

Chapitre 4

Les mémoires à semi-conducteur

Une mémoire est un circuit à semi-conducteur permettant d'enregistrer, de conserver et de restituer des informations (instructions et variables). C'est cette capacité de mémorisation qui explique la polyvalence des systèmes numériques et leur adaptabilité à de nombreuses situations. Les informations peuvent être écrites ou lues. Il y a écriture lorsqu'on enregistre des informations en mémoire, lecture lorsqu'on récupère des informations précédemment enregistrées.

Avec une adresse de n bits il est possible de référencer au plus 2^n cases mémoire. Chaque case est remplie par **un mot** de données (sa longueur m est toujours une puissance de 2). Le nombre de fils d'adresses d'un boîtier mémoire définit donc le nombre de cases mémoire que comprend le boîtier. Le nombre de fils de données définit la taille des données que l'on peut sauvegarder dans chaque case mémoire.

En plus du bus d'adresses et du bus de données, un boîtier mémoire comprend une entrée de commande qui permet de définir le type d'action que l'on effectue avec la mémoire (lecture/écriture) et une entrée de sélection qui permet de mettre les entrées/sorties du boîtier en haute impédance.

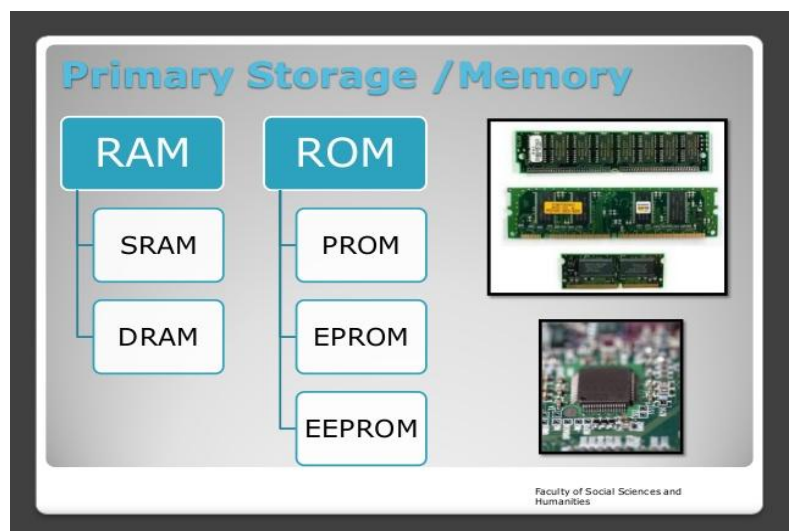


Figure 4.1 Type de mémoires dans des cartes à microprocesseur

4.1 Caractéristiques d'une mémoire

La capacité : c'est le nombre total de bits que contient la mémoire. Elle s'exprime aussi souvent en octet.

Le format des données : c'est le nombre de bits que l'on peut mémoriser par case mémoire.

On dit aussi que c'est la largeur du mot mémorisable.

Le temps d'accès : c'est le temps qui s'écoule entre l'instant où a été lancée une opération de lecture/écriture en mémoire et l'instant où la première information est disponible sur le bus de données.

Le temps de cycle : il représente l'intervalle minimum qui doit séparer deux demandes successives de lecture ou d'écriture.

Le débit : c'est le nombre maximum d'informations lues ou écrites par seconde.

Volatilité : elle caractérise la permanence des informations dans la mémoire. L'information stockée est volatile si elle risque d'être altérée par un défaut d'alimentation électrique et non volatile dans le cas contraire.

4.2 Les mémoires vives (RAM)

Une mémoire vive sert au stockage temporaire de données. Elle doit avoir un temps de cycle 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 Acces Memory : mémoire à accès aléatoire) :

- Les RAM statiques
- Les RAM dynamiques

4.2.1 Les RAM statiques

Le bit mémoire d'une RAM statique (SRAM) est composé d'une bascule. Chaque bascule contient entre 4 et 6 transistors.

4.2.2 Les RAM dynamiques

Dans les RAM dynamiques (DRAM), l'information est mémorisée sous la forme d'une charge électrique stockée dans un condensateur. Cette technique permet une plus grande densité d'intégration, car un point mémoire nécessite environ quatre fois moins de transistors que dans une mémoire statique. Sa consommation s'en retrouve donc aussi très réduite.

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.

4.3 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, et contrairement aux RAM, ne peuvent être que lues. L'inscription en mémoire des données reste possible mais est appelée programmation.

