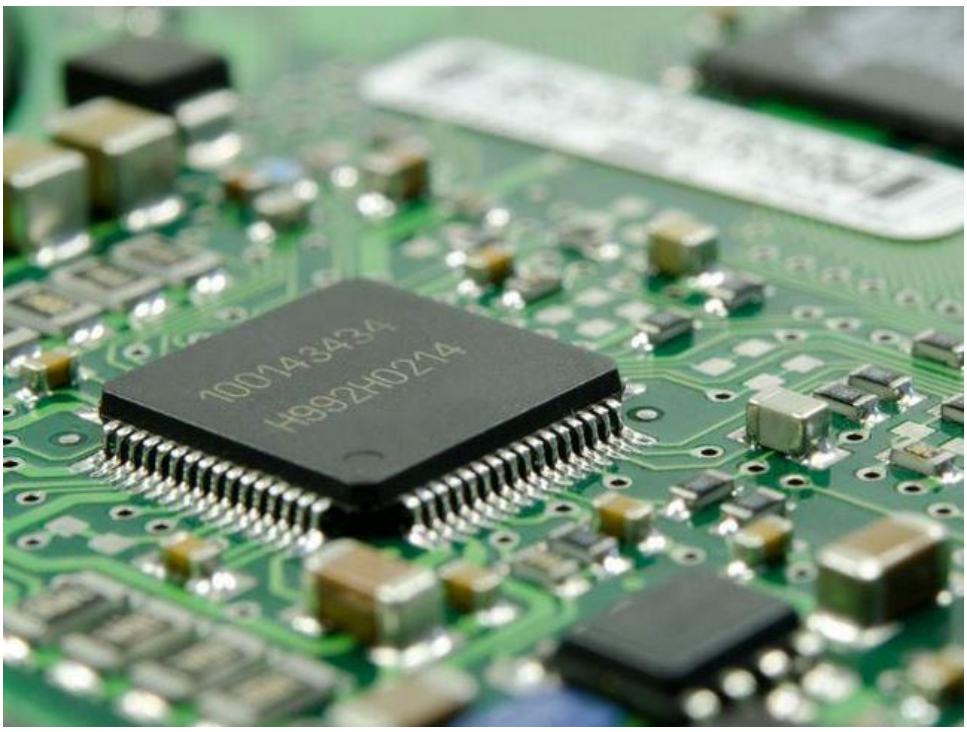
جامعة الحمدانية
كلية التربية

Lecture 15

Computer Architecture



Prepared By :

Firas Abdulrahman Yosif

Replacement Algorithms & Write Policy

Replacement algorithms are used in order to attain optimized usage of cache. When cache is full, then replacement policies decide which piece of data is replaced in order to make space for new data that is currently being used. An efficient algorithm is that which can take less time and number of cache misses are low and also balancing cost. Following are some of the algorithms.

يتم استخدام خوارزميات الاستبدال من أجل تحقيق الاستخدام الأمثل لذاكرة التخزين المؤقت الـ cache memory. عندما تكون ذاكرة الـ cache ممتلئة ، تحدد سياسات الاستبدال (replacement policies) أي block داخل الـ cache سوف يلغى ويتم استبداله بالـ block القادم من الـ main memory وذلك لتوفير مساحة للبيانات الجديدة المستخدمة حاليًا. الخوارزمية الفعالة هي تلك التي يمكن أن تستغرق وقتًا أقل وكذلك التكلفة.

عندما ترسل الـذاكرة الرئيسية block جديد الى الـ cache memory وقبل ان يدخل الى الـ cache يجب ان يستبدل بأحد الـ blocks الموجودة داخل الـ cache . بالنسبة لطريقة الـ direct mapping لايحتاج الى استخدام خوارزمية الاستبدال replacement algorithm لان الـ block القادم من الـ main memory سوف يدخل الـ cache في slot واحد محدد.

بينما باستخدام طريقة الـset associative mapping & associative mapping فأن خوارزمية الاستبدال تكون ضرورية لان الـ block القادم من الـ main memory يمكن ان يدخل في اي slot داخل الـcache memory .

❖ **Types of replacement algorithm:**

1. **LRU** (Least Recently Used): This algorithm discards the least recently used item from the cache in order to make space for the new data item. In order to achieve this, history of all data items that is which data item is

used when, is kept. Variants of LRU are the most popular among all other algorithms. Key advantage of this policy is its simple implementation, time and space overhead is constant.

خوارزمية LRU (الأقل استخدامًا مؤخرًا):

خوارزمية LRU وهي اختصار لـ (Least Recently Used) تستبدل هذه الخوارزمية الـ block الأقل استخدامًا مؤخرًا الموجود في الـ set داخل ذاكرة الـ cache الذي مضى على بقائه فترة طويلة ولم يستخدم وذلك لتوفير مساحة للبيانات الجديدة القادمة من الذاكرة الرئيسية.

تعد خوارزمية LRU هي الأكثر شيوعًا بين جميع الخوارزميات الأخرى. الميزة الرئيسية لهذه الخوارزمية هي التنفيذ البسيط ، والوقت والمساحة الزائدة تكون ثابتة وهي عامل رئيسي في هذه الخوارزمية.

اي ان الـ block القادم من الـ main memory يستبدل مع الـ slot الأقل من استخدم مؤخرا داخل الـ cache. ان نسبة الـ Hit ratio في ايجاد المعلومة تكون اكثر من خوارزمية الـ FIFO واقل من خوارزمية الـ LFU لان الـ block استخدمناه ولكن منذ فترة زمنية قصيرة.

2. **FIFO** (first in first out): The first in first out algorithm removes the page that has not been used for a long time. It treats the pages as a circular buffer, and pages are removed in a round robin fashion. It causes early page fault.

خوارزمية FIFO هي اختصار لـ (first in first out) تنص هذه الخوارزمية على ان الـ block القادم من الـ main memory يستبدل مع اقدم slot موجود في الـ cache الذي قضى فترة طويلة في الـ cache . نسبة الـ Hit ratio في هذه الخوارزمية قليلة لانه ممكن ان نحتاج هذا الـ block الذي الغي مؤخرا.

3. **LFU** Algorithm (Least Frequently Used): This algorithm counts how often data items have been used. The data items which are used less are deleted from the cache first. If all objects have same frequency, then this algorithm randomly discards any data item.

خوارزمية LFU (الأقل استخدامًا): خوارزمية LFU هي اختصار (Least Frequently Used) اي تمثل الـ slot الاقل عدد مرات استخداما داخل الـ cache اي تحسب هذه الخوارزمية عدد مرات استخدام البيانات الموجودة في ذلك الـ slot. حيث يتم حذف عناصر البيانات من ذاكرة الـ cache التي يتم استخدامها بشكل أقل من غيرها. هنا نسبة الـ Hit ratio تكون عالية وذلك لانه يبقى الـ slot الاكثر استخداما ويلغى الـ slot الاقل استخداما.

حيث تنص هذه الخوارزمية على ان الـ block القادم من الـ main memory يستبدل مع الـ slot الذي عدد مرات استخدامه قليل في الـcache memory .

ملاحظة: خوارزمية الاستبدال (replacement algorithm) تطبق على طريقة الـ associative cache mapping والـ set – associative cache mapping ولكن هذه خوارزمية الاستبدال لاتطبق على طريقة direct cache mapping .

❖ Write Policy in cache memory

A cache's *write policy* is the behavior of a cache while performing a write operation. A cache's write policy plays a central part in all the variety of different characteristics exposed by the cache.

سياسة الكتابة في الذاكرة cache هي سلوك التي تتخذه ذاكرة الـ cache أثناء إجراء عملية كتابة عليها. تلعب سياسة الكتابة في ذاكرة التخزين المؤقت الـ cache دورًا مركزيًا ومهما في مجموعة متنوعة من الخصائص المختلفة التي تعرضها ذاكرة الـ cache.

Types of write policy in cache memory:

1. **Write-through:**

   write-through provides us the best consistency, it does not help us with write operation latency – the cache returns a response to the client only after the backing store is updated.

   We may take advantage of our fast cache storage to streamline this as well. To do this, we would have to return the response before updating the backing store.

Write through هذه الالية تنص على ان الـ CPU يقوم بتحديث المعلومات والبيانات على كل من الـ cache والـ main memory في نفس الوقت وذلك للتأكد من ان كل المعلومات التي تم تحديثها موجودة في الـ main memory كذلك.

- **Advantage of write through:**

1. قيم البيانات في الـ main memory والبيانات في الـ cache تكون متشابهة ومتطابقة وغير مختلفة.

2. هذه الطريقة ملائمة للحاسبات المتوازية (parallel computer) أي حاسبات التي تكون من نوع (MIMD&SIMD).

أي في الحاسبات التي تحتوي على (multiple instruction stream& multiple data stream) والتي تسمى (MIMD).

وكذلك في الحاسبات التي تحتوي على (single instruction stream and multiple data stream) والتي تسمى (SIMD).

- **Disadvantage of write through:**

1. توجد عمليات كتابة غير ضرورية تجري في الـ main memory.
2. السرعة تكون بطيئة لان كتابة البيانات من الـ (input/output) الى الـ RAM والـ Cache تتم في وقت واحد حيث اي تحديث في الـ cache يرسل التحديث الى الـ RAM فلذلك تكون سرعة كتابة البيانات في الـ Cache تكون بطيئة.
3. عدد مرات الرجوع الى الذاكرة main memory تكون كبيرة.

Question) Justify the write through method suitable for MIMD & SIMD computer?

Answer)

لان في هذه الطريقة كل processor يحتوي على cache خاص به ،فعند طلب احد المعالجات معلومة معينة او قراءة معلومة من الذاكرة Ram سوف يقوم هذا المعالج بتحديث هذه المعلومة في كل من الـ cacheالخاصة به والـ Ram في نفس الوقت , لذلك فعند طلب المعلومة من معالج اخر سيأخذ المعلومة وقراءتها بشكل محدث.

2. Write-back:

In a write back scheme, only the cache memory is updated during a write operation. This method states that when writing in the cache several times, the wizard only updates the cache, and after updates are completed in the cache completely, all updates are sent at once to the cache and Ram at the same time.

Write back: تنص هذه الطريقة انه عند الكتابة في الـ cache عدة مرات يقوم المعالج بتحديث الـ cache فقط ،وبعد ان تتم التحديثات في الـ cache بصورة كاملة يتم ارسال التحديثات جميعها مرة واحدة الى الـ Ram.

- **Advantage of write back:**

1. تقليل عدد مرات الرجوع الى الذاكرة RAM .
2. زيادة سرعة الكتابة وتحديث البيانات في الـ cache.
3. هذه الطريقة تكون مناسبة للحاسبات التي تحتوي على single processor.

- **Disadvantage of write back:**

1. لايوجد دائما تطابق بين معلومات الـ cache والـ RAM .
2. هذه الطريقة غير ملائمة للحاسبات المتوازية (MIMD&SIMD).

Q) Justify the write back method not suitable for MIMD & SIMD computer?

لان عند قيام الـ CPU بأخذ القيمة او المعلومة من الـ RAM سوف يقوم الـ CPU بتحديث تلك المعلومة في الـ cache التي داخل هذا المعالج فقط الخاصة به وعند طلب نفس المعلومة من قبل معالج اخر من الـ RAM سيأخذ المعلومة القديمة (الغير محدثة) وقراءتها بشكل غيرمحدث. هذه الطريقة مناسبة للحاسبات التي تحتوي على CPU واحد فقط.

Difference Between Write Through and Write Back Methods:

- **Write Through Method**

1. في هذه الطريقة يتم تحديث الذاكرة الرئيسية مع كل عملية كتابة للذاكرة cache وكذلك يتم تحديث ذاكرة التخزين المؤقت الـ cache بالتوازي مع الـ RAM.
2. تحتوي الذاكرة الرئيسية RAM دائمًا على نفس البيانات الموجودة في ذاكرة التخزين المؤقت الـ cache.
3. عند توصيل جهاز الإدخال / الإخراج عبر DMA(data memory access) ، سيتلقى أحدث البيانات.
4. عملية الكتابة في ذاكرة التخزين المؤقت والذاكرة الرئيسية تكون في وقت واحد.
5. هذه الطريقة ملائمة للحاسبات المتوازية (MIMD&SIMD).
6. عدد مرات الرجوع الى الذاكرة main memory تكون كبيرة.
7. السرعة تكون بطيئة لان كتابة البيانات من الـ (input/output) الى الـ RAM والـ Cache تتم في وقت واحد.

- **Write Back Methods**

1. في هذه الطريقة يتم تحديث موقع ذاكرة التخزين المؤقت الـ cache فقط أثناء عملية الكتابة.
2. قد تحتوي الذاكرة الرئيسية وذاكرة التخزين المؤقت الـ cache على بيانات مختلفة.
3. عند اتصال جهاز الإدخال / الإخراج عبر DMA (data memory access) لن يتلقى أحدث البيانات.
4. عملية الكتابة في ذاكرة التخزين المؤقت والذاكرة الرئيسية لاتكون في وقت واحد.
5. هذه الطريقة تكون مناسبة للحاسبات التي تحتوي على single processor.
6. تقليل عدد مرات الرجوع الى الذاكرة RAM .
7. زيادة سرعة الكتابة وتحديث البيانات في الـ cache.