

Организация БИС ЗУ с произвольным доступом

Запоминающие устройства (ЗУ) характеризуются рядом параметров, определяющих возможные области применения различных типов таких устройств. К основным параметрам, по которым производится наиболее общая оценка ЗУ, относятся их информационная емкость (E), время обращения (T) и стоимость (C).

Состав и структура микросхем оперативных ЗУ в процессе совершенствования технологий их изготовления подверглись определенным изменениям. **Первые полупроводниковые оперативные ЗУ строились на схемах малой и средней степени интеграции и включали в себя несколько различных типов микросхем: собственно матрицы элементов памяти, усилители чтения-записи, дешифраторы и, при необходимости, регистры (адреса и данных).**

Позднее, с появлением больших интегральных схем (БИС) и повышением частоты их работы, использование отдельных типов микросхем перестало быть оправданным по следующим причинам. *Во-первых*, количество элементов памяти в матрицах возросло настолько, что число выводов, требующееся для выбора элемента памяти и равное сумме количества строк и количества столбцов матрицы, стало очень большим (несколько тысяч). *Во-вторых*, длина соединений между микросхемами больше, чем длина соединений внутри микросхемы, что увеличивает время прохождения сигнала и реактивные составляющие (емкость и индуктивность, или перекрестные помехи), а следовательно, уменьшает быстродействие памяти.

Поэтому микросхемы памяти стали включать в себя не только элементы памяти, но и всю остальную электронику управления: дешифраторы, усилители, буферные регистры, схемы управления. Такой состав БИС памяти приводил к известной аппаратной избыточности строящихся на их основе модулей памяти.

Действительно, первые БИС памяти имели логическую организацию вида N одноразрядных слов, или $N \times 1$, где N - количество адресов (одноразрядных слов) микросхемы. Следовательно, каждый разряд модуля памяти, построенного на таких микросхемах, включал свои собственные дешифраторы, буферные регистры и схемы управления, одного комплекта которых при традиционной организации было достаточно для целого модуля. Однако такое дублирование оправдывалось достигаемыми характеристиками памяти, а его стоимость не была чрезмерной (электроника обрания составляла не более 5-15% от общей площади кристалла микросхемы).

Впоследствии разрядность хранимых в микросхеме слов была увеличена и составляет на сегодня от 4-х до 16-ти разрядов, т.е. $N \times 4$, $N \times 8$, $N \times 16$. Понятно, что относительная доля избыточных схем обрания при этом падает.

На функциональных схемах микросхема памяти изображается обычным прямоугольником с левым и правым полями, как показано на рис. 9. Микросхема имеет три группы входов: адресные входы, вход(ы) данных и управляющие входы.

Количество адресных входов ($A_0 \div A_k$) определяется емкостью и организацией микросхемы памяти, а также способом подачи адреса. Нетрудно видеть, что емкость микросхемы E_{Cx} , равная произведению количества адресов (слов) N на разрядность хранимых слов n , не определяет однозначно требуемое число адресных входов. Для адресации любого из N слов требуется адрес разрядностью $\log_2 N$. Например, для адресации микросхемы емкостью $E_{Cx} = 128$ Мбит, имеющей организацию $16M \times 8$ (адресов \times бит), достаточно $\log_2 16M = \log_2 (2^4 \times 2^{20}) = 24$ разряда. Способ подачи адреса также оказывает влияние на количество адресных входов микросхемы. **Так,**

распространенный в динамических оперативных ЗУ прием мультиплексирования адресных входов, состоящий в поочередной подаче на одни и те же адресные входы сначала старшей части (половины) адреса - адреса строки (*Row Address*), а затем - младшей части - адреса столбца (*Column Address*), позволяет уменьшить вдвое количество требующихся адресных входов. Конечно, это несколько увеличивает время обращения к памяти, но оказывается экономически (да и схемотехнически) оправданным.

