

ARSITEKTUR DAN ORGANISASI KOMPUTER

MEMORY CACHE

Karakteristik System Memori

1. Location
2. Capacity
3. Unit of transfer
4. Access method
5. Performance
6. Physical type
7. Physical characteristics
8. Organisation

Location Memory

- Register
: memori yang berada dalam CPU, sebagai memori sementara dalam perhitungan maupun pengolahan data dalam prosesor
- Internal
: berada diluar chip CPU pengaksesan langsung ke CPU, memori utama dan cache memori
- External
: perangkat penyimpan tambahan seperti disk, pita yang dapat diakses oleh CPU melalui I/O

Capacity/Kapasitas Memori

- Memori Internal
 - kapasitas dinyatakan dalam **byte** (1 byte = 8 bit) atau **word**.
 - panjang satu word 8, 16, 32 bit.
- Memori External
 - kapasitas dinyatakan dalam byte.
 - biasanya lebih besar kapasitasnya daripada memori internal, hal ini disebabkan karena teknologi dan sifat penggunaannya yang berbeda.

Unit Of Transfer/Satuan Transfer

⊙ Internal

- Unit transfer sama dengan banyaknya saluran data ke dalam dan keluar modul memori.

⊙ External

- Jumlah blok transfer selalu lebih besar dari word
- 64, 128, 256 bit

⊙ Addressable unit

- Dalam beberapa sistem sering disebut word.
- Sistem lain memperbolehkan pengalamatan pada tingkatan byte.

Access Methods

a. Sequential

- diorganisir dalam unit-unit data/RECORD
- akses harus dibuat dalam urutan linier spesifik
- pembacaan akan dilakukan secara urut
- contoh : unit pita

b. Direct

- Setiap blok dan record memiliki alamat unik berdasarkan lokasi fisiknya
- waktu akses variabel
- akses dilakukan secara langsung pada alamat memori
- contoh : unit disk

Lanjutan:

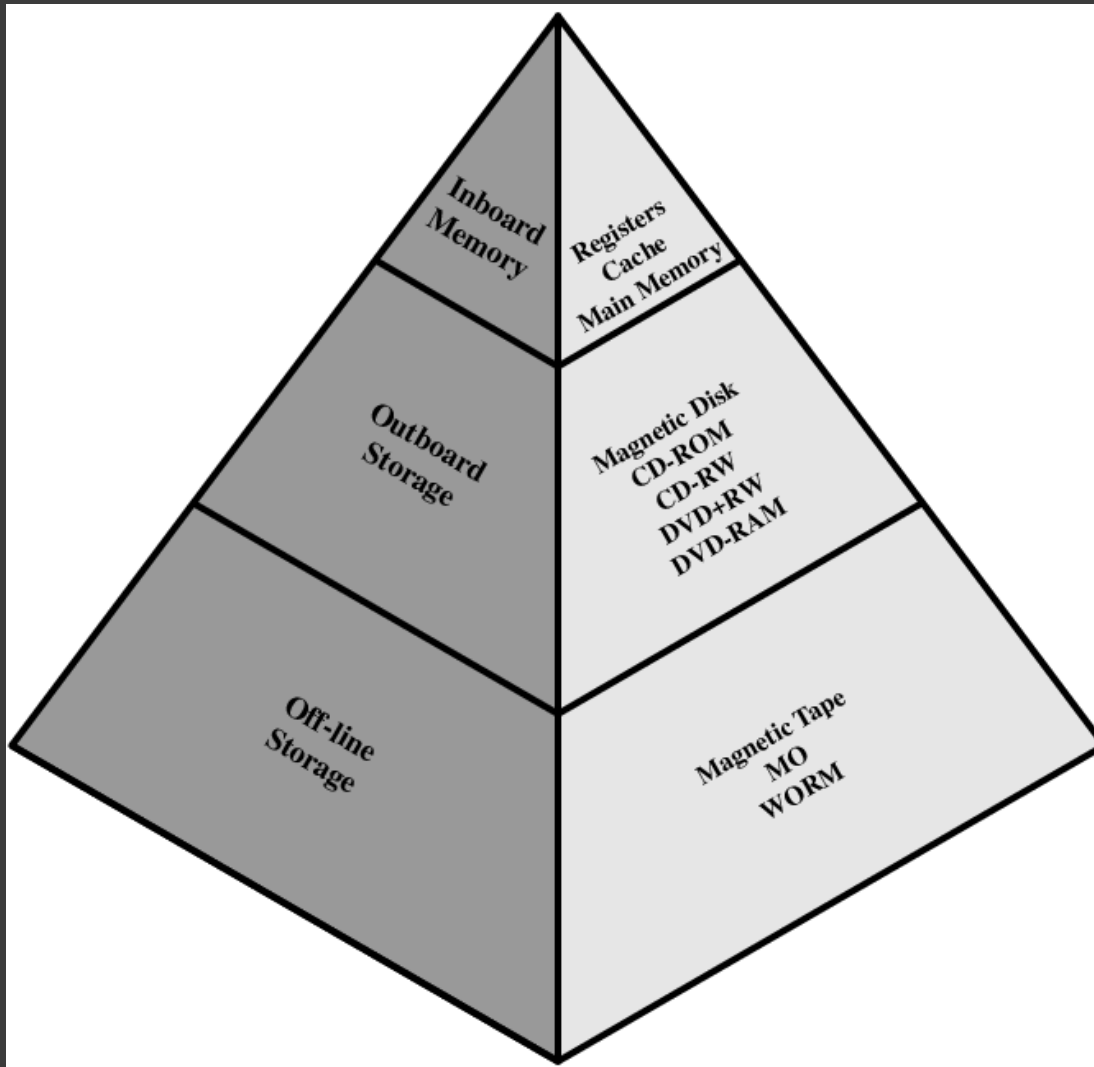
c. Random / acak

- pengalamatan dengan sistem fisik wired-in / pengalamatan langsung.
- waktu akses memori bersifat konstan.
- tidak tergantung pada urutan akses.
- contoh : sistem memori utama dan cache.

d. Associative

- merupakan jenis random akses
- memungkinkan perbandingan lokasi bit yang diinginkan untuk pencocokan
- data dicari berdasarkan isinya bukan alamatnya dalam memori
- contoh : cache memori

Memory Hierarchy - Diagram



Kinerja (Performance)

- *Access time*

- pada memori random waktu access adalah waktu yang dibutuhkan untuk melakukan operasi baca atau tulis.
- pada memori non random waktu access merupakan waktu yang dibutuhkan untuk melakukan operasi baca atau tulis pada lokasi tertentu.

