

Память современных МС

Материалы по дисциплине «Микропроцессорные системы»
Специальность «Компьютерные системы и комплексы»
Составитель: Торгашин Р.Г

ГБПОУ ВО "Борисоглебский техникум промышленных и информационных технологий"

2016 год

Оглавление

Память программы начального запуска.....	3
Стек.....	3
Параметры оперативной памяти.....	4
Достоверность хранения данных.....	5
Статическая память (Static Random Access Memory - SRAM).....	6
Конструкция.....	6
Применение.....	8
Динамическая память.....	10
Типы DRAM.....	14
Страничная память (PM DRAM).....	14
Быстрая страничная память (FPM-DRAM).....	15
Память с усовершенствованным выходом (EDO).....	16
Пакетная EDO RAM (BEDO).....	17
Синхронная DRAM (SDRAM).....	17
DDR SDRAM.....	18
DDR2 SDRAM.....	20
DDR3 SDRAM.....	23
DDR3L SDRAM.....	25
DDR4 SDRAM.....	25
Форм-факторы модулей памяти.....	27
SIMM-30.....	27
SIMM-72.....	28
DIMM-168.....	29
DIMM-184.....	30
DIMM-240.....	30
DIMM-288.....	32
SO-DIMM -144.....	32
SO-DIMM -200.....	33
SO-DIMM -204.....	33
Двухканальный режим Dual-channel.....	34
Выбор модулей оперативной памяти.....	35
Источники.....	36

Память программы начального запуска

Память программы начального запуска всегда выполняется на ПЗУ или флэш-памяти.

Именно с этой области процессор начинает работу после включения питания и после сброса его с помощью сигнала RESET

Стек

Память для *стека* или *стек* (Stack) — это часть оперативной памяти, предназначенная для временного хранения данных в режиме *LIFO* (Last In — First Out). Особенность *стека* по сравнению с другой оперативной памятью — это заданный и неизменяемый *способ адресации*. При записи любого числа (кода) в *стек* число записывается по адресу, определяемому как содержимое *регистра* указателя *стека*, предварительно уменьшенное (декрементированное) на единицу (или на два, если 16-разрядные слова расположены в памяти почетным адресам). При чтении из *стека* число читается из адреса, определяемого содержимым указателя *стека*, после чего это содержимое указателя *стека* увеличивается (инкрементируется) на единицу (или на два). В результате получается, что число, записанное последним, будет прочитано первым, а число, записанное первым, будет прочитано последним.

Такая память называется *LIFO* или памятью магазинного типа (например, в магазине автомата патрон, установленный последним, будет извлечен первым).

Принцип действия *стека* показан на рисунке (адреса ячеек памяти выбраны условно).

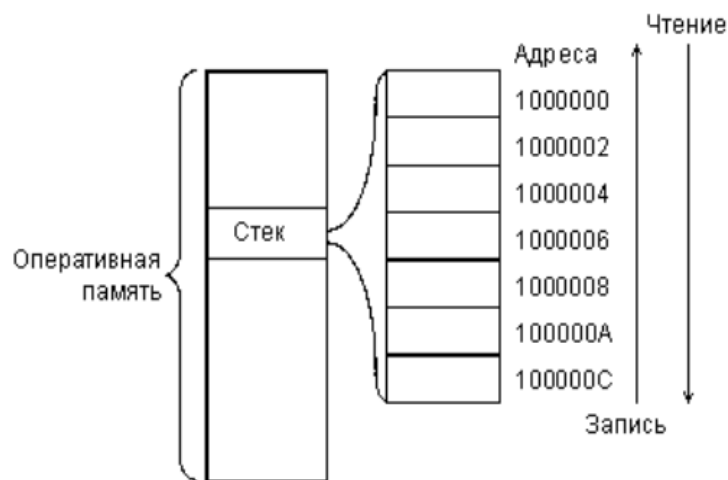


Рисунок 1: Принцип работы стека

Пусть, например, текущее состояние указателя *стека* 1000008, и в него надо записать два числа (слова).

Первое слово будет записано по адресу 1000006 (перед записью указатель *стека* уменьшится на два).

Второе — по адресу 1000004. После записи содержимое указателя *стека* — 1000004.

Если затем прочитать из *стека* два слова, то первым будет прочитано слово из адреса 1000004, а после чтения указатель *стека* *станет* равным 1000006.

Вторым будет прочитано слово из адреса 1000006, а указатель *стека* *станет* равным 1000008. Все вернулось к исходному состоянию.

Первое записанное слово читается вторым, а второе — первым.

Необходимость такой адресации становится очевидной в случае многократно вложенных подпрограмм.

Пусть, например, выполняется основная программа, и из нее вызывается подпрограмма 1.

Если нам надо сохранить значения данных и внутренних *регистров* основной программы на время выполнения подпрограммы, мы перед вызовом подпрограммы сохраним их в *стеке* (запишем в *стек*), а после ее окончания извлечем (прочитаем) их из *стека*. Если же из подпрограммы 1 вызывается подпрограмма 2, то ту же самую операцию мы сделаем с данными и содержимым внутренних *регистров* подпрограммы 1. Понятно, что внутри подпрограммы 2 крайними в *стеке* (читаемыми в первую очередь) будут данные из подпрограммы 1, а данные из основной программы будут глубже. При этом в случае чтения из *стека* автоматически будет соблюдаться нужный порядок читаемой информации. То же самое будет и в случае, когда таких уровней вложения подпрограмм гораздо больше. То есть то, что надо хранить подольше, прячется поглубже, а то, что скоро может потребоваться — с краю.

В системе команд любого процессора для обмена информацией со *стеком* предусмотрены специальные команды записи в *стек* (PUSH) и чтения из *стека* (POP). В *стеке* можно прятать не только содержимое всех внутренних *регистров* процессоров, но и содержимое *регистра* признаков (слово состояния процессора, PSW). Это позволяет, например, при возвращении из подпрограммы контролировать результат последней команды, выполненной непосредственно перед вызовом этой подпрограммой. Можно также хранить в *стеке* и данные, для того чтобы удобнее было передавать их между программами и подпрограммами. В общем случае, чем больше область памяти, отведенная под *стек*, тем больше свободы у программиста и тем более сложные программы могут выполняться.

Параметры оперативной памяти

Быстродействие памяти определяется временем выполнения операций записи и считывания данных. Основными параметрами любых элементов памяти является минимальное время доступа и длительность цикла обращения.

Время доступа (access time) определяется как задержка появления действительных данных на выходе памяти относительно начала цикла чтения, длительность цикла — как минимальный период следующих друг за другом обращений к памяти, причем циклы чтения и записи могут требовать различных затрат времени. В цикл обращения помимо активной фазы самого доступа входит и фаза восстановления (возврата памяти к исходному состоянию), которая соизмерима по времени с активной фазой. Временные характеристики самих запоминающих элементов определяются их принципом действия и технологией изготовления.

Производительность памяти можно характеризовать как скорость потока записываемых или считываемых данных и измерять в мегабайтах в секунду. Производительность подсистемы памяти наравне с производительностью процессора существенным образом определяет производительность компьютера. Выполняя определенный фрагмент программы, процессору придется, во-первых, загрузить из памяти соответствующий программный код, а во-вторых, произвести требуемые обмены данными, и чем меньше времени потребуется подсистеме памяти на обслуживание этих операций, тем лучше.

Производительность памяти, как основной, так и кэша второго уровня, обычно характеризуют длительностью пакетных циклов чтения (memory burst read cycle).

Производительность подсистемы памяти зависит от типа и быстродействия применяемых запоминающих элементов, разрядности шины памяти и некоторых «хитростей» архитектуры.

Современные типы памяти обеспечивают высокую скорость передачи внутри пакета, используя двойную и даже четырехкратную синхронизацию. При этом параметром шины, по которой передаются данные, может быть как частота тактового сигнала, так и частота передачи данных. Последняя может в 2 (DDR SDRAM) или в 4 (DDR2 SDRAM, шина Pentium 4) раза превышать тактовую частоту. Задержка получения данных чтения процессорным ядром в современных компьютерах может составлять от 45 до нескольких сотен наносекунд в зависимости от способа подключения памяти.

Производительность микросхем или модулей памяти повышают применением различных вариантов конвейеризации.

Разрядность шины памяти — это количество байтов (или битов), с которыми операция чтения или записи может быть выполнена одновременно. Разрядность основной памяти обычно согласуется с разрядностью внешней шины процессора {1 байт — для 8088; 2 байта — для 8086, 80286, 386SX; 4 байта — для 386DX, 486; 8 байт — для Pentium и выше}. Вполне очевидно, что при одинаковом быстродействии микросхем или модулей памяти производительность блока с большей разрядностью будет выше, чем у малоразрядного. Именно с целью повышения производительности у 32-битных (по внутренним регистрам) процессоров класса Pentium и выше внешняя шина, связывающая процессор с памятью, имеет разрядность 64 бита. У современных процессоров пропускная способность системной шины превышает пропускную способность шины памяти.

Достоверность хранения данных

В любой из многих миллионов ячеек памяти возможен случайный сбой или окончательный отказ, приводящий к ошибке. Вероятность ошибки, естественно, возрастает с увеличением объема памяти. Современные технологии позволяют выпускать высоконадежные микросхемы памяти, у которых при корректной эксплуатации (то есть при соблюдении заданных характеристик напряжения питания, температуры, временной диаграммы, уровней сигналов, нагрузки на выходные шины...) вероятность ошибки довольно мала, но все-таки она не нулевая.

Отказ ячейки памяти — потеря ее работоспособности, обычно требующая замены элемента памяти. Отказ может быть устойчивым, но возможно и самопроизвольное восстановление работоспособности, например после повторного включения питания. Часто причиной отказов является неисправность контакта или нарушение условий эксплуатации.

Случайный сбой может произойти и в исправной микросхеме памяти, например при пролете через нее ионизирующей частицы (по этой причине в условиях высокого уровня радиации обычные электронные элементы неработоспособны). После сбоя следующая же запись в ячейку произойдет нормально.

В первых моделях РС, когда микросхемы памяти имели существенно худшие характеристики надежности по сравнению с современными, обязательно применялся контроль четности. В этом случае каждый байт памяти сопровождается битом четности (parity bit), дополняющим количество единиц в байте до нечетного.

При обнаружении ошибки четности схемой контроля вырабатывается немаскируемое прерывание (NMI), его обработчик обычно выводит на экран сообщение «Parity Check Error» (ошибка четности) с указанием адреса сбойной ячейки и останавливает процессор командой Halt. Команда Halt (останов) останавливает выполнение текущего потока команд процессора. Из состояния останова процессор может выйти только по прерываниям, которые обработчиком ошибки четности могут и блокироваться. Останов не позволяет процессору «перепахать» всю память (в том числе и внешнюю), что, в принципе, возможно при

возникновении ошибки в памяти. «Сдвинуть с места» процессор, остановленный таким образом, можно нажатием кнопки Reset, выключением и последующим включением питания, а иногда и нажатием клавиш Ctrl+Alt+ Del

Со временем качество применяемых микросхем памяти улучшилось, и в целях удешевления модулей памяти от контроля четности стали отказываться — сначала через установку параметра CMOS Setup предлагали на выбор, проверять или не проверять четность, а потом появилась масса моделей системных плат, в которых контроля четности нет вообще. Модули памяти (SIMM, SIPP) стали выпускать как с битом четности, так и без него, а для «ублажения» (точнее— обмана) плат, требующих наличия такого бита, стали выпускать модули с «подделкой» четности (fake parity). В этих модулях вместо дополнительной микросхемы памяти используется генератор четности (Parity Generator, PG) — логическая схема сумматора по модулю 2, формирующая всегда «хороший» бит четности независимо от ошибок в самой памяти.

В компьютерах особо ответственного применения используют память с обнаружением и коррекцией ошибок (Error Checking and Correcting, ECC). В этом случае для каждого записываемого информационного слова памяти (а не байта, как при контроле четности) по определенным правилам вычисляется функция свертки, результат которой разрядностью в несколько битов также хранится в памяти. Для 64-битного слова обычно используют 7 — 8 дополнительных битов. При считывании схема контроля с использованием этих избыточных битов способна обнаруживать ошибки различной кратности и/или исправлять однократные ошибки. Обнаружение ошибки выполняется «на лету» и само по себе дополнительного времени не требует. Однако при исправлении ошибок требуются дополнительные такты для срабатывания аппаратной логики исправления и фиксации скорректированного результата.

Даже при отсутствии ошибки нужен по крайней мере один лишний такт для срабатывания схемы контроля, разрешающей использование считанных данных. Таким образом, ECC-память работает несколько медленнее неконтролируемой, и при наличии исправимых ошибок замедление становится заметнее.

Возможны различные варианты поведения, например:

- автоматически исправлять ошибки, не уведомляя об этом систему;
- исправлять однократные ошибки, уведомляя систему только о многократных;
- не исправлять ошибки, а только уведомлять об их обнаружении (самый

достоверный контроль).

Статическая память (Static Random Access Memory - SRAM)

Конструкция

Статическая память выполняется обычно на основе ТТЛ-, КМОП- или БиКМОП-микросхем и по способу доступа к данным может быть как асинхронной, так и синхронной. Асинхронным называется доступ к данным, который можно осуществлять в произвольный момент времени. Асинхронная SRAM применялась на материнских платах для третьего - пятого поколения процессоров. Время доступа к ячейкам такой памяти составляло от 15 нс (33 МГц) до 8 нс (66 МГц).

Ядро микросхемы статической оперативной памяти представляет собой совокупность триггеров - логических устройств, имеющих два устойчивых состояния, одно из которых условно соответствует логическому нулю, а другое - логической единице. Другими словами, каждый триггер хранит один бит информации, - ровно столько же, сколько и ячейка динамической памяти

Между тем, триггер как минимум по двум позициям обыгрывает конденсатор (являющийся основой динамической памяти):

- а) состояния триггера устойчивы и при наличии питания могут сохраняться бесконечно долго, в то время как конденсатор требует периодической регенерации;
- б) триггер, обладая мизерной инертностью, без проблем работает на частотах вплоть до нескольких ГГц, тогда как конденсаторы "сваливаются" уже на 75-100 МГц.

К недостаткам триггеров следует отнести их высокую стоимость и низкую плотность хранения информации.

Триггеры объединяются в единую матрицу, состоящую из строк (row) и столбцов (column), последние из которых так же называются битами (bit).

В отличие от ячейки динамической памяти, для управления которой достаточно всего одного ключевого транзистора, ячейка статической памяти управляется как минимум двумя. Это не покажется удивительным, если вспомнить, что триггер, в отличие от конденсатора, имеет отдельные входы для записи логического нуля и единицы соответственно. Таким образом, на ячейку статической памяти расходуется целых восемь транзисторов - четыре идут, собственно, на сам триггер и еще два — на управляющие "защелки".

Основной недостаток шести транзисторной ячейки заключается в том, что в каждый момент времени может обрабатываться всего лишь одна строка матрицы памяти. Параллельное чтение ячеек, расположенных в различных строках одного и того же банка невозможно, равно как невозможно и чтение одной ячейки одновременно с записью другой.

