МАТЕРИАЛЫ САЙТА

Куда пропадает память?

Журнал «Хакер» , 29.10.2008 10 мин на чтение 0 0 8142

Мобильная версия статьи

Содержание статьи

Сравнивая объемы оперативной памяти, выдаваемые BIOS и операционной системой с физическим объемом установленной памяти, практически всегда можно видеть, что доступно меньше памяти, чем установлено. "Пропажа" обычно составляет единицы мегабайт, но иногда достигает более существенных размеров. В предлагаемом материале перечисляются и детально рассматриваются причины данного явления. Также приведены рекомендации по оптимизации использования адресного пространства и оперативной памяти. Речь пойдет исключительно об объеме памяти, который BIOS сообщает операционной системе и о том, почему он меньше физического объема. Управление памятью внутри ОС – тема отдельной статьи.

Первая проверка

Разумеется, не всегда вопрос о "пропавшей" памяти задается исключительно из исследовательского интереса. И не всегда ответ лежит в области архитектуры и схемотехники материнской платы. Если после приобретения нового компьютера или переустановки модулей DIMM, мы видим, что памяти значительно меньше, чем заявлено поставщиком оборудования, возникает вполне обоснованное желание проверить комплектность нашей системы. Встречаются и случаи, когда надписи на наклейках модулей DIMM не соответствуют действительности. Анализ маркировки самих микросхем памяти, установленных на DIMM, также не всегда эффективен, так как не все производители придерживаются регулярной системы обозначений. Поэтому, перед тем, как перейти к главной теме статьи, напомним один рецепт для выявления банального подлога. Причем воспользоваться им можно, даже не открывая корпус компьютера.

Известно, что в современных системах, идентификация модулей оперативной памяти основана на использовании протокола SPD (Serial Presence Detect). На каждом модуле DIMM, вместе с микросхемами оперативной памяти, установлена микросхема постоянного запоминающего устройства (EPROM), объемом 256 байт. В нее производителем модуля записаны его параметры. При старте компьютера, BIOS считывает эти параметры и использует их для инициализации контроллера памяти. Диагностические программы, запускаемые в сеансе ОС (например, Astra32, Everest), также могут прочитать информацию SPD, таким образом она доступна для просмотра пользователем. Если по информации SPD объем памяти (сумма объемов модулей) соответствует значению, заявленному поставщиком, но вместе с тем, операционной системе доступно меньше памяти, то причина в особенностях архитектуры и схемотехники материнской платы, рассмотренных ниже, часть памяти выделена для использования различными устройствами или недоступна вследствие ограничений контроллера DRAM. Рассмотрению именно таких ситуаций посвящена данная статья. Если же, объем памяти, определенный на основании SPD, меньше ожидаемого, то все гораздо прозаичнее – нужно предъявлять претензию поставщику.

Ниже рассмотрены причины, по которым в распоряжение операционной системы попадает меньше оперативной памяти, чем физически установлено на плате. Очевидно, каждая из причин относится к одному из трех типов:

1. Часть памяти используется для внутренних нужд BIOS или устройств системной платы.
2. Часть памяти физически недоступна из-за ограничений контроллера памяти.
3. Часть памяти физически доступна в адресном пространстве, но не используется из-за ограничений операционной системы.

Выделение памяти для System Management RAM

System Management RAM – это память, используемая BIOS для собственных нужд. Физически, это часть оперативной памяти. Она "вырезана" из адресного пространства с помощью картирующей логики, входящей в состав "северного моста" чипсета. Данный вопрос детально рассмотрен в ранее опубликованной статье "SMM и SMRAM или 128Кб потусторонней памяти. Исследовательская работа № 5 и 6". Сколько памяти
будет "отрезано" для SMRAM зависит от реализации BIOS. В большинстве платформ это 128 Кбайт, используется диапазон 000A0000h-000BFFFFh, разделяемый с видео адаптером. В некоторых платформах также используется Extended SMRAM, расположенная выше 1MB и ее объем достигает нескольких мегабайт.

Выделение памяти для Shadow RAM

Shadow RAM или "теневая" память — область оперативной памяти, в которую переписывается или распаковывается содержимое микросхемы ROM BIOS материнской платы, а также дополнительные BIOS периферийных адаптеров. Первоначально это было задумано как опция, исключительно для повышения производительности, так как скорость работы RAM существенно выше, чем скорость работы ROM. Современные реализации BIOS, используют хранение основного блока в упакованном виде, при старте он распаковывается в Shadow RAM. Таким образом, операция Shadow из опциональной превратилась в обязательную. Упаковка позволяет использовать микросхему ROM меньшего объема, следовательно, более дешевую. Для корректной эмуляции ПЗУ, картирующая логика, входящая в состав "северного моста" чипсета, блокирует запись в данную область RAM. Распакованный блок BIOS, помещаемый в Shadow RAM, иногда называют Runtime-блоком.

