# Advanced Systems Lab

**Homework 4**

Albert Cerfeda (20-980-272)

## Contents



ETH zürich
9.04.2024
Eidgenössische Technische Hochschule Zürich
D-INFK Department
Switzerland

## 4.1 Stride access (30 pts)

### a) Direct-mapped cache, $s = 1$ $n = 8$

The cache is direct-mapped, has blocks of size 32 bytes and a total capacity of 1 KiB. This means that there are $\frac{2^{10}}{32} = 32$ cache blocks. Each cache block/line contains 32 bytes (8 doubles).

Since $n = 8$, we can infer how the cache is able to hold both matrix $A$ and $O$ in their entirety.

The access pattern of matrix $O$ is the following: Each element is accessed sequentially from row 0 to row $n - 3 = 5$. Therefore in total there are 6 compulsory cache misses, one for each row of $O$.

Regarding the access pattern of $A$ instead, at each iteration, the corresponding elements at row $i - 1$ and $i + 1$ are accessed. Similar to before we only have 8 compulsory cache misses, one for each row of $A$.

We have a total of 192 memory accesses. 15 of which are compulsory cache misses, after which all the rows of $A$ and $O$ are in the cache. Therefore the cache miss rate is $\frac{15}{192} = 0.078125\%$.

### b) Direct-mapped cache, $s = 2$ $n = 16$

The cache structure is analagous as before. The main difference is that since matrix $A$ and $O$ are $16 \times 16$, the cache can only store exactly one of them at a time. In other words, elements from $A$ and $O$ with the same coordinates map to the same cache location.
By increasing the stride to 2, not every element of $A$ is accessed. In fact, the columns are covered in an alternating pattern, resulting in only half of the elements being accessed. Matrix $O$ is still accessed in a sequential manner.

In the first 4 iterations, the cache gets repeatedly evicted because the first 4 elements of $O$ share the same cache block as the first elements of $A$. After the initial 4 iterations, this light form of thrashing stops, resulting only in compulsory cache misses.

We have 112 iterations and a total of 448 memory accesses, 66 of which are cache misses, therefore the cache miss rate is $\frac{66}{448} = 0.14732143\%$.

### c) 2-way associatice cache, $s = 2$ $n = 16$

Since the cache is 2-way associative, it has 16 sets with 2 cache lines each. Each row of $A$ and $O$ is 16 doubles, meaning that two cache blocks are needed to store one full row.

## 4.2 Cache mechanics (20 pts)

### a) Access patterns and state of direct-mapped cache

i. Cache at line 18
**- i = 0**
x: **MMM**
y: **MMM**

| Set | Block 1 | Block 2 |
|---|---|---|
| 0 | $y_2.a$ | $y_2.b$ |
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | $x_2.a$ | $x_2.b$ |
| 5 | | |
| 6 | | |
| 7 | | |

**- i = 1**

x: **HHH**
y: **MHH**

| Set | Block 1 | Block 2 |
|---|---|---|
| 0 | $y_2.a$ | $y_2.b$ |
| 1 | | |
| 2 | $y_3.a$ | $y_3.b$ |
| 3 | | |
| 4 | $x_2.a$ | $x_2.b$ |
| 5 | | |
| 6 | | |
| 7 | | |

ii. Cache at line 30
**- i = 0**
x: **MMM**
y: **MMM**

| Set | Block 1 | Block 2 |
|---|---|---|
| 0 | $y_2.a$ | $y_2.b$ |
| 1 | $y_6.c$ | $y_6.d$ |
| 2 | $y_3.a$ | $y_3.b$ |
| 3 | $x_5.c$ | $x_5.d$ |
| 4 | $x_2.a$ | $x_2.b$ |
| 5 | $y_0.c$ | $y_0.d$ |
| 6 | | |
| 7 | | |

**- i = 1**
x: **MMM**
y: **MHM**

| Set | Block 1 | Block 2 |
|---|---|---|
| 0 | $y_2.a$ | $y_2.b$ |
| 1 | $y_6.c$ | $y_6.d$ |
| 2 | $y_3.a$ | $y_3.b$ |
| 3 | $y_3.c$ | $y_3.d$ |
| 4 | $x_2.a$ | $x_2.b$ |
| 5 | $x_2.c$ | $x_2.d$ |
| 6 | | |
| 7 | | |

### b) Access patterns and state of 2-way associative cache

i. Cache at line 18
**- i = 0**
x: **MMM**
y: **MMM**

| Set | Line 1 - Block 1 | Line 1 - Block 2 | Line 2 - Block 1 | Line 2 - Block 2 |
|---|---|---|---|---|
| 0 | $x_2.a$ | $x_2.b$ | $y_2.a$ | $y_2.b$ |
| 1 | | | | |
| 2 | | | | |
| 3 | | | | |

**- i = 1**
x: **HHH**
y: **MHH**

| Set | Line 1 - Block 1 | Line 1 - Block 2 | Line 2 - Block 1 | Line 2 - Block 2 |
|---|---|---|---|---|
| 0 | $x_2.a$ | $x_2.b$ | $y_2.a$ | $y_2.b$ |
| 1 | | | | |
| 2 | $y_3.a$ | $y_3.b$ | | |
| 3 | | | | |

ii. Cache at line 30
**- i = 0**
x: **MMM**
y: **MMM**

| Set | Line 1 - Block 1 | Line 1 - Block 2 | Line 2 - Block 1 | Line 2 - Block 2 |
|---|---|---|---|---|
| 0 | $x_2.a$ | $x_2.b$ | $y_2.a$ | $y_2.b$ |
| 1 | $y_6.c$ | $y_6.d$ | $y_0.c$ | $y_0.d$ |
| 2 | $y_3.a$ | $y_3.b$ | | |
| 3 | $x_5.c$ | $x_5.d$ | | |

**- i = 1**
x: **MMH**
y: **MMH**

| Set | Line 1 - Block 1 | Line 1 - Block 2 | Line 2 - Block 1 | Line 2 - Block 2 |
|---|---|---|---|---|
| 0 | $x_2.a$ | $x_2.b$ | $y_2.a$ | $y_2.b$ |
| 1 | $x_2.c$ | $x_2.d$ | $y_6.c$ | $y_6.d$ |
| 2 | $y_3.a$ | $y_3.b$ | | |
| 3 | $x_5.c$ | $x_5.d$ | $y_3.c$ | $y_3.d$ |

## 4.3 Rooflines (40 pts)

## 4.4 Cache miss analysis (25 pts)