

Support de cours de la matière : Architecture des Ordinateurs (AO)

Chapitre II : Principaux composants d'un ordinateur

II.1 Schéma global d'une architecture d'un ordinateur

A partir de cette section nous nous intéressons uniquement aux ordinateurs suivant le modèle de Von Neumann.

Ces ordinateurs sont constitués d'une mémoire centrale (RAM) et d'un processeur, ce dernier est lui-même constitué d'une Unité arithmétique et logique (ALU), d'une unité de contrôle (CU) et de registres généraux et registres spéciaux. La figure II.1 illustre ce schéma.

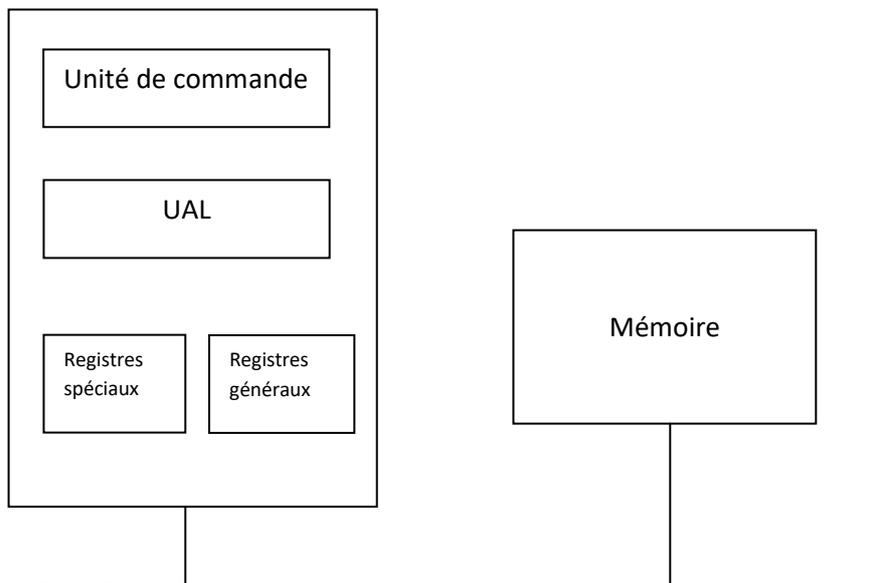


Figure II.1 Schéma général d'un ordinateur

II.2 Le processeur :

Le processeur (UCT pour Unité centrale de traitement) ou CPU (Central Processing Unit) est l'élément moteur de l'ordinateur qui interprète et exécute les instructions du programme. Son mode de fonctionnement est le suivant :

- Extraire et décoder une instruction (phase de recherche ou fetch).
- Exécuter l'instruction extraite (phase d'exécution).
- Identifier la prochaine instruction.

Avec cette simple séquence d'opérations répétée à l'infini, un processeur est capable d'exécuter un programme, quelque soit sa complexité. C'est ce qui fait l'attrait du modèle de Von Neumann. Ce mécanisme d'interprétation est resté inchangé.

Le processeur se compose de deux unités fonctionnellement séparées :

1. Unité arithmétique et logique (UAL).
2. Unité de commande (UC).

II.2.1 Unité arithmétique et logique :

L'UAL est la zone du CPU où les opérations arithmétiques et logiques sont réalisées. L'UAL opère sur les registres généraux du processeur ainsi que sur les registres spéciaux (invisibles pour les programmes) utilisés pour les calculs, par exemple le registre d'état contient des indicateurs (flags) sur le résultat de la dernière opération exécutée. Par exemple, l'indicateur Z (comme zéro) est positionné à 1 si le résultat de la dernière est nul.

Tout traitement de données a lieu dans l'UAL. Cette partie du CPU, où se trouvent tous les circuits capables d'effectuer les opérations élémentaires qui sont à la base de tout algorithme, est totalement asservie à l'unité de commande. C'est précisément cette dernière qui déclenche, contrôle et synchronise toute activité de l'UAL.

