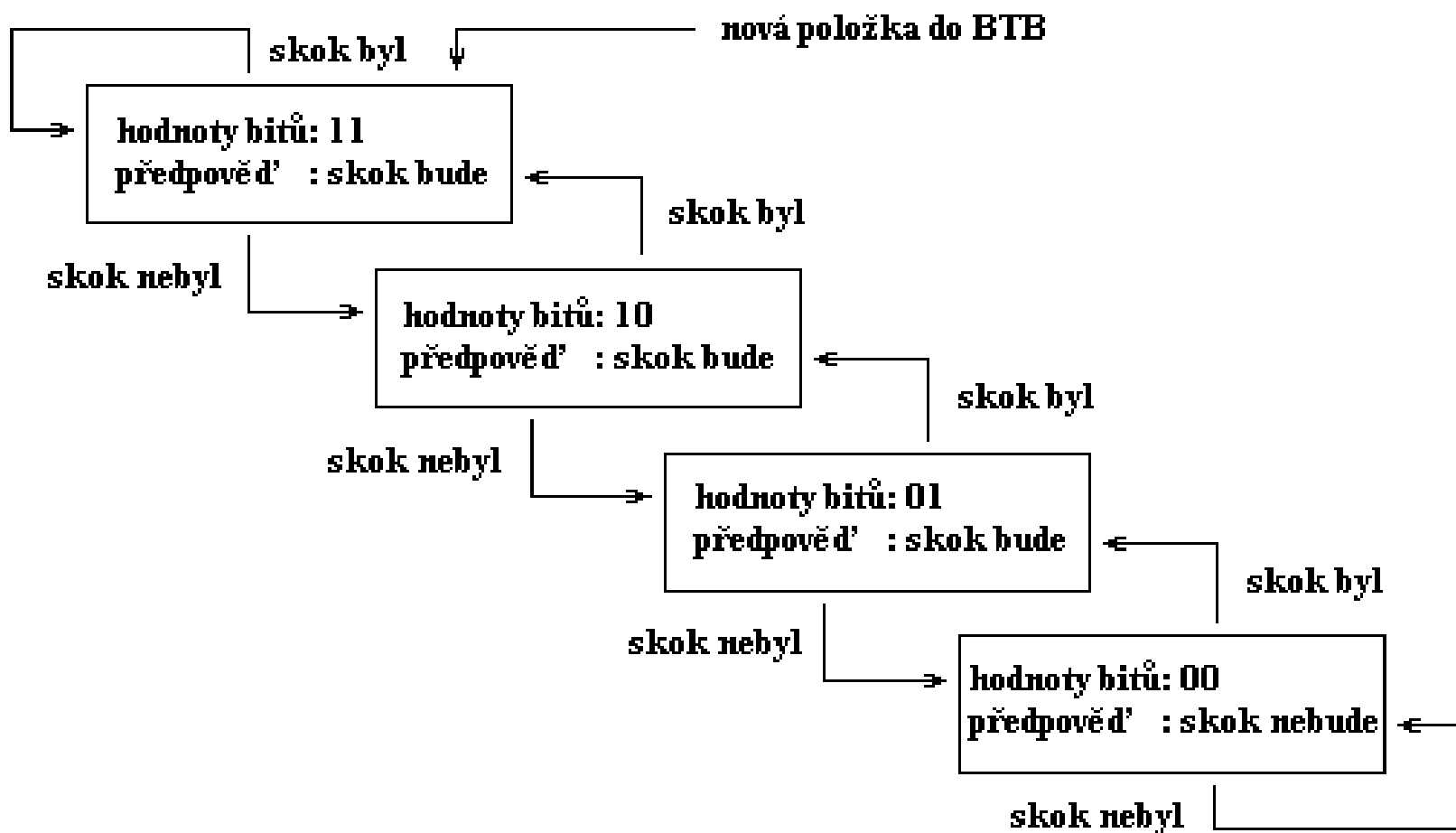
- Největším nepřítelem plynulého pipeliningu je **podmíněný skok**
- Dopředu se neví, jak bude program po podmíněném skoku dále pokračovat
- Aby bylo možné rozpracovávat nějaké další instrukce dopředu, má procesor Intel Pentium zabudováno tzv. **dynamické předvídání skoků**
- podle dosavadního chování programu se snaží **odhadnout**, zda při dalším průchodu instrukce podmíněného skoku způsobí skok nebo ne
- K této činnosti Pentium používá paměť označovanou jako **BTB** (**B**ran**T**arget **B**uffer), v níž jsou uchovány informace o posledních prováděných instrukcích podmíněného skoku, spolu s dvoubitovou informací o historii provedení skoku
- Podle hodnot těchto bitů je také dána předpověď, zda instrukce skok způsobí či ne
- ... a podle toho se postupuje při rozpracování dalších instrukcí za podmíněným skokem