- *Memory cycle time*

: konsep ini digunakan pada random access memory dan terdiri dari access time ditambah dengan waktu yang diperlukan transient agar hilang pada saluran sinyal.

- *Transfer rate*

: kecepatan data transfer ke unit memori atau dari unit memori.

Tipe Fisik

- Memori Semikonduktor teknologi VLSI
contoh : RAM
- Memori magnetik
contoh : disk dan pita.
- Optik
contoh : CD dan DVD
- Magneto-optik.
contoh : buble, hologram

Karakteristik Fisik

- ⊙ *volatile memory*,

 - : informasi akan hilang apabila daya listriknya dimatikan

- ⊙ *non-volatile*

 - : memory tidak hilang walau daya listriknya hilang.

- Memori permukaan magnetik → non volatile

- Semikonduktor → volatile dan non-volatile

Keandalan Memori

⦿ *Berapa banyak ?*

: sulit dijawab → karena berapapun kapasitas memori tentu aplikasi akan menggunakannya.

⦿ *Berapa cepat?*

: memori harus mampu mengikuti kecepatan CPU sehingga terjadi sinkronisasi kerja antar CPU dan memori

⦿ *Berapa mahal?*

: sangatlah relatif

Hubungan harga, kapasitas dan waktu akses adalah :

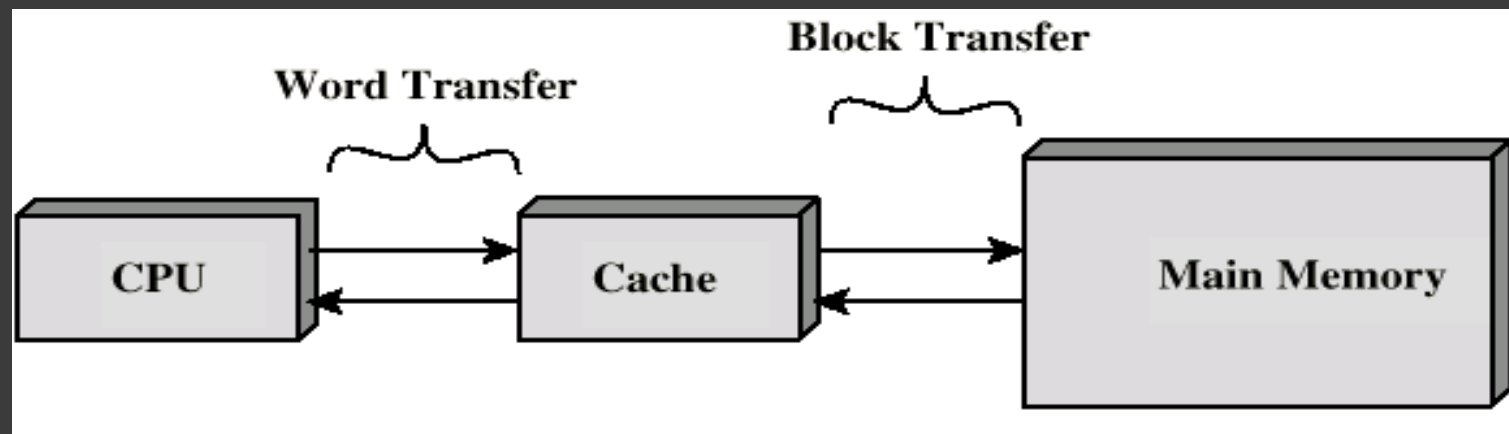
- ⦿ Semakin kecil waktu akses, semakin besar harga per bitnya.
- ⦿ Semakin besar kapasitas, semakin kecil harga per bitnya.
- ⦿ Semakin besar kapasitas, semakin besar waktu aksesnya.

Tabel 4.2 Tabel spesifikasi memori

Tipe memori	Teknologi	Ukuran Waktu	akses
Cache Memory	semikonduktor RAM	128 – 512 MB	KB 10 ns
Memori Utama	semikonduktor RAM	4 – 128 MB	50 ns
Disk magnetik	Hard Disk	Gigabyte	10 ms, 10MB/det
Disk Optik	CD-ROM	Gigabyte	300ms, 600KB/det
Pita magnetik	Tape	100 MB	Det -mnt, 10MB/mnt