II.3 Les Bus

- Un bus est un dispositif destiné à assurer le transfert simultané d'informations entre les divers composants d'un ordinateur.
- On distingue trois catégories de Bus :
 - Bus d'adresses (unidirectionnel)
 - il permet à l'unité de commande de transmettre les adresses à rechercher et à stocker.
 - **Bus de données** (bi-directionnel)
 - sur lequel circulent les instructions ou les données à traiter ou déjà traitées en vue de leur rangement.
 - **Bus de contrôle** (bi-directionnel)
 - transporte les ordres et les signaux de synchronisation provenant de l'unité de commande vers les divers organes de la machine. Il véhicule aussi les divers signaux de réponse des composants.

Largeur du bus

Pour certains Bus on désigne par largeur du Bus, le nombre de bits qui peuvent être transportés en même temps par le Bus, on dit aussi transportés en parallèle.

Les principaux Bus de données récents de micro-ordinateur

Les Bus de données sont essentiellement des bus "synchrones", c'est à dire qu'ils sont cadencés par une horloge spécifique qui fonctionne à une fréquence fixée. Entre autres informations commerciales, les constructeurs de Bus donnent en plus de la fréquence et pour des raisons psychologiques, le débit du Bus qui est en fait la valeur du produit de la fréquence par la largeur du Bus, ce débit correspond au nombre de bits par seconde transportés par le Bus.

| BUS | Largeur | Fréquence | Débit | Utilisation |
|------|---------|------------|----------|---------------------------------------|
| PCI | 64 bits | 66 MHz | 528 Mo/s | Processeur/périphérique non graphique |
| AGP | 32 bits | 66 MHz x 8 | 4 Go/s | Processeur/carte graphique |
| SCSI | 16 bits | 40 MHz | 80 Mo/s | Echanges entre périphériques |

Il existe aussi des "Bus série" (Bus qui transportent les bits les uns à la suite des autres, contrairement aux Bus parallèles), les deux plus récents concurrents équipent les matériels de grande consommation : USB et Firewire.

| BUS | Débit | Nombre de périphériques acceptés | Ces Bus évitent de connecter des périphériques divers comme les souris, les lecteurs de DVD, les GSM, les scanners, les imprimantes, les appareils photo, ..., sur des ports spécifiques de la machine |
|-----------|----------|----------------------------------|--|
| USB | 1,5 Mo/s | 127 | |
| USB2 | 60 Mo/s | 127 | |
| Firewire | 50 Mo/s | 63 | |
| FirewireB | 200 Mo/s | 63 | |

II.4 Mémoire :

La mémoire est la partie de l'ordinateur dans laquelle programmes (instructions) et données sont stockées.

Les instructions sont stockées sous forme de code machine. Par exemple, une instruction d'addition, dans un Intel 8086, est codée : 10000001. Les données sont stockées selon d'autres codes. Par exemple, dans le code ASCII (American Standard Code for Information Interchange) le caractère Y est codé ainsi : 1011001.

II.5 Hiérarchie des mémoires :

Dans un ordinateur il y a différents niveaux (types) de mémoire organisés selon la hiérarchie schématisée dans la figure suivante :

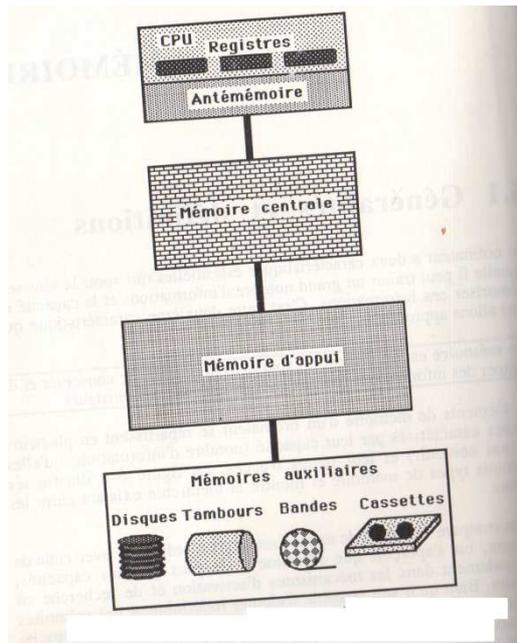


Figure I.3 : Principaux niveaux de mémoire d'un ordinateur.

