

Кеш памет. Виртуална памет

Публикувано от legolas на 13.02.2010, 19.04

Кеш памет. Виртуална памет

1. Кеш памет

Основната памет на компютърните системи (RAM) се реализира с относително бавна динамична памет (DRAM). Обръщението към тази памет обикновено води до престой на микропроцесора за изчакване на зареждането на информацията от паметта. Статичната памет (SRAM) е изпълнена от тригерни елементи и може да работи с честотата на микропроцесора. Тя не се използва като масова оперативна памет поради високата си цена. Освен това, данните от основната памет се трансформират до микропроцесора посредством системната шина, чиято честота е доста по-ниска от честота с която работи микропроцесора. Разумен компромис за изграждане на икономични и производителни системи е използването на йерархичен способ на организация на паметта. Идеята е да се съчетае голям обем DRAM памет с относително неголяма, но бърза SRAM памет, работеща с честотата на микропроцесора. Такава памет се нарича кеш памет.

Значението на думата 'кеш' е 'таен склад'. В компютърните системи този склад е 'прозрачен' – за програмите той не представлява допълнителна адресируема памет. Кеш паметта се явява допълнителен бързодействащ склад за адресируемата памет в компютърната система.

Вътрешната кеш памет представлява бърза буферна памет, която обикновено се вгражда в схемата на микропроцесора. В процеса на работа отделни блокове информация (данни или програмен код) се копират от основната памет (RAM) в кеш паметта. Тази процедура на копиране се нарича кеширане. В съответствие с алгоритъма на кеширане определени блокове от паметта се копират в кеш паметта и заместват други блокове, които вече са били използвани от микропроцесора.

Когато микропроцесорът се обръща за прочитане на команда или данни, специален кеш контролер проверява дали те не се намират в кеш паметта. Ако необходимата информация е там, тя се извлича много бързо, тъй като процеса се извършва с тактовата честота на микропроцесора. Такъв случай на обръщение към паметта се нарича кеш попадение (cache hit). Ако необходимата информация не се намира в кеш паметта тя се зарежда от основната памет (RAM) и едновременно

се записва в кеш паметта. Такъв случай се нарича кеш пропуск (cache miss).

Повишаването на бързодействието на компютърните системи се получава, когато кеш попадения се реализират много по-често отколкото кеш пропуските. Висок процент на кеш попадения се получава, благодарение на това, че в по-голямата част от случаите обръщенията към клетки от паметта се извършва в области, които са използвани преди това.

В съвременните компютри кеш паметта обикновено се изпълнява на две нива (в най-новите микропроцесори и на три нива). Първичният кеш (L1 cache) се нарича вътрешен кеш и се вгражда в схемата на микропроцесора. Неговият обем е малък (8 – 32 kB). В най-новите модели микропроцесори L1 cache е 64 kB и повече. За повишаване на производителността на микропроцесорите, за данни и програмен код понякога се използва отделна кеш памет - **Харвардска архитектура**. Когато първичната кеш памет е обща и за данни и програмен код микропроцесорът е с **Принстънска архитектура**.

Вторичната кеш памет (L2 cache) се разполага или в кристала (чипа) на микропроцесора или отделно от него на обща платка (Pentium P6). Тази памет се свързва с микропроцесора посредством вътрешна шина на микропроцесора (back side bus) – задна шина. Капацитетът на тази памет е значително по-голям – 256 kB, 512 kB или 1 MB.

Кеш контролерът обезпечава така наречената кохерентност (съгласуваност) на данните от кеш паметта и данните в основната памет. Това е особено важно, тъй като към тези данни може да се обръща не само микропроцесорът но и други активни устройства (адаптери) включени към различните шини в компютърната система. Трябва да се има пред вид и това, че микропроцесорите могат да бъдат няколко и всеки да има собствена кеш памет.

Контролерът на кеш паметта оперира с *низове (cache line)* с определена дължина. Низовете съхраняват блок от паметта имащ дължина каквато е дължината на низа. С всеки низ е свързана информация за адреса на блока в основната памет, който се съхранява в низа.

Информацията за съхранявания блок от паметта в низа се нарича *tag (tag)* и е записана в специална памет (памет на таговете).

В зависимост от начина на установяване на взаимно съответствие между низовете и блоковете от паметта се различават три вида архитектура на кеш паметта:

- Ø кеш с пряко изображение (direct mapped cache);
- Ø напълно асоциативен кеш (fully associative cache) ;

Ø наборно асоциативен кеш (set-associative cache).

2. Виртуална памет

Едно от най-важните предимства на 386 е, че добавя нов режим на действие – виртуален режим 86. Той дава възможността на процесора да превключва повече от една задача в средата на DOS, всяка от които счита, че работи в своя собствена машина 8086.

