

Характеристики систем памяти.

Место расположения:

- Процессорные (РОНЫ, КЭШ I)
- Внутренние (Основная память, КЭШ II, КЭШ IV)
- Внешние (магнитные/оптические диски, магнитные ленты)

Ёмкость и организация. Ёмкость измеряется в байтах:

- Килобайтах (2^{10}), мегабайтах (2^{20}), гигабайтах (2^{30}), терабайтах (2^{40}), петабайтах (2^{50}), эксабайтах (2^{60}).

Единица пересылки. Определяется шириной шины данных (слово). Применительно к внешней памяти данные передаются блоками.

Метод доступа:

- Последовательный
- Прямой
- Произвольный
- Ассоциативный

Быстродействие:

- Время доступа.
- Длительность цикла памяти.
- Скорость передачи.

Физический тип:

- Полупроводниковая память
- Память с магнитными носителями
- Память с оптическими носителями

Физические особенности ЗУ. Например энергозависимость.

Стоимость. Принято оценивать стоимость хранения 1 бита информации

Методы доступа к данным ЗУ.

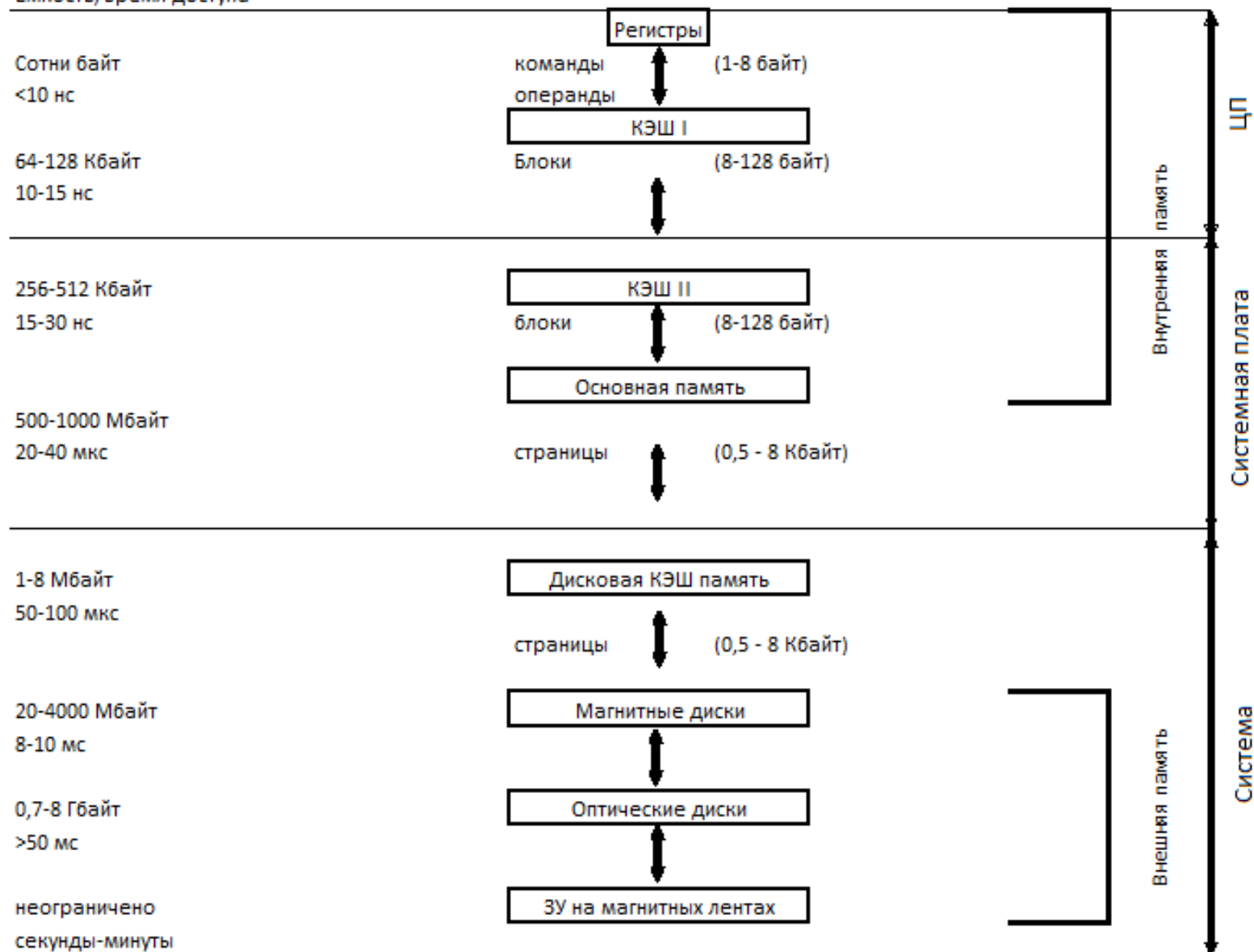
- Последовательный. Информация хранится в виде последовательности блоков данных, называемых «записями». Для доступа к нужному элементу необходимо прочитать все предшествующие ему данные. Пример: магнитная лента.
- Прямой доступ. Каждая запись имеет уникальный адрес, отражающий её физическое размещение на носителе. Обращение осуществляется как адресный доступ к записи с последующим последовательным доступом к единице информации.
- Произвольный доступ. Каждая ячейка имеет уникальный физический адрес. Обращения к любой ячейке занимает одно и тоже время и может производиться в последовательности.
- Ассоциативный доступ. Осуществляется поиск информации не по адресу, а по признаку (тегу). Примером является кэш-память