Les différentes mémoires sont ordonnées en fonction des critères suivants : temps d'accès, capacité et coût par bit.

Plus on s'éloigne du CPU vers les mémoires auxiliaires, on constate que le temps d'accès et la capacité des mémoires augmentent, mais le coût par bit diminue.

Les éléments de mémoire situés dans le CPU sont les **registres** qui sont caractérisés par une grande vitesse et servent principalement au stockage des opérandes et des résultats intermédiaires.

L'**antémémoire** ou **mémoire cache** est une mémoire rapide de faible capacité (par rapport à la mémoire centrale) utilisée comme mémoire tampon entre le CPU et la mémoire centrale. Cette mémoire permet au CPU de faire moins d'accès à la mémoire centrale et ainsi de gagner du temps.

La **mémoire centrale** est l'organe principal de rangement des informations utilisées par le CPU. Pour exécuter un programme, il faut le charger (instructions + données) en mémoire centrale. Cette mémoire est une mémoire à semi-conducteurs, mais son temps d'accès est beaucoup plus grand que celui des registres et du cache.

La **mémoire d'appui** sert de mémoire intermédiaire entre la mémoire centrale et les mémoires auxiliaires. Elle est présente dans les ordinateurs les plus évolués et permet d'augmenter la vitesse d'échange des informations entre ces deux niveaux.

Les **mémoires auxiliaires**, appelées aussi **mémoires de masse** (dont les **mémoires d'archivage**), elles servent d'éléments de stockage permanent et utilisent pour cela des supports magnétiques (disques, tambours, bandes) et des supports optiques (disques optiques).

II.6 Caractéristiques des mémoires :

- **Adresse** : La mémoire centrale est divisée physiquement en cellules. Chaque cellule correspond à un mot-mémoire et possède une adresse (valeur numérique) qui lui est propre (ex : adresse d'un mot mémoire centrale). Ainsi les cellules peuvent être adressées séparément pour une opération de lecture ou d'écriture.
- **Capacité** ou taille mémoire : elle correspond au nombre d'informations qu'elle peut contenir. On peut exprimer cette valeur en fonction du nombre de bits, de bytes ou de mots (ex : mémoire centrale de 8 Méga mots de 32 bits ou disque magnétique de 2 Gbytes). La longueur d'un mot-mémoire varie d'une machine à l'autre, par exemple : 1, 4, 8, 12, 16, 24, 32, 48, 60, 64 bits. La valeur 32 tend à se généraliser dans la plupart des ordinateurs. La longueur du mot-mémoire est une caractéristique importante de l'architecture d'un ordinateur et reflète la structure des différents composants fonctionnels (principalement de l'unité centrale).
- **Temps d'accès** : c'est le temps qui s'écoule entre le lancement d'une opération d'accès (lecture ou écriture) et son accomplissement.
- **Cycle mémoire** : c'est le temps minimal s'écoulant entre deux accès successifs à la mémoire. Il est plus long que le temps d'accès, car le bon fonctionnement de la mémoire nécessite quelques opérations de maintien, de stabilisation des signaux dans les circuits, de synchronisation, etc.
- **Débit** : c'est le nombre d'informations lues ou écrites par seconde (par exemple, unité de bande magnétique avec un débit de 3 Mbytes/seconde).
- **Volatilité** : elle caractérise la permanence des informations dans une mémoire. Une mémoire volatile perd son contenu lorsque l'on coupe le courant, celle-ci a donc besoin d'un apport constant d'énergie électrique pour conserver ses informations. La mémoire centrale à semi-conducteurs est volatile alors que les mémoires auxiliaires magnétiques ne le sont pas. On peut réaliser des mémoires non volatiles à semi-conducteurs, moyennant une petite batterie.