В большинстве платформ, для Runtime-блоков BIOS периферийных адаптеров отводится диапазон 000C0000h-000EFFFFh. Для Runtime-блока системного BIOS – диапазон 000F0000h-000FFFFFh. Отметим, что даже если указанные диапазоны используются частично или не используются, весь 256-Кбайтный блок 000C0000h-000FFFFFh "отрезается" от оперативной памяти. Практически все современные чипсеты позволяют его использовать только как Shadow RAM.

Примечание

Утверждение о том, что RAM (ОЗУ) существенно быстрее, чем ROM (ПЗУ) справедливо для частного случая — применительно к элементной базе и схемотехнике персональных компьютеров, так как используются медленные микросхемы ROM и быстрые микросхемы RAM, к тому же, разрядность шины данных RAM на материнской плате значительно больше. К физическим принципам работы ячеек RAM и ROM это утверждение не относится.

Выделение памяти для таблиц ACPI

Спецификация ACPI, которая используется для передачи от BIOS к ОС информации о конфигурации платформы, а также для оптимизации энергопотребления, представляет собой альтернативный подход к взаимодействию BIOS и ОС. Напомним, что в "классических" функциях BIOS, например, в функциях дискового сервиса, доступных через программное прерывание INT 13h, операционная система или другая программа, для выполнения заданной операции, должна вызывать подпрограммы, входящие в состав BIOS. Взаимодействие ОС и платформы посредством ACPI выполняется принципиально по-другому. BIOS при старте платформы, перед загрузкой ОС, записывает в специальную область памяти набор таблиц, описывающих выполнение ряда операций. Упрощенно говоря, таблицы содержат информацию о том, какие данные в какой регистр записывать для выполнения заданной операции. ОС считывает эту информацию и использует при взаимодействии с оборудованием. Одно из преимуществ такого подхода, в том, что независимо от системы команд процессора или текущего режима работы (например, 16- 32- или 64-битный), можно использовать одни и те же таблицы, так как построение таблиц ACPI, в отличие от выполняемых процедур BIOS, не привязано к архитектуре процессора.

Объем памяти, выделяемый для хранения таблиц ACPI, зависит от реализации BIOS. Обычно это сотни килобайт, часто BIOS округляет размер резервируемой области до 1 Мбайта. Заметим, что в отличие от SMRAM (которая доступна только в режиме SMM) и Shadow RAM (которая имеет защиту от записи), область памяти, содержащая таблицы ACPI не имеет специального статуса с точки зрения контроллера памяти. Факт ее резервирования состоит только в том, что BIOS при передаче ОС информации об объеме памяти, передает значение с вычетом размера этой области. Для таблиц ACPI используется диапазон адресов, непосредственно примыкающий к верхней границе Extended памяти. Подробности в [14].

Выделение памяти для USB RAM

Как известно, контроллер USB является интеллектуальным устройством, способным взаимодействовать с оперативной памятью в обход процессора (в режиме Bus Master). Это взаимодействие состоит не только в передаче данных между устройствами, подключенными к USB и буферами в оперативной памяти. Для работы

контроллера USB требуется достаточно много вспомогательной информации в памяти, например расписание транзакций. Так как BIOS должен взаимодействовать с устройствами USB до загрузки ОС (например, ввод с USB клавиатуры, загрузка с Flash и т.п.), резервировать память должен BIOS, а не ОС. Обычно, резервируются десятки килобайт.

Заметим, что такие устройства, как например, контроллер жестких дисков, также поддерживают режим Bus Master и используют управляющую информацию, располагаемую в оперативной памяти. Но разница в том, что контроллер дисков, в отличие от контроллера USB, можно также использовать в режиме программного обмена (PIO Mode), что BIOS и делает при передаче управления на загрузку ОС. Переход в режим Bus Master (синоним DMA) и резервирование памяти под управляющие блоки, в этом случае является обязанностью ОС, а не BIOS.

Выделение памяти для интегрированного видео адаптера

Если на материнской плате имеется интегрированный видео адаптер, реализованный в составе "северного моста" чипсета, в качестве видео памяти обычно используется часть оперативной памяти. Перед загрузкой ОС, BIOS резервирует под видео память блок, размером единицы-десятки мегабайт. На некоторых платах, в BIOS Setup есть возможность управлять размером выделяемого блока. При старте ОС и загрузке видео драйвера, происходит инициализация графического процессора и в распоряжение видео адаптера динамически может быть выделено больше памяти.

Заметим, что существуют платы, у которых резервирование памяти для интегрированного видео адаптера происходит даже в том случае, когда он не используется. Одна из причин этого – небрежно написанный BIOS.

