

Provokující otázka? Samozřejmě. Snaha být lepší, nebo alespoň mít něco lepšího než ti druzí, je lidem zřejmě zakódována v genech. Pokud přijmeme výzvu a na tuto hru přistoupíme, měli bychom se také ptát po nějakých pravidlech. Jinými slovy řečeno, musíme otázku z nadpisu doplnit a ptát se, jak se pozná dobrý počítač pro to, co právě potřebujeme. Tím se ale vše komplikuje...

Jak se pozná dobrý počítač?

Není možné jednoduše říci, že některý počítač je lepší nežli jiný. Podobně jako některé auto lépe zabírá do kopce a jiné je vynikající na rovině, je některý počítač lepší pro zpracování velkých objemů dat a jiný může být zase vynikající počtář. Podívejme se, zda je možné vše nějak kvantifikovat a tím alespoň částečně, zejména z důvodu srovnávání, objektivizovat.

Úlevou při tomto hledání nám budiž fakt, že se všechny počítače skládají z vesměs podobných dílů, a že jejich srovnávání je tedy podstatně jednodušší. Můžeme srovnávat výkon procesoru, dobu přístupu procesoru do paměti, vybavovací dobu pevného disku aj. Jak spolu jednotlivé komponenty ladí, to už tak snadno posuzovatelné není. Sebelepší procesor pracující na sběrnici, na které se tlačí pět dalších sebelepších procesorů, funguje spíš jako ústřední topení než jako výpočetní prvek.

Komponenty mající zásadní vliv na výkon celého systému jsou zejména procesor, vyrovnávací paměti procesoru, druh hlavní paměti a způsob jejího připojení k procesoru a dalším částem počítače, dále pak výkon vstupně-výstupních systémů a v neposlední řadě použitý program. Zvláště ten má zcela zásadní význam pro určení rychlosti systému, neboť nevhodně napsaný program může například téměř eliminovat výhodu vyrovnávacích pamětí.

Teoretické hodnoty aneb Data pro reklamní kampaň

Teoretické hodnoty výkonu (výkonem se rozumí počet operací, které je počítač či procesor schopen provést za sekundu) se udávají zejména pro procesory, a to celkem oprávněně. Procesor je zřejmě nejdůležitějším prvkem, který podstatným způsobem určuje výkon celého systému.

Teoretický výkon (*peak performance*) procesoru se určí tak, že hodnota hodinové frekvence (např. pro 500MHz Pentium 500 milionů) se vynásobí počtem instrukcí, které je procesor schopen provádět v jednom hodinovém cyklu. Tak se jednoduše dojde k závratným číslům, například k miliardě operací za sekundu (1 GFLOPS). To, že je procesor schopen vykonat více operací v jednom cyklu, je dáno tzv. zřetěžením některých specifických posloupností instrukcí.

Chceme-li znát teoretický výkon celého systému, je nutné toto číslo vynásobit počtem procesorů v počítači. Výsledná čísla jsou však skutečně pouze teoretická. Špičkových výsledků totiž počítač může dosáhnout jen v naprosto výjimečných případech. Skutečný výkon, a to jak procesoru, tak celého

počítače, se obvykle pohybuje někde na 30 až 50 % těchto hodnot. V případě paralelních počítačů tomu bývá ještě méně, asi 10 %.

Vyrovnávací paměti

Zcela zásadním problémem je v současné době vybavovací doba pamětí, která je neúměrně dlouhá ve srovnání s rychlostí procesoru. Zatímco procesor je schopen pracovat velice rychle, paměť není schopna tak rychle dodávat instrukce a především data. Existují sice rychlé paměti, ale ty jsou dosti drahé, a tak se tato situace řeší kompromisem. Mezi procesor a paměť se dá rychlá, tzv. vyrovnávací paměť (*cache memory*). Její funkce je v podstatě jednoduchá. Když se čtou data z paměti, s největší pravděpodobností se brzy budou číst data, která leží těsně vedle. Při čtení se tedy přenesou do vyrovnávací paměti celá *stránka*, nejen požadované místo, ale ještě jeho bezprostřední okolí. Při dalším čtení se pak většinou nemusí sahát do pomalé hlavní paměti a data jsou k dispozici velice rychle. Navíc údaj, který se do vyrovnávací paměti jednou uložil, bude s velkou pravděpodobností za okamžik použit znovu, a tak se tam nějakou chvíli drží.

Problém je, jsou-li data roztroušena po paměti tak, že se výhod vyrovnávací paměti nevyužije. Dojde k tzv. zanesení paměti (*cache trashing*) a k celkovému snížení výkonu. Podobný jev způsobuje přenášení velkých souvislých objemů dat mezi paměti a procesorem. Stránky z vyrovnávací paměti se pak navzájem vystrkují a tím se snižuje efektivita této techniky. Obecně však vyrovnávací paměti velmi výrazně zvyšují výkon systému.

