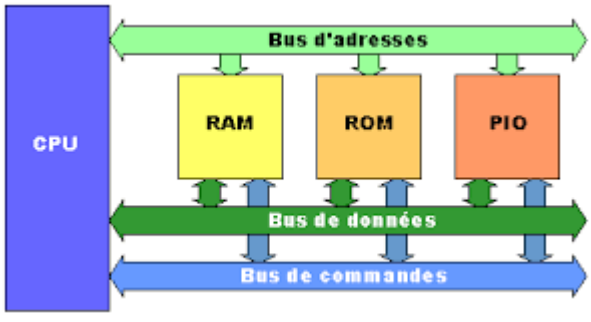
Architecture de base d'un ordinateur

Table des matières

L'architecture dite architecture de von Neumann est un modèle pour un ordinateur qui utilise une structure de stockage unique pour conserver à la fois les instructions et les données demandées ou produites par le calcul. De telles machines sont aussi connues sous le nom d'ordinateur à programme enregistré. La séparation entre le stockage et le processeur est implicite dans ce modèle.



1. Introduction

Un ordinateur peut se décrire à plusieurs échelles :

1. échelle du nanomètre: plusieurs millions de transistors qui produisent des tensions électriques hautes (bit = 1) ou basses (bit = 0)
2. échelle plus grande du microprocesseur qui effectue des opérations arithmétiques (addition, multiplication) et logiques (et, ou, etc.) à partir de données enregistrées dans sa mémoire vive (RAM).
3. échelle encore plus grande, l'ordinateur est une machine capable d'exécuter des programmes écrits dans un langage simple appelé le langage machine. Les langages machines sont difficiles à comprendre par les humains, car les programmes exprimés dans ce langage sont des suites de 0 ou de 1. On a créé 2 niveaux de langage plus simples :
   - les langages d'assemblage
   - les langages de haut niveau

Le langage d'assemblage reprend la structure du langage machine mais utilise des symboles ou étiquette plus facile à comprendre.

Par exemple, un processeur de la famille **x86** reconnaît une instruction suivante en langage machine : 10110000 01100001

En langage assembleur, cette instruction est représentée par un équivalent plus facile à comprendre pour le programmeur :

movb $0x61,%al

(10110000 = movb %al et 01100001 = $0x61)

Ce qui signifie : « écrire le nombre 97 (la valeur est donnée en **hexadécimal** : $61_{16} = 97_{10}$) dans le **registre** AL ».

Le langage d'assemblage (ou assembleur) est trop compliqué pour le commun des mortels, on utilise des langages de programmation plus simples dit de haut niveau comme le langage Java ou C. Ces langages sont traduits par un compilateur en assembleur puis en langage machine.

2. Généralités

2.1. Introduction

L'informatique, contraction d'information et automatique, est la science du traitement de l'information. Apparue au milieu du 20ème siècle, elle a connu une évolution extrêmement rapide. A sa motivation initiale qui était de faciliter et d'accélérer le calcul, se sont ajoutées de nombreuses fonctionnalités, comme l'automatisation, le contrôle et la commande de processus, la communication ou le partage de l'information. Le cours d'architecture des systèmes à microprocesseurs expose les principes de base du traitement programmé de l'information. La mise en œuvre de ces systèmes s'appuie sur deux modes de réalisation distincts, le matériel et le logiciel. Le matériel (**hardware**) correspond à l'aspect concret du système : unité centrale, mémoire, organes d'entrées-sorties, etc… Le logiciel (**software**) correspond à un ensemble d'instructions, appelé

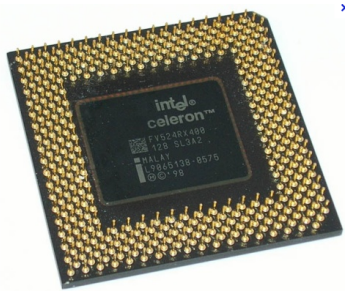
programme, qui sont contenues dans les différentes mémoires du système et qui définissent les actions effectuées par le matériel.

2.2. Qu'entend-t-on par architecture ?

L'architecture d'un système à microprocesseur représente l'organisation de ses différentes unités et de leurs interconnexions. Le choix d'une architecture est toujours le résultat d'un compromis

- entre performances et coûts
- entre efficacité et facilité de construction
- entre performances d'ensemble et facilité de programmation
- etc …

2.3. Qu'est ce qu'un microprocesseur ?



Un microprocesseur est un circuit intégré complexe. Il résulte de l'intégration sur une puce de fonctions logiques combinatoires (logiques et/ou arithmétique) et séquentielles (registres, compteur, etc…). Il est capable d'interpréter et d'exécuter les instructions d'un programme..Le concept de microprocesseur a été créé par la Société Intel, qui donna naissance, en 1971, au premier microprocesseur, le 4004, qui était une unité de calcul 4 bits fonctionnant à 108 kHz. Il résultait de l'intégration d'environ 2300 transistors.