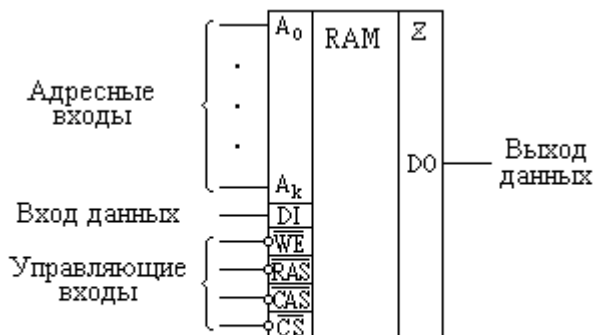


Рис. 9. Условное изображение ОЗУ на функциональных схемах

В статических ЗУ все разряды адреса подаются на адресные входы одновременно.

Количество входов данных (*DI - Data Input*) равно разрядности хранимых слов. Количество выходов данных (*DO - Data Output*) также равно разрядности хранимых слов. Однако во многих случаях входы и выходы данных объединяются, что позволяет уменьшить вдвое количество выводов данных у микросхем памяти, а также упростить их подключение к шинам данных.

Для этого выходы микросхем памяти (или объединенные входы/выходы) обычно имеют специальный выходной каскад, позволяющий подключать к одной шине выходы нескольких микросхем без использования дополнительных сборок ИЛИ. Есть два варианта организации таких выходов: выход с тремя устойчивыми состояниями (или z-выход) и выход с открытым коллектором. Тип выхода отмечается специальным значком в верхней части правого поля изображения микросхемы. На рис. 9 показан z-выход. Выход данных, реализованный по схеме с открытым коллектором, как правило, инверсный.

Управляющие входы могут заметно различаться как по назначению, так и по обозначениям для разных типов микросхем памяти.

Во всех случаях присутствует вход управления режимом обращения: чтение или запись. Частым его обозначением является *WE#* (*Write Enable - разрешение записи*). Вход этот обычно инверсный (это и обозначает символ #), т.е. режим записи включается при нулевом значении сигнала на данном входе, а при единице на входе производится чтение.

Другим общим сигналом, имеющимся почти во всех микросхемах, является сигнал выбора микросхемы - *CS#* (*Chip Select*). Этот вход также обычно является инверсным и при единичном значении на нем микросхема переходит в "выключенное" состояние (выход данных микросхемы переходит в состояние высокого выходного сопротивления, если он является z-выходом, или в состояние "1", если это инверсный выход с открытым коллектором). При нулевом значении сигнала на входе *CS#* микросхема находится в активном состоянии.

В динамических ОЗУ при мультиплексировании адресных входов используются два управляющих входа сигналов строба: *RAS#* (*Row Address Strobe* - строб адреса строки) и *CAS#* (*Column Address Strobe* - строб адреса столбца, или колонки). Сигналы на этих входах переводятся в активное состояние (в "0") в тот момент, когда на адресных входах установлен адрес строки или адрес столбца соответственно.

Структурная схема БИС динамического ОЗУ показана на рис. 10. Основными ее компонентами являются четыре банка памяти, представляющих собой матрицы элементов памяти с дешифраторами строк и столбцов и усилителями чтения-записи. Кроме собственно банков памяти в состав ОЗУ входят:

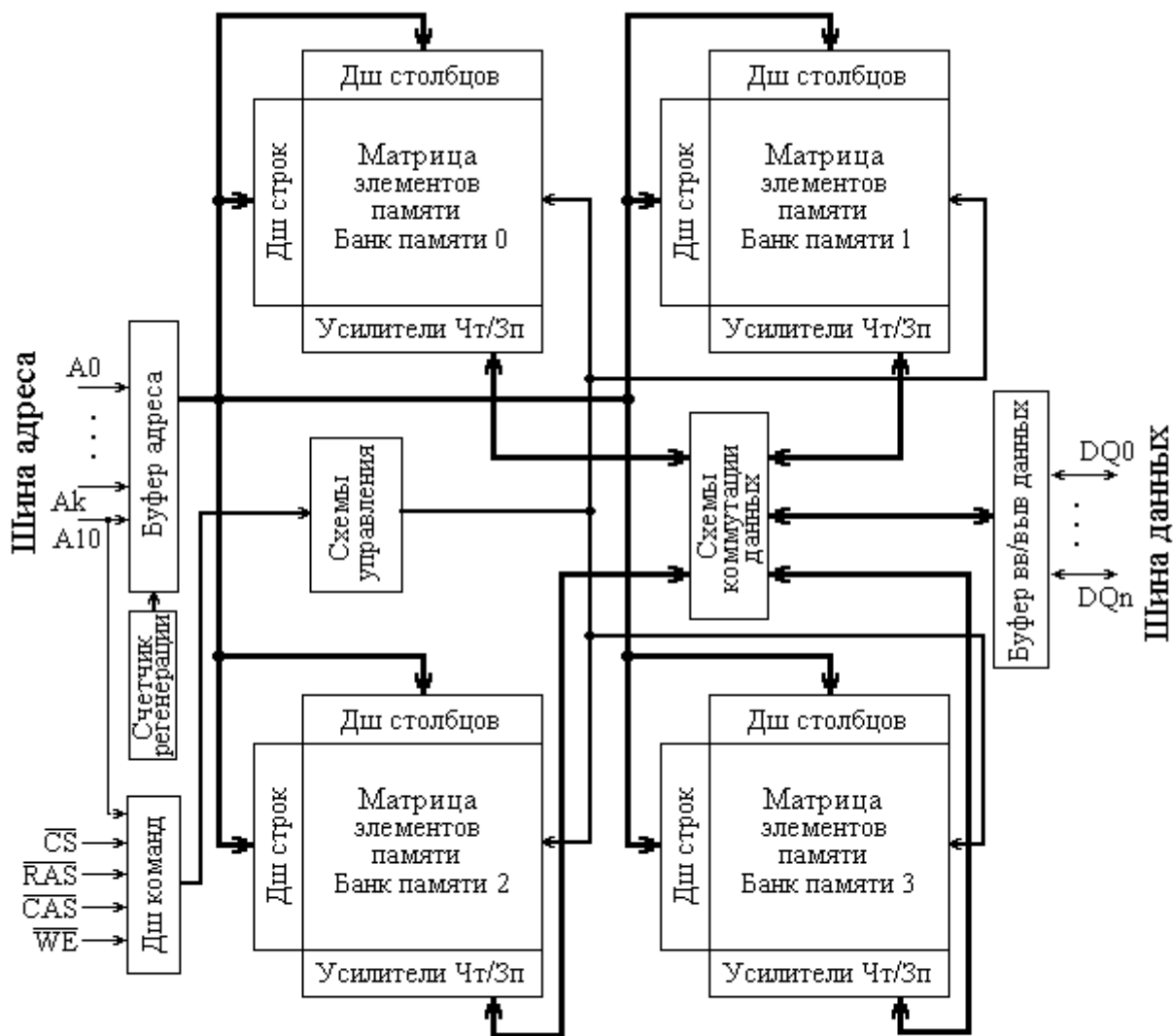


Рис. 10. Структурная схема БИС ОЗУ

- буфер адреса, фиксирующий адреса строки и столбца;
- счетчик регенерации, формирующий адрес строки, в которой должна выполняться очередная регенерация;
- дешифратор команд, определяющий, какое действие (команду) должна выполнить микросхема в соответствии с поданными управляющими сигналами (и сигналом A_{10});
- схемы управления, формирующие управляющие сигналы для остальных узлов микросхемы;
- схемы коммутации данных, передающие читаемые или записываемые данные из/в банки памяти;
- буфер ввода/вывода данных, обеспечивающий связь микросхемы памяти с шиной данных. Независимо от того, какова конкретная модификация

динамической памяти, запоминающие конденсаторы ее запоминающих элементов разряжаются из-за наличия токов утечки. Постоянная разряда, как известно, зависит от емкости запоминающего конденсатора и сопротивления цепи тока утечки и может различаться для разных модификаций. Время, в течение которого информация сохраняется в элементе памяти, составляет до нескольких десятков миллисекунд. Это приводит к необходимости периодического (с периодом не больше, чем время сохранения информации) восстановления зарядов емкостей. Такая процедура и получила название регенерации (*refresh*) динамической памяти. Выполняется она одновременно для целой строки матрицы (банка) элементов памяти, поскольку регенерировать информацию по элементам или по словам (по 8 байт) слишком долго.