II.7 différents types d'accès aux mémoires :

- Accès **séquentiel** : c'est l'accès le plus lent, pour accéder à une information particulière, on est obligé de parcourir toutes celles qui la précèdent (par exemple, les bandes magnétiques).
- Accès **direct** : les informations ont une adresse propre, ce qui permet de les accéder directement (par exemple : mémoire centrale, registres).
- Accès **semi-séquentiel** : c'est une combinaison des accès direct et séquentiel (par exemple, pour un disque magnétique, l'accès au cylindre est direct et l'accès au secteur séquentiel).
- Accès **par le contenu** (mémoire associative) : les informations sont identifiées par une clé et la recherche s'effectue sur cette clé de façon simultanée sur toutes les positions de la mémoire (par exemple : l'antémémoire).

Chaque case d'une mémoire associative comprend deux champs correspondant à la clé et à l'information associée à cette clé (figure I.4). Dans le cas de l'antémémoire, la clé est constituée par l'adresse en mémoire centrale de l'instruction ou la donnée cherchée et l'information associée est constituée de l'instruction ou la donnée elle-même.

La recherche par clé dans la mémoire associative ne s'effectue pas de manière séquentielle, mais en parallèle sur toutes les cases de la mémoire associative. En un accès, on sait si l'instruction cherchée se trouve ou non dans l'antémémoire. Si l'on considère la mémoire associative de la figure I.4, les clés sont des noms des pays, alors que l'information associée à chaque clé est une ville de ce pays.

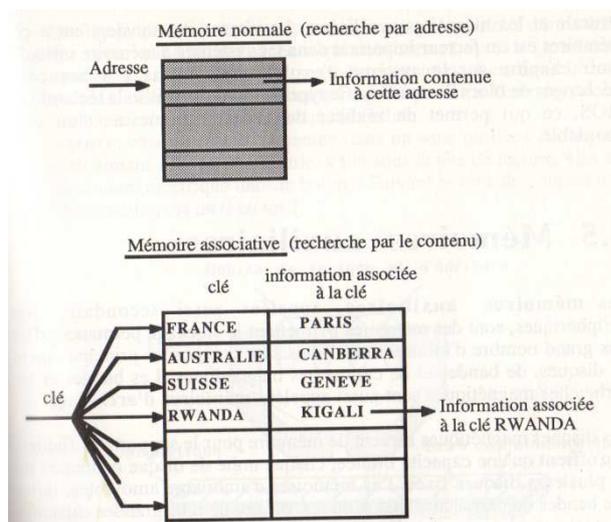


Figure I.4 Schéma comparatif de l'adressage de la mémoire centrale et de l'antémémoire

Si l'on interroge cette mémoire avec la clé RWANDA, on compare ce nom avec tous les noms contenus, chaque cellule a son propre circuit comparateur, et ainsi en un seul accès, on a l'information associée à cette clé, dans notre exemple la ville KIGALI.

II.8 Mémoires à semi conducteurs :

Les mémoires à semi conducteurs constituent les éléments de base de toute mémoire centrale. Les mémoires à semi conducteur peuvent être divisées en deux grandes parties : les mémoires vives et les mémoires mortes (ROM).

Mémoire Vive ou RAM (Random Access Memory) : est une mémoire à accès aléatoire, le temps d'accès est indépendant du numéro de la cellule adressée. On distingue deux types de mémoires RAM, les SRAM et les DRAM :

1. SRAM (Static RAM) :

Chaque bit d'une SRAM est formé par une bascule constituée par 4 à 6 transistors.

L'information stockée peut être maintenue sans dégradation pendant une centaine d'heures.

L'intérêt de ce type de mémoire est sa vitesse (quelques ns) mais son coût est prohibitif, en effet on utilisera la SRAM lorsque le facteur vitesse est critique et notamment pour des mémoires de petite taille comme la mémoire cache.

2. DRAM (Dynamic RAM) :

Chaque bit d'une DRAM est constitué par un transistor et un condensateur.

L'inconvénient des DRAM est que le condensateur possède une tendance naturelle à se décharger. Pour que l'information reste cohérente, on va devoir réaliser un **rafraîchissement** de la mémoire toutes les quelques millisecondes. Ce rafraîchissement consiste à lire et à réécrire la donnée.