4.3.1 La ROM

Elle est programmée par le fabricant et son contenu ne peut plus être ni modifiée, ni effacée par l'utilisateur.

Avantages :

- Non volatile
- Mémoire rapide

Inconvénients :

- Écriture impossible
- Modification impossible (toute erreur est fatale)
- Délai de fabrication (3 à 6 semaines)
- Obligation de grandes quantités en raison du coût élevé qu'entraîne la production du masque et le processus de fabrication.

3.3.2 La PROM

C'est une ROM qui peut être programmée une seule fois par l'utilisateur (Programmable ROM). La programmation est réalisée à partir d'un programmeur spécifique.

Avantages :

- Claquage en quelques minutes
- Coût relativement faible

Inconvénients :

4.3.3 L'EPROM ou l'UV-EPROM

Pour faciliter la mise au point d'un programme ou tout simplement permettre une erreur de programmation, il est intéressant de pouvoir reprogrammer une PROM. La technique de claquage utilisée dans celles-ci ne le permet évidemment pas. L'EPROM (Erasable Programmable ROM) est une PROM qui peut être effacée.

Avantages :

- Reprogrammable et non Volatile

Inconvénients :

- Impossible de sélectionner une seule cellule à effacer
- Impossible d'effacer la mémoire in-situ
- l'écriture est beaucoup plus lente que sur une RAM (environ 1000x).

4.3.4 L'EEPROM

L'EEPROM (Electrically EPROM) est une mémoire programmable et effaçable électriquement. Elle répond ainsi à l'inconvénient principal de l'EPROM et peut être programmée in situ.

Avantages :

- Comportement d'une RAM non Volatile.
- Programmation et effacement mot par mot possible.

Inconvénients :

- Très lente pour une utilisation en RAM.

4.3.5 La FLASH EPROM

La mémoire Flash s'apparente à la technologie de l'EEPROM. Elle est programmable et effaçable électriquement comme les EEPROM.

Avantages :

- Comportement d'une RAM non Volatile.
- Programmation et effacement mot par mot possible.

Inconvénients :

- Très lente pour une utilisation en RAM.
- Coût de réalisation.

4.4 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.

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.

4.5 Notion de cache mémoire

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.

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 principale 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.
- 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.

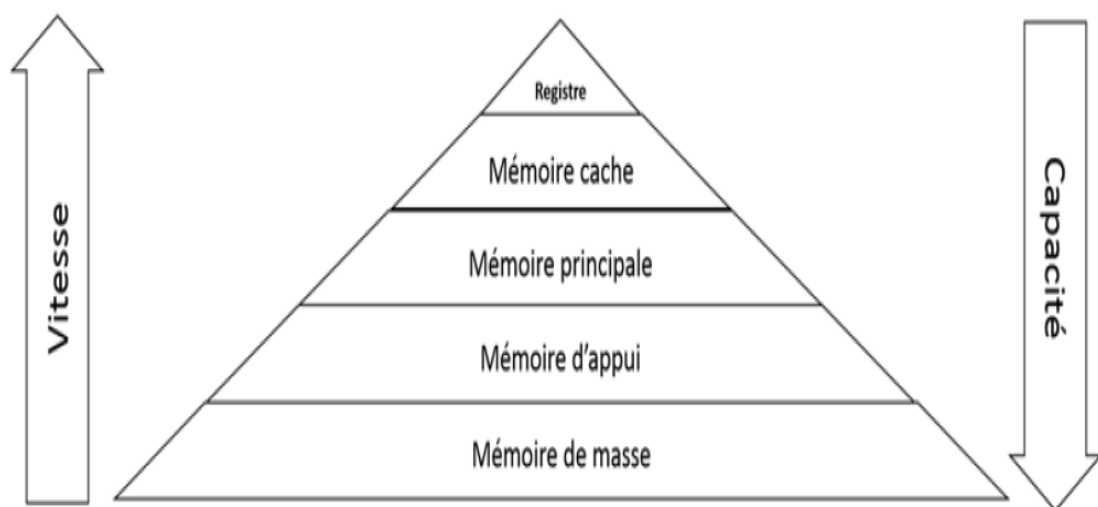


Figure 4.2 Hiérarchie mémoire

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 principale : 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 électroniques (Flash disk, SD-cards), magnétiques (disque dur) ou optiques (CDROM, DVDROM).