CACHE MEMORY

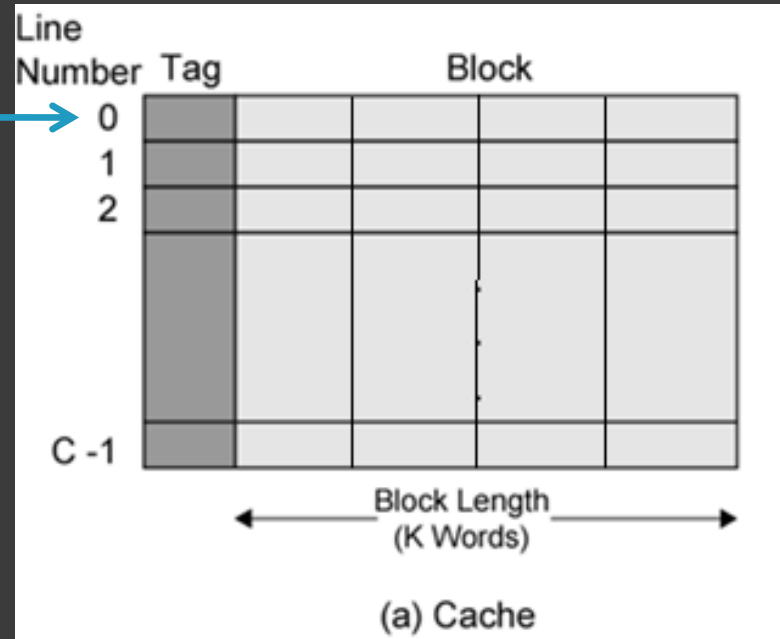
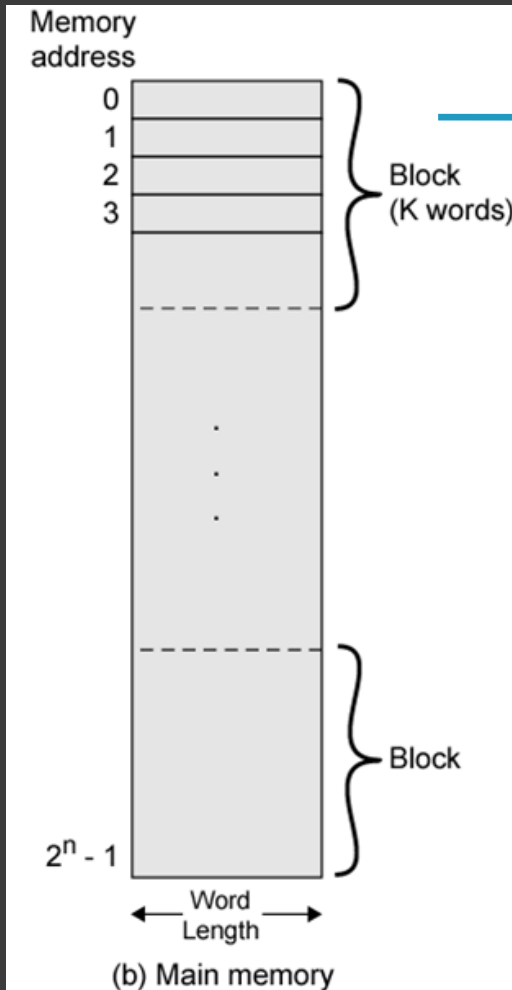
- Berfungsi mempercepat kerja memori sehingga mendekati kecepatan prosesor.
- Cache memori berisi salinan memori utama.
- Memori utama lebih besar kapasitasnya namun lambat operasinya, sedangkan cache memori berukuran kecil namun lebih cepat.



Cara pembacaan pada cache :

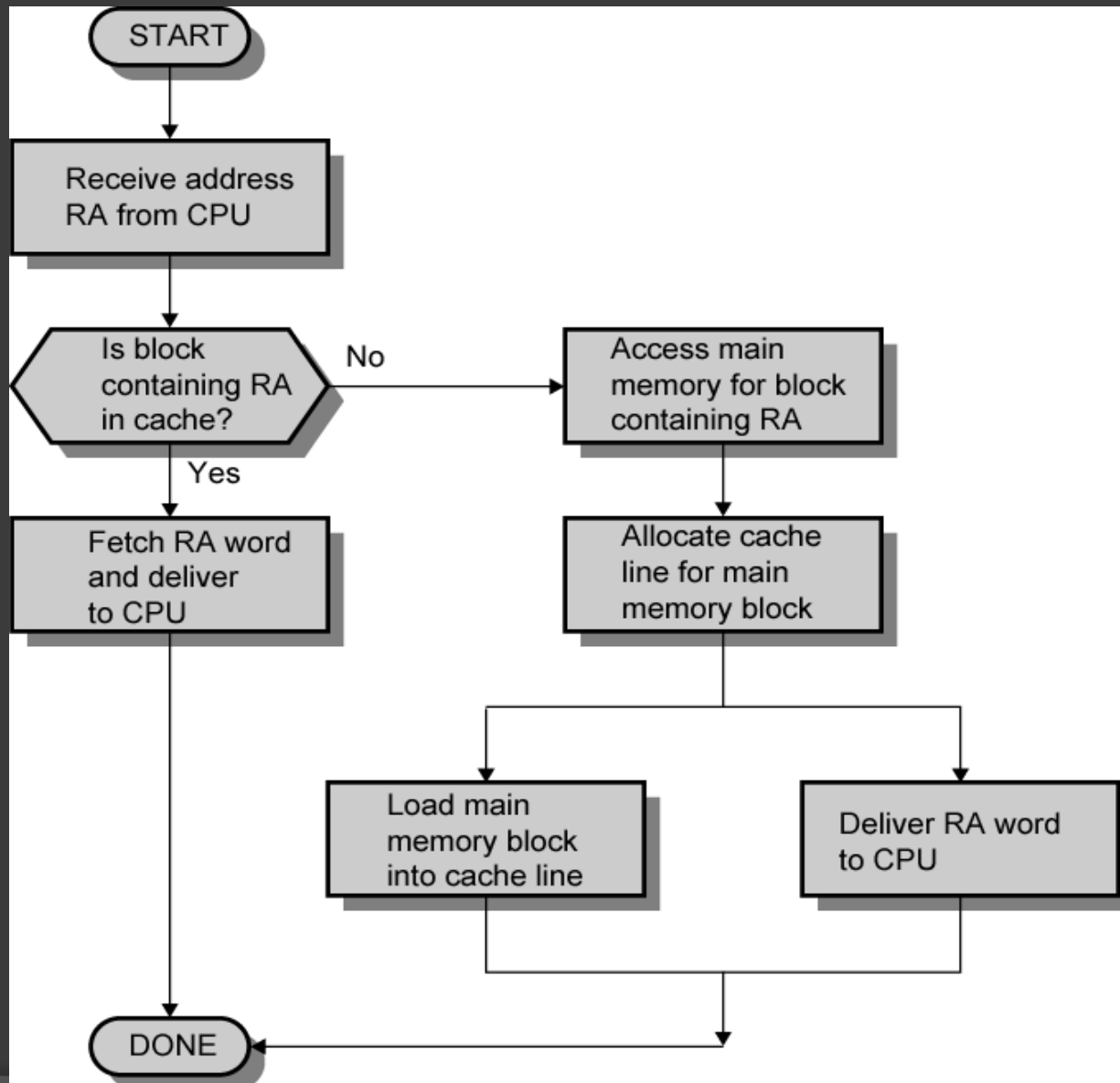
- CPU meminta data 1 alamat
- Data akan dicari di lokasi cache
- Jika ada maka akan langsung dikirim ke CPU
- Jika tidak ditemukan, cache akan meminta atau mengambil 1 blok data yang mengandung alamat yang diminta dari main memori.
- Akan dikirim satu blok data ke cache, cache akan mengirim 1 alamat yang diminta CPU
- Sisa data dari main memori akan disimpan di cache