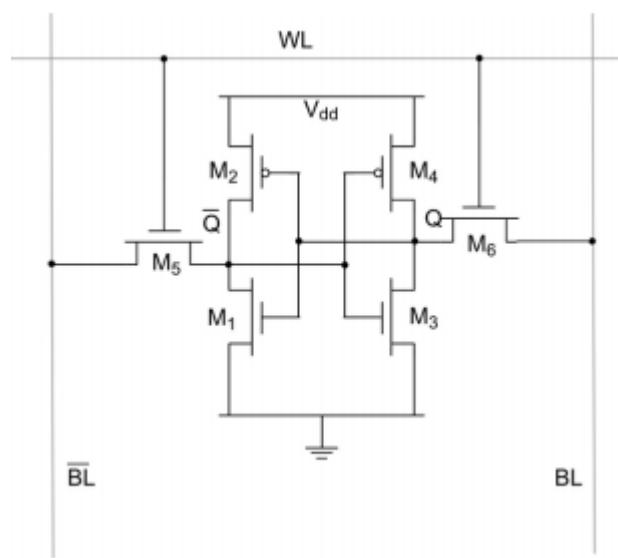


Рисунок 2: Шеститранзисторная ячейка

Этого ограничения лишена многопортовая память. Каждая ячейка многопортовой памяти содержит один-единственный триггер, но имеет несколько комплектов управляющих транзисторов, каждый из которых подключен к "своим" линиям ROW и BIT, благодаря чему различные ячейки матрицы могут обрабатываться независимо. Такой подход намного более прогрессивен, чем деление памяти на банки. Ведь, в последнем случае параллелизм достигается лишь при обращении к ячейкам различных банков, что не всегда выполнимо, а много портовая память допускает одновременную обработку любых ячеек, избавляя программиста от необходимости вникать в особенности ее архитектуры.

Наиболее часто встречается двух - портовая память, Нетрудно подсчитать, что для создания одной ячейки двух - портовой памяти расходуется аж восемь транзисторов. Пусть емкость кэш-памяти составляет 32 Кб, тогда только на одно ядро уйдет свыше двух миллионов транзисторов.

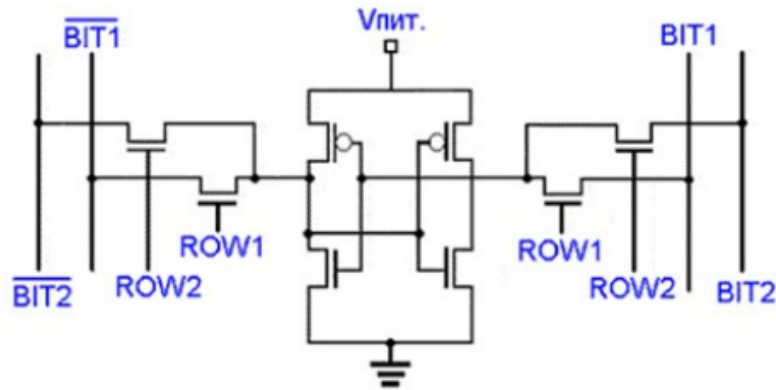


Рисунок 3: Многопортовая статическая память

Асинхронная статическая память работает независимо от контроллера и потому, контроллер не может быть уверен, что окончание цикла обмена совпадет с началом очередного тактового импульса. В результате, цикл обмена удлиняется по крайней мере на один такт, снижая тем самым эффективную производительность. "Благодаря" последнему обстоятельству, в настоящее время асинхронная память практически нигде не применяется (последними компьютерами, на которых она еще использовались в качестве кэша второго уровня, стали машины, построенные на базе процессора Intel 80386).

Синхронная статическая память выполняет все операции одновременно с тактовыми сигналами, в результате чего время доступа к ячейке укладывается в один-единственный такт. Именно на синхронной статической памяти реализуется кэш первого уровня современных процессоров. Конвейерная статическая память представляет собой синхронную статическую память, оснащенную специальными "защелками", удерживающими линии данных, что позволяет читать (записывать) содержимое одной ячейки параллельно с передачей адреса другой.

Так же, конвейерная память может обрабатывать несколько смежных ячеек за один рабочий цикл. За счет большей аппаратной сложности конвейерной памяти, время доступа к первой ячейке пакета увеличивается на один такт, однако, это практически не снижает производительности.

Применение

SRAM применяется в микроконтроллерах и ПЛИС¹, в которых объём ОЗУ невелик (единицы килобайт), зато нужны низкое энергопотребление (за счёт отсутствия сложного контроллера динамической памяти), предсказываемое с точностью до такта время работы подпрограмм и отладка прямо на устройстве.

В устройствах с большим объёмом ОЗУ рабочая память выполняется как DRAM. SRAM же применяется для **регистров и кэш-памяти**.

Первоначально, кэш память устанавливалась на системных плаках в виде отдельных модулей. Сейчас такая память интегрируется в кристалл CPU.

1 Программируемая логическая интегральная схема

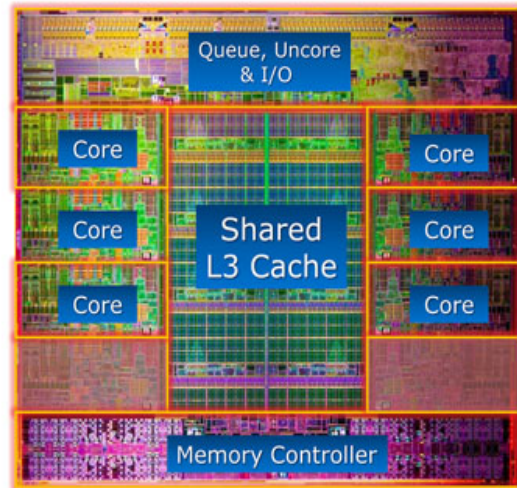
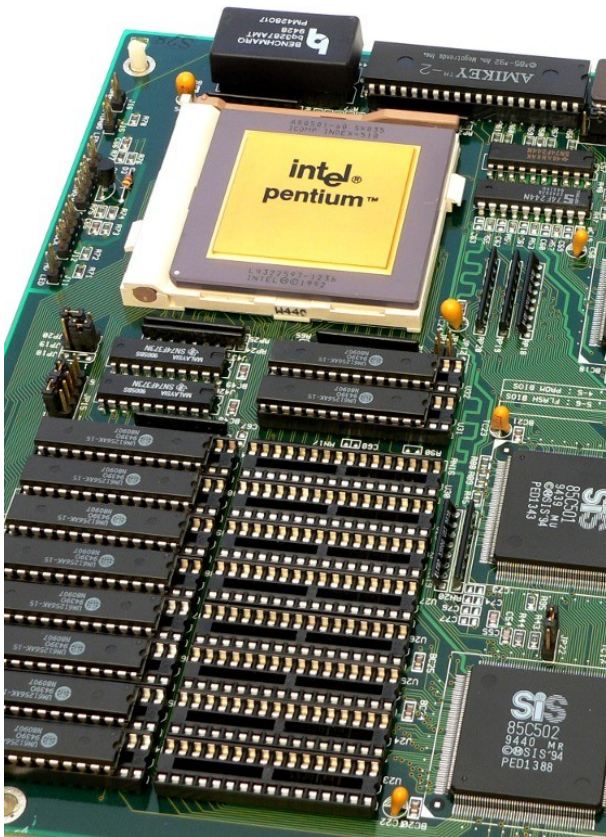


Рисунок 4: КЭШ — память на системной плате Elitegroup SI5PI 1994 года выпуска (слева) и в процессоре Intel Core-i7

Микросхемы SRAM могут использоваться в качестве кэша или буфера в различных устройствах, например жестких дисках.

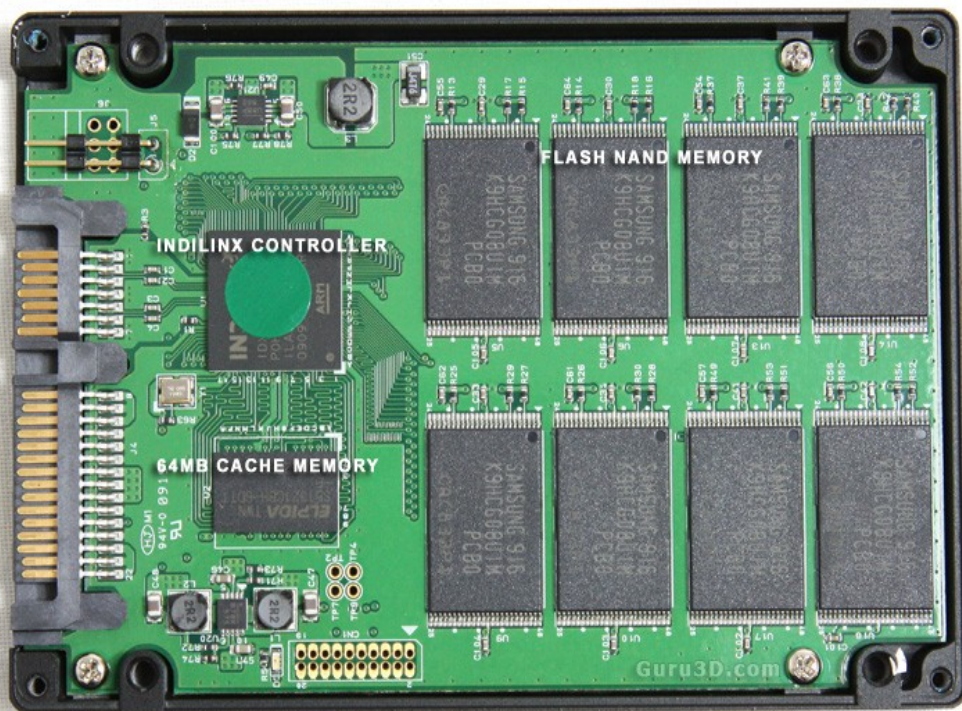


Рисунок 5: КЭШ на плате контроллера накопителя Corsair X128 SSD

Динамическая память

Динамическая память (Dynamic RAM, DRAM) получила свое название от принципа действия ее запоминающих ячеек, которые выполнены в виде конденсаторов, образованных элементами полупроводниковых микросхем. Несколько упрощая описание физических процессов, можно сказать, что при записи логической единицы в ячейку конденсатор заряжается, при записи нуля — разряжается.

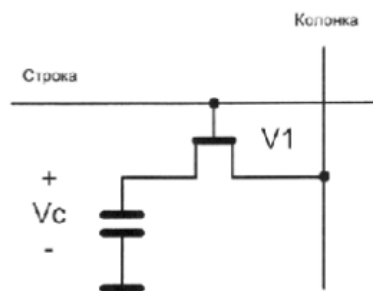


Рисунок 6: Ячейка DRAM

Схема считывания разряжает через себя этот конденсатор, и если заряд был ненулевым, выставляет на своем выходе единичное значение и подзаряжает конденсатор до прежнего уровня. При отсутствии обращения к ячейке со временем за счет токов утечки конденсатор разряжается и информация теряется, поэтому такая память требует постоянной периодической подзарядки конденсаторов (обращения к каждой ячейке), то есть память может работать только в динамическом режиме. Этим она принципиально отличается от статической памяти, реализуемой на триггерных ячейках и хранящей информацию без обращений к ней сколь угодно долго (при включенном питании).

Благодаря относительной простоте ячейки динамической памяти на одном кристалле удастся размещать миллионы ячеек и получать самую дешевую полупроводниковую память достаточно высокого быстродействия с умеренным энергопотреблением, используемую в качестве основной памяти компьютера. Расплатой за низкую цену являются некоторые сложности в управлении динамической памятью.

Запоминающие ячейки микросхем DRAM организованы в виде двухмерной матрицы.

Адрес строки и столбца передается по мультиплексированной шине адреса MA (Multiplexed Address) и обрабатывается по импульсам RAS (Row Access Strobe — строб адреса строки) и CAS (Column Access Strobe — строб адреса столбца).

Выбранной микросхемой памяти является та, на которую во время активности (низкого уровня) сигнала RAS приходит сигнал CAS (тоже низким уровнем)

Функционирование ячейки DRAM

В спокойном состоянии на обоих выводах поддерживается высокий уровень сигнала, При попытке прочесть содержимое некоторой ячейки памяти:

1. Контроллер преобразует физический адрес в пару чисел - номер строки и номер столбца, а затем посылает первый из них на адресные линии.
2. Дождавшись, когда сигнал стабилизируется, контроллер сбрасывает сигнал RAS в низкий уровень, сообщая микросхеме памяти о наличии информации на линии.
3. Микросхема считывает этот адрес и подает на соответствующую строку матрицы электрический сигнал.

а) Все транзисторы, подключенные к этой строке, открываются и конденсатор разряжается на входы чувствительного усилителя.

б) Чувствительный усилитель декодирует всю строку, преобразуя ее в последовательность нулей и единиц, и сохраняет полученную информацию в специальном буфере.

Пока микросхема обрабатывает адрес, контроллер выдерживает паузу - "RAS to CAS delay" (на сухом официальном языке — t_{RCD}).

4. Контроллер подает на адресные линии номер колонки и, дав сигналу стабилизироваться, сбрасывает CAS в низкое состояние.

5. Микросхема номер колонки в смещение ячейки внутри буфера. Остается всего лишь прочесть ее содержимое и выдать его на линии данных.

Это занимает еще какое-то время, в течение которого контроллер ждет запрошенную информацию- "CAS delay" (или t_{CAS})

6. На финальной стадии цикла обмена контроллер считывает состояние линий данных, деактивирует сигналы RAS и CAS, устанавливая их в высокое состояние, а микросхема берет определенный тайм-аут на перезарядку внутренних цепей и восстановительную перезапись строки.

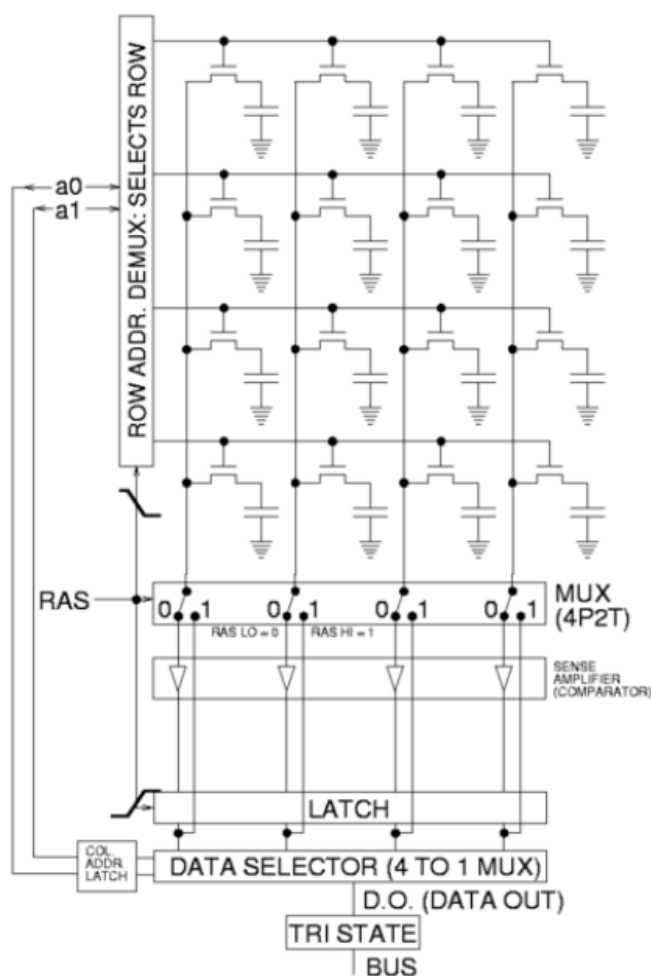


Рисунок 7: Массив ячеек DRAM

Пауза в работе контроллера между чтением последней ячейки и подачей номера новой строки - "RAS precharge" (t_{RP}).