Адресный сигнал A_{10} выделен среди других адресных линий, так как он имеет специальное назначение: при подаче адреса столбца этот сигнал указывает на особенности выполнения последующего (пакетного) чтения или записи, задавая (при единичном значении) режим так называемого автоматического подзаряда банка памяти. <http://www.ord.com.ru/files/book3/p21.html>

Статические ЗУ с произвольным доступом

В статических ЗУ (*Static Random Access Memory – SRAM*) в качестве элемента памяти используется триггер, что, конечно, сложнее, чем конденсатор с транзисторным ключом динамического ЗУ. Поэтому статические ЗУ обладают меньшей плотностью хранения информации: емкость микросхем статических ЗУ начала 2000-х годов не превосходила 16 Мбит.

Однако триггер со времен первых компьютеров был и остается самым быстродействующим элементом памяти. **Поэтому статическая память позволяет достичь наибольшего быстродействия, обеспечивая время доступа в единицы и даже десятые доли наносекунд, что и обуславливает ее использование в ЭВМ, главным образом, в высших ступенях памяти – кэш-памяти всех уровней.** Главными недостатками статической памяти являются ее относительно высокие стоимость и энергопотребление.

Конечно, в зависимости от используемой технологии, память обладает различным сочетанием параметров быстродействия и потребляемой мощности. Например, статическая память, изготовленная по КМОП-технологии (CMOS), имеет низкую скорость доступа, со временем порядка 100 нс, но зато отличается очень малым энергопотреблением. В ПЭВМ такую память применяют для хранения конфигурационной информации компьютера при выключенном напряжении сети (в этой же микросхеме размещают и часы, отсчитывающие реальное время). Питание такой памяти осуществляется от небольшой батарейки, которая может служить несколько лет.

Основными разновидностями статической памяти (SRAM) с точки зрения организации ее функционирования являются асинхронная (*Asynchronous*), синхронная пакетная (*Synchronous Burst*) и синхронная конвейерно-пакетная (*Pipeline Burst*) память.

Первой появилась асинхронная память, Интерфейс этой памяти включает шины данных, адреса и управления. В состав сигналов последней входят:

CS# (*Chip Select*) – сигнал выбора микросхемы;

WE# (*Write Enable*) – сигнал разрешения записи;

OE# (*Output Enable*) – сигнал включения выходов для выдачи данных.

Все сигналы управления инверсные, т.е. их активный (вызывающий соответствующее действие) уровень низкий. При единичном значении сигнала *OE#* выход микросхемы переходит в состояние высокого выходного сопротивления.

Время доступа t_{AC} у типовых микросхем составляет порядка 10 нс. Поэтому реально такие микросхемы могут работать на частотах, близких к частоте системной шины.

Несколько позже появилась синхронная пакетная статическая память (SBSRAM), ориентированная на выполнение пакетного обмена информацией, который характерен для кэш-памяти. Эта память включает в себя внутренний счетчик адреса, предназначенный для перебора адресов пакета, и использует сигналы синхронизации *CLK*, как и синхронная DRAM память

Для организации пакетного обмена, помимо имеющихся у асинхронной памяти управляющих сигналов *CS#*, *OE#* и *WE#*, в синхронную память также введены сигналы *ADSP#* (*Address Status of Processor*) и *CADS#* (*Cache Address Strobe*), сопровождающие передачу адреса нового пакета, а также сигнал *ADV#* (*Advance*) продвижения на следующий адрес пакета. Пакетный цикл всегда предусматривает передачу четырех элементов, так как внутренний счетчик имеет всего 2 бита, причем перебор адресов в пределах пакета может быть последовательным или с расслоением (чередованием) по банкам (при использовании процессоров семейства x86).

Следующим шагом в развитии статической памяти явилась конвейерно-пакетная память PBSRAM, обеспечивающая более высокое быстродействие, чем SBSRAM. В нее были введены дополнительные внутренние буферные регистры данных (здесь можно провести аналогию с EDO DRAM памятью) адреса, а в ряде модификаций предусмотрена возможность передачи данных на двойной скорости по переднему и заднему фронтам синхросигнала и используются удвоенные внутренние тракты записи и чтения. Это позволило получить время обращения порядка 2-3 нс и обеспечить передачу данных пакета без задержек на частотах шины более 400 МГц. Внутренняя логика позволяет переключаться с циклов чтения на циклы записи и наоборот без дополнительных задержек, кроме того, анализируется совпадение адресов записи и чтения для исключения избыточных операций. Структурная схема такой памяти приведена на рис. 13, где ФАП – блок формирования адресов пакета, МП – мультиплексоры, переключающие внутренние тракты чтения и записи в соответствии со значением младшего разряда адреса A0.