Параметры быстродействия ЗУ.

- Время доступа (T_d). Для памяти с произвольным доступом она соответствует интервалу времени от момента поступления адреса до момента когда данные заносятся или становятся доступными.
 - Длительность цикла памяти или период обращения ($T_{ц}$). Применяется к памяти с произвольным доступом. Означает максимальное время между двумя последовательными обращениями к памяти.
 - Скорость передачи. Скорость с которой данные могут передаваться в память или из неё. Для памяти с произвольным доступом она равна $1 / T_{ц}$.
- Для современных ЗУ характерно обращение к группе слов (**пакету**). Для них дополнительно используются следующие параметры:
- Время доступа при первом обращении (latency).
 - Темп передачи для последующих слов пакета (band words): предельный и усредненный (с учетом latency).

Иерархия запоминающих устройств.

Емкость/время доступа



Закономерность типов ЗУ.

- Чем меньше время доступа, тем выше стоимость хранения бита;
- Чем больше емкость, тем ниже стоимость хранения бита, но больше время доступа.

По мере движения вниз по иерархической структуре памяти:

1. Уменьшается соотношения «стоимость/бит».
2. Возрастает емкость.
3. Растёт время доступа.
4. Уменьшается частота обращения к памяти со стороны центрального процессора.

Принцип локальности по обращению.

- Пространственная локальность программ. С очень высокой вероятностью адрес очередной команды программы следует непоследовательно за адресом по которому была считана текущая команда.
- Пространственная локальность данных. Обрабатываемые данные как, правило, структурированы, и такие структуры обычно хранятся в последовательных ячейках памяти.
- Временная локальность. Программы, как правило, содержат множество циклов и подпрограмм, которые могут многократно повторяться в течение некоторого интервала времени.

Принцип локальности часто обличают в численную форму и представляют в виде правила 90/10: 90% времени работы программы связано с обращению с обращением к 10% адресного пространства этой программы.

Характеристики оценки эффективности иерархической организации ОП.

- Коэффициент попадания (hit rate) – отношение числа обращений к памяти, при котором произошло попадание, к общему числу обращение к запоминающему устройству данного уровня иерархии.
- Коэффициент промахов (miss rate) – отношение числа обращений к памяти, при которых имел место промах, к общему числу обращение к запоминающим устройствам данного уровня иерархии.
- Время обращения при попадании (hit time) – время, необходимое для поиска нужной информации в памяти верхнего уровня (включая выяснение является ли обращение попаданием), и время на фактическое считывание данных.
- Потери на промах (miss penalty) – время требуемое для замены блоков в памяти более высокого уровня на блок с нужными данными в запоминающем устройстве следующего уровня.

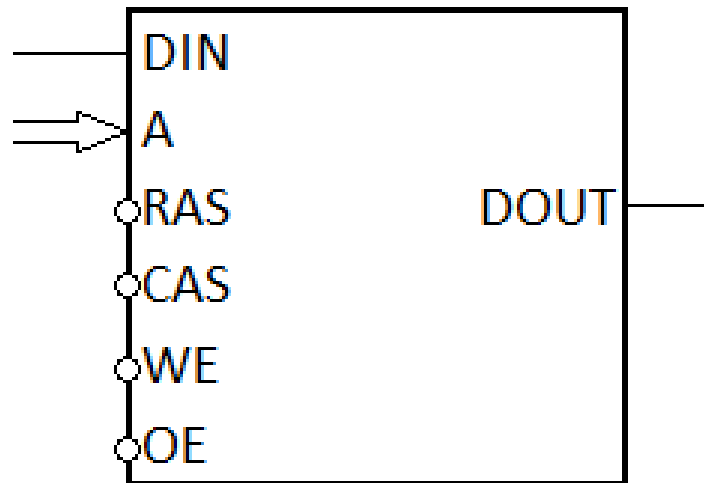
Описание уровней иерархии запоминающего устройства предполагает конкретизацию 4 моментов:

- Размещение блока – допустимого места расположения блока на примыкающим сверху уровни иерархии.
- Идентификация блока – способа нахождения блока на примыкающем сверху уровне.
- Замещение блока – выбора блока, замещаемого при промахе с целью освобождения места для нового блока.
- Согласования копий (стратегии записи) – обеспечение согласованности копий одних и тех же блоков, расположенных на разных уровнях, при записи новой информации в копию, находящуюся на более высоком уровне.

