

Introduction à l'informatique

Table des matières

1. Histoire de l'informatique.....	3
1.1. Automatiser les calculs.....	3
1.1.1. La naissance du nombre et du calcul.....	3
1.1.2. Du boulier aux machines à calculer mécaniques.....	3
1.1.3. L'automatisation du travail.....	4
1.1.4. Des machines programmables à cartes perforées.....	5
1.1.5. Les grands développements scientifiques.....	7
1.1.6. Des machines électromécaniques binaires.....	12
1.2. La 1ère génération d'ordinateurs : le passage à l'électronique.....	13
1.3. La 2ème génération d'ordinateurs : l'ère du transistor.....	16
1.4. La 3ème génération d'ordinateurs : l'ère des circuits intégrés.....	18
1.5. La 4ème génération d'ordinateurs : l'ère des microprocesseurs.....	19
2. Les micro-ordinateurs modernes.....	23
2.1. Introduction à la notion d'ordinateur.....	23
2.1.1. Présentation de l'ordinateur.....	23
2.1.2. Types d'ordinateurs.....	23
2.1.3. Constitution de l'ordinateur.....	24
2.2. La carte mère.....	25
2.2.1. Présentation de la carte-mère.....	25
2.2.2. Facteur d'encombrement d'une carte-mère.....	26
2.2.3. Le chipset.....	27
2.2.4. L'horloge et la pile du CMOS.....	27
2.2.5. Le BIOS.....	28
2.2.6. Le processeur.....	28
2.2.7. La mémoire-cache.....	28
2.2.8. La mémoire vive.....	29
2.2.9. Les connecteurs d'extension.....	29
2.2.10. Le bus système.....	29
2.3. Le processeur.....	29
2.3.1. Qu'est-ce qu'un processeur?.....	29
2.3.2. A quoi ressemble une instruction?.....	30
2.3.3. Les registres.....	30
2.3.4. Les signaux de commande.....	30
2.3.5. Qu'est-ce qu'un microprocesseur ?.....	31
2.3.6. Le parallélisme.....	31
2.3.7. Le pipelining.....	31
2.3.8. L'architecture CISC.....	32
2.3.9. L'architecture RISC.....	32
2.3.10. CISC ou RISC?.....	32
2.4. La mémoire.....	32
2.4.1. Rôle de la mémoire vive (RAM).....	32
2.4.2. Fonctionnement de la mémoire vive.....	33

2.4.3. La correction d'erreurs.....	34
2.4.4. La mémoire morte (ROM).....	34
2.4.5. Les types de ROM.....	35
2.5. Les périphériques externes.....	36
2.5.1. Les ports série.....	36
2.5.2. Les ports parallèle.....	37
2.5.3. Les ports USB.....	38
2.5.4. Le port FireWire.....	39
2.6. Le système d'affichage.....	40
2.6.1. Le moniteur.....	40
2.6.2. La carte graphique.....	41
2.7. Les disques durs.....	43
2.7.1. Le fonctionnement interne.....	43
2.7.2. La lecture et l'écriture.....	44
2.7.3. Le mode bloc des disques durs.....	44
2.7.4. Le mode 32 bits des disques durs.....	45
2.7.5. L'interface SCSI.....	46
2.7.6. Les caractéristiques du disque.....	46
2.8. Les disques compacts.....	47
2.8.1. La composition d'un CD-ROM.....	47
2.8.2. Le lecteur de CD-ROM.....	47
2.8.3. Le DVD-ROM.....	48
2.8.4. Les zones.....	48
3. Le système d'exploitation.....	49
3.1. Description.....	49
3.1.1. Systèmes multi-tâches.....	50
3.2. Systèmes multi-processeurs.....	50
3.3. Les types de systèmes d'exploitation.....	50
3.4. Organisation des données persistantes.....	51
3.4.1. Les fichiers.....	51
3.4.2. Les répertoires.....	51



1. Histoire de l'informatique

1.1. Automatiser les calculs

L'histoire des ordinateurs est étroitement liée aux découvertes théoriques dans le domaine des mathématiques et de la logique et aux développements technologiques. L'histoire de l'ordinateur est également marquée par la volonté de l'homme d'automatiser les calculs afin de les rendre plus précis tout en accélérant cette tâche fastidieuse. Cette volonté va de pair avec celle de traiter l'information pour la communiquer et la contrôler. D'ailleurs, plus on progressera dans l'automatisation des opérations arithmétiques et logiques, plus grande sera la nécessité de trouver des moyens sophistiqués pour communiquer avec la machine, afin de lui donner les instructions nécessaires pour qu'elle effectue ces opérations. Alors que les premières machines à calculer pouvaient tenir dans la main de l'homme, les premiers ordinateurs étaient des monstres mécaniques et électriques qui occupaient des pièces entières d'un immeuble. On assiste aujourd'hui à un retour à l'échelle humaine avec les petits ordinateurs personnels, grâce au développement technologique qui est allé dans le sens de la miniaturisation et de la plus grande puissance de calcul. Nous verrons dans ce qui suit, les grandes lignes de cette évolution.

