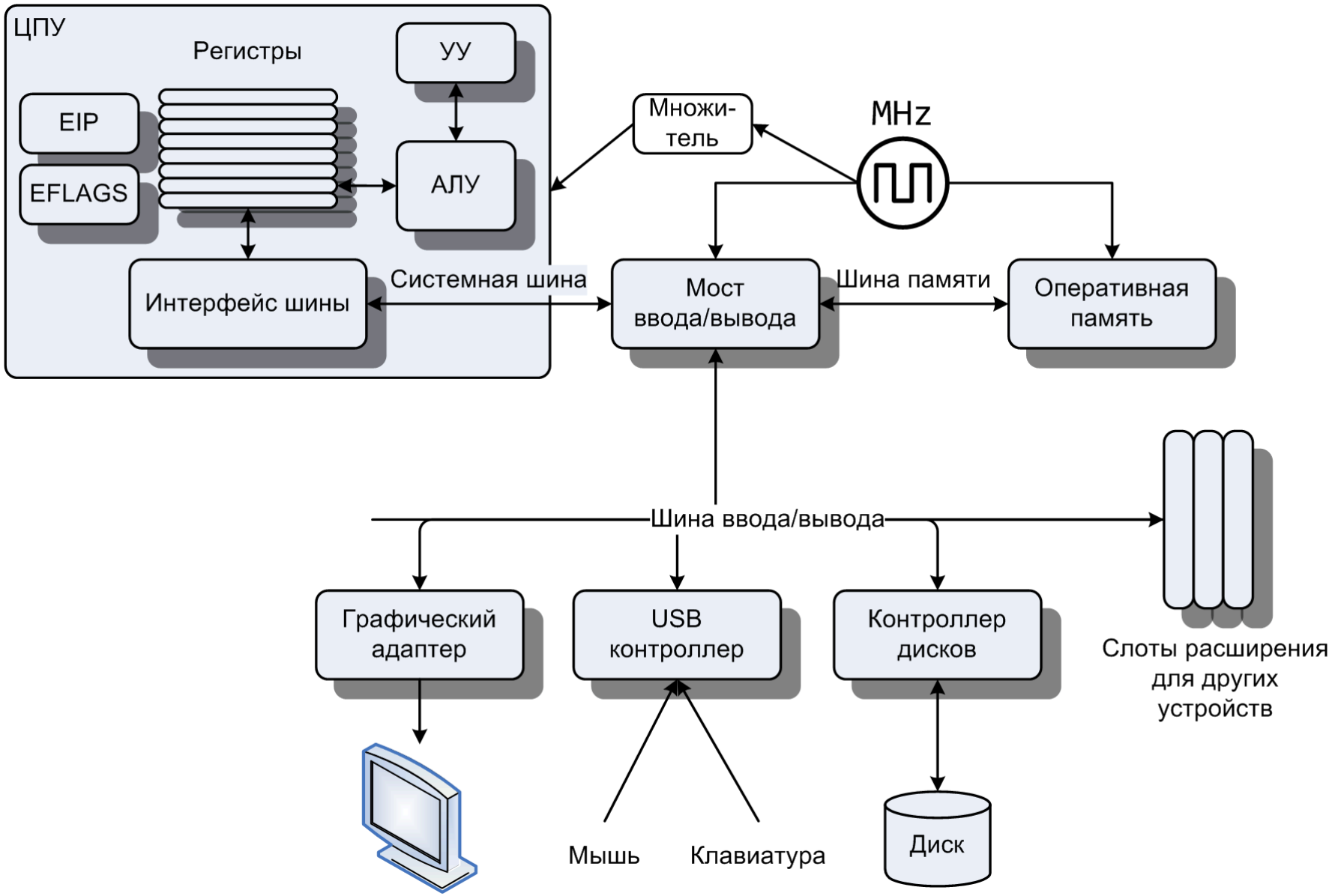
© 2014 МГУ/ВМК/СП

Лекция 20

19 апреля

© 2014 МГУ/ВМК/СП

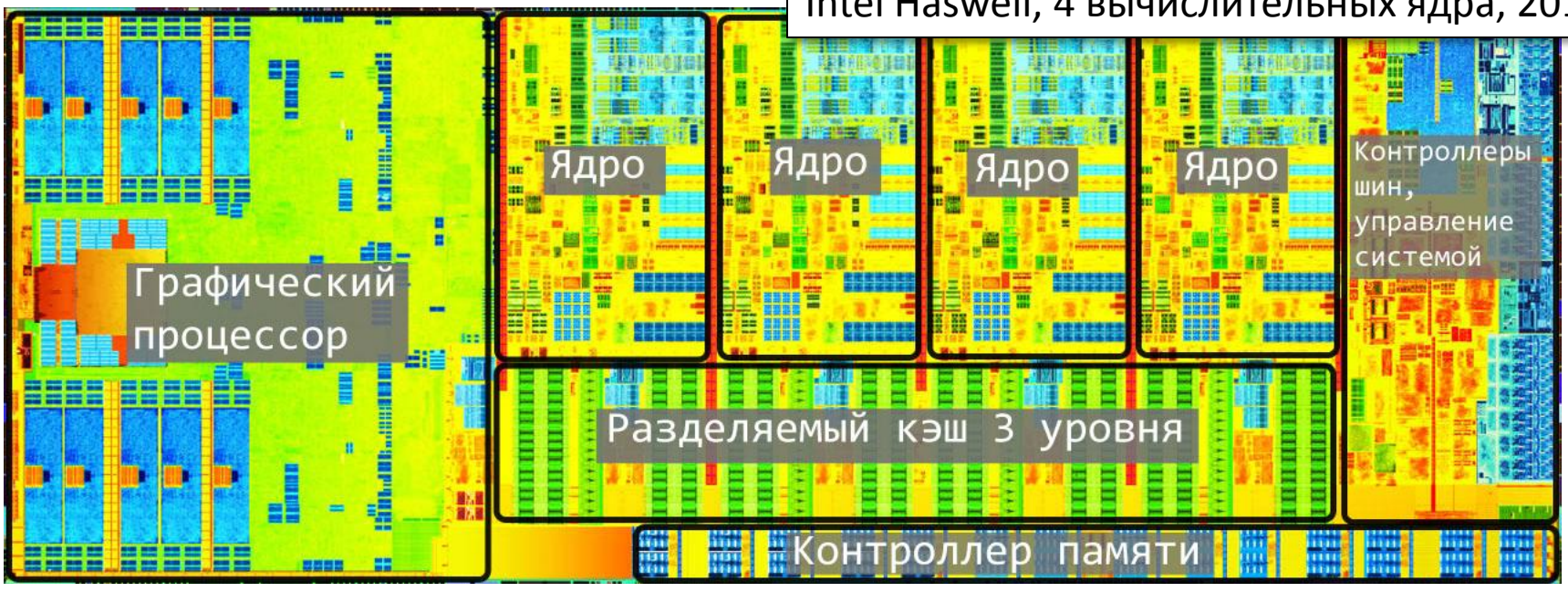


© 2014 МГУ/ВМК/СП

Закон Мура (Moore's law)

- Число транзисторов на кристалле будет удваиваться каждые 24 месяца
- Гипотеза выдвинута в 1965 году Гордоном Муром (один из основателей Intel)
- Ограничения
  - Атомарная природа вещества
  - Скорость света
- Негативная сторона – предельно быстрое устаревание вычислительной техники
- Открытый вопрос: область применимости

Intel Haswell, 4 вычислительных ядра, 2013



© 2014 МГУ/ВМК/СП

Закон Гроша (Grosch's law)

- Производительность компьютера увеличивается как квадрат стоимости
  - Емкость мирового рынка компьютеров – 5 машин
- Гипотеза выдвинута в 1965 году Хербом Грошем (второй ведущий ученый IBM, после Эккерта)
- 1997: закон полностью опровергнут
- Применимость к определенному классу машин
  - Рабочая станция
  - Майнфрейм
  - Суперкомпьютер
- Новые вычислительные/информационные ресурсы
  - Поисковые системы
  - Облачные вычисления

© 2014 МГУ/ВМК/СП

Различные классы компьютеров

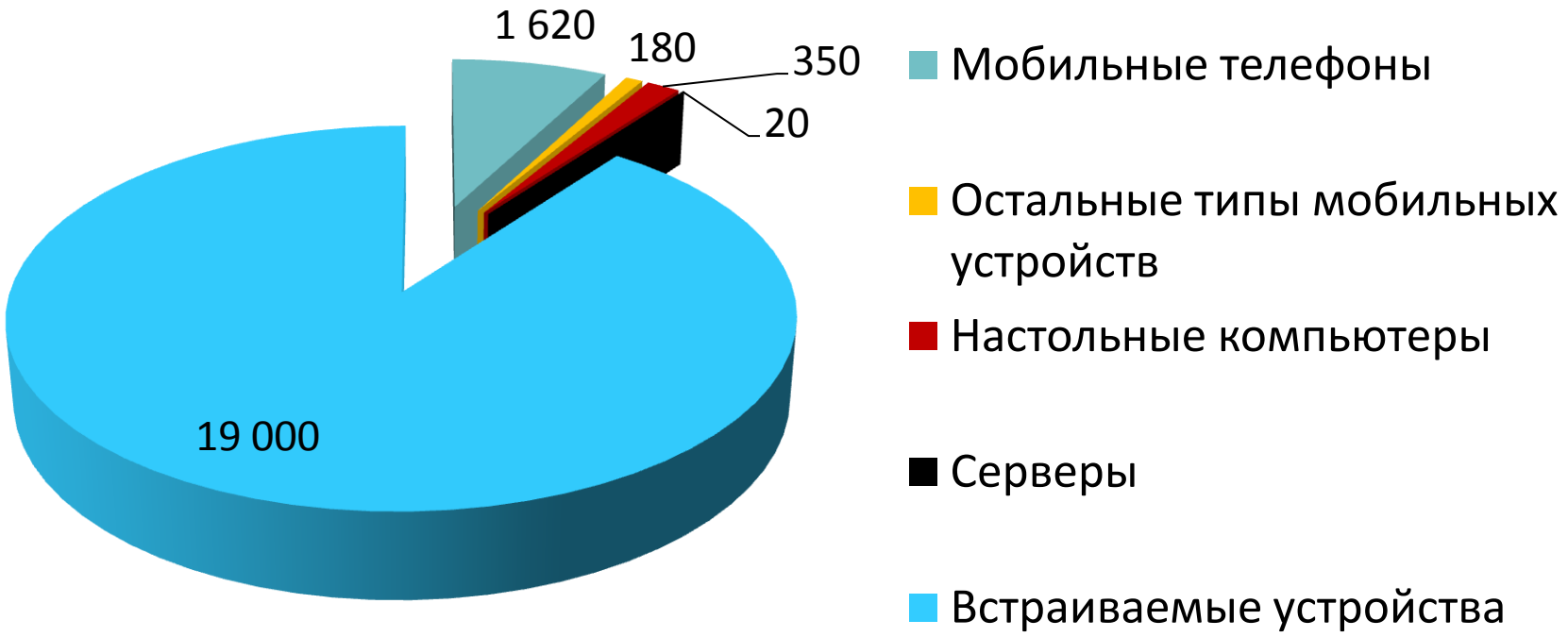
RFID

Характе-ристика	Персональ-ные мобиль-ные устройства	Настольные компьютеры	Серверы	Кластеры / ЦОД	Встраива-емые системы
Цена всей системы	$100 - $1000	$300 - $2500	$5000 - $10000000	$100000 - $200000000	$10 - $100000
Цена процессора	$10 - $100	$50 - $500	$200 - $2000	$50 - $250	$0.01-$100
Критические требования	Стоимость, энергопо-требление , производи-тельность при работе с аудио/видео, быстрота реакции	Цена/произ-водитель-ность, энергопо-требление, производи-тельность графики	Пропускная способность, готовность, масштаби-руемость, энергопо-требление	Цена/произ-водитель-ность, пропускная способность, энергети-чекая пропорци-ональность	Цена, энергопо-требление , производи-тельность при решении прикладных задач

© 2014 МГУ/ВМК/СП

Различные классы компьютеров: кого больше?