Терминът „**виртуална памет**“ обикновено се свързва с **възможността да се адресира обем памет**, който значително **превишава размера на реалната (физическата) оперативна памет** на съответния компютър.

За първи път виртуална памет е реализирана през **1960 г.** в машината **Атлас (Atlas)**, създадена в **Манчестърския университет**.

Характерна черта на системите с виртуална памет е, че **генерираните от изпълняваната програма адреси не са длъжни да съвпадат с реално съществуващи адреси на ОП**.

При 32 разрядните микропроцесори адресното пространство е : **$2^{32} = 4 \text{ GB}$** , но благодарение на сегментирането на паметта логическия адрес се състои от 2 части – номер на сегмента + отместване в сегмента. При 80386 логическия адрес е с размер 46 бита т.е. адресното пространство, което може да се адресира е: **$N = 2^{46} = 64 \text{ TB}$** (терабайта). Това е извън пределите на оперативната памет, която е **16 MB**. За осигуряване на допълнителен обем памет се използват външни запаметяващи устройства. За външни запаметяващи устройства се използват магнитен диск(уинчестър, флопи диск).Възникналото противоречие между голямото адресно пространство и малката по обем оперативна памет се решава чрез виртуална памет. Тя осигурява апаратно програмно решаване на проблема, като обхваща в адресното пространство и паметта на външните запаметяващи устройства. Виртуалната памет автоматично прехвърля необходимите програмни данни от външни запаметяващи устройства в оперативната памет и обратно. Виртуалната памет е програмно апаратно средство, като апаратната част е вградена в МП, а програмната се разработва заедно със системното програмно осигуряване. При адресиране на несъществуващ участък от оперативната памет се генерира прекъсване, чрез което се активира програмната част, като се запазва състоянието на процесора и инструкцията, така, че след извършване на обмяна между памет и външни запаметяващи устройства да се повтори неосъществения цикъл на шината.

ИЗОБРАЗЯВАНЕ НА АДРЕСА

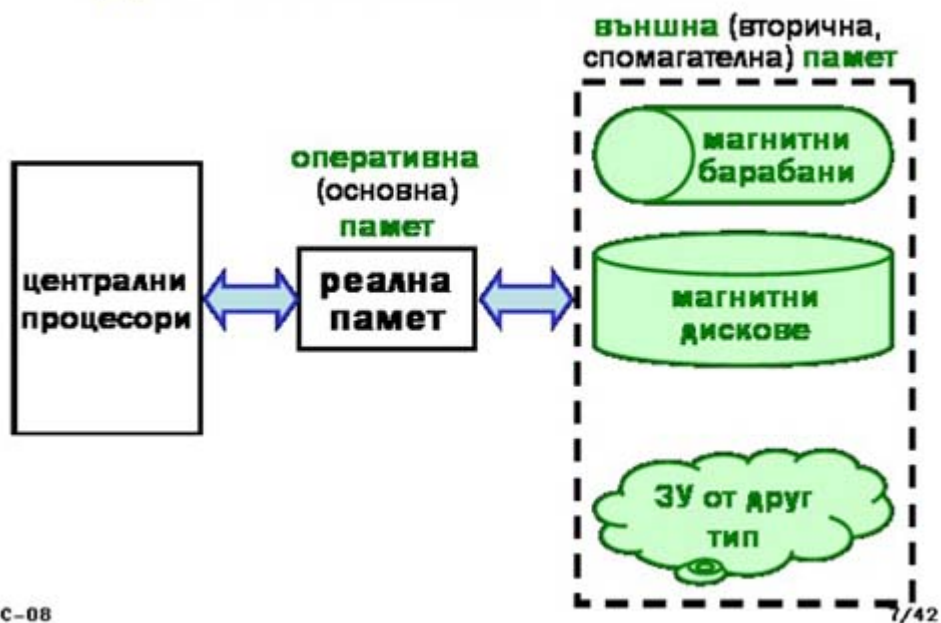
Разработени са **различни способности за установяване на съответствие** между виртуални и реални адреси.

Механизмът за динамично преобразуване (превод, трансляция) трансформира виртуалните адреси в реални (Dynamic Address Translation, DAT) по време на изпълнение на съответния процес.

Всички системи за преобразуване имат едно общо свойство: съседните виртуални адреси не са длъжни да бъдат съседи и в реалната памет.

Това освобождава потребителите от необходимостта да отчитат разположението на своите процедури и данни в R и дава възможност да пишат програми по най-естествения начин: внимание към алгоритъма и данните, вместо към апаратните особености.

ДВУСТЕПЕННА СХЕМА



МНОЖЕСТВО ПРОЦЕСИ