1.1.1. La naissance du nombre et du calcul

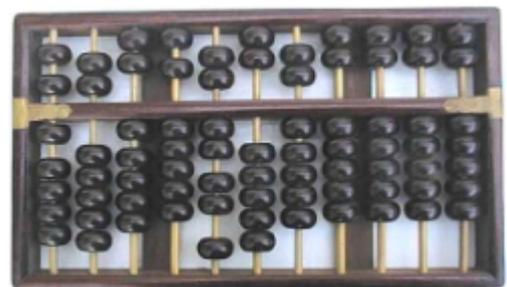
Dans l'antiquité, l'homme compte avec des grains de blé ou des cailloux. La grosseur du grain ou du caillou est proportionnelle à la quantité d'objets qu'on veut représenter : plus le caillou est gros, plus il y a d'objets. L'origine latine du mot calcul signifie d'ailleurs petit caillou ou grain.

Antiquité : 3000 ans av. JC, on écrit les chiffres avec un stylet sur une tablette de bois recouverte de sable. Le système de numération babylonien est à base 60 sans zéro. Les égyptiens écrivent les chiffres avec des barres : III = 3. Les Grecs écrivent les nombres avec les lettres de l'alphabet. Les Hébreux et les Arabes adaptent ce système à leur langue. Les Romains écrivent les chiffres avec des lettres majuscules représentant les doigts de la main : III = 3, V = 5, X = 10, C = 100. Il était extrêmement difficile de faire des calculs avec ces chiffres romains.

4ème siècle : Les indiens inventent le zéro. Zéro signifie rien en sanskrit. Les chiffres arabes de 0 à 9 apparaissent aux Indes vers le 5e siècle. Il se répand grâce aux voyages de savants et de papes. Ce système de numération est définitivement fixé en 1440 avec l'invention de l'imprimerie. Il permet de faire des calculs faciles et entraîne l'invention de l'arithmétique.

1.1.2. Du boulier aux machines à calculer mécaniques

De la tablette de sable, on est passé à l'abacus (ou abaque) pour calculer, dont le nom d'origine sémite signifie d'ailleurs poussière rappelant ainsi cette tablette. L'abacus d'origine babylonienne, est un boulier qui contient des rangées de boules qui se déplacent sur un axe; la position des boules renvoie à un nombre. On peut ainsi additionner, soustraire, multiplier et diviser; l'utilisateur possède le plein contrôle de la machine et il peut voir tout ce qui se passe. L'abaque est utilisé en Chine sous le nom de suan pan, en Russie sous le nom de tshoty et au Japon sous le nom de soroban. Mais il faut attendre le XVIIIe siècle avant de voir renaître l'ambition d'automatiser les calculs.



1617 : John Napier (1550-1617), un écossais, inventeur des logarithmes, émet la théorie que la multiplication est une suite d'additions et la division une suite de soustractions. Cela ouvre la porte au calcul mécanisé car il suffit de répéter des additions pour multiplier ou des soustractions pour diviser.

1624 : Wilhelm Schikard (1592-1635) est ce professeur allemand de l'Université de Heidelberg qui, s'inspirant des mécanismes d'horlogerie, a l'idée de construire une machine à calculer en utilisant les roues dentelées d'horloge comme engrenages. Mais il ne la construira jamais.

