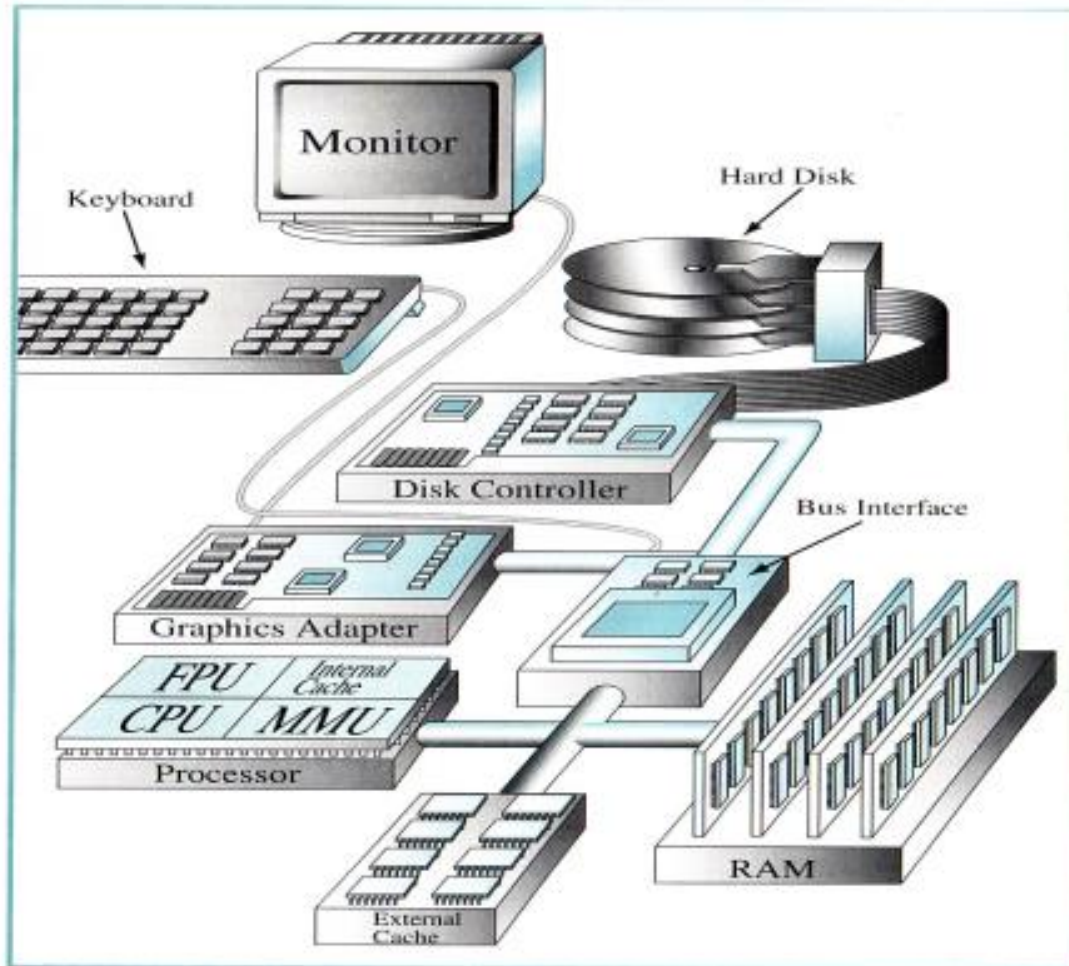


ORGANISASI ARSITEKTUR KOMPUTER CACHE MEMORY

**STT TERPADU NURUL FIKRI
TEKNIK INFORMATIKA
2017**

SISTEM KOMPUTER SAAT INI



LOKASI MEMORY KOMPUTER

1. **Memori Internal / Primary Memory**

Memori yang ada di dalam internal sistem prosesor dan motherboard.