На жаргоне эти параметры RAS to CAS delay, CAS delay, RAS precharge принято называть таймингами²

Часто тайминги записывают в виде серии цифр: 2-2-3. Мера таймингов — такт. Таким образом, каждая цифра в формуле 2-2-3 означает задержку сигнала для обработки, измеряемая в тактах системной шины.

² В микросхемах памяти SDRAM – DDR, DDR2 и DDR3 добавляются другие тайминги

Иногда формула таймингов для памяти может состоять из четырёх цифр, например 2-2-2-6. Последний параметр называется «DRAM Cycle Time Tras/Trc» и характеризует быстродействие всей микросхемы памяти. Он определяет отношение интервала, в течение которого строка открыта для переноса данных (tRAS — RAS# Active time), к периоду, в течение которого завершается полный цикл открытия и обновления ряда (tRC — Row Cycle time), также называемого циклом банка (Bank CycleTime).

Теоретически, чем меньше величина этих таймингов, тем быстрее работает оперативная память.

Пример из практики: система с памятью на частоте 100 МГц с таймингами 2-2-2 обладает примерно такой же производительностью, как та же система на частоте 112 МГц, но с задержками 3-3-3. Другими словами, в зависимости от задержек, разница в производительности может достигать 10 %.

Однако слишком маленькие тайминги могут привести к тому, что микросхема не будет успевать выполнять операции. В результате, из-за частых ошибок чтения/записи производительность банка памяти снизится или он перестанет функционировать.

Рекомендованные производителем тайминги записываются в микросхему SPD на плате памяти. При включении компьютера BIOS может прочитать эти параметры и автоматически установить тайминги для данной платы памяти. Некоторые производители предлагают на выбор несколько наборов параметров. Но ни один не указывает минимально допустимы тайминги в качестве рекомендованных.

Использование микросхем в таком форсированном режиме делает их работу нестабильной и снижает ресурс. Поэтому производители перестраховываются.

Однако, в ручную можно подобрать такие параметры. Следует только помнить, что использование такой, разогнанной, памяти может привести к сбоям в работе системы и потребовать дополнительного охлаждения плат RAM

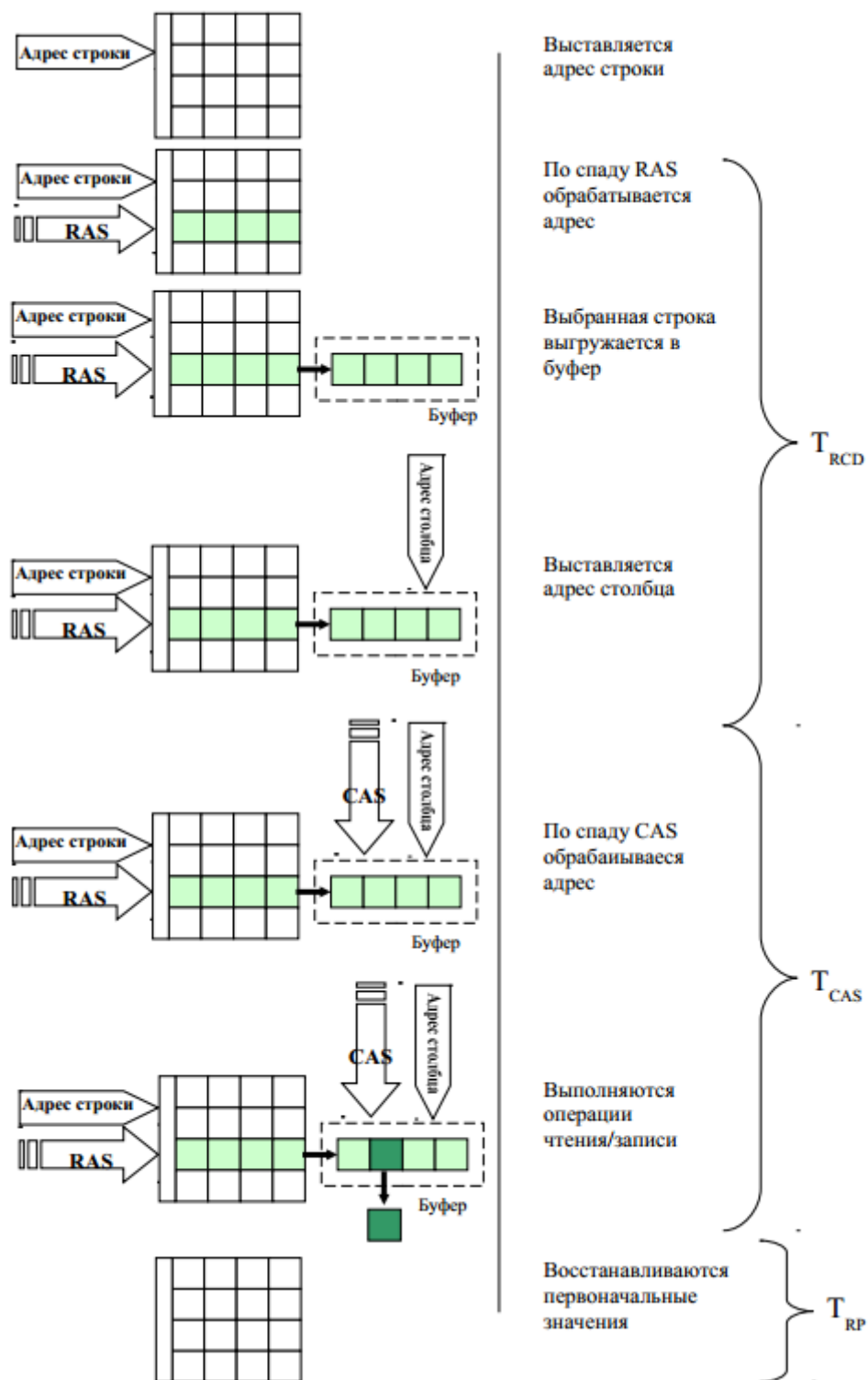


Рисунок 8: Алгоритм работы ячейки DRAM

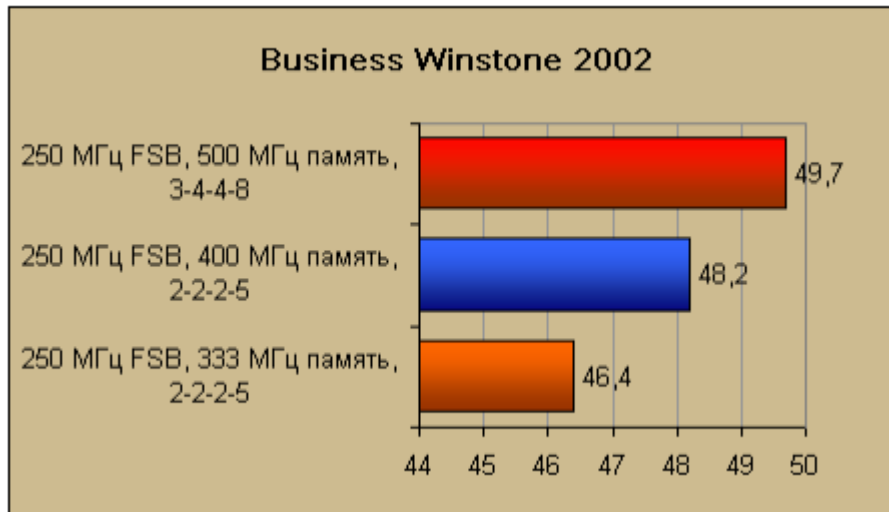


Рисунок 9: Зависимость производительности RAM от таймингов

Регенерация

Поскольку обращения (запись или чтение) к различным ячейкам памяти обычно происходят в случайном порядке, для поддержания сохранности данных применяется *регенерация памяти* (memory refresh) — регулярный циклический перебор ее ячеек (обращение к ним) с холостыми циклами.

Регенерация в микросхеме происходит одновременно по всей строке матрицы при обращении к любой из ее ячеек.

При *распределенной регенерации* (distributed refresh) одиночные циклы регенерации выполняются равномерно с периодом t_{RF} который для стандартной памяти принимается равным 15,6 мкс. Период этих циклов называют *частотой регенерации* (refresh rate).

Для памяти с расширенной регенерацией (extended refresh) допустим период циклов до 125 мкс.

Возможен также и вариант *пакетной регенерации* (burst refresh), когда все циклы регенерации собираются в пакет, во время которого обращение к памяти по чтению и записи блокируется. При количестве циклов 1024 эти пакеты будут периодически занимать шину памяти примерно на 130 мкс, что далеко не всегда допустимо. По этой причине, как правило, выполняется распределенная регенерация, хотя возможен и промежуточный вариант — пакетами по несколько (например, 4) циклов.

Типы DRAM

Страничная память (PM DRAM)

Страничная память (англ. page mode DRAM, PM DRAM) являлась одним из первых типов выпускаемой компьютерной оперативной памяти. Алгоритм ее работы наиболее близок к описанному в предыдущей главе.

Память такого типа выпускалась в начале 90-х годов, но с ростом производительности центральных процессоров и ресурсоемкости приложений требовалось увеличивать не только объем памяти, но и скорость её работы.

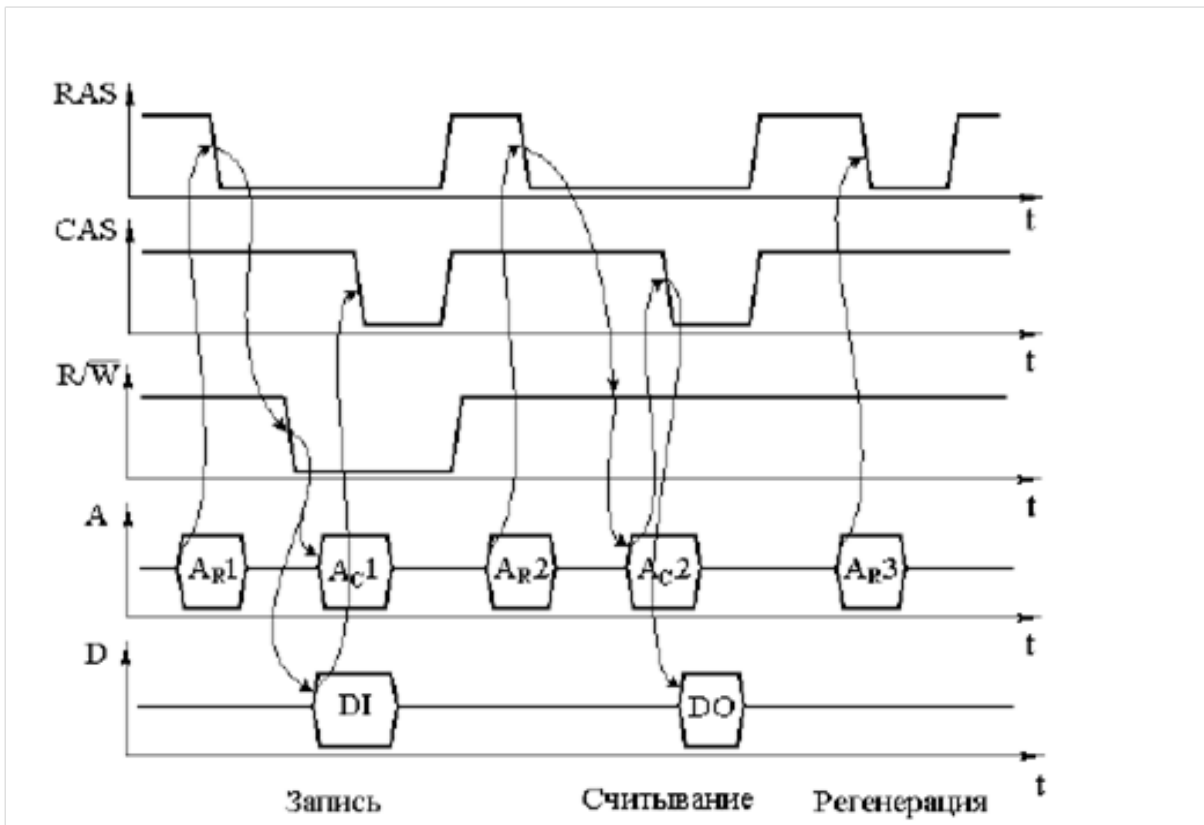


Рисунок 10: Работа PM DRAM

Быстрая страничная память (FPM-DRAM)

Основным отличием от памяти предыдущего поколения стала поддержка сокращенных адресов. Если очередная запрашиваемая ячейка находится в той же самой строке, что и предыдущая, ее адрес однозначно определяется одним лишь номером столбца и передача номера строки уже не требуется. За счет чего это достигается? Обратимся к диаграмме, изображенной на рисунке.

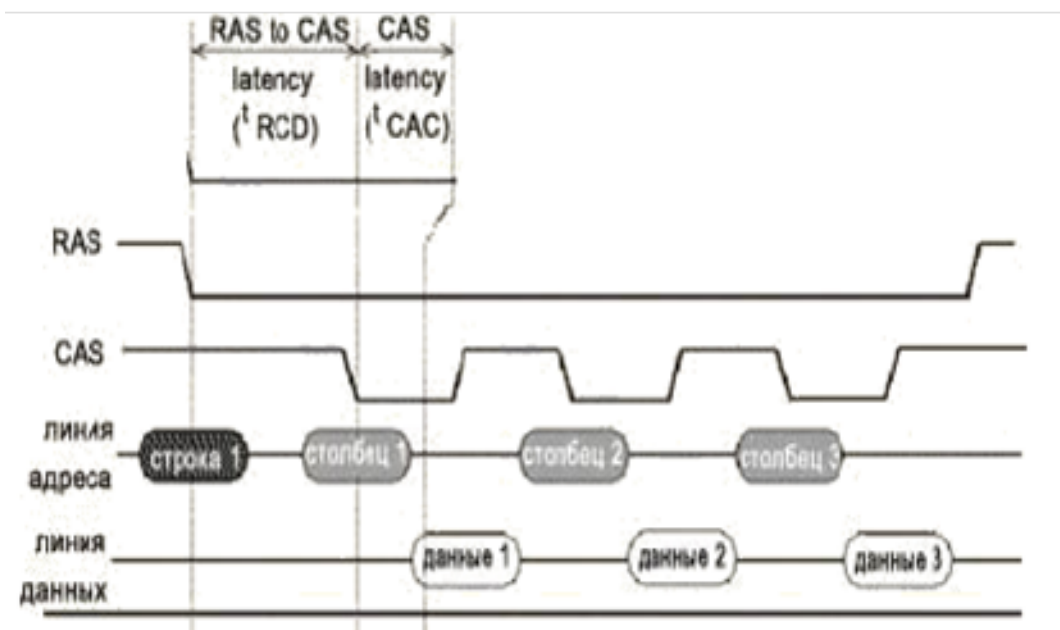


Рисунок 11: Работа FPM DRAM

В то время как при работе с обычной DRAM (верхняя диаграмма) после считывания данных сигнал RAS деактивируется, подготавливая микросхему к новому циклу обмена, контроллер FPMDRAM удерживает RAS в низком состоянии, избавляясь от повторной пересылки номера строки.

