Lecture 22: Cache Hierarchies

• Today’s topics:
  - Cache access details
  - Examples
Accessing the Cache

Direct-mapped cache: each address maps to a unique location in cache

Byte address

101000

Offset

Data array

8-byte words

Sets

8 words: 3 index bits
The Tag Array

Direct-mapped cache: each address maps to a unique address

Byte address

Tag

Compare

101000

Tag array

Data array

8-byte words

8-byte words
Example Access Pattern

Assume that addresses are 8 bits long.
How many of the following address requests are hits/misses?
4, 7, 10, 13, 16, 68, 73, 78, 83, 88, 4, 7, 10...

Direct-mapped cache: each address maps to a unique address.
Increasing Line Size

A large cache line size $\rightarrow$ smaller tag array, fewer misses because of spatial locality

32-byte cache line size or block size

Tag array

Data array
Associativity

- Byte address
- Tag

Set associativity $\rightarrow$ fewer conflicts; wasted power because multiple data and tags are read

Tag array

Compare

10100000

Way-1

Way-2

Data array
Associativity

How many offset/index/tag bits if the cache has 64 sets, each set has 64 bytes, 4 ways
Example

• 32 KB 4-way set-associative data cache array with 32 byte line sizes

• How many sets?

• How many index bits, offset bits, tag bits?

• How large is the tag array?

\[
\text{Cache size} = \#\text{sets} \times \#\text{ways} \times \text{blocksize}
\]

\[
\text{Index bits} = \log_2(\text{sets})
\]

\[
\text{Offset bits} = \log_2(\text{blocksize})
\]

\[
\text{Addr width} = \text{tag} + \text{index} + \text{offset}
\]
Example 1

- 32 KB 4-way set-associative data cache array with 32 byte line sizes

\[
\text{cache size} = \#\text{sets} \times \#\text{ways} \times \text{block size}
\]

- How many sets? 256

- How many index bits, offset bits, tag bits?

\[
\begin{align*}
8 & \quad \log_2(\text{sets}) \\
5 & \quad \log_2(\text{blksize}) \\
19 & \quad \text{addrsizer-index-offset}
\end{align*}
\]

- How large is the tag array?

\[
\text{tag array size} = \#\text{sets} \times \#\text{ways} \times \text{tag size} = 19 \text{ Kb} = 2.375 \text{ KB}
\]
Example 2

- A pipeline has CPI 1 if all loads/stores are L1 cache hits
- 40% of all instructions are loads/stores
- 85% of all loads/stores hit in 1-cycle L1
- 50% of all (10-cycle) L2 accesses are misses
- Memory access takes 100 cycles
What is the CPI?
Example 2

- A pipeline has CPI 1 if all loads/stores are L1 cache hits
  - 40% of all instructions are loads/stores
  - 85% of all loads/stores hit in 1-cycle L1
  - 50% of all (10-cycle) L2 accesses are misses
Memory access takes 100 cycles
What is the CPI?

Start with 1000 instructions
1000 cycles (includes all 400 L1 accesses)
+ 400 (ld/st) x 15% x 10 cycles (the L2 accesses)
+ 400 x 15% x 50% x 100 cycles (the mem accesses)
= 4,600 cycles
CPI = 4.6
Example 3

Assume that addresses are 8 bits long.

How many of the following address requests are hits/misses?
4, 7, 10, 13, 16, 24, 36, 4, 48, 64, 4, 36, 64, 4

Way-1  
Way-2  
8-byte blocks

Data array
Assume that addresses are 8 bits long. How many of the following address requests are hits/misses?
4, 7, 10, 13, 16, 24, 36, 4, 48, 64, 4, 36, 64, 4
M H M H M M H M M H M M M

8-byte blocks

Way-1
Way-2

Data array