

1 КЛАССИФИКАЦИЯ И ХАРАКТЕРИСТИКИ ПАМЯТИ

Важной частью вычислительных систем является **память**. Память соединяется с управляющим памятью контроллером (устройством управления памятью) по шине адреса, шине данных и шине управления. Разрядность шины данных определяет, сколько двоичных разрядов одновременно (параллельно) может быть считано из памяти. Каждый двоичный разряд (1 бит) хранится элементом памяти. Элементы для памяти различного типа строятся на основе различных физических принципов записи и хранения информации. Элементы памяти объединяются в ячейки памяти. При этом все элементы ячейки адресуются одновременно, одинаково и организованы так, что одновременно могут выдавать данные на шину данных. Такие объединенные ячейки образуют слово. Количество разрядов данных, считываемых из памяти одновременно, называют длиной выборки. Для хранения 1 байта используется 8 элементов памяти, восьмибитные ячейки памяти организованы с использованием шины данных шириной 8 линий.

Из микросхем памяти (чипов) создаются модули памяти, которые устанавливаются в специальные слоты (разъемы) вычислительной системы. Сейчас наиболее распространены DIMM модули - модули памяти с двумя рядами контактов.

Разрядность адресной шины определяет адресное пространство, то есть количество ячеек памяти, которые могут адресоваться непосредственно. Если разрядность адресной шины - n , то количество всех возможных двоичных комбинаций (количество адресов) определится как $N = 2^n$.

Память вычислительного устройства может выполнять три операции:

- а) хранение информации;
- б) запись информации;
- в) чтение информации.

1.1 Характеристики памяти

Память характеризуется такими параметрами, как емкость, удельная емкость, плотность записи, время доступа к памяти, пропускная способность, надежность, стоимость единицы информации

Емкость памяти определяет максимальное количество хранимой в памяти информации и измеряется в битах, байтах, килобайтах (кБ), мегабайтах (МБ), гигабайтах (ГБ), терабайтах (ТБ) и т.д.

$1 \text{ кБ} = 2^{10} \text{ Б}$; $1 \text{ МБ} = 2^{20} \text{ Б}$; $1 \text{ ГБ} = 2^{30} \text{ Б}$; $1 \text{ ТБ} = 2^{40} \text{ Б}$.

Удельная емкость определяется как отношение емкости памяти к физически занимаемому ею объему.

Плотность записи информации определяется как количество информации, приходящееся на единицу площади носителя информации или на единицу длины носителя информации.

Время доступа к памяти

Быстродействие памяти определяется продолжительностью выполнения операций при обращении к памяти. Время обращения при записи и время обращения при чтении складывается из времени поиска ячейки памяти по заданному адресу и собственно записи или чтения соответственно.

$$t_{\text{обращ.зап}} = t_{\text{поиска}} + t_{\text{записи}}$$

$$t_{\text{обращ.чтен}} = t_{\text{поиска}} + t_{\text{чтения}}$$

Пропускная способность памяти (Bandwidth)

Пропускная способность памяти характеризует скорость передачи данных, которую может обеспечить организация памяти. Скорость передачи данных могут называть еще и производительностью памяти. Измеряется пропускная способность памяти в килобайтах в секунду (кБ/с), мегабайтах в секунду (МБ/с), гигабайтах в секунду (ГБ/с).

Надежность запоминающего устройства определяется вероятностью его безотказной работы в заданный интервал времени.

Стоимость памяти определяется стоимостью одного бита информации.

1.2 Классификация памяти

1) по способу доступа к информации

память с произвольным доступом (оперативная)

Для памяти с произвольным доступом (электронной памяти) время обращения не зависит от местоположения искомого участка памяти. Выбор ячейки происходит по адресу при помощи электронных схем.

память с прямым циклическим доступом (дисковая)

При обращении к дисковой памяти используется прямой циклический доступ. Носитель информации непрерывно вращается, поэтому возможность обращения к одному и тому же участку памяти является циклической.

память с последовательным доступом (ленточная)

Последовательный доступ к данным возможен при использовании в качестве носителя магнитной ленты, где последовательный просмотр участков носителя необходим для нахождения нужных данных.

