

# КЭШ-память

- 1 Многоуровневая организация памяти (виды и характеристики памяти).
- 2 КЭШ-память: основные определения и характеристика.
- 3 Основные виды КЭШ-памяти.
- 4 Эффективность применения.

**Знать:** место и роль кэш-памяти в иерархии памяти ЭВМ, способы отображения оперативной памяти на кэш-память и соответствующие им виды организации кэш-памяти (с ассоциативным, прямым, частично ассоциативным отображением, с отображением секторов), стратегии записи и замещения строк в кэш-памяти, преимущества и недостатки различных видов кэш-памяти.

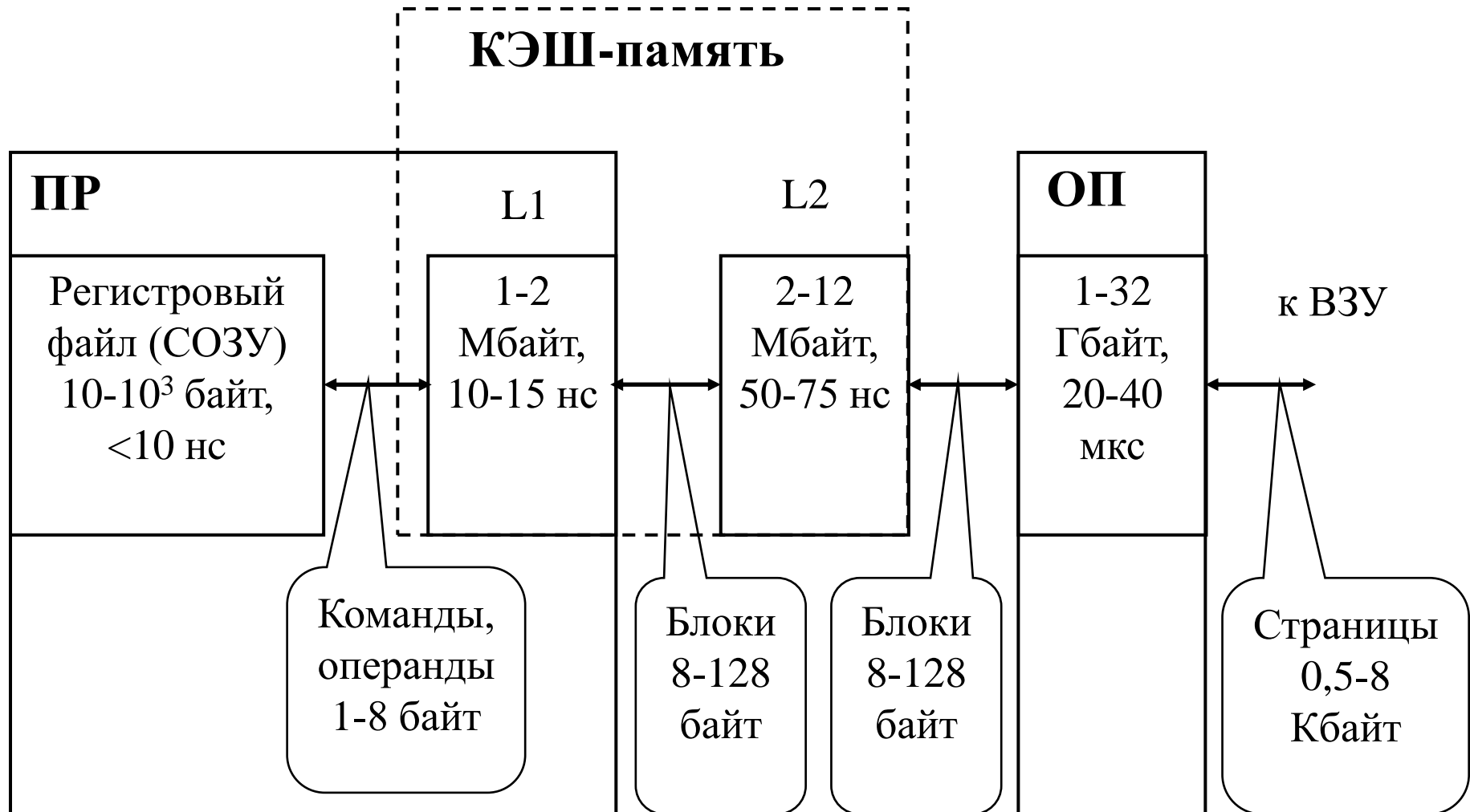
**Уметь:** разработать структурную схему и микропрограмму работы кэш-памяти заданного вида.