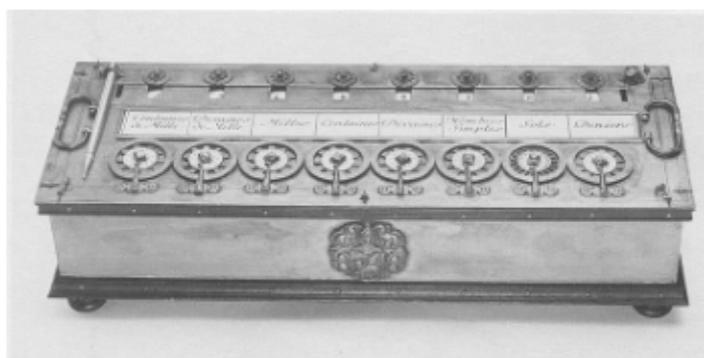


Fig. 1.2: Pascaline

1642 : Le philosophe et mathématicien français Pascal (1623-1662) présente à Paris sa Pascaline, machine à additionner et à soustraire qui fait appel à un mécanisme d'horlogerie : huit roues de neuf dents. La première roue tourne de neuf crans puis elle reprend sa position initiale pendant que la seconde roue avance d'un cran à la fois et ainsi de suite. On pouvait y entrer deux nombres à la fois en bougeant des roues avec un stylos et le résultat, pas toujours exact, apparaissait sur des cylindres rotatifs; le mécanisme de l'addition est dès lors automatisé et il s'effectue à l'intérieur de la machine.

1672 : Le philosophe et mathématicien allemand Leibniz (1646-1716) a passé 4 ans à Paris à étudier les mathématiques et, en particulier, les théories de Descartes et de Pascal. Il est fasciné par la machine à calculer mécanique de Pascal; il va d'ailleurs la perfectionner en automatisant les opérations de multiplication et de division. La machine de Leibniz de même que celle de Pascal sont des machines pas à pas. Les machines de Pascal et de Leibniz utilisent le système décimal. C'est à Leibniz que l'on doit également l'idée de concevoir une calculatrice en binaire. Mais la technologie de l'époque n'en permet pas la réalisation.

1888 : William Burroughs (1857-1898) obtient le premier brevet américain pour une machine à additionner. En 1892, il fabrique aux états-Unis, la première machine à calculer de l'ère moderne distribuée commercialement par une compagnie qui porte son nom. Celle-ci deviendra, à un moment donné, un géant des machines de bureaux et des ordinateurs. La Burroughs est une machine essentiellement mécanique qui soustrait, additionne et imprime les résultats, mais elle est non programmable et si peu fiable que Burroughs dût reprendre les machines vendues. Mais il n'abandonne pas et perfectionne sa machine qui acquerra avant la fin du siècle, la fiabilité qui lui manquait.

1.1.3. L'automatisation du travail

Si l'être humain a, depuis des siècles, l'ambition de calculer puis d'automatiser les calculs, à l'aube de la révolution industrielle du XIXe siècle, il a l'ambition d'automatiser les tâches.

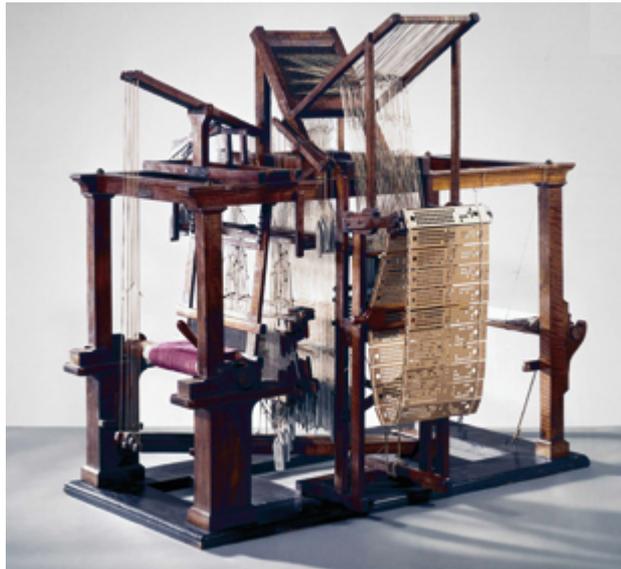


Fig. 1.3: Métier à tisser de Joseph-Marie Jacquard

1728 : L'Anglais Falcon a imaginé un système de plaquettes de bois perforées pour contrôler les opérations d'un métier à tisser. Cette idée servira également pour les pianos et les orgues mécaniques.

1805 : C'est à Joseph-Marie Jacquard (1752-1834), à peine un siècle plus tard, que l'on reconnaît généralement l'invention du métier à tisser automatisé. S'inspirant de l'idée de Falcon, il invente un métier contrôlé par une série de cartes perforées en boucle continue, comportant toutes les instructions du motif à tisser répétées plusieurs fois sur une même longueur de tissu. On peut assez facilement changer le motif à tisser en changeant la série de cartes. Sans s'en douter, il vient d'inventer la carte perforée qui servira plus tard à donner des instructions aux ordinateurs. Mais les tisserands de l'époque sentent leur travail menacé et se révoltent en détruisant les métiers de Jacquard. Mais l'idée va néanmoins faire son chemin.