Мировые продажи в 2010 году (млн. шт.)



Процессоры архитектуры ARM: отгружено с заводов 6.1 миллиарда шт.

© 2014 МГУ/ВМК/СП

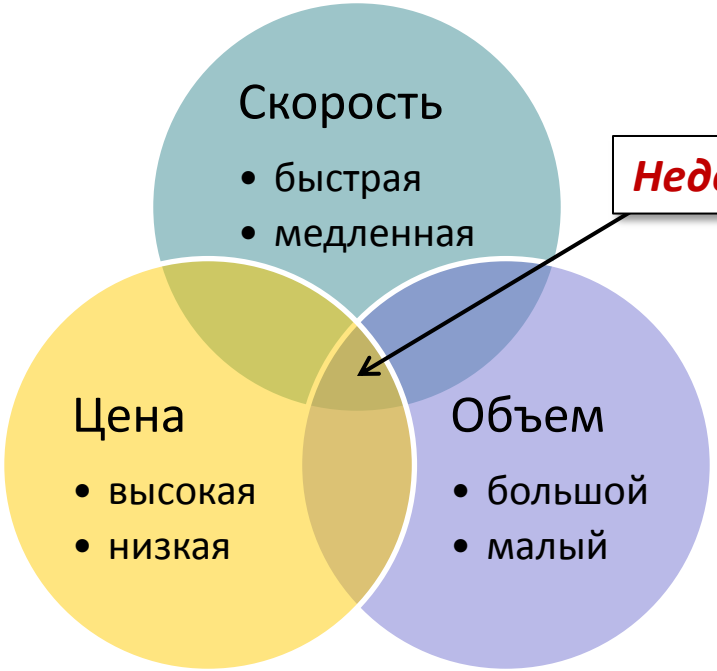
Оперативная память (RAM)

- Основные свойства
  - RAM традиционно оформляется в виде отдельного чипа.
  - Единица хранения – клетка/ячейка (один бит на ячейку).
  - Оперативная память состоит из нескольких чипов RAM.
- Статическая память (SRAM)
  - Каждая ячейка хранит значение одного бита с помощью схемы из 4 или 6 транзисторов.
  - При наличии питания, сохраняет значение неограниченно долго.
  - Относительно устойчива к радиации, ЭМП
  - Быстрее и дороже чем DRAM.
- Динамическая память (DRAM)
  - Состоит из конденсатора и транзистора.
  - Сохраняемое значение должно обновляться каждые 10-100 мс.
  - Более чувствительная к воздействиям (ЭМП, радиация,…) чем SRAM.
  - Медленней и дешевле чем SRAM.

© 2014 МГУ/ВМК/СП

SRAM vs DRAM

Тип памяти	Транз. на 1 бит	Относ. время доступа	Устойчивая	Контроль	Относ. стоимость	Применение
SRAM	4 или 6	1×	Да	Нет	100×	Кэш
DRAM	1	10×	Нет	Да	1×	Основная оперативная память



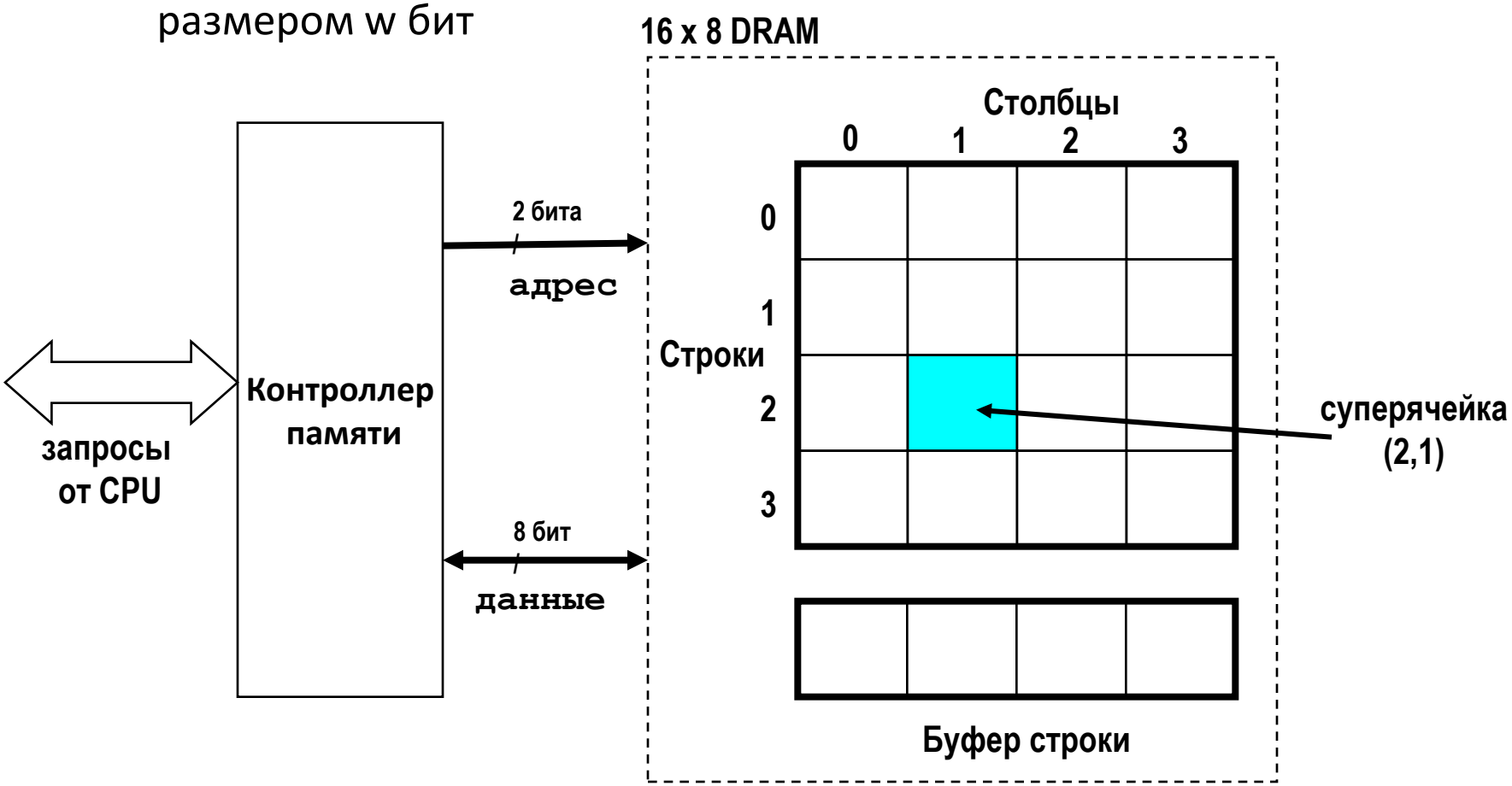
Недостижимо

Можно выбрать преимущества только двух характеристик в ущерб третьей

© 2014 МГУ/ВМК/СП

Типовая организация DRAM

- d × w DRAM:
  - Общий объем данных dw бит организован как d суперячеек размером w бит

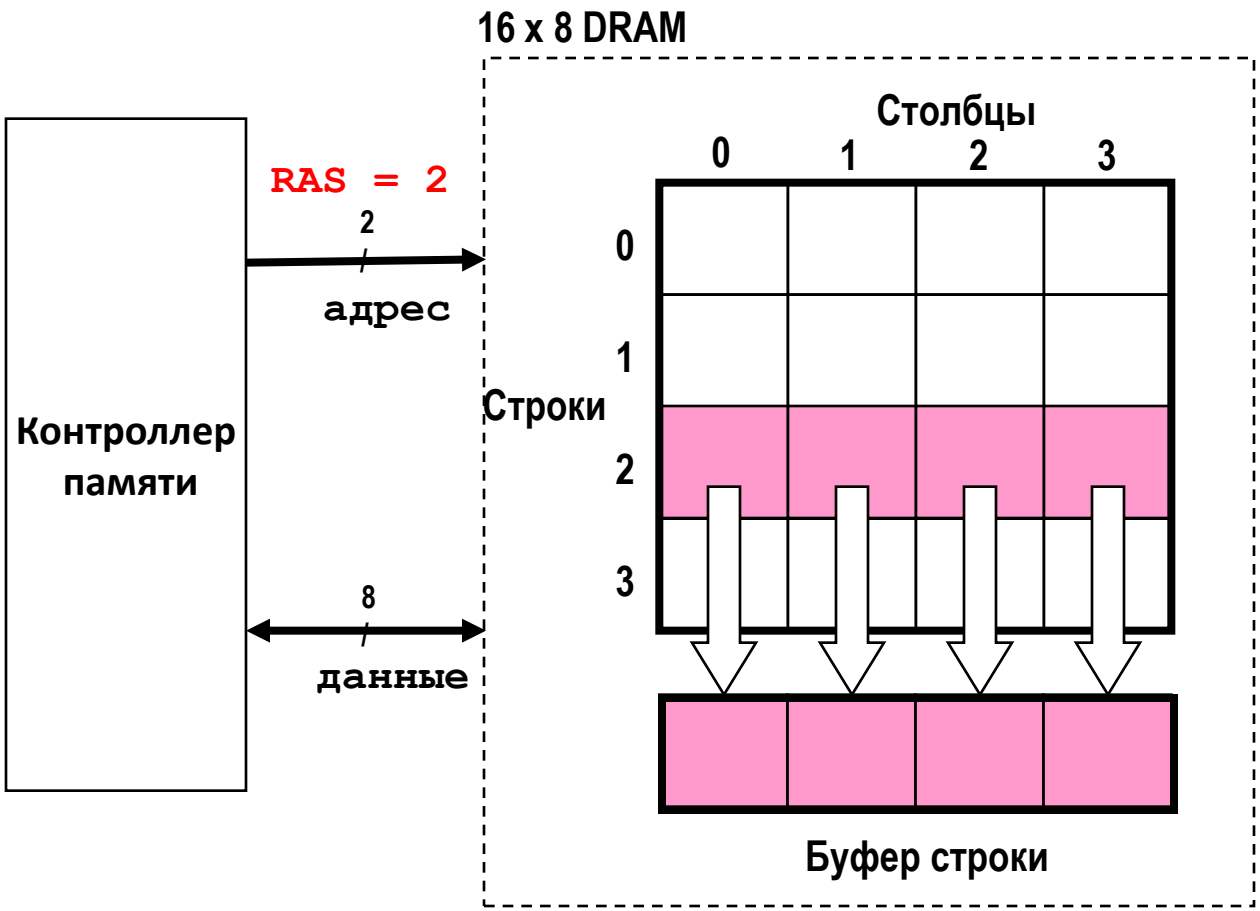


© 2014 МГУ/ВМК/СП

Чтение суперячейки DRAM (2,1)

Шаг 1(а): Строб адреса строки (RAS) указывает строку 2.

Шаг 1(б): Строка 2 копируется из DRAM в буфер строки.