При последовательном чтении ячеек памяти, (равно как и обработке компактных одно-двух килобайтовых структур данных), время доступа сокращается на 40%, а то и больше, ведь обрабатываемая строка находится во внутреннем буфере микросхемы, и обращаться к матрице памяти нет никакой необходимости.

Память с усовершенствованным выходом (EDO)

С появлением процессоров Intel Pentium II память FPM DRAM оказалась неэффективной.

Следующим шагом стала память с усовершенствованным выходом (англ. extended data out DRAM, EDO DRAM). Её производительность оказалась на 10-15 % выше. Её рабочая частота была 40 МГц и 50 МГц, соответственно, время полного доступа — 60 нс и 50 нс, а время рабочего цикла — 25 нс и 20 нс.

Микросхему снабдили специальным триггером-защелкой, удерживающим линии данных после исчезновения сигнала CAS, становя возможным деактивировать CAS до окончания чтения данных, подготавливая в это время микросхему к приему номера следующего столбца.

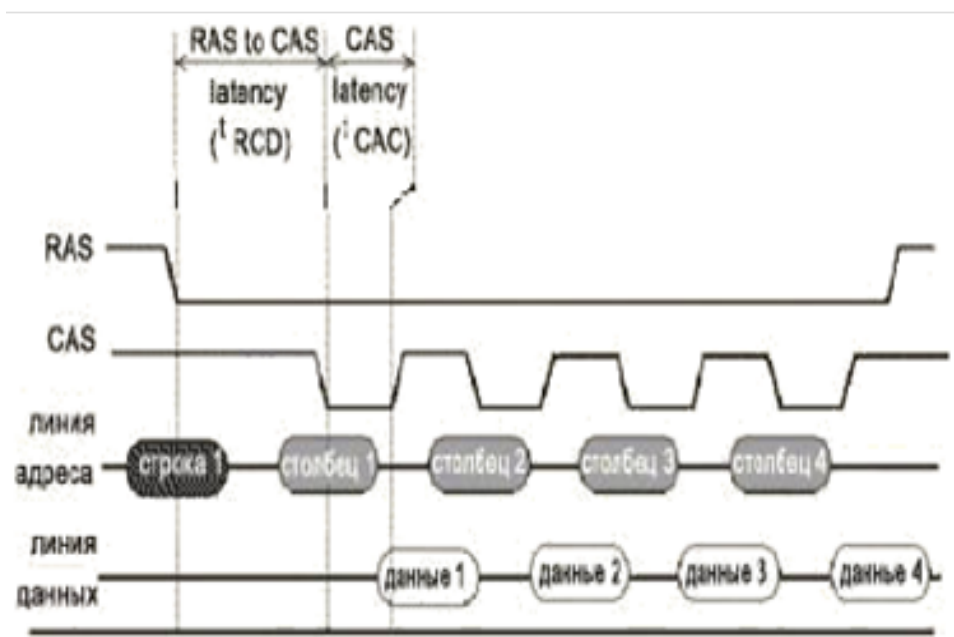


Рисунок 12: Работа EDO DRAM

Продолжительность рабочего цикла EDO-DRAM (в зависимости от качества микросхемы) составляла 30, 25 и 20 нс., что соответствовало всего двум тактам в 66 МГц системе.

Совершенствование производственных технологий сократило и полное время доступа. На частоте 66 МГц формула лучших EDO-микросхем выглядела так: 5-2-x-x. Простой расчет позволяет установить, что пиковый прирост производительности (в сравнении с FPM-DRAM) составляет около 30%.

Пакетная EDO RAM (BEDO)

Двукратное увеличение производительности EDO было достигнуто лишь в BEDO-DRAM (Burst EDO). Добавив в микросхему генератор номера столбца, конструкторы ликвидировали задержку CAS Delay, сократив время цикла до 15 нс.

После обращения к произвольной ячейке микросхема BEDO автоматически, без указаний со стороны контроллера, увеличивает номер столбца на единицу, не требуя его явной передачи. По причине ограниченной разрядности адресного счетчика (конструкторы отвели под него всего лишь два бита) максимальная длина пакета не могла превышать четырех ячеек ($2^2=4$).

Процессоры Intel 80486 и Pentium в силу пакетного режима обмена с памятью никогда не обрабатывают менее четырех смежных ячеек за раз. Поэтому, независимо от порядка обращения к данным, BEDO всегда работает на максимально возможной скорости и для частоты 66 МГц ее формула выглядит так: 5-1-1-1, что на ~40% быстрее EDO-DRAM!

Все же, несмотря на свои скоростные показатели, BEDO оказалась не конкурентоспособной и не получила практически никакого распространения.

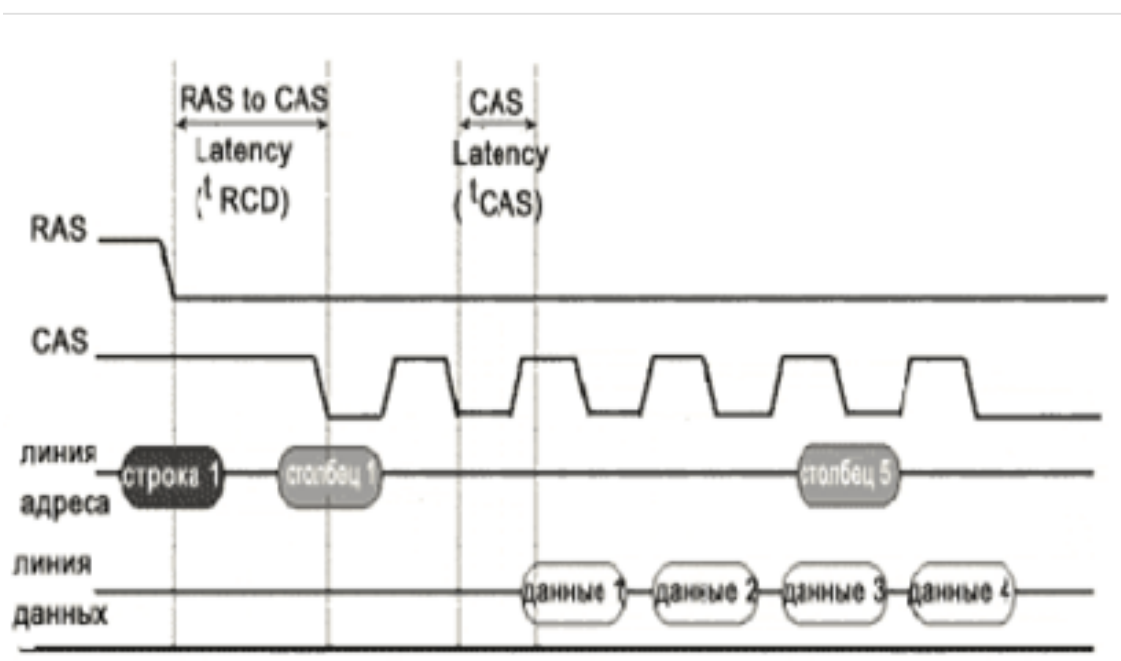


Рисунок 13: Работа BEDO DRAM

Синхронная DRAM (SDRAM)

Появление микропроцессоров с шинами на 100 MHz привело к радикальному пересмотру механизма управления памятью, и подтолкнуло конструкторов к созданию синхронной динамической памяти - SDRAM (Synchronous-DRAM). Как и следует из ее названия, микросхемы SDRAM памяти работают синхронно с контроллером, что гарантирует завершение цикла в строго заданный срок.

Кроме того, номера строк и столбцов подаются одновременно, с таким расчетом, чтобы к приходу следующего тактового импульса сигналы уже успели стабилизироваться и были готовы к считыванию.

Так же, в SDRAM реализован усовершенствованный пакетный режим обмена. Контроллер может запросить как одну, так и несколько последовательных ячеек памяти, а при желании - всю строку целиком! Это стало возможным благодаря использованию полноразрядного адресного счетчика уже не ограниченного, как в BEDO, двумя битами.

Количество матриц (банков) памяти в SDRAM увеличено с одного до двух (а, в некоторых моделях, и четырех). Это позволяет обращаться к ячейкам одного банка параллельно с перезарядкой внутренних цепей другого, что вдвое увеличивает предельно допустимую тактовую частоту.

Помимо этого появилась возможность одновременного открытия двух (четырех) страниц памяти, причем открытие одной страницы (т.е. передача номера строки) может происходить во время считывания информации с другой, что позволяет обращаться по новому адресу столбца ячейки памяти на каждом тактовом цикле.

В отличие от FPM-DRAM\EDO-DRAM\BEDO, выполняющих перезарядку внутренних цепей при закрытии страницы (т.е. при деактивации сигнала RAS), синхронная память прodelывает эту операцию автоматически, позволяя держать страницы открытыми столь долго, сколько это угодно.

Наконец, разрядность линий данных увеличилась с 32 до 64 бит, что еще вдвое увеличило ее производительность! Формула чтения произвольной ячейки из закрытой строки для SDRAM обычно выглядит так: 5-1-x-x, а открытой так: 3-1-x-x. В настоящее время все компьютеры оснащаются SDRAM и ее модификациями

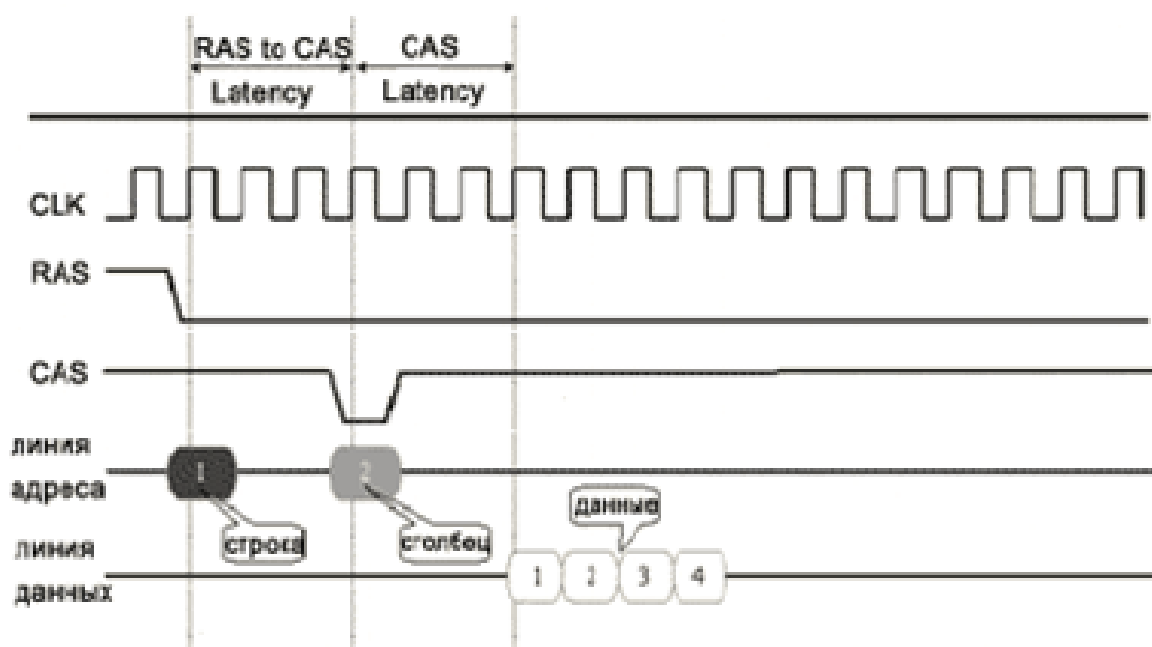


Рисунок 14: Работа SDRAM

DDR SDRAM

По мере дальнейшего увеличения быстродействия микропроцессоров эффективное быстродействие элементов памяти становится узким местом компьютерных систем. Если частоты микропроцессоров давно превысили частоты системной шины, то шина памяти до самого последнего времени работала синхронно с последней.

Значения этих частот явно указывали в спецификации на типы синхронной динамической памяти SDRAM: PC66, затем PC100 и наконец PC133. Однако в то время как частота шины памяти увеличилась на треть и соответственно на столько же возросла ее пропускная способность (с 800 Мб/с до 1064 Мб/с), частоты микропроцессоров повысились в два с половиной раза - с 400 МГц до 1 ГГц.

Действительно, если пропускная способность PC133 SDRAM составляет лишь 1064 Мб/с, то современным ПК требуется вдвое большая полоса пропускания (1 Гб/с для

микропроцессора с частотой системной шины 133 МГц, столько же для графической шины AGP 4x и 132 Мб/с для 33 МГц шины PCI, что в сумме составляет более 2 Гб/с).

Однако дальнейшее повышение тактовой частоты для микросхем SDRAM было неперспективным. Уже с тактовой частоты 166 МГц память SDRAM получается слишком дорогой, особенно с учетом объемов ОЗУ современных компьютеров. Требовалось найти простое технологическое решение, которое и воплотилось в DDR SDRAM.

Память DDR SDRAM сохранила архитектуру, количество банков, тонкие малогабаритные корпуса типа TSOP (Thin Small-Outline Package) и сам технологический процесс производства SDRAM, однако включает некоторые схмотехнические усовершенствования, которые позволяют существенно увеличить ее быстродействие.

Применена более жесткая синхронизация работы устройства.

Введены схемы DDL (Delay Locked Loop), обеспечивающие для сигналов стробирования данных цикл с фиксированной задержкой. Используя эти сигналы, контроллер памяти более точно осуществляет синхронизацию данных, поступающих от разных модулей памяти одного банка.

Микросхемы DDR SDRAM по сравнению с SDRAM фактически увеличивают скорость доступа к данным вдвое при одной и той же частоте. Дело в том, что применение DDR SDRAM дает возможность **читать информацию как по переднему, так и по заднему фронту тактового сигнала.**

Кроме того, частота работы повышается за счет применения интерфейсных логических схем с еще более пониженным уровнем питания. Если для SDRAM обычно используются схмотехнические решения на базе LVTTTL (Low Volt Transistor-to-Transistor Logic) с напряжением питания 3,3 В, то в DDR SDRAM - на базе SSTL (Stub Series Terminated Logic) с напряжением 2,5 В (а в перспективе и SSTL-2 с напряжением 1,25 В).

Форм-фактор модулей DDR SDRAM изменился, число контактов увеличилось со 168 до 184, а иное положение ключа не позволяет вставить модули DIMM DDR в разъемы для SDRAM.