BTB



- Instrukce **podmíněného skoku**, která způsobila skok, je uložena do **BTB** spolu se dvěma bity, jejichž hodnoty jsou rovny 1
- Tyto hodnoty při příštím průchodu přes tuto instrukci signalizují předpověď, že skok se asi bude konat (podmínka bude splněna)
- Pokud podmíněný skok skutečně byl znovu proveden, hodnoty bitů zůstanou nezměněny (11)
- Pokud byla předpověď mylná a skok nebyl proveden, jsou bity nastaveny na hodnotu 10, která opět signalizuje, že skok asi bude
- Podle toho, zda v dalším průběhu programu skok skutečně následuje nebo ne, jsou pak příslušným způsobem bity dále modifikovány a jejich hodnota signalizuje předpověď skoku
- Programy je vhodné optimalizovat tak, aby podmínka všech podmíněných skoků byla formulována tak, aby byla častěji splněna než nesplněna (výchozí předpoklad je totiž vždy ten, že podmíněný skok skočí)

BTB

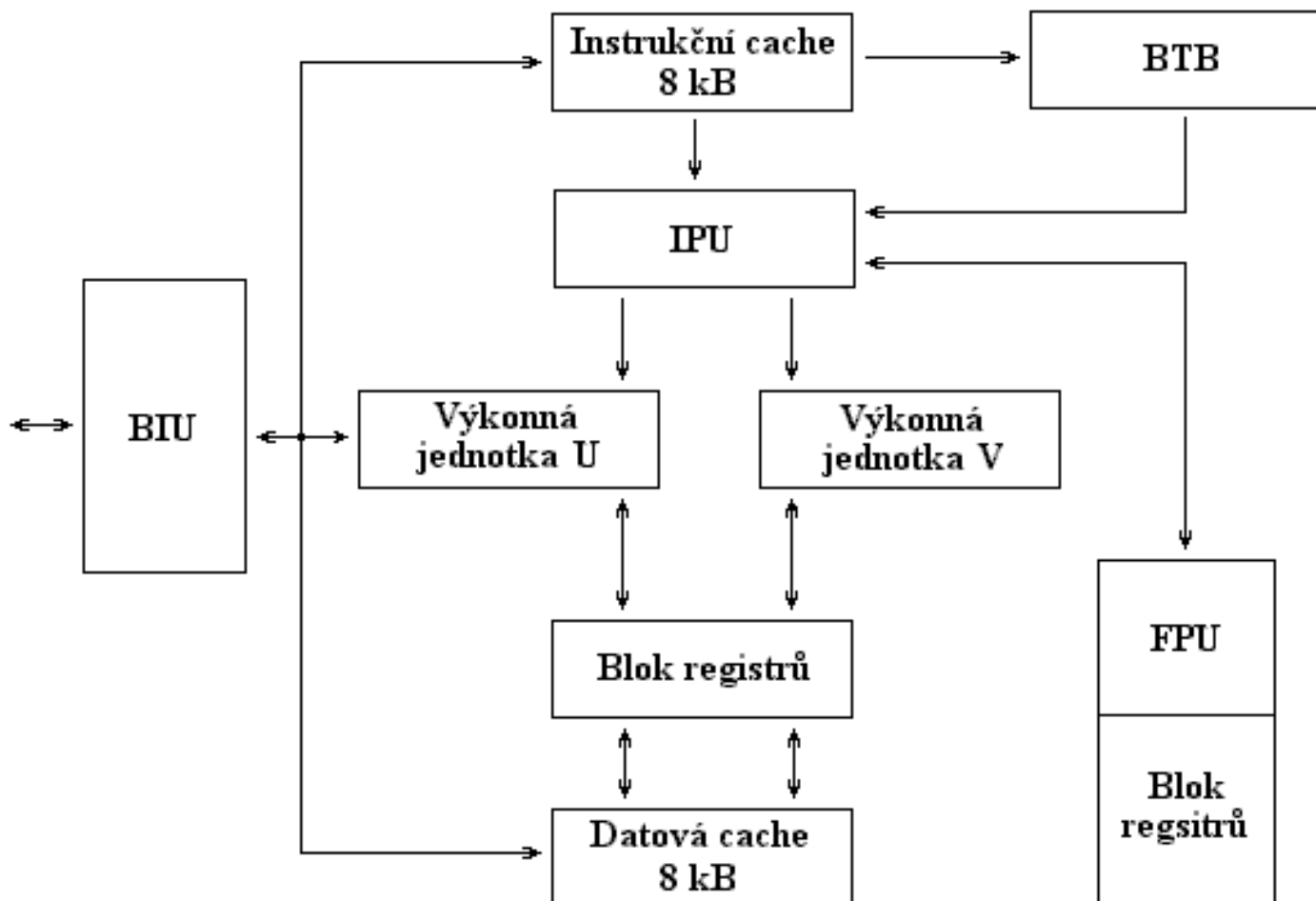


FPU



- **Floating point unit** – jednotka pro výpočty s reálnými čísly (čísla se znaménkem, desetinou čárkou a exponentem)
- ALU neumí s takto zakódovanými čísly provádět aritmetické operace
- Mikroprocesory 8086, 80286 a 80386 obsahovaly pouze ALU a pro urychlení výpočtů s reálnými čísly bylo nutné připojit externí koprocessor (8087, 80287, 80387)
- Mikroprocesor 80486 se vyráběl ve dvou verzích
 - 80486DX plnohodnotná verze s integrovanou FPU
 - 80486SX čip bez funkční FPU
- Procesor Pentium a všechny následující mikroprocesory obsahují jednotku FPU (nebo i více jednotek FPU) a pojem „koprocessor“ se přestává používat
