KAPITOLA 1

Pamäťový model

Pri časovej analýze algoritmov sa zvyčajne používa takzvaný *RAM model* (skratka z anglického *Random-Access Machine*, stroj s náhodným prístupom k pamäti), v ktorom sa predpokladá možnosť pristupovať k ľubovoľnému úseku pamäte v konštantnom čase. To znamená, že vo výslednej asymptotickej analýze počítame len počet vykonaných operácií.

V skutočnosti však moderné počítače využívajú niekoľko úrovňovú pamäťovú hierarchiu [7]. Tá sa typicky skladá z registrov a troch úrovní *cache* (vyrovnávacej pamäte) priamo na procesore, následne z hlavnej operačnej pamäte a disku. V tomto poradí referencie sú tieto úrovne zoradené od najrýchlejšej a najmenšej (odozva rádovo 1 ns, kapacita 64 KiB) až po najpomalšiu ale najväčšiu (odozva od 100 µs po 10 ms podľa typu[1], kapacita rádovo 1 TiB). To znamená, že rozdiel v prístupovej dobe je 10 000 000-násobný a len ťažko môžeme hovoriť o konštantnom čase. Približné hodnoty pre všetky úrovne sú v tabuľke 1.1.

ram blok - dobre?

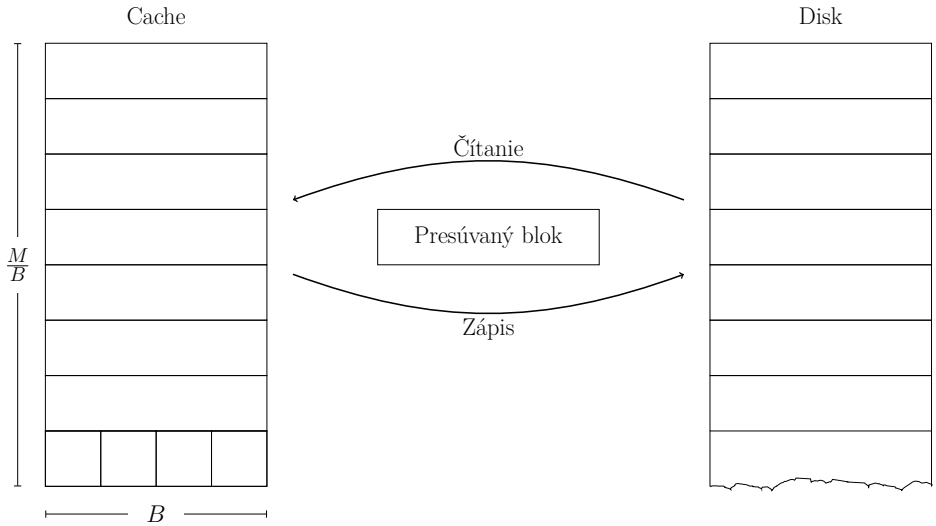
registre? tazko zratat velkost

Tabuľka 1.1: Približné parametre rôznych úrovní pamäte. Hodnoty troch úrovní *cache* na procesore (L1, L2 a L3) sú pre mikroarchitektúru Intel Haswell. [9, 10]

Úroveň	Veľkosť	Odozva		Asociativita	Veľkosť bloku
L1	64 KiB[1]	4clk[2]	≈ 1 ns	8	64 B
L2	256 KiB[1]	11clk[2]	≈ 4 ns	8	64 B
L3	2–20 MiB	36clk[2]	≈ 12 ns		64 B
RAM	≈ 8 GiB		≈ 100 ns		16 B
Disk	≈ 1 TiB		≈ 0.1–10 ms		4 KiB–2 MiB

[1] Hodnota pre jedno jadro procesora.
[2] Počet cyklov procesora, uvedené časové aproximácie pri 3 GHz.

