

Компютърни архитектури CSCB008

Организация на паметта. Йерархия на паметта

доц. д-р Ясен Горбунов
2021

Какво представлява компютърната памет?

- основна функция – съхраняване на данни и програми
- масив от запомнящи клетки, достъпни чрез уникален адрес

Основни функционални характеристики:

- **бързодействие** – определено от
 - времето за достъп (ns)
 - скоростта на трансфер (MB)
 - **латентност** – генериране на известен брой “паразитни” тактови импулси, за да може изискваната информация да бъде доставена до изходните шини на магистралата за данни (15-20 пъти по-висока от тази на процесорната памет)
- **капацитет** – брой единици информация, които могат да се съхраняват

SI →	1000 kB	kilobyte	1024 KiB	kibibyte	← IEC
	1000 ² MB	megabyte	1024 ² MiB	mebibyte	
	1000 ³ GB	gigabyte	1024 ³ GiB	gibibyte	
	1000 ⁴ TB	terabyte	1024 ⁴ TiB	tebibyte	
	1000 ⁵ PB	petabyte	1024 ⁵ PiB	pebibyte	

Ние ще използваме системата, в която 1kB = 1024 bytes, 1MB = 1024kB и т.н.

Какво представлява компютърната памет?

според мястото

- главна памет

- статична памет (тригери – малък обем, висока цена)
- динамична памет (кондензатори – по-трудно управление, голям обем, ниска цена)

- кеш памет (скрита памет)

- междинно ниво м/у процесора и основната памет
- малък капацитет и бързодействие, съизмеримо с това на процесора
- прозрачна за потребителя

- външна памет

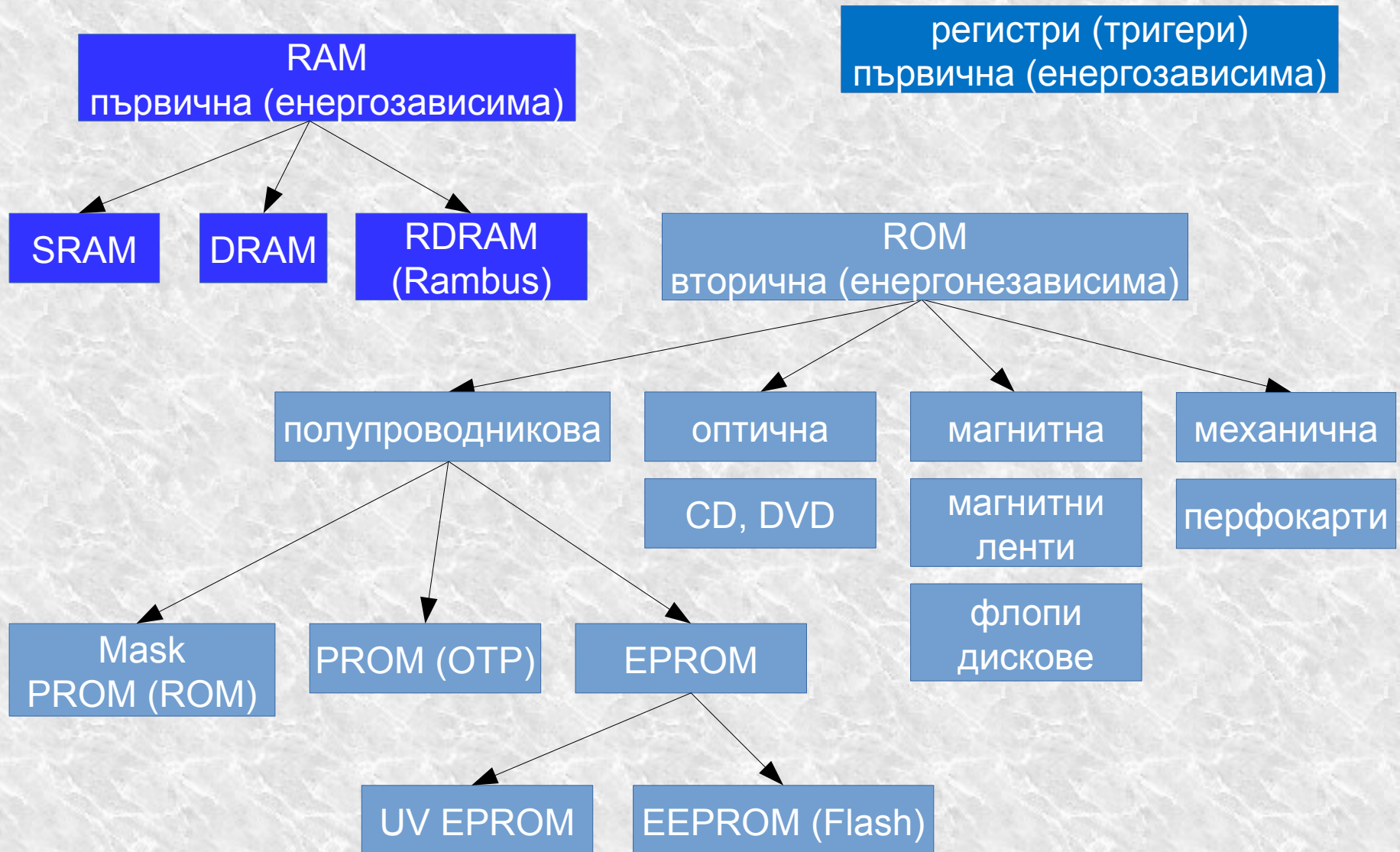
- HDD, SSD, CD-ROM, Floppy и др.