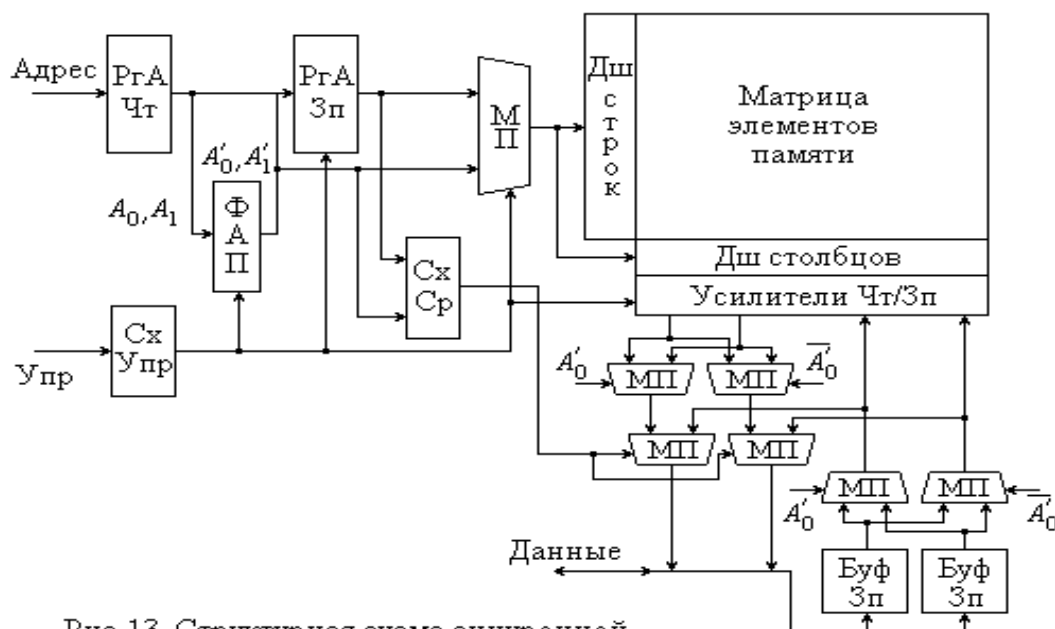


Рис. 13. Структурная схема синхронной статической пакетно-конвейерной (PBSRAM) памяти

Временная диаграмма, показывающая запись и чтение как на одиночной (SDR – *Single Data Rate*), так и на двойной скорости (DDR – *Double Data Rate*) передачи. Сигналы *CQ* и *CQ#* – дифференциальные выходные сигналы синхронизации, близкие по времени к моменту появления данных на шине при чтении и используемые для синхронизации принимающих устройств. Сигналы *SA* и *B* – адресные и управляющие сигналы соответственно, причем последние используются для задания типа цикла. В режиме чтения с двойной скоростью передачи формируются два набора данных, начиная со второго переднего и заднего фронтов синхросигнала, если по его первому переднему фронту передается начальный адрес пакета. Первый набор данных (DOUT-A) формируется для заданного адреса, а второй (DOUT-A') – для следующего адреса пакета, в соответствии с определенным для пакета порядком.

Аналогичным образом запись с двойной скоростью передачи требует установки данных для записи, начиная со второго переднего и заднего фронтов сигнала синхронизации. **В режиме чтения с одиночной скоростью формируется только один набор данных по переднему фронту синхросигнала, начиная со второго сигнала, а при записи с такой же скоростью данные должны выставляться на шину, начиная с переднего фронта второго синхросигнала после передачи адреса.** Переключение из режима двойной скорости в режим одиночной (и наоборот) производится при подаче соответствующего управляющего сигнала.