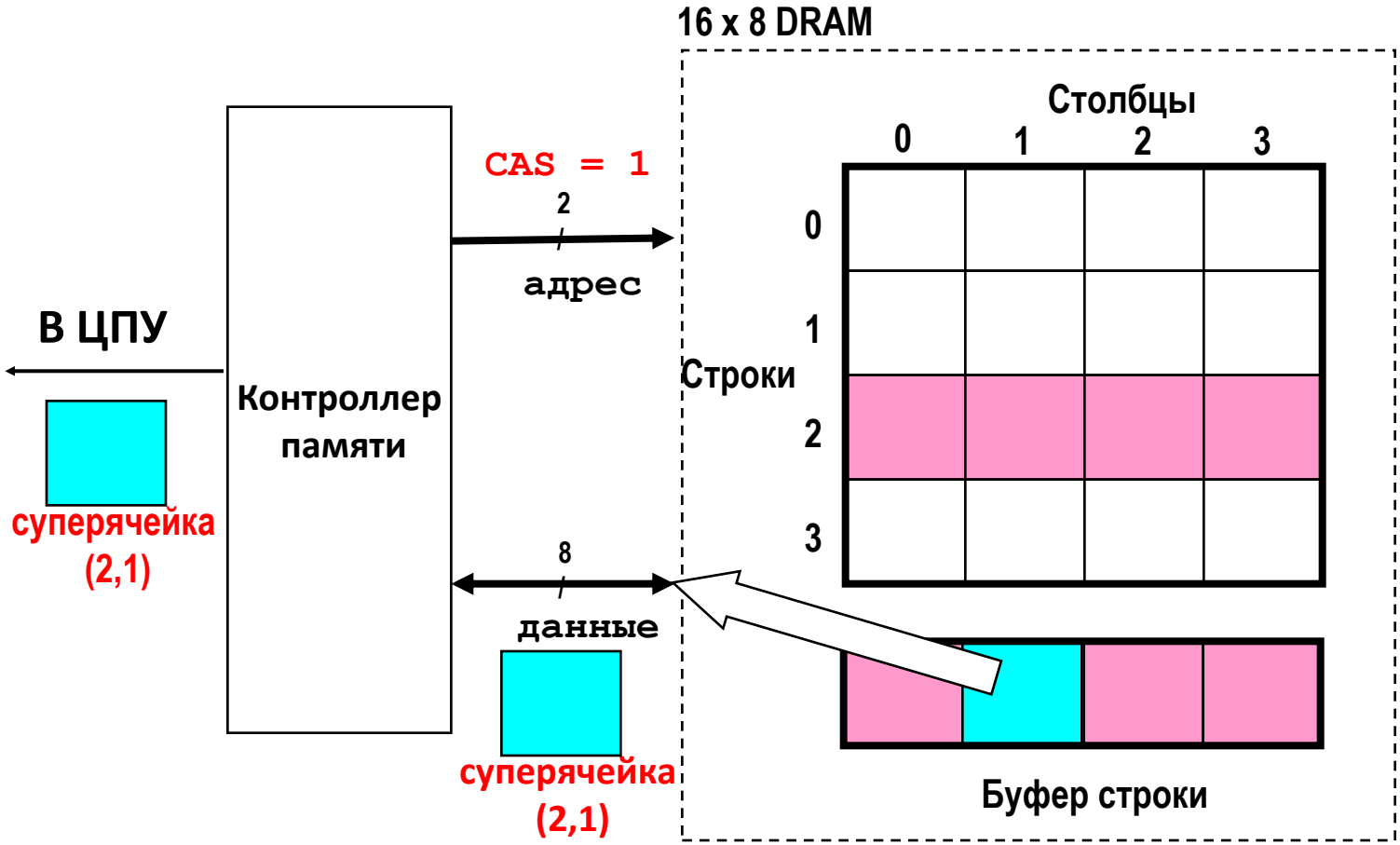


© 2014 МГУ/ВМК/СП

Чтение суперячейки DRAM (2,1)

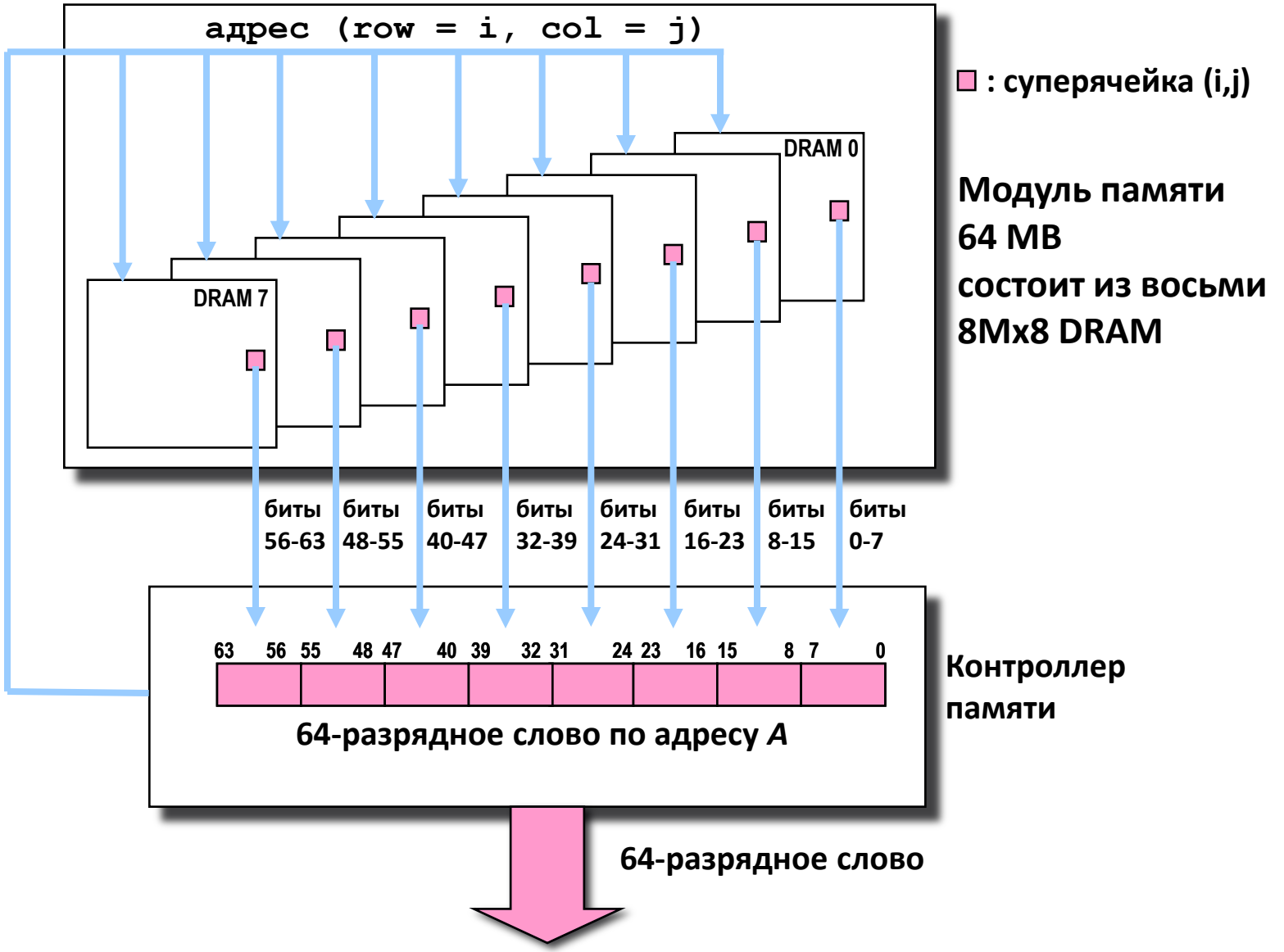
Шаг 2(а): Строб адреса столбца (CAS) указывает столбец 1.

Шаг 2(б): Суперячейка (2,1) копируется из буфера на линии шины данных и далее в ЦПУ.



© 2014 МГУ/ВМК/СП

Расслоение памяти



© 2014 МГУ/ВМК/СП

Развитие DRAM

- Организация ячейки DRAM принципиально не менялась с момента изобретения в 1966 году.
  - Коммерческий выпуск начат Intel в 1970.
- Модули DRAM с улучшенным интерфейсом:
  - Синхронная DRAM (SDRAM)
    - Синхронизируется с системными часами
    - Позволяет повторно использовать адрес строки (т.е., RAS, CAS, CAS, CAS)
  - Синхронная DRAM с удвоенной частотой (DDR SDRAM)
    - Управляется фронтами – две посылки данных за один такт
    - В 2014 году, стандартная память для большинства серверов и настольных компьютеров
      - Intel Core i7 поддерживает только DDR3 SDRAM
      - Готовится к выпуску DDR4 SDRAM

© 2014 МГУ/ВМК/СП

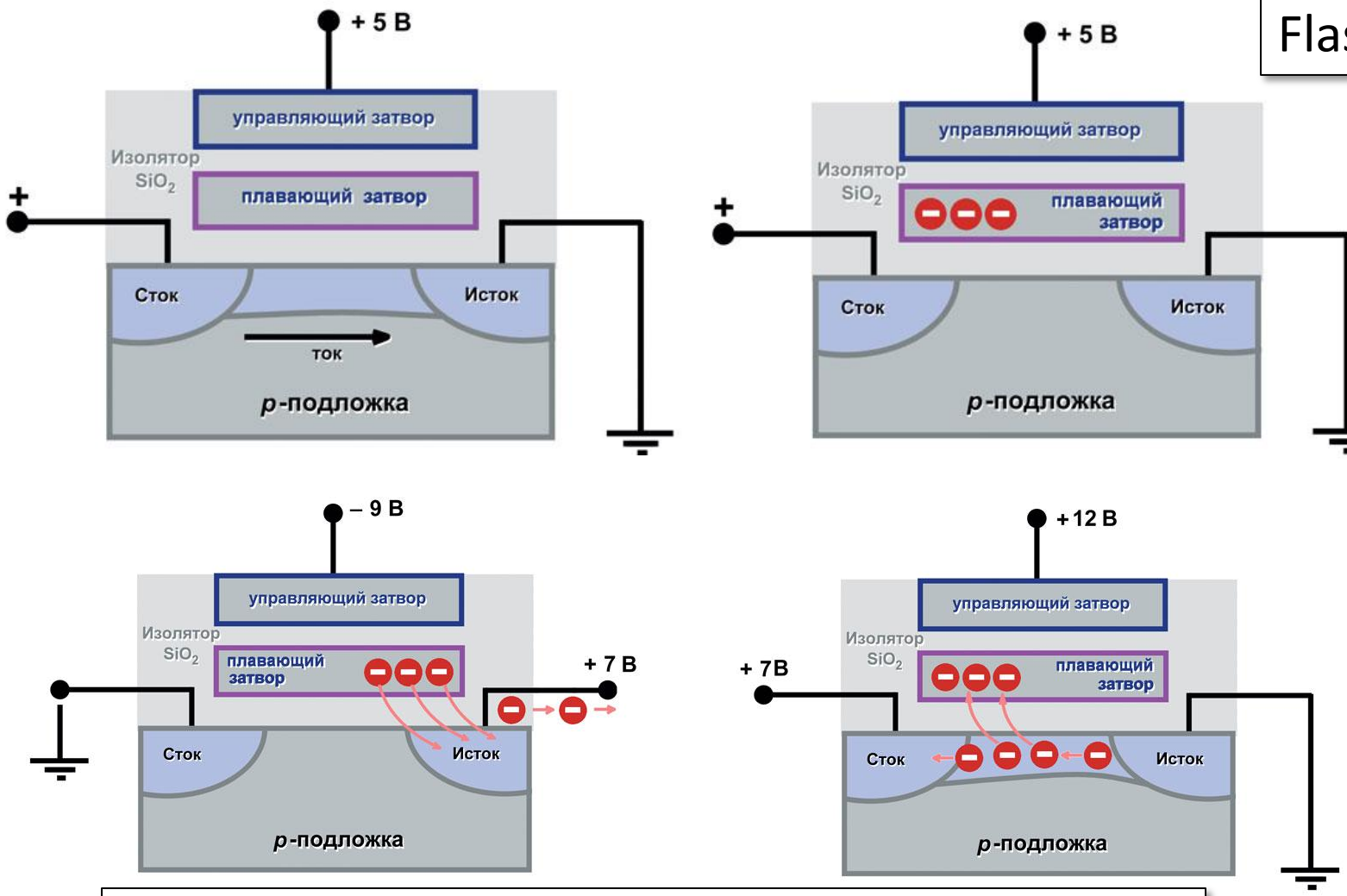
Энергонезависимая память

- DRAM и SRAM – энергозависимы
  - Информация теряется при отключении электропитания.
- Энергонезависимая память сохраняет свое состояние даже при отключении питания
  - ROM: программируется на этапе производства
  - Программируемая ROM (PROM): может быть запрограммирована пользователем один раз
  - Стираемая PROM (EPROM): может быть стерта (УФ, рентген)
  - Электрически стираемая PROM (EEPROM): стирание происходит через подачу электрического сигнала
  - Флеш-память: EEPROM с частичной возможностью стирания (по секторам)
    - Выдерживает порядка 100,000 циклов перезаписи.
- Сфера применения энергонезависимой памяти
  - Встраиваемые программы размещаются в ROM (BIOS, контроллеры дисков, сетевых и графических адаптеров, аппаратно-криптографические средства,…)
  - Твердотельные диски (заменяют обычные диски в переносных накопителях, смартфонах, плеерах, и т.д.)
  - Кеш в обычных дисковых накопителях.