- Moderní mikroprocesory se začínají značně vzdalovat od původní koncepce, kdy platilo mikroprocesor=řadič+ALU
- Mikroprocesory dnes obsahují několik ALU, které navíc pracují proudově v několika fázích a vedle nich pracují nezávisle jednotky FPU.
- Další jednotky mikroprocesoru zrychlují přístup do paměti (TLB) nebo odhadují budoucí větvení programu (BTB), starají se o transformaci adres a chráněný režim (SU, PU, AU...)
- Další velmi složitou částí moderního mikroprocesoru je blok provádějící překlad „vnějších“ CISC instrukcí na uvnitř používané mikrooperace pro RISC jádra (každá ALU nebo FPU je vlastně malý RISCový procesor uvnitř procesoru, který navenek předstírá, že je CISC)
- Znázornit koncepci a fungování moderního mikroprocesoru blokovým schématem je téměř nemožné

Pentium





Problém zahřívání

- S příchodem **Pentia** se poprvé objevují problémy se značným **zahříváním procesoru**
- Procesor 486 byl posledním, který se dal provozovat bez chlazení, ačkoliv rychlejší verze procesoru již vyžadovala alespoň pasivní chlazení
- Procesor Pentium je nutné vždy ochlazovat **chladičem** umístěným přímo na procesor
- U pozdějších verzí Pentia sníženo napájecí napětí z **5 V** na **3.3 V**
- Tepelný výkon klesá kvadraticky ($P=U^2/R$), takže tímto snížením napětí se příkon sníží hned 2,5 krát
- U všech dalších mikroprocesorů uvidíme další postupné snižování napájecího napětí
- První generace: Pentium 60 MHz, 66 MHz, napětí **5 V** - výkon **13 W**
- Druhá generace: Pentium 75, 90, 100, 120, 133, 150, 166, 200 MHz, napětí **3.3 V** - výkon **5 W**