Etant donné les caractéristiques des SRAM et DRAM, on peut en déduire les propriétés suivantes :

- la SRAM est rapide mais chère
- la DRAM est lente mais bon marché

Mémoire morte : on peut distinguer les mémoires suivantes :

ROM (Read Only Memory) est une mémoire morte. C'est-à-dire une mémoire où l'on peut lire uniquement, l'écriture étant impossible. Ce sont des mémoires non volatiles, rogrammées par le fabricant. Elles permettent un stockage permanent même en l'absence d'alimentation électrique ;

L'information est stockée au moment de la conception du circuit d'une manière définitive.

PROM (Programmable ROM) Permet une écriture unique mais faite par l'utilisateur au moyen d'une machine spéciale s'appelant un programmeur de PROM.

EPROM (Erasable Programmable ROM) sont aussi appelées REEPROM (REProgrammable ROM). Le principe est le même que celui des PROM, mais elles offrent en plus la possibilité de pouvoir être effacées un certain nombre de fois. L'effacement se fait par exposition aux rayons ultra-violet. Le temps d'exposition est de l'ordre de 30 minutes. Ces mémoires sont utilisées lors de la mise au point de programmes.

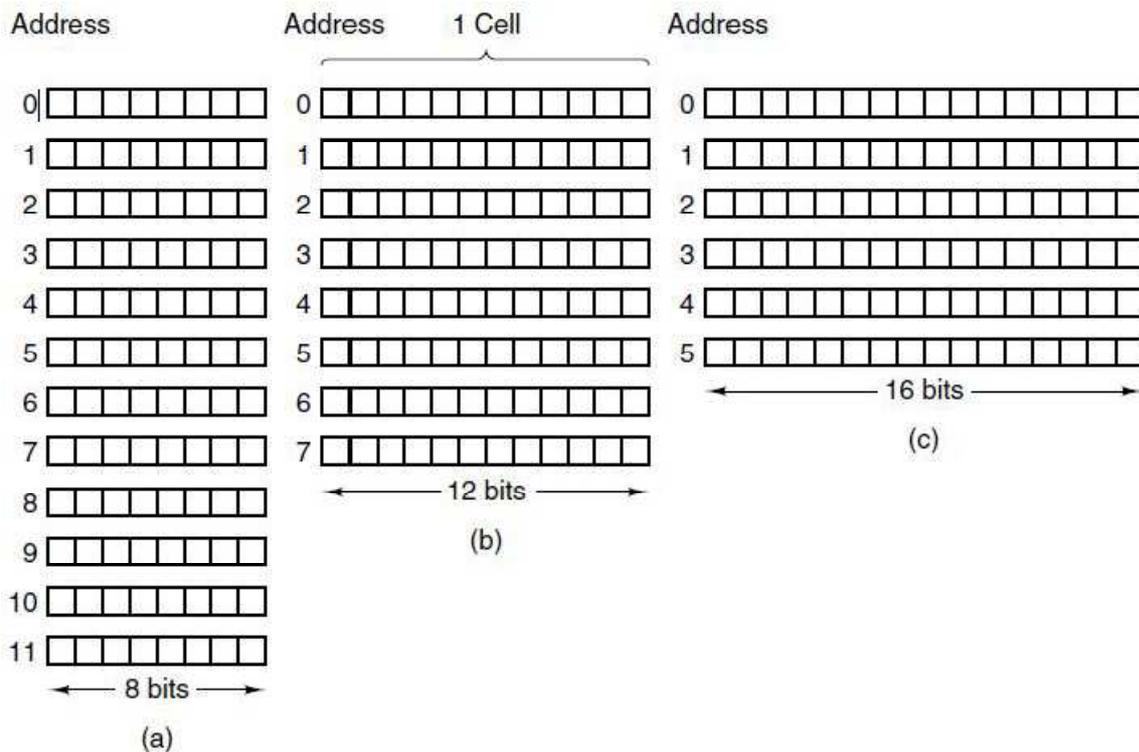
EAROM (electrically Alterable ROM) suivant le même principe que celui des EPROM, elles sont reprogrammables par l'utilisateur mais elles sont à effacement électrique. De plus l'effacement est sélectif et l'effacement total ne demande pas plus d'une minute. On trouve aussi les appellations EROM (electrically Erasable ROM) et EEPROM.