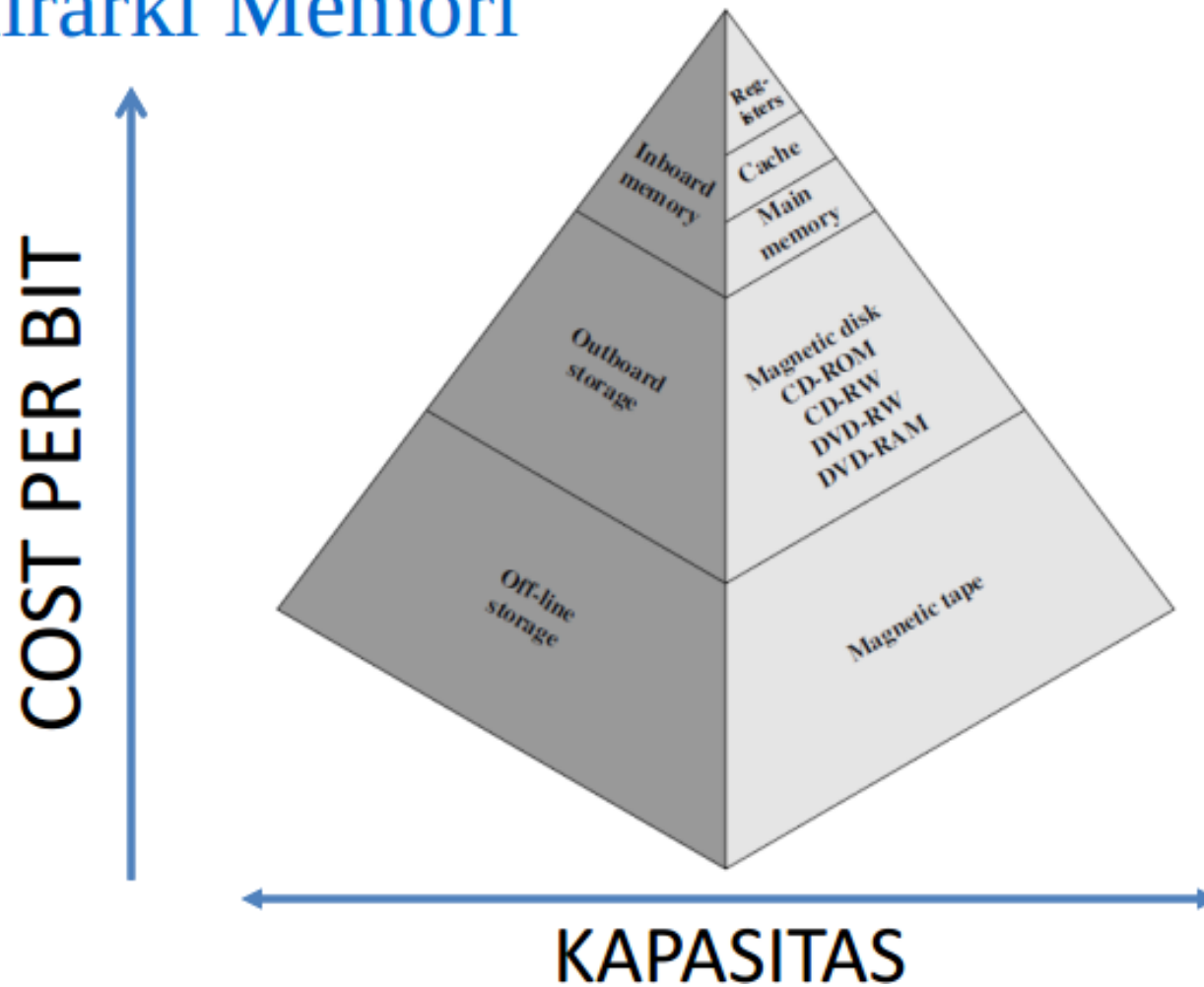
Contoh: register, cache memory, main memory/RAM

2. **Memori Eksternal / Secondary Memory**

Memori yang ada di luar sistem prosesor dan motherboard.