Závislost příkonu na napětí

- $P = U \cdot I$
- $I = \frac{U}{R}$
- $P = U \cdot \frac{U}{R} = \frac{U^2}{R}$
- Závislost výkonu na napětí **je kvadratická**
- Jestliže se napětí zvýší na dvojnásobek, výkon se zvýší 4x
- Jestliže napětí klesne na polovinu, výkon se sníží 4x – tím se také čtyřikrát sníží zahřívání procesoru



Závislost příkonu na napětí

- Příklad
 - Mikroprocesor má příkon **10W** a napájecí napětí **3,3 V**
 - Zlepšením výrobní technologie procesoru se podařilo snížit napájecí napětí na **2,3 V**
 - Jaký má modernizovaná verze procesoru příkon?
-
- Napětí se snížilo $\frac{3,3}{2,3} = 1,434$ krát
 - Příkon mikroprocesoru je na napětí závislý kvadraticky
 - Jestliže se napětí snížilo 1,434 krát, příkon se sníží $(1,434)^2 = \mathbf{2,05}$ krát
-
- Moderní verze procesoru se sníženým napětím bude mít poloviční příkon



Závislost příkonu na napětí

- Příklad
- Moderní mikroprocesor má napájecí napětí **1,15 V** a příkon **30 W**
- Jaký by byl příkon tohoto mikroprocesoru, kdybychom stále používali napájecí napětí 5V jako kdysi.... ?
- Pokud by se stále používalo napájecí napětí 5V, bylo by napájecí napětí $\frac{5}{1,15}$
= 4,348 *krát* větší
- Příkon závisí na napětí kvadraticky
- Kdyby bylo napětí 4,348 krát vyšší, byl by příkon $(4,348)^2 = 18,9$ krát vyšší
- Mikroprocesor by tedy měl příkon $18,9 \times 30 \text{ W} = \mathbf{567 \text{ W}}$
- Takový mikroprocesor by se nedal uchladiť vzduchem

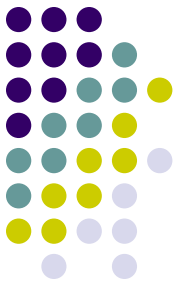
80486DX4



- Přichází na trh po procesoru **Pentium** jako levnější procesor, avšak výkonnější než **80486DX2**
- **Trojnásobná** vnitřní frekvence (overdrive)
- **16 kB** interní cache paměť (dvojnásobná oproti původní 486)
- Procesor **80486DX4 100 MHz** pracuje vnitřně s frekvencí 100 MHz, ale externě s frekvencí 33 MHz
- Procesor **80486DX4 75 MHz** pracuje vnitřně s frekvencí 75 MHz, ale externě s frekvencí 25 MHz
- Snížené napájecí napětí na **3,3 V**
- Tento procesor není superskalární (jde pouze o vylepšenou 486)

Pentium PRO

- Uvedeno na trh v roce **1995**
- pouzdro **387** vývodů
- **5,5 milionů** tranzistorů
- napájecí napětí **2,9 V**
- spotřeba **20 W**



- **16 kB** interní L1 **Cache**
- Novinkou je **integrace L2 cache** paměti o kapacitě **256 kB (512 kB, 1 MB)** přímo do pouzdra procesoru
- L2 cache je větší než L1, ale je o něco pomalejší
- L2 cache se používali již dříve, ale pouze jako samostatný externí chip mezi mikroprocesorem a pamětí
- L2 cache je zde tvořena čipem umístěným v jednom pouzdru s čipem procesoru
- Integrací této cache přibyl v procesoru **dalších 15,5 milionů tranzistorů**
- Varianty Pentium PRO
 - 166 MHz (66 MHz sběrnice, 512 kB L2 Cache)
 - 180 MHz (60 MHz sběrnice, 256 kB L2 Cache)
 - 200 MHz (66 MHz sběrnice, 256 kB L2 Cache)
 - 200 MHz (66 MHz sběrnice, 512 kB L2 Cache)
 - 200 MHz (66 MHz sběrnice, 1 MB L2 Cache)



