Name: Yi Yao

ID: 011743488

**Part 2: Find CPI from equation**

Given an L1 miss penalty of 5 cycles, L2 miss penalty of 50 cycles, and one cycle cache hit/instruction execution

CPI = 1 + [ 5 \* (IL1misses + DL1misses) + 50 \* L2misses ] / Total instructions

1. **From the “401.bzip2”**

IL1misses: 656



DL1misses: 617680



L2misses: 283838



So, **CPI = 1 + [ 5 \* (656 + 617680) + 50 \* 283838 ] / 100000000 = 1.1728358**

1. **From the “429.mcf”**

IL1misses: 486



DL1misses: 97288



L2misses: 92093



So, **CPI = 1 + [ 5 \* (486 + 97288) + 50 \* 92093 ] / 100000000 = 1.0509352**

1. **From the “456.hmmer”**

IL1misses: 1401



DL1misses: 32893



L2misses: 6071



So, **CPI = 1 + [ 5 \* (1401 + 32893) + 50 \* 6071 ] / 100000000 = 1.0047502**

1. **From the “458.sjeng”**

IL1misses: 2183



DL1misses: 8298580



L2misses: 8297070



So, **CPI = 1 + [ 5 \* (2183 + 8298580) + 50 \* 8297070 ] / 100000000 = 5.56357315**

1. **From the “470.lbm”**

IL1misses: 481



DL1misses: 1533680



L2misses: 1534158



So, **CPI = 1 + [ 5 \* (481 + 1533680) + 50 \* 1534158 ] / 100000000 = 1.84378705**

**Part 3: Find optimal configuration**

This section explores the impact of different factors, which include associativity, block size, and size allocation for L1 instruction cache and L1 data cache. Table 1 shows the Domains of Cache Parameters and 512KB available for L1 cache (for L1 d-cache and i-cache together) and 4MB available for L2 cache.

**Table 1:** Domains of Cache Parameters.

|  |  |
| --- | --- |
| Parameter | Domains |
| L1 d-cache | 128KB, 256KB |
| L1 i-cache | 64KB, 128KB, 256KB |
| L2 cache | 1MB, 4MB |
| associativity | 1, 2, 4, 8 |
| block size | 64bytes, 128bytes |

**Block size**

When a data block is retrieved and put into the Cache, the required words and adjacent words are all taken out. When the block size changes from small to large, the hit rate will increase. This is because of the principle of locality: data near the quoted word has a high probability of being quoted later. When the block size increases, more useful data is loaded into the Cache. However, when the block becomes quite large and the probability of using the newly retrieved information becomes less than the probability of reusing the replaced information, the hit rate begins to decrease.

**Cache capacity**

When designing a Cache, it is usually hoped that the capacity of the Cache is small enough to save cost, power consumption and area(L1 cache cost a high price). At the same time, it is hoped that the capacity of the Cache is large enough so that the average time of the entire storage system is close to the access time of the entire Cache.

**Associativity**

The choice of the method of association between main memory and cache is the result of a compromise between multiple factors. For example, when a certain replacement strategy allows a main memory data block to have 10 cache lines corresponding to it, the processor must find a way to find these 10 locations to see if it hits one of them. The more search, the greater the power consumption, the larger the chip area, and the longer the time; from another point of view, the more associativity can have lower misses. It takes less time to wait for data to be fetched from low-speed main memory.

In the design of Cache, there are many factors that affect Cache performance, such as design elements such as Cache capacity and associativity. The capacity of the Cache is a very important parameter, and its size directly affects the performance of the Cache. Under the same conditions, the larger the cache capacity, the higher the hit rate. Therefore, if you want to improve the performance of the Cache, you can increase the capacity of the Cache. However, the capacity of the Cache cannot be increased indefinitely, because increasing the capacity of the Cache will cause the Cache area to be too large, and the power consumption of the Cache will increase with the increase in capacity.

However, the improvement of Cache capacity and the improvement of Cache performance are not linear. When the Cache capacity is increased to a certain level, the growth rate of Cache performance will decrease. Therefore, the method of excessively increasing the Cache capacity is not desirable. For group-associated Caches, increasing the degree of the association will also increase the hit rate of the Cache, but after increasing the degree of association, the hit time of the Cache will also increase. Therefore, increasing the association degree of the Cache, and the performance of the Cache may not promote. Increasing the block size of the Cache, the more data is read from the memory when the Cache is missing, but as the block size of the Cache increases, since the Cache capacity is constant, the number of Cache groups will be reduced, thereby increasing the conflict and missing.

1. **401.bzip2**

When IL1 cache\_Associativity 2-way; DL1 cache\_Associativity 4-way; L2 cache\_Associativity 8-way; L2 1MB; Block Size=64bytes; IL1 cache 256KB or DL1 cache 256KB

Fig.1 shows that as the cache capacity increasing the CPI of IL 1 declines nonlinear. However, as the cache capacity increasing the CPI of IL1 declines almost linear.

When IL1 cache 256KB; DL1 cache256KB; L2 1MB; IL1 cache\_Associativity 8-way; DL1 cache\_Associativity 8-way

Fig.2 reveals that as L2 Cache associativity increase, CPI decreases slowly. Simultaneously, the CPI decreased as the capacity of block size increase.

When IL1 cache 256KB; DL1 cache256KB; L2 1MB; Block Size=64bytes; L2 cache\_Associativity 8-way; DL1 cache\_Associativity 2-way

Fig.3 states that as an increase in the number of IL 1 cache associativity, the CPI declines first then not change.

When IL1 cache 256KB; DL1 cache256KB; L2 1MB; Block Size=64bytes; L2 cache\_Associativity 8-way; IL1 cache\_Associativity 2-way

Fig.4 shows that as an increase in the number of DL1 cache associativity, the CPI declines first then increased.

When IL1 cache 256KB; DL1 cache256KB; IL1 cache\_Associativity 8-way; DL1 cache\_Associativity 8-way; Block Size=128bytes

Fig.5 shows that as the capacity of L2 increased, the CPI will decline. However as the number of L2 cache associativity increases, CPI almost unchanged.

1. **429.mcf**

When IL1 cache\_Associativity 2-way; DL1 cache\_Associativity 4-way; L2 cache\_Associativity 8-way; L2 1MB; Block Size=64bytes; IL1 cache 256KB or DL1 cache 256KB

Fig.6 reveals that as the capacity of IL 1 cache increased the CPI declines dramatically then unchanged. However, as the capacity of the DL1 cache increased the CPI almost unchanged.

When IL1 cache 256KB; DL1 cache256KB; L2 1MB; IL1 cache\_Associativity 8-way; DL1 cache\_Associativity 8-way

Fig.7 reveals that as L2 Cache associativity increase, CPI almost unchanged. At the same time, the CPI decreased as the capacity of block size increase.

When IL1 cache 256KB; DL1 cache256KB; L2 1MB; Block Size=64bytes; L2 cache\_Associativity 8-way; DL1 cache\_Associativity 2-way

Fig.8 states that as an increase in the number of IL 1 cache associativity, the CPI almost unchanged.

When IL1 cache 256KB; DL1 cache256KB; L2 1MB; Block Size=64bytes; L2 cache\_Associativity 8-way; IL1 cache\_Associativity 2-way

Fig.9 states that as an increase in the number of DL 1 cache associativity, the CPI almost unchanged.

When IL1 cache 256KB; DL1 cache256KB; IL1 cache\_Associativity 8-way; DL1 cache\_Associativity 8-way; Block Size=128bytes

Fig.10 reveals that as the capacity of L2 increased, the CPI will decline. When the L2 cache capacity is 1MB, the CPI decreases slowly with the associativity increase. But when the L2 cache capacity is 4MB, the CPI first decreases slowly with the associativity increase then rise up.

1. **456.hmmer**

When IL1 cache\_Associativity 2-way; DL1 cache\_Associativity 4-way; L2 cache\_Associativity 4-way; L2 1MB; Block Size=64bytes; IL1 cache 256KB or DL1 cache 256KB

Fig.11 reveals that as the capacity of DL 1 cache increased the CPI declines dramatically. However, as the capacity of the IL1 cache increased the CPI almost unchanged.

When IL1 cache 256KB; DL1 cache256KB; L2 1MB; IL1 cache\_Associativity 8-way; DL1 cache\_Associativity 8-way

Fig.12 shows that as L2 Cache associativity increase, CPI almost unchanged. At the same time, the CPI decreased as the capacity of block size increase.

When IL1 cache 256KB; DL1 cache256KB; L2 1MB; Block Size=64bytes; L2 cache\_Associativity 4-way; DL1 cache\_Associativity 2-way

Fig.13 shows that the CPI will not change as the number of associativities greater than 2.

When IL1 cache 256KB; DL1 cache256KB; L2 1MB; Block Size=64bytes; L2 cache\_Associativity 4-way; IL1 cache\_Associativity 2-way

Fig.14 shows that the CPI will not change as the number of associativities greater than 2.

When IL1 cache 256KB; DL1 cache256KB; IL1 cache\_Associativity 8-way; DL1 cache\_Associativity 8-way; Block Size=128bytes

Fig.15 reveals that as the capacity of the L2 cache increased the CPI unchanged. Simultaneously, as the number of L2 associativity increased the CPI unchanged.

1. **458.sjeng**

When IL1 cache\_Associativity 2-way; DL1 cache\_Associativity 4-way; L2 cache\_Associativity 8-way; L2 1MB; Block Size=64bytes; IL1 cache 256KB or DL1 cache 256KB

Fig.16 shows that as the capacity of IL1 cache increased the CPI declined. However, as the capacity of the DL1 cache increased the CPI declined not significantly.

When IL1 cache 256KB; DL1 cache256KB; L2 1MB; IL1 cache\_Associativity 8-way; DL1 cache\_Associativity 8-way

Fig.17 states that as the number of L2 Cache associativity increases, CPI almost unchanged. At the same time, the CPI decreased significantly as the capacity of block size increase.

When IL1 cache 256KB; DL1 cache256KB; L2 1MB; Block Size=64bytes; L2 cache\_Associativity 8-way; DL1 cache\_Associativity 2-way

Fig.18 states that as the number of IL1 cache associativity increases, the CPI will not change as the number of associativities greater than 2.

When IL1 cache 256KB; DL1 cache256KB; L2 1MB; Block Size=64bytes; L2 cache\_Associativity 8-way; IL1 cache\_Associativity 2-way

Fig.19 shows that as the number of DL1 associativity increases, the CPI changed a little bit.

When IL1 cache 256KB; DL1 cache256KB; IL1 cache\_Associativity 8-way; DL1 cache\_Associativity 8-way; Block Size=128bytes

Fig.20 reveals that as the capacity of the L2 cache increased the CPI almost unchanged. Simultaneously, as the number of L2 associativity increased the CPI unchanged.

1. **470.lbm**

When IL1 cache\_Associativity 2-way; DL1 cache\_Associativity 4-way; L2 cache\_Associativity 4-way; L2 1MB; Block Size=64bytes; IL1 cache 256KB or DL1 cache 256KB

Fig.21 shows that as the capacity of IL1 cache increased the CPI declined. While the capacity of IL1 cache larger than 128KB then CPI unchanged. However, as the capacity of the DL1 cache increased the CPI unchanged.

When IL1 cache 128KB; DL1 cache256KB; L2 1MB; IL1 cache\_Associativity 8-way; DL1 cache\_Associativity 8-way

Fig.22 states that as the number of L2 Cache associativity increases, the CPI unchanged. At the same time, the CPI decreased as the capacity of block size increase.

When IL1 cache 128KB; DL1 cache256KB; L2 1MB; Block Size=64bytes; L2 cache\_Associativity 4-way; DL1 cache\_Associativity 2-way

Fig.23 shows that the CPI unchanged as the number of IL1 cache associativity greater than 2.

When IL1 cache 128KB; DL1 cache256KB; L2 1MB; Block Size=64bytes; L2 cache\_Associativity 4-way; IL1 cache\_Associativity 2-way

Fig.24 shows that the CPI unchanged as the number of DL1 cache associativity greater than 2.

When IL1 cache 128KB; DL1 cache256KB; IL1 cache\_Associativity 8-way; DL1 cache\_Associativity 8-way; Block Size=128bytes

Fig.25 reveals that as the capacity of the L2 cache increased the CPI unchanged. Simultaneously, as the number of L2 associativity increased the CPI unchanged.

**From the simulation of the data, the following configurations are the optimum for each benchmark.**

**401.bzip2**

IL1 cache 128KB

DL1 cache 256KB

IL1 cache\_Associativity 2-way

DL1 cache\_Associativity 4-way

L2 4MB

L2 cache\_Associativity 8-way

Block size 128bytes

Optimum CPI = 1.0741179

**429.mcf**

IL1 cache 64KB

DL1 cache 256KB

IL1 cache\_Associativity 4-way

DL1 cache\_Associativity 4-way

L2 4MB

L2 cache\_Associativity 4-way

Block size 128bytes

Optimum CPI = 1.02355395

**456.hmmer**

IL1 cache 256KB

DL1 cache 256KB

IL1 cache\_Associativity 4-way

DL1 cache\_Associativity 8-way

L2 1MB

L2 cache\_Associativity 1-way

Block size 128bytes

Optimum CPI = 1.00177965

**458.sjeng**

IL1 cache 256KB

DL1 cache 256KB

IL1 cache\_Associativity 4-way

DL1 cache\_Associativity 8-way

L2 4MB

L2 cache\_Associativity 2-way

Block Size 128bytes

Optimum CPI = 3.28306745

**470.lbm**

IL1 cache 128KB

DL1 cache 128KB

IL1 cache\_Associativity 2-way

DL1 cache\_Associativity 4-way

L2 1MB

L2 cache\_Associativity 1-way

Block size 128bytes

Optimum CPI = 1.42194685

**Part 4: Define cost function**

Assumption:

L1 d-cache capacity: 4/KB

L1 i-cache capacity: 4/KB

L2 cache capacity: 2/MB

Associativity 1-way: 10

Associativity 2-way: 20

Associativity 4-way: 40

Associativity 8-way: 80

**Cost** = L1 i-cache size \* 4 + L1 d-cache size \* 4 + L2 cache size \* 2 + Associativity of IL1 cache + Associativity of DL1 cache + Associativity of L2

**Part 5: Evaluation function**

Total Cost = CPI \* Cost

The best configuration:

IL1 cache 64KB

DL1 cache 256KB

IL1 cache\_Associativity 2-way

DL1 cache\_Associativity 2-way

L2 4MB

L2 cache\_Associativity 2-way

Block Size 128bytes

CPI \* Cost = 1448.619236

The best configuration:

IL1 cache 32KB

DL1 cache 256KB

IL1 cache\_Associativity 2-way

DL1 cache\_Associativity 2-way

L2 1MB

L2 cache\_Associativity 2-way

Block Size 128bytes

CPI \* Cost = 1247.424394

The best configuration:

IL1 cache 32KB

DL1 cache 256KB

IL1 cache\_Associativity 2-way

DL1 cache\_Associativity 2-way

L2 1MB

L2 cache\_Associativity 4-way

Block Size 64bytes

CPI \* Cost = 1208.074095

The best configuration:

IL1 cache 64KB

DL1 cache 256KB

IL1 cache\_Associativity 2-way

DL1 cache\_Associativity 2-way

L2 1MB

L2 cache\_Associativity 2-way

Block Size 128bytes

CPI \* Cost = 4408.261118

The best configuration:

IL1 cache 128KB

DL1 cache 128KB

IL1 cache\_Associativity 2-way

DL1 cache\_Associativity 2-way

L2 1MB

L2 cache\_Associativity 2-way

Block Size 128bytes

CPI \* Cost = 1530.014918

**Reference:**

Sun, J. (2016). Cache design space exploration for android applications (Master's thesis, Southeast university, 2016) (pp. 7-13). Nanjing: Wanfang data.