1.1.4. Des machines programmables à cartes perforées

Bâtissant sur l'idée des cartes de Jacquard pour contrôler une mécanique, l'Anglais Babbage utilise le même principe dans les machines à calculer qu'il a imaginées.

1822 : L'Anglais Babbage reçoit en 1821, la première médaille d'or de la British Astronomical Society pour un article intitulé *Observations on the Application of Machinery to the Computation of Mathematical Tables*. En 1822, il invente une machine mécanique à calculer la position des planètes et du soleil. Elle pèse deux tonnes. Son dispositif se veut capable de calculer des tables de fonctions mathématiques y compris les logarithmes; cependant, la mécanique du temps n'est pas à la hauteur de son génie et sa découverte a été plus importante sur le plan de la théorie. Il a somme toute donné naissance au concept de registre (dispositif de mémorisation temporaire) : la suite d'instructions pour effectuer les opérations était portée sur une bande perforée dans le style des cartes perforées de Jacquard. Il suffisait de manipuler la machine pour obtenir les résultats qui pouvaient même être gravés sur une plaque de cuivre à partir de laquelle on pouvait les imprimer. Deux Suédois, père et fils, Per Georg et Edvard Scheuz, ont réussi par la suite à produire une machine semblable plus simple mais fonctionnelle.