Pentium PRO

● Varianty

- 166 MHz (66 MHz sběrnice, 512 kB L2 Cache)
- 180 MHz (60 MHz sběrnice, 256 kB L2 Cache)
- 200 MHz (66 MHz sběrnice, 256 kB L2 Cache)
- 200 MHz (66 MHz sběrnice, 512 kB L2 Cache)
- 200 MHz (66 MHz sběrnice, 1 MB L2 Cache)

U mikroprocesorů Pentium se poprvé objevuje velký rozdíl ve vnitřní frekvenci procesoru a vnější frekvenci použité pro komunikaci s okolím. Frekvencí 200 MHz je možné přenášet 64b data uvnitř mikroprocesoru, ale pro přenos dat mezi mikroprocesorem a jeho vnějším vzdáleným okolím je již tato frekvence příliš vysoká.



- Pentium PRO je jedním z nejvydařenějších procesorů firmy Intel a z jeho architektury vychází řada dalších procesorů

- Generace procesorů Intel x86

1. 8086
2. 80286
3. 80386
4. 80486
5. Pentium
6. Pentium Pro
 - Pentium MMX
 - Pentium II
 - Pentium III
7. Pentium IV
8. Intel Core
 - Penryn
 - Nehalem
 - Westmere
 - Sandy Bridge

Tyto procesory jsou v podstatě pouze zdokonalené Pentium PRO a nejsou považovány za novou samostatnou generaci

Architektura Intel Core nevychází z architektury procesoru Pentium IV, ale opět se vrací k Pentium PRO

Superskalární provádění instrukcí



- Procesor obsahuje **3 ALU**, na kterých běží **14-stupňový pipelining** a jednu (také 14-stupňovou) jednotku pro výpočty v pohyblivé řádové čárce (FPU)
- Během jediného taktu mohou být v ALU dokončeny **až 3 instrukce** současně
- Jsou-li ALU plně vytíženy, může být v jednom okamžiku **rozpracováno až 42** (3×14) instrukcí
- Paralelně s tím mohou být prováděny jednotkou **FPU** operace s reálnými čísly
- Takže v jednom taktu můžou být někdy dokončeny **3+1 instrukce** (takový ideální případ se ale stává málokdy)
- Jednotlivé jednotky nezpracovávají přímo instrukce instrukčního souboru 80x86, ale pracují se svým **vlastním souborem jednoduchých instrukcí** označovaných jako **mikrooperace (micro-ops)**
- Jednotlivé instrukce souboru instrukcí 80x86 jsou tak **překládány** do jedné nebo několika **jednoduchých mikrooperací**, které jsou předávány ke zpracování jednotlivým prováděcím jednotkám (CISC -> RISC)
- Mikroprocesor se tedy **navenek tváří jako CISC** a předstírá bohatou instrukční sadu (aby byla zachována **kompatibilita** s předchozími procesory), ale ve skutečnosti je **uvnitř RISC** (aby fungoval plynulý **pipelining**)



Register renaming

- Při provádění **mikrooperací** se využívá techniky zvané **přejmenování registrů** (register renaming).
- Procesor Pentium Pro obsahuje **40 skutečných a navenek skrytých** 32-bitových registrů, které mohou být předstírat libovolný z datových registrů (být jakoby přejmenovány)
- Registry AX, BX, CX atd. nám tedy mikroprocesor pouze předstírá navenek a ve skutečnosti jeho vnitřní RISCové jednotky pracují se 40 skutečnými registry, které jsou ovšem pro programátora neviditelné
- Uživatel procesoru má dojem, že pracuje s registrem AX, EAX, BX.... Ve skutečnosti se data uvnitř procesoru ukládají do jiných registrů, se kterými pracuje RISCový mikroprogram
- Vše je tedy navenek jedna velká iluze – kvůli kompatibilitě s předchozími procesory řady x86