**Помнить:**

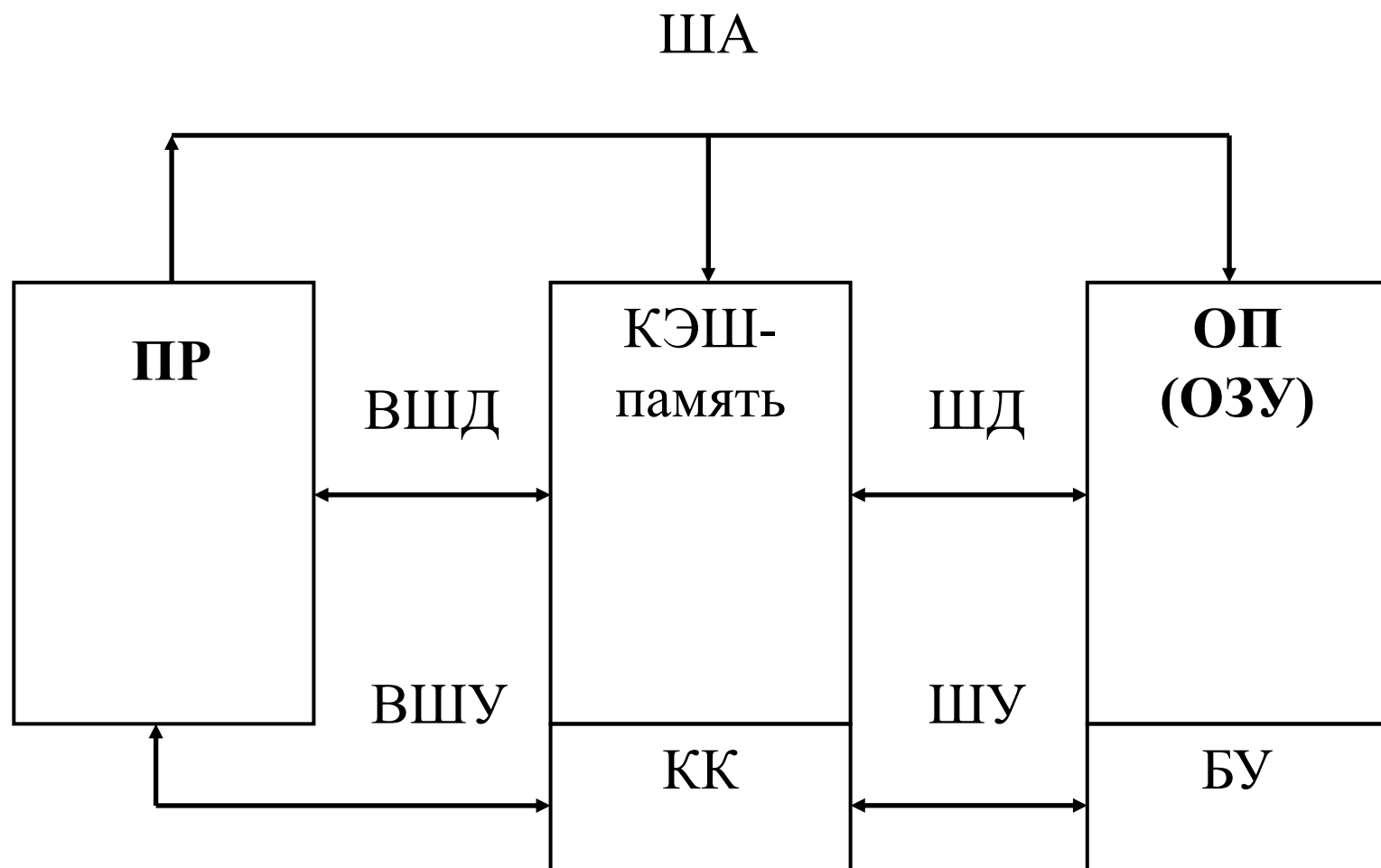
**Литература:**

- Цилькер Б.Я., Орлов С.А. Организация ЭВМ и систем. Учебник для вузов. – СПб.: Питер, 2004. – 668 с. (с. 249-263).

