

# 15-213

*“The course that gives CMU its Zip!”*

## Cache Memories

### October 6, 2006

#### Topics

- Generic cache memory organization
- Direct mapped caches
- Set associative caches
- Impact of caches on performance
- The memory mountain

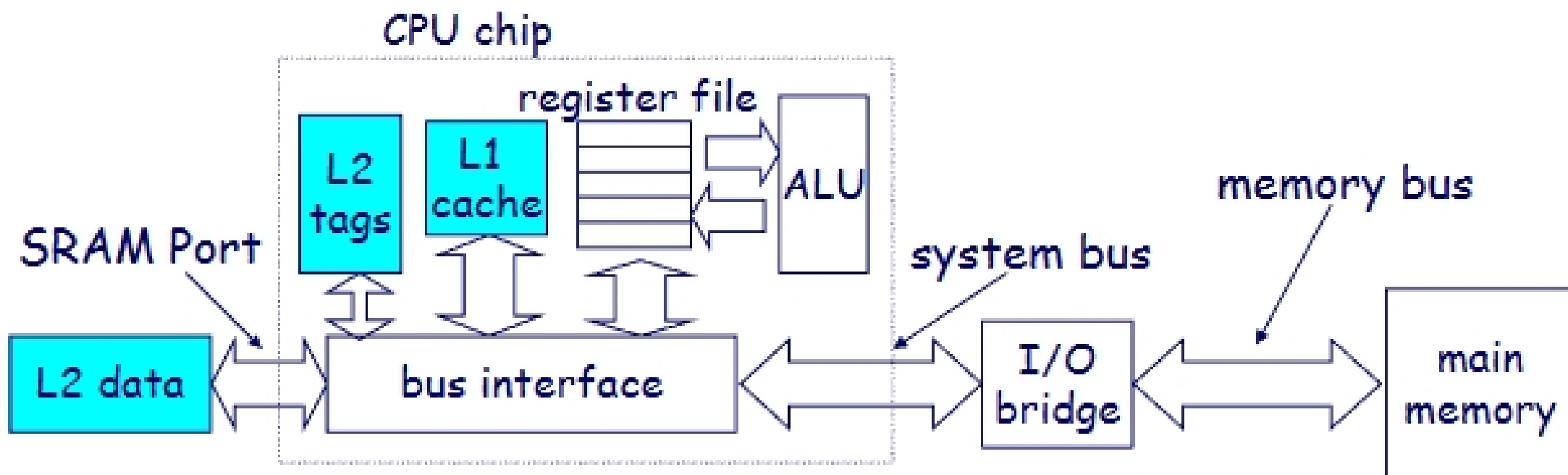
# Cache Memories

Cache memories are small, fast SRAM-based memories managed automatically in hardware.

- Hold frequently accessed blocks of main memory

CPU looks first for data in L1, then in L2, then in main memory.

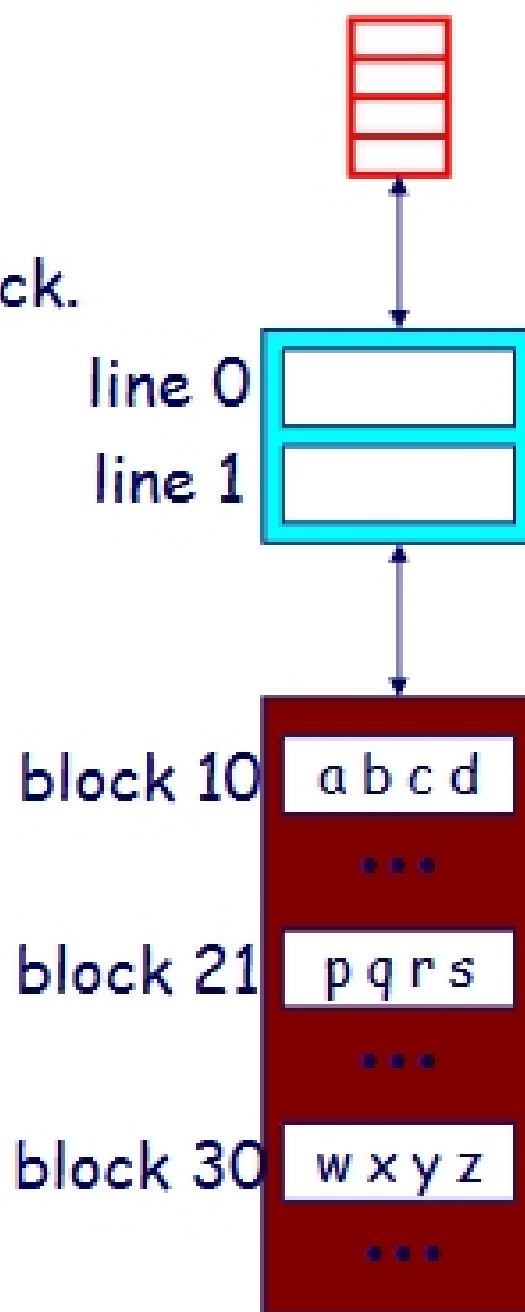
Typical system structure:



# Inserting an L1 Cache Between the CPU and Main Memory

The transfer unit between the CPU **register file** and the **cache** is a 4-byte block.

The transfer unit between the **cache** and **main memory** is a 4-word block (16 bytes).



The tiny, very fast CPU **register file** has room for four 4-byte words.

The small fast **L1 cache** has room for two 4-word blocks.

The big slow **main memory** has room for many 4-word blocks.