© 2014 МГУ/ВМК/СП

Энергонезависимая память

Flash

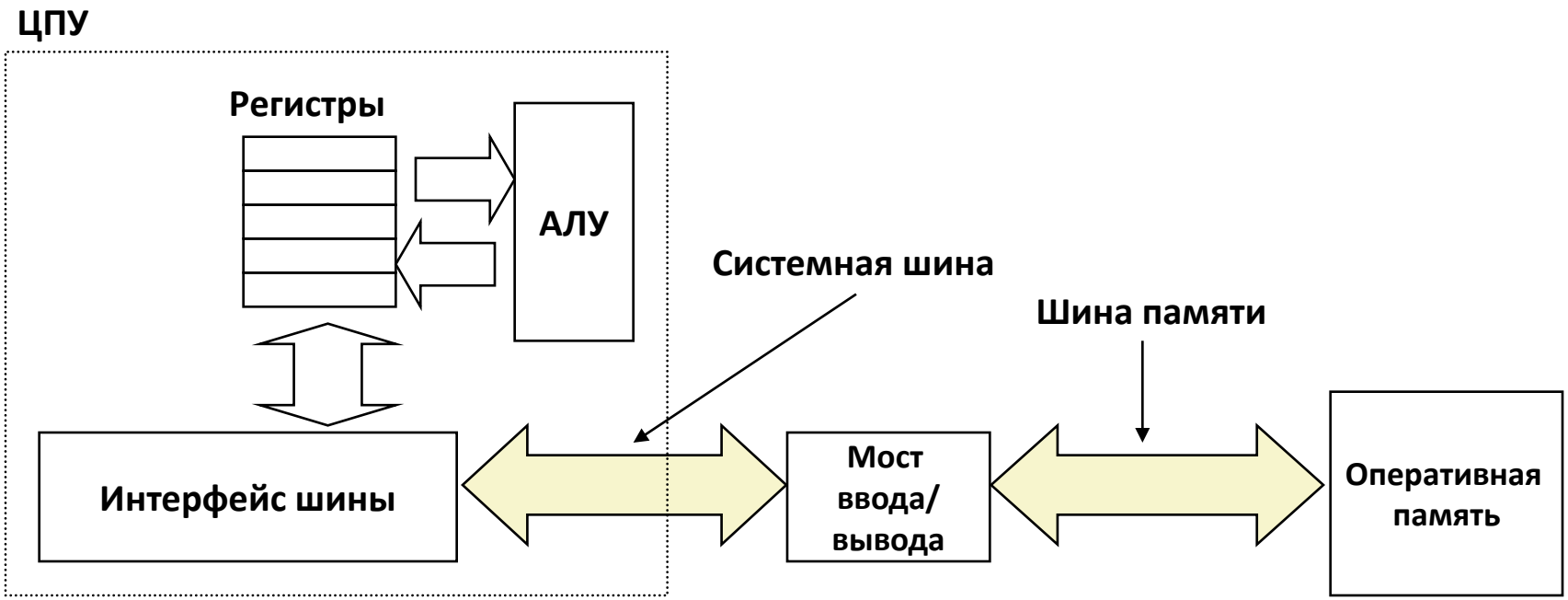


Для записи и стирания используется
«квантовый эффект туннелирования Фаулера-Нордхейма»

© 2014 МГУ/ВМК/СП

Типовое соединение ЦПУ и оперативной памяти

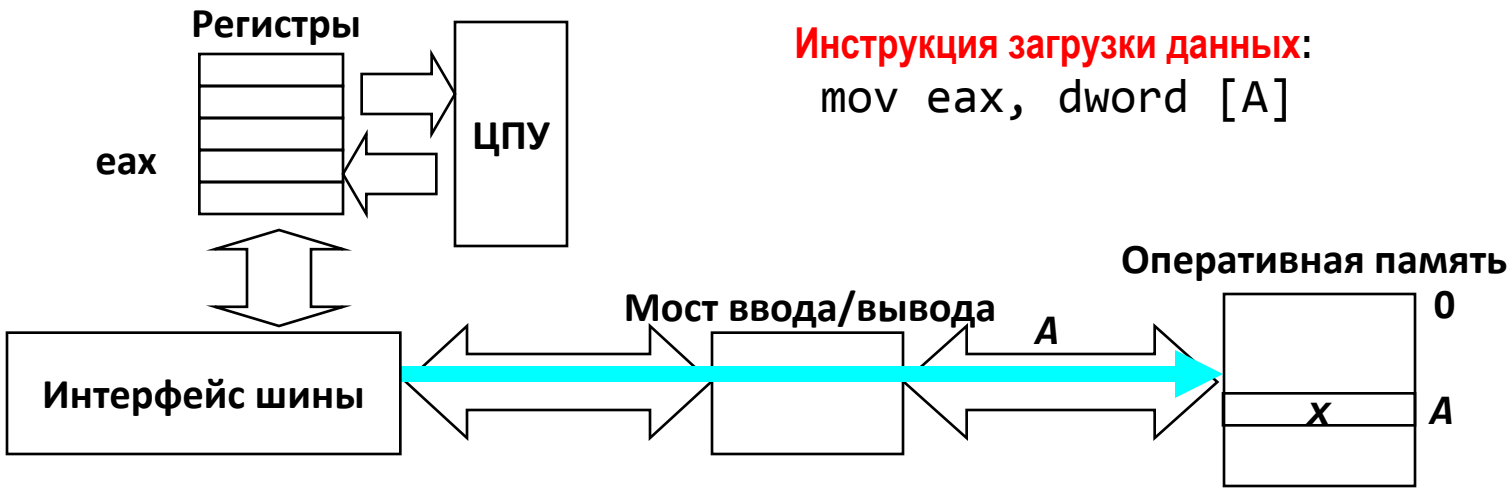
- Шина – набор проводов используемых для передачи данных, адресов, управляющих сигналов.
- Шины, как правило, используются несколькими устройствами.



© 2014 МГУ/ВМК/СП

Чтение данных из памяти (1)

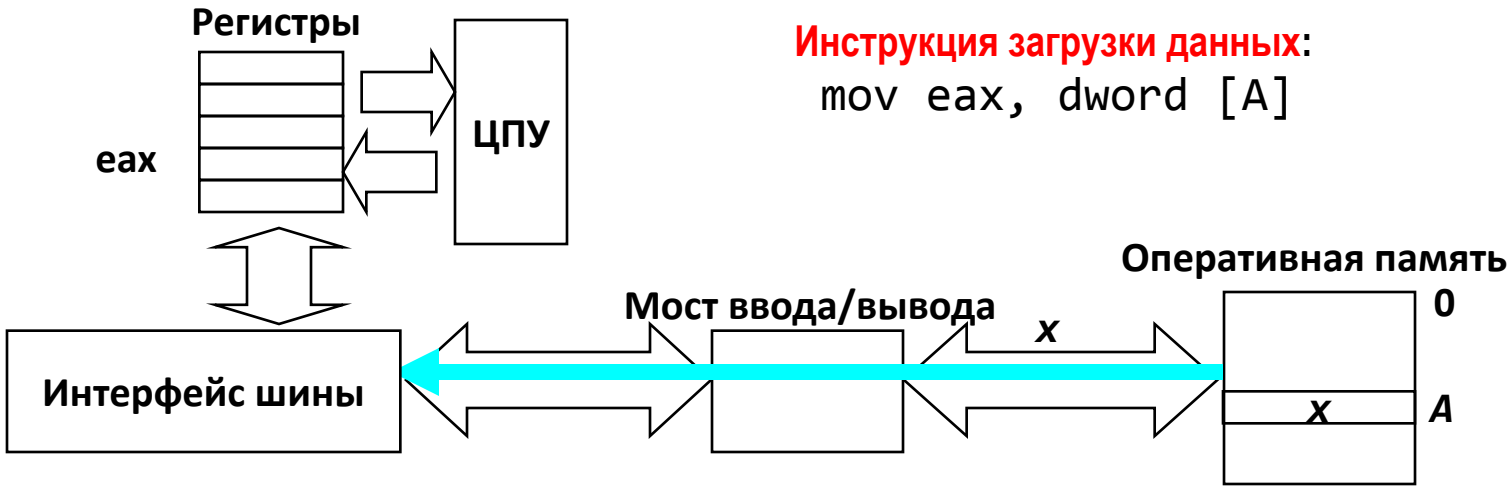
- ЦПУ передает адрес A интерфейсу шины памяти.



© 2014 МГУ/ВМК/СП

Чтение данных из памяти (2)

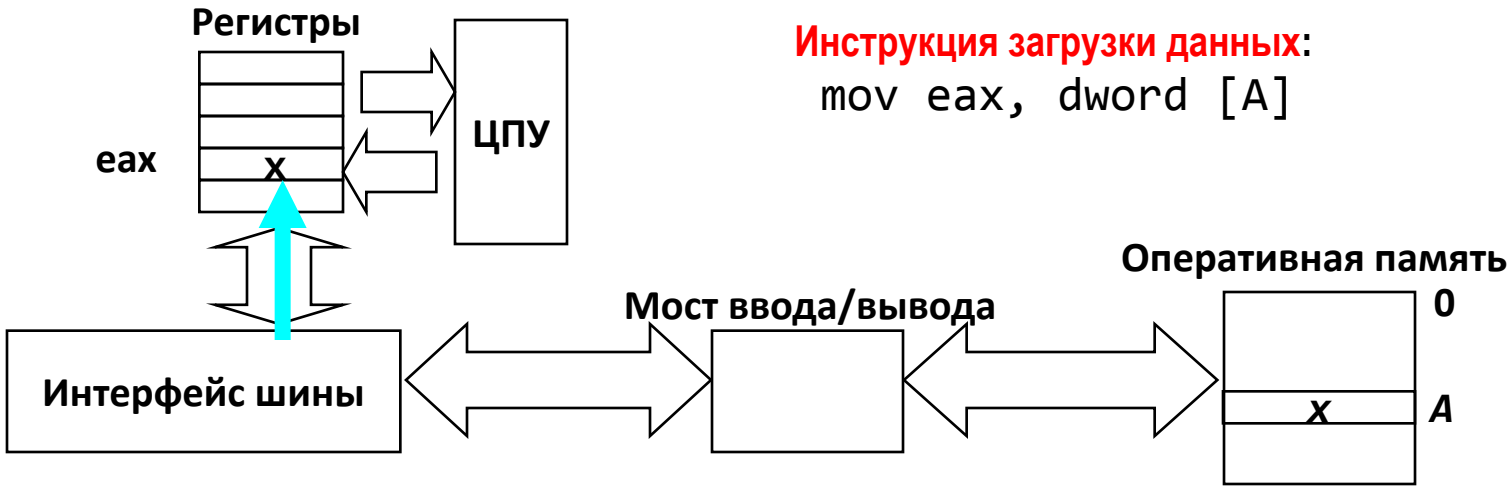
- Оперативная память получает запрос на выборку данных по адресу A из шины, осуществляет выборку значения x, и отправляет его назад, в шину.