Je jasné, že reálný výkon procesoru závisí zejména na struktuře úlohy, která je řešena, a na tom, jak dobře je využito vyrovnávacích pamětí. Reálný výkon je samozřejmě menší, v nejlepším případě rovný teoretickému výkonu procesoru.

Klasické benchmarky

V reálných úlohách zřídka dochází k tomu, že jsou procesoru pravidelně a nepřetržitě dodávány instrukce, které může zpracovávat tak, aby dosahoval teoretického výkonu. Obvykle dochází k nějakým prodlevám a celkový výkon je menší. Z toho také plyne otázka, jak reálný výkon změřit.

Nejjednodušší metodou je spustit na počítači nějakou typickou úlohu a změřit dobu jejího běhu. Typická úloha se však hledá obtížně, a tak se pro tyto účely vytvářejí "umělé typické úlohy" – benchmarky. Benchmark je tedy syntetická úloha, která reprezentuje typické úlohy z reality. Získá se obvykle tak, že se statisticky změří, v jakém pořadí a jak často které instrukce přicházejí do procesoru, a vytvoří se umělá úloha, která tuto statistiku dodržuje.

U procesoru se uvádějí dva důležité údaje. Počet operací s celými čísly za sekundu, jednotkou je IPS – počet instrukcí za sekundu. Nejčastěji se setkáme s veličinou MIPS, která reprezentuje milion instrukcí za sekundu. Druhým důležitým údajem je počet operací s reálnými čísly za sekundu. Zde je jednotkou FLOPS.

Pro oba tyto údaje jsou vyvinuty různé úlohy. Pro MIPS se již dlouhou dobu používá celkem dobře zavedený test, který se jmenuje DHRYSTONE podle anglického města, kde byl vyvinut. V současné době je k dispozici na internetu verze 2.1 a je volně ke stažení ve zdrojovém kódu. A zde je nutné upozornit na jednu závažnou vlastnost všech benchmarků.

Každý překladač, který k překladu programu použijeme, má k dispozici tzv. optimalizace. Překladač se podívá na některé konstrukce a pokusí se je připravit procesoru tak, aby je provedl pokud možno co nejrychleji. Optimalizace můžeme podle potřeby zapnout či vypnout, případně nastavit jejich určitou úroveň. Proto se při měření výkonu obvykle udávají údaje dva – počty operací pro neoptimalizovanou a optimalizovanou verzi. Překladače však bývají vybaveny stále sofistikovanějšími konstrukcemi, a tak například pro běh konkrétního programu jsem s různými stupni optimalizace dosáhl dokonce dvojnásobného výkonu. Z tohoto i z mnoha dalších důvodů se benchmark DHRYSTONE používá stále méně.

V souvislosti s optimalizací je vhodné uvést ještě jeden fakt. Pro některé úlohy se uvádí speciální benchmark, který se jmenuje TPP (*toward peak performance*). Není to nic mimořádného, ve skutečnosti jde o klasický benchmark, který je optimalizován člověkem. Po automatickém překladu si k němu sedl nějaký programátor, podíval se do přeloženého kódu a optimalizoval jej ještě více. Výkon dosažený pomocí TPP bývá až dvakrát vyšší nežli výkon dosažený nejlepší automatickou optimalizací.

Pro získání údaje o výkonu v operacích s reálnými čísly (FLOPS) se již od roku 1979 používá benchmark LINPACK, který reprezentuje klasické úlohy z lineární algebry. Tento program je rovněž volně ke stažení z internetu a pro jeho optimalizace platí zhruba totéž, co bylo řečeno o programu DHRYSTONE.

SPEC

Protože u benchmarků nejde o nic jiného než o peníze, bylo silnou snahou vše nějak objektivizovat. Různé firmy dosahovaly u svých počítačů závratných výsledků a leckteré jednotky, zejména grafické karty, byly dokonce konstruovány tak, aby co nejlépe vyhověly známým testům. Z těchto důvodů vzniklo konsorcium zvané SPEC – *Standard Performance Evaluation Corporation* (www.spec.org), které sdružuje snad všechny důležité firmy, jež nějak souvisejí s počítači. SPEC má tři důležité skupiny:

- + *SPEC High Performance Group (SPEC Hpg)*, která se zabývá především měřením výkonu superpočítačů, se zvláštním důrazem na paralelní počítače, které se hodnotí obzvlášť obtížně;

- + *SPEC Graphics Performance Characterization Group (SPEC Gpc)* pro testování výkonu v počítačové grafice;

- + *SPEC Open Systems Group (SPEC Osg)*, v podstatě původní organizace – dvě výše uvedené jsou nově vzniklé, lépe řečeno odštěpené.