според метода за достъп

- с последователен достъп – модел на магнитната лента
- с директен достъп – модел на твърдия диск
- с произволен достъп – достъп по адрес (RAM)
- с асоциативен достъп – достъп по признак (съдържание)

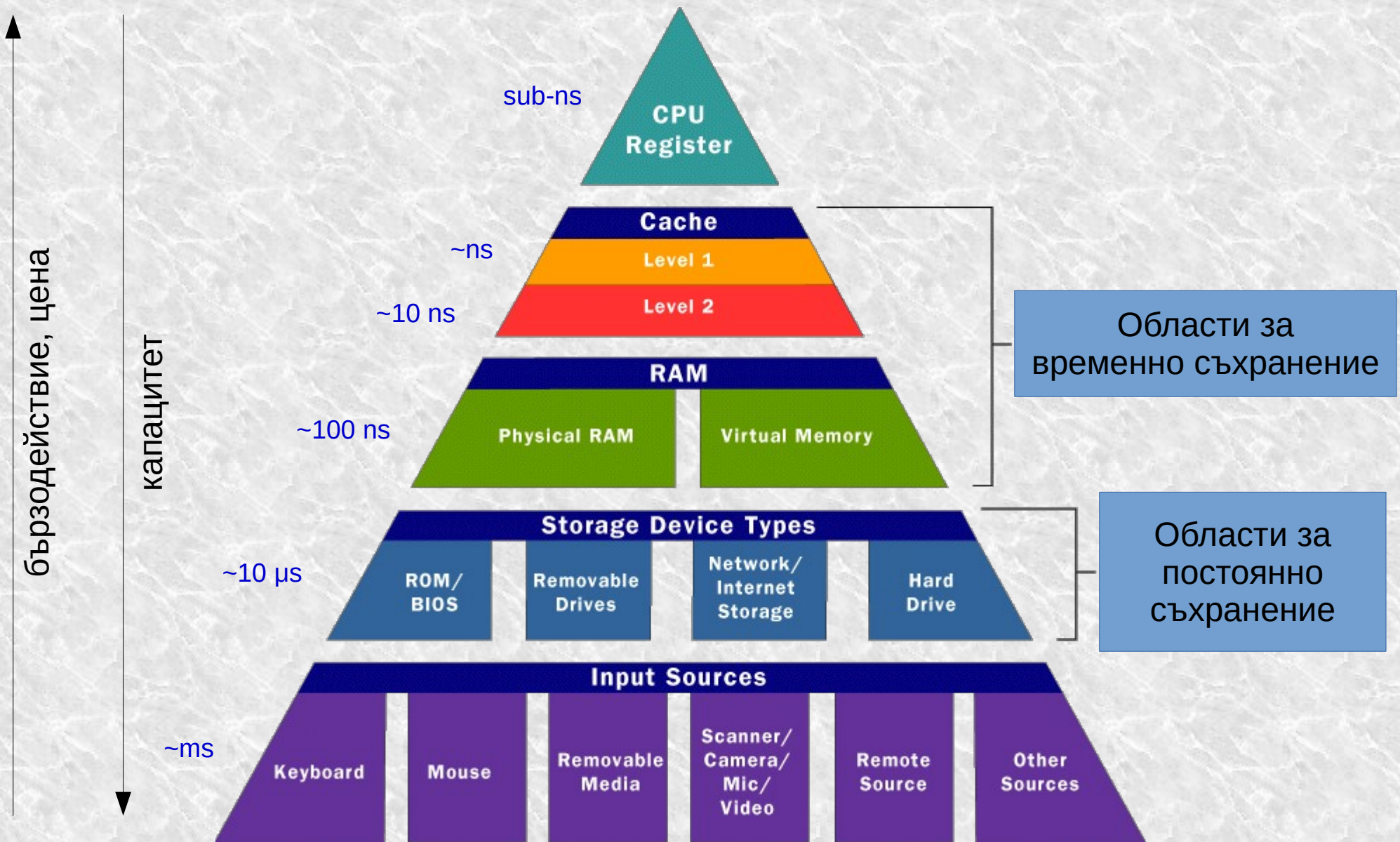
Йерархия на паметта

Видове памет – обща класификация



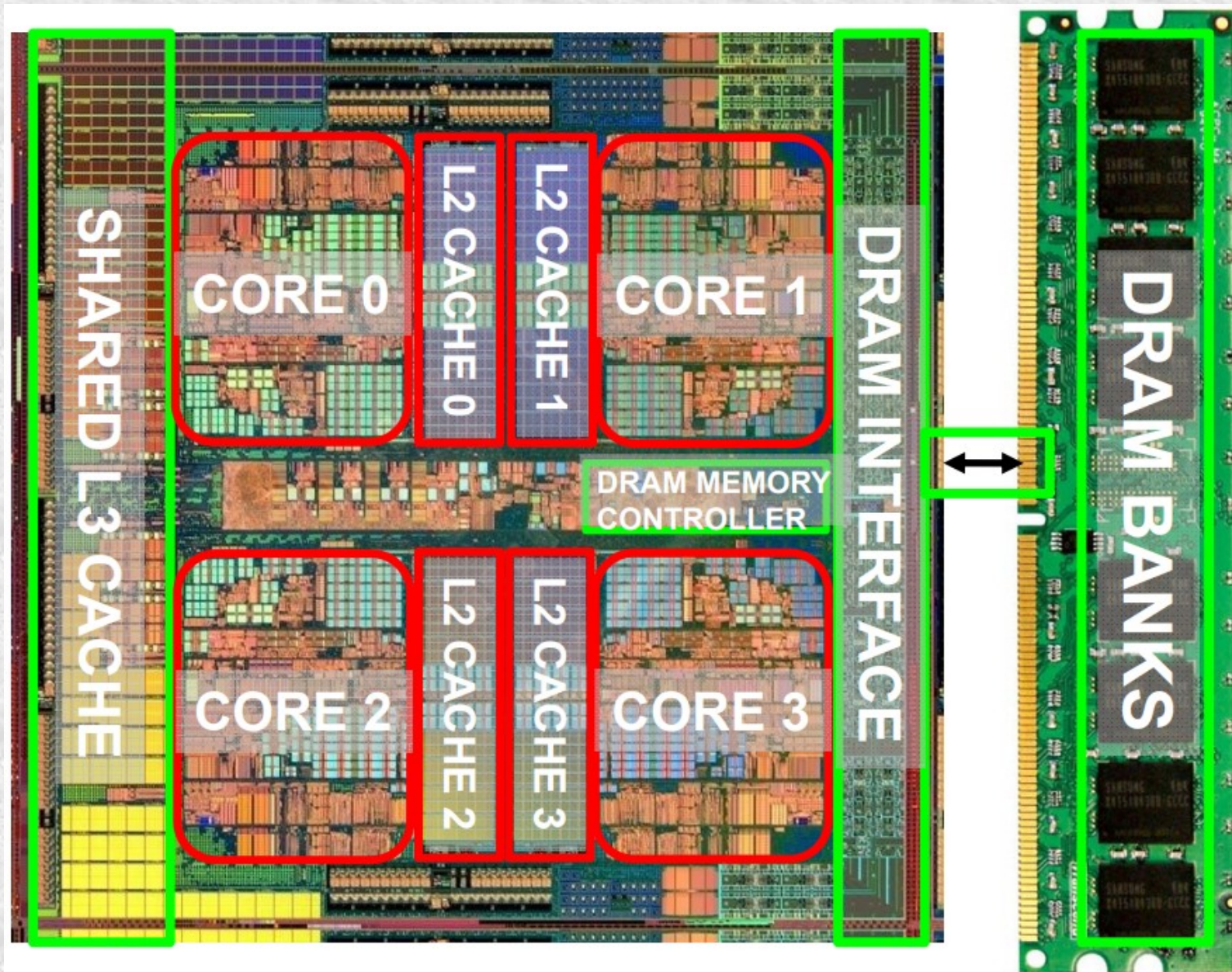
