Review

1. Writes
   - Write-back
   - Write-through
   - Immediately

   Write-back
   - If evicted then written to lower level storage.

2. Writes
   - Write-allocate
   - No-write-allocate

Read Cache Hierarchy

L1 - d - i
L2 - unified
L3 - "
Cache Parameters

1. miss rate = \frac{\# misses}{\# references}

2. hit rate

3. hit time = t_{set} + t_{line} + t_{block\ offset}

4. miss penalty:
   - L1: 1-3 cycles
   - L2: 10 cycles
   - L3: 30-40 cycles
   - L4(MM): 100 cycles
   - HDD: 100,000 cycles
Cache attributes

C - cache size
B - block size
E - associativity
S - sets

1. Impact of cache size (C)

large C $\Rightarrow$ high hit rate.

... $\Rightarrow$ high hit time.

eg. L1 $\Rightarrow$ 1. low hit time.
2. low hit rate.
2. Impact of block size

- 64 bytes

large block size $\Rightarrow$ increase hit rate

(arrays, stride-1 ref., spatial locality)

For a given $C$, larger $B \Rightarrow$ less cache lines.
3. Impact of $E$

- Larger $E$ $\Rightarrow$ Hit rate $\uparrow$
- Conflict misses $\downarrow$

$L_1, L_2 \Rightarrow E = 8$

$L_3 \Rightarrow E = 16$

Diagram:

- $L_0$
- $L_1$
- $L_2$