[1]Klasické pevné disky (HDD) alebo disky bez pohyblivých častí (SSD)



Obrázok 1.1: External-memory model

Dôsledkom tejto hierarchie je závislosť výslednej rýchlosti algoritmu od jeho prístupov k pamäti. Operácie, ktoré využívajú dáta uložene na disku potrvajú dlhšie ako tie, ktoré využívajú iba dáta v registroch. Dôvodom je, že pri prístupe k dátam na disku sa v skutočnosti sa tieto dáta postupne presunú z disku do hlavnej pamäte, do *cache* na procesore a napokon do registrov. Následne sa môže požadovaná operácia vykonať rovnako rýchlo ako v druhom prípade (keď už je v registroch), avšak nejaký čas bol algoritmus nečinný a čakal na presun dát. Pre všetky susedné dvojice pamäťo- rewrite vých úrovní teda slúži tá menšia a rýchlejšia ako vyrovnávacia pamäť pre tú väčšiu a pomalšiu.

1.1 External-memory model

Spôsobom ako zohľadniť tieto skutočnosti pri analýze algoritmov je takzvaný *external-memory model* (model externej pamäte), nazývaný tiež *I/O model* alebo *cache-aware model* [1]. Ten popisuje pamäť skladajúcu sa z dvoch častí (obrázok 1.1), ktoré voláme *cache* a *disk*.

Všetky výpočty prebiehajú nad dátami v *cache*, ktorá má obmedzenú veľkosť. Ostatné dáta sú uložené na disku neobmedzenej veľkosti, no nemôžeme s nimi priamo manipulovať a je ich potrebné najskôr preniesť do *cache*. V samotnej analýze algoritmov potom počítame počet týchto prenosov z disku do *cache* a naopak.

Prístup k dátam, ktoré sa práve nachádzajú v *cache* voláme *cache hit* (zásah *cache*). Naopak prístup, ktorý vyžaduje dáta najskôr presunúť z *disk*-u do *cache* voláme *cache miss* (minutie *cache*). Tieto presuny sú realizované v *blokoch* pamäte veľkosti B. Disk aj *cache* sa skladajú z takýchto blokov a za jednu operáciu považujeme presun jedného bloku medzi nimi. *Cache* má obmedzenú veľkosť M a skladá sa teda z $\frac{M}{B}$ blokov.

1.1.1 Cache-aware algoritmy

Pokiaľ poznáme parametre B a M, môžeme skonštruovať algoritmus, ktorý bude túto dvojicu pamätí využívať efektívne. Takýto algoritmus voláme *cache-aware* (uvedomujúci si *cache*). Súčasťou tohto algoritmu by bolo explicitné spravovanie presunov pamäte - je potrebné riešiť čítanie blokov z disku a ich umiestňovanie do *cache*, nahrádzanie blokov v *cache* pri zaplnení a spätný zápis blokov na disk.

Tento model popisuje len dve úrovne pamäte a teda funguje efektívne len pre danú susednú dvojicu, pre ktorú ho na základe znalosti parametrov optimalizujeme. V moderných systémoch ale máme takýchto dvojíc niekoľko. Keby sme poznali parametre pre všetky tieto dvojice, môžeme zovšeobecniť tento model pre viac úrovní a explicitne riešiť presun blokov medzi nimi. Takto sa však samotná správa pamäte potenciálne stáva komplikovanejším problémom ako pôvodný algoritmus.

1.2 Cache-oblivious model

Riešením týchto problémov je *cache-oblivious model* (na *cache* nedbajúci), v ktorom uvažujeme rovnakú dvoj-úrovňovú pamäť zloženú z disku a *cache* [8, 15]. Na rozdiel od *cache-aware* modelu však algoritmus nepozná parametre B a M. Pokiaľ sa nám napriek tomu podarí navrhnúť algoritmus, ktorý vykonáva (asymptoticky) rovnaký počet pamäťových presunov ako *cache-aware* algoritmus, bude bežať efektívne pre ľubovolné takéto parametre.