© 2014 МГУ/ВМК/СП

Чтение данных из памяти (3)

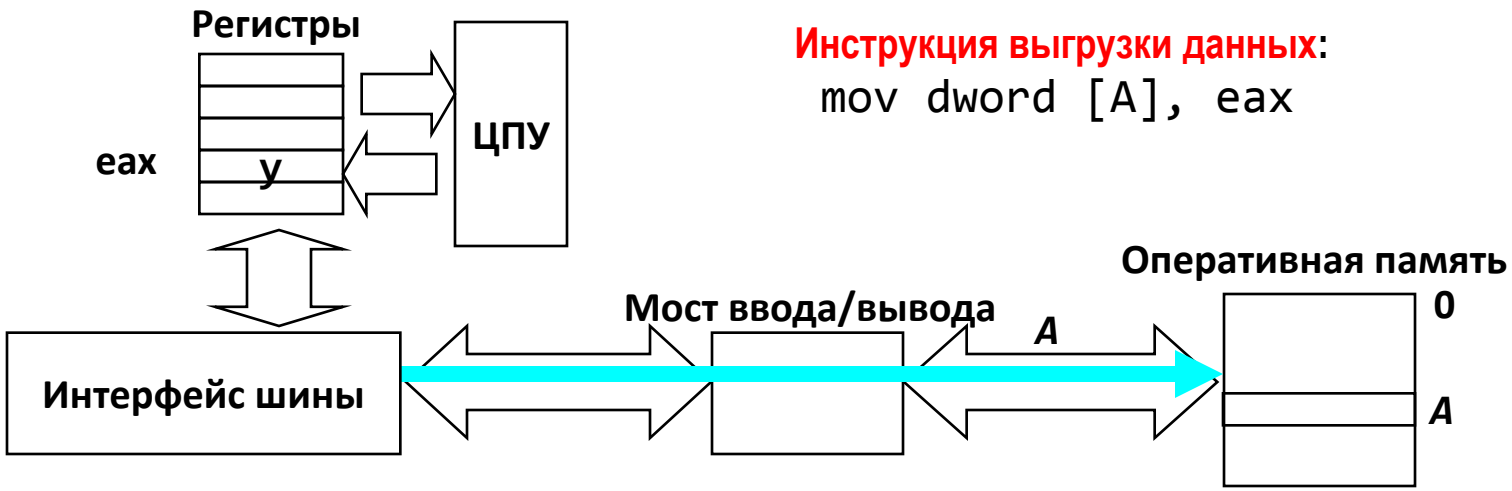
- ЦПУ считывает двойное слово x из шины и пересылает его в регистр eax.



© 2014 МГУ/ВМК/СП

Запись данных в память (1)

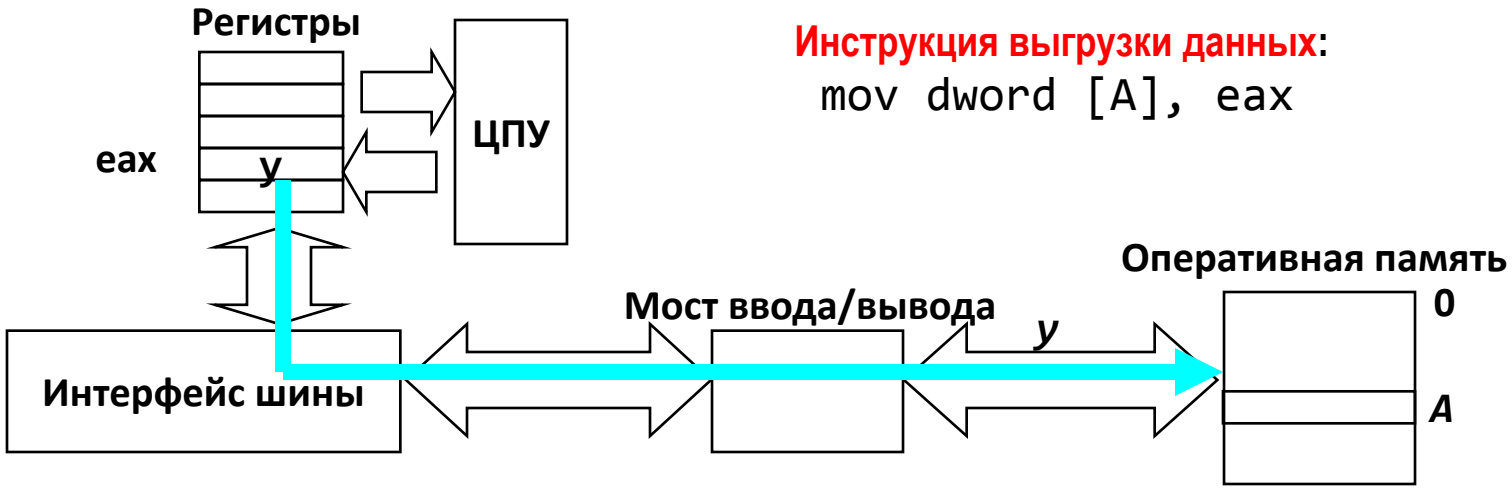
- ЦПУ передает адрес A интерфейсу шины. Оперативная память считывает адрес и ждет посылки соответствующего значения.



© 2014 МГУ/ВМК/СП

Запись данных в память (2)

- ЦПУ передает значение y интерфейсу шины.

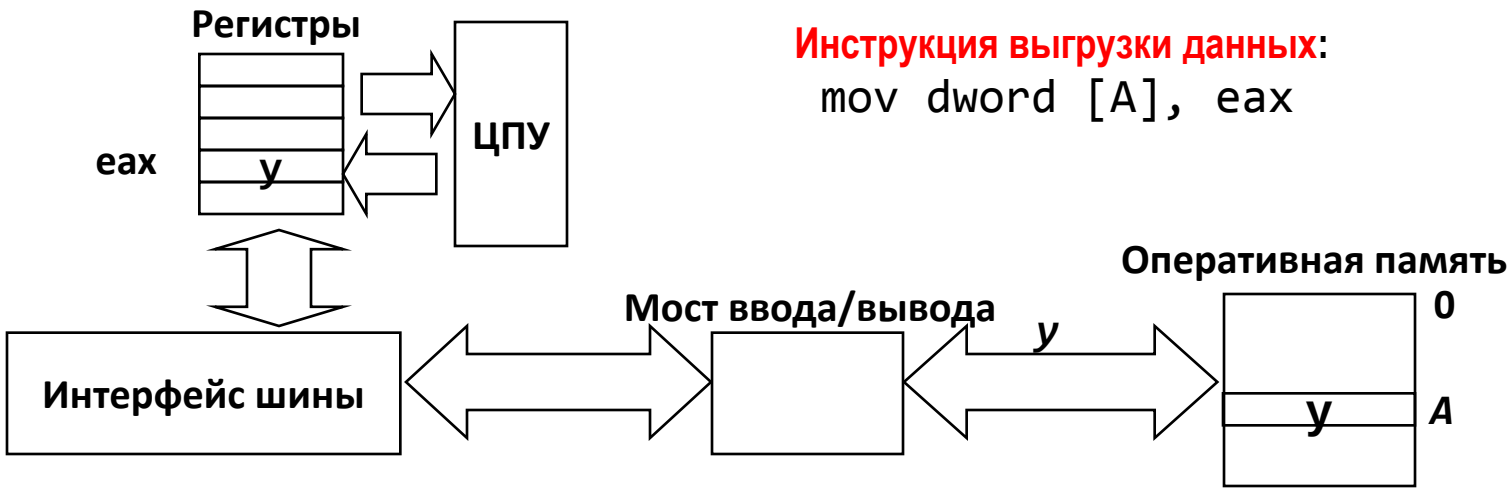


21

© 2014 МГУ/ВМК/СП

Запись данных в память (3)

- Оперативная память получает двойное слово y из шины и сохраняет его по адресу A.

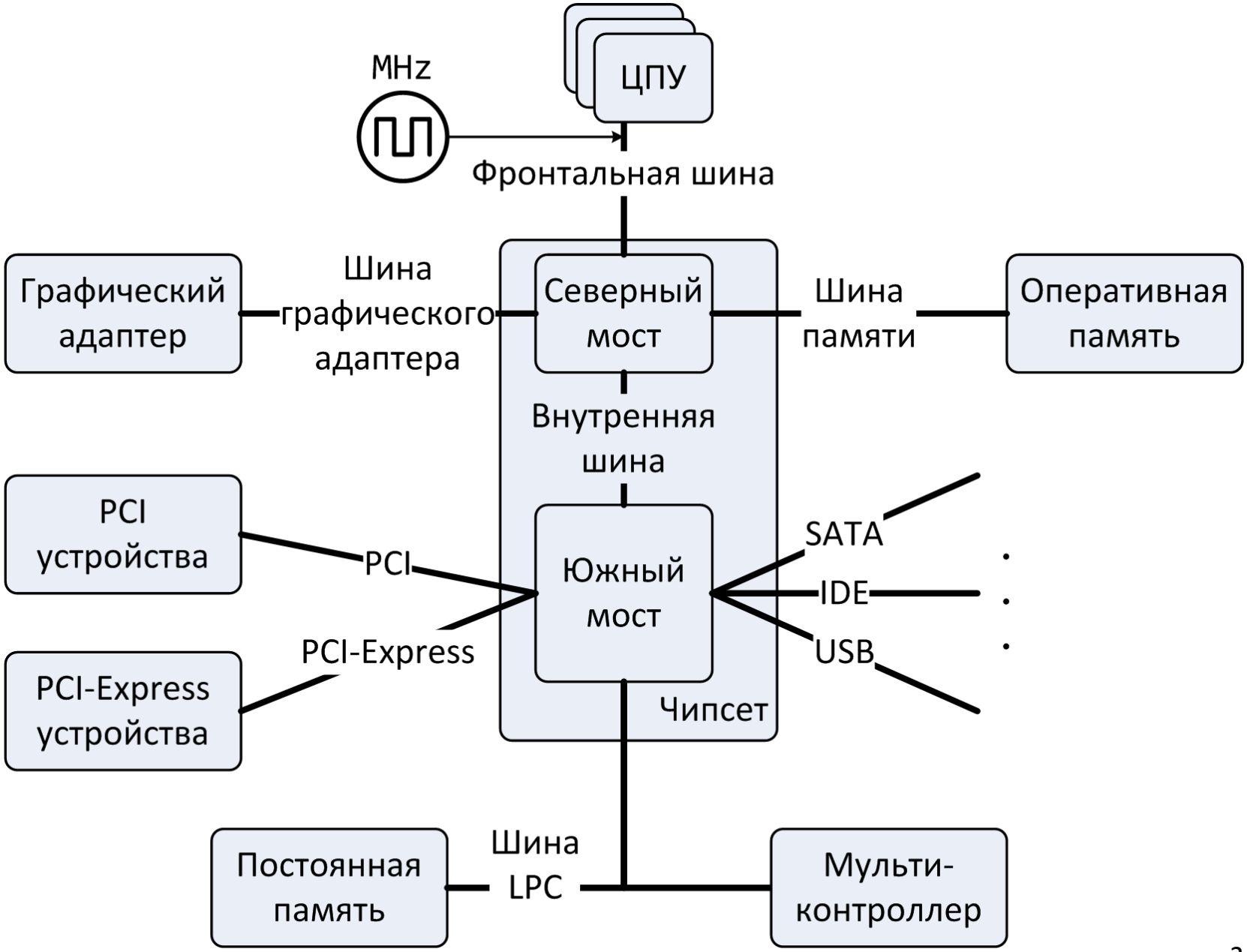


© 2014 МГУ/ВМК/СП

Характеристики шин

- Ширина
  - Количество линий, линии адреса, линии данных, мультиплексирование адресов и данных
- Частота
- Пиковая пропускная способность
  - Ширина (байты или биты) × Частота (1/сек)
- Арбитраж
  - Централизованный
    - Линии запроса и захвата шины
  - Децентрализованный
- Возможность горячей замены устройства
- Физическая организация
  - Выделенные линии данных, адресов, команд
  - Мультиплексированные линии
  - Топология связей
  - Синхронная/асинхронная
  - Ограничения по длине линии