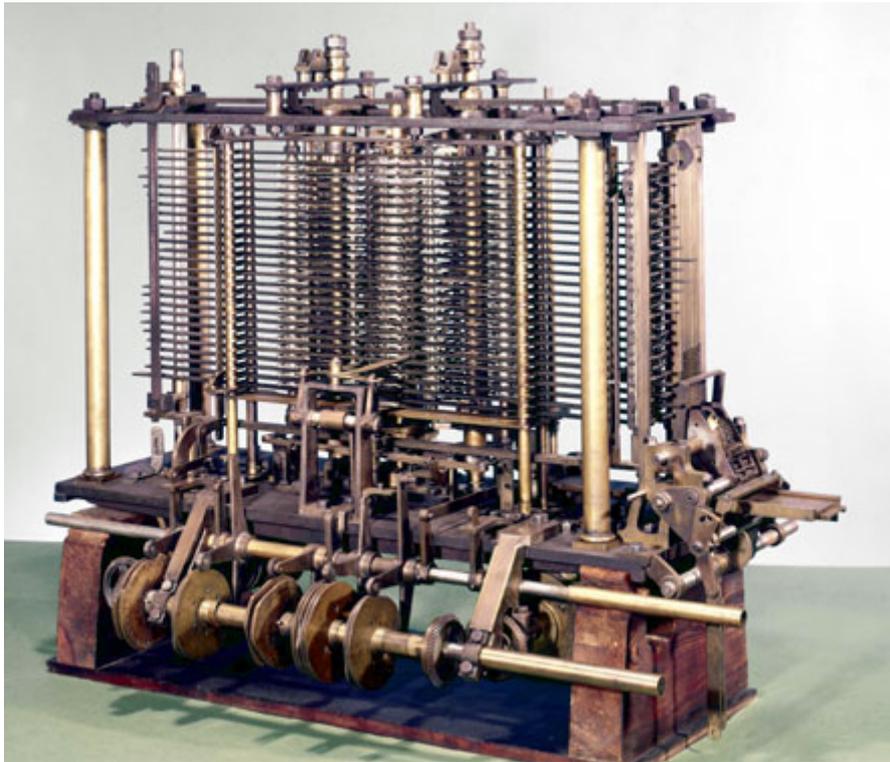


Fig. 1.4: Première machine de Babbage

1832 : La première machine de Babbage ne pouvait produire que des tables et ne permettait qu'une lecture séquentielle des données. Babbage laisse en plan cette machine pour en inventer une autre avec l'appui du gouvernement anglais qui s'est par la suite retiré : la machine analytique, une machine à calculer à usage général et entièrement mécanique ne sera jamais terminée. Elle peut effectuer toutes sortes de calculs; on peut y emmagasiner des données qu'elle peut lire dans n'importe quel ordre; on peut aussi y emmagasiner les nombres générés en cours de calcul. Alors que la différence machine avait des instructions fixes, la machine analytique pouvait recevoir différentes instructions. Dans son design, Babbage avait aussi prévu l'équivalent d'une unité de contrôle pour faire effectuer les instructions dans le bon ordre et un moulin, mill, comme il l'appelait, sorte d'unité arithmétique et logique qui effectuait les opérations nécessaires aux calculs qui devaient être précis à 50 décimales près; 1 000 nombres devaient pouvoir y être emmagasinés. Babbage n'a jamais terminé la construction de sa machine qui devait comprendre 50 000 pièces mobiles dans sa version la plus simple, car il avait en tête un troisième type de machine de laquelle il n'a laissé aucune description. Mais 150 ans plus tard, Allan Bromlym, un chercheur australien, la construisit et l'exposa au Musée des sciences à Londres; elle pèse 3 tonnes et comporte 4 000 pièces en bronze et en acier. L'un des mérites de Babbage a été d'imaginer un système de cartes perforées pour emmagasiner les instructions données à la machine, une carte par instruction; d'autres cartes devaient contenir les adresses du registre, les données devaient être enregistrées; ces cartes devaient être aussi réutilisables. Babbage n'avait cependant pas imaginé qu'une instruction puisse contenir à la fois une opération à effectuer et une adresse; il n'avait pas imaginé la programmation comme on la connaît aujourd'hui ni l'idée de conserver un programme d'instructions en mémoire.

1890 : Le statisticien Herman Hollerith (1860-1929) est responsable du recensement de la population américaine au Bureau of the Census des états-Unis. Il a l'idée de coder sur des cartes perforées, l'information provenant du recensement, ce qui permet d'en faire l'analyse en des temps records pour l'époque. Les cartes perforées ne servent pas à programmer la machine mais à trier et à

conserver les données. Les cartes sont lues par un lecteur électrique; un trou laisse passer le courant et l'absence de trou l'arrête. La machine de Hollerith est une machine mécanique mais qui, pour la première fois dans le domaine des calculateurs, a recours à l'électricité pour incrémenter les compteurs. Le système utilisé pour le recensement de 1890 est un succès immédiat et il se répand dans plusieurs pays; on l'utilise également pour effectuer des opérations comptables. En 1900, Hollerith invente un dispositif automatique pour alimenter sa machine de cartes perforées. Hollerith est à l'origine d'une compagnie, la Tabulating Machine Company, qui fusionnée à deux autres, devient en 1924 la International Business Machines Corporation (IBM) sous la gouverne de Thomas J. Watson.

A la suite d'une dispute concernant les frais de location de la machine d'Hollerith, le Bureau of Census demande à l'un de ses ingénieurs, James Power, de développer une nouvelle machine. Elle sera dotée de cartes perforées mais à lecture mécanique. Mais Power quitte aussi le Bureau et forme sa propre compagnie, The Powers Accounting Machine Company qui deviendra en 1927, la Remington Rand Corporation et suite à des regroupements, la SperryRand Corporation, en 1955, l'un des grands compétiteurs de IBM dans les années 1950.

