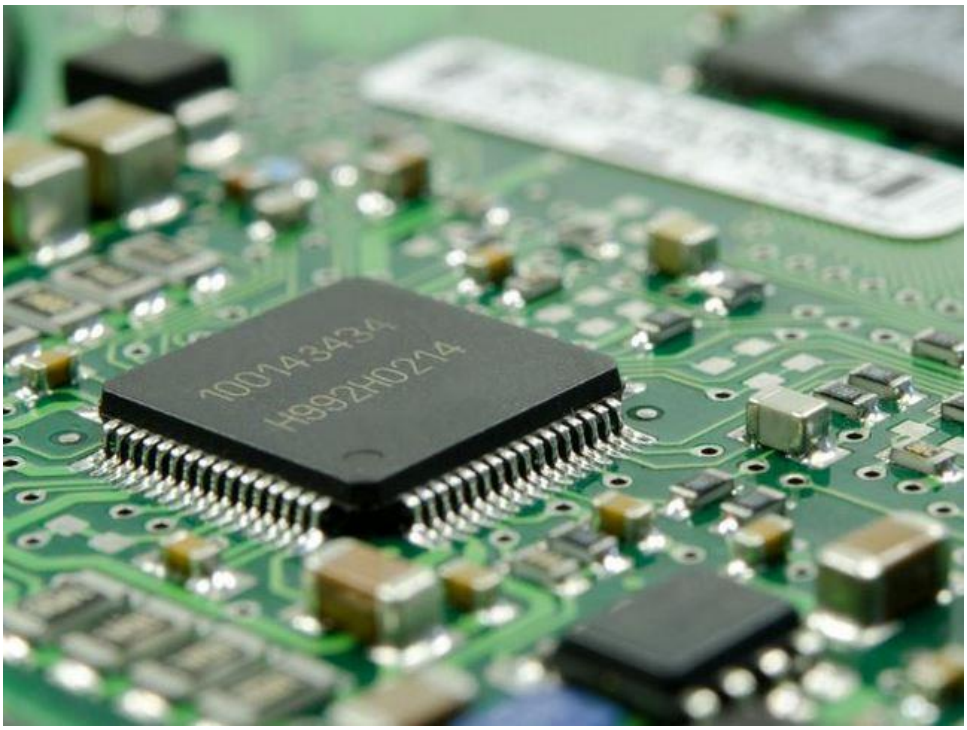
جامعة الحمدانية
كلية التربية

Lecture 9

Computer Architecture



Prepared By :

Firas Abdulrahman Yosif

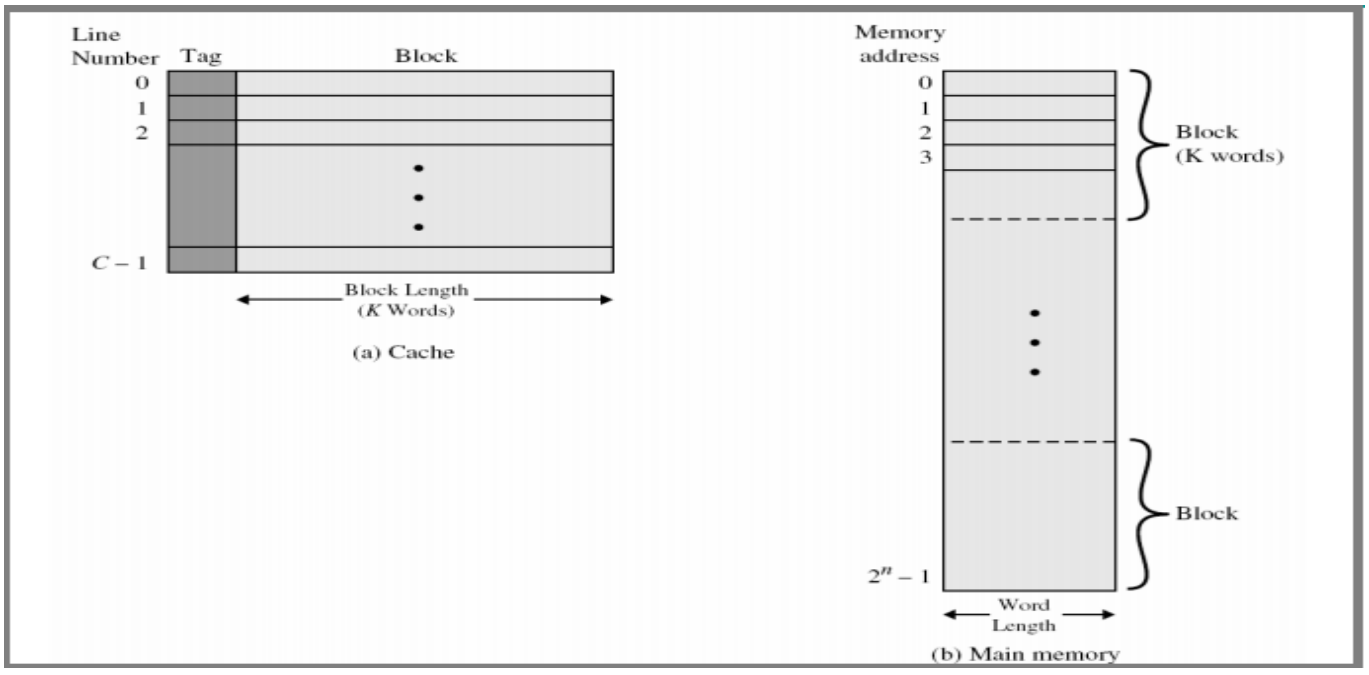
Design of Cache Memory Structure

تخزن الـ cache مجموعة من عناوين الـ main memory وماتحتويه تلك العناوين من words والتي تسمى (blocks) كل واحدة من هذه البلوك جزء من صفحات الذاكرة الرئيسية (memory pages) والتي تقابل العناوين المخزونة في الـ Tag address وبذلك تكون محتويات الـ cache هي نسخ من مجموعة صغيرة من blocks الذاكرة الرئيسية الغير المتسلسلة و المرتبطة بعناوينها, و block الذاكرة الرئيسية الموجودة داخل الـ cache تسمى الـ block وهي تحتوي على مجموعة من الـ words التي تحتوي بداخلها على بيانات.

العنوان الرئيسي يقسم الى جزئين الـ Tag والـ word وهذا يمثل الـ main memory address والـ Tag يمثل عنوان الـ block الذي يحتوي على word المقابل للعنوان, لاحظ الشكل ادناه :

Tag	word

والمخطط ادناه يوضح التركيب الداخلي(الـblock والـ Tag)بالنسبة للـ cache و الـmain memory :



Cache & Main Memory Structure

وتقسم الـ cache الى نوعين وذلك حسب عمليات البحث عن البيانات (data):

- **Data Cache** : حيث يقوم الـ processor في البحث عن البيانات ضمن مجموعة البيانات ضمن الـ cache .
- **Instruction Cache** : حيث يقوم الـ processor في البحث عن الايعازات ضمن مجموعة الايعازات ضمن الـ cache memory .

Why is the instruction memory designed?

لان الايعازات لاتشبه البيانات وذلك لان الايعازات لاتتغير اثناء التنفيذ بينما البيانات تحتاج الى تحديث دائما (update) ولهذا نقلل عدد مرات الرجوع الى الذاكرة الرئيسية لعمل نسخ من الـ cache الى الـ main اذ لاتوجد حاجة لاعادة كتابة الايعاز على الذاكرة الرئيسية مرة ثانية.

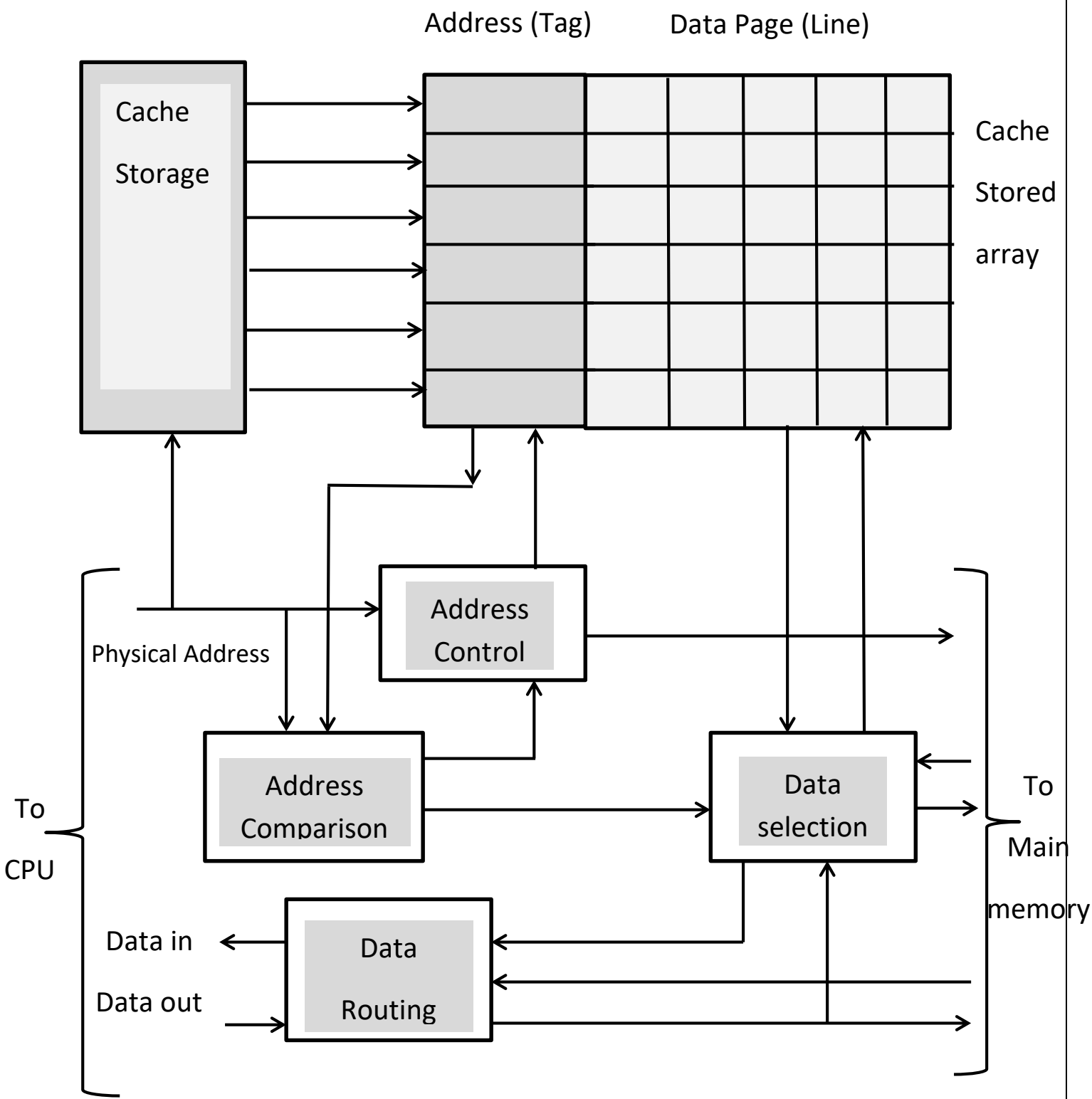
How to increase write & read operation speed in cache?

- عن طريق جعل معدل زمن المطلوب للوصول الى الذاكرة الرئيسية يرى من قبل الـ CPU وكأنه يساوي معدل زمن الوصول الى الـ cache .
- عن طريق: cache access time for data = main access time for data
  اي كلما الـ CPU يطلب معلومة يجدها في الـ cache اي ان Hit ratio =1 .

Cache design:

يتكون التصميم الداخلي للـ cache memory من مجموعة من البايتات التي تمثل مواقع الذاكرة والتي تسمى بالـ blocks او (lines) والتي تقابل الـ blocks في الذاكرة الرئيسية حيث لها نفس الحجم ولكن اقل عدد من الـ block في الـ RAM وكل block داخل الـ cache يحتوي على words وهذه الـ words تحتوي بيانات ولها عناوين خاصة تسمى Tag . ولكي يستطيع الـ processor القراءة والكتابة داخل الـ cache يحتاج الى بعض الدوائر الالكترونية مثل address comparison(وهي المسؤلة عن مقارنة العناوين بين العنوان القادم من المعالج وبين العنوان الموجود في الـ cache), data selectionوهي المسؤلة عن جلب البيانات من الـ data page داخل الـ cache , address , control , data routing .

والمخطط يوضح التصميم الداخلي للـ cache memory :



Block diagram of Basic Cache Design

Operation in cache memory:

عندما يريد المعالج قراءة معلومة معينة فأنه في البداية يرسل العنوان الى الـ cache memory وفي هذه الحالة ستكون لدينا حالتين اما reading hit او reading miss وفيما يلي شرح الية القراءة داخل الـ cache memory من قبل الـ processor :

1. Reading Hit:

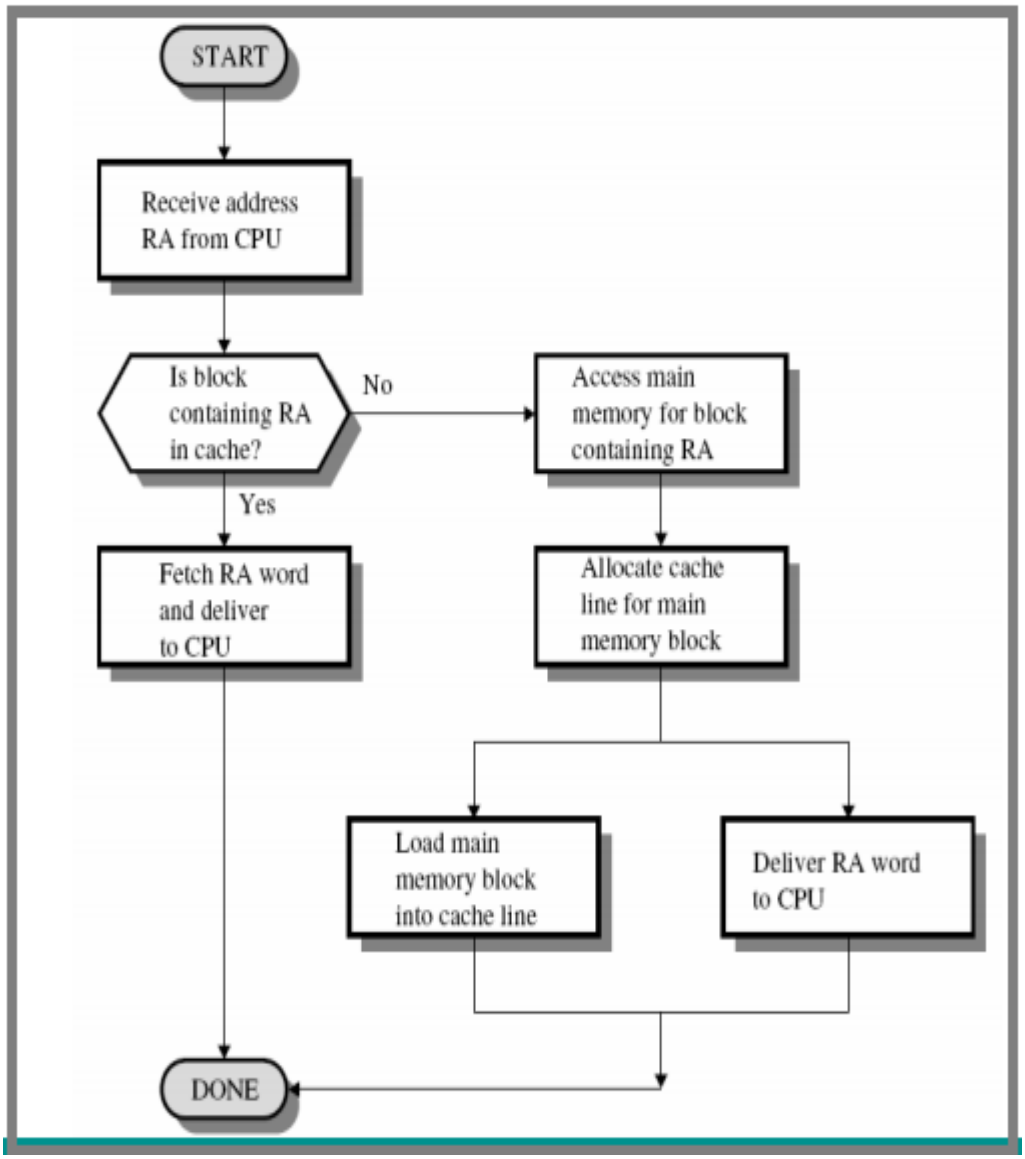
الـ CPU في البداية يرسل عنوان فيزيائي physical address الى الـ cache بعد ذلك يدخل العنوان الى دائرة مقارنة العناوين (address comparison) التي تقوم بمقارنة جزء الـ Tag من العنوان الذي ارسل من الـ CPU مع جميع الـ Tag الخاص بالعناوين داخل الـ cache, فأذا حدث تطابق بين عنوان القادم من الـ CPU مع العنوان الموجود داخل الـ cache معنى هذا ان العملية Hit اي ان المعلومة موجودة داخل الـ cache , بعد ذلك تقوم دائرة الـ address comparison بأرسال signal الى دائرة الـ data selection هذه الدائرة تقوم بأختيار البيانات من الـ block الموجود داخل الـ cache حيث تقوم بأخذ المعلومة المطلوبة من الـ block المقابل للعنوان داخل الـ cache وارسال المعلومة الى السجل data reading register حيث يخزن المعلومات بالتسلسل ثم يرسل المعلومة عن طريق دائرة الـ data routing الى الـ CPU لغرض قراءة المعلومة.

2. Reading Miss:

الـ CPU في البداية يرسل عنوان فيزيائي physical address الى الـ cache بعد ذلك يدخل العنوان الى دائرة مقارنة العناوين (address comparison) التي تقوم بمقارنة جزء الـ Tag من العنوان الذي ارسل من الـ CPU مع جميع الـ Tag الخاص بالعناوين داخل الـ cache, فأذا لم يحدث تطابق بين العناوين معنى هذا ان المعلومة التي يبحث عنها الـ processor داخل الـ cache غير موجودة وهذه العملية تسمى cache Miss, عندئذ تقوم دائرة الـ address comparison بأرسال signal الى دائرة الـ address control حيث تقوم بأرسال العنوان القادم من الـ CPU الى main memory

حيث تقوم الذاكرة الرئيسية الـ RAM بأرسال المعلومة المطلوبة مباشرة الى الـ CPU عن طريق دائرة الـ data routing وفي نفس الوقت تقوم الـ main memory بتحميل الـ block المرافق لهذه المعلومة الى الـ cache memory عن طريق دائرة الـ data selection .

المخطط ادناه يوضح الية القراءة في الذاكرة cache memory :



Flow Chart of Cache Read Operation

Writing Operation in cache memory:

عندما يريد المعالج كتابة معلومة معينة فأنه في البداية يرسل العنوان الى الـ cache memory وفي هذه الحالة ستكون لدينا حالتين اما writing Hit او writing Miss وفيما يلي شرح الية الكتابة داخل الـ cache memory من قبل الـ processor :

1. Writing Hit:

الـ CPU في البداية يرسل عنوان فيزيائي physical address الى الـ cache بعد ذلك يدخل العنوان الى دائرة مقارنة العناوين (address comparison) التي تقوم بمقارنة جزء الـ Tag من العنوان الذي ارسل من الـ CPU مع جميع الـ Tag الخاص بالعناوين داخل الـ cache, فأذا حدث تطابق بين عنوان الذي ارسله الـ Processor وبين عنوان الموجود في الـ cache memory فعندئذ تسمى العملية Cache Hit , بعد ذلك تقوم دائرة مقارنة العناوين (address comparison) بأرسال اشارة الى data selection التي تقوم بأخذ المعلومة عن طريق دائرة data routing والبيانات الداخلة اليها تأتي عن طريق الـ data bus القادم من الـ CPU, والـ data selection بدورها ترسل المعلومة الى الـ block المقابل للعنوان الموجود داخل الـ cache وعندئذ تتم عملية الكتابة write operation داخل الـ cache memory .

1. Writing Miss:

الـ CPU في البداية يرسل عنوان فيزيائي physical address الى الـ cache بعد ذلك يدخل العنوان الى دائرة مقارنة العناوين (address comparison) التي تقوم بمقارنة جزء الـ Tag من العنوان الذي ارسل من الـ CPU مع جميع الـ Tag الخاص بالعناوين داخل الـ cache, , فأذا لم يحدث تطابق بين عنوان الذي ارسله الـ Processor وبين عنوان الموجود في الـ cache memory فعندئذ تسمى العملية Cache Miss , فعندئذ تقوم دائرة مقارنة العناوين (address comparison) بأرسال

اشارة الى دائرة Address control التي تقوم بتمرير العنوان القادم من الـ processor الى ذاكرة الرئيسية الـ main memory التي تقوم بدورها بأخذ المعلومة المطلوبة عن طريق دائرة

الـ data routing حيث تستقبل المعلومة عن طريق الـ data bus القادم من الـ CPU . بعد ذلك ترسل الـ main memory الـ block المجاورللمعلومة الى الـ cache عن طريق الـ data selection وبذلك تتم عملية الكتابة.