2) по функциональному назначению:

ПЗУ - постоянные запоминающие устройства или ROM (Read Only Memory), служат для хранения постоянных данных и служебных программ;

СОЗУ- сверхоперативное запоминающее устройство, это набор регистров общего назначения - РОН, предназначенных для хранения операндов и результатов выполнения операции в процессоре. Сверхоперативная память (кэш память) обладает высоким быстродействием и согласует работу процессора с более медленной внешней памятью;

ОЗУ - оперативное запоминающее устройство или RAM (Random Access Memory - память с произвольной выборкой), служит для хранения выполняемой программы и оперативных данных.

3) по способу адресации:

Адресная память

Любая ячейка памяти имеет свой адрес, по которому можно выделить ячейку как для записи, так и для чтения информации.

Ассоциативная память

При чтении информации нужную ячейку находят не по адресу, а по её содержимому, задавая ассоциативный поиск. Запись информации производится в любую свободную ячейку.

Стековая память

Также является безадресной памятью. Может функционировать по принципу "первым вошел - первым вышел" (FIFO-First Input, First Output) или по принципу "первым вошел - последним вышел" (FILO-First Input, Last Output).

Стековая память получила широкое распространение. Для ее реализации в ОЗУ посредством программ операционной системы выделяется часть адресной памяти.

4) по способу хранения информации:

По способу хранения информации память делится на статическую, динамическую, постоянную, голографическую, биологическую, внешнюю на магнитных носителях, внешнюю оптическую и магнитооптическую.

Статическая память

В статических запоминающих устройствах БИС выполнены на триггерных элементах памяти (имеющих два стабильных состояния).

Динамическая память

В динамических запоминающих устройствах используются более дешевые БИС, в которых запоминающим элементом является конденсатор. Конденсатор со временем разряжается (в этом - динамика), поэтому необходимо поддерживать значение потенциала, подзаряжая конденсатор. Этот процесс называют регенерацией.

Постоянная память

В постоянных запоминающих устройствах запоминающим элементом является пережигаемая плавкая перемычка или полупроводниковый диод, играющий роль разрушаемой перемычки. В перепрограммируемых ПЗУ для записи и хранения информации применяются ячейки, выполненные на МОП транзисторах с плавающим и изолированным затвором.

Голографическая память

В голографических запоминающих устройствах информация хранится в объеме голографического кристалла в виде снимка интерференции двух волн, опорной и информационной. Этот перспективный вид запоминающих устройств имеет большую плотность записи информации.

Биологическая память

В биологических запоминающих устройствах для записи информации используется изменение состояния органических молекул, обладающих свойством хранить заряд и обмениваться электронами.

Память на магнитных носителях

Во внешних запоминающих устройствах на магнитных носителях информация хранится в виде намагниченных в определенном направлении участков ферромагнитной поверхности диска или магнитной ленты.

Оптическая память В оптических внешних запоминающих устройствах информация записывается в виде участков, имеющих разные коэффициенты рассеяния света направленного луча лазера.

2 ДИНАМИЧЕСКАЯ И СТАТИЧЕСКАЯ ПАМЯТЬ

2.1 Динамическая память

Динамическая память - DRAM (Dynamic RAM) - получила свое название от принципа действия элемента памяти, хранящего один бит информации. Запоминающим

элементом служит конденсатор, являющийся емкостью перехода исток - сток полевого транзистора. При записи логической единицы в элемент памяти конденсатор заряжается, при записи нуля - разряжается. При чтении информации конденсатор разряжается, поэтому специальной схемой перезаписи конденсатор нужно снова зарядить до исходного потенциала. При отсутствии обращения к элементу динамической памяти со временем, за счет токов утечки, конденсатор разряжается и информация теряется, поэтому такая память требует постоянного периодического перезаряда конденсаторов методом перезаписи каждого элемента (регенерации). Следовательно, память может работать только в динамическом режиме, чем принципиально отличается от статической памяти, реализуемой на триггерных ячейках и хранящей информацию без обращений к ней статично, сколь угодно долго при включенном питании. Благодаря относительной простоте ячейки динамической памяти, на одном кристалле удастся размещать миллионы ячеек и получать сравнительно недорогую полупроводниковую память, пускай и недостаточно высокого быстродействия, но с умеренным энергопотреблением. Используется динамическая память в качестве **оперативной памяти** компьютера и оперативной памяти видеоадаптера.

Главные особенности микросхем динамической памяти заключаются в следующем:

- отсутствуют источники для питания элементов запоминающих ячеек;
- необходимы логические схемы, обеспечивающие регенерацию информации;
- максимально проста схема накопителя, обеспечивающая минимум занимаемой площади;

2.2 Статическая память.

