**CS 219 – Assignment #7**

Purpose: Become familiar with MIPS cache implementation

Points: 100

**Reading/References:** Chapter 5 (5.1, 5.2, 5.3, 5.4)

**Assignment:**

Answer the following questions:

1. Define cache? Discuss whether the cache size should be small or large to overcome misses like compulsory, capacity, conflict? [10 pts]
2. Explain the Temporal and Spatial locality principle and give examples for each locality principle. [10 pts]
3. Explain memory hierarchy level and talk about their speeds. Give example for each level-memory type and where they are placed in a typical computer architecture. [10 pts]

[Hint: Refer to A Typical Memory Hierarchy on Ch5 slides on Canvas]

1. In general, cache access time is proportional to capacity. Assume that main memory accesses take 70 ns. The following table shows data for L1 caches attached to each of two processors, P1 and P2. [10 pts]

|  |  |  |  |
| --- | --- | --- | --- |
|  | L1 Size | L1 Miss rate | L1 Hit time |
| P1 | 1KB | 11.4% | 0.62ns |
| P2 | 2KB | 8.0% | 0.66ns |

1. If the L1 hit time determines the cycle times for P1 and P2, what are their respective clock rates? [Hint: clock rate= 1/L1 hit time]
2. What is the AMAT for each of P1 and P2?

AMAT= Hit time + Miss rate x Miss penalty

Table 1: Reference address to cache mapping

|  |  |  |  |
| --- | --- | --- | --- |
| **Address**  **Decimal** | **Address**  **Binary** | **Line ID/**  **Cache Set** | **Hit/Miss** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| **Hit Ratio** |  | | |
| **Comments** |  | | |

**Note: For Q5 and Q6, fill/use Table 1 for each part (a, b, c, d) in addition to the provided tables.**

Hints for direct-mapped cache (one-word blocks)

(**Hint 1:** Read section 5.3 and map the cache locations similar to Figure 5.8,

**Hint2:** Convert each address that is given in decimal to binary, then based on last 3 digits, map into one of the 8 cache entries.

**Hint 3:** For the first time, any address is considered as miss and if the address is repeated from the given list then it is considered as hit. To consider hit the previous entry should be the same in that cache. For example: 1, 9, 1, 2, 1 are the addresses, in binary 0001, 1001, 0001, 0010, 0001 the addresses (1, 9, 1, 1) will be in 001 cache entry and address 2 in 010 cache entry. In this (1, 9, 1, 2, 1) for the first time 1-miss, then 9-miss, 1-miss (1 is miss here because previous entry is 9), 2-miss, 1-hit(1 is hit here because previous entry in that cache list is 1)

**Hint 4:** To find hit ratio: number of hits divided by total entries, in this example it is 1/5 or 20%.)

Diagram

Description automatically generated

1. Here is a series of address references given as words addresses: [15 pts]

**5, 18, 1, 1, 2, 3, 11, 10, 21, 18, 17, 10, 12, 3, 11, 7, 10, 22, 4, 22.**

* 1. Assuming a direct-mapped cache with 8 one-word blocks that is initially empty, label each reference in the list as a hit or a miss and show the contents of the cache (including previous, over-written values). You do not need to show the tag field. When done, include the hit ratio. If you have entry in cache, then valid bit is 1 else zero.

|  |  |  |
| --- | --- | --- |
| **Cache Set** | **valid** | **Address** |
| **000** |  |  |
| **001** |  |  |
| **010** |  |  |
| **011** |  |  |
| **100** |  |  |
| **101** |  |  |
| **110** |  |  |
| **111** |  |  |

1. Assuming a direct-mapped cache with 16 one-word blocks that is initially empty, label each reference in the list as a hit or a miss and show the contents of the cache (including previous, over-written values). You do not need to show the tag field. When done, include the hit ratio. [15 pts]

|  |  |  |
| --- | --- | --- |
| **Cache Set** | **Valid** | **Address** |
| **0000** |  |  |
| **0001** |  |  |
| **0010** |  |  |
| **0011** |  |  |
| **0100** |  |  |
| **0101** |  |  |
| **0110** |  |  |
| **0111** |  |  |
| **1000** |  |  |
| **1001** |  |  |
| **1010** |  |  |
| **1011** |  |  |
| **1100** |  |  |
| **1101** |  |  |
| **1110** |  |  |
| **1111** |  |  |

1. Show the hits and misses and cache contents (including previous, overwritten values) for a direct-mapped cache with four-word blocks and a total size of 8 words. You do not need to show the tag field. When done, include the hit ratio. [10 pts]

|  |  |  |
| --- | --- | --- |
| **Cache Set** | **valid** | **Address** |
| **00/ 10** |  |  |
| **01/11** |  |  |

1. Show the hits and misses and cache contents (including previous, overwritten values) for a direct-mapped cache with four-word blocks and a total size of 16 words. You do not need to show the tag field. When done, include the hit ratio. [10 pts]

|  |  |  |
| --- | --- | --- |
| **Cache Set** | **valid** | **Address** |
| **00** |  |  |
| **01** |  |  |
| **10** |  |  |
| **11** |  |  |

1. Compare the mapping techniques used above (a, b, c, and d) and write your observations that includes hit ratio and the reasons why the hit ratio is more compared to others.

[Hint: Improvement due to more one-word blocks; Improvement due to multiple words in each block, etc.] [10 pts]