© 2014 МГУ/ВМК/СП



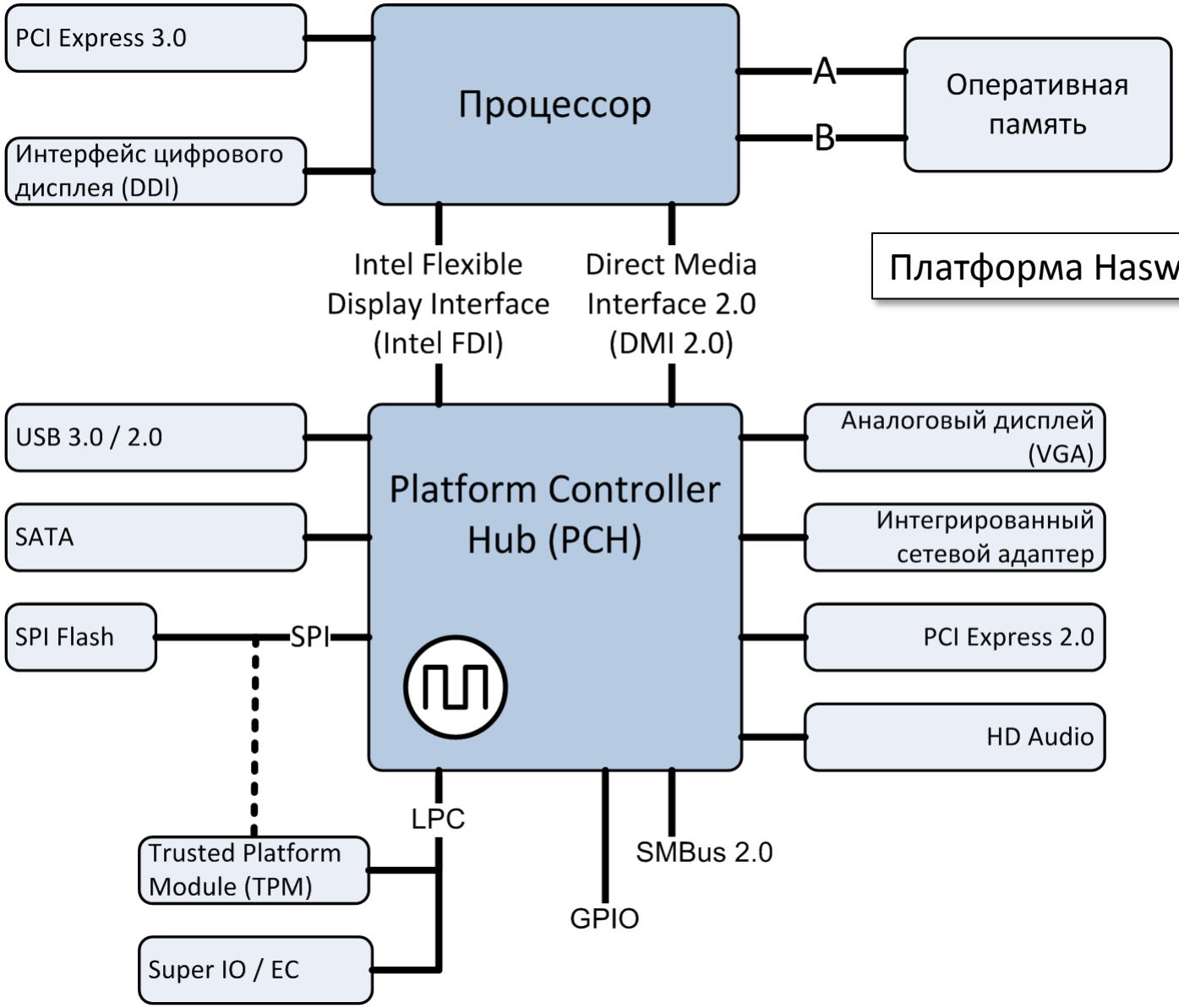
© 2014 МГУ/ВМК/СП

Примеры шин (1/3)

- Фронтальная шина
  - HyperTransport (HT), апрель 2001, AMD. Открытый стандарт - HyperTransport Technology
    - 2 – 32 разряда, двунаправленная
    - 200 – 2600 MHz, DDR
  - QuickPath Interconnect (QPI), ноябрь 2008, Intel
    - 20 линий, двунаправленная, 4 такта = 64 бита
    - 2.4, 2.93, 3.2 GHz, DDR
  - Соединение точка-точка.
    - Гарантированные физические каналы: один отправитель – один приемник.
    - Не требуется арбитраж.

© 2014 МГУ/ВМК/СП

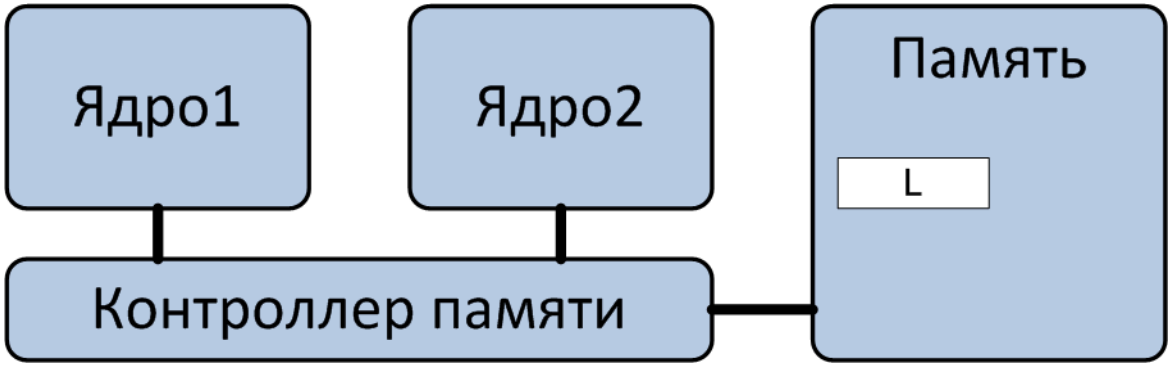
Пример реализации



© 2014 МГУ/ВМК/СП

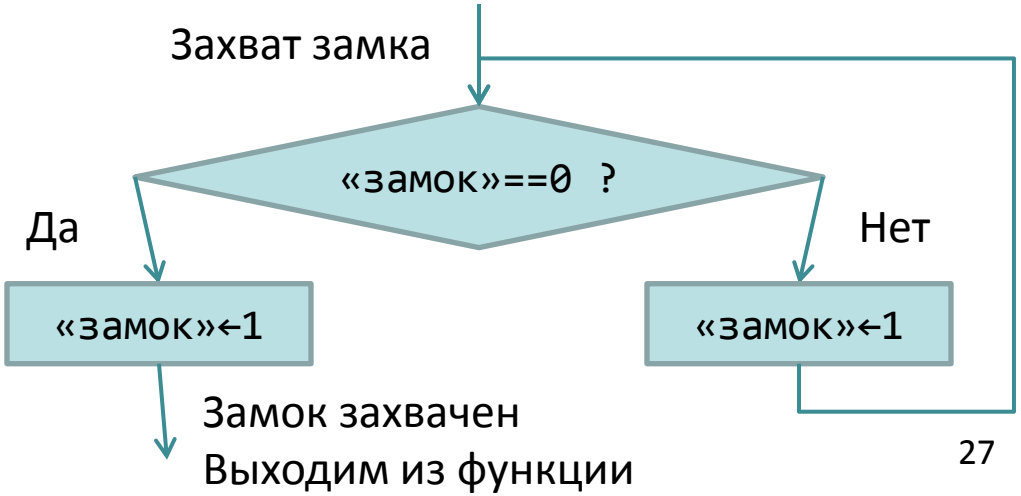
Синхронизация обращений к памяти

- Захват блокировки
- Работа с общими (разделяемыми) переменными
- Освобождение блокировки



Замок – статическая переменная L (младший разряд)
- 0 – доступ открыт
- 1 – доступ закрыт

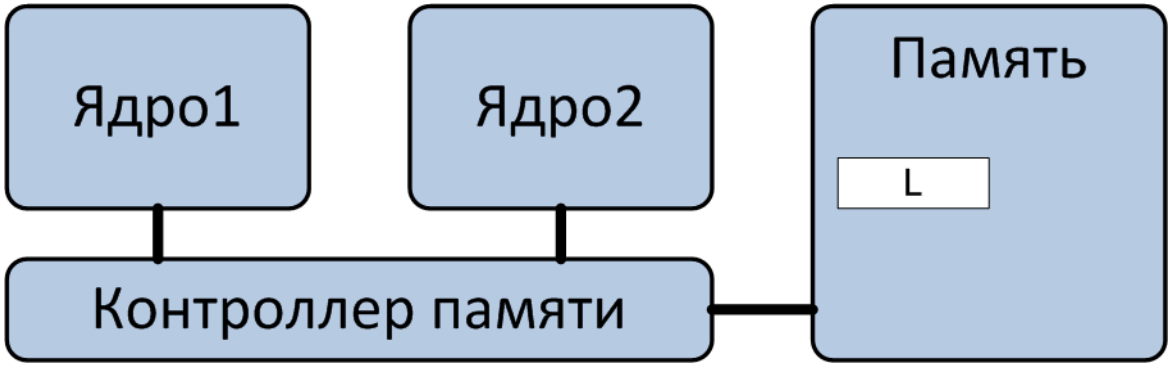
```
acquireLock:
 .retry:
   bts byte [L], 0
   jc .retry
   ret

releaseLock:
   btr byte [L], 0
   ret
```



© 2014 МГУ/ВМК/СП

Синхронизация обращений к памяти

- Захват блокировки
- Работа с общими (разделяемыми) переменными
- Освобождение блокировки

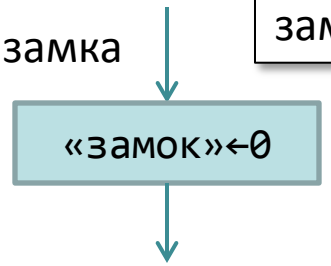


Замок – статическая переменная L (младший разряд)
- 0 – доступ открыт
- 1 – доступ закрыт

Начальное состояние замка – 0

```
acquireLock:
 .retry:
   bts byte [L], 0
   jc .retry
   ret

releaseLock:
   btr byte [L], 0
   ret
```