Йерархия на паметта

Видове памет – обща класификация

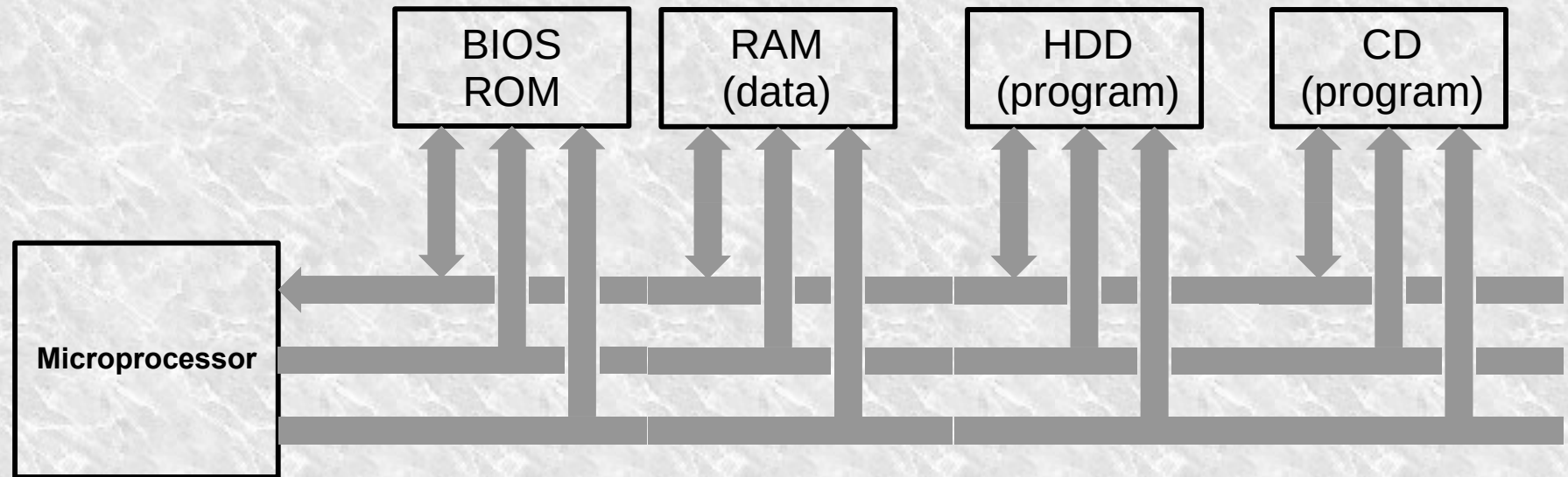


Йерархия на паметта

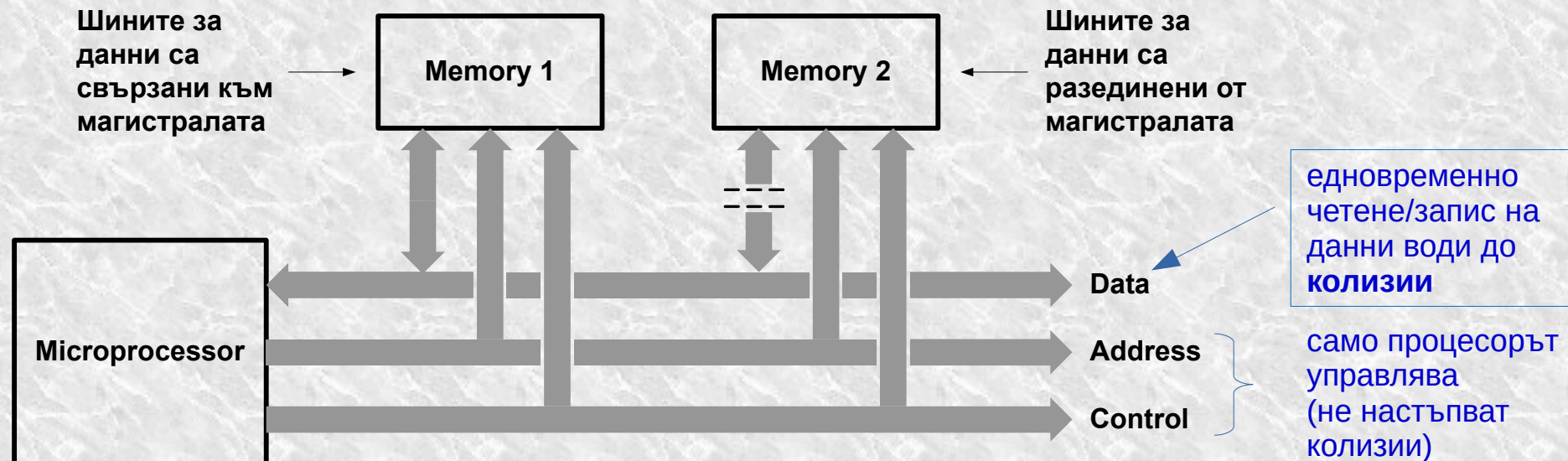
Място на главната памет в компютърната архитектура



Свързване на паметта – шинна организация



Свързване на паметта – шинна организация



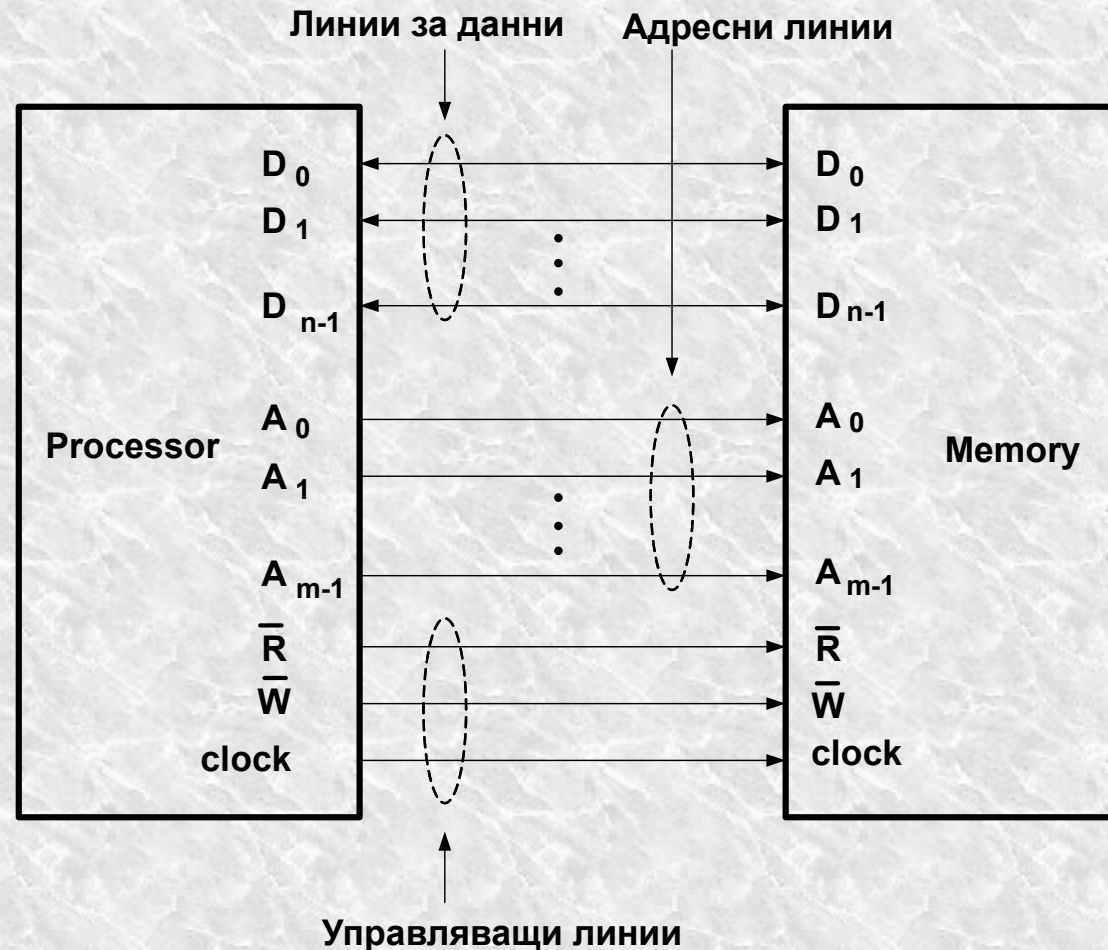
избор на активно устройство

	nCS ₀	nCS ₁	nCS ₂	nCS ₃
memory 1	0	1	1	1
memory 2	1	0	1	1
memory 3	1	1	0	1
memory 4	1	1	1	0
всички устройства са разединени	1	1	1	1

one-cold
кодиране

Свързване на паметта – шинна организация

- данни → индивидуален проводник за всеки бит
- адреси → двоично число без знак
- управление → генерирано от процесора

