COMPUTER ORGANIZATION AND ARCHITECTURE (COA)

EET 2211

4TH SEMESTER – CSE & CSIT

CHAPTER 4, LECTURE 16

CHAPTER 4 – CACHE MEMORY

TOPICS TO BE COVERED

- Computer Memory System Overview
- Cache Memory Principles
- Elements of Cache Design
- Cache Organization

LEARNING OBJECTIVES

After studying this chapter, you should be able to:

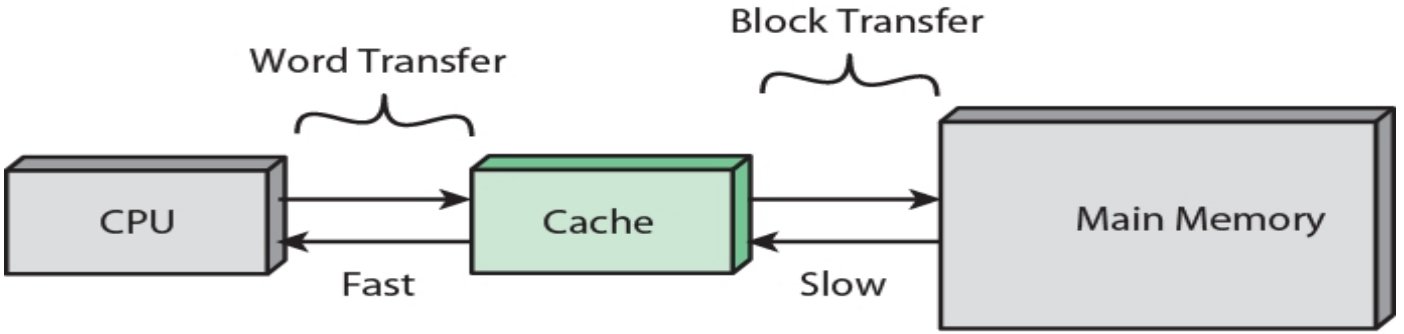
- Present an overview of the main characteristics of computer memory systems and the use of a memory hierarchy.
- Describe the basic concepts and intent of cache memory.
- Discuss the key elements of cache design.
- Distinguish among direct mapping, associative mapping and set-associative mapping.
- Explain the reasons for using multiple levels of cache.
- Understand the performance implications of multiple levels of memory.

4.2 CACHE MEMORY PRINCIPLES

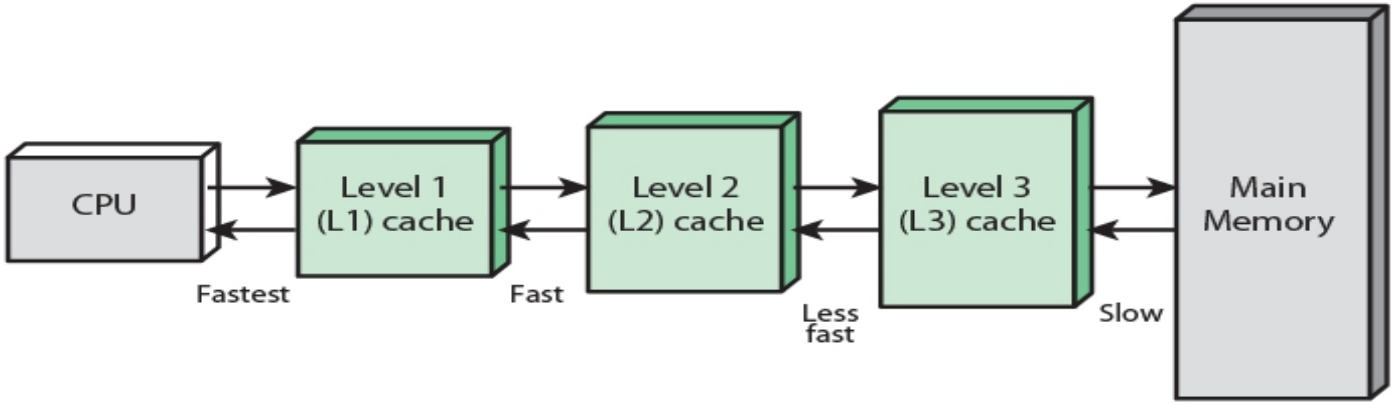
Cache memory is designed to combine the memory access time of expensive, high speed memory combined with large memory size of less expensive, lower speed memory.