Освобождение замка

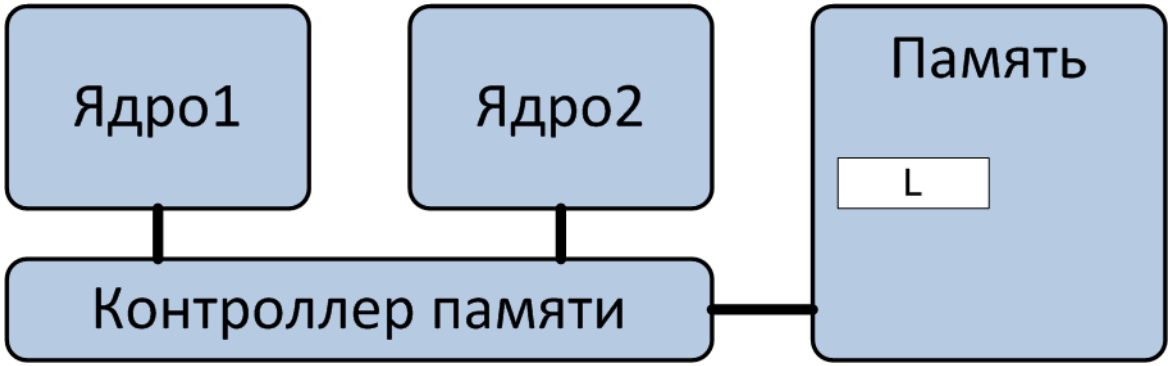


Замок освобожден
Выходим из функции

© 2014 МГУ/ВМК/СП

Синхронизация обращений к памяти

- Захват блокировки
- Работа с общими (разделяемыми) переменными
- Освобождение блокировки

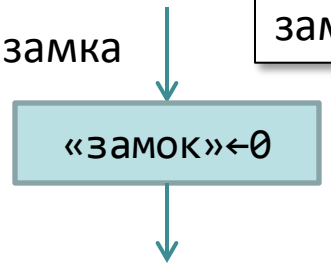


Замок – статическая переменная L (младший разряд)
- 0 – доступ открыт
- 1 – доступ закрыт

Начальное состояние замка – 0

```
acquireLock:
 .retry:
   bts byte [L], 0
   jc .retry
   ret

releaseLock:
   btr byte [L], 0
   ret
```

Освобождение замка

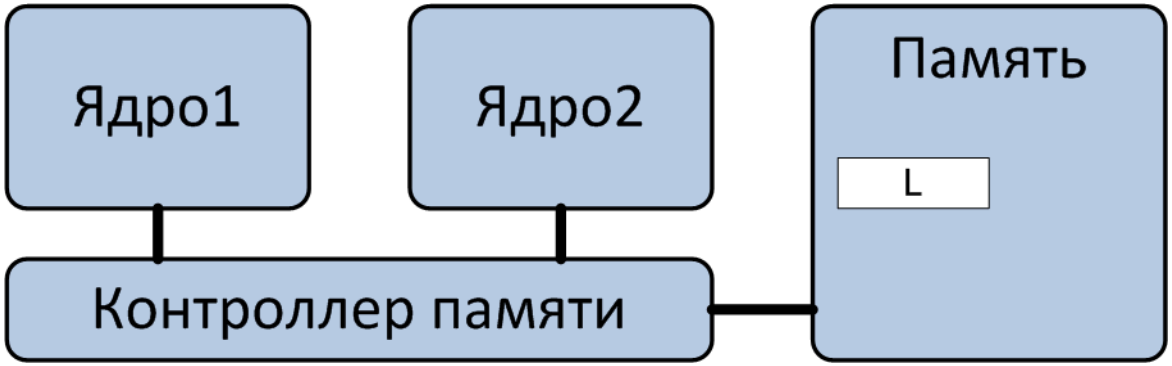


Замок освобожден
Выходим из функции

© 2014 МГУ/ВМК/СП

Синхронизация обращений к памяти

- Захват блокировки
- Работа с общими (разделяемыми) переменными
- Освобождение блокировки



Оба ядра попытались захватить замок «***почти***» одновременно

```
acquireLock:
 .retry:
   bts byte [L], 0
   jc .retry
   ret

releaseLock:
   btr byte [L], 0
   ret
```

Может возникнуть ***Состояние Гонки***

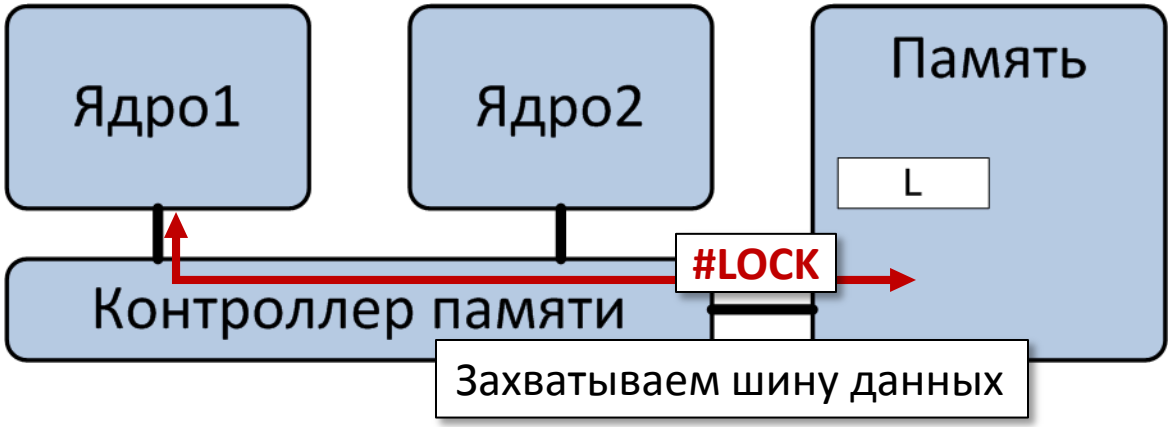
Ядро1		Ядро2	
ЦПУ ← память	BTS		
CF ← ЦПУ:L|0		ЦПУ ← память	BTS
память:L ← 1		CF ← ЦПУ:L|0	
		память:L ← 1	

Время

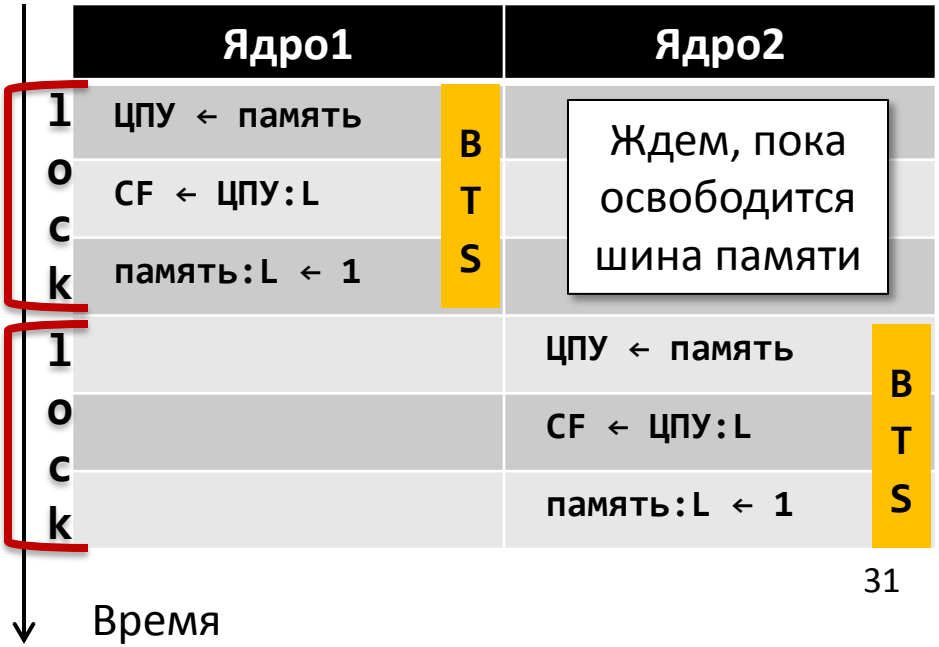
© 2014 МГУ/ВМК/СП

Синхронизация обращений к памяти

- Захват блокировки
- Работа с общими (разделяемыми) переменными
- Освобождение блокировки




```
acquireLock:
 .retry:
   lock bts byte [L], 0
   jc .retry
   ret

releaseLock:
   lock btr byte [L], 0
   ret
```



© 2014 МГУ/ВМК/СП

Синхронизация обращений к памяти

- Явно указываемый префикс `lock`
  - Применим только к некоторым командам ADD, AND, BTC, BTR, BTS, CMPXCHG, …
    - Первый операнд команды – память
- Все время выполнения команды процессор удерживает шину памяти, посылая на нее сигнал #LOCK
- Постоянный захват шины негативно сказывается на производительности

```
acquireLock:
   lock bts byte [L], 0
   jc .retry
   ret
 .retry:
   pause
   lock bts byte [L], 0
   jc .retry
   ret
```

В современных процессорах команда `pause` используется как подсказка, что выполнение находится в цикле активного ожидания (busy wait)

ОС Linux использует для реализации активного ожидания не `bts/btr`, а гораздо более быструю команду `cmpxchg`
http://heather.cs.ucdavis.edu/~matloff/50/PLN/lock.pdf (на английском)

© 2014 МГУ/ВМК/СП

Примеры шин (2/3)

- Общая шина ввода/вывода для подключения периферийных устройств Peripheral component interconnect (PCI), 1992, Intel, открытый стандарт.
  - PCI 1.0 / 2.0
    - Топология - общая шина, децентрализованный арбитраж
    - 32 линии, общие для адресов и данных
      - Передача данных транзакциями, возможна приостановка
    - Частота 33 MHz
  - Расширения
    - PCI 64, PCI 66, PCI 64/66, PCI-X (266 и 533 МГц)
  - PCI Express (PCI-E), июль 2002, Intel
    - Топология – звезда.

© 2014 МГУ/ВМК/СП

Примеры шин (3/3)

- AGP
  - Непосредственный доступ к оперативной памяти
- USB
  - Топология – звезда, до 127 устройств, разветвители, оконечные точки (15/15 + 1/1 управляющие).
  - 4 типа передач: управляющие, поточные, прерывания, изохронные
  - Open Host Controller Interface
    - Сам контроллер является PCI-устройством
- Serial ATA
  - 7 линий, 2 – прием, 2 – передача.
  - 1.5, 3, 6 GBit/s
  - Возможность горячей замены