Модули динамических оперативных ЗУ

Оперативные запоминающие устройства всегда были ресурсом, допускающим увеличение емкости, а иногда и сокращение времени обращения, с целью повышения общей производительности ЭВМ. Совершенствование технологий изготовления оперативных ЗУ привело к тому, что стойки памяти середины 1970-х годов емкостью 512 Кбайт, размером с двусторчатый платяной шкаф, сменили маленькие платы с микросхемами размером с зажигалку. Более того, если модернизация (или, как теперь принято говорить, “апгрейд”) ЭВМ тех времен, связанная с наращиванием или заменой оперативной памяти, предполагала проведение достаточно серьезных монтажных работ, **то теперь, в стандартной ситуации, замену памяти в течение 5 – 10 минут может провести даже пользователь, не являющийся специалистом по компьютерной технике.**

Стандартизация интерфейсов оперативной памяти сделала возможным использование модулей памяти одних производителей в системах, собранных из компонент других производителей. Хотя, конечно, параметры и качество их могут быть различны.

Модули динамической полупроводниковой памяти прошли эволюцию от набора микросхем, устанавливаемых на системной плате и заметных по своему регулярному расположению (несколько смежных рядов одинаковых микросхем) до отдельных небольших плат, вставляемых в стандартный разъем (слот) системной платы. Первенство в создании таких модулей памяти обычно относят к фирме IBM.

Основными разновидностями модулей динамических оперативных ЗУ с момента их оформления в виде самостоятельных единиц были:

- **30-контактные** однобайтные модули **SIMM (DRAM)**
- **72-контактные** четырехбайтные модули **SIMM (DRAM)**
- **168-контактные** восьмибайтные модули **DIMM (SDRAM)**
- **184-контактные** восьмибайтные модули **DIMM (DDR SDRAM)**
- **184-контактные (20 из них не заняты)** двухбайтные модули **RIMM RDDRAM).**

Сокращение SIMM означает *Single In-Line Memory Module* – модуль памяти с одним рядом контактов, так как контакты краевого разъема модуля, расположенные в одинаковых позициях с двух сторон платы, электрически соединены. Соответственно DIMM значит *Dual In-Line Memory Module* – модуль памяти с двумя рядами контактов. А вот RIMM означает *Rambus Memory Module* – модуль памяти типа Rambus. Кроме этих модулей имеются также варианты для малогабаритных компьютеров, для графических карт и некоторые другие.

Если микросхемы памяти физически располагаются только с одной стороны платы, то такой модуль называют *односторонним*, а если с двух сторон – то *двухсторонним*. При равной емкости модулей, у двухстороннего модуля количество микросхем больше, поэтому на каждую линию шины данных приходится большая нагрузка, чем при использовании одностороннего. С этой точки зрения односторонние модули предпочтительней двухсторонних. Однако количество банков в двухсторонних модулях вдвое больше, чем в односторонних, поэтому при определенных условиях и хорошем контроллере памяти двухсторонний модуль может обеспечить несколько большую производительность.

На рис. 24 изображен односторонний 168-контактный 8-байтный модуль DIMM памяти SDRAM. Размеры этого модуля 133x35 мм. Микросхема, обозначенная звездочкой и пунктиром, используется в модулях, имеющих контрольные разряды и схемы контроля (см. далее). Размеры модулей SIMM несколько меньше. Модули RIMM имеют также металлическую пластину радиатора, закрывающую микросхемы.

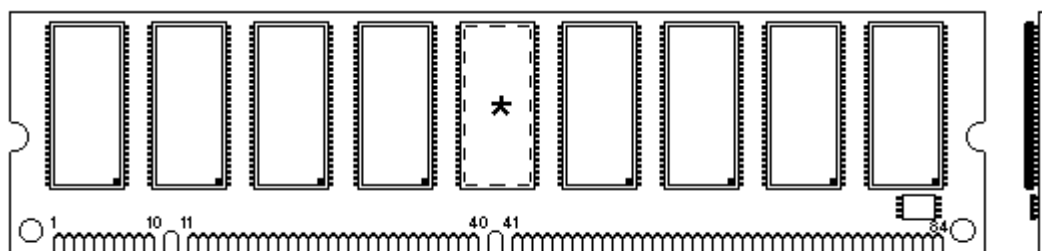


Рис. 24. Модуль оперативной памяти DIMM со 168 контактами

Помимо собственно конструктивной организации и типа памяти, модули имеют также и некоторые другие различия. Одним из таких различий является возможность (или ее отсутствие) *контроля хранимых данных*.