Takéto algoritmy majú na rozdiel od *cache-aware* algoritmov v *external-memory* modely mnohé výhody. Samotná implementácia algoritmu nemôže explicitne riešiť presun blokov pokiaľ nepozná veľkosť bloku ani koľko blokov môže do *cache* uložiť. Táto úloha zostane ponechaná na nižšiu vrstvu (operačný systém resp. hardware v prípade *cache* na procesore) - algoritmus bude pristupovať k pamäti priamo bez ohľadu na to, či sa nachádza v *cache* alebo nie a v prípade potreby prebehnú nutné prenosy na nižšej úrovni (z pohľadu algoritmu) automaticky.

Ďalšou výhodou je automatická optimalizácia pre dané parametre. V prípade *cache-aware* algoritmov môže byť problémom získať presné hodnoty týchto parametrov a potrebné pri ich zmene upraviť algoritmus. Vývoj algoritmu, ktorý bude fungovať na rozličných architektúrach, môže byť problematický.

V neposlednom rade bude takýto *cache-oblivious* algoritmus efektívny medzi každou dvojicou susedných úrovní. Vzhľadom na to, že hodnoty parametrov nepozná, musí pre ľubovolnú takú dvojicu pracovať rovnako efektívne.

1.2.1 Správa pamäte

V momente, keď sa *cache-oblivious* algoritmus pokúsi o vykonanie operácie, ktorá potrebuje dáta mimo cache je potrebné ich najskôr z disku skopírovať. V prípade, že je v cache voľný blok, je možné presunúť dáta bez nutnosti nahradenia. V opačnom prípade je však potrebné uvoľniť miesto tým, že sa vyberie blok z cache, v prípade, že bol upravený sa jeho obsah zapíše späť na disk a následne sa požadovaný blok z disku zapíše na jeho miesto v cache. Tento proces sa nazýva výmena stránok (*page replacement*), a algoritmus rozhodujúci, ktorý blok z cache odstrániť, voláme stratégia výmeny stránok (*page-replacement strategy*).

Ak by sa táto stratégia správala tak, že vždy odstráni blok, ktorý bude potrebný v ďalšom kroku algoritmu, tak by bolo zakaždým presúvať bloky medzi diskom a cache. To by znamenalo, že počet blokov, s ktorými v cache môžeme pracovať je $M/B = 1$. Ďalším problémom je *asociatívnosť* cache - z praktických dôvodov je často možné daný blok z disku uložiť len na niekoľko pozícií v cache. Inak by bolo potrebné ukladať spolu s každým blokom jeho plnú adresu na disku, čo by redukovalo celkový počet blokov, ktoré sa do cache zmestia. Znížením asociativity je možné ukladať iba časť adresy, pričom zvyšok je implicitne určený pozíciou v cache. V prípade nízkej asociativity však môžu opäť nastať situácie, kedy je algoritmus schopný využiť iba malý počet blokov v cache.

Tieto problémy *cache-oblivious* model obchádza predpokladom ideálnej cache - cache, ktorá je plne asociatívna (každý blok disku je možné uložiť v každom bloku cache) a používa optimálnu stratégiu výmeny stránok, ktorá vždy odstráni blok, ktorý bude potrebný najneskôr. Prvý predpoklad je síce v reálnych systémoch nepraktický, no z teoretického hľadiska je v poriadku. Druhý predpoklad je však nerealizovateľný, keďže by stratégia výmeny stránok musela predpovedať budúce kroky algoritmu. Nasledovné lemy však ukazujú, že bez týchto predpokladov na reálnom systéme s nízkou asociativitou a jednoduchou stratégiou výmeny stránok sa algoritmus zhorší len o konštantný faktor.