Статическая память - SRAM (Static Random Access Memory), способна хранить информацию в статическом режиме - то есть сколь угодно долго при отсутствии обращений (но при наличии питающего напряжения). Ячейки статической памяти реализуются на триггерах - элементах с двумя устойчивыми состояниями. По сравнению с динамической памятью эти ячейки более сложны и занимают больше места в кристалле, однако они проще в управлении и не требуют регенерации. Быстродействие и энергопотребление статической памяти определяется технологией изготовления и схемотехникой запоминающих ячеек.

Самая экономичная статическая память КМОП (или CMOS Memory) в тоже время и самая медленная память такого типа, имеет время доступа более 100 наносекунд, но зато пригодна для длительного хранения информации при питании от маломощной батареи. Применяется CMOS память в персональных компьютерах для хранения данных о конфигурации и для реализации внутренних часов.

Самая быстродействующая статическая память имеет время доступа в несколько наносекунд, что позволяет ей работать на частоте системной шины процессора, не требуя от него тактов ожидания. Относительно высокая удельная стоимость хранения информации и высокое энергопотребление при низкой плотности упаковки элементов не позволяет использовать SRAM в качестве оперативной памяти компьютеров.

Статические запоминающие устройства (SRAM) имеют перед динамическими то преимущество, что у них высокое быстродействие. Главным ограничением в

использовании статической памяти является стоимость. При равной емкости с динамической, статическая память примерно в четыре раза дороже. Поэтому данный вид памяти получил распространение в высокопроизводительных системах в качестве внешней (относительно процессора) кэш памяти.

Асинхронная и синхронная статическая память

Asynchronous SRAM - обычная (стандартная) статическая память. Асинхронная статическая память работает независимо от синхронизации системной шины и поэтому, окончание цикла обмена может не совпадать с началом очередного тактового импульса. В результате, цикл обмена удлиняется, снижая тем самым эффективную производительность. В настоящее время асинхронная память практически нигде не используется. Микросхемы этого типа имеют простейший асинхронный интерфейс, включающий шину адреса, шину данных и сигналы управления.

Sync Burst SRAM - синхронная статическая память, оптимизированная под выполнение пакетных (burst) операций обмена, свойственных работе кэш-памяти. Синхронная статическая память выполняет все операции одновременно с тактовыми сигналами, в результате чего время доступа к ячейке укладывается в один-единственный такт.

3 СВЕРХОПЕРАТИВНАЯ (КЭШ) ПАМЯТЬ

Оперативная память вычислительных систем реализуется на относительно медленных микросхемах динамической памяти. Если процессор работает непосредственно с памятью такого типа, требуется вводить такты ожидания для процессора. В этом случае процессор частично простаивает. Статическая память, построенная на триггерах, имеет то же быстродействие, что и процессоры и использование такой памяти позволило бы работать процессору без тактов ожидания, но параметр цена/производительность диктует поиски другого выхода из этой ситуации. Разумным компромиссом для построения экономичных и производительных систем явился иерархический способ построения памяти. Обмен процессора с оперативной памятью производится при помощи промежуточной относительно небольшой кэш-памяти на быстродействующих микросхемах SRAM.

В переводе слово кэш (cache) означает склад или тайник. Для выполняемой программы он не представляет собой дополнительной адресуемой памяти. Он выполняет роль быстродействующего хранилища копий блоков информации основной памяти, к которым в ближайшее время ожидается обращение. Объем кэш во много раз меньше объема оперативной памяти. Контроллер кэш памяти ведет каталог и при каждом обращении к кэш памяти по каталогу проверяет, есть ли действительная копия затребованных данных в кэше. Если она там есть, то это случай кэш попадания (cache hit), и обращение за данными происходит только к кэш-памяти. Если действительной копии там нет, то это случай кэш-промаха (cache miss), и данные берутся из основной памяти.

Отношение числа попаданий к общему числу обращений к памяти называется коэффициентом эффективности поиска. Если этот коэффициент достаточно высок, большая часть времени затрачивается на обращение к кэш-памяти, что эквивалентно использованию быстродействующей основной памяти при выполнении программы. Обычно коэффициент эффективности поиска превышает 95%.

Обращение к основной памяти может начинаться одновременно с поиском данных в каталоге кэш, и в случае попадания - прерывается (архитектура Look Aside). При параллельном поиске из-за параллельного обращения к оперативной памяти возрастают затраты энергии.

В другом варианте обращение к внешней памяти начинается только после фиксации случая промаха (архитектура Look Through), на этом теряется, по крайней мере, один такт процессора, зато значительно экономится энергия.

В современных процессорах кэш обычно строится по двухуровневой схеме. Первичный кэш (L1 Cache) имеет сравнительно небольшой объем и для повышения производительности обычно используется отдельный кэш для данных и команд (так называемая Гарвардская архитектура - противоположность Принстонской, использующей общий кэш для команд и данных). Вторичный кэш (L2 Cache) тоже включается в микросхему микропроцессора, но не разделяется на кэш команд и данных. Работает встроенный кэш на внутренней тактовой частоте процессора.

Принципы взаимодействия кэш памяти с оперативной памятью.

Кэш память может быть построена как синхронная или асинхронная память.

Асинхронный кэш

При обращении к памяти контроллер асинхронного кэш производит поиск адреса данных в таблицах, поэтому при каждом обращении необходим дополнительный цикл для просмотра таблиц (тегов), можно реализовать кэш с быстрым циклом 3-2-2-2, но чаще цикл асинхронного кэш составляет 4-2-2-2 такта.

Синхронный кэш

Контроллер синхронного кэш помещает поступающие адреса в буфер. В первом такте запрашиваемый адрес запоминается в регистре. Во втором такте данные устанавливаются на шине данных и читаются процессором, в это же время осуществляется чтение в буфер следующего адреса. Для последовательных элементов данных (без дополнительной дешифрации адреса) в оптимальном варианте может быть обеспечен цикл 2-1-1-1 по высокоскоростной шине.

Контроллер кэш

Контроллер кэш памяти оперирует со строками данных (cache line) фиксированной длины (или блоками данных). С каждой строкой кэш (блоком) связана информация об адресе скопированного в нее блока основной памяти и признаки ее состояния. Строка может быть действительной (valid) - это означает, что в текущий момент времени она достоверно отражает соответствующий блок основной памяти, или недействительной (пустой).

Строки кэш памяти обычно выделяются только при операциях чтения. Запись блока, не имеющего копии в кэш памяти, производится только в основную память. Поведение кэш-контроллера при операции записи в память, когда копия затребованной области находится в некоторой строке кэш памяти, определяется его политикой записи (Write Policy).

Существуют два основных алгоритма записи данных из кэш памяти в оперативную память: сквозная запись WT (Write Through) и обратная запись WB (Write Back).

Алгоритм сквозной записи (WT) предусматривает выполнение каждой операции записи (даже однобайтной) одновременно и в строку кэш памяти, и в основную память. Алгоритм достаточно прост в реализации и легко обеспечивает целостность данных за счет постоянного совпадения копий данных в кэш и в основной памяти. Но эта простота оплачивается низкой эффективностью записи.

Алгоритм обратной записи (WB) позволяет уменьшить количество операций записи в оперативную память. Если блок памяти, в который должна производиться запись, отображен и в кэш памяти, то физическая запись сначала будет произведена в эту действительную строку кэш памяти, и будет отмечена как грязная (dirty), или модифицированная, то есть требующая записи в оперативную память. Только после записи в оперативную память строка станет чистой (clean). Данный алгоритм сложнее в реализации, но существенно эффективнее, чем алгоритм сквозной записи (WT).

В зависимости от способа определения взаимного соответствия строки кэш памяти и области основной памяти различают три архитектуры кэш-памяти: кэш прямого отображения (direct-mapped cache), полностью ассоциативный кэш (fully associative cache) и их комбинация - частично или наборно-ассоциативный кэш (set-associative cache).

4 ПОСТОЯННАЯ ПАМЯТЬ

Постоянная память является энергонезависимой, информация в ней сохраняется и после выключения питания компьютера. Существует множество типов энергонезависимой памяти: ROM, PROM, EPROM, EEPROM, Flash Memory, различающихся по своим свойствам, обусловленным способом построения запоминающих ячеек, и сферам применения. Запись информации в энергонезависимую память, называемая программированием, обычно существенно сложнее и требует больших затрат времени и энергии, чем считывание. Основным режимом работы такой памяти является считывание данных, а некоторые типы после программирования допускают только считывание, что и обуславливает их общее название ROM (Read Only Memory - память только для чтения) или ПЗУ (постоянное запоминающее устройство). Самые первые постоянные запоминающие устройства выполнялись на магнитных сердечниках, где информация заносилась их прошивкой проводниками считывания. С тех пор применительно к программированию ПЗУ укоренилось понятие "прошивка".

По возможности программирования различают:

- Микросхемы, программируемые при изготовлении (масочные ПЗУ) - ROM.
- Микросхемы, программируемые однократно после изготовления перед установкой в целевое устройство (прожигаемые ПЗУ, программируемые на специальных программаторах) - *PROM* (Programmable ROM) или ППЗУ (программируемые ПЗУ).
- Микросхемы, стираемые и программируемые многократно - *EPROM* (Erasable PROM - стираемые ПЗУ) или РПЗУ (репрограммируемые ПЗУ).

Запоминающие ячейки энергонезависимой памяти обычно несимметричны по своей природе и позволяют записывать только нули (реже - только единицы) в предварительно стертые (чистые) ячейки. Однократно программируемые микросхемы

позволяют изменять только исходное (после изготовления) состояние ячеек. Для многократно программируемых микросхем возможно стирание, которое выполняется либо для всей микросхемы, либо для определенного блока, либо для одной ячейки (байта). Стирание приводит все биты стираемой области в одно состояние (обычно во все единицы, реже - во все нули). Процедура стирания обычно существенно дольше записи.

В зависимости от способа стирания различают:

- микросхемы, стираемые ультрафиолетовым облучением, - их обычно называют просто *EPROM* (Eraseable ROM - стираемые микросхемы) или *UV-EPROM* (Ultra-Violet EPROM, УФППЗУ);
- электрически стираемые микросхемы *EEPROM* (Electrical Eraseable PROM, ЭСПЗУ), в том числе и флэш-память.

Процедура программирования многих типов памяти требует наличия относительно высокого напряжения программирования (12-26 В). Стирание или программирование микросхем может выполняться либо в специальном устройстве - программаторе, либо в самом целевом устройстве, если у него предусмотрены соответствующие средства.

Энергонезависимая память в основном применяется для хранения неизменяемой (или редко изменяемой) информации - системного программного обеспечения (BIOS), таблиц (например, знакогенераторов графических адаптеров), памяти конфигурации устройств. Эта информация обычно является ключевой для функционирования PC, поэтому весьма существенна забота о ее сохранности и предотвращении несанкционированного изменения. Нежелательное (ошибочное или под воздействием вируса) изменение содержимого становится возможным при использовании для хранения BIOS флэш-памяти, программируемой в целевом устройстве (на системной плате PC). Флэш-память используется также в качестве внешней памяти, позволяющей как считывать, так и оперативно записывать новые данные.

Важными параметрами энергонезависимой памяти является время хранения и устойчивость к электромагнитным воздействиям, а для перепрограммируемой памяти еще и гарантированное количество циклов перепрограммирования.

Программирование масочного ПЗУ производится на заводе изготовителе, что очень неудобно для мелких и средних серий производства, не говоря уже о стадии разработки устройства. Естественно, что для крупносерийного производства масочные ПЗУ являются самым дешевым видом ПЗУ. Для мелких и средних серий производства радиоаппаратуры были разработаны микросхемы, которые можно программировать в специальных устройствах - программаторах. В этих микросхемах постоянное соединение проводников в запоминаящей матрице заменяется плавкими перемычками, изготовленными из поликристаллического кремния. При производстве микросхемы изготавливаются все перемычки, что эквивалентно записи во все ячейки памяти логических единиц. В процессе программирования перемычки пережигаются, т.е. в определённые разряды ячеек памяти записываются нули.

В репрограммируемых ПЗУ стирание ранее записанной информации осуществляется ультрафиолетовым излучением. Для того, чтобы этот свет мог беспрепятственно проходить к полупроводниковому кристаллу, в корпус микросхемы встраивается окошко из кварцевого стекла. Время стирания микросхемы колеблется в пределах 10 - 30 минут.

Количество циклов записи - стирания микросхем находится в диапазоне от 10 до 100 раз, после чего микросхема выходит из строя.

Флэш-память по определению относится к классу EEPROM, но использует особую технологию построения запоминающих ячеек. Стирание во флэш-памяти производится сразу для целой области ячеек (блоками или полностью всей микросхемы). Это позволило существенно повысить производительность в режиме записи (программирования). Флэш-память обладает сочетанием высокой плотности упаковки (ее ячейки на 30% меньше ячеек DRAM), энергонезависимого хранения, электрического стирания и записи, низкого потребления, высокой надежности и невысокой стоимости.