Контроль может основываться на использовании дополнительных (по одному на каждый хранимый байт) битов четности (*Parity bits*), т.е. в этом случае каждый байт занимает в памяти по 9 бит. Такой контроль позволяет выявить ошибки при считывании хранимой информации из памяти, но не исправить их. Причем, строго говоря, нельзя и установить, когда возникла ошибка: при записи или при чтении. Кроме того, парная ошибка, хотя ее вероятность и очень мала, обнаружена не будет.

Более сложный контроль предполагает использование кодов, корректирующих ошибки – ЕСС (*Error Correcting Codes*), которые позволяют обнаруживать ошибки большей кратности, чем одиночные, а одиночные ошибки могут быть исправлены. Подобные схемы используются в серверных конфигурациях, когда требуется повышенная надежность.

Память, устанавливаемая в настольные ПЭВМ, обычно не имеет никакого контроля. Кроме того, известны также различные модификации схем контроля, вплоть до просто имитирующих контрольные функции, но не осуществляющие их, например с генерацией всегда верного бита четности.

Модули DIMM также различаются по наличию или отсутствию в них буферных схем на шинах адреса и управляющих сигналов. Не буферизованные (*unbuffered*) модули больше нагружают эти шины, но более быстродействующие и дешевые. Их обычно применяют в настольных ПЭВМ. Буферизованные (*registered*) имеют буферные регистры и, обеспечивая меньшую нагрузку на шины, позволяют подключить к ней большее количество модулей. Однако эти регистры несколько снижают быстродействие памяти, требуя лишнего такта задержки. Применяют буферизованные модули обычно в серверных системах.

Кроме того, в последнее время, для получения более дешевых конфигураций памяти большого объема (в частности, для серверов) на модули устанавливают большое количество микросхем памяти. Так, серверный модуль DIMM может иметь до 64 (72 с ЕСС) чипов памяти. Это позволяет использовать более дешевые микросхемы меньшей плотности (объема), что оказывается заметно дешевле, по сравнению с применением более емких, но и более дорогих чипов последних выпусков. Например, используя микросхемы емкостью 1 Гбит с организацией 256М x 4 бита, можно, установив на модуль 64 (72) микросхемы, получить его емкость 8 Гбайт, не используя более дорогие, но более емкие чипы памяти.

В таких модулях приходится разделять установленные микросхемы на группы, каждая из которых имеет свой сигнал выбора кристалла CS#. Так, в приведенном примере групп окажется четыре и каждая будет включать в себя 16 (18) микросхем, дающих полноразрядное (8 байт) слово памяти. Такие группы называют “ранк” (*rank*) по аналогии с банками, на которые разделены сами микросхемы памяти (см. рис. 10). Этим групп бывает одна, две или четыре. Соответственно, говорят об одноранковых (*single rank - SR*), двухранковых (*dual rank - DR*) и четырехранковых (*quadro rank - QR*) модулях памяти DIMM.

На рис. 24-1 показан модуль памяти емкостью 8 Гбайт, построенный на 64 (72) микросхемах емкостью 1 Гбит и с организацией 256М x 4 бита, образующих 4 ранка (каждый из 16 микросхем) по 2 Гбайта. Для управления таких модулей контроллер памяти и материнская плата должны обеспечивать подачу на каждый слот, в который вставляется соответствующая

плата DIMM, одного, двух или четырех сигналов выбора кристалла, включающих микросхемы, относящиеся к одному из ранков.