2.4. Où trouve-t-on des systèmes à microprocesseur ?

Les applications des systèmes à microprocesseurs sont multiples et variées :

- Ordinateur, PDA
- console de jeux
- calculatrice
- télévision
- téléphone portable
- distributeur automatique d'argent
- robotique
- lecteur carte à puce, code barre
- automobile
- instrumentation
- etc…

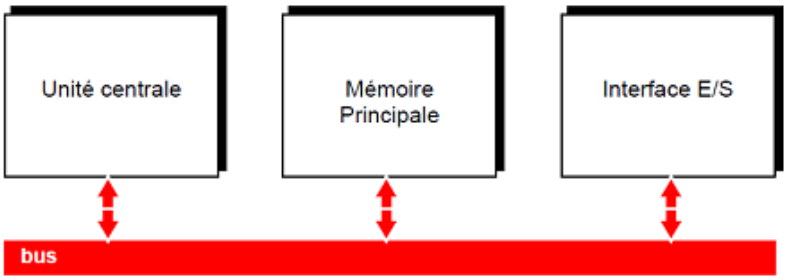
3. Architecture de base

3.1. Modèle de von Neumann

Pour traiter une information, un microprocesseur seul ne suffit pas, il faut l'insérer au sein d'un système minimum de traitement programmé de l'information. John Von Neumann est à l'origine d'un modèle de machine universelle de traitement programmé de l'information (1946). Cette architecture sert de base à la plupart des systèmes à microprocesseur actuel. Elle est composée des éléments suivants :

- une unité centrale
- une mémoire principale
- des interfaces d'entrées/sorties

Les différents organes du système sont reliés par des voies de communication appelées **bus**.



3.2. L'unité centrale

Elle est composée par le microprocesseur qui est chargé d'interpréter et d'exécuter les instructions d'un programme, de lire ou de sauvegarder les résultats dans la mémoire et de communiquer avec les unités d'échange. Toutes les activités du microprocesseur sont cadencées par une horloge.

On caractérise le microprocesseur par :

– sa fréquence d'horloge : en MHz ou GHz

– le nombre d'instructions par secondes qu'il est capable d'exécuter

– la taille des données qu'il est capable de traiter : en bits

3.3. La mémoire principale

Elle contient les instructions du ou des programmes en cours d'exécution et les données associées à ce programme. Physiquement, elle se décompose souvent en :

- une mémoire morte (**ROM** = Read Only Memory) chargée de stocker le programme. C'est une mémoire à lecture seule.
- une mémoire vive (**RAM** = Random Access Memory) chargée de stocker les données intermédiaires ou les résultats de calculs. On peut lire ou écrire des données dedans, ces données sont perdues à la mise hors tension.

Remarque : Les disques durs, disquettes, CDROM, etc… sont des périphériques de stockage et sont considérés comme des mémoires secondaires.

3.4. Les interfaces d'entrées/sorties

Elles permettent d'assurer la communication entre le microprocesseur et les périphériques. (capteur, clavier, moniteur ou afficheur, imprimante, modem, etc…).