Cache/Main Memory Structure



1 blok data pada main memori menjadi 1 line pada cache memori yang terbagi menjadi K word dan 1 tag.

Cache Read Operation - Flowchart



Elemen Rancangan Cache :

- a. Size / Kapasitas
- b. Mapping Function/Fungsi Pemetaan
- c. Replacement Algorithm/Algoritma Pengganti
- d. Write Policy
- e. Block Size / Ukuran Blok
- f. Number of Caches / Jumlah Cache

Size / Kapasitas

- Ukuran memori cache sangatlah penting untuk mendongkrak kinerja komputer.
- Semakin besar kapasitas cache tidak berarti semakin cepat prosesnya,
- dengan ukuran besar akan terlalu banyak gate pengalamatannya sehingga akan memperlambat proses.

Mapping Function

Cache memori yang besar akan mengurangi kecepatan pada saat pencarian data.

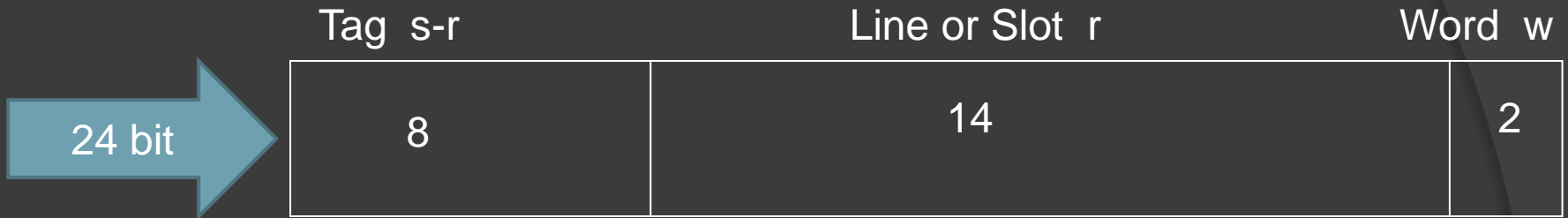
Terdapat tiga metode, yaitu

- a. pemetaan langsung,
- b. pemetaan asosiatif, dan
- c. pemetaan asosiatif set.

Direct Mapping/Pemetaan Langsung

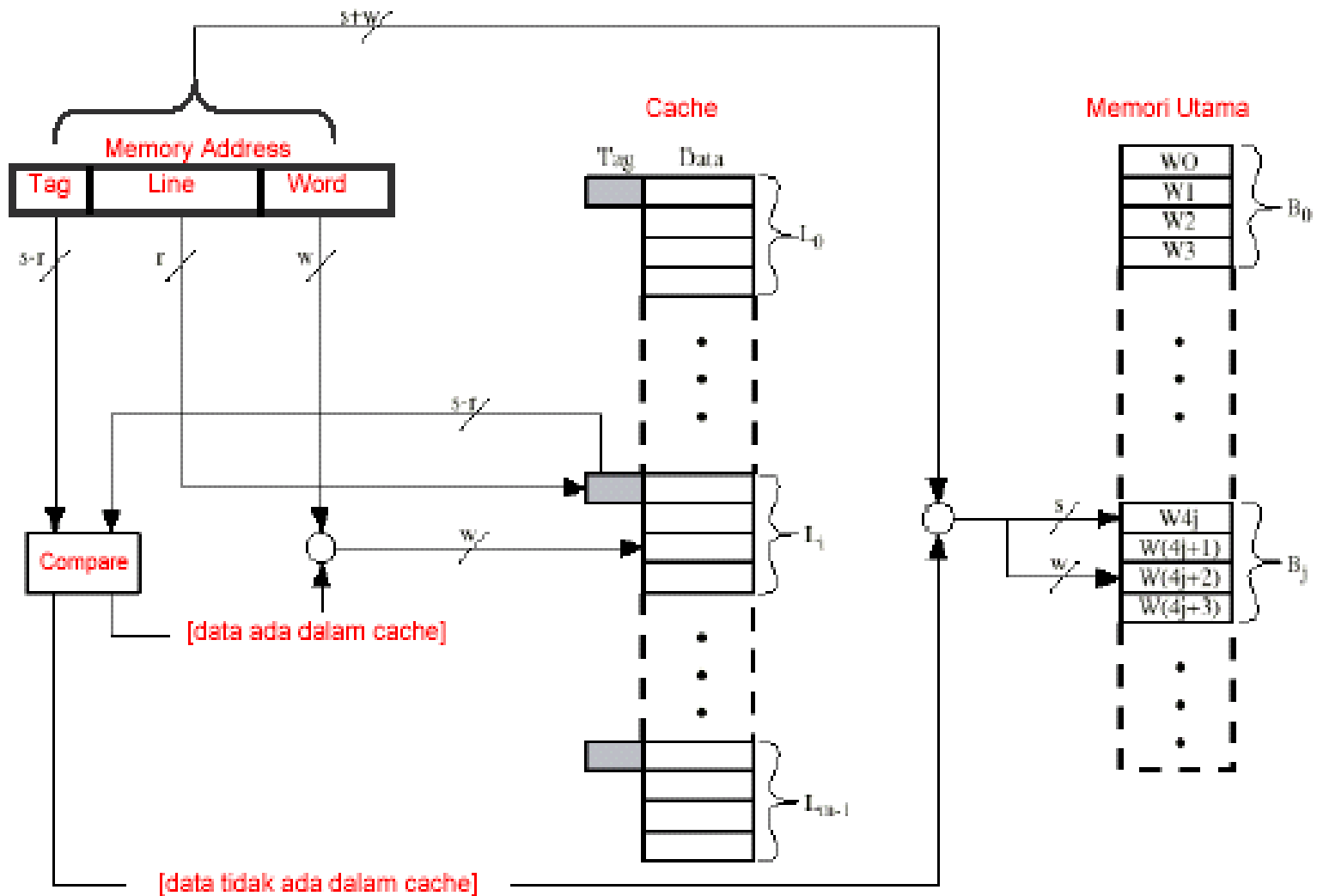
- Setiap blok pada main memori terpetakan hanya satu baris pada cache
- Alamat dari CPU akan dibagi 2 bagian untuk menentukan word berapa dan sisa blok.
- Fungsi pemetaan diimplementasikan dengan menggunakan alamat, yang terdiri dari tiga field (tag, line, word)