Contoh: harddisk, CD, DVD, flashdisk

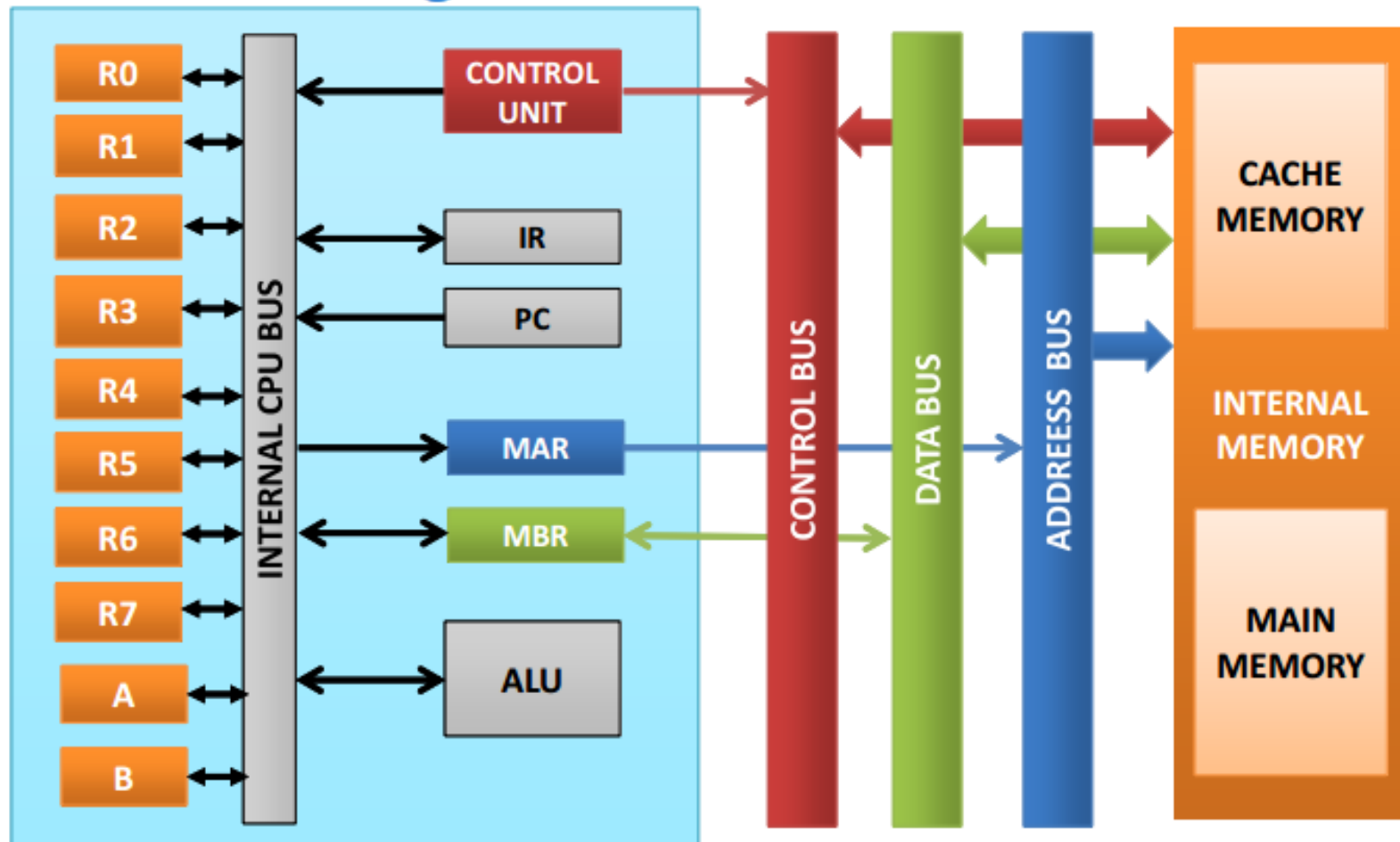
Hirarki Memori



FAKTOR PERTIMBANGAN PENGGUNAAN MEMORI

1. Kapasitas
adalah jumlah bit/byte yang dapat ditampung dalam suatu memori.
2. Access time
Yaitu waktu yang diperlukan untuk membaca atau menulis ke memori.
3. Transfer rate
adalah kecepatan transfer data yang dapat dilakukan dari/ke suatu memori.
4. Cost per bit
adalah biaya per bit yang dibutuhkan untuk memproduksi terkait bahan dan teknologi dari suatu memori, dimana akan mempengaruhi harga jual.

Bagan Interkoneksi CPU



MEMORI INTERNAL

Adalah memori yang kerjanya berhubungan dengan CPU dalam pemrosesan data.

- CPU memiliki kecepatan yang sangat tinggi. Misal Dual Core dengan clock 1GHz dapat memiliki instruksi yang bisa selesai dalam 1 ns, sehingga dalam selang waktu 1 ns bisa diperlukan akses untuk membaca/menulis ke memori.
- Diperlukan memori yang memiliki access time yang cepat, dengan kapasitas yang mencukupi. Karena itu digunakan RAM.
- Memori internal terdiri atas main memory dan cache memory.

JENIS MEMORI RAM

Dynamic RAM (DRAM)	Static RAM (SRAM)
<ol style="list-style-type: none">1. Sel tersusun atas transistor dan kapasitor2. Perlu dilakukan refresh secara periodik untuk menjaga isi logic di sel3. Memiliki kerapatan sel lebih tinggi4. Kapasitas lebih tinggi5. Access time cepat6. Harga lebih rendah7. Dipakai untuk main memory	<ol style="list-style-type: none">1. Sel tersusun atas transistor-transistor2. Tidak perlu dilakukan refresh terhadap isi sel memori3. Kerapatan sel tidak dapat lebih rapat dari DRAM4. Kapasitas lebih rendah5. Access time sangat cepat6. Harga sangat mahal7. Dipakai untuk cache memory

Perbandingan Teknologi DRAM

	Clock Frequency (MHz)	Transfer Rate (GB/s)	Access Time (ns)	Pin Count
SDRAM	166	1.3	18	168
DDR	200	3.2	12.5	184
RDRAM	600	4.8	12	162

Sebelum tahun 2000, teknologi DDR dan RDRAM bersaing sebagai standar memori yang digunakan untuk komputer.