SPEC Osg

Je nejdůležitější skupinou SPEC a její jednotlivé podskupiny poskytují následující testy:

- + *SPEC Cpu*, skupina testů, která obsahuje testy SPECint95 a SPECfp95 pro měření MIPS a FLOPS. Jde o nejlepší známé a nejlépe zavedené testy, připravuje se balík SPEC2000.

- + *SPEC jvm*, test pro měření kvality JAVA Virtual Machine.

- + *SPEC sdm* pro testování kvality implementace příkazů Unixu.

- + SPECsfs pro měření kvality systémů souborů.
- + SPECweb, poměrně nový test, který porovnává kvalitu WWW serverů.

Všechny testy jsou k dispozici za peníze. Ceny nejsou nijak závratné, SPECjvm98 je za 100, SPECsfs97 za 900, SPECweb99 za 800, SPECcpu95 za 600, SPECsdm za 1450 a SPECchpc96 za 1200 USD. Pokud vás zajímají výsledky, jsou k dispozici na stránkách www.spec.org a jsou průběžně aktualizovány. Jednou za čtvrt roku vychází přehledný časopis, kde jsou veškeré výsledky uvedeny. Internetová verze však je daleko lepší, už kvůli možnosti automatického vyhledávání. SPEC zásadně nedává doporučení ani nijak výsledky nekomentuje. Čísla však obvykle hovoří sama za sebe.

SPECgpc

Je poměrně nová, ale o to aktivnější skupina ve sdružení SPEC. Zabývá se měřením rychlosti počítačové grafiky a dělí se na tři skupiny.

SPECmedia je zcela nová skupina, která se zaměřuje především na kvalitu implementace MPEG animací, zejména pro jejich přítomnost v DVD. Další aktivity směřují k měření kvality trojrozměrného zvuku a grafiky orientované na hry a zábavu. Očekává se, že první testy budou k dispozici někdy v roce 2000.

SPECapc se zabývá aplikacemi pro trojrozměrné modelování a zde jsou k dispozici dva testy – jeden pro SolidWorks98 a druhý pro Pro/ENGINEER verze 20. Oba testy jsou za peníze a výsledky jednotlivých počítačových sestav jsou k dispozici zadarmo na již zmíněné webové stránce www.spec.org.

SPECopc pro měření rychlosti implementace OpenGL, standardu pro rychlou trojrozměrnou grafiku. Testy jsou k dispozici dva: *SPECviewperf* a *SPECglperf*.

SPECviewperf je orientován na poměrně vysokou úroveň OpenGL a koresponduje s komerčními produkty, jejichž nejčastěji používané grafické operace měří. Například test AWadv odpovídá Alias|Wavefrontu, ProCDRS programu Pro/Designer aj. Příklady výsledků jsou na uvedených obrázcích, více je na Chip CD 12/98 a aktuální výsledky jsou na www.spec.org/gpc/opc.static/viewresults.htm.

Druhý, *SPECglperf*, měří kvalitu implementace OpenGL na podstatně nižší úrovni, v podstatě na úrovni hardwaru. Zahrnuje třináct měření, například rychlost osvětlování, texturování, rychlost generování vyplněných trojúhelníků aj.

Závěr

Zdá se, že situace na poli běžných počítačů, tedy těch, které jsou nasazovány v aplikacích od domácností až po velké podniky, je celkem přehledná. Existuje konsorcium, které vyvíjí důkladné a kvalitní benchmarky, počítače jsou aktivně testovány a je jen na uživatelích, aby se před investicí do počítače podívali na stránky SPEC. Testy jsou celkem objektivní a jejich výsledky, vzhledem k tomu, kdo všechno je v konsorciu zastoupen, jsou zřejmě nezmanipulované. Konkurenční firmy se totiž navzájem hlídají samy. Bohužel s výsledky se v reklamních materiálech setkáme jen tehdy, je-li příslušný počítač v nějakém testu na prvním místě. Možná právě proto nejsou testy moc známé.

Výhodou těchto testů je, že nemusíme za nic platit. Vše platí ti, kdo si nechávají své počítače změřit. Stále však platí pravidlo, že pokud máme nějakou nestandardní úlohu, je dobré, pokud to lze, provést vlastní měření. Pro běžné záležitosti, jako je WWW server, koupě počítače, na kterém se bude pracovat v programu Pro/ENGINEER, či koupě fileserveru, stačí pouze otevřít příslušné webové stránky a podívat se, jak se věci mají.

Bedřich Beneš