

Виртуальная память

Материал из Википедии — свободной энциклопедии

Виртуáльная па́мять (англ. *virtual memory*) — метод управления памятью компьютера, позволяющий выполнять программы, требующие больше оперативной памяти, чем имеется в компьютере, путём автоматического перемещения частей программы между основной памятью и вторичным хранилищем (например, жёстким диском)^{[1][2][3]}. Для выполняющейся программы данный метод полностью прозрачен и не требует дополнительных усилий со стороны программиста, однако реализация этого метода требует как аппаратной поддержки, так и поддержки со стороны операционной системы.

В системе с виртуальной памятью используемые программами адреса, называемые виртуальными адресами, транслируются в физические адреса в памяти компьютера. Трансляцию виртуальных адресов в физические выполняет аппаратное обеспечение, называемое блоком управления памятью. Для программы основная память выглядит как доступное и непрерывное адресное пространство либо как набор непрерывных сегментов, вне зависимости от наличия у компьютера соответствующего объёма оперативной памяти. Управление виртуальными адресными пространствами, соотнесение физической и виртуальной памяти, а также перемещение фрагментов памяти между основным и вторичным хранилищами выполняет операционная система (см. подкачка страниц).

Применение виртуальной памяти позволяет:

- освободить программиста от необходимости вручную управлять загрузкой частей программы в память и согласовывать использование памяти с другими программами
- предоставлять программам больше памяти, чем физически установлено в системе
- в многозадачных системах изолировать выполняющиеся программы друг от друга путём назначения им непересекающихся адресных пространств (см. защита памяти)

В настоящее время виртуальная память аппаратно поддерживается в большинстве современных процессоров^[1]. В то же время в микроконтроллерах и в системах специального назначения, где требуется либо очень быстрая работа, либо есть ограничения на длительность отклика (системы реального времени), виртуальная память используется относительно редко. Также в таких системах реже встречается многозадачность и сложные иерархии памяти.

Содержание

История

Страничная организация виртуальной памяти

Сегментная организация виртуальной памяти

См. также

Примечания

Литература

История

В 1940-е и 1950-е годы все программы, размер которых превышал объём оперативной памяти, должны были содержать логику управления оперативной и внешней памятью, такую как оверлей. В связи с этим виртуальная память была введена не только для увеличения объёма оперативной памяти, но и для того, чтобы сделать такое увеличение наиболее простым в использовании для программистов.^[4] Для поддержки мультипрограммирования и многозадачности во многих ранних компьютерных системах разделение памяти между несколькими программами осуществлялось без виртуальной памяти, с помощью сегментации (например, в компьютерах PDP-10).

Понятие виртуальной памяти было впервые введено в 1956 году немецким физиком Фрицем-Рудольфом Гюнтшем (нем. *Fritz-Rudolf Güntsch*) из Берлинского технического университета в его докторской диссертации «Логическое проектирование цифровой вычислительной машины с несколькими асинхронными вращающимися барабанами и автоматическим высокоскоростным управлением памятью». В ней описывается машина с 6 блоками памяти на магнитных сердечниках ёмкостью 100 слов каждый и адресным пространством в 1000 блоков по 100 слов. Аппаратное обеспечение машины автоматически перемещает блоки между основной и внешней памятью на магнитных барабанах.^{[5][6]} Подкачка страниц впервые была реализована в Манчестерском университете как способ расширения памяти компьютера Atlas путём объединения его основной памяти на магнитных сердечниках объёмом 16000 слов с памятью на магнитных барабанах ёмкостью 96000 слов. Первая поставка машины Atlas состоялась в 1962 году, но работающие прототипы подкачки страниц были разработаны уже в 1959 году.^{[4][7][8]} В 1961 году Burroughs Corporation независимо выпустила первый промышленно производившийся компьютер с виртуальной памятью B5000, использующий сегментацию вместо подкачки страниц.^{[9][10]} В 1965 году виртуальная память была реализована в отечественной машине БЭСМ-6, хотя реализованный механизм не позволял расширить 15-разрядное адресное пространство процесса и ограничивал удобство программирования.^[11]

Прежде чем виртуальная память смогла быть использована в распространённых операционных системах, должен был быть решен ряд проблем. Динамическая трансляция адресов требовала дорогого и сложного в изготовлении специального оборудования. Первые реализации замедляли доступ к памяти.^[4] Существовали опасения, что новые общесистемные алгоритмы, использующие внешнюю память, будут менее эффективны, чем ранее использовавшиеся алгоритмы конкретных приложений. К 1969 году дебаты вокруг виртуальной памяти для промышленно выпускаемых компьютеров были завершены: команда исследователей из IBM под руководством Дэвида Сейра (англ. *David Sayre*) показала, что созданная ими система виртуальной памяти во всех отношениях превосходит лучшие системы, управляемые вручную.^{[4][12]} Первым мини-компьютером, в котором была использована виртуальная память, был норвежский Norsk Data (но.) Nord-1 (но.). В 1970-е годы реализации виртуальной памяти появились в других мини-компьютерах, наиболее известным является VAX, работающий под управлением операционной системы VMS.