- Small amount of fast memory
- Expensive
- Sits between normal main memory and CPU
- May be located on CPU chip or module

Cache and Main Memory

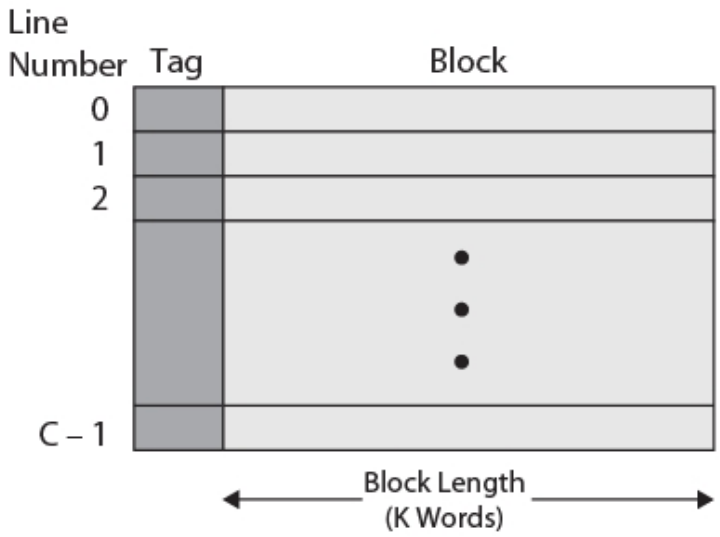


(a) Single cache

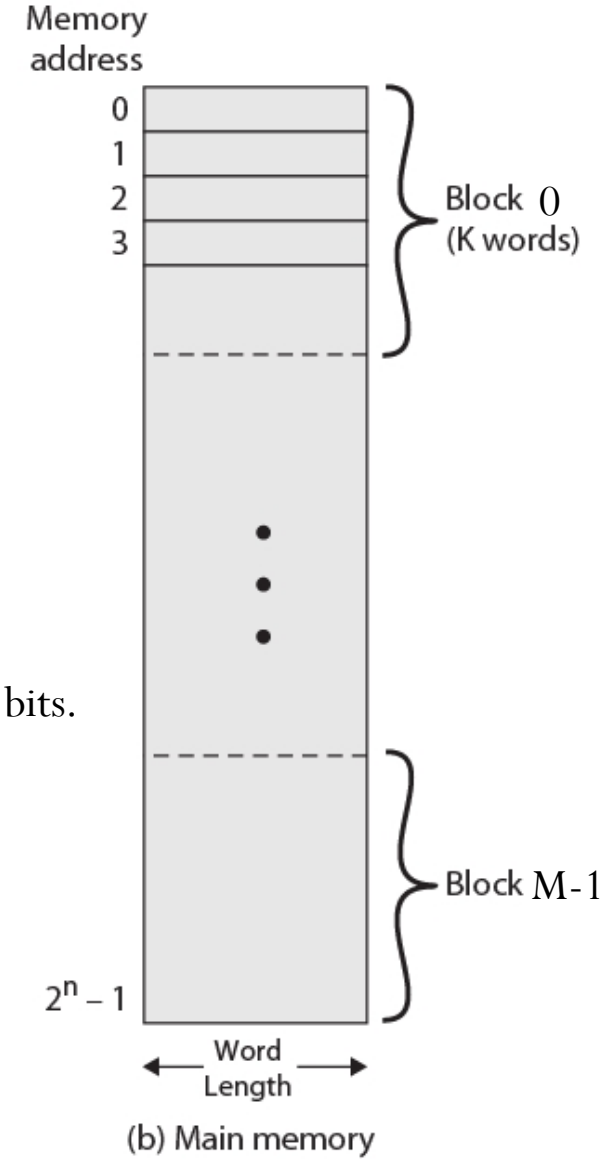


(b) Three-level cache organization

Cache/Main Memory Structure



(a) Cache



(b) Main memory

The length of line(line size) not include tag and control bits.