Динамическая память DRAM.

В динамических запоминающих устройствах данные хранятся в виде зарядов емкостей МОП-структур. Основой запоминающего элемента является конденсатор небольшой емкости. За счёт этого в кристалле ИС можно разместить в 4-5 раз больше ЗЭ чем в кристалле SRAM.

Недостаток – необходимость регенерации.



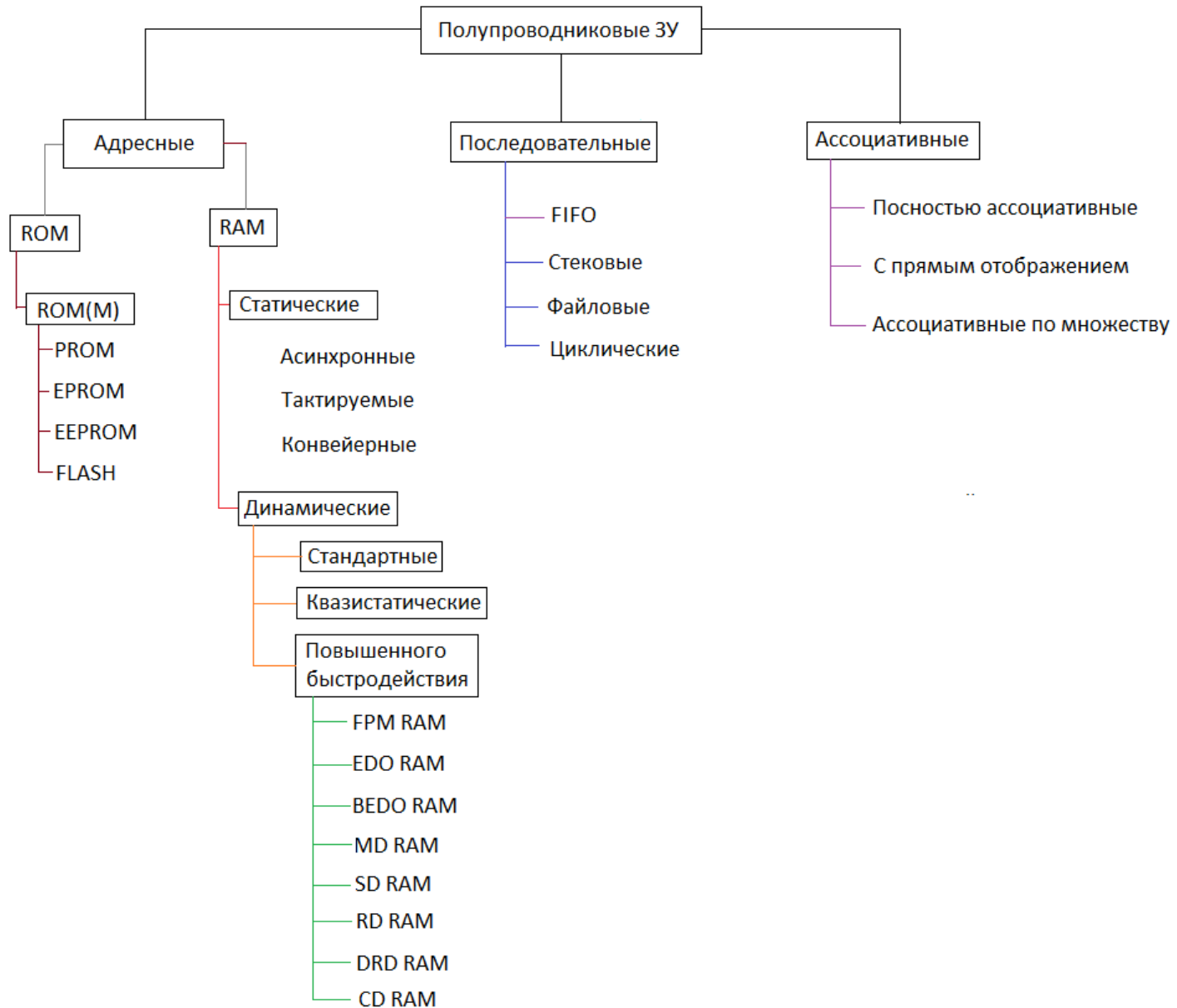
RAS (row access strobe) – строб выборки адреса строки.