Struktur alamat Direct Mapping



- ⊙ 24 bit alamat \rightarrow 16Mbyte dari main memori
- ⊙ 2 bit word (4 byte block)
- ⊙ 22 bit block
 - 8 bit tag ($=22-14$)
 - 14 bit slot or line
- ⊙ Tidak ada dua buah blok yang dipetakan ke nomor saluran yang memiliki tag sama

Direct Mapping Cache Organization



Ket.:

- Panjang Alamat = $(s + w)$ bit
- Number of addressable units = 2^{s+w} words or bytes
- block = line = 2^w words or bytes
- Nilai block di main memory = $2^{s+w} / 2^w = 2^s$
- Besar line pada cachec = $m = 2^r$
- Lenar tag = $(s - r)$ bit

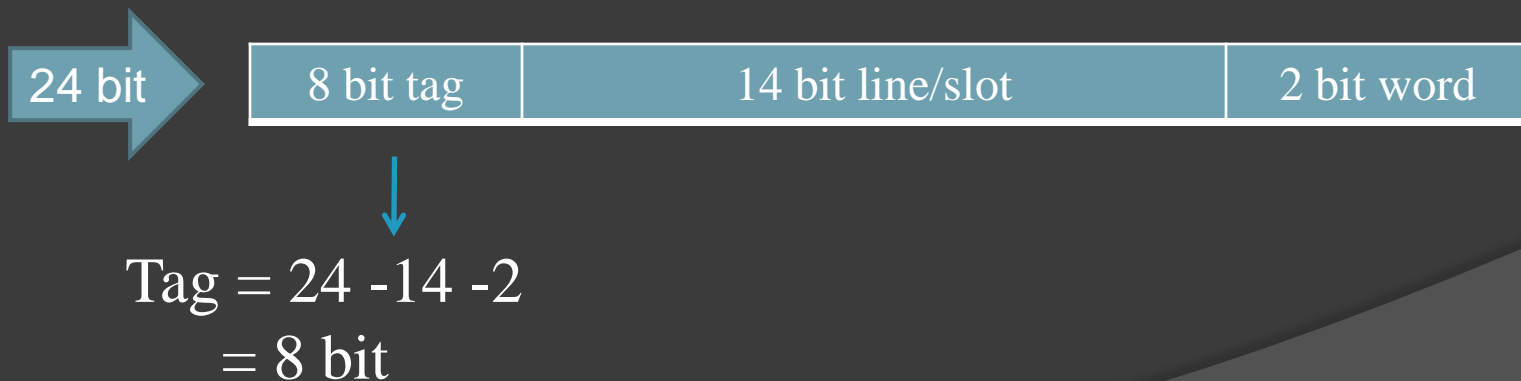
Contoh :

1. Suatu sistem komputer memiliki memori utama sebesar 16 Mbyte, memiliki cache sebesar 64 Kbyte. Transfer data antara memori utama dengan cache dalam ukuran blok besarnya 4 byte. Jika digunakan pemetaan langsung (*direct mapping*) tunjukkan masing-masing berapa bit untuk tag, slot/line dan word.