3.5. Les bus

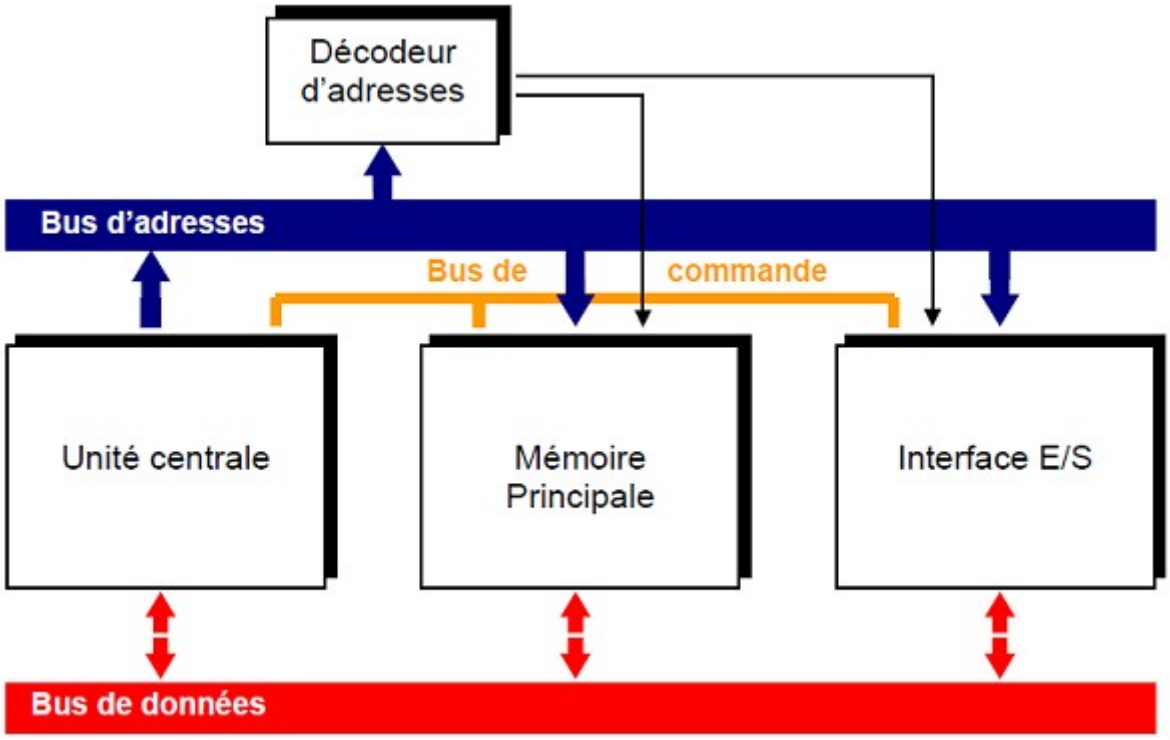
Adresse	Case mémoire
7 = 111	
6 = 110	
5 = 101	
4 = 100	
3 = 011	
2 = 010	
1 = 001	
0 = 000	0001 1010

Un bus est un ensemble de fils qui assure la transmission du même type d'information. On retrouve trois types de bus véhiculant des informations en parallèle dans un système de traitement programmé de l'information :

- **un bus de données** : bidirectionnel qui assure le transfert des informations entre le microprocesseur et son environnement, et inversement. Son nombre de lignes est égal à la capacité de traitement du microprocesseur.
- **un bus d'adresses**: unidirectionnel qui permet la sélection des informations à traiter dans un *espace mémoire* (ou *espace adressable*) qui peut avoir 2^n emplacements, avec n = nombre de fils conducteurs du bus d'adresses.
- **un bus de commande**: constitué par quelques conducteurs qui assurent la synchronisation des flux d'informations sur les bus des données et des adresses.

3.6. Décodage d'adresses

La multiplication des périphériques autour du microprocesseur oblige la présence d'un **décodeur d'adresse** chargé d'aiguiller les données présentes sur le bus de données. En effet, le microprocesseur peut communiquer avec les différentes mémoires et les différents boîtier d'interface. Ceux-ci sont tous reliés sur le même bus de données et afin d'éviter des conflits, un seul composant doit être sélectionné à la fois. Lorsqu'on réalise un système microprogrammé, on attribue donc à chaque périphérique une zone d'adresse et une fonction « décodage d'adresse » est donc nécessaire afin de fournir les signaux de sélection de chacun des composants.



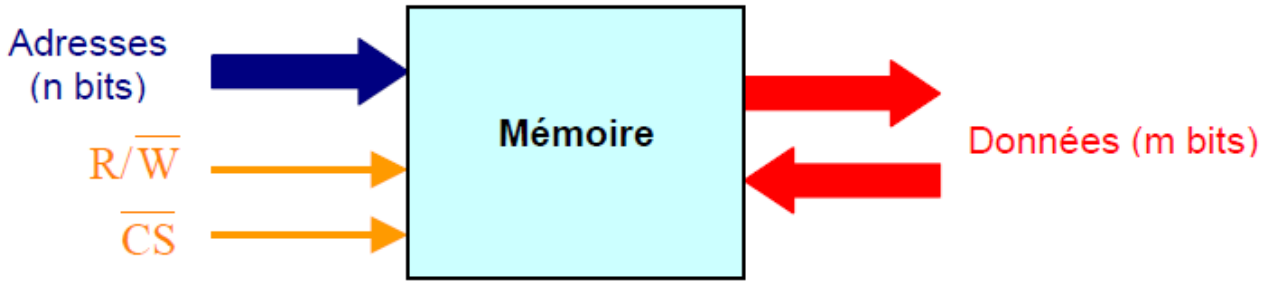
4. Les mémoires

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.