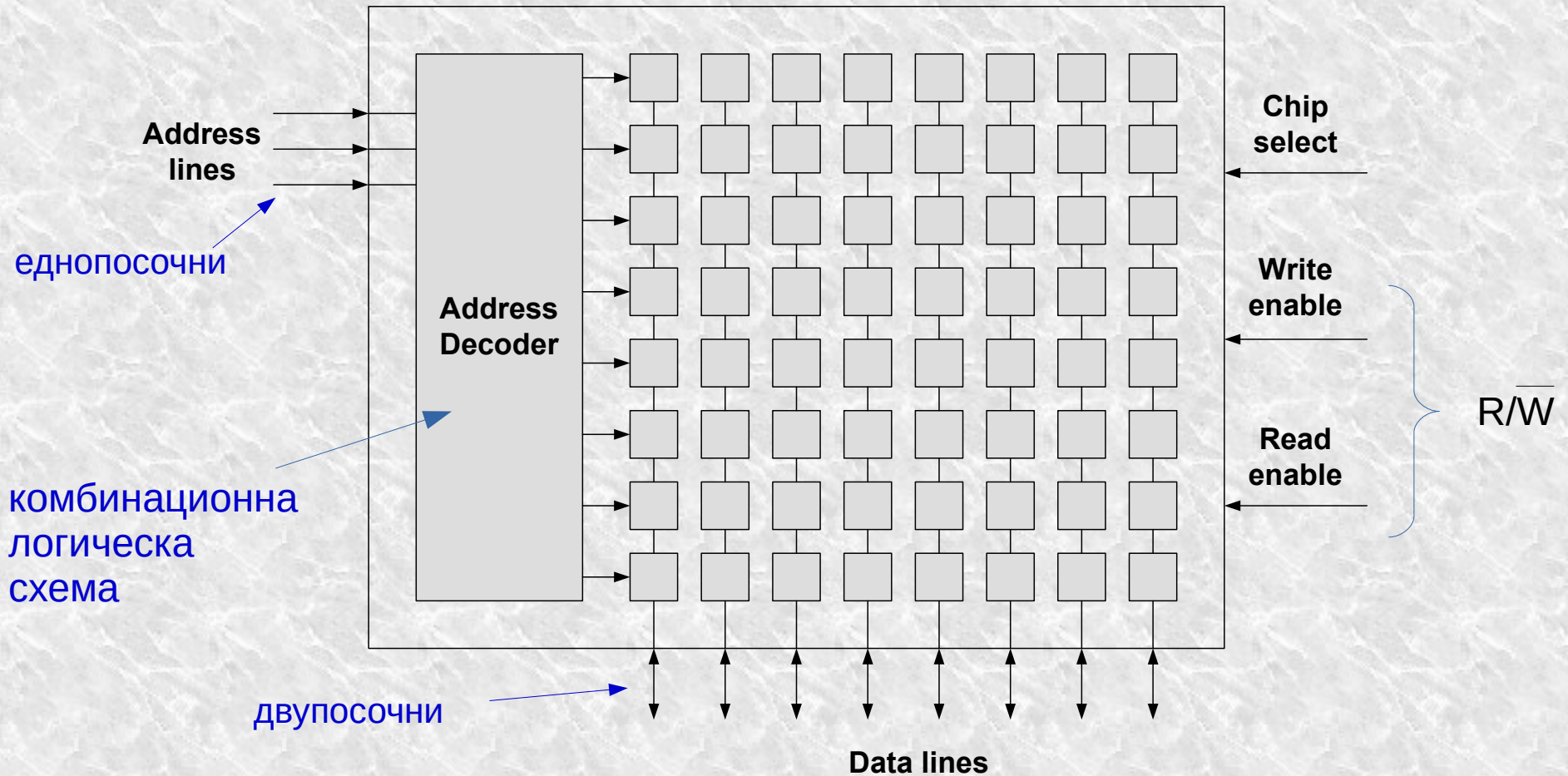


Основна организация на схема памет

матрична организация

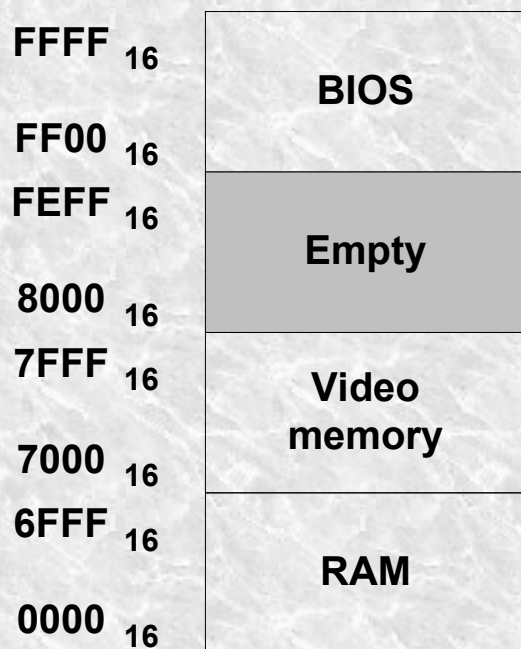
редове → адреси

колони → данни (брой битове)

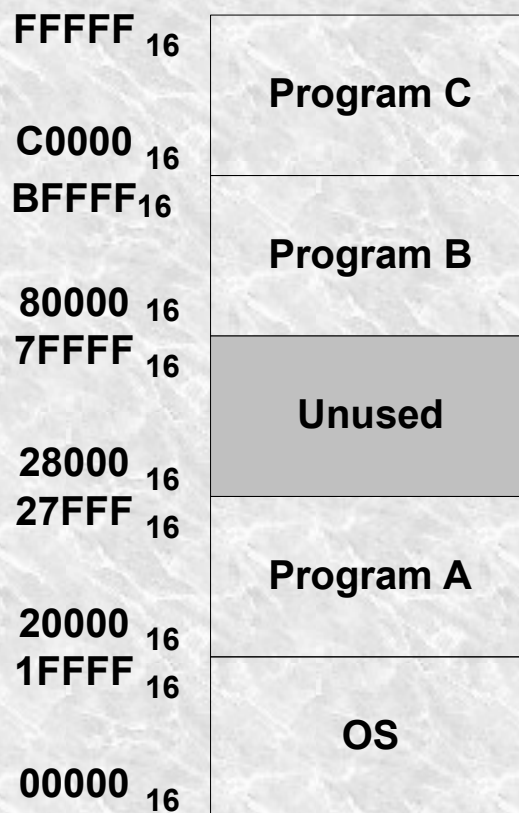


Карти на разпределението на паметта (Memory Maps)

memory map – дълга вертикална колона, в която редовете отговарят на отделните области памет, често разположени в отделни чипове – **пълен адресен обхват на процесора**