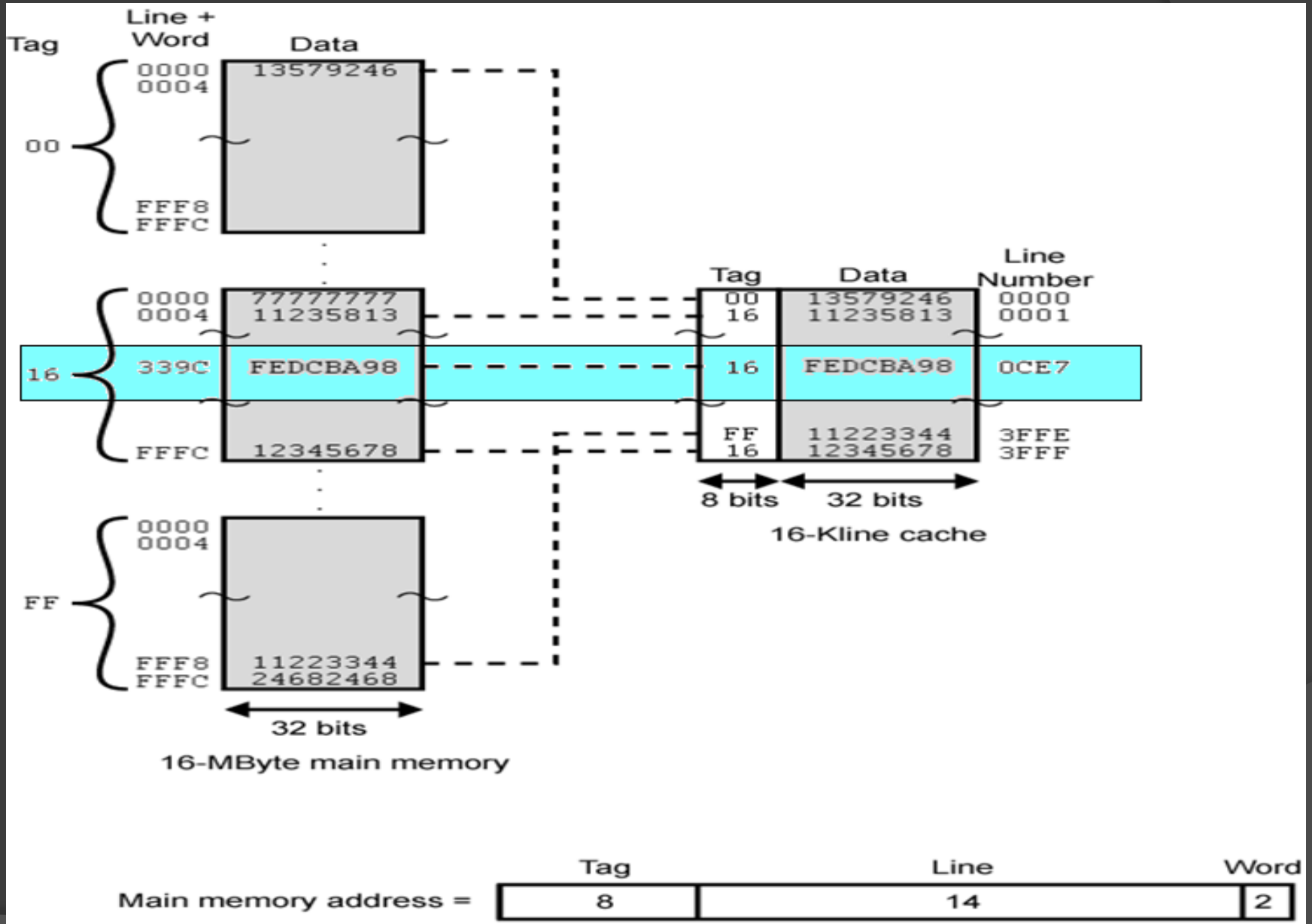
Penyelesaian:

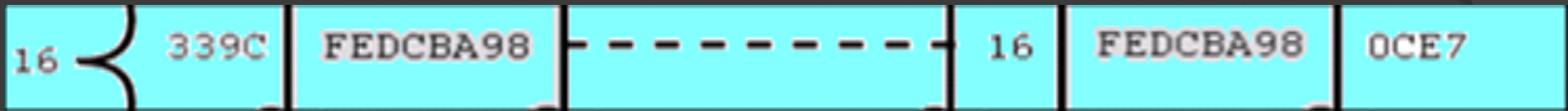
- Memori utama : $16 \text{ MB} = 2^{20} + 2^4 = 2^{24}$ → lebar alamat = 24 bit
- Size blok : 4 B = 2^2 → lebar word = 2 bit
- Cache : $64 \text{ KB} = \frac{64 \text{ KB}}{4 \text{ B}} = 16 \text{ K}$
 $16 \text{ K} = 2^{10} + 2^4 = 2^{14}$ → lebar Line/slot = 14 bit

Jadi direct mapping untuk main memori 16MB, cache 64KB dan ukuran blok=4B menjadi



2. Pemetaan langsung/ direct mapping :





Main memori

cache memori

Tag

data

data

16339C

FE
DC
BA
98



tag

0

1

2

3

16	FE	DC	BA	98
----	----	----	----	----

Bagaimana Penyelesaiannya??????

Step Penyelesaian:

1. Cari pemetaan alamatnya (jika menggunakan direct mapping) cari: berapa tag, line dan word
2. Alamat hexa **16339C** di konversi ke biner
3. Ambil LSB untuk bit word.
4. Sisa digit biner diambil sebesar line dan dihitung ulang.
5. Sisa digit dari word dan line adalah tag.

Penyelesaian :

8 bit tag	14 bit line/slot	2 bit word
-----------	------------------	------------

16339C → dihexakan

1	6		3	3	9		c	
-----			-----					-----
0001	0110		0011	0011	1001		1100	

tag = 16

line = 0CE7

word = 0

Jadi masukan data 16339C dengan direct mapping dihasilkan alamat tag = 16, line/slot = 0CE7 dan word = 0

3. Konversikan alamat memori dibawah ini

a. 256 MB =.....

b. 512 MB =

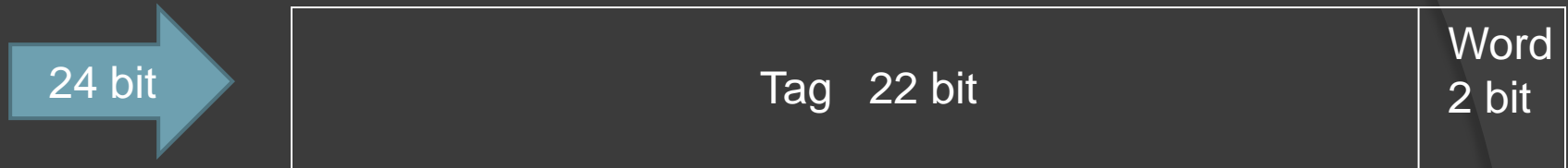
c. 512 KB =.....

4. Suatu sistem komputer memiliki memori utama sebesar 512 Mbyte, memiliki cache sebesar 512 Kbyte. Transfer data antara memori utama dengan cache dalam ukuran blok besarnya 32 byte. Jika digunakan pemetaan langsung (*direct mapping*) tunjukkan masing-masing berapa bit untuk tag, slot/line dan word.