$C << M$

CACHE MEMORY

LINE NO.	TAG	BLOCK
0	3	384
1	1	129
2	0	2
…		
…		
126		
127	31	4096

MAIN MEMORY

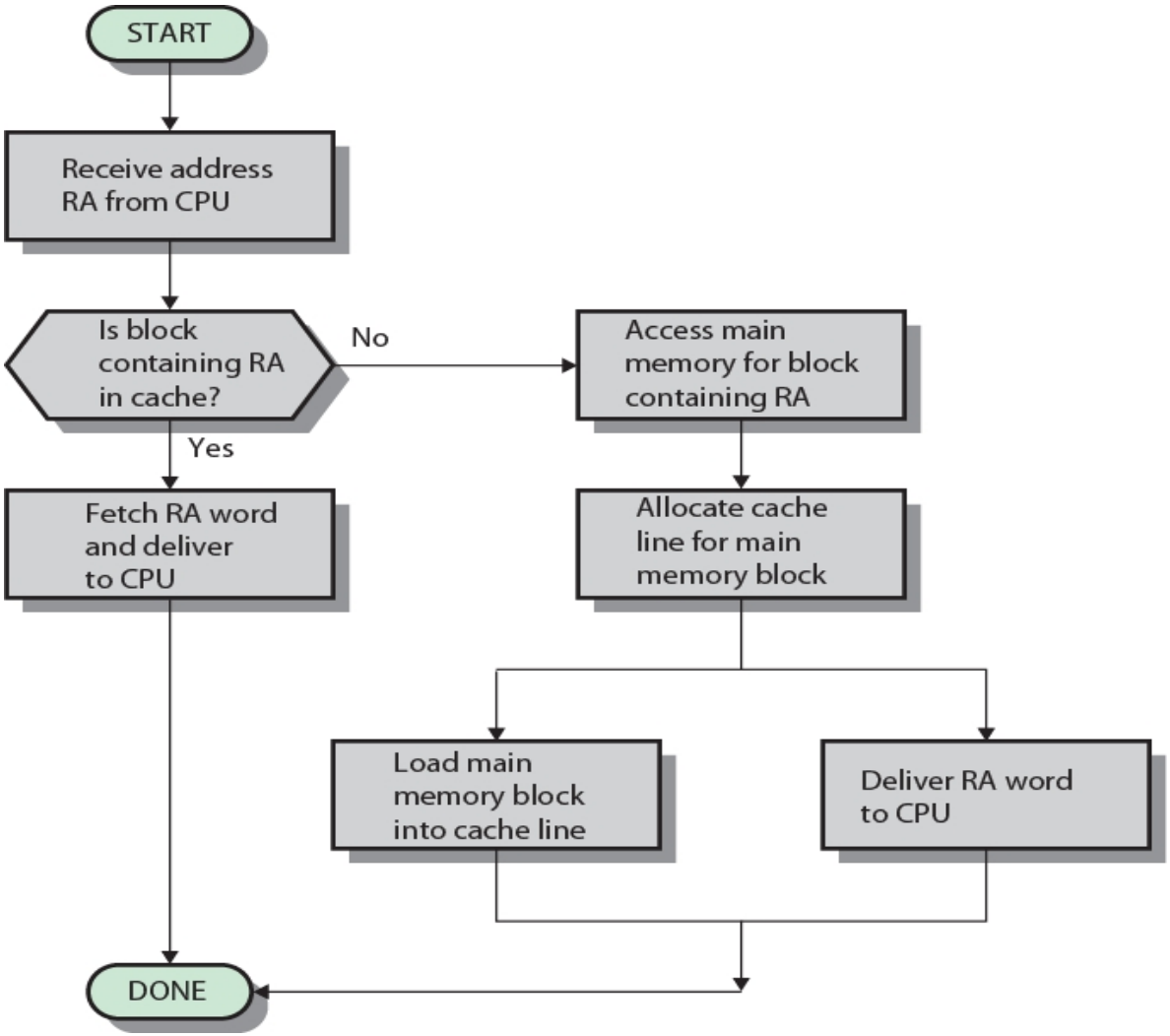
TAG NO. --	0	1	2	3	…	31
0	0	128	256	384	…	3098
1	1	129	257	385	…	3099
2	2	130	258	386	…	3100
3	3	131	259	387	…	…
….	…	…	….	…	…	….
126	126	254	382	…	…	4094
127	127	255	383	…	…	4095