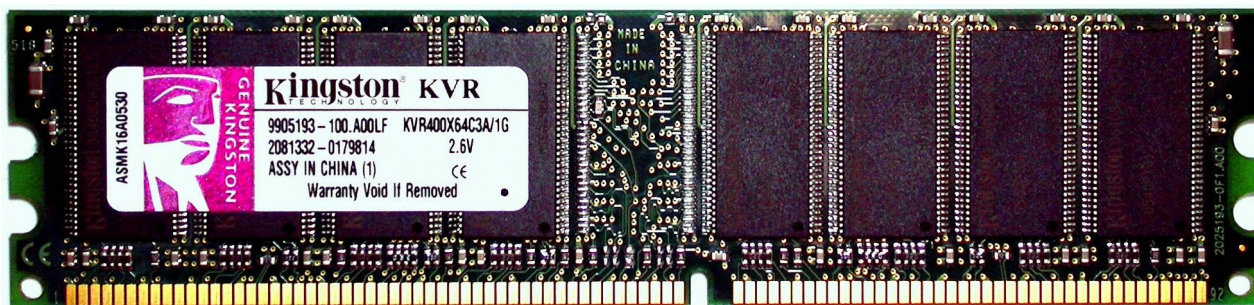
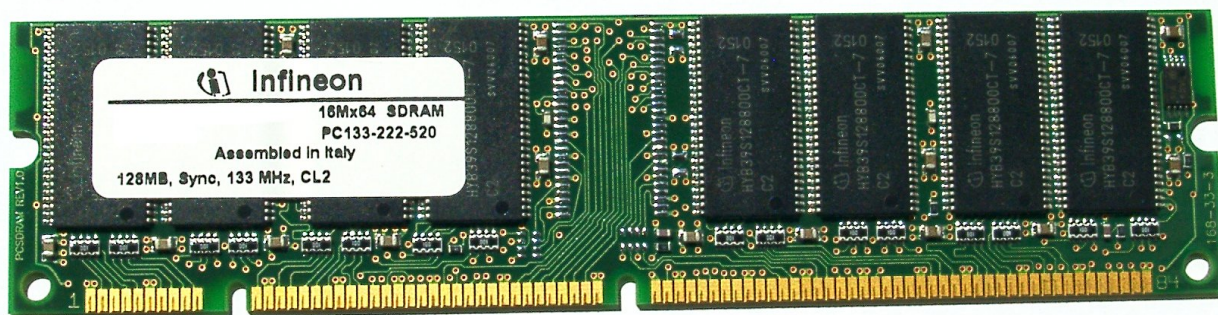


Рисунок 15: Модули памяти SDRAM (DIMM 168) и DDR SDRAM (DIMM 184)

Компании, разрабатывавшие DDR SDRAM, выбрали для ее обозначения пиковую пропускную способность.

Т.е. DDR-1600 работает вовсе не 1.6 GHz (что пока является недостижимым идеалом), а всего лишь на 100 MHz. Соответственно, DDR- 2100 работает на частоте 133 MHz.

Спецификация	Тактовая частота шины памяти	Максимальная теоретическая пропускная способность памяти	
		в одноканальном режиме	в двухканальном режиме
PC1600* (DDR200)	100 МГц	1600 Мбайт/сек	3200 Мбайт/сек
PC2100* (DDR266)	133 МГц	2133 Мбайт/сек	4267 Мбайт/сек
PC2400 (DDR300)	150 МГц	2400 Мбайт/сек	4800 Мбайт/сек
PC2700* (DDR333)	166 МГц	2667 Мбайт/сек	5333 Мбайт/сек
PC3200* (DDR400)	200 МГц	3200 Мбайт/сек	6400 Мбайт/сек
PC3500 (DDR433)	217 МГц	3467 Мбайт/сек	6933 Мбайт/сек
PC3700 (DDR466)	233 МГц	3733 Мбайт/сек	7467 Мбайт/сек
PC4000 (DDR500)	250 МГц	4000 Мбайт/сек	8000 Мбайт/сек
PC4300 (DDR533)	267 МГц	4267 Мбайт/сек	8533 Мбайт/сек

Рисунок 16: Спецификации модулей DDR SDRAM

DDR2 SDRAM

DDR2 SDRAM пришла на смену памяти DDR SDRAM.

Как и DDR SDRAM, DDR2 SDRAM использует передачу данных по обоим фронтам тактового сигнала, за счёт чего при такой же частоте шины памяти, как и в обычной SDRAM, можно фактически удвоить скорость передачи данных (например, при работе DDR2 на частоте 100 МГц эквивалентная эффективная частота для SDRAM получается 200 МГц).

Основное отличие DDR2 от DDR — **вдвое большая частота работы шины**, по которой данные передаются в буфер микросхемы памяти. При этом, чтобы обеспечить необходимый поток данных, осуществляется **передача на шину четырёх бит одновременно**.

Итоговые задержки оказываются выше, чем для DDR.

Внешнее отличие модулей памяти DDR2 от DDR - 240 контактов (по 120 с каждой стороны). DDR2 не является обратно совместимой с DDR, поэтому ключ на модулях DDR2 расположен в другом месте по сравнению с DDR и вставить модуль DDR2 в разъем DDR, не повредив последний (или первый), невозможно. Существуют переходники для установки модулей DDR2 в слоты DDR, но их можно рассматривать скорее как технологический курьез. Дело в том, что для функционирования такого переходника необходим контроллер памяти, обладающий способностью работать как с памятью типа DDR, так и DDR2 — например, Intel 915 Express.

Стандарт DDR2 включает в себя и ряд других усовершенствований, улучшающих различные характеристики нового типа памяти, в том числе — электрические.

Одним из таких новшеств является **внутричиповое терминирование сигнала**. Суть его заключается в том, что для устранения излишнего электрического шума (вследствие отражения сигнала от конца линии) на шине памяти для нагрузки линии используются резисторы не на материнской плате (как это было с предыдущими поколениями памяти), а внутри самих чипов. Эти резисторы деактивируются, когда чип находится в работе и, наоборот, активируются, как только микросхема входит в состояние ожидания. Поскольку

гашение сигнала теперь осуществляется намного ближе к его источнику, это позволяет устранить электрические помехи внутри чипа памяти при передаче данных.

Добавочная задержка (также известная как «отложенная выдача CAS») — еще одно усовершенствование, введенное в стандарт DDR2, которое призвано минимизировать простои планировщика команд при передаче данных из памяти/в память.

Чтобы проиллюстрировать это (на примере чтения), рассмотрим для начала чтение данных с чередованием банков (Bank Interleave) из устройства типа DDR2 с добавочной задержкой, равной нулю, что эквивалентно чтению из обычной памяти типа DDR. При этом может возникнуть ситуация, когда следующая команда открытия банков не может быть отправлена на исполнение, поскольку в данный момент времени еще не закончилось исполнение других команд. В этом случае активация 2-го банка должна быть отложена на один такт. В конечном счете, это приводит к разрыву в последовательности выдачи данных по внешней шине, что снижает реальную пропускную способность памяти.

Для устранения подобной ситуации и увеличения эффективности работы планировщика команд в DDR2 вводится понятие добавочной (дополнительной) задержки. При ненулевом значении устройство памяти отслеживает команды READ (RD_AP) и WRITE (WR_AP), но откладывает их исполнение на время, равное величине добавочной задержки.

Слабым местом технологии DDR2, является более высокая латентность. Именно поэтому основным направлением, на котором сосредоточили свои усилия производители памяти, заинтересованные в популяризации DDR2 SDRAM, стало уменьшение латентности.

Первые модули DDR2-533 SDRAM, которые были доступны с момента анонса чипсетов семейств i925 и i915, имели тайминги 4-4-4 (CAS Latency - RAS to CAS Delay - RAS Prechage Time). Сегодня же многие производители памяти, в особенности производители модулей для энтузиастов, такие как Corsair или OCZ, предлагают модели DDR2 SDRAM, способные работать при частоте 533 МГц с таймингами 3-3-3. Причём, это не разгон, а официально утверждённый JEDEC стандарт. В рамках официально утверждённой спецификации предусмотрена модификация DDR2-533 с таймингами 3-3-3, но с повышенным до 1.9В напряжением питания.

Память	Тайминги	Латентность	Пропускная способность в двухканальном режиме
DDR400 SDRAM	2,5-3-3	12,5 нс	6,4 Гбайт/сек
DDR400 SDRAM	2-3-2	10 нс	6,4 Гбайт/сек
DDR533 SDRAM	3-4-4	11,2 нс	8,5 Гбайт/сек
DDR533 SDRAM	2,5-3-3	9,4 нс	8,5 Гбайт/сек
DDR2-533 SDRAM	5-5-5	18,8 нс	8,5 Гбайт/сек
DDR2-533 SDRAM	4-4-4	15 нс	8,5 Гбайт/сек
DDR2-533 SDRAM	3-3-3	11,2 нс	8,5 Гбайт/сек
DDR2-600 SDRAM	5-5-5	16,6 нс	9,6 Гбайт/сек
DDR2-600 SDRAM	4-4-4	13,3 нс	9,6 Гбайт/сек

Рисунок 17: Сравнение DDR и DDR2

Преимущества DDR2 по сравнению с DDR:

- Более высокая полоса пропускания
- Как правило, меньшее энергопотребление
- Улучшенная конструкция, способствующая охлаждению

Недостатки по сравнению с DDR:

- Обычно более высокая CAS-латентность (от 3 до 6)
- Итоговые задержки при одинаковых (или даже более высоких) частотах оказываются выше

Для использования в ПК, DDR2 RAM поставляется в модулях DIMM с 240 контактами и одним ключом (прорезью в полосе контактов).

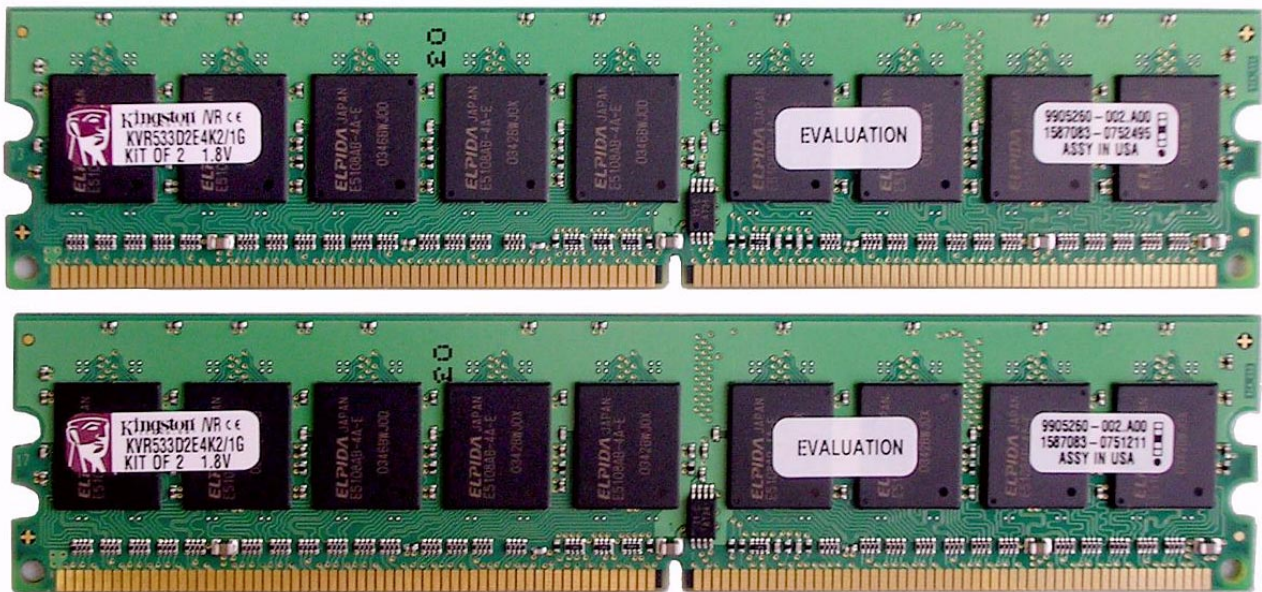


Рисунок 18: Модули памяти DDR2

DIMM'ы различаются по максимальной скорости передачи данных (часто называемой пропускной способностью).

Название модуля	Частота шины	Тип	Пиковая скорость передачи данных
PC2-3200	200 МГц	DDR2-400	3200 МБ/с или 3,2 ГБ/с
PC2-4200	266 МГц	DDR2-533	4200 МБ/с или 4,2 ГБ/с
PC2-5300	333 МГц	DDR2-667	5300 МБ/с или 5,3 ГБ/с ¹
PC2-5400	337 МГц	DDR2-675	5400 МБ/с или 5,4 ГБ/с
PC2-5600	350 МГц	DDR2-700	5600 МБ/с или 5,6 ГБ/с
PC2-5700	355 МГц	DDR2-711	5700 МБ/с или 5,7 ГБ/с
PC2-6000	375 МГц	DDR2-750	6000 МБ/с или 6,0 ГБ/с
PC2-6400	400 МГц	DDR2-800	6400 МБ/с или 6,4 ГБ/с
PC2-7100	444 МГц	DDR2-888	7100 МБ/с или 7,1 ГБ/с
PC2-7200	450 МГц	DDR2-900	7200 МБ/с или 7,2 ГБ/с
PC2-8000	500 МГц	DDR2-1000	8000 МБ/с или 8,0 ГБ/с
PC2-8500	533 МГц	DDR2-1066	8500 МБ/с или 8,5 ГБ/с
PC2-9200	575 МГц	DDR2-1150	9200 МБ/с или 9,2 ГБ/с
PC2-9600	600 МГц	DDR2-1200	9600 МБ/с или 9,6 ГБ/с

Рисунок 19: Спецификации модулей DDR2 SDRAM

DDR3 SDRAM

Уже из названия следует, что DDR3 SDRAM по своему строению и принципам работы не сильно отличается от DDR памяти предыдущих поколений. Собственно, так оно и есть, DDR3 SDRAM – своего рода третье воплощение принципов, заложенных ещё в DDR SDRAM.

Основная идея, позволившая нарастить частоты DDR3 памяти по сравнению с DDR2, заключается в удвоении размера выборки данных, выполняемой непосредственно из устройств хранения информации в буфера ввода-вывода. В то время как в DDR2 SDRAM используется 4-битная выборка, в DDR3 SDRAM применяется **выборка размером 8 бит (называемая также 8n-prefetch)**.

Иными словами, технология DDR3 SDRAM подразумевает двукратное увеличение ширины внутренней шины, соединяющей собственно устройства хранения данных и буфера ввода вывода. В результате, увеличение эффективной частоты передачи данных, происходящее с вводом DDR3 SDRAM, не требует ускорения работы ядра памяти. Возрастает лишь скорость работы внешних буферов. Частота же ядра чипов памяти оказывается в 8 раз меньше частоты внешней шины и буферов DDR3 (в DDR2 эта частота была в 4 раза меньше частоты внешней шины).

Таким образом, достижение DDR3 памятью более высоких эффективных частот по сравнению с DDR2 SDRAM становится возможно практически сразу, без внесения каких-либо изменений и усовершенствований в полупроводниковый технологический процесс. Впрочем, применение описанной техники имеет и обратную сторону – вполне очевидным образом возрастает не только пропускная способность памяти, но и, к сожалению, её латентность. В результате, ожидать от DDR3 SDRAM более высокой скорости работы, чем у DDR2 SDRAM, можно не всегда, даже в том случае, если DDR3 превосходит DDR2 по частоте.