Виртуальная память в архитектуре x86 была реализована с появлением защищенного режима процессора 80286, однако она использовала сегментацию памяти, и метод подкачки сегментов плохо масштабировался для больших размеров сегментов. В процессоре 80386 была введена поддержка подкачки страниц, не приводящая к возникновению двойной ошибки, если во время обработки другого исключения возникло исключение ошибки страницы. Поверх системы подкачки страниц работал существующий механизм сегментации памяти. Однако загрузка дескрипторов сегментов являлась дорогостоящей операцией, заставляя разработчиков операционных систем полагаться только на механизм подкачки страниц, а не на комбинацию подкачки страниц и сегментации.

Страничная организация виртуальной памяти

В большинстве современных операционных систем виртуальная память организуется с помощью страничной адресации. Оперативная память делится на страницы: области памяти фиксированной длины (например, 4096 байт^[13]), которые являются минимальной единицей выделяемой памяти (то есть даже запрос на 1 байт от приложения приведёт к выделению ему страницы памяти). Исполняемый процессором пользовательский поток обращается к памяти с помощью адреса виртуальной памяти, который делится на номер страницы и смещение внутри страницы. Процессор преобразует номер виртуальной страницы в адрес соответствующей ей физической страницы при помощи буфера ассоциативной трансляции (TLB). Если ему не удалось это сделать, то требуется дозаполнение буфера путём обращения к таблице страниц (так называемый *Page Walk*), что может сделать либо сам процессор, либо операционная система (в зависимости от архитектуры)^[14]. Если страница была выгружена из оперативной памяти, то операционная система подкачивает страницу с жёсткого диска в ходе обработки события *Page fault* (см. свопинг, подкачка страниц). При запросе на выделение памяти операционная система может «сбросить» на жёсткий диск страницы, к которым давно не было обращений. Критические данные (например, код запущенных и работающих программ, код и память ядра системы) обычно находятся в оперативной памяти (исключения существуют, однако они не касаются тех частей, которые отвечают за обработку аппаратных прерываний, работу с таблицей страниц и использование файла подкачки).

Сегментная организация виртуальной памяти

Механизм организации виртуальной памяти, при котором виртуальное пространство делится на части произвольного размера — сегменты. Этот механизм позволяет, к примеру, разбить данные процесса на логические блоки.^[15] Для каждого сегмента, как и для страницы, могут быть назначены права доступа к нему пользователя и его процессов. При загрузке процесса часть сегментов помещается в оперативную память (при этом для каждого из этих сегментов операционная система подыскивает подходящий участок свободной памяти), а часть сегментов размещается в дисковой памяти. Сегменты одной программы могут занимать в оперативной памяти несмежные участки. Во время загрузки система создает таблицу сегментов процесса (аналогичную таблице страниц), в которой для каждого сегмента указывается начальный физический адрес сегмента в оперативной памяти, размер сегмента, правила доступа, признак модификации, признак обращения к данному сегменту за последний интервал времени и некоторая другая информация. Если виртуальные адресные пространства нескольких процессов включают один и тот же сегмент, то в таблицах сегментов этих процессов делаются ссылки на один и тот же участок оперативной памяти, в который данный сегмент загружается в единственном экземпляре. Система с сегментной организацией функционирует аналогично системе со страничной организацией: время от времени происходят прерывания, связанные с отсутствием нужных сегментов в памяти, при необходимости освобождения памяти некоторые сегменты выгружаются, при каждом обращении к оперативной памяти выполняется преобразование виртуального адреса в физический. Кроме того, при обращении к памяти проверяется, разрешен ли доступ требуемого типа к данному сегменту.

Виртуальный адрес при сегментной организации памяти может быть представлен парой (g , s), где g — номер сегмента, а s — смещение в сегменте. Физический адрес получается путём сложения начального физического адреса сегмента, найденного в таблице сегментов по номеру g , и смещения s .



Недостатком данного метода распределения памяти является фрагментация на уровне сегментов и более медленное по сравнению со страничной организацией преобразование адреса.

Существует также гибридная странично-сегментная организация виртуальной памяти^[16].

См. также

- Разделяемая память
- Кэш-память
- Менеджер виртуальной памяти
- Блок управления памятью (Memory Management Unit)
- Утечка памяти