# 1 Многоуровневая организация памяти



# *Связь КЭШ-памяти с процессором и ОЗУ*



## ***2 КЭШ-память: основные понятия и характеристика***

- *Основные понятия:* принцип временной и пространственной локальности обращений к основной памяти (ОП); строка КЭШ-памяти; блок данных ОП; попадание; промах; тег; отображение блока ОП на кэш-память.
- *Виды кэш-памяти в зависимости от способа отображения.*
  - Ассоциативная.
  - С прямым отображением.
  - С частично-ассоциативным (множественно-ассоциативным) отображением.
  - С отображением секторов.

# Стратегии записи (прямая запись)

- Прямая (сквозная) запись:
  - с распределением (отображением);
  - без распределения (отображения).
- Введем обозначения:  $s[a]$  – слово с адресом  $a$ ;  $ST[A]$  – строка, содержащая слово  $s[a]$ .

	ОП	КЭШ-память
$s[a]$ есть в кэш-памяти	Запись	Запись
$s[a]$ нет в кэш-памяти		
1) Сквозная с распределением	Запись	Пересылка блока, содержащего $s[a]$ из ОП в кэш-память
2) Сквозная без распределения	Запись	---

- Используется также сквозная буферизированная запись. Информация в начале записывается в кэш-память и специальный буфер, работающий по схеме FIFO.

## *Стратегии записи (обратная запись)*

- При обратной записи слово заносится только в кэш-память. Обратная запись подразделяется на *простую* и *флаговую* (при изменении строки КЭШ-памяти устанавливается связанный с ней бит – флажок).
- При флаговой обратной записи в процессе замещения строка из кэш-памяти переписывается только тогда, когда ее флажок установлен в «1».
- В среднем обратная запись на 10% эффективнее сквозной, но операции записи составляют от 5 до 34%, поэтому ее преимущество невелико.

## *Простая обратная запись*

	ОП	КЭШ-память
$s[a]$ есть в кэш-памяти	Нет записи	Запись
$s[a]$ нет в кэш-памяти	Пересылка блока, содержащего $s[a]$ из ОП в кэш-память	Запись
Чтение $s[a]$ из ОП при обратной записи	Пересылка блока, содержащего $s[a]$ из ОП в кэш-память	Чтение строки $ST[A]$