Assosiatif Mapping

- Setiap blok memori utama dapat dimuat ke sembarang saluran cache.
- Alamat memori utama diinterpretasikan dalam field tag dan field word oleh kontrol logika cache.
- Alamat memori utama diinterpretasikan dalam field tag dan field word oleh kontrol logika cache

Associative Mapping Address Structure



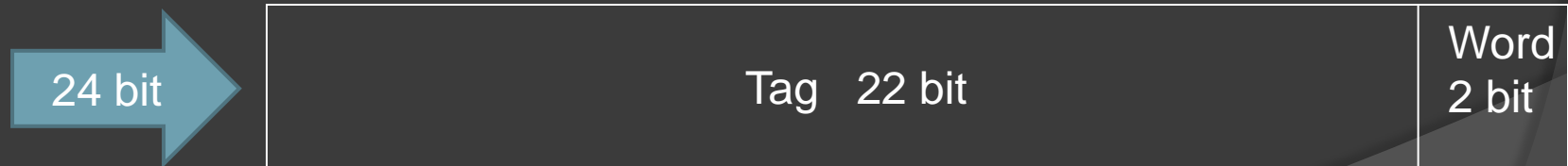
- 24 bit alamat \rightarrow 16Mbyte dari main memori
- 2 bit word (4 byte block)
- 22 bit tag
- Tag 22 bit harus disimpan dalam blok data 32 bit pada cache.

Contoh :

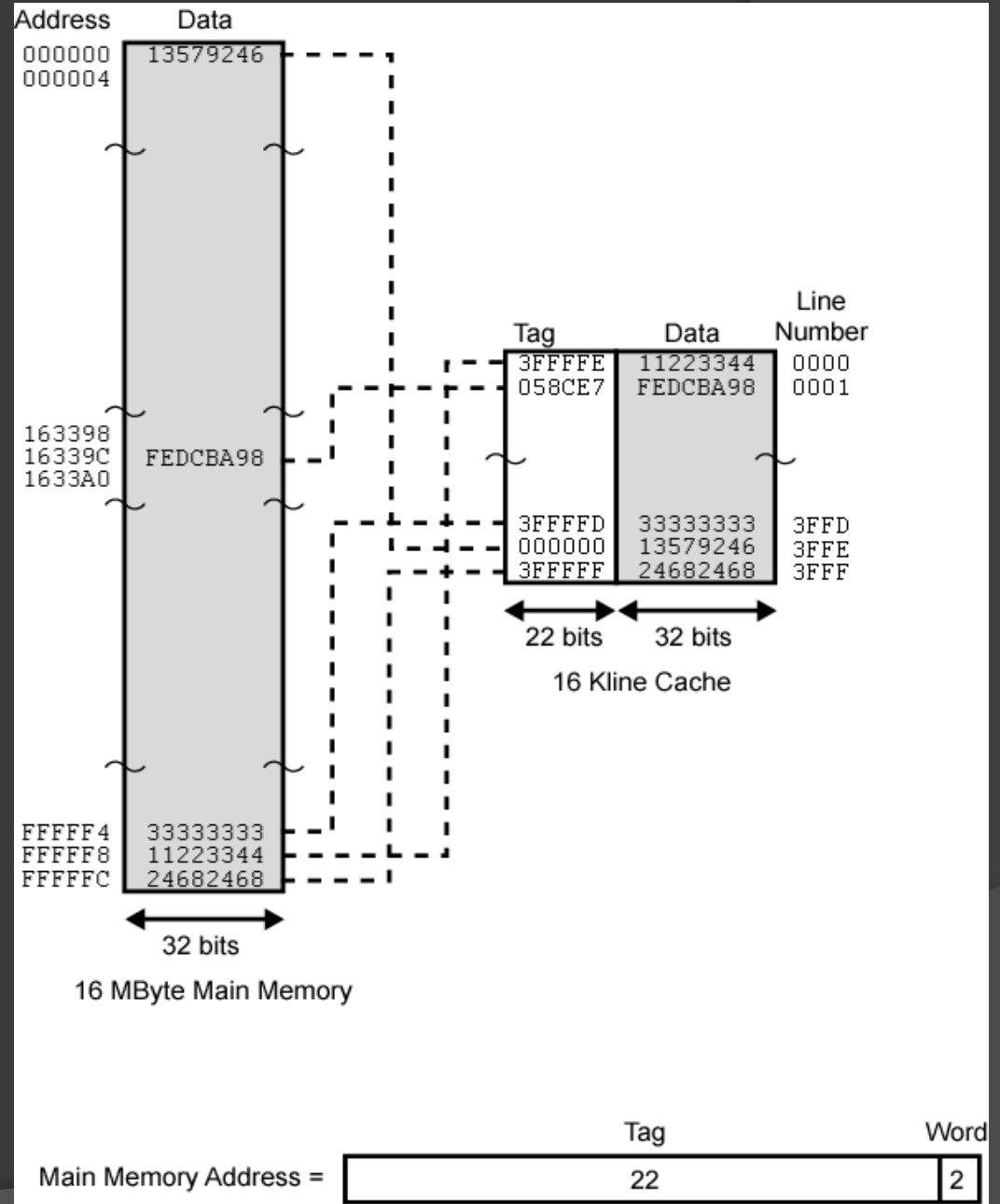
1. Diket : memori = 16 MB
cache = 64 KB
blok = 4 B

- Memori : 16 MB = $2^{20} + 2^4 = 2^{24} \rightarrow$ lebar alamat = 24 bit
- Size blok : 4 B = $2^2 \rightarrow$ lebar word = 2 bit

Jadi associative mapping untuk main memori 16MB, cache 64KB dan ukuran blok=4B menjadi



2. Dengan diberikan data sama : main memori 16 MB, alamat data 16339C



3. Suatu sistem komputer memiliki memori utama sebesar 512 Mbyte, memiliki cache sebesar 526 Kbyte. Transfer data antara memori utama dengan cache dalam ukuran blok besarnya 32 byte. Jika digunakan pemetaan Asosiatif (*Assosiative mapping*) tunjukkan masing-masing berapa bit untuk tag dan word.

Set Associative Mapping

- Terbagi dari tag, set dan word
- Cache dibagi beberapa set
- Setiap set berisi beberapa line

Macam Set Associative mapping :

- a. Two-way set (2-way set) \rightarrow dalam 1 line terdapat 2 set.
- b. Four-way set (4-way set) \rightarrow dalam 1 line terdapat 4 set.