Примечания

1. Э. Танненбаум. Архитектура компьютера = Structured Computer Organization. — 5-е изд. — СПб.: Питер, 2013. — С. 476. — 884 с. — ISBN 978-5-469-01274-0.
2. Ч. Кэнс. Р. Стаффорд. VAX: Программирование на языке ассемблера и архитектура = VAX assembly language and architecture. — М.: Радио и связь, 1991. — С. 310. — 416 с. — ISBN 5-256-00706-8.
3. Морс С.П., Алберт Д.Д. Архитектура микропроцессора 80286 (https://archive.org/details/isbn_5256004662) = The 80286 architecture. — М.: Радио и связь, 1990. — С. 167 (https://archive.org/details/isbn_5256004662/page/n166). — 304 с. — ISBN 5-256-00466-2.
4. Peter J. Denning. Before Memory Was Virtual (<http://denninginstitute.com/pjd/PUBS/bvm.pdf>)  // In the Beginning: Recollections of Software Pioneers. — 1997.
5. Elke Jessen. Origin of the Virtual Memory Concept // IEEE Annals of the History of Computing. — 2004. — Т. 26, вып. 4. — С. 71—72.
6. Elke Jessen. Die Entwicklung des virtuellen Speichers // Informatik-Spektrum. — Springer-Verlag, 1996. — Т. 19, вып. 4. — С. 216—219. — ISSN 0170-6012 (<https://www.worldcat.org/search?fq=x0:jrnl&q=n2:0170-6012>). — doi:10.1007/s002870050034 (<https://dx.doi.org/10.1007/2Fs002870050034>).
7. R. J. Creasy. The origin of the VM/370 time-sharing system (http://pages.cs.wisc.edu/~stjones/roj/vm_reading/ibmrd2505M.pdf)  // IBM Journal of Research & Development. — 1981. — Т. 25, № 5. — С. 486.
8. Atlas design includes virtual memory (<http://curation.cs.manchester.ac.uk/computer50/www.computer50.org/kgill/atlas/atlas.html>) — Computer50, University of Manchester 1996
9. Ian Joyner on Burroughs B5000 (<https://web.archive.org/web/20080213220156/http://web.mac.com/joynerian/iWeb/Ian%20Joyner/Burroughs.html>) (англ.)
10. Harvey G. Cragon. Memory Systems and Pipelined Processors (<https://books.google.ru/books?id=q2w3JSFD7l4C&pg=PA113>). — Jones and Bartlett Publishers, 1996. — С. 113. — ISBN 0-86720-474-5.
11. Королёв Л. Н. Архитектура ЭВМ БЭСМ-6 (<http://www.parallel.ru/history/besm6.html>). Из книги Л. Н. Королёва «Структуры ЭВМ и их математическое обеспечение» (1978). PARALLEL.RU - Информационно-аналитический центр по параллельным вычислениям:. Дата обращения: 21 ноября 2016. Архивировано (<https://web.archive.org/web/20010413010855/http://parallel.ru/history/besm6.html>) 13 апреля 2001 года. ()
12. D. Sayre. Is automatic "folding" of programs efficient enough to displace manual? // Communications of the ACM. — ACM, 1969. — Т. 12, вып. 12. — С. 656—660. — ISSN 0001-0782 (<https://www.worldcat.org/search?fq=x0:jrnl&q=n2:0001-0782>). — doi:10.1145/363626.363629 (<https://dx.doi.org/10.1145/363626.363629>).
13. Harvey G. Cragon. Memory Systems and Pipelined Processors (<https://books.google.ru/books?id=q2w3JSFD7l4C&pg=PA121>). — Jones and Bartlett Publishers, 1996. — С. 121. — ISBN 0-86720-474-5.

14. Patterson, Hennessy «Computer organisation and Design. Hardware Software Interface.» 4th edition. ISBN 9780080886138 Chapter 5.4, page 503 (англ.)
15. В семействе операционных систем Microsoft Windows сегментная адресация не используется и заменена делением памяти на секции, что в сущности *почти* одно и то же.
16. Harvey G. Cragon. Memory Systems and Pipelined Processors (<https://books.google.ru/books?id=q2w3JSFD7l4C&pg=PA121>). — Jones and Bartlett Publishers, 1996. — С. 121. — ISBN 0-86720-474-5.

Литература

- Вахалия Ю. UNIX изнутри. — СПб.: Питер, 2003. С.844. ISBN 5-94723-013-5
- Иртегов Д. Введение в операционные системы. СПб.: Питер, 2002.
- Bruce Jacob, Trevor Mudge. Virtual Memory: Issues of Implementation. — Los Alamitos, CA, USA: IEEE Computer Society Press, 1998. ISSN:0018-9162
- Лекция 9: Виртуальная память. Архитектурные средства поддержки виртуальной памяти (<http://www.intuit.ru/studies/courses/2192/31/lecture/984>) / Основы операционных систем - Академия Intel, Московский физико-технический институт, 2004 - Интуит ISBN 978-5-9556-0044-4

Ссылки

- Virtual Memory: Issues of Implementation (<http://portal.acm.org/citation.cfm?id=621004>) / Computer Volume 31 Issue 6, June 1998 doi:10.1109/2.683005 (англ.)
- Bo Brantén, Windows driver examples: SwapFs (<http://www.acc.umu.se/~bosse/>) - Драйвер виртуальной памяти для Windows, использующий Linux swap раздел (англ.)
- FAQ по файлу подкачки на WinBlog.ru (<http://www.winblog.ru/2007/09/04/04090701.html>), 2007

Источник — https://ru.wikipedia.org/w/index.php?title=Виртуальная_память&oldid=111966811

Эта страница в последний раз была отредактирована 26 января 2021 в 00:32.

Текст доступен по лицензии Creative Commons Attribution-ShareAlike; в отдельных случаях могут действовать дополнительные условия.

Wikipedia® — зарегистрированный товарный знак некоммерческой организации Wikimedia Foundation, Inc.