а) Хардуерно представяне



б) Софтуерно представяне

Колко адресни линии има в това адресно пространство?

$$2^{20} = \text{FFFFFF}_{\text{HEX}}$$

20 линии

Колко е адресното пространство (в MB)?

$$\text{FFFFFF}_{\text{HEX}} = 1048575_{\text{DEC}}$$

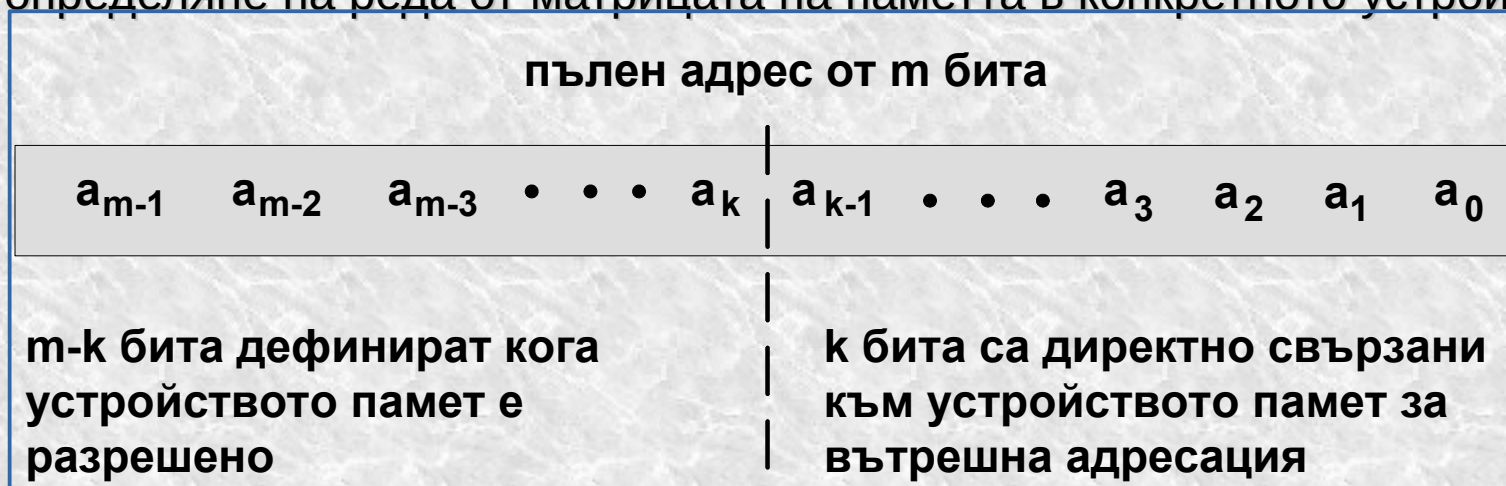
2^{20} бита = 1 MB

Декодиране на адреси (Address Decoding)

метод на използване на адресите за разрешаване само на едно устройство памет

пълен адрес – две групи

- идентифициране на конкретно устройство на паметта
- определяне на реда от матрицата на паметта в конкретното устройство



Големина на паметта	Брой необходими адресни линии		Големина на паметта	Брой необходими адресни линии
1 KB	10		256MB	28
256 KB	18		1GB	30
1 MB	20		4GB	32
16 MB	24		64GB	36

Пример:

Процесор с 256 MB адресно пространство използва адрес $35E3C03_{16}$ за достъп до 16 MB устройство памет.

общ обем памет 256 MB → 28 адресни линии
достъп до 16 MB → 24 адресни линии

$$35E3C03_{16} = \mathbf{0011} \ 0101 \ 1110 \ 0011 \ 1100 \ 0000 \ 0011_2$$

разрешаване на схемата → битове $a_{27}=0$, $a_{26}=0$, $a_{25}=1$, $a_{24}=1$

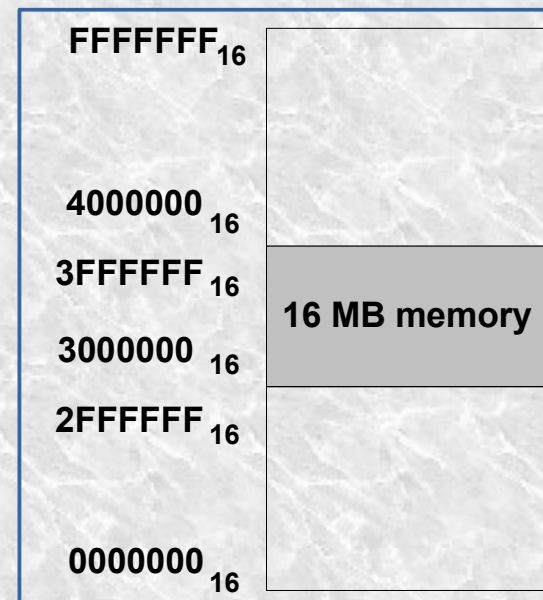
от гледна точка на картата на паметта

старши адрес

$$0011 \ \mathbf{1111 \ 1111 \ 1111 \ 1111 \ 1111 \ 1111}_2 = \mathbf{3FFFFFF}_{16}$$

младши адрес

$$0011 \ \mathbf{0000 \ 0000 \ 0000 \ 0000 \ 0000 \ 0000}_2 = \mathbf{3000000}_{16}$$