CAS (column access strobe) – строб выборки адреса столбца.

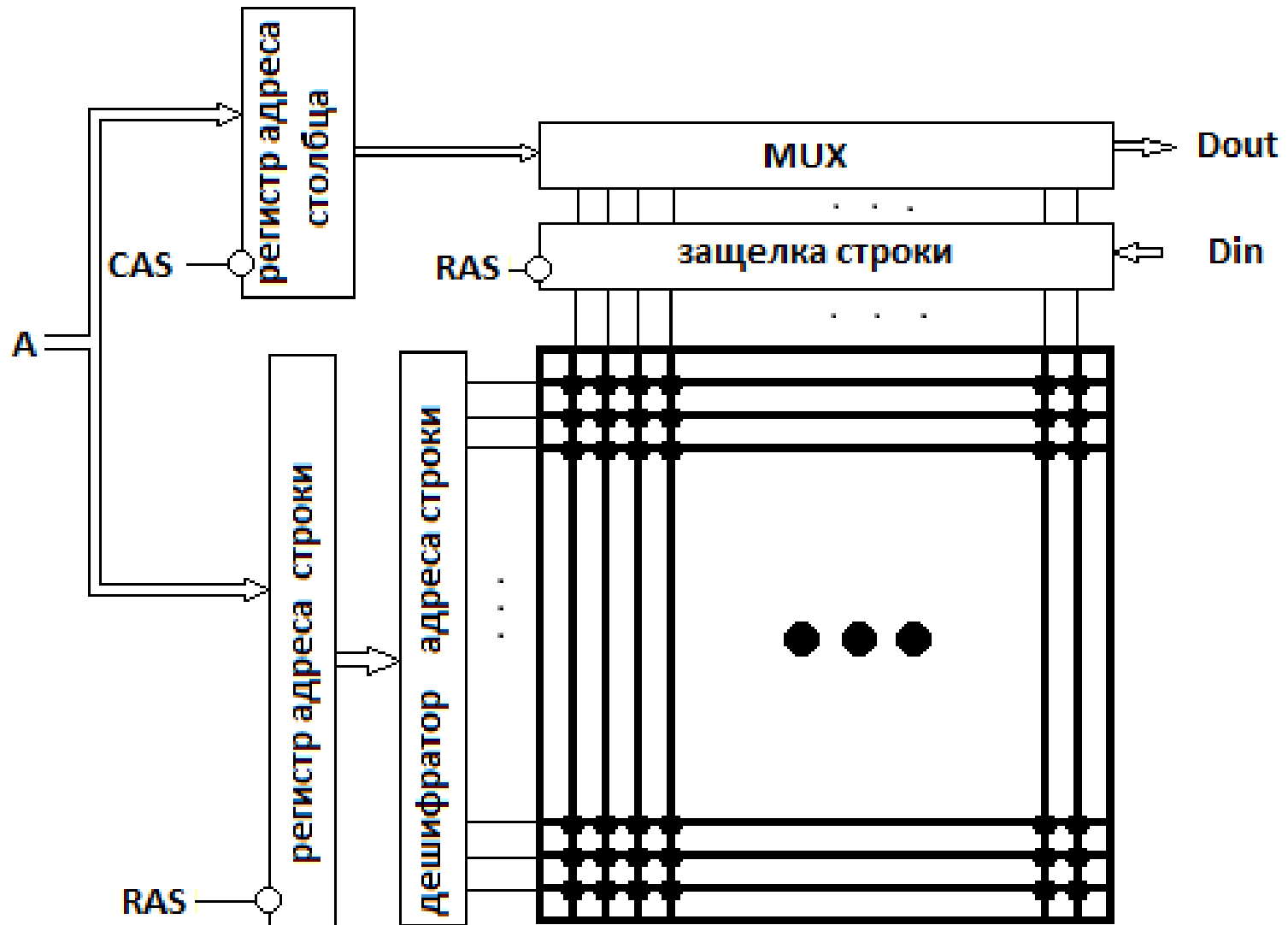
WE (write enable) – размещение записи.

OE (output enable) – размещения выхода.