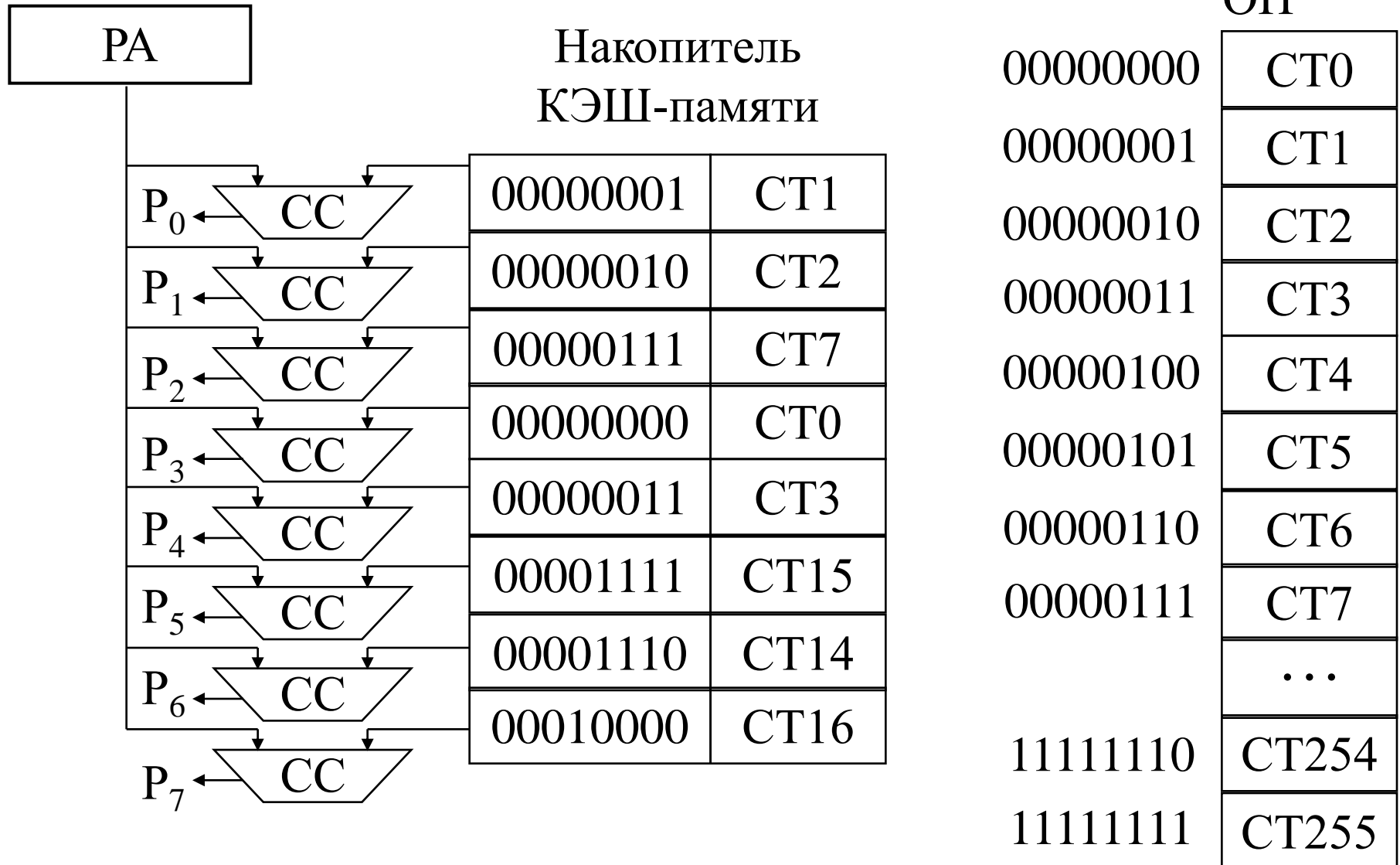


# *Алгоритмы замещения информации в заполненной кэш-памяти*

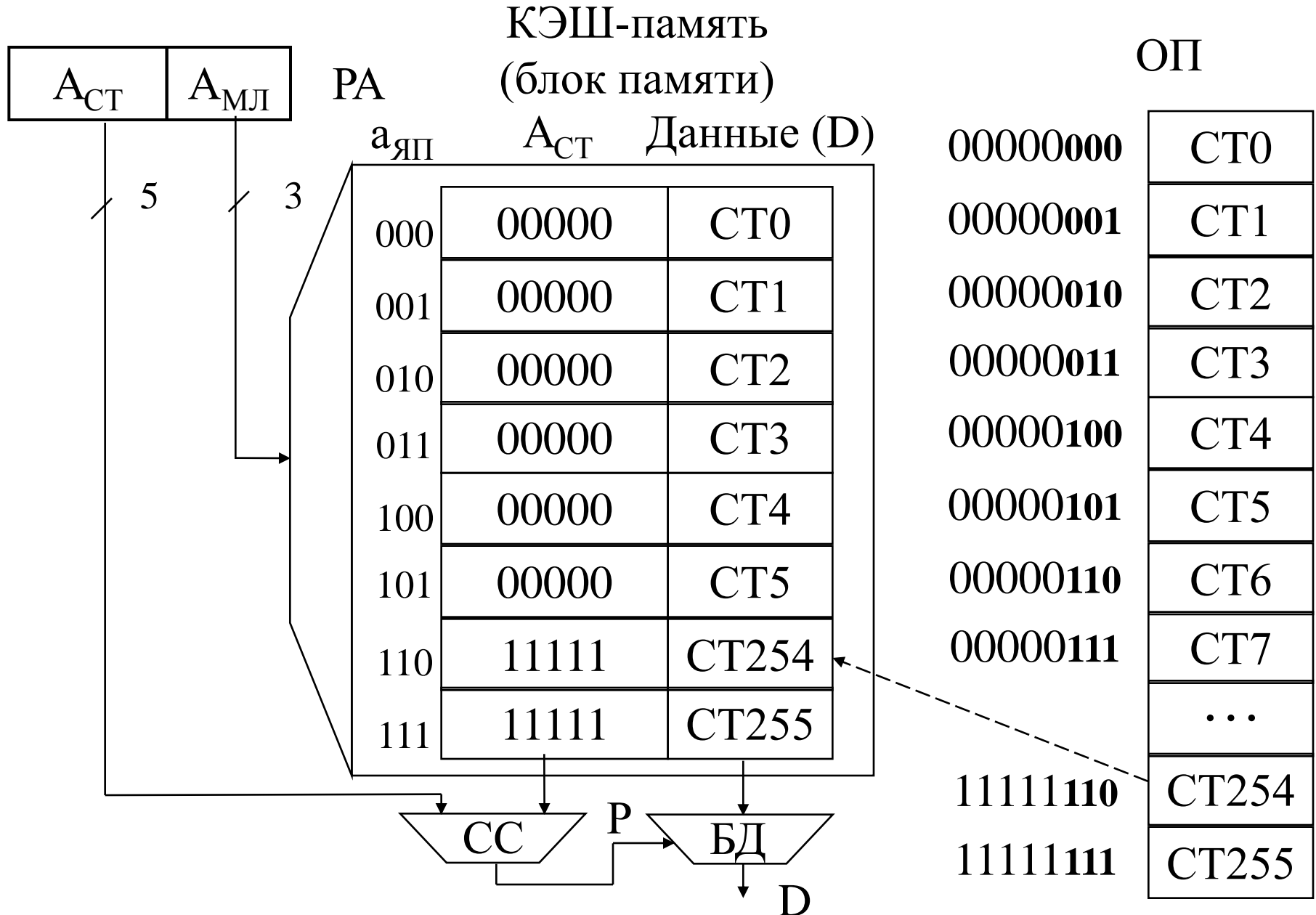
- При прямом отображении выбор замещаемой строки не требуется, она определяется однозначно.
- Алгоритм замещения на основе наиболее давнего использования (Least Recently Used – LRU).
  - Исходные строки:                      a              b              c              d
  - Использована строка c:              c              a              b              d
  - Использована строка d:              d              c              a              **b**
  - Записывается строка e:              e              d              c              a
- Алгоритм, работающий по принципу FIFO.
- Алгоритм замены наименее часто использованной строки (Least Frequently Used – LFU).
- Простейший алгоритм случайного выбора замещаемой строки.