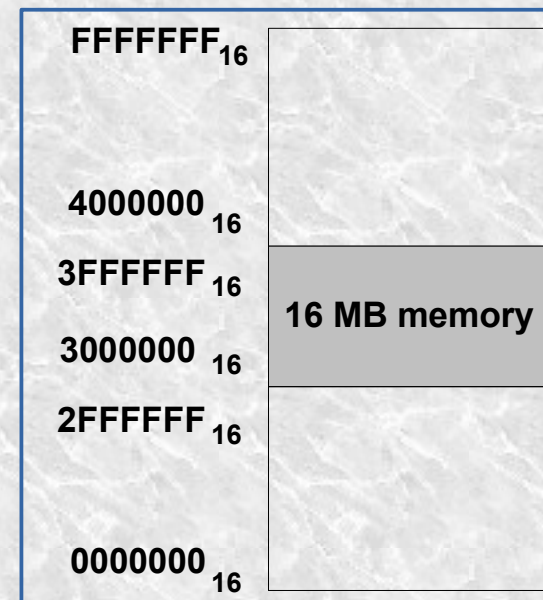
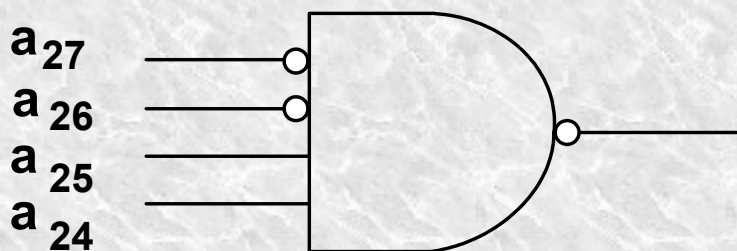
Пример:

Процесор с 256 MB адресно пространство използва адрес $35E3C03_{16}$ за достъп до 16 MB устройство памет.

$$35E3C03_{16} = \mathbf{0011} \ 0101 \ 1110 \ 0011 \ 1100 \ 0000 \ 0011_2$$

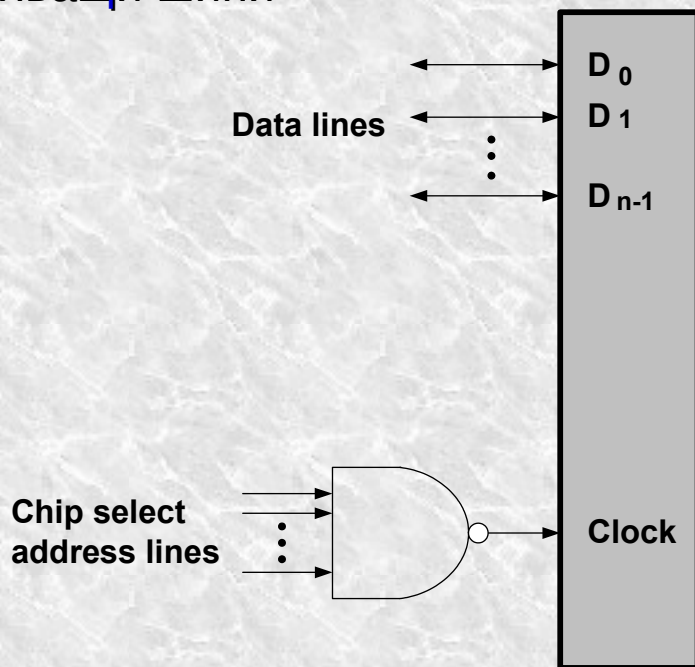
разрешаване на схемата → битове $a_{27}=0$, $a_{26}=0$, $a_{25}=1$, $a_{24}=1$

за синтез на хардуера, определящ Chip Select (nCS , \overline{CS})

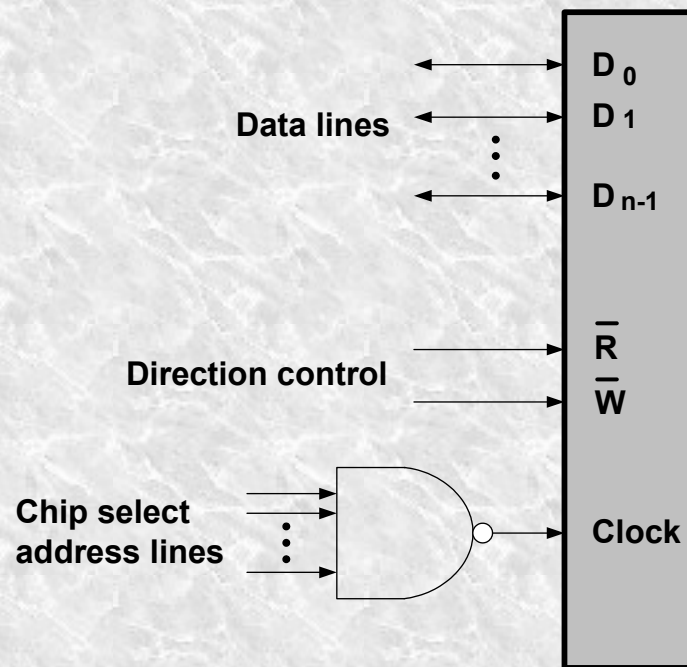


Входно-изходни устройства, свързани като памет (Memory Mapped Input/Output)

някои устройства нямат матрица памет, но интерфейстът към CPU използва
шини за данни
управляващи шини



а) Изходно устройство
свързано като памет



б) Входно-изходно устройство
свързано като памет

Памет с произволен достъп (RAM – Random Access Memory)

- **достъп до всеки адрес в произволен момент от време**
(за разлика от паметта с последователен достъп – лентова памет)
- **съдържанието се губи при отпадане на захранването (volatile)**
- **висока скорост на четене/запис**
- **оперативна памет**

статична RAM
(SRAM – тригери)

динамична RAM
(DRAM – кондензатори)

синхронни памети – тактувани

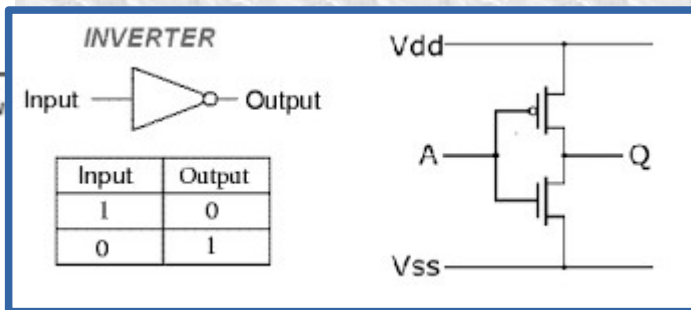
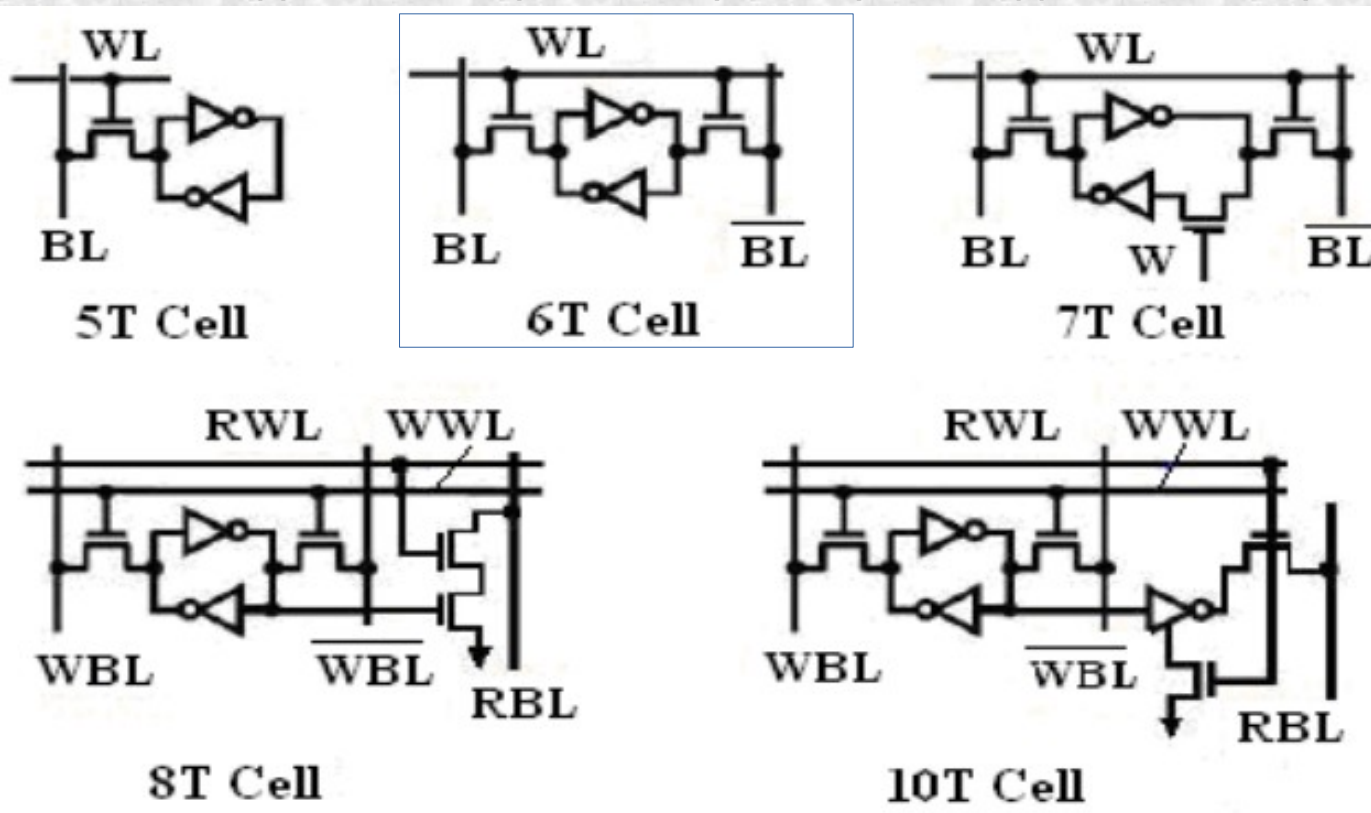
асинхронни памети – при постъпване на данни на входовете на паметта, CPU трябва да изчака паметта да отговори

Памет с произволен достъп – статична RAM памет – SRAM

- Помни данните в транзисторни вериги подобни на D тригери

Т – брой транзистори

CMOS инвертор



Operation	BL	BL bar
Write ("1")	1	0
Write ("0")	0	1
Read	0	0

- Използва се за много бързи приложения, като например RAM кеш (cache)
- Запомня малко данни при много бърз достъп и има проста декодираща логика
- Данните се помнят до прекъсване на захранването (volatile)

Памет с произволен достъп – статична RAM памет – SRAM

Колко „тригера“ има в адресно пространство 256 KB?

$$256 \cdot 1024 \cdot 8 = 2^8 \cdot 2^{10} \cdot 2^3 = 2^{21} = 2\,097\,152$$

Колко транзистора означава това?

при клетка с 6 транзистора

$$6T - 2\,097\,152 \cdot 6 = 12\,582\,912$$

при клетка с 10 транзистора

$$10T - 2\,097\,152 \cdot 10 = 20\,971\,520$$