Lema 1.2.1. *Algoritmus, ktorý v ideálnej cache veľkosti M s blokmi veľkosti B vykoná T pamäťových operácií, vykoná najviac $2T$ pamäťových operácií v cache veľkosti $2M$ s blokmi veľkosti B pri použití stratégie* LRU *alebo* FIFO. *[8, Lemma 12]*

Lema 1.2.2. *Plne asociatívna cache veľkosti M sa dá simulovať s použitím $\mathcal{O}(M)$ pamäte tak, že prístup ku každému bloku v cache zaberie v priemernom prípade $\mathcal{O}(1)$ času. [8, Lemma 16]*

Stratégia *LRU* (least recently used) vyberá vždy blok, ktorý bol najdlhšie nepoužitý. Implementácia vyžaduje udržiavať si ku každému bloku počítadlo, ktoré sa pri prístupe nastaví na nulu a pri prístupe k iným blokom zvýši o jedna. Pri potrebe uvoľniť miesto

move abo- ve,rewrit

v cache vyberieme blok s najväčšou hodnotou počítadla - ten, ku ktorému najdlhšie nebol prístup.

Stratégia *FIFO* (first in, first out) je ešte jednoduchšia - bloky sú udržiavame zoradené podla poradia ich vloženia do cache. Keď vyberáme blok na odstránenie tak vezmeme ten, ktorý bol pridaný najskôr a teda je na začiatku tohto usporiadania.

Literatúra

[1] Aggarwal, Alok ; Vitter, Jeffrey u. a.: The input/output complexity of sorting and related problems. In: *Communications of the ACM* 31 (1988), Nr. 9, S. 1116–1127

[2] Bayer, Rudolf: Binary B-trees for Virtual Memory. In: *Proceedings of the 1971 ACM SIGFIDET (Now SIGMOD) Workshop on Data Description, Access and Control*. New York, NY, USA : ACM, 1971 (SIGFIDET '71), 219–235

[3] Bender, Michael A. ; Demaine, Erik D. ; Farach-Colton, Martin: Cache-Oblivious B-Trees. In: *Proceedings of the 41st Annual Symposium on Foundations of Computer Science (FOCS 2000)*. Redondo Beach, California, November 12–14 2000, S. 399–409

[4] Bender, Michael A. ; Demaine, Erik D. ; Farach-Colton, Martin: Cache-Oblivious B-Trees. In: *SIAM Journal on Computing* 35 (2005), Nr. 2, S. 341–358

[5] Bender, Michael A. ; Duan, Ziyang ; Iacono, John ; Wu, Jing: A locality-preserving cache-oblivious dynamic dictionary. In: *Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms* Society for Industrial and Applied Mathematics, 2002, S. 29–38

[6] Demaine, Erik D.: Cache-Oblivious Algorithms and Data Structures. In: *Lecture Notes from the EEF Summer School on Massive Data Sets*. BRICS, University of Aarhus, Denmark, June 27–July 1 2002

[7] Drepper, Ulrich: What every programmer should know about memory. In: *Red Hat, Inc* 11 (2007)

[8] Frigo, Matteo ; Leiserson, Charles E. ; Prokop, Harald ; Ramachandran, Sridhar: Cache-oblivious algorithms. In: *Foundations of Computer Science, 1999. 40th Annual Symposium on* IEEE, 1999, S. 285–297

[9] Intel Corporation: *Intel® 64 and IA-32 Architectures Optimization Reference Manual*. 2014 (248966-029)

[10] Intel Corporation: *Intel® 64 and IA-32 Architectures Software Developer's Manual*. 2014 (325462-050US)

[11] Kasheff, Zardosht: *Cache-oblivious dynamic search trees*, Massachusetts Institute of Technology, Diss., 2004

[12] Kotrlová, Katarína: *Vizualizácia háld a intervalových stromov*, Univerzita Komenského v Bratislave, bakalárska práca, 2012

[13] Kováč, Jakub: *Vyhľadávacie stromy a ich vizualizácia*, Univerzita Komenského v Bratislave, bakalárska práca, 2007

[14] Lukča, Pavol: *Perzistentné dátové štruktúry a ich vizualizácia*, Univerzita Komenského v Bratislave, bakalárska práca, 2013

[15] Prokop, Harald: *Cache-oblivious algorithms*, Massachusetts Institute of Technology, Diss., 1999

[16] Tomkovič, Viktor: *Vizualizácia stromových dátových štruktúr*, Univerzita Komenského v Bratislave, bakalárska práca, 2012