4096/128 = 32 TAGS IN TOTAL

128= 2^7 = FRAMES = K WORDS

32 NO. OF BLOCKS

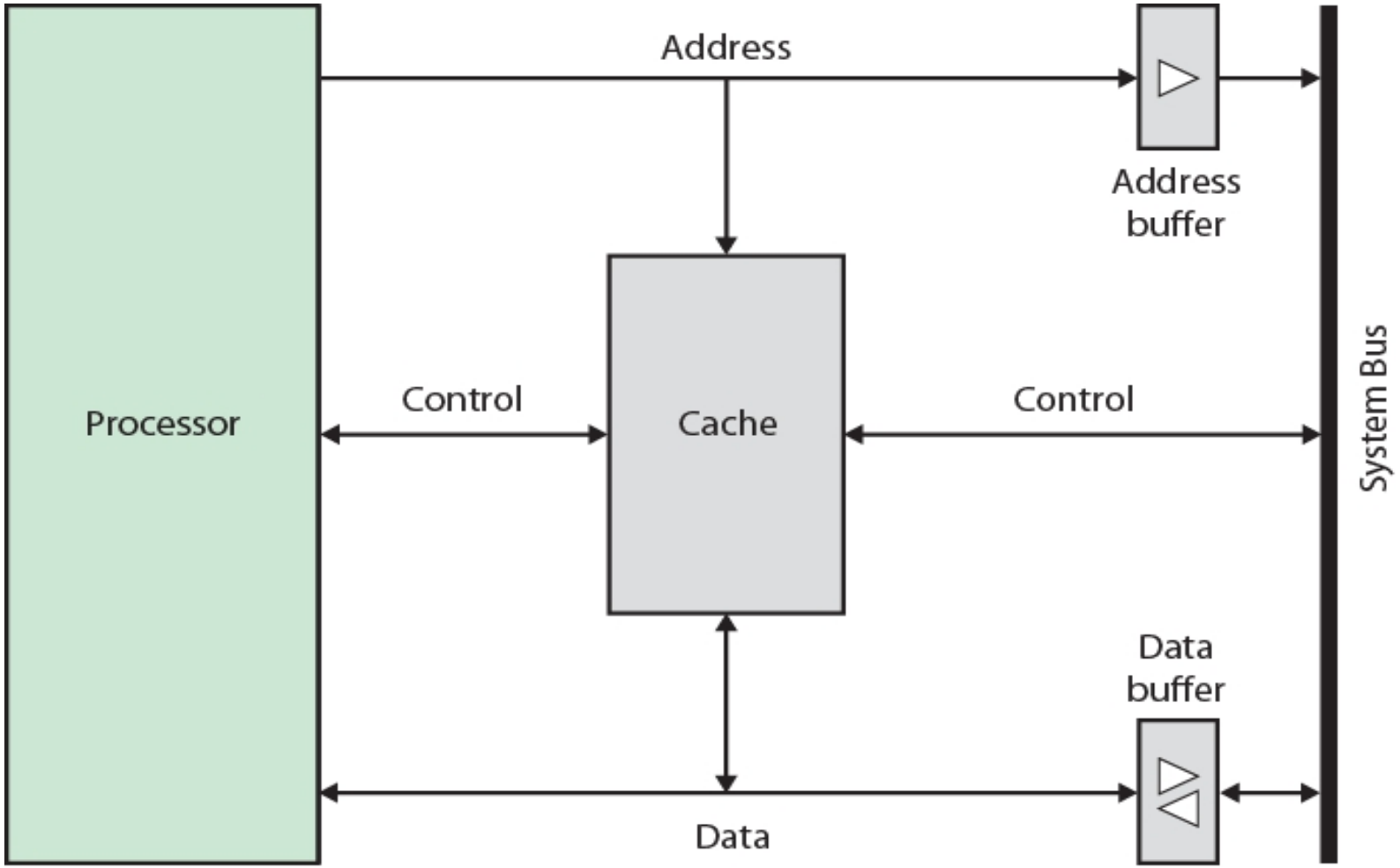
Cache Read Operation - Flowchart



Cache Operation – Overview

- CPU requests contents of memory location
- Check cache for this data
- If present, get from cache (fast)
- If not present, read required block from main memory to cache(Cache Organization)
- Then deliver from cache to CPU
- Cache includes tags to identify which block of main memory is in each cache slot

Typical Cache Organization



Cache Organization – Overview

- The cache connects to the processor via data, control and address lines
- The data and address lines attach to data and address buffers through system bus to reach main memory
- Hit occurs-communication only between processor and cache(disable data and address buffer)
- Miss occurs-the data are return through the data buffer to both cache and the processor(desired address is loaded onto the system bus)

THANK YOU