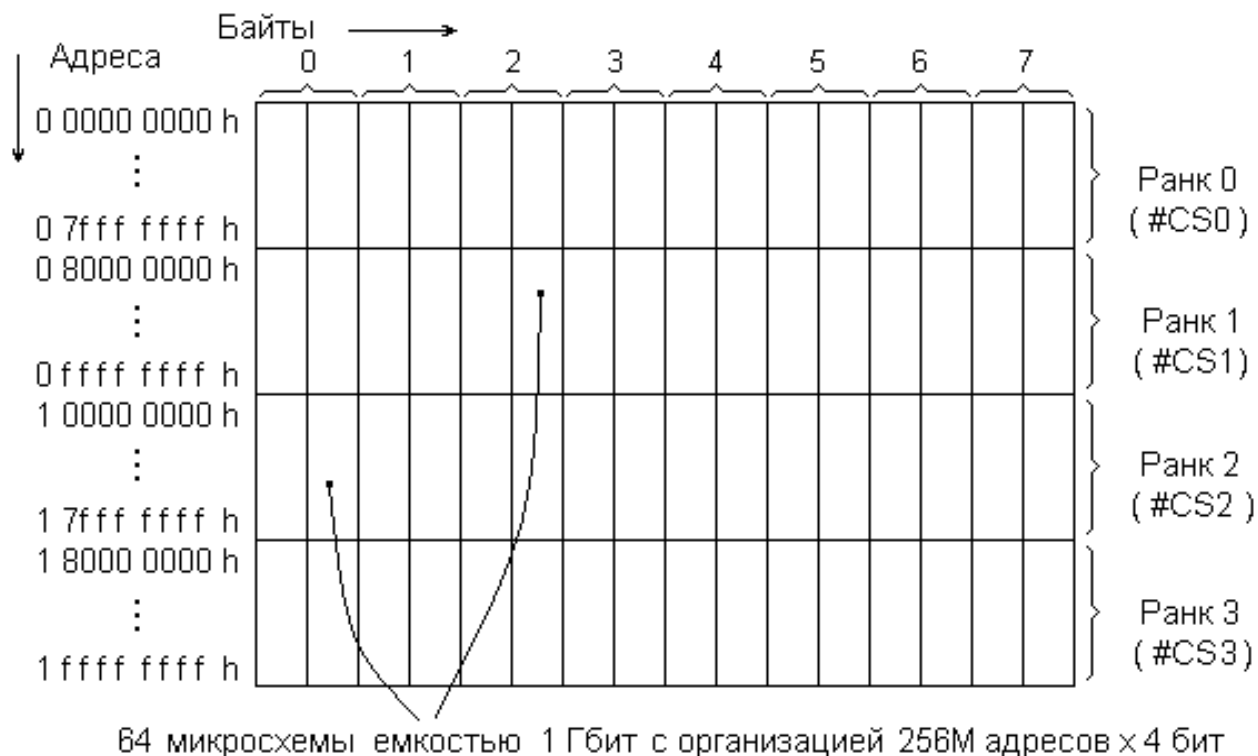


Рис. 24-1. Распределение адресов в четырехранковом модуле емкостью 8 Гбайт

Еще одной особенностью, различающей модули динамической памяти, является способ, посредством которого после включения компьютера определяется объем и тип установленной в нем памяти. В первых ПЭВМ объем и быстродействие установленной памяти задавались переключателями (джамперами – *jumper*s), расположенными на системной плате.

С появлением модулей SIMM (существовали также похожие на них модули SIPP) стал использоваться так называемый параллельный метод идентификации (*parallel presence detect*), при котором краевой разъем модуля имел дополнительные контакты, используемые только для целей указания присутствия модуля в том слоте, где он установлен, его объема и времени обращения. В самых первых (30-контактных) модулях таких дополнительных контактов было только два, в 72-контактных модулях их стало четыре: два указывали на объем модуля и два – на время обращения. Эти контакты могли заземляться непосредственно на модуле, что позволяло различить четыре вида модулей по объему и четыре – по времени доступа.

Попытки использовать этот же прием в последующих модулях потребовали увеличения количества таких контактов, но решить все проблемы идентификации не смогли. Поэтому, начиная с модулей DIMM, используют так называемый последовательный способ идентификации (*Serial Presence Detect* - SPD), при котором на плату модуля устанавливается специальная дополнительная микросхема, так называемый SPD-чип, представляющая собой небольшую постоянную память на 128 или 256 байт с последовательным (I²C) интерфейсом доступа. На рис. 24 эта микросхема показана в правом нижнем углу, хотя она может располагаться и иначе. В этой микросхеме записана основная информация об

изготовителе микросхемы и ее параметрах. Формат этих данных стандартный, определенный советом JEDEC (*Joint Electron Devices Engineering Council*), стандартов которого придерживаются все изготовители полупроводниковой памяти.