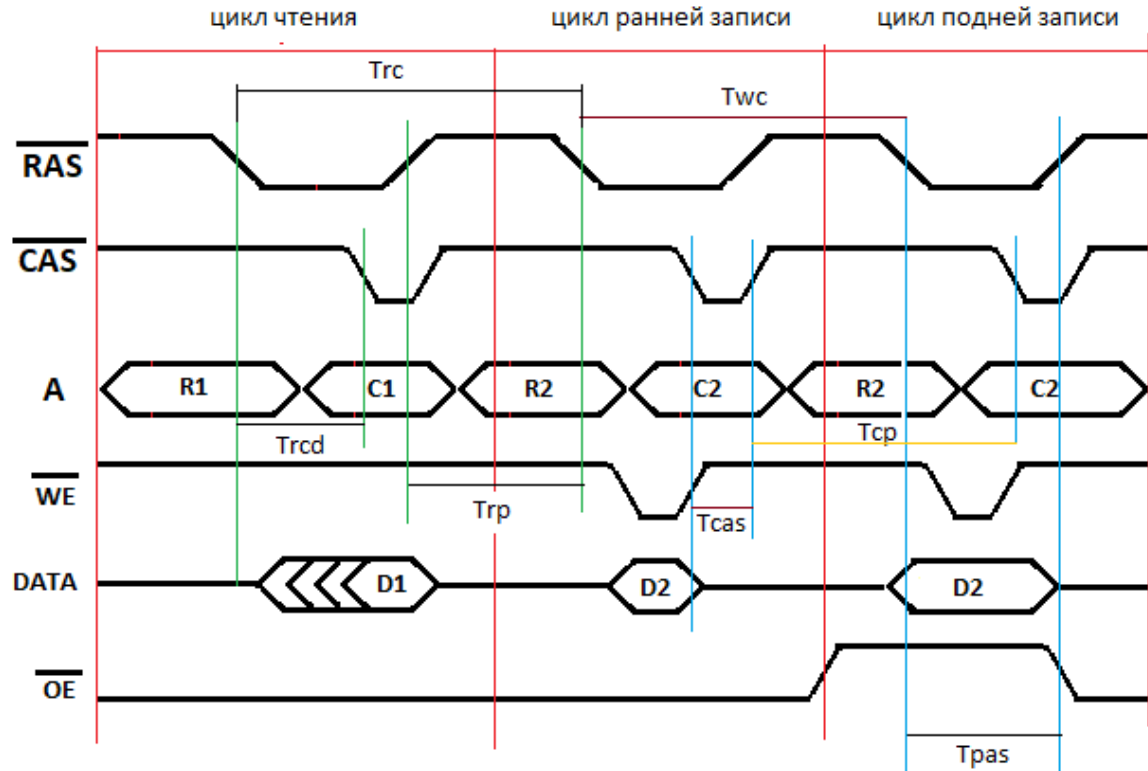
Классификация полупроводниковых ЗУ.



Структура микросхемы DRAM.



Временные диаграммы работы DRAM.



T_{ras} – время доступа. Задержка появления данных на выходе относительно RAS.

T_{rc} – цикл чтения.

T_{wc} – цикл записи.

T_{trd} – минимальная длительность активной части RAS.

T_{trc} – минимальная длительность активной части CAS.

T_{cp} – время предварительного заряда CAS. Минимальное время высокого состояния.

T_{rd} – время предварительного заряда RAS. Минимальное время нахождения сигнала в высоком состоянии.

T_{trc} – задержка между импульсами RAS и CAS.