Spekulativní provádění instrukcí



- **Spekulativní provádění** instrukcí mimo pořadí - např. v případě, kdy požadovaná data nejsou v cache paměti, nečeká se až budou načtena z pomalejší operační paměti, ale začne provádět další nezávislé instrukce do té doby, než budou data z paměti k dispozici
- **Out-of-order** – instrukce jsou prováděny v pořadí tak, aby se pokud možno vykonalo co nejvíce nezávislých instrukcí naráz
- Pentium PRO se aktivně snaží vyhledávat trojice nezávislých mikroinstrukcí, které by se daly vykonat paralelně
- Takto může Pentium Pro **odložit** provádění až čtyř instrukcí a místo nich provádět libovolné další instrukce, které právě lze provést
- Podobně procesor postupuje i v případě **závislosti** instrukcí (instrukce potřebuje výsledek předchozí instrukce), kdy provádí instrukce mimo pořadí, dokud nedojde k výpočtu požadovaného operandu
- V případě **větvení** lze spekulativně provádět nezávislé instrukce obou větví (jako by podmínka, dle které se program rozvětví byla splněna i nesplněna) a po zjištění výsledku podmínky skoku zrušit výpočty neplatné větve



Out-of-order CISC-RISC

Původní program (CISC)

Instrukce 1
Instrukce 2
Instrukce 3
Instrukce 4



Mikroprogram (RISC)

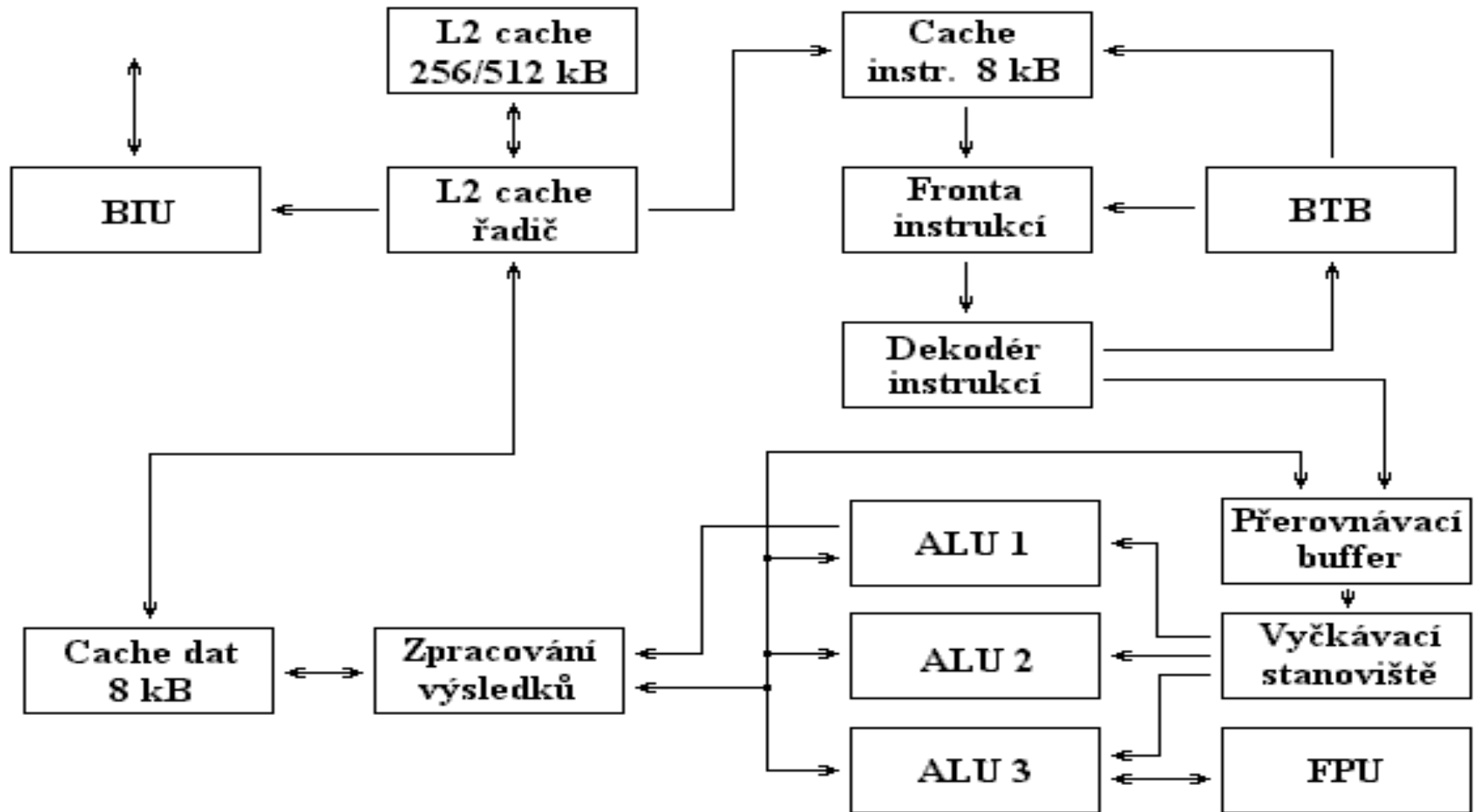
Mikro-operace 1a
Mikro-operace 1b
Mikro-operace 2a
Mikro-operace 2b
Mikro-operace 2c
Mikro-operace 3a
Mikro-operace 3b
Mikro-operace 4