Также заметим, что не всегда интегрированный видео адаптер реализуется в составе "северного моста" чипсета. Существуют материнские платы, содержащие "полноценный" видео адаптер в виде отдельной микросхемы графического контроллера со своими микросхемами памяти. В этом случае системная память для нужд видео адаптера не резервируется.

Лимит 4Гб и Memory-Mapped I/O

Данный фактор иногда отнимает больше памяти, чем все остальные, вместе взятые. Причем, когда мы говорили о таких вещах как SMRAM, Shadow RAM, ACPI, то речь шла о том, что память, которую BIOS "утаил" от операционной системы, использовалась для внутренних нужд платформы. Здесь же, часть памяти просто пропадает. Когда и почему это происходит?

Возьмем реальный пример. Платформа класса Intel Socket 775. Процессор Intel Pentium 4 650 3.4 ГГц (ядро Prescott-2M), чипсет Intel 925XE. Устанавливаем 4 Гб памяти и видим, что операционной системе доступно около 3.5 Гб. Куда пропало около 0.5 Гб?

Расследование начнем с процессора. Читая документ [1] и просматривая назначение сигналов на Socket 775, видим, что процессор поддерживает 36-битную адресацию. Старший разряд адреса — A35# (считая от нулевого). Для справки, это контакт с координатами AJ6 на Socket 775. Количество адресуемых байтов равно 2 в степени 36, то есть, наш процессор может адресовать 64 Гб памяти. Благодаря механизму страничной трансляции, использование 36-битного адреса возможно как в 32- так и в 64-битном режиме. Подробности в [2]. Таким образом, в цепочке, которую мы прослеживаем, "слабым звеном" является явно не процессор.

Следующим компонентом, на пути от процессора к памяти, является "северный мост" чипсета, в нашем примере это микросхема Intel 82925XE, описанная в [3]. Из документа [3] следует, что чипсет поддерживает 32-битную адресацию, следовательно, объем адресного пространства памяти равен (2 в степени 32) байт, то есть 4 Гб. Причем, все 4 Гб нельзя отдать под оперативную память, требуется разместить еще ряд устройств, доступ к которым также осуществляется через пространство памяти. Именно поэтому, доступный объем оперативной памяти будет существенно меньше 4 Гб. Полный список таких устройств можно узнать из документов [3-4]. Для рассматриваемой платформы, наибольший объем отнимают видео память и окно для доступа к конфигурационным регистрам PCI Express. Рассмотрим их подробнее.

Классические адаптеры VGA, выпускавшиеся еще во времена шины ISA, используют постраничный доступ к видео памяти через окно, размер которого не превышает 128 KB (000A0000h-000BFFFFh). Современные адаптеры, поддерживая этот режим для совместимости, также поддерживают линейный доступ к видео памяти. При этом адаптеру с 256 MB видео памяти требуется выделить столько же адресного пространства. Из-за унификации при производстве видео адаптеров можно встретить и такие ситуации, когда адаптер со 128 MB видео памяти требует выделения окна размером 256 MB.

Классический механизм доступа к конфигурационному пространству шины PCI, описанный в [11-13] использует 256 байт конфигурационных регистров на устройство. Спецификация PCI Express использует блоки регистров, размером 4 KB, поэтому возникла необходимость в новом механизме доступа к ним. Новый механизм использует регион адресного пространства, размером 256 MB, через который конфигурационные регистры всех устройств адресуются как ячейки памяти. Подробности в [3-6].

Вопросы организации регистров, отображенных на память (Memory-Mapped I/O) рассмотрены в ранее опубликованной статье "Устройства системной поддержки. Исследовательская работа № 7,

8 и
9".

Операция Memory Remap

Начиная с чипсета Intel 955, лимит 4 Гб был преодолен. Разумеется, в модельных рядах чипсетов для серверов и рабочих станций это произошло значительно раньше.

Микросхема Intel 82955X принимает от процессора 36-битный адрес и поддерживает адресное пространство 64 Гб. Максимальный объем оперативной памяти – 8 Гб, на этот раз ограничение связано не с разрядностью адреса, который "северный мост" способен принять от процессора, а с возможностями контроллера DRAM.

Обычно, при использовании операции Memory Remap, диапазон 0-4 Гб форматирован так же, как и раньше. Там находится оперативная память, фрагмент которой недоступен из-за необходимости размещения других устройств. Новшество в том, что указанный фрагмент не пропадает, а размещается по адресам выше 4 Гб. Соответственно, если у нас памяти больше, чем 4 Гб, все, что не поместилось в диапазоне 0-4 Гб, размещается выше.