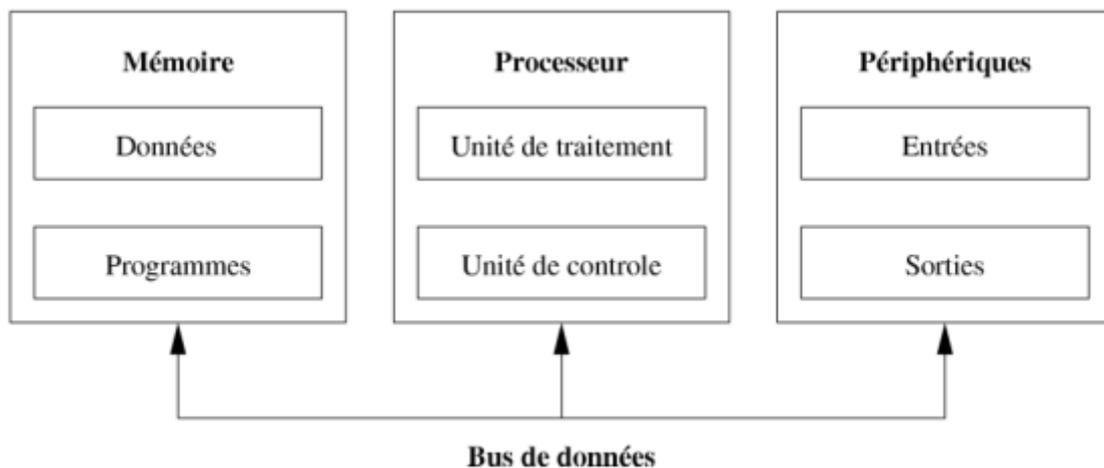


Fig. 1.5: L'architecture de Von Neumann

Toutes ces machines traitent des données et fournissent un résultat. Dans les plus complexes d'entre elles, avant même l'utilisation de l'électricité, on voit déjà poindre les principes de base de l'informatique moderne. Ces principes théoriques ont été énoncés en 1840 par Ada Lovelace. Ils ont été formalisés à l'ère des calculateurs électriques par John Von Neumann en 1946. Selon ces principes, les quatre éléments fondamentaux d'un ordinateur sont :

- la mémoire qui stocke les données et les programmes;
- l'unité logique qui traite et modifie les données;
- l'unité de contrôle qui organise le fonctionnement interne de la machine;
- les différents organes d'entrée et de sortie (roues dentées, cartes perforées, claviers, écrans, imprimantes etc) .

1.1.5. Les grands développements scientifiques

La naissance de l'informatique moderne doit beaucoup au développement des mathématiques et de la logique.

330 av JC : Aristote propose les premiers éléments de la logique.

1614 : John Neper montre la multiplication et la division peuvent être ramenées à deux opérations très simples effectuées en séquence : l'addition et la soustraction.

1697 : Gottfried Leibnitz introduit le calcul binaire en Europe. Au lieu d'utiliser 9 chiffres, on n'en utilise plus que 2 : le 0 et le 1. Par exemple, 9 s'écrit 101 et 4 s'écrit 10. Compter en binaire revient à énoncer 0, 1, 10, 11, 100, 101, 110, 111. Les opérations arithmétiques simples telles que l'addition, la soustraction et la multiplication sont faciles à effectuer en binaire. L'addition en binaire : L'addition en binaire se fait avec les mêmes règles qu'en décimal : On commence à additionner les bits de poids faibles (les bits de droite) puis on a des retenues lorsque la somme de deux bits de mêmes poids dépasse la valeur de l'unité la plus grande (dans le cas du binaire: 1), cette retenue est reportée sur le bit de poids plus fort suivant. Par exemple:

$$\begin{array}{r} 01101 \\ + 01110 \\ \hline 11011 \end{array}$$

La multiplication en binaire : La table de multiplication en binaire est très simple : $0 \times 0 = 0$, $0 \times 1 = 0$, $1 \times 0 = 0$, $1 \times 1 = 1$.

La multiplication se fait en formant un produit partiel pour chaque digit du multiplieur (seul les bits non nuls donneront un résultat non nul). Lorsque le bit du multiplieur est nul, le produit partiel est nul, lorsqu'il vaut un, le produit partiel est constitué du multiplicande décalé du nombre de positions égal au poids du bit du multiplieur.

Conversion du binaire en décimal : Exemple avec 49 :

$$\begin{aligned} 49/2 &= 24, \text{ reste } 1 \\ 24/2 &= 12, \text{ reste } 0 \\ 12/2 &= 6, \text{ reste } 0 \\ 6/2 &= 3, \text{ reste } 0 \\ 3/2 &= 1, \text{ reste } 1 \\ 1/2 &= 0, \text{ reste } 1 \end{aligned}$$

donc, $49 = \%11\ 0001$

Conversion du décimal en binaire : Exemple avec $\%1\ 1011$:

$$1 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 27.$$

Exercice : Convertir $\%1\ 0010$ en décimal et 43 en binaire.

1854 : George Boole (1815-1864), professeur au Queen's College en Irlande, définit une nouvelle forme de logique, fondée sur trois opérateurs ET, OU, NON, dans un ouvrage intitulé An Investigation into the Laws of Thought. Il fait ainsi passer la logique de la philosophie aux mathématiques. Une proposition vraie se voit affecter la valeur 1 tandis qu'une proposition fausse, la valeur 0. On peut ainsi effectuer des opérations sur des propositions unies par l'un des trois opérateurs logiques. Cette forme de logique s'accorde bien avec le calcul binaire et elle sera au centre des futurs ordinateurs.