# 3 Основные виды КЭШ-памяти

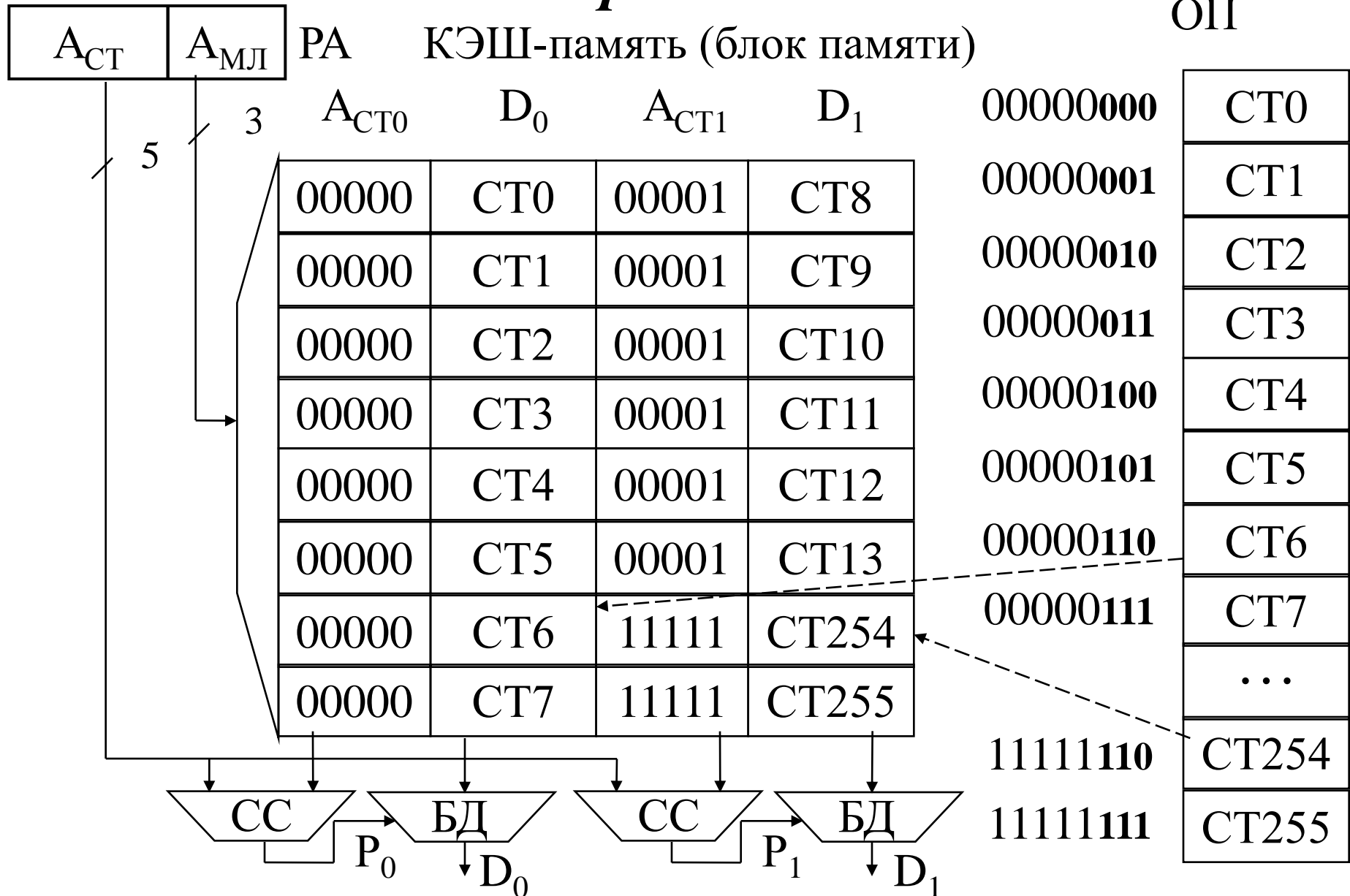
## Ассоциативная КЭШ-память



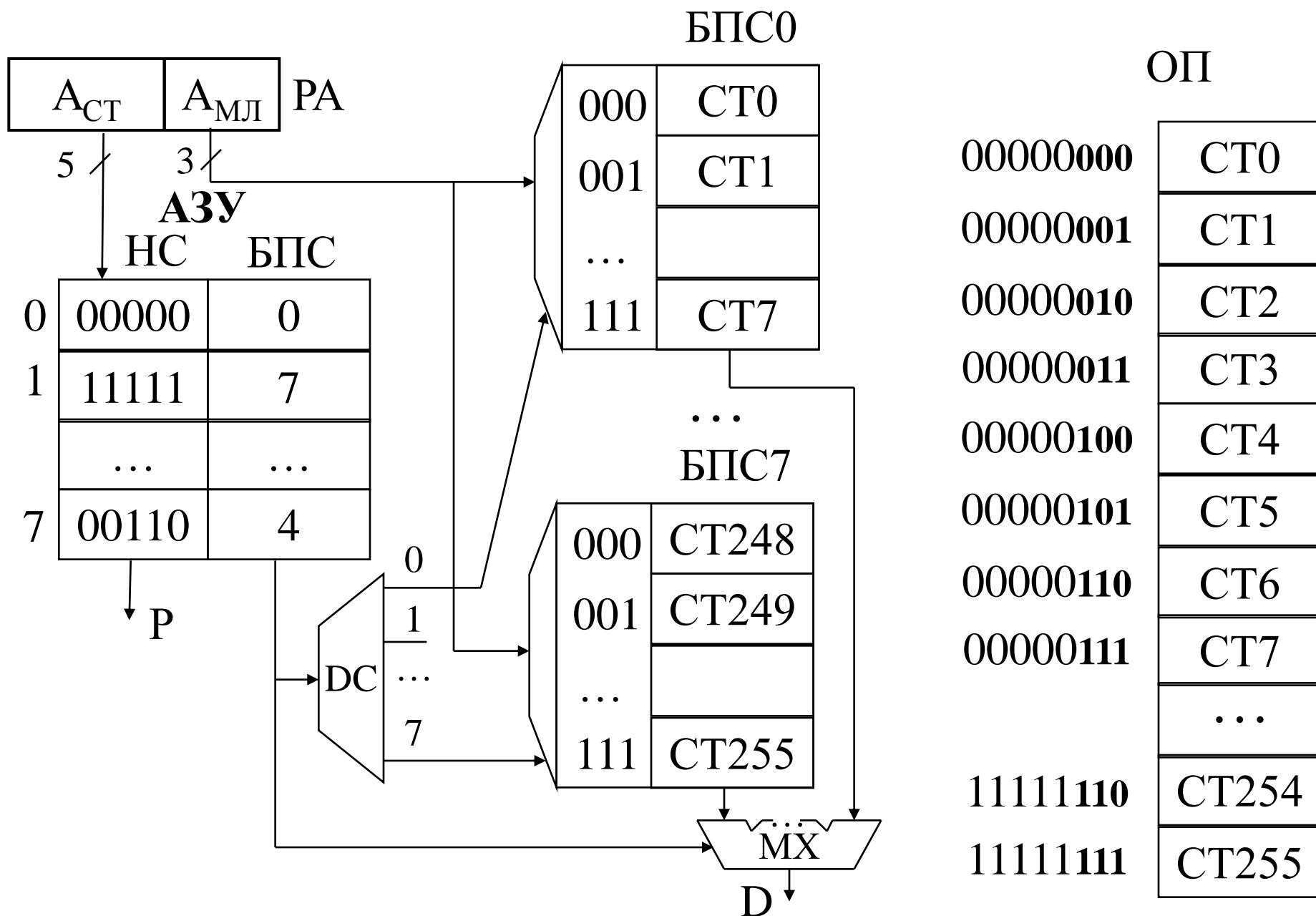
# КЭШ-память с прямым отображением



# ***КЭШ-память с частично ассоциативным отображением***



# КЭШ-память с отображением секторов



## *Обращение к секторной КЭШ-памяти*

- Основная память делится на сектора. Данные внутри сектора отображаются с помощью прямого, а номер сектора – с помощью ассоциативного отображения.
- При обращении можно выделить три случая.
  - Попадание – номер запрашиваемого сектора есть в АЗУ КЭШ-памяти и бит присутствия строки в выбранном блоке строк равен единице. Обращение считается успешным.
  - Запрашиваемого номера сектора в АЗУ нет. В этом случае адрес запрашиваемого сектора записывается в АЗУ и в БПС пересылается запрашиваемая строка. У данной строки бит присутствия устанавливается в единицу, у остальных строк этого БПС биты присутствия устанавливаются в ноль.
  - Номер запрашиваемого сектора в АЗУ есть, но бит присутствия требуемой строки равен нулю. Строка вводится и бит ее присутствия устанавливается в единицу.

# 4 Эффективность применения

## Возможности по отображению оперативной памяти на КЭШ-память различного вида

- Введем следующие обозначения:  $N$  – разрядность адреса ОП, а  $n$  – КЭШ-памяти ( $N \gg n$ ).