Out-of-order superskalární vykonání RISC mikroprogramu

Mikro-operace 1a	Mikro-operace 1b	Mikro-operace 4
Mikro-operace 2a	Mikro-operace 3a	Mikro-operace 3b
Mikro-operace 2b	Mikro-operace 2c	

Mikroprocesor se snaží v každém taktu najít a začít zpracovávat tři nezávislé mikrooperace mimo pořadí, v jakém byly vygenerovány při převodu na mikroprogram



Pentium PRO



- Pentium PRO je dosti odlišné od původního Pentia
- V podstatě lze říct, že mezi Pentiem PRO a původním Pentiem je stejný rozdíl jako mezi Pentiem a 80486

Pentium MMX



- V roce **1997** uvádí firma Intel na trh další zdokonalenou verzi procesoru Intel **Pentium MMX**
- Pentium PRO je doplněno o skupinu instrukcí **MMX** – viz další prezentace
- MMX = **M**ulti**M**edia **eX**tension
- Tyto instrukce jsou určeny především pro podporu multimediálních aplikací, mohou však být použity pro libovolné jiné aplikace
- Instrukce MMX využívají osmi **64bitových MMX registrů (MM0, MM1.. MM7)**
- Technologie MMX využívá techniky **SIMD** (Single Instruction Multiple Data), která dovoluje jednou instrukcí manipulovat s více daty naráz (například jednou instrukcí dát povel ke dvěma nezávislým součtům, které proběhnou paralelně)
- Možnosti MMX technologie využívají především aplikace
 - 2D / 3D grafika
 - zpracování zvuku
 - zpracování videa
 - komprese dat



Kontrolní otázky

- Co je to superskalární procesor ?
- Jak funguje BTB ?
- Popište cache mikroprocesoru Pentium
- Jakým způsobem docházelo ke zvyšování výkonu mikroprocesorů 8086-80486 a jakými novými způsoby dochází ke zvyšování výkonu modernějších procesorů Pentium – Pentium IV ?
- První mikroprocesory Pentium se nechvalně proslavily jednou závažnou chybou. Pokuste se vyhledat, o jaký problém se jednalo
- V jakém rozmezí se pohybovaly taktovací frekvence mikroprocesorů Pentium 1?
- Proč bylo výhodné snížit napájecí napětí mikroprocesoru Pentium ?
- Co jsou to micro-ops ?
- Jak funguje CISC->RISC architektura mikroprocesoru Pentium Pro ?
- Kolik instrukcí v jednom taktu může dokončit Pentium ?
- Kolik instrukcí v jednom taktu může dokončit Pentium Pro ?
- Co je to FPU
- Jaký je rozdíl mezi L1 a L2 cache?