A	B	$A \times B$
0	0	0
0	1	0
1	0	0
1	1	1

A	B	$A + B$
0	0	0
0	1	1
1	0	1
1	1	1

A	\bar{A}
0	1
1	0

Tab. 1.1: Tables de vérité des opérateurs booléens ET, OU et NON

Exercice : Ecrire la table de vérité de l'opérateur ou exclusif ?. On définit ce nouvel opérateur comme suit :

$$A \oplus B = A \times \bar{B} + \bar{A} \times B$$

1879 : L'algèbre booléenne a pris sa forme actuelle en 1879, grâce aux travaux de Gottlob Frege qui a proposé une forme de logique qui rompt avec la logique aristotélicienne dominante, une logique des symboles abstraits. Frege propose un système de notation pour le raisonnement mécanique. C'est le philosophe anglais Bertrand Russell qui a principalement fait connaître les travaux de Frege.

1900 : David Hilbert (1862-1943) est un mathématicien allemand qui a publié en 1899, un ouvrage sur les fondements de la géométrie. Il présente à la Conférence internationale des mathématiques de Paris de l'année 1900, une proposition d'agenda de recherche qui comprend une liste des 33 problèmes les plus urgents à résoudre en mathématiques pour le prochain siècle. Le 33e de ces problèmes questionne l'existence ou la non existence d'un algorithme pour déclarer vraie ou fausse une proposition logique, dans un système de logique assez puissant pour représenter les nombres naturels. Ce problème va attirer l'attention de bien des mathématiciens et logiciens qui suivront.

1903 : Bertrand Russell est un jeune mathématicien et philosophe anglais du début du XXe siècle. Il publie en 1903 l'ouvrage majeur *The Principles of Mathematics* qui propose de nouveaux fondements à la logique et la théorie des ensembles : les lois de base de l'arithmétique sont réductibles à des propositions de logique élémentaire. Puis, en collaboration avec son tuteur de l'époque, North Whitehead (1861-1947), il entreprend d'appliquer ses théories aux mathématiques.

<i>Associativité</i>	$(A \times B) \times C = A \times (B \times C)$ $(A + B) + C = A + (B + C)$
<i>Absorption</i>	$A \times (A + B) = A$ $A + A \times B = A$
<i>Commutativité</i>	$A \times B = B \times A$ $A + B = B + A$
<i>Distributivité</i>	$A + (B \times C) = (A + B) \times (A + C)$ $A \times (B + C) = A \times B + A \times C$
<i>Idempotence</i>	$A \times A = A$ $A + A = A$
<i>Element neutre</i>	$1 \times A = A$ $0 + A = A$
<i>Inversion</i>	$A \times \bar{A} = 0$ $A + \bar{A} = 1$
<i>Element absorbant</i>	$0 \times A = 0$ $1 + A = 1$
<i>Théorème de Morgan</i>	$\overline{(A + B)} = \bar{A} \times \bar{B}$ $\overline{(A \times B)} = \bar{A} + \bar{B}$

Tab. 1.2: Opérations de composition dans une algèbre booléenne

Cela donnera une somme publiée en trois volumes en 1910, 1912 et 1913, la Principia Mathematica. On reconnaît à Russell le mérite d’avoir conçu sur papier, une machine qui ressemble étrangement à un ordinateur.

1931 : Kurt Gödel (1906-1978) est ce logicien autrichien qui publie en 1931, un théorème qui s’est avéré majeur dans l’histoire de la logique et des mathématiques, le théorème de l’incomplétude : il y a des problèmes en mathématiques et en logique impossible à résoudre, il en est ainsi du problème 33 de Hilbert. Dans de tels systèmes, propose Gödel, il y a des propositions qui ne peuvent être ni démontrées ni rejetées. Ces propositions ne sont pas indéterminées, elles sont effectivement ou vraies ou fausses, mais c’est juste qu’on ne peut pas dire dans quel sens elles vont.



Fig. 1.6: Alan Turing

1937 : Alan Turing (1912-1954) connaît bien les travaux de Russell et ceux de Gödel. Il s'en inspire dans un article célèbre intitulé *On Computable Numbers, with an Application to the Entscheidungs problem*, dans lequel il propose le modèle théorique d'une machine universelle capable de réaliser n'importe quel calcul mathématique. Son modèle est fondé sur l'idée que tout problème humain peut être résolu par une suite d'algorithmes comme ceux définis par Ada Lovelace. Il en fait la démonstration avec sa machine théorique, dite « Machine de Turing ». Celle-ci serait dotée d'une unité logique de traitement et d'une unité de contrôle qui puisse entraîner plusieurs rubans. Chacun de ces rubans contient des informations codées comme une suite de 0 et de 1. Le premier ruban représente les données, et le second représente le programme. Si on change le programme, on change la manière dont le système de contrôle modifie les données, ce qui lui donne son caractère universel. La machine de Turing démontre également que certains problèmes de mathématiques et de logique sont insolubles : bien qu'ils soient clairement formulés, il n'est pas possible de trouver des solutions avec sa machine. Mais ces problèmes ne sont pas plus solubles par l'esprit humain. Il rejoint ainsi le théorème d'incomplétude de Kurt Gödel (1931). Néanmoins, la machine de Turing est si simple mais si puissante qu'elle va inspirer fortement les constructeurs des premiers ordinateurs.