JEDEC в финальной спецификации DDR3 SDRAM определяет несколько версий такой памяти, с частотами от 800 до 1600 МГц. В таблице приведено описание основных параметров перечисленных в спецификации вариантов.

Частота	Тайминги	Пропускная способность в двухканальном режиме	Латентность (T _{ясд})	Срок освоения (год)
DDR3-800	6-6-6	12.8 Гбайт/сек	15.0 нс	2007
DDR3-800	5-5-5	12.8 Гбайт/сек	12.5 нс	2007
DDR3-1066	8-8-8	17.1 Гбайт/сек	15.0 нс	2007
DDR3-1066	7-7-7	17.1 Гбайт/сек	13.1 нс	2007
DDR3-1066	6-6-6	17.1 Гбайт/сек	11.2 нс	2007-2008
DDR3-1333	9-9-9	21.3 Гбайт/сек	13.5 нс	2008
DDR3-1333	8-8-8	21.3 Гбайт/сек	12.0 нс	2008
DDR3-1333	7-7-7	21.3 Гбайт/сек	10.5 нс	2008-2009
DDR3-1600	10-10-10	25.6 Гбайт/сек	12.5 нс	2009
DDR3-1600	9-9-9	25.6 Гбайт/сек	11.3 нс	2009-2010
DDR3-1600	8-8-8	25.6 Гбайт/сек	10.0 нс	2009-2010

Рисунок 20: Параметры памяти DDR3

Если учесть, что латентность распространённой DDR2-800 с таймингами 4-4-4 составляет 10 нс, то эффективность DDR3 SDRAM действительно можно поставить под вопрос. Получается, что эта память способна выигрывать у предшественницы исключительно за счёт увеличения пропускной способности, которая должна компенсировать ухудшающуюся латентность.

К сожалению, переход на использование DDR3 SDRAM – мера вынужденная. DDR2 память уже исчерпала свой частотный потенциал: если увеличение её частоты до 1066 МГц ещё возможно при некоторых допущениях, то дальнейший рост скорости резко снижает

выход годных чипов и значительно повышает стоимость модулей DDR2 SDRAM. Именно поэтому JEDEC не стал стандартизировать DDR2 память с частотами выше 800 МГц, ратуя за переход к DDR3 технологии.

Впрочем, к счастью, DDR3 SDRAM привносит и ряд других полезных усовершенствований, которые должны позволить сделать более однозначный вывод в пользу новинки не только с позиции производителей, но и с точки зрения конечного пользователя. Среди плюсов в первую очередь следует отметить снизившееся напряжение питания модулей DDR3 SDRAM, достигшее 1.5 В. Это на 20% ниже напряжения DDR2 SDRAM, что в конечном итоге выливается в примерно 30-процентное падение энергопотребления при сравнении с DDR2 памятью, работающей на аналогичной тактовой частоте. Этот эффект достигается и благодаря внедрению производителями чипов памяти более современных технологических процессов.

Некоторые изменения претерпел и сигнальный протокол DDR3 SDRAM, усовершенствованный в связи с очередным и значительным ростом частоты шины памяти. Теперь для передачи адресов и команд, а также управляющих и стробирующих команд используется fly-by архитектура с терминированием сигналов непосредственно на модуле. Это значит, что сигналы подаются на все микросхемы модуля не одновременно, а последовательно.

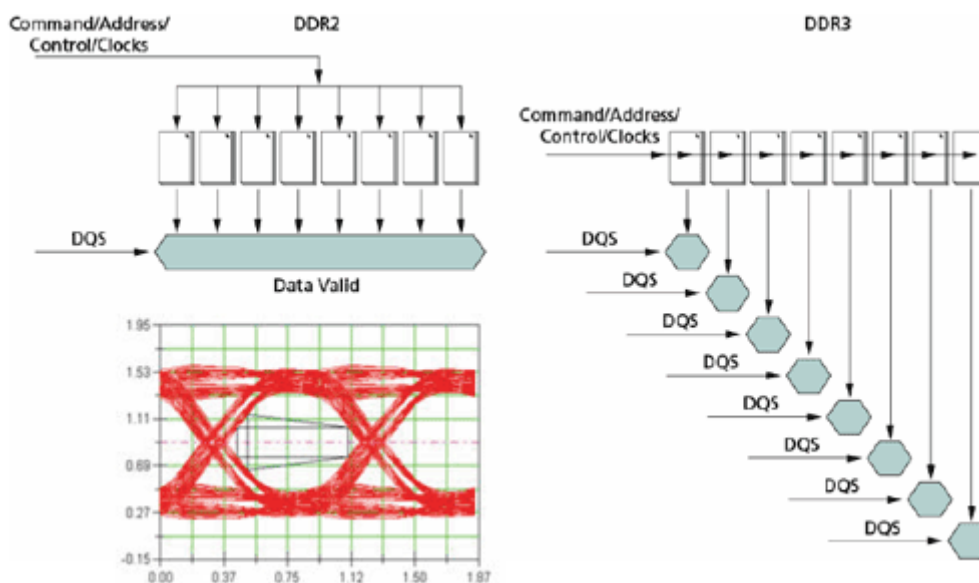


Рисунок 21: Передача команд микросхемам памяти DDR2 и DDR3

Соответственно, изменилась и тактика чтения/записи данных. Контроллер DDR3 должен быть способен успешно распознавать и обрабатывать временные смещения при поступлении данных с чипов, вызванные применением fly-by архитектуры передачи команд.

	DDR2 SDRAM	DDR3 SDRAM
Скорость передачи данных	400-800 Мбит/сек	800-1600 Мбит/сек
Напряжение питания	1.8 В ± 0.1 В	1.5 В ± 0.075 В
Число банков	4/8	8
Упреждающая выборка	4 бита	8 бит
Burst Length	BL=4, 8	BL =4, 8
Топология	Conventional T	Fly-by
CAS Latency	3-5	6-10
Упаковка чипов	60 BGA для x4/x8 84 BGA для x16	78 BGA для x4/x8 96 BGA для x16
Ёмкость чипов	256 Мбит – 4 Гбита	512 Мбит – 8 Гбит

Рисунок 22: Параметры DDR2 и DDR3

Учитывая различное напряжение питания и отличающиеся протоколы DDR2 и DDR3 SDRAM, память этих двух типов логически не совместима друг с другом. Хотя число контактов у модулей DDR2 и DDR3 одинаково и равно 240, слоты, предназначенные для памяти разного типа, отличаются расположением "ключа". Поэтому, установить DDR3 SDRAM в DDR2 DIMM и наоборот не удастся.



Рисунок 23: Модули памяти DDR3 (вверху) и DDR2

DDR3L SDRAM

Прогресс в полупроводниковой отрасли позволил проектировать более экономичные микросхемы оперативной памяти, отличающиеся пониженным напряжением питания ядра и схем ввода/вывода. Ассоциация JEDEC, определяющая ряд стандартов в электронной промышленности, приняла стандарт памяти **DDR3L, который предусматривает питание от шины 1,35 В**, что позволяет снизить потребляемую мощность примерно на 20% по сравнению с традиционными модулями DDR3 (1,5 В).

Память DDR3L полностью совместима с устройствами, ориентированными на работу с DDR3. Пользователи могут безо всякого риска использовать новые модули в современных материнских платах – стандарт DDR3 предусматривает использование так называемых SPD-устройств, которые представляют собой микросхему EEPROM с «защитными» данными о характеристиках модуля памяти, в том числе и максимально допустимом напряжении питания.

DDR4 SDRAM

Ассоциация JEDEC начала разработку стандарта еще в 2005 году. Как говорится в пресс-релизе JEDEC, при разработке инженеры пытались достичь наибольшей производительности и надежности, увеличив при этом энергоэффективность новых модулей.

Как и предшественники, DDR4 переняла архитектуру 2ⁿ-prefetch, которую JEDEC в данном случае называет 8n-Prefetch. **Любой чип памяти может состоять из двух или четырех отдельных групп банков.**

8Gb x4 DDR4 чип обычно состоит из 4 групп банков, по 4 банка в каждой группе. Каждый банк такого чипа содержит 131.072 (2¹⁷) строк (rows), по 512 байтов каждая. Для сравнения 8Gb x4 DDR3 чип содержит 8 независимых банков, 65.536 (2¹⁶) строк на банк, по 2048 байтов в каждой строке. **При равном объеме, у DDR4 чипа в два раза больше банков и гораздо короче строки памяти.** Это означает, что новая память может переключаться между банками памяти гораздо быстрее, чем это делала DDR3. В частности, для 8Gb x4 DDR4 чипов, заявленных как 1600 MT/s compatible, показатель tFAW(Four-bank Activation

Window) равен 20ns, что вдвое меньше, чем у DDR3 (40ns). Это означает, что **DDR4 чипы памяти могут открывать произвольные строки в разных банках в два раза быстрее, чем DDR3.**

Изменена архитектура доступа к модулям. Раньше шина Multi-Drop имела всего два канала и даже при организации работы с четырьмя модулями памяти, они висели попарно на одном канале, что не всегда положительно сказывалось на производительности.

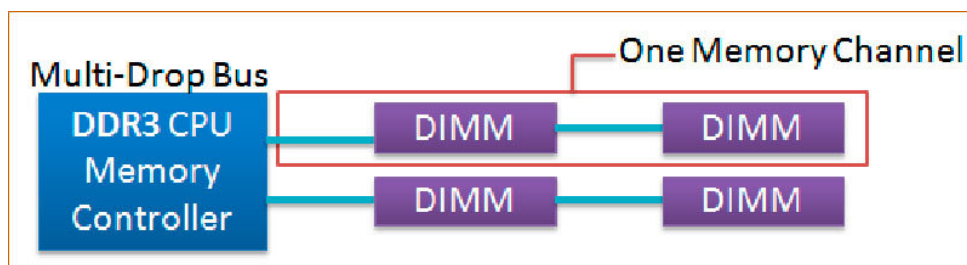


Рисунок 24: Шина Multi-Drop

Новая шина **Point-to-Point** будет связывать каждый канал с одним модулем памяти. То есть при наличии в процессоре двухканального контроллера памяти — будут доступны два слота, а при наличии четырехканального — четыре. Для системных плат с 8-ю слотами под модули памяти применяются цифровые коммутаторы — аналогичные по смыслу, тем, что разветвляют линии PCI Express. Таким образом, оперативная память переходит на использование параллельного доступа.

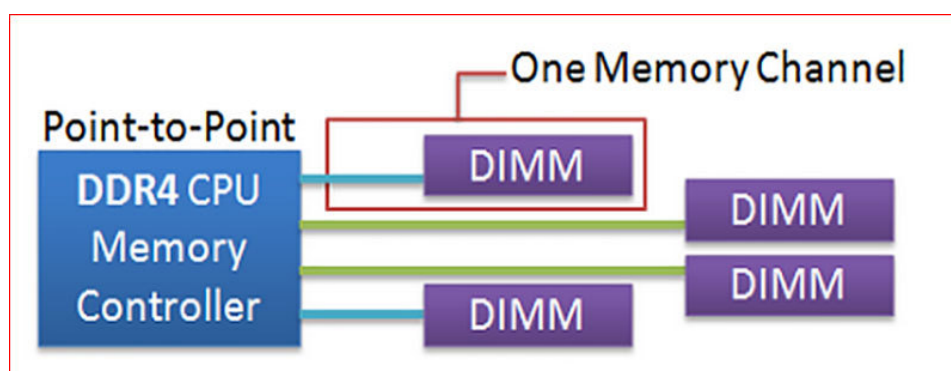


Рисунок 25: Шина Point-to-Point

Для DDR4, используя 16Gb кристаллы, и восьмислойную упаковку кристаллов в чип, теоретически можно создать модуль памяти объемом до 512GB. Количество контактов на модулях DDR4 увеличилось до 284, чтобы адресовать такой объем памяти. Каждый чип DDR4 памяти может представлять собой стек из 2, 4 или 8 кристаллов DRAM. Стек из 8 слоёв описан в дополнениях к спецификации и скорее всего потребует использования TSV (through silicon via) для своей практической реализации. Другая важная часть спецификации DDR4 – **повышение энергоэффективности** по сравнению с DDR3. Кроме снижения напряжения на I/O с 1.35V до 1.2V, новый стандарт также специфицирует использование более высокого уровня напряжения внутри чипов (DRAM word line 2.5V), что обеспечивает быстрый доступ в активном режиме и малый ток утечки в пассивном.

Изменилась и электрическая реализация интерфейса ввода-вывода данных. Новый интерфейс носит название **pseudo-open drain (POD, «псевдо-открытый сток»)** и его основное отличие в том, что в схеме не протекает ток, когда на линии установлен высокий уровень напряжения.

Уменьшение напряжения на I/O, изменения электрического интерфейса и уменьшение длины строк в банках памяти приводят к существенному сокращению энергопотребления по сравнению с DDR3. Предварительные оценки говорят о 30% выигрыше. Хотя, разумеется, это зависит от характера обращений к памяти, техпроцесса и многих других факторов. Такой выигрыш может использоваться для того, чтобы увеличить тактовую частоту и, соответственно, скорость работы, или для того, чтобы сэкономить немного энергии при той же производительности.

Много немаловажных изменений относятся к надежности DDR4. Например, спецификация говорит об обнаружении и коррекции чипами памяти ошибок, связанных с контролем четности команд и адресов.

Другой пример – то, что чипы DDR4 имеют режим тестирования соединений. Этот режим позволяет контроллеру памяти проверять электрические связи (и находить «обрыв» линий), гораздо быстрее, чем раньше, и без использования инициализирующих последовательностей.

Также модуль DDR4 может быть сконфигурирован так, чтобы отбрасывать команды, содержащие ошибки контроля четности. В DDR3 такие команды пропускались и доходили до чипов памяти, многократно усложняя восстановление после сбоев.

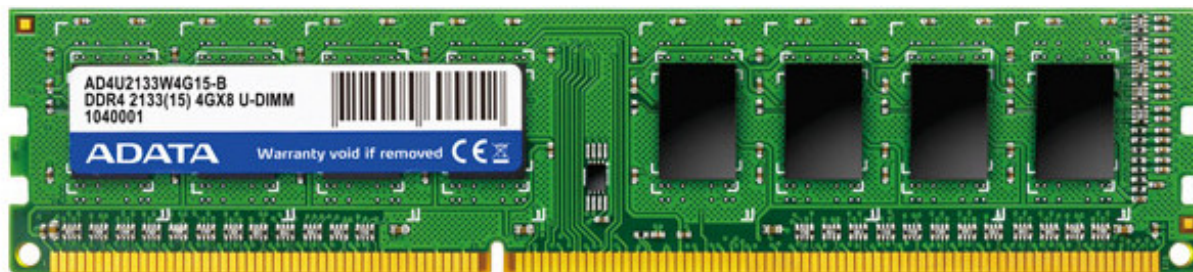


Рисунок 26: Модуль памяти DDR4 SDRAM

Форм-факторы модулей памяти

SIMM-30

Число контактов: 30

Разрядность шины данных: 8 бит (9 бит у модулей с контролем четности)

Тип применяемых микросхем динамической памяти: FPM

Стандартные значения объема памяти модулей: 64 Кб, 256 Кб, 1 Мб, 4 Мб, 16 Мб

Физические размеры модуля: 90 на 13 или 90 на 25 мм

Шаг расположения контактных площадок — 0,1"

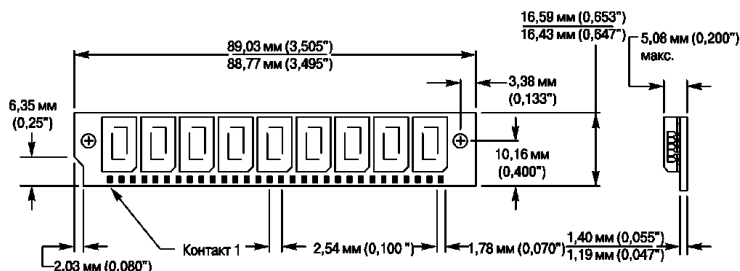


Рисунок 27: Параметры модулей форм-фактора SIMM-30



Рисунок 28: Модуль SIMM-30

SIMM-72

Число контактов: 72

Разрядность шины данных: 32 бита (36 бит у модулей с контролем четности)

Тип применяемых микросхем динамической памяти: FPM, EDO

Стандартные значения объема памяти модулей: 1 Мб, 2 Мб, 4 Мб, 8 Мб, 16 Мб, 32 Мб, 64 Мб, 128 Мб

Физические размеры модуля: в большинстве случаев 107,95 на 25 мм, иногда 107,95 на 39 мм

Шаг расположения контактных площадок — 0,05"; между выводами 36 и 37 увеличенное расстояние и вырез (ключ)

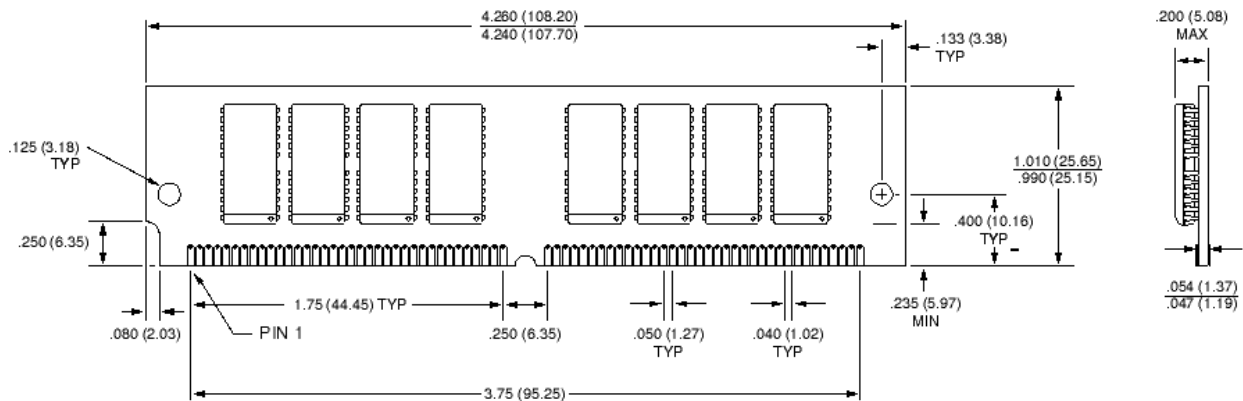


Рисунок 29: Параметры модулей форм-фактора SIMM-72



Рисунок 30: Модуль SIMM-72

Для установки модуля SIMM его нужно:

1. правильно ориентировать относительно слота
2. под углом вставить модуль в слот
3. повернуть до щелчка

Для извлечения нужно:

1. отжать боковые защелки
2. повернуть модуль



Рисунок 31: Установка модуля SIMM

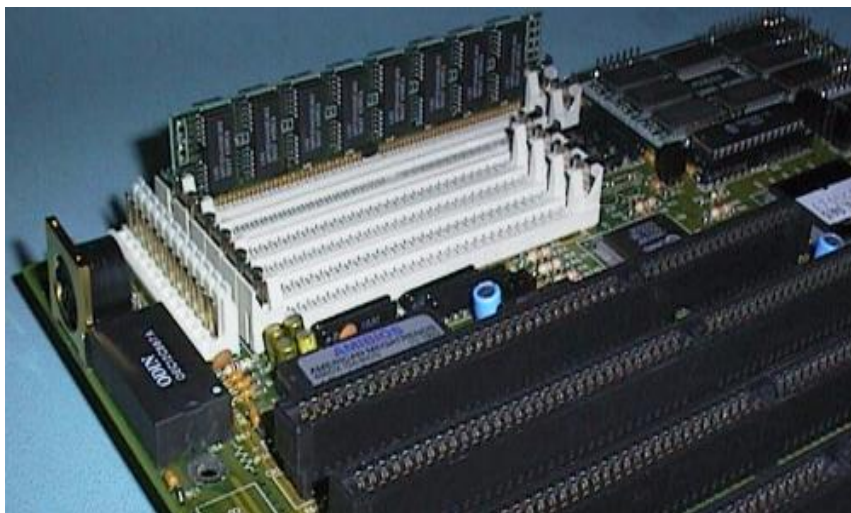


Рисунок 32: Фрагмент системной платы с слотами для модулей SIMM-30 и SIMM-72

DIMM-168

Число контактов: 168

Используется для SDR SDRAM (реже для FPM/EDO DRAM в рабочих станциях/серверах) Модули могут быть рассчитаны на 5V или 3,3V. Для того, чтобы избежать путаницы, у этих модулей различается расположение среднего паза — Voltage Key.

Второй паз (ключ) показывает с буфером или без буфера модуль памяти.

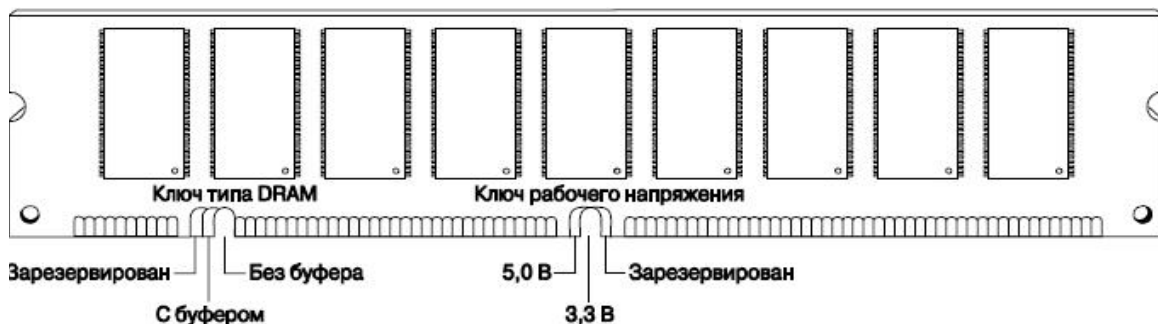


Рисунок 33: Ключи модуля памяти DIMM-168

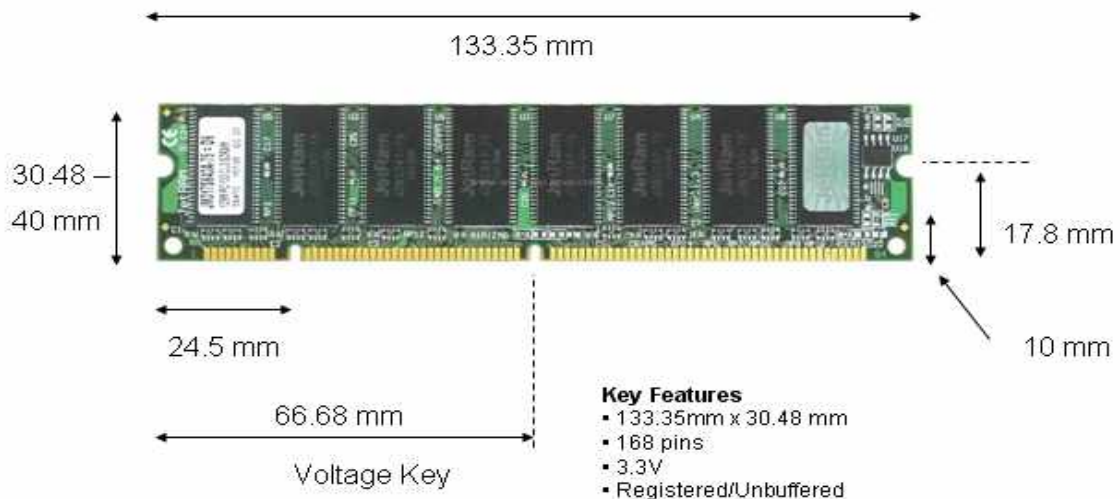
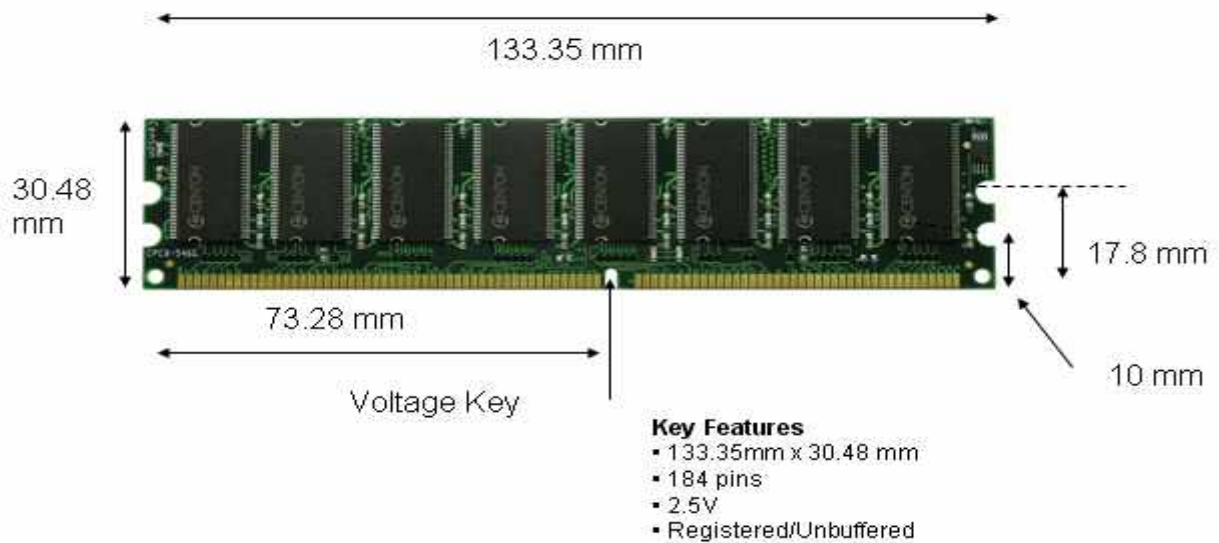