II.9 Les adresses mémoires :

Une mémoire est formée d'un certain nombre de mots (ou cases). Un mot mémoire est un ensemble de bits (contenu : instruction ou donnée). Chacun de ces mots a un numéro que nous appellerons son *adresse*, qui permet à un programme de le référencer. Si une mémoire a n mots, les adresses iront de 0 à $n-1$. Toutes les cases de la mémoire contiennent le même nombre de bits, par exemple k bits, ce qui permet de représenter 2^k combinaisons différentes.

Exemple :

Trois organisations différentes d'une mémoire à 96 bits



Dans (a) on a 12 mots (ou unités adressables) de 8 bits.

Dans (b) on a 8 mots (ou unités adressables) de 12 bits.

Dans (c) on a 6 mots (ou unités adressables) de 16 bits.

Les adresses sont exprimées en binaire et par convention commencent par 0.

Si une adresse a m bits, le nombre maximum de mots mémoire directement adressable est 2^m .

Le nombre de bits de l'adresse ne dépend que du nombre de cellules (mots) adressables et non de leur taille.

Par exemple : une mémoire de 2^{12} cellules de 8 bits chacune et une mémoire de 2^{12} cellules de 64 bits chacune nécessite toutes les deux des adresses à 12 bits.

Si n =nombre de bits d'adresse taille de la MC= 2^n unités adressables (UA).

Si la taille de la MC= 2^n UA la taille de l'adresse = $n=\log_2 2^n$.

Un registre est une cellule mémoire ayant une fonction particulière. Dans la mémoire centrale on trouve deux types de registres ; le *registre d'adresse* qui contient l'adresse d'un mot-mémoire et le *registre mot* qui contient le contenu d'un mot-mémoire. Un registre mot a la même taille qu'un mot-mémoire, alors qu'un registre d'adresse doit permettre d'adresser tous les mots de la mémoire.

Exemple :

Si la mémoire comporte 256 mots, le registre d'adresse doit avoir $\log_2(256) = \log_2(2^8) = 8$ bits. Un registre d'adresse de 32 bits permet d'adresser $2^{32} = 4\,294\,967\,296$ mots différents (4 G mots).