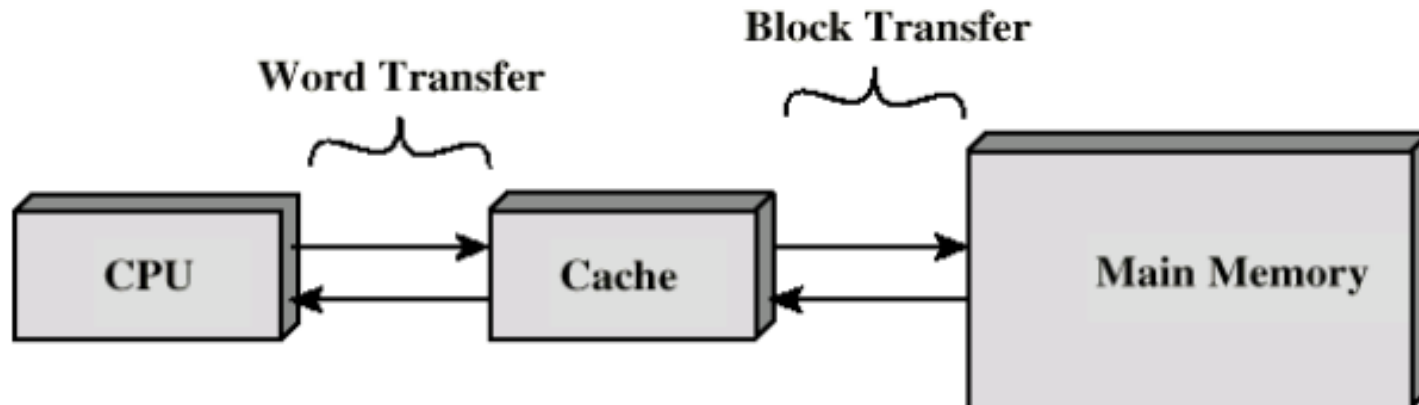
Pada perkembangannya teknologi DDR yang kemudian banyak digunakan dan terus ditingkatkan sehingga menjadi menjadi DDR2, DDR3 dan seterusnya.

PERMASALAHAN “CPU-MEMORY BOTTLENECK”

- Main memory jauh lebih lambat dibandingkan dengan CPU
- Bisa saja mengurangi selisih kecepatan dengan menggunakan Static RAM untuk main memory
 - Akan jauh lebih cepat
 - Tapi biaya bahan & pembuatan memori jadi sangat sangat tinggi
 - Siapa yang mau beli?

Cache Memory

- Merupakan sejumlah kecil memori cepat SRAM
- Ada di antara jalur main memory dan CPU
- Secara fisik ada dalam chip atau modul CPU
- Menduplikasi sebagian block data di main memory
- Sehingga dapat memberikan data ke CPU jika membutuhkan tanpa CPU harus menunggu lama main memory

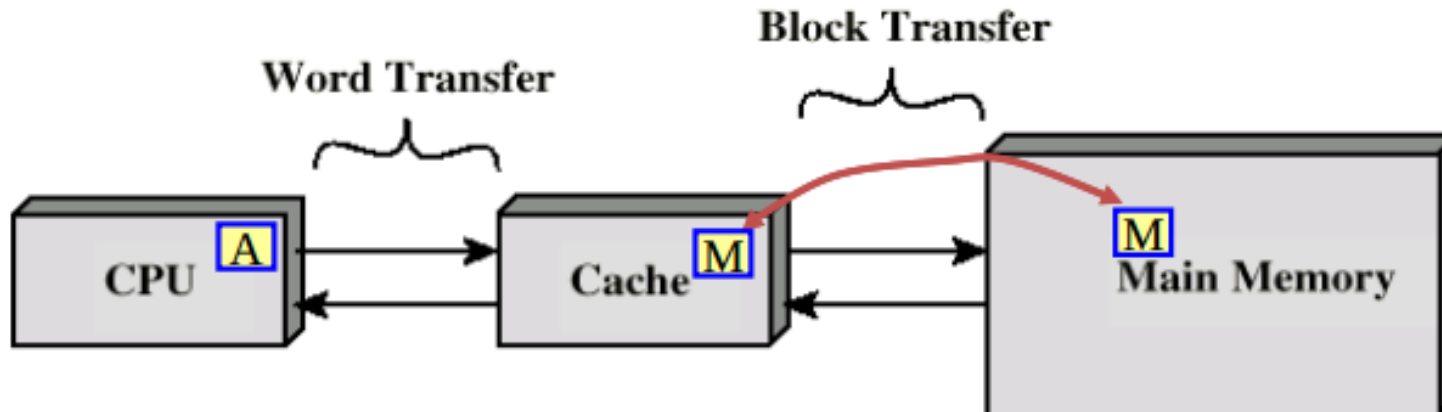


Ilustrasi Cache Memory

Loop: MOV M,A
...
MOV A,M
...
MOV M,A
...
MOV A,M
...

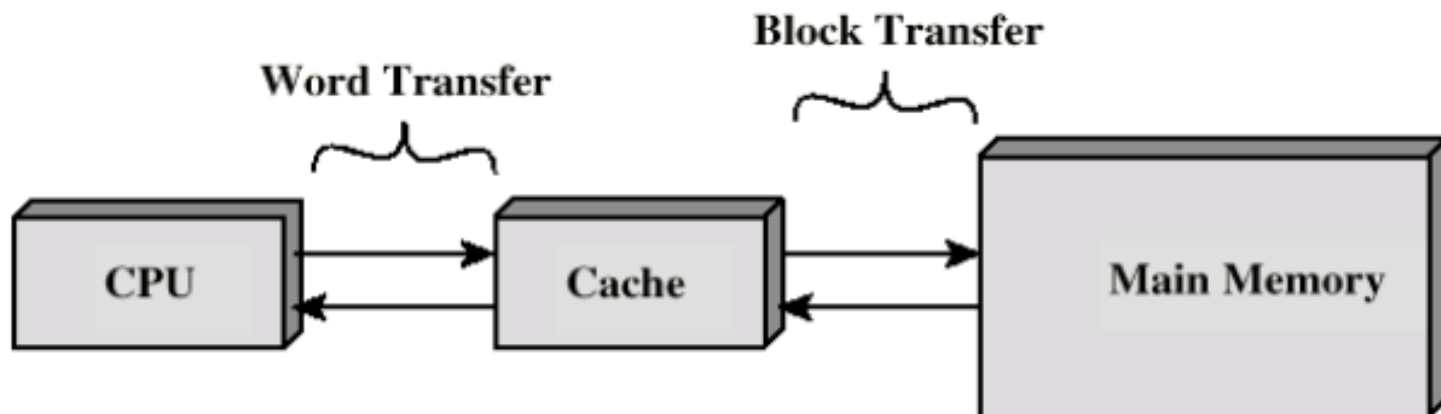
Locality of Reference:

Baris program yang dibuat manusia hampir selalu mengambil/mengolah data yang sebelumnya juga baru diambil/diolah



Cara Kerja Cache Memory

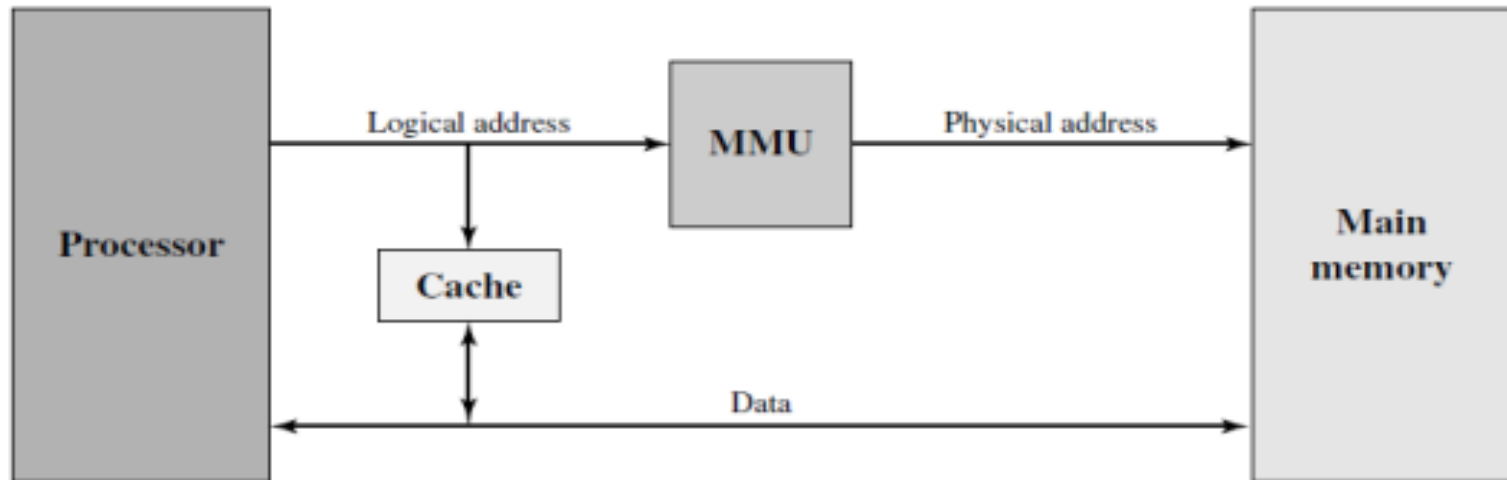
- CPU meminta data dari suatu lokasi di main memory
- Cache mengecek data tersebut apakah ada di dalam dirinya
- Jika ada, maka ambil dari cache (disebut HIT, akan menjadi cepat)
- Jika tidak ada, ambil dari main memory untuk dikopi ke cache dan dibawa ke CPU (disebut MISS, akan menjadi lambat)



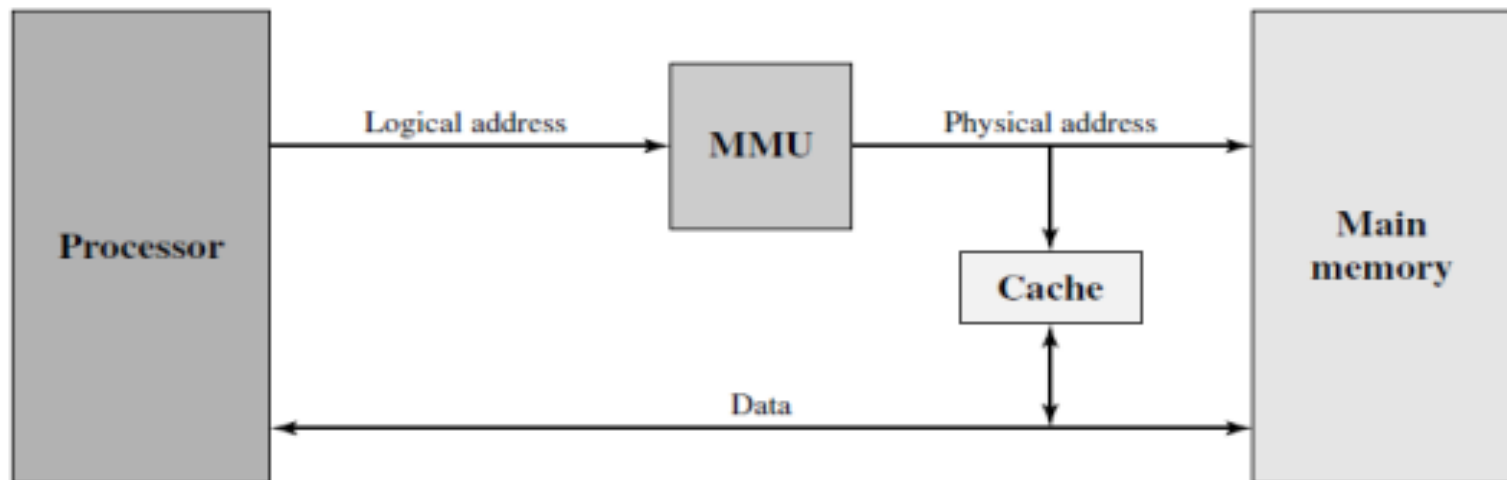
PERTIMBANGAN DALAM DESAIN CACHE MEMORY

1. Alamat Cache: logical atau physical
2. Ukuran Cache: seberapa besar
3. Level Cache: 1, 2, 3 atau berapa level
4. Algoritma Penggantian: LRU, FIFO, LFU, Random
5. Cara Mapping: Direct Mapping atau Set Associative
6. Cara Write: Write Through atau Write Back

Alamat Cache



(a) Logical cache



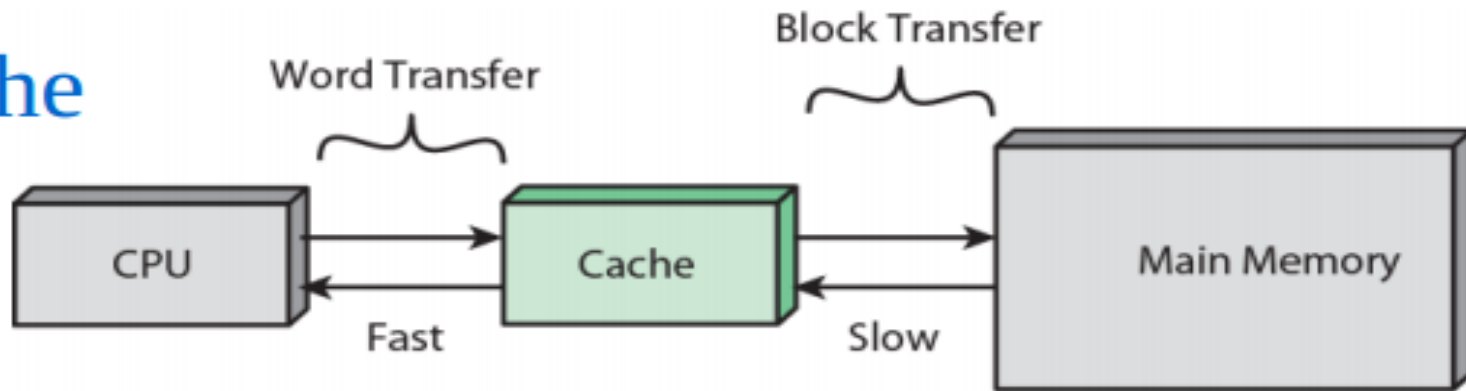
(b) Physical cache

Ukuran Cache Memory

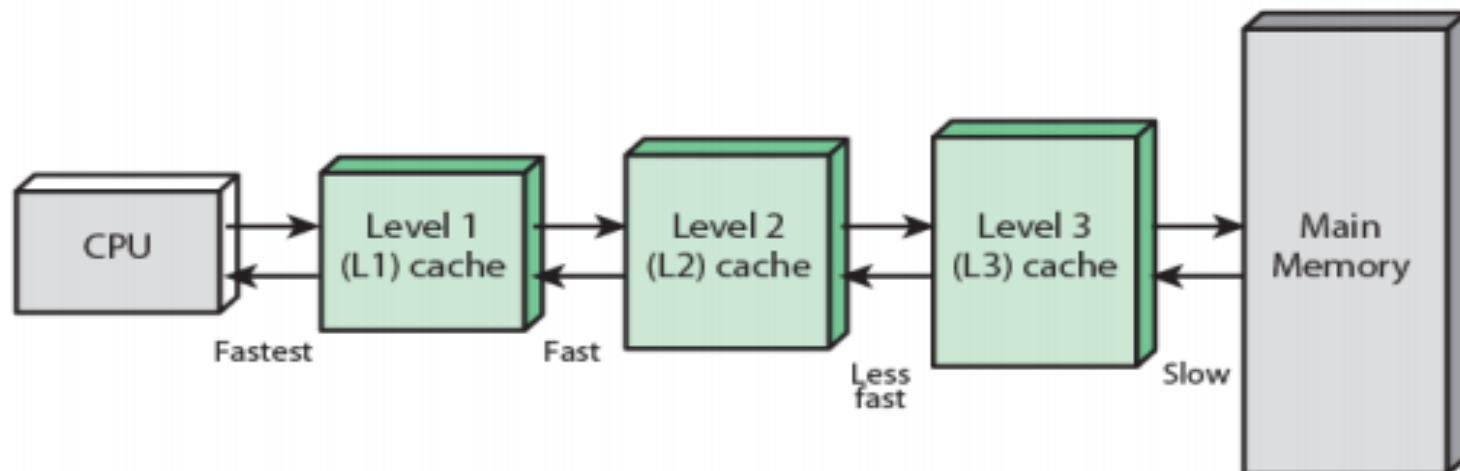
Processor	Type	Year of Introduction	L1 cache ^a	L2 cache	L3 cache
IBM 360/85	Mainframe	1968	16 to 32 KB	—	—
PDP-11/70	Minicomputer	1975	1 KB	—	—
VAX 11/780	Minicomputer	1978	16 KB	—	—
IBM 3033	Mainframe	1978	64 KB	—	—
IBM 3090	Mainframe	1985	128 to 256 KB	—	—
Intel 80486	PC	1989	8 KB	—	—
Pentium	PC	1993	8 KB/8 KB	256 to 512 KB	—
PowerPC 601	PC	1993	32 KB	—	—
PowerPC 620	PC	1996	32 KB/32 KB	—	—
PowerPC G4	PC/server	1999	32 KB/32 KB	256 KB to 1 MB	2 MB
IBM S/390 G4	Mainframe	1997	32 KB	256 KB	2 MB
IBM S/390 G6	Mainframe	1999	256 KB	8 MB	—
Pentium 4	PC/server	2000	8 KB/8 KB	256 KB	—
IBM SP	High-end server/ supercomputer	2000	64 KB/32 KB	8 MB	—
CRAY MTA ^b	Supercomputer	2000	8 KB	2 MB	—
Itanium	PC/server	2001	16 KB/16 KB	96 KB	4 MB
SGI Origin 2001	High-end server	2001	32 KB/32 KB	4 MB	—

Makin besar makin mahal!

Level Cache



(a) Single cache

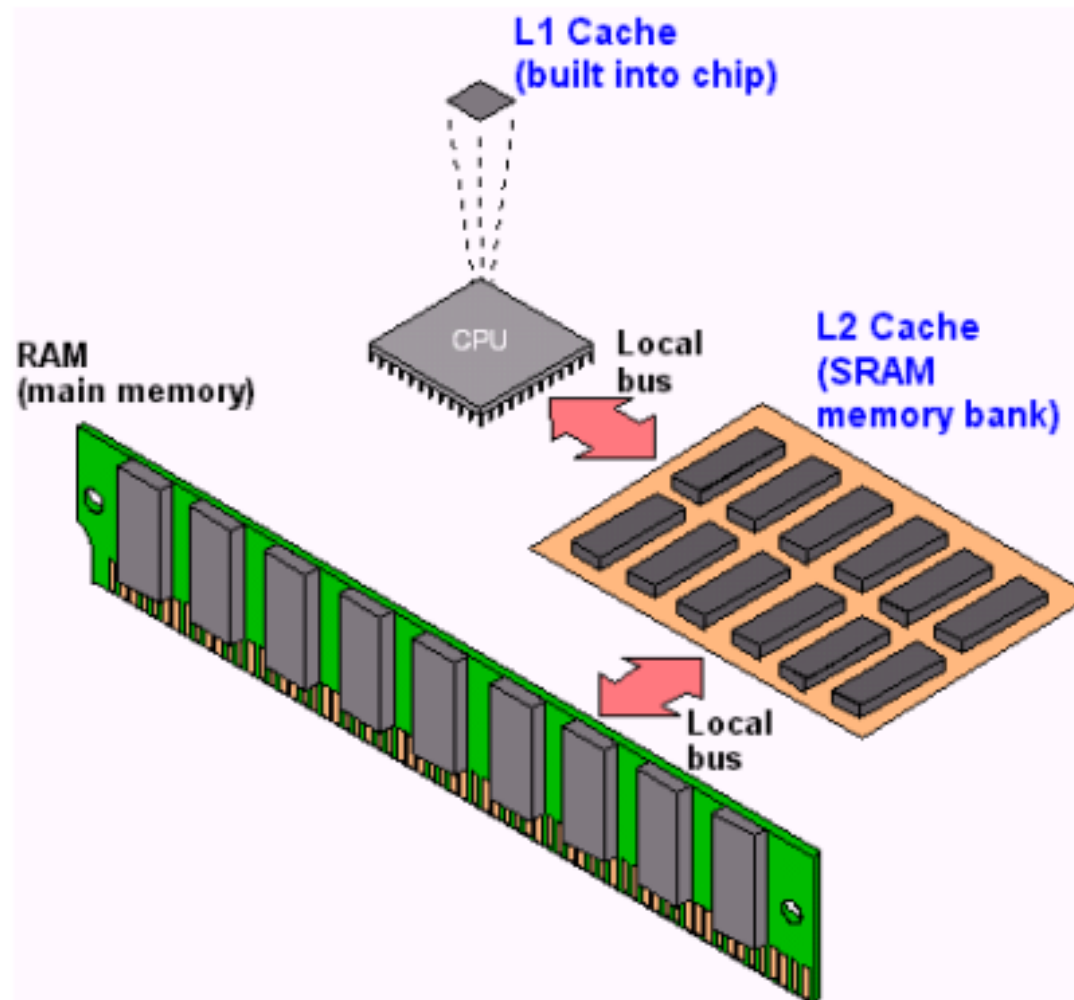


(b) Three-level cache organization



From Computer Desktop Encyclopedia
© 1999 The Computer Language Co. Inc.

STT TERPADU
NURUL FIKRI



ALGORITMA PENGANTIAN DATA DALAM CACHE

1. Least Recently Used (LRU)
Data yang sudah paling lama tidak digunakan diganti dengan yang baru
2. First In First Out (FIFO)
Data yang terbaru masuk menggantikan data yang paling lama
3. Least Frequently Used (LFU)
Data yang paling jarang digunakan diganti dengan yang baru
4. Random
Penggantian secara acak

CARA MAPPING CACHE

Bagaimana menata duplikasi data block memory di cache?

1. Direct Mapping
Memetakan bagian-bagian awal secara berurutan. Sederhana dan tidak mahal.
2. Associative Mapping
Memetakan dimana saja, tidak perlu berurutan. Lebih canggih namun rumit.

CARA WRITE CACHE

Bagaimana menuliskan ke main memory jika terjadi perubahan data?

1. Write Through

Setiap kali ada perubahan data oleh CPU terhadap data cache, maka ditulis perubahannya juga di main memory. Memastikan isi main memory selalu valid pada setiap waktu. Namun mengakibatkan traffic penulisan data yang tinggi.

2. Write Back

Baru akan ditulis ke main memory ketika data di cache akan digantikan dengan yang baru. Mengurangi traffic penulisan data ke main memory namun rumit.

TERIMA KASIH



Thank you very much for your kind attention