Contoh:

1. Soal sama...

diket : memori 16 MB

cache 64 KB

blok 4 B

Buat dalam two-way set ?

2. Dari main memori seperti soal satu diberikan masukan data : 16339C cari dengan 2-way set ?

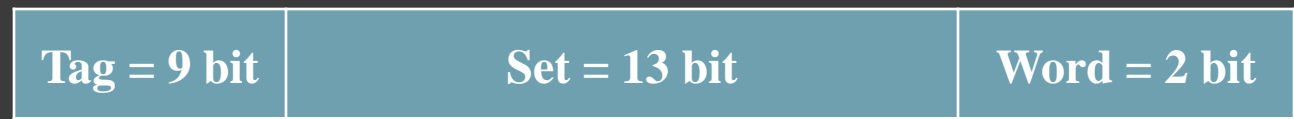
Penyelesaian:

1. $16 \text{ MB} = 2^{20} + 2^4 = 2^{24} \rightarrow \text{alamat masuk} = 24 \text{ bit}$

$4 \text{ KB} = 2^2 \rightarrow \text{lebar word} = 2 \text{ bit}$

$64 = \frac{64 \text{ KB}}{4 \text{ B}} = 16 \text{ K} = \frac{16 \text{ K}}{2} = 8 \text{ K}$

$8 \text{ K} = 2^{10} + 2^3 = 2^{13} \rightarrow \text{set } 13 \text{ bit}$



2. 16339C → di binerkan

0001 0110 0011 0011 1001 1100

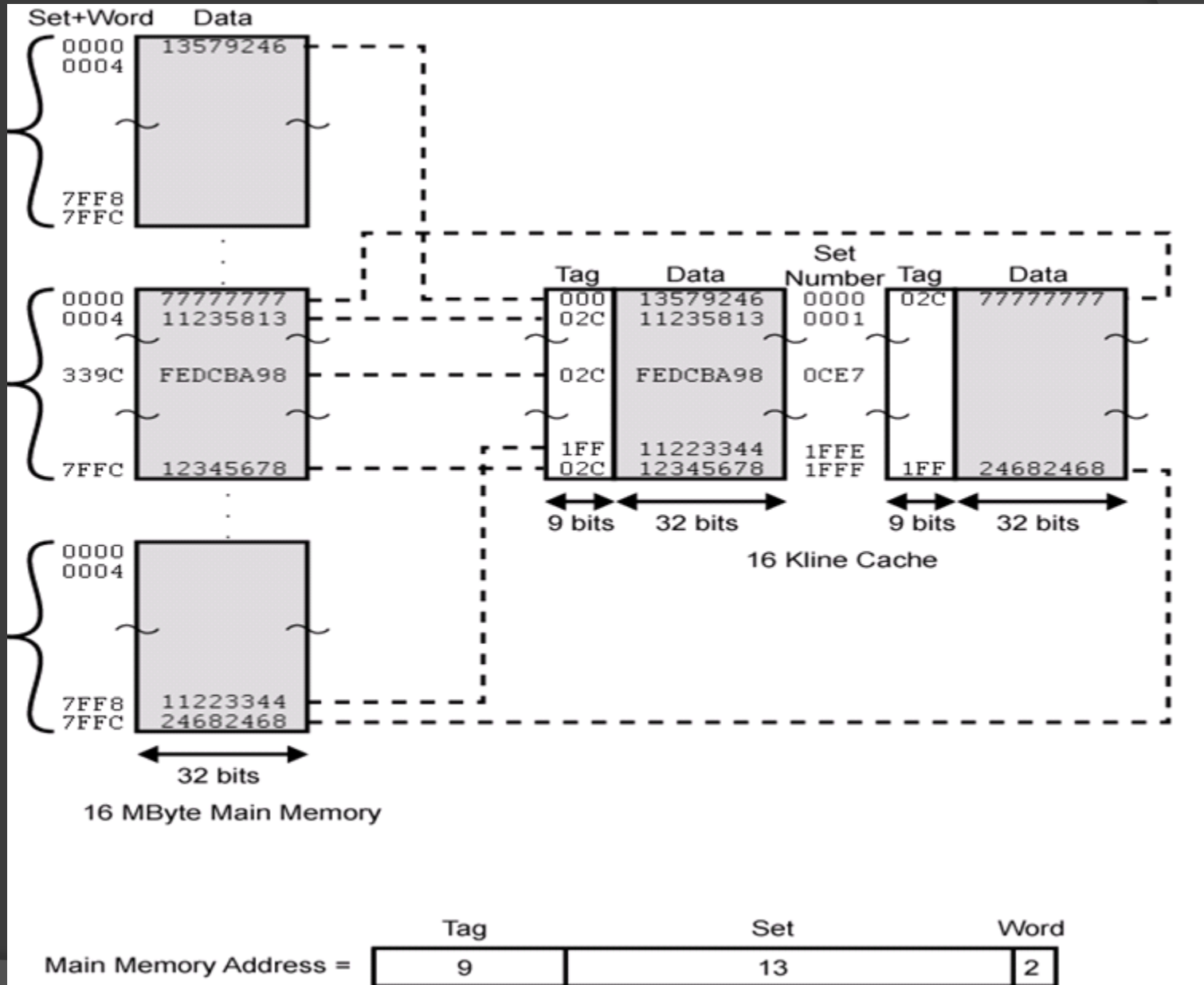
tag = 02C

set = 0CE7

w = 0

Jadi 16339C dengan two-way set data masuk pada alamat 0CE7 dengan $t = 02C$. $S = 0CE7$ dan $w = 0$.

Gambar contoh two way set



TERIMA KASIH

Kuis Memori Cache – Pertemuan ke-5

1. Terangkan apa yang dimaksud memori cache?
2. Jelaskan proses terjadinya cache hit dan cache miss pada memori cache?
3. Uraikan tentang cache memory L3?

Note: Kerjakan di selembar kertas/buku catatan mata kuliah Arsikom kalian!