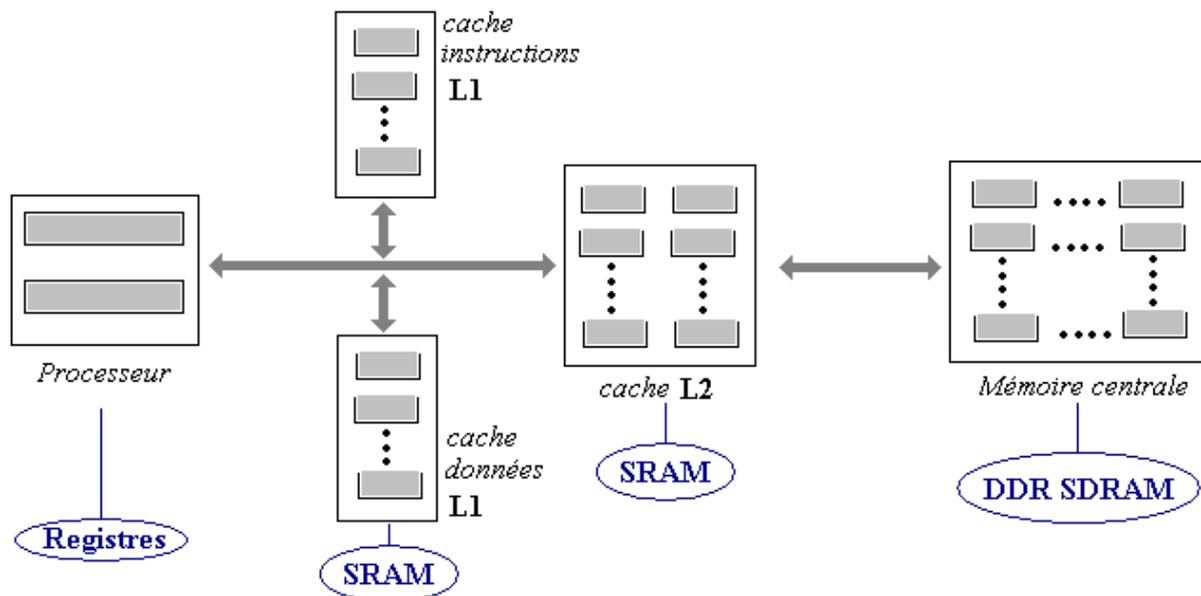
II.10 Mémoire cache

La caractéristique fonctionnelle du cache est de servir à stocker des instructions et des données provenant de la mémoire centrale et qui ont déjà été utilisées les plus récemment par le processeur central.

Actuellement le cache des micro-processeurs récents du marché est composé de deux niveaux de mémoires de type SRAM la plus rapide (type de mémoire RAM statique semblable à celle des registres) : le cache de niveau un est noté **L1**, le cache de niveau deux est noté **L2**.

Le principe est le suivant :

Le cache L1 est formé de deux blocs séparés, l'un servant au stockage des données, l'autre servant au stockage des instructions.



- Si un étage du processeur cherche une donnée, elle va être d'abord recherchée dans le cache de donnée L1 et rapatriée dans un registre adéquat, si la donnée n'est pas présente dans le cache L1, elle sera recherchée dans le cache L2.
- Si la donnée est présente dans L2, elle est alors **rapatriée** dans un registre adéquat et **recopiée** dans le bloc de donnée du cache L1. Il en va de même lorsque la donnée n'est pas présente dans le cache L2, elle est alors **rapatriée** depuis la mémoire centrale dans le registre adéquat et **recopiée** dans le cache L2.

Généralement la mémoire cache de niveau L1 et celle de niveau L2 sont regroupées dans la même puce que le processeur (**cache interne**).

II.11 Registres

Le processeur traite des informations ; les registres sont utilisés pour stocker l'information qui va être traitée ou qui vient d'être traitée. Ils stockent les informations relatives à une instruction :

Les opérandes nécessaires à l'instruction, les résultats produits par l'instruction.

Les registres sont au cœur du processeur, on peut donc en mettre peu (< 20). Ils doivent être très rapide (cadencées à la vitesse du processeur). Ce sont en fait les mémoires les plus rapides et les plus chères.

Registre 1-bit = 1 bascule RS (ou D)

Registre n-bits = n bascules RS (ou D) en parallèle

Registre 4 bits :

On appelle le registre 4 bits un registre parallèle-parallèle car on écrit et on lit en parallèle 4 bits (4 entrées et 4 sorties).

Types de registres

Il existe plusieurs types de registres dans un processeur, nous pouvons citer les suivants :

1. **Compteur de programme** : ce registre contient l'adresse mémoire de l'instruction en cours d'exécution.
2. **Accumulateur** : ce registre est utilisé pour stocker les données en cours de traitement par l'UAL.
3. **Registre d'adresses** : il contient toujours l'adresse de la prochaine information à lire par l'UAL : soit la suite de l'instruction en cours, soit la prochaine instruction.
4. **Registre d'instructions** : il contient l'instruction en cours de Traitement.
5. **Registre d'état** : il sert à stocker le contexte du processeur, ce qui veut dire que les différents bits de ce registre sont des " drapeaux " (flags) servant à stocker des informations concernant le résultat de la dernière instruction exécutée.
6. **Pointeurs de pile** : ce type de registre, dont le nombre varie en fonction du type de processeur, contient l'adresse du sommet de la pile (ou des piles).
7. **Registres généraux** : ces registres sont disponibles pour les calculs.