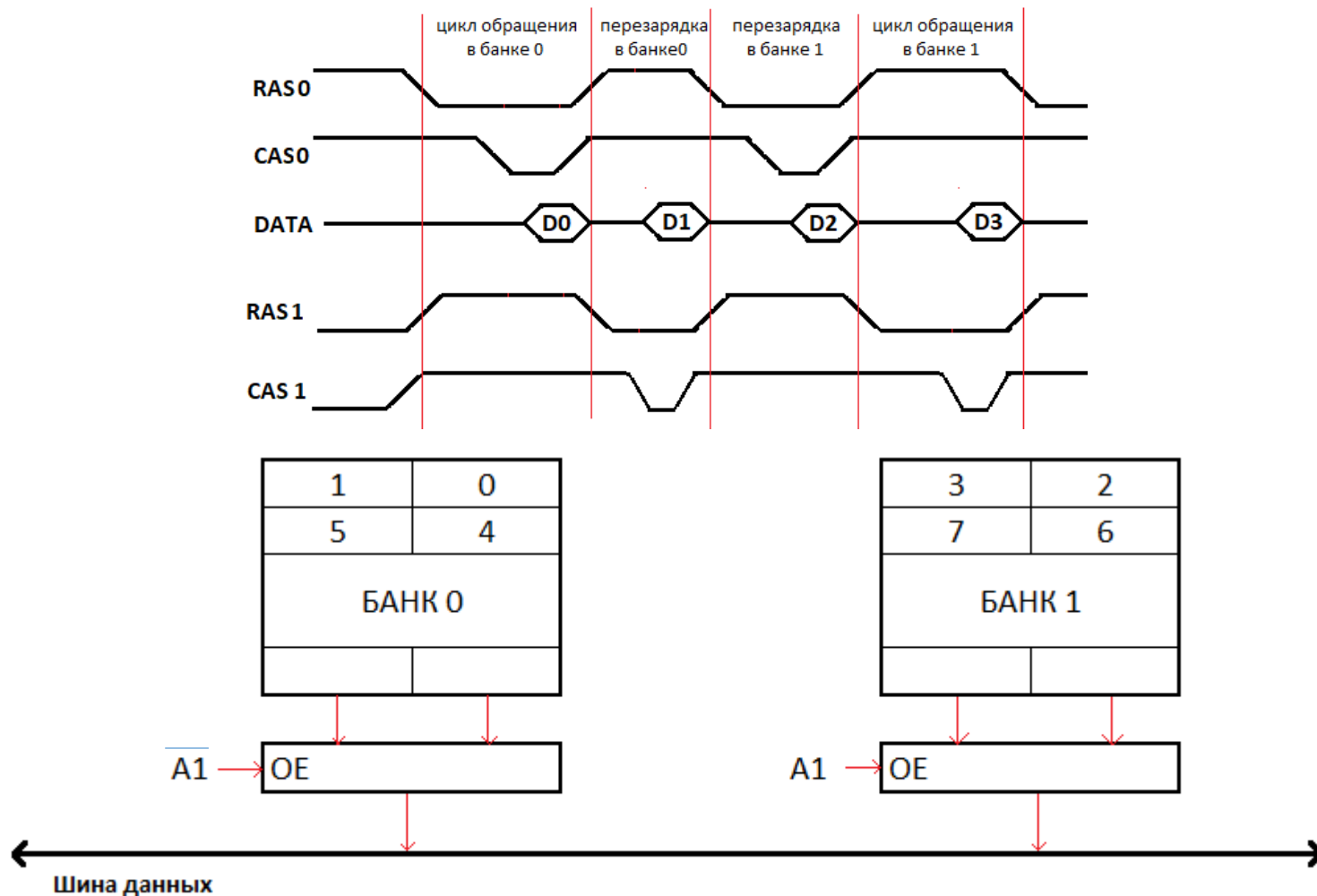
Особенности DRAM.

По спаду RAS в регистре адреса строки фиксируется адрес строки и элементы строки фиксируются в защелке строки.

Опрос элементов строки разрушает их содержимое, поэтому микросхема восстанавливает элементы строки из защелки одновременно с выдачей информации в систему. Это время называется временем перезаряда строки. Следующее обращение к памяти может начаться только после перезаряда строки. Цикл памяти завершается когда сигнал RAS становится пассивным.

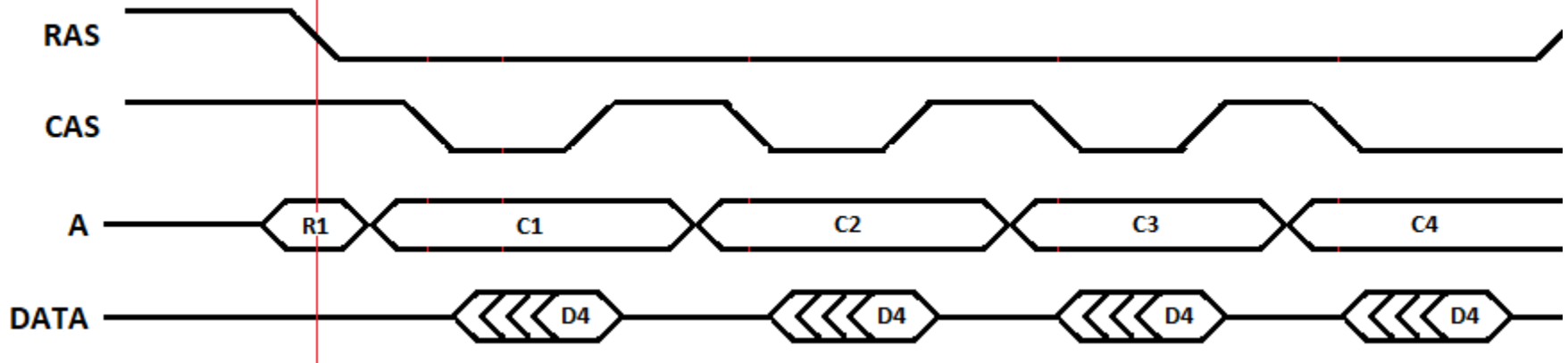
Способы повышение фактического быстродействия DRAM.