© 2014 МГУ/ВМК/СП

Внутреннее устройство HDD

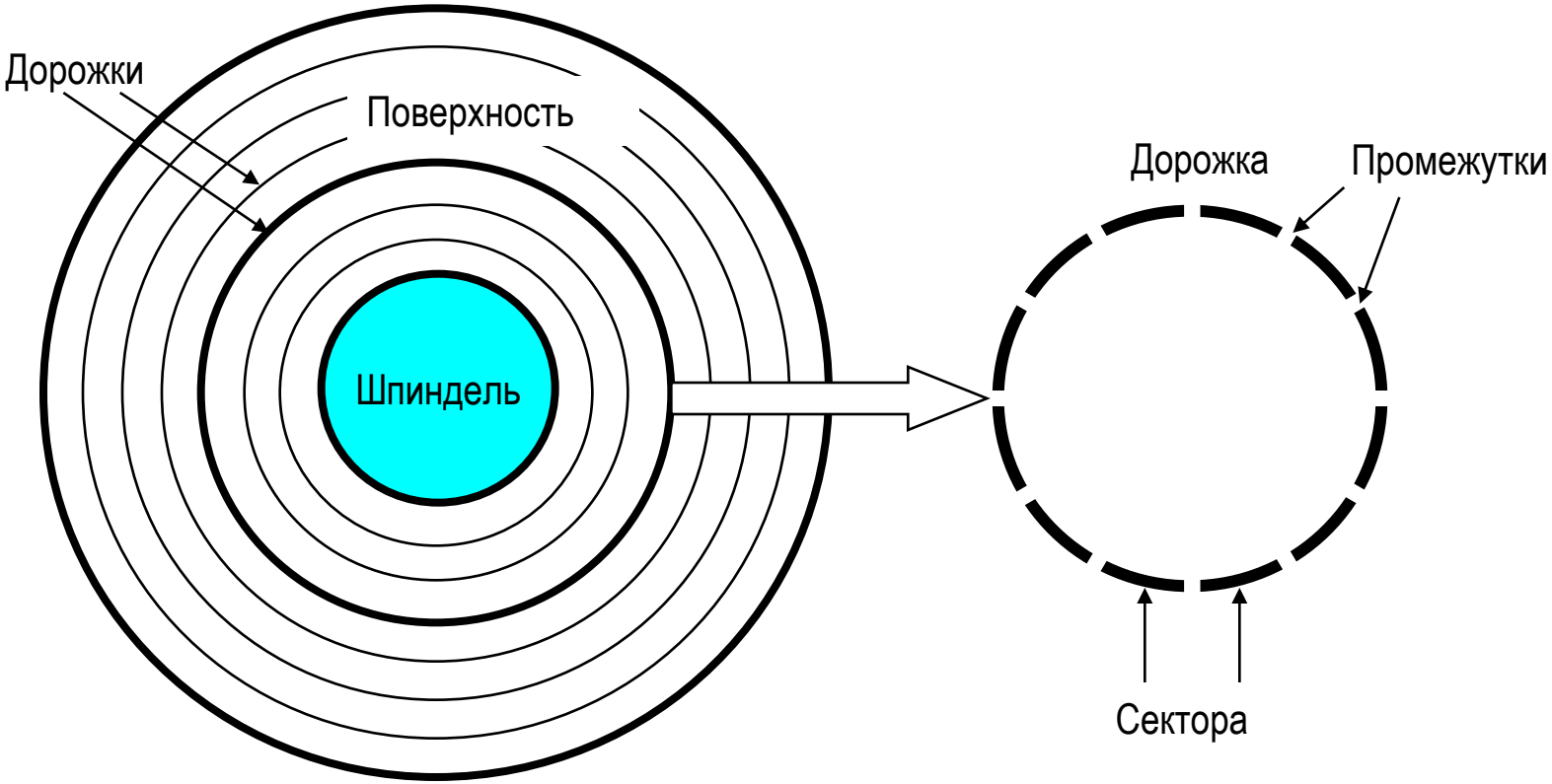
+ различная электроника, включая процессор и память

Изображение является собственностью Seagate Technology

© 2014 МГУ/ВМК/СП

Геометрия диска

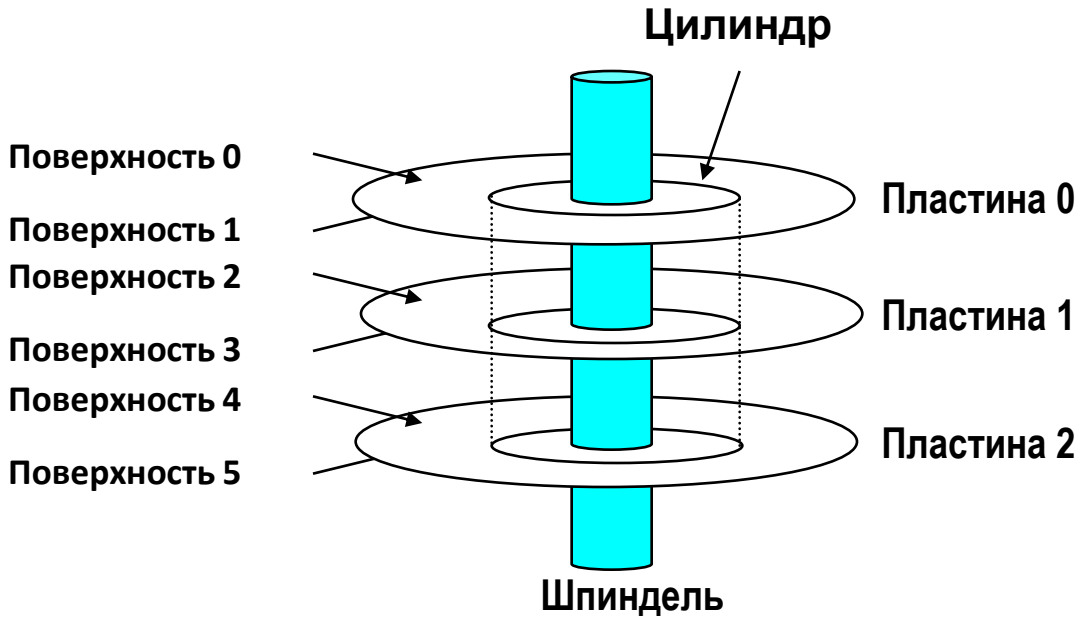
- Диск состоит из пластин, каждая обладает двумя поверхностями.
- Каждая поверхность состоит из концентрических кругов, называемых дорожки.
- Каждая дорожка состоит из секторов, разделенных промежутками.



© 2014 МГУ/ВМК/СП

Геометрия диска (несколько пластин)

- Равноудаленные от шпинделя дорожки образуют цилиндр.



© 2014 МГУ/ВМК/СП

Емкость диска

- Емкость: максимальное количество сохраняемых бит.
  - Производители выражают емкость в «необычных» гигабайтах, 1 ГБ = 10^9 байтам.
- Емкость определяется следующими технологическими факторами:
  - Плотность записи / линейная плотность (биты/дюймы – BPI): сколько битов может быть размещено на одном дюйме дорожки.
  - Трековая плотность (треки/дюйм – TPI): сколько треков может быть размещено на одном дюйме радиуса.
  - Поверхностная плотность (биты/дюймы^2): произведение линейной плотности на трековую плотность.
- Современные диски группируют дорожки в несколько зон записи
  - Каждая дорожка в зоне состоит из одного и того же количества секторов, определяемого длиной самой короткой дорожки.
  - У каждой зоны различное количество дорожек/секторов

© 2014 МГУ/ВМК/СП

Вычисление емкости диска

Емкость = (#байт/сектор) x (среднее # сектор/дорожка) x (# дорожка/поверхность) x (# поверхность/пластина) x (# пластина/диск)

Пример:

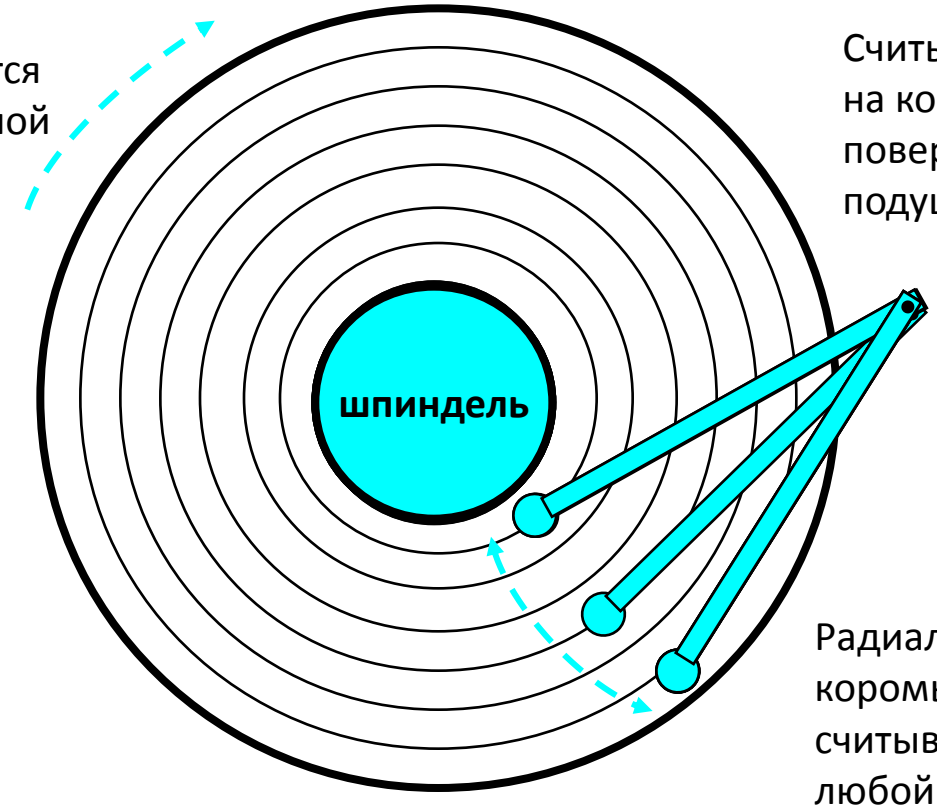
- 512 байт/сектор
- 300 сектор/дорожка (в среднем)
- 20,000 дорожка/поверхность
- 2 поверхность/пластина
- 5 пластина/диск

Емкость = 512 x 300 x 20000 x 2 x 5 = 30,720,000,000 = 30.72ГБ

© 2014 МГУ/ВМК/СП

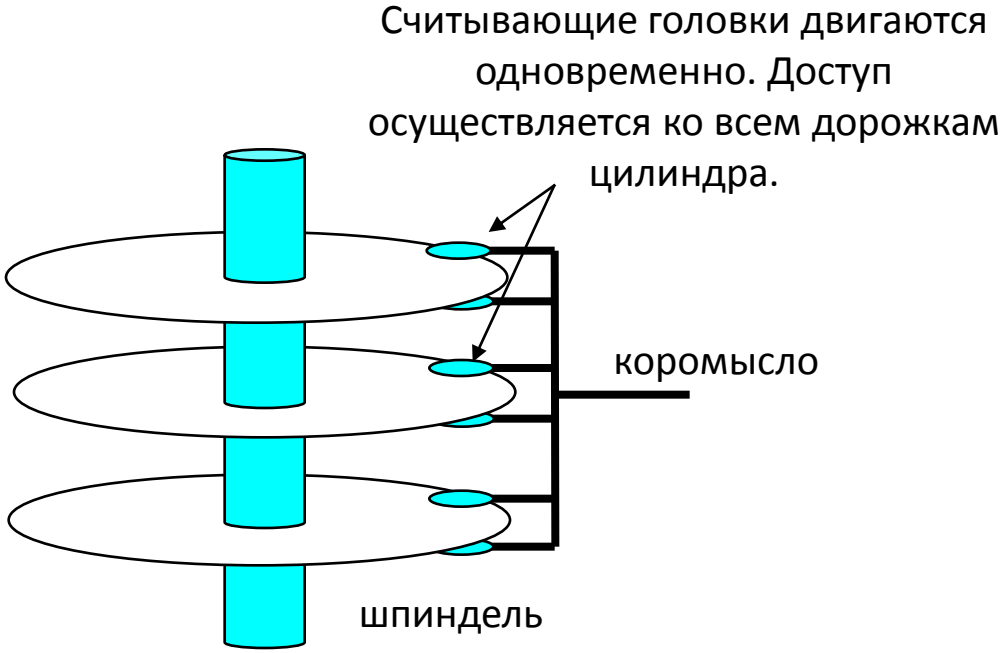
Работа с диском (одна пластина)

Поверхность диска вращается с фиксированной скоростью.



Считывающая головка закреплена на конце коромысла и парит над поверхностью на тонкой воздушной подушке.

Радиально перемещаясь, коромысло может выставить считывающую головку над любой дорожкой.

© 2014 МГУ/ВМК/СП

Работа с диском (несколько пластин)