4.1. Organisation d'une mémoire

Une mémoire peut être représentée comme une armoire de rangement constituée de différents tiroirs. Chaque tiroir représente alors une case mémoire qui peut contenir un seul élément : des **données**. Le nombre de cases mémoires pouvant être très élevé, il est alors nécessaire de pouvoir les identifier par un numéro. Ce numéro est appelé **adresse**. Chaque donnée devient alors accessible grâce à son adresse. 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. On peut donc schématiser un circuit mémoire par la figure suivante où l'on peut distinguer :



- les entrées d'adresses
- les entrées de données
- les sorties de données
- les entrées de commandes :
  - ➔ une entrée de sélection de lecture (Read) ou d'écriture(write). (R/W)
  - ➔ une entrée de sélection du circuit. (CS)

Une opération de lecture ou d'écriture de la mémoire suit toujours le même cycle :

1. sélection de l'adresse
2. choix de l'opération à effectuer (R/W)
3. sélection de la mémoire (CS = 0)
4. lecture ou écriture de la donnée

Remarque : lorsqu'un composant n'est pas sélectionné, ses sorties sont mises à l'état « haute impédance » afin de ne pas perturber les données circulant sur le bus. (elle présente une impédance de sortie très élevée = circuit ouvert).

4.2. 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 (mot de 8, 16, 32 ou 64 bits).
- **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.

Remarque : Les mémoires utilisées pour réaliser la mémoire principale d'un système à microprocesseur sont des mémoires à semi-conducteur. On a vu que dans ce type de mémoire, on accède directement à n'importe quelle information dont on connaît l'adresse et que le temps mis pour obtenir cette information ne dépend pas de l'adresse. On dira que l'accès à une telle mémoire

est aléatoire ou direct. A l'inverse, pour accéder à une information sur bande magnétique, il faut dérouler la bande en repérant tous les enregistrements jusqu'à ce que l'on trouve celui que l'on désire. On dit alors que l'accès à l'information est séquentiel. Le temps d'accès est variable selon la position de l'information recherchée. L'accès peut encore être semi-séquentiel : combinaison des accès direct et séquentiel.

Pour un disque magnétique par exemple l'accès à la piste est direct, puis l'accès au secteur est séquentiel.

4.3. Différents types de mémoire

4.3.1. 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

4.3.2. Les mémoires mortes (ROM)

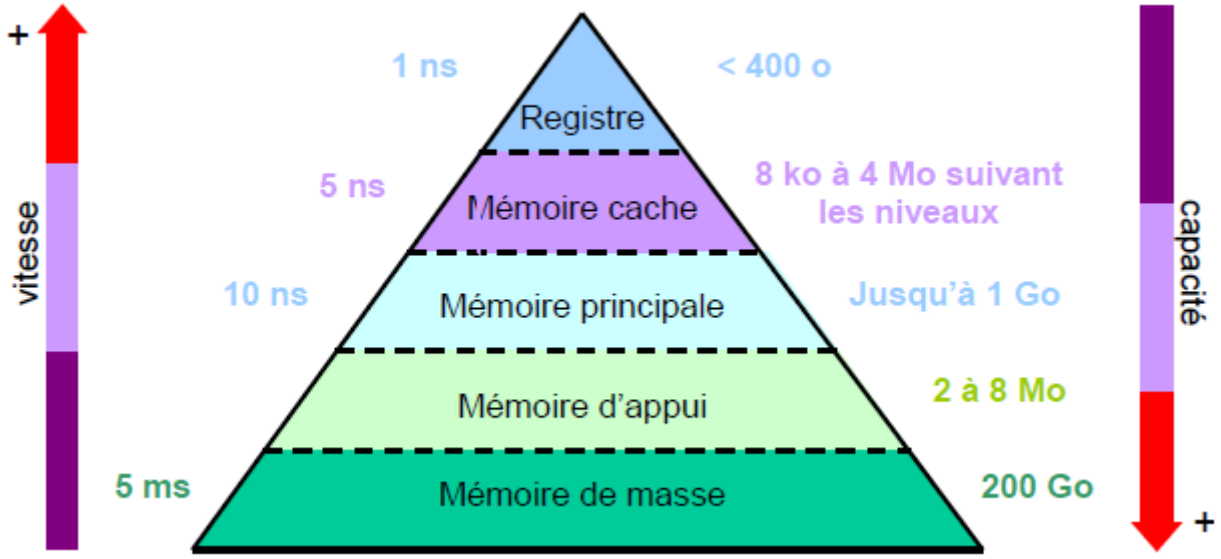
Pour certaines applications, il est nécessaire de pouvoir conserver des informations de façon permanente même lorsque l'alimentation électrique est interrompue. On utilise alors des mémoires mortes ou mémoires à lecture seule (ROM : Read Only Memory). Ces mémoires sont non volatiles. Ces mémoires, contrairement aux RAM, ne peuvent être que lue. L'inscription en mémoire des données restent 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

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 mémoire 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 actuel 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 définie 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 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