Чередование банков. Последовательные обращения осуществляются к разным банкам памяти. В то время пока первый банк выполняет перезаряд строки, и его трогать нельзя, по второму банку можно обращаться, а потом наоборот

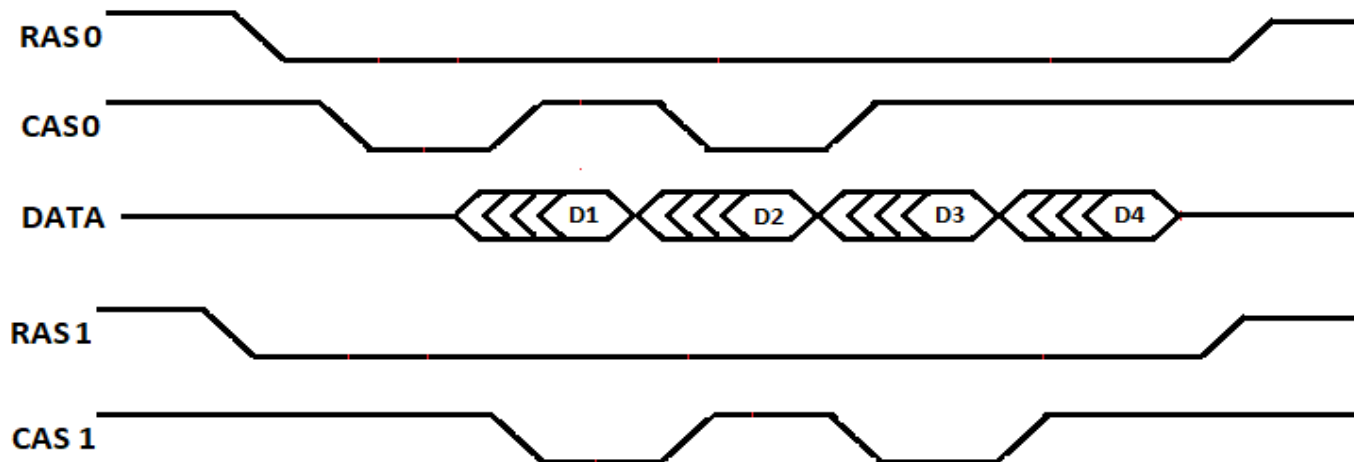


Быстрый страничный режим (Fast Page Mode).

Так как при обращении к памяти по адресу RAS происходит считывание всей строки в защелку строки, то целесообразно последующие обращения выполнять к элементам этой же строки.



Возможно совместное применение страничного режима и чередования банков.

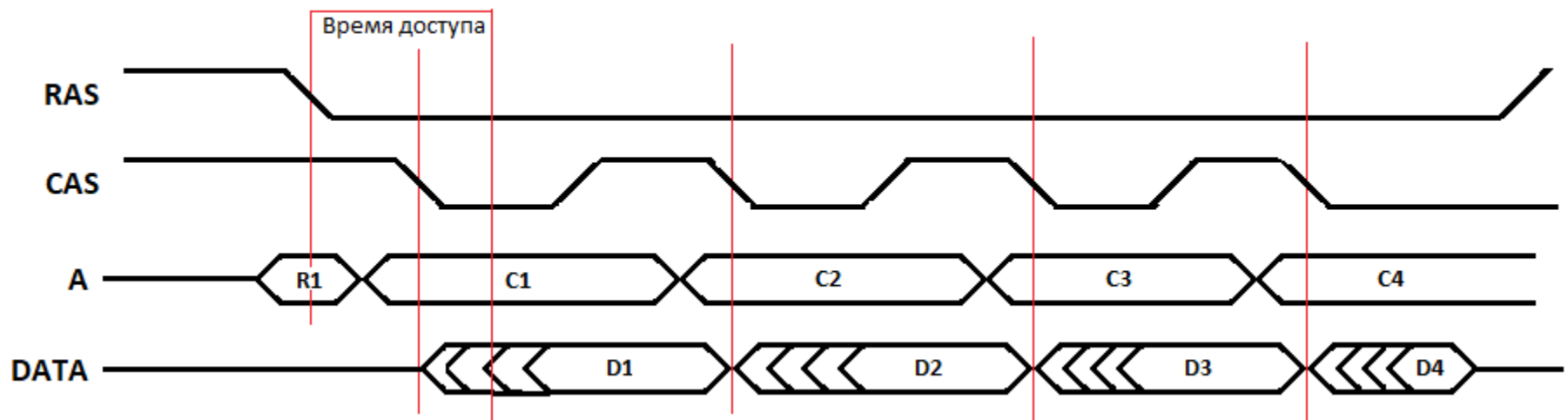


EDO (Extend Data output) DRAM.