Разумеется, польза от физической доступности памяти выше 4 Гб будет только тогда, когда операционная система поддерживает адресацию выше 4 Гб. Это обеспечивается в 64-битном режиме, а также в 32-битном режиме при использовании PAE (Physical Address Extension). Если ОС не поддерживает адресацию выше 4 Гб, перемещенная память будет недоступна. Подробности в [2].

Следует помнить и о том, что обращения к памяти инициируются не только центральным процессором, но и другими устройствами, использующими технологию Bus Master, например контроллером жестких дисков. Если контроллер поддерживает только 32-битную адресацию при чтении и записи данных, то при размещении данных выше 4 Гб, потребуется дополнительно использовать транзитный буфер, расположенный ниже 4 Гб, так как контроллер дисков "не умеет" адресовать память выше 4 Гб. Пересылку между транзитным и целевым буфером должен выполнить центральный процессор. Это снижает производительность и отнимает память. Поэтому, "истинно 64-битной" платформу можно считать только тогда, когда не только процессор, но и Bus Master контроллеры поддерживают 64-битную адресацию.

Заключение

Логическим продолжением данного материала является изложение методов и фрагментов кода, позволяющих для заданной платформы "с точностью до бита"

определить, как используется память, которую BIOS "утаил" от операционной системы. Поэтому, при наличии читательского интереса, автор планирует продолжение. Задача осложняется тем, что для получения ответов на многие из поставленных вопросов, потребуется анализировать содержимое системных конфигурационных регистров, архитектура которых не определяется единым для всех платформ стандартом. Такие регистры по-своему реализованы в каждом чипсете. К сожалению, подробная документация доступна далеко не на все чипсеты. Поэтому, универсальных рецептов здесь не существует. Раскрывая данную тему, автор изложил основные принципы, используя которые, заинтересованный читатель может провести собственное исследование, для своей конкретной платформы.

Источники информации

Электронные документы, доступные на сайте developer.intel.com:

1) Intel Pentium 4 Processor 660, 650, 640, and 630 and Intel Pentium 4 Processor Extreme Edition Datasheet. Document Number: 306382-001.
2) TLBs, Paging-Structure Caches, and Their Invalidation. Application Note. Document Number 317080-001.
3) Intel 925X/925XE Express Chipset Datasheet. Document Number: 301464-003.
4) Intel I/O Controller Hub 6 (ICH6) Family Datasheet. Document Number 301473-001.
5) Intel 955X Express Chipset Datasheet. Document Number 306828-001.
6) Intel I/O Controller Hub 7 (ICH7) Family Datasheet. Document Number 307013-002.
7) AGP V3.0 Interface Specification (без номера).

Электронные документы, доступные на сайте developer.amd.com:

8) AMD Functional Data Sheet, 754 Pin Package. Publication # 31410.
9) AMD Functional Data Sheet, 939 Pin Package. Publication # 31411.
10) AMD Functional Data Sheet, 940 Pin Package. Publication # 31412.

Электронные документы, доступные на сайте pcisig.com:

Документы [12], [13] на сайте pcisig.com доступны только для членов PCI Special Interest Group. Воспользовавшись поисковыми системами, можно найти данные документы для свободной загрузки.

11) PCI BIOS Specification. Revision 2.1.
12) PCI Local Bus Specification. Revision 3.0.

13) PCI-to-PCI Bridge Architecture Specification. Revision 1.1.

Электронные документы, доступные на сайте acpi.info:

14) Advanced Configuration and Power Interface Specification. Hewlett-Packard Corporation, Intel Corporation, Microsoft Corporation, Phoenix Technologies Ltd., Toshiba Corporation. Revision 3.0.:

Книги:

15) В.Л. Григорьев. Микропроцессор i486. Архитектура и программирование. Москва ТОО "ГРАНАЛ" 1993.
16) Ю.М. Казаринов, В.Н. Номоконов, Г.С. Подклетнов, Ф.В. Филиппов. Микропроцессорный комплект К1810. Структура, программирование, применение. Справочная книга. Москва "Высшая школа" 1990.
17) М. Гук. Аппаратные средства IBM PC. Энциклопедия. Санкт-Петербург, издательство "Питер" 2006.

Теги: Статьи

Журнал «Хакер»

ДАЛЕЕ ПО ЭТОЙ ТЕМЕ | РАНЕЕ ПО ЭТОЙ ТЕМЕ

Вирус в Shadow RAM
03.12.2008

Проникновение в BIOS ROM №1
29.12.2008

Проникновение в BIOS ROM: осваиваем SPI Flash №1
07.05.2009

Проникновение в BIOS ROM: осваиваем SPI Flash №2
19.05.2009

Термальная угроза: мифы и реальность №1
15.07.2009

Еще одна аппаратная угроза или проникновение в SPD ROM №1
18.02.2009