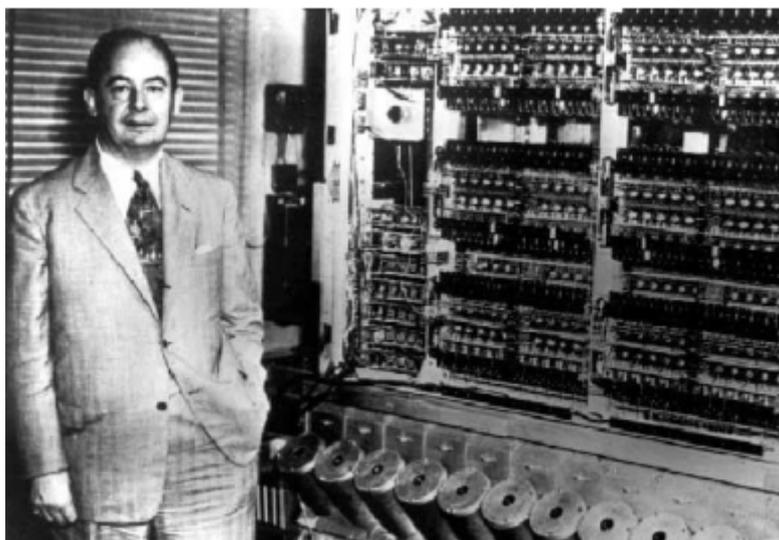


Fig. 1.7: John Von Neumann

1939 : A cette époque, Shannon est étudiant de maîtrise en génie (électricité) au M.I.T. à Boston. Dans son mémoire, il propose une idée géniale : on peut décrire les états d'un relais électromagnétique en utilisant la logique booléenne.

1946 : John Von Neumann (1903-1957) est un mathématicien d'origine hongroise qui a travaillé aux états-Unis à l'Institute for Advanced Study de Princeton University. Il est le premier, au XXe siècle, à publier un article sur la possibilité de loger un programme dans la mémoire d'un ordinateur. Au siècle précédent, Babbage avait évoqué cette idée mais sans l'appliquer à sa Machine analytique.

1.1.6. Des machines électromécaniques binaires

1930 : Vannevar Bush du M.I.T. et ses collègues conçoivent, en 1925, un prototype de machine analogique pour calculer des équations différentielles. Elle sera construite en 1930, sous le nom de The Differential Analyser. Elle doit servir à calculer les trajectoires de l'artillerie durant la Deuxième guerre mondiale.

1931 : Konrad Zuse, un Allemand, s'inspirant des idées de Leibniz et s'appuyant sur la logique de Boole, réalise la Z1, une machine mécanique binaire. Il a eu accès à la description de la machine analytique de Babbage qu'en a faite Lady Lovelace. Les programmes ne sont pas intégrés à la machine. De 1935 à 1945, il développera d'autres prototypes : la Z2 était une machine essentiellement mécanique alors que la Z3 était un calculateur à relais capable d'emmagasiner des nombres binaires de 64 décimales flottantes; la Z3 était une machine programmable, mais elle n'a pas survécu au bombardement de Berlin. En 1940, Zuse veut faire de la Z4 une machine à tubes électroniques mais les fonds lui font défaut car Hitler est sûr de remporter la deuxième guerre mondiale; la Z4 sera donc une machine à relais. Les machines de Zuse sont à la limite des technologies de l'avant-guerre.

1939 : On doit le premier calculateur binaire à l'Américain George R. Stibitz des Laboratoires Bell. Il se nomme le Model 1 Relay Computer ou Complex Number Calculator. Il s'agit d'une machine à relais téléphonique ON/OFF, bricolée, en une fin de semaine, à partir de rebuts.

1940 : C'est aux Anglais que l'on doit ce qui pourrait être le premier ordinateur opérationnel du monde, le Robinson. Il est fait de relais électromagnétiques et de tubes à vide. Il n'est pas programmable. Il est issu d'un effort de guerre du gouvernