

## Chapitre 3 : Différents types de mémoire

### 1. Introduction

Dans le chapitre précédent nous avons présenté la mémoire centrale. Dans ce chapitre on va voir les différents types de mémoires utilisés dans les ordinateurs.

Nous nous intéressons maintenant au fonctionnement des mémoires vives (ou volatiles), qui ne conservent leur contenu que lorsqu'elles sont sous tension.

Ce type de mémoire est souvent désigné par l'acronyme **RAM, Random Access Memory**, signifiant que la mémoire adressable (on peut accéder à n'importe quelle information dans la mémoire) par opposition aux mémoires secondaires séquentielles comme les bandes (pour accéder à une information il faut passer sur toutes les informations qu'elles précèdent).

Nous mentionnerons aussi différents types de mémoires mortes, qui sont des circuits accessibles uniquement en lecture (**ROM, Read Only Memory**).

### 2. Les mémoires vives (RAM)

Une mémoire vive sert au stockage temporaire de données. Elle doit avoir un temps de lecture et écriture très court pour ne pas ralentir le microprocesseur. Les mémoires vives sont en général volatiles : elles perdent leurs informations en cas de coupure d'alimentation. Certaines d'entre elles, ayant une faible consommation, peuvent être rendues non volatiles par l'adjonction d'une batterie. Il existe deux grandes familles de mémoires RAM (**Random Access Memory** : mémoire à accès aléatoire) :

- Les RAM statiques
- Les RAM dynamiques

#### 4.1 Les RAM statiques

Le bit mémoire d'une RAM statique (**SRAM**) est composé d'une bascule. Chaque bascule contient entre 4 et 6 transistors. Les **SRAM** permettent des temps d'accès courts à l'information.

#### 4.2 Les RAM dynamiques

Dans les RAM dynamiques (**DRAM**), chaque bit est réalisé à partir d'un transistor relié à un petit condensateur. L'état chargé ou déchargé du condensateur permet de distinguer deux états (bit 0 ou bit 1).

#### 4.3 Comparaison entre SRAM et DRAM

##### Avantages du DRAM:

La technique **DRAM** permet une plus grande densité d'intégration, car un bit nécessite environ quatre fois moins de transistors que dans une mémoire statique.

##### Inconvénients du DRAM:

La présence de courants de fuite dans le condensateur contribue à sa décharge. Ainsi, l'information est perdue si on ne la régénère pas périodiquement (charge du condensateur). Les RAM dynamiques doivent donc être rafraîchies régulièrement pour entretenir la mémorisation : il s'agit de lire l'information et de la recharger. Ce rafraîchissement indispensable a plusieurs conséquences :

- il complique la gestion des mémoires dynamiques car il faut tenir compte des actions de rafraîchissement qui sont prioritaires.
- la durée de ces actions augmente le temps d'accès aux informations.

D'autre part, la lecture de l'information est destructive. En effet, elle se fait par décharge de la capacité du point mémoire lorsque celle-ci est chargée. Donc toute lecture doit être suivie d'une réécriture.

#### 4.4 Conclusions

En général les mémoires dynamiques, qui offrent une plus grande densité d'information et un coût par bit plus faible, sont utilisées pour la mémoire centrale, alors que les mémoires statiques, plus rapides, sont utilisées lorsque le facteur vitesse est critique, notamment pour des mémoires de petite taille comme les caches et les registres.

### 3. Les mémoires mortes (ROM)

Pour certaines applications, il est nécessaire de pouvoir conserver des informations de façon permanente même lorsque l'alimentation électrique est interrompue. On utilise alors des mémoires mortes ou mémoires à lecture seule (**ROM : Read Only Memory**). Ces mémoires sont non volatiles. Ces mémoires, contrairement aux **RAM**, ne peuvent être que lues. L'inscription en mémoire des données reste possible mais est appelée programmation. Suivant le type de ROM, la méthode de programmation changera. Il existe donc plusieurs types de ROM :

- ROM
- PROM
- EPROM
- EEPROM
- FLASH EPROM.

#### 3.1 ROM

ROM est un circuit intégré dont le contenu est déterminé une fois pour toute au moment de la fabrication. Le coût relativement élevé de leur fabrication impose une fabrication en grandes séries, ce qui complique la mise à jour de leur contenu. Au départ, ces mémoires étaient utilisées pour stocker les parties bas-niveau du système d'exploitation de l'ordinateur (BIOS du PC par exemple).

#### 3.2 PROM (Programmable ROM)

Alors que la mémoire ROM est enregistrée de manière irréversible lors de sa fabrication, la mémoire PROM est configurée par l'utilisateur en utilisant un programmeur de PROM, utilisé pour enregistrer son contenu. Le circuit PROM ne peut plus être modifié par la suite.

#### 3.3 EPROM (Erasable PROM)

Les mémoires EPROM sont des PROM reconfigurables : il est possible de les effacer pour les reprogrammer. L'effaçage se produit en exposant le boîtier à un rayonnement ultraviolet (UV).

#### EEPROM (Electrically Erasable PROM)

Même principe qu'une EPROM, mais l'effacement se fait à l'aide de signaux électriques, ce qui est plus rapide et pratique.

#### 3.4 FLASH EPROM (Flash disque et carte mémoire)

Les mémoires FLASH sont similaires aux mémoires EEPROM, mais l'effacement peut se faire par sélectivement par blocs et ne nécessite pas le démontage du circuit. La Flash EPROM a connu un essor très important ces dernières années avec le boom de la téléphonie portable et des appareils multimédia (PDA, appareil photo numérique, lecteur MP3, etc...).

## 4. Cache mémoire

### 4.1 Problème posé

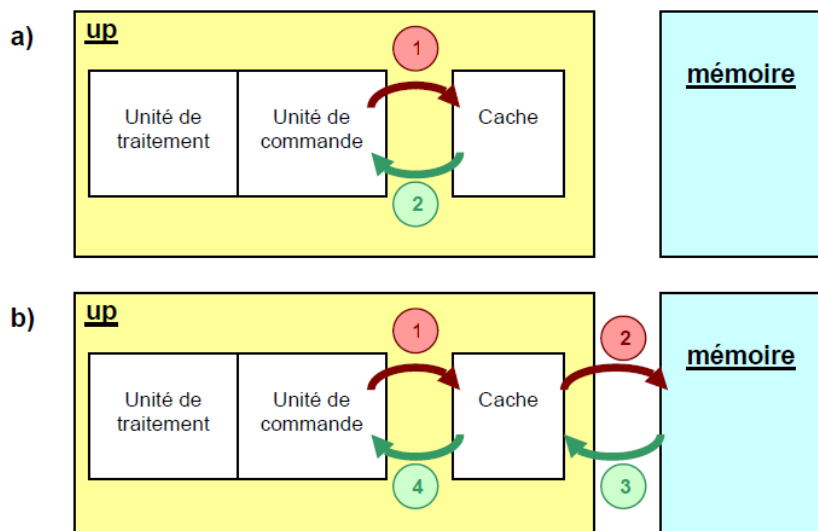
L'écart de performance entre le microprocesseur et la mémoire ne cesse de s'accroître. En effet, les composants mémoire bénéficient des mêmes progrès technologique que les microprocesseurs mais le décodage des adresses et la lecture/écriture d'une données sont des étapes difficiles à accélérer. Ainsi, le temps de cycle processeur décroît plus vite que le temps d'accès mémoire entraînant un goulot d'étranglement. La mémoire n'est plus en mesure de délivrer des informations aussi rapidement que le processeur est capable de les traiter. Il existe donc une latence d'accès entre ces deux organes.

### 4.2 Principe

Depuis le début des années 80, une des solutions utilisées pour masquer cette latence est de disposer une mémoire très rapide entre le microprocesseur et la mémoire. Elle est appelée cache mémoire. On compense ainsi la faible vitesse relative de la mémoire en permettant au microprocesseur d'acquérir les données à sa vitesse propre. On la réalise à partir de cellule SRAM de taille réduite (à cause du coût). Sa capacité mémoire est donc très inférieure à celle de la mémoire centrale et sa fonction est de stocker les informations les plus récentes ou les plus souvent utilisées par le microprocesseur. Au départ cette mémoire était intégrée en dehors du microprocesseur mais elle fait maintenant partie intégrante du microprocesseur et se décline même sur plusieurs niveaux.

Le principe de cache est très simple : le microprocesseur n'a pas conscience de sa présence et lui envoie toutes ses requêtes comme s'il agissait de la mémoire principale :

- Soit la donnée ou l'instruction requise est présente dans le cache et elle est alors envoyée directement au microprocesseur. On parle de succès de cache. (a)
- Soit la donnée ou l'instruction n'est pas dans le cache, et le contrôleur de cache envoie alors une requête à la mémoire principale. Une fois l'information récupérée, il la renvoie au microprocesseur tout en la stockant dans le cache. On parle de défaut de cache. (b)



Bien entendu, le cache mémoire n'apporte un gain de performance que dans le premier cas.

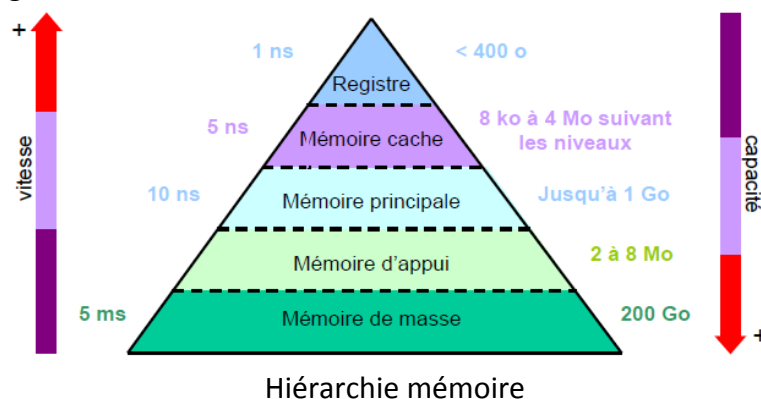
#### Remarques :

- Un cache utilisera une carte pour savoir quels sont les mots de la mémoire centrale dont il possède une copie.

- Il existe dans le système deux copies de la même information : l'originale dans la mémoire principale et la copie dans le cache. Si le microprocesseur modifie la donnée présente dans le cache, il faudra prévoir une mise à jour de la mémoire principale.
- Lorsque le cache doit stocker une donnée, il est amené à en effacer une autre. Il existe donc un contrôleur permettant de savoir quand les données ont été utilisées pour la dernière fois. La plus ancienne non utilisée est alors remplacée par la nouvelle.

## 5. Notion de hiérarchie mémoire

Une mémoire idéale serait une mémoire de grande capacité, capable de stocker un maximum d'informations et possédant un temps d'accès très faible afin de pouvoir travailler rapidement sur ces informations. Mais il se trouve que les mémoires de grande capacité sont souvent très lente et que les mémoires rapides sont très chères. Et pourtant, la vitesse d'accès à la mémoire conditionne dans une large mesure les performances d'un système. En effet, c'est là que se trouve le goulot d'étranglement entre un microprocesseur capable de traiter des informations très rapidement et une mémoire beaucoup plus lente (ex : processeur à 3Ghz et mémoire à 400MHz). Or, on n'a jamais besoin de toutes les informations au même moment. Afin d'obtenir le meilleur compromis coût-performance, on définit donc une hiérarchie mémoire. On utilise des mémoires de faible capacité mais très rapide pour stocker les informations dont le microprocesseur se sert le plus et on utilise des mémoires de capacité importante mais beaucoup plus lente pour stocker les informations dont le microprocesseur se sert le moins. Ainsi, plus on s'éloigne du microprocesseur et plus la capacité et le temps d'accès des mémoires vont augmenter.



- **Les registres** sont les éléments de mémoire les plus rapides. Ils sont situés au niveau du processeur et servent au stockage des opérandes et des résultats intermédiaires.
- **La mémoire cache** est une mémoire rapide de faible capacité destinée à accélérer l'accès à la mémoire centrale en stockant les données les plus utilisées.
- **La mémoire centrale** est l'organe principal de rangement des informations. Elle contient les programmes (instructions et données) et est plus lente que les deux mémoires précédentes.
- **La mémoire d'appui** sert de mémoire intermédiaire entre la mémoire centrale et les mémoires de masse. Elle joue le même rôle que la mémoire cache.
- **La mémoire de masse** est une mémoire périphérique de grande capacité utilisée pour le stockage permanent ou la sauvegarde des informations. Elle utilise pour cela des supports magnétiques (disque dur) ou optiques (CDROM, DVDROM).