Рисунок 34: Модуль DIMM-168

DIMM-184

Число контактов: 184

Используется для DDR SDRAM



DIMM-240

Число контактов: 240

Используется для DDR2 SDRAM, DDR3 SDRAM и FB-DIMM DRAM

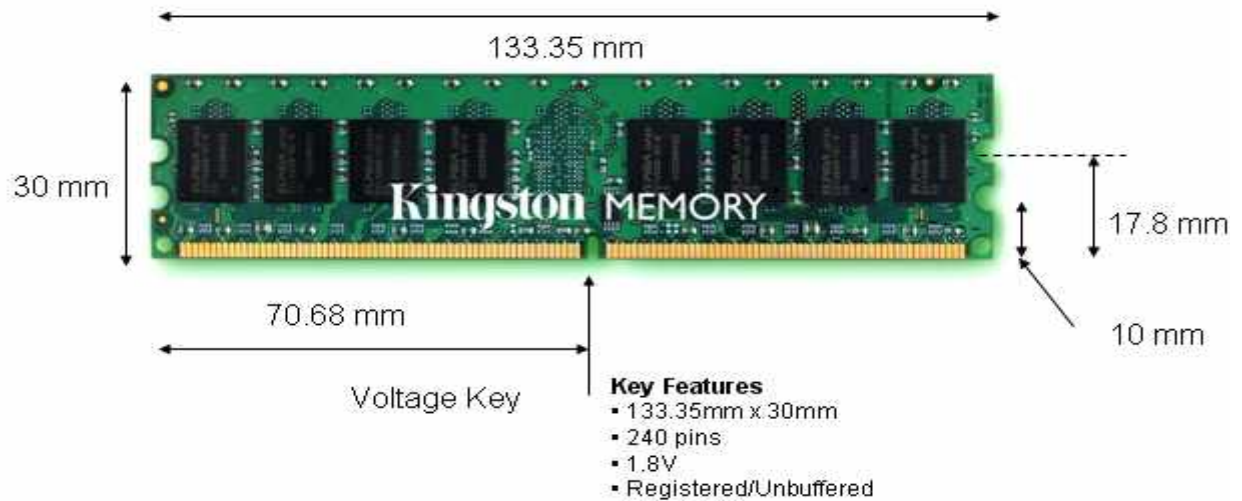


Рисунок 35: Модуль DIMM-240 для DDR2 SDRAM

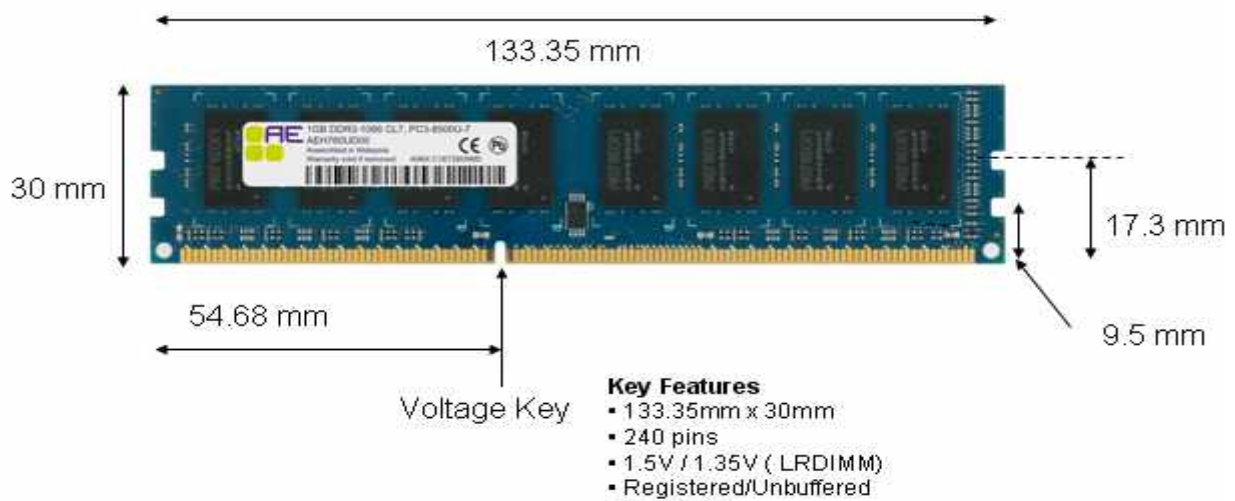


Рисунок 36: Модуль DIMM-240 для DDR3 SDRAM

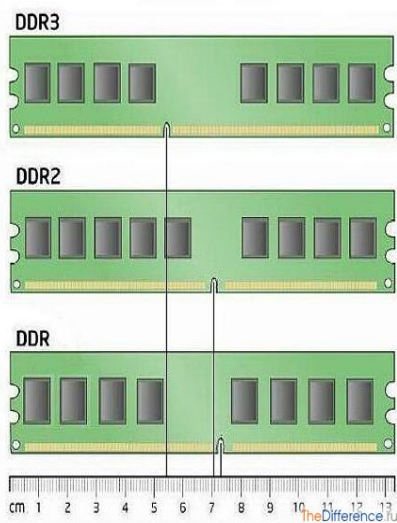


Рисунок 37: Сравнение модулей для памяти DDR, DDR2, DDR3

DIMM-288

Число контактов: 288

Используется для DDR4 SDRAM

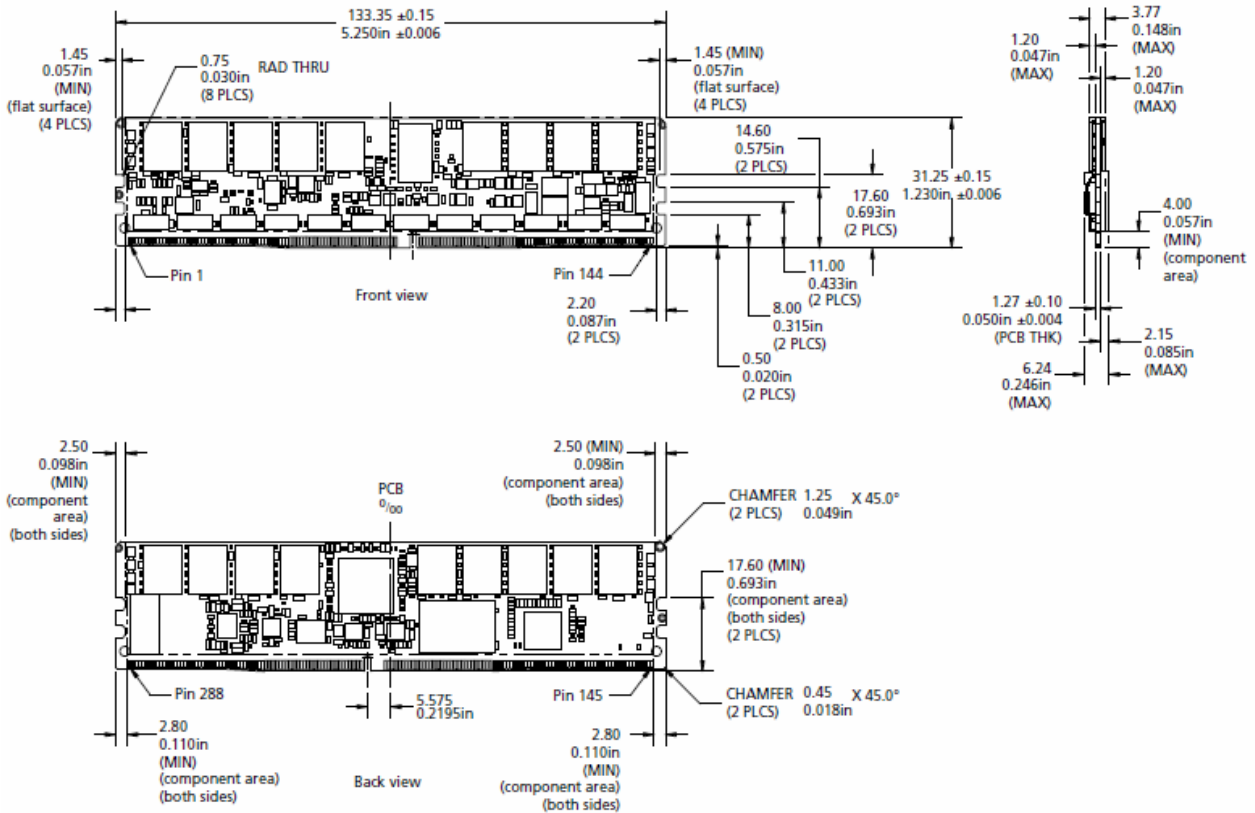


Рисунок 38: Параметры модуля форм фактора DIMM-288

SO-DIMM -144

Используется для SDR SDRAM в портативных компьютерах

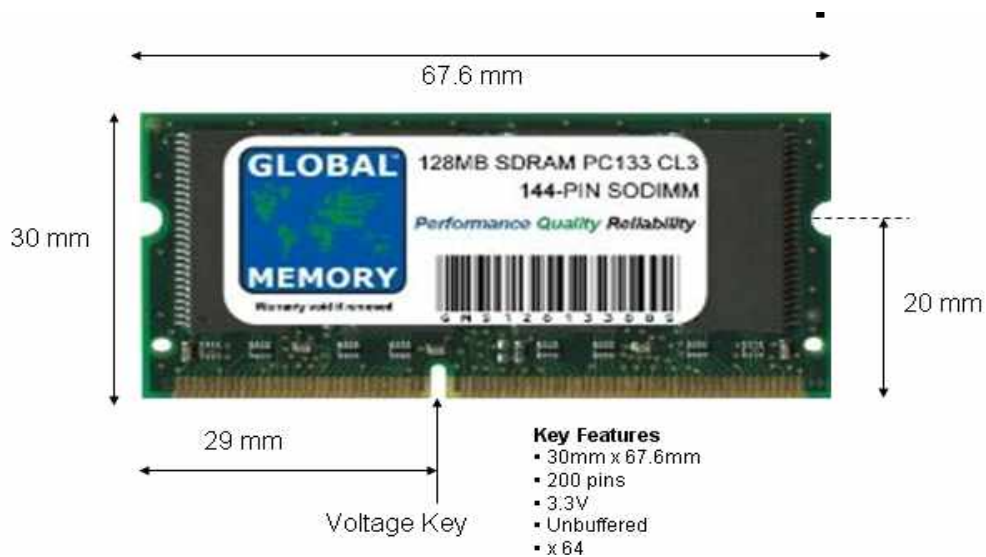


Рисунок 39: SO-DIMM 144

SO-DIMM -200

Используется для DDR SDRAM и DDR2 SDRAM в портативных компьютерах

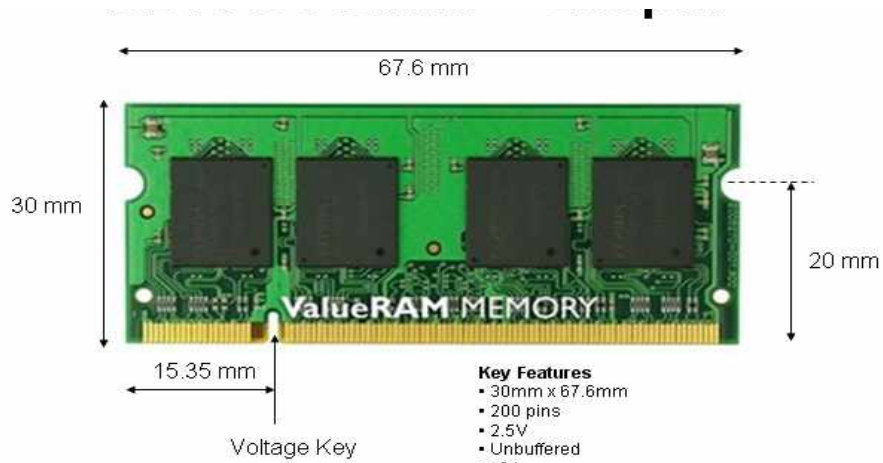


Рисунок 40: SO-DIMM 200 для DDR SDRAM

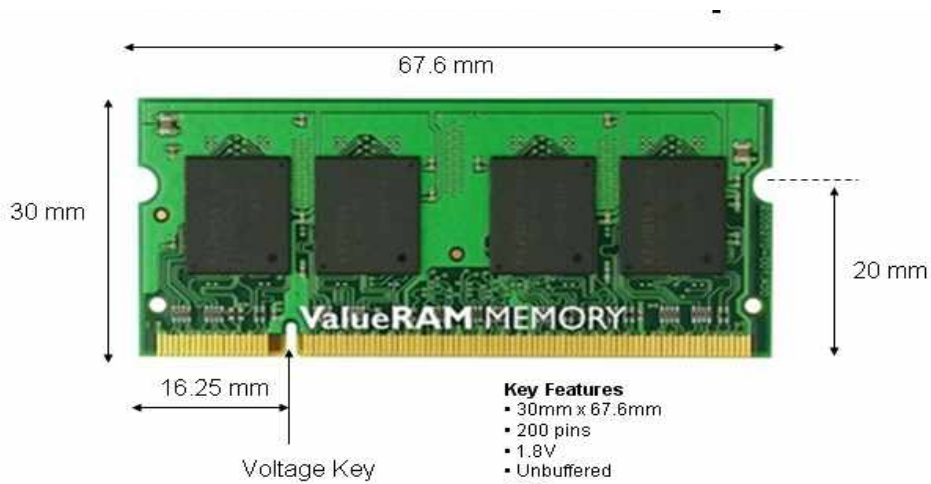


Рисунок 41: SO-DIMM 200 для DDR2 SDRAM

SO-DIMM -204

Используется для DDR3 SDRAM

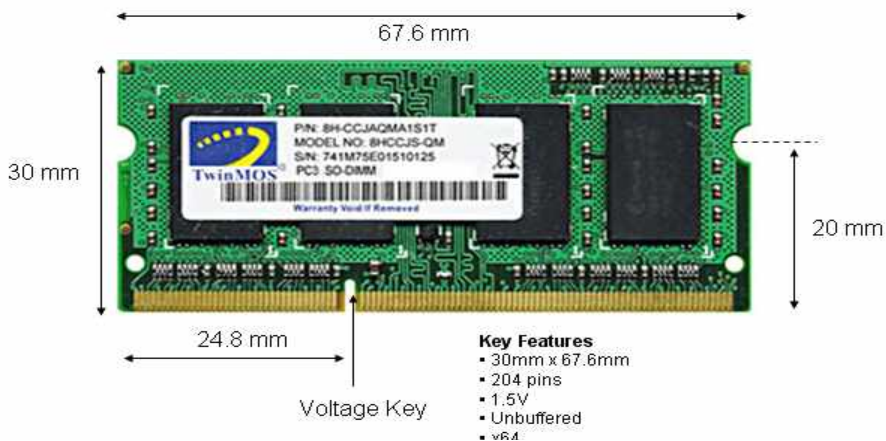


Рисунок 42: O-DIMM 204 для DDR3 SDRAM

Двухканальный режим Dual-channel

Двухканальный режим (англ. *Dual-channel architecture*) — режим работы оперативной памяти (RAM) и её взаимодействия с материнской платой, процессором и другими компонентами компьютера, при котором может быть увеличена скорость передачи данных между ними за счёт использования двух каналов для доступа к объединённому банку памяти (это можно проиллюстрировать на примере ёмкостей, через горлышко одной из которых жидкость может выливаться дольше, чем из двух других с такими же общим суммарным объёмом и горлышками, но с большей пропускной способностью — двумя горлышками).

Таким образом, система при использовании, например, двух модулей памяти в двухканальном режиме может работать быстрее, чем при использовании одного модуля, равного их суммарному объёму.

При теоретическом увеличении пропускной способности памяти в 2 раза, тесты показывают, что на практике прирост составляет порядка 5—10 % в играх, и от 20 % до 70 % в тяжёлых графических приложениях, которые по максимуму используют оперативную память и обрабатывают графику в больших разрешениях (Photoshop, CorelDRAW и другие программы).

Правила включения двухканального режима

Двухканальный режим может быть получен при использовании чётного числа модулей DIMM.

Для включения двухканального режима необходимо выполнить следующие условия:

- одинаковая конфигурация модулей DIMM на каждом канале;
- одинаковая ёмкость (128 МБ, 256 МБ, 512 МБ и т. п.). Необязательное условие;
- каналы памяти А и В должны быть идентичны;
- на большинстве материнских плат (за редким исключением) должны быть заполнены симметричные разъемы памяти (разъем 0 или разъем 1)

То есть в двухканальном режиме будет работать память одной частоты, одного типа. Память разного объёма будет работать, если выполняются все условия.

Примечание:

Прирост производительности от использования двухканального режима работы памяти зависит от:

- типа памяти;
 - таймингов, задержек памяти;
 - типа чипсета материнской платы или типа контроллера памяти;
 - частоты работы памяти
- и ряда других факторов.

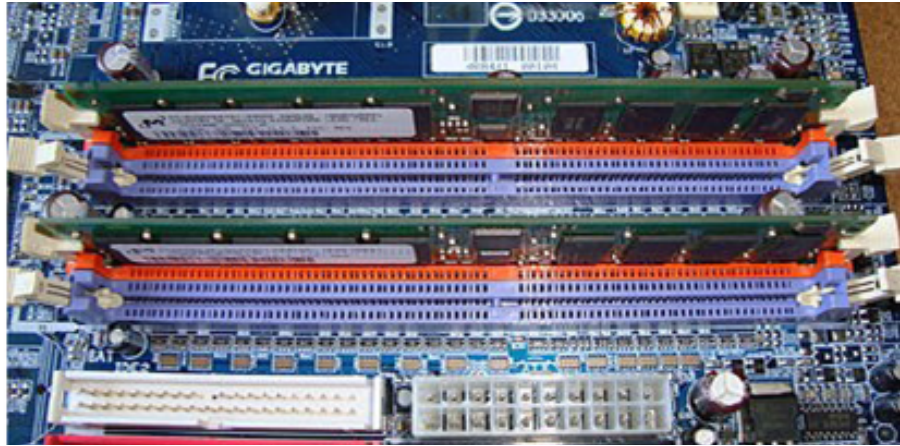


Рисунок 43: Фрагмент системной платы с поддержкой двухканального режима

Выбор модулей оперативной памяти

При выборе модулей RAM следует учесть следующие параметры:

- Тип памяти поддерживаемый контролером в вашей системе – точное совпадение
- Тип разъема на системной плате– точное совпадение
- Частоту шины памяти - желательно чтобы у модулей была не ниже чем у системной платы. У модулей установленных в один канал частота шины должна совпадать
- Производительность модулей – желательно выбирать модули с одинаковой производительностью
- Максимальный объем на которые рассчитан контроллер памяти
- Тайминги
- Нагрев и необходимость охлаждения
- Производитель

Источники

Андрей Борзенко DDR SDRAM: удвоение скорости <http://www.computer-museum.ru/technlgy/ddrsdram.htm>

DDR2 против DDR: реванш состоялся <http://fcenter.ru/online/hardarticles/motherboards/11076>

DDR3 SDRAM: революция или эволюция?

<http://fcenter.ru/online/hardarticles/motherboards/21754>

Разработка → Зачем переходить на DDR4?: <https://habrahabr.ru/company/intel/blog/205608/>

DDR3 против DDR4. Теоретические различия

https://geektimes.ru/company/kingston_technology/blog/247670/

Немного об оперативной памяти: <http://sysadm.pp.ua/interestnoe/ram.html>