- Число  $K$  различных способов отображения ОП на КЭШ-память, влияющее на вероятность промаха приведено в таблице ( $N=20$ ,  $n=12$ )

Кэш-память	Формула	Пример оценки
Ассоциативная	$K=(2^N)^{2n}$	$K=3,12 \times 10^{144}$
С прямым отображением	$K=(2^{N-n})^{2n}$	$K=6,28 \times 10^{57}$
С множественно-ассоциативным отображением	$K=(2^{N-m})^{2n}$ ( $m=\log M$ , где $M$ – число строк в множестве)	$K=1,77 \times 10^{72}$ ( $m=10$ )
Секторная	$K=(2^{N-n+s})^{2s}$ ( $s$ – число секторов)	$K=1,16 \times 10^{77}$ ( $s=8$ )

# ***Время считывания данных***

<b>Считывание из ОП</b>		<b>Считывание из КЭШ-памяти (попадание)</b>		<b>Считывание из КЭШ-памяти (промах)</b>	
Действие	Число тактов	Действие	Число тактов	Действие	Число тактов
Выдача адреса	1	Выдача адреса	1	Выдача адреса	1
Ожидание RAM	2	Проверка КЭШ		Проверка КЭШ (нет)	1
Прием данных	1	Прием данных	1	Выдача адреса (ОП)	1
ВСЕГО	4	ВСЕГО	2	Ожидание RAM	2
				Прием данных	1
				ВСЕГО	6



## *Пример оценки эффективности применения КЭШ-памяти с прямым отображением*

Размер КЭШ-памяти	Попадание (%)	Выигрыш (%)
Нет КЭШ-памяти, динамическая память (2 такта ожидания)	—	—
16 Кб	81	35
32 Кб	86	38
64 Кб	88	39
128 Кб	89	39
Нет КЭШ-памяти, статическая память вместо динамической	100	47

- Емкость КЭШ-памяти составляет порядка 1/100-1/500 емкости основной памяти, а быстродействие в 5-10 раз выше.