Информация на выходе стандартной памяти DRAM появляется с некоторой задержкой относительно спада импульса CAS и держится только во время низкого уровня этого сигнала. После подъёма CAS выходной буфер переводится в третье состояние.

Память EDO DRAM содержит регистр-защелку выходных данных, что обеспечивает конвейеризацию работы для повышения производительности при чтении. Регистр прозрачен при низком CAS, а по подъёму фиксируют текущее значение выходных данных до следующего спада.

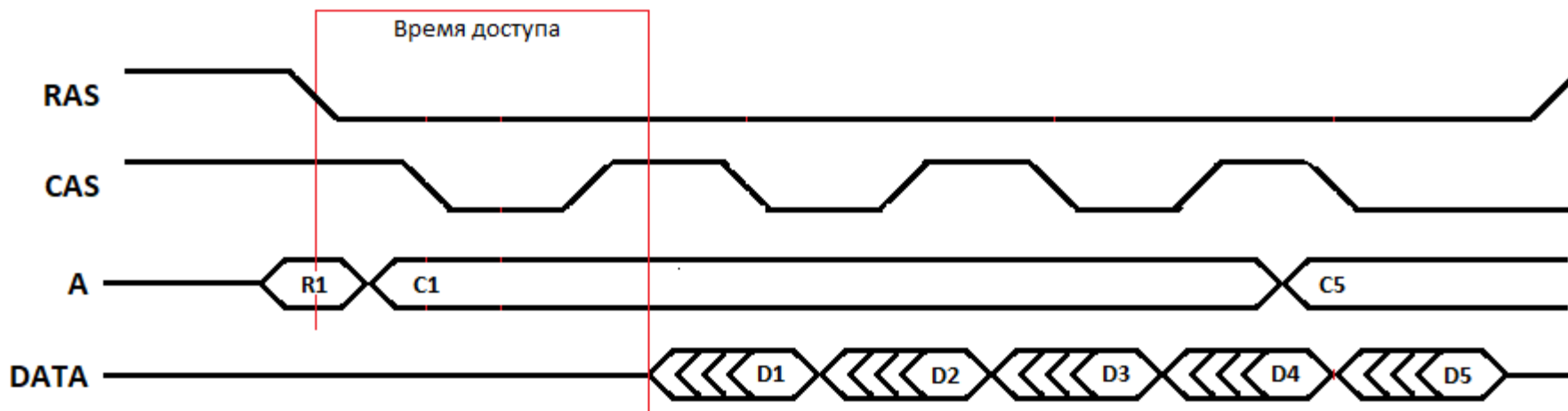
HPM (Hyper Page Mode) – страничный режим обмена. Реализуется когда память EDO DRAM работает в режиме FPM.



BEDO (Burst EDO) DRAM

В микросхемах BEDO DRAM дополнительно содержится счётчик адресов столбцов. При обращении к группе слов (пакету) адрес столбца формируется обычным способом только в начале пакетного цикла. Для последующих передач, адреса образуются с помощью инкрементирования счётчика во 2-й, 3-й и 4-й передачах импульсы CAS только запрашивают очередные данные.

Память BEDO не получила широкого распространение из-за появления сильного конкурента – синхронных DRAM (SDRAM).



SDRAM (Synchronous DRAM) – синхронная DRAM.

Появление микропроцессоров с шинами на 100 МГц привело к созданию **SDRAM**. SDRAM работает синхронно с контроллером это гарантирует завершение цикла в строго заданный срок.

В SDRAM используется полноразрядный внутренний счетчик. Количество матриц (банков памяти) увеличено до 2-ух, а в некоторых моделях до 4-х. Это позволяет обращаться к ячейкам одного банка параллельно с перезарядкой внутренних цепей другого банка.

Появилась возможность одновременного открытия двух (четырёх) страниц, причём открытие одной страницы может происходить во время считывания информации с другой страницы, что позволяет обращаться по новому адресу столбца ячейки памяти на каждом тактовом цикле.

DDR SDRAM (Double Data Rate).

SDRAM с удвоенной скоростью передачи данных.

- Удвоение скорости достигается за счёт передачи данных и по фронту и походу тактового импульса.
- Количество матриц (банков памяти) увеличилось до 4-х. Каждый банк имеет свой персональный контроллер.
- В результате вместо одной микросхемы имеем как бы 4, работающие независимо друг от друга.
- Максимальное количество ячеек обрабатываемых за один такт, возросло с одной до четырех.

RSRAM (Rambus DRAM).

Отличие от памяти предыдущих поколений:

1. Увеличение тактовой частоты за счёт сокращения разрядности шины.
2. Одновременная передача номеров строки и столбца ячейки.
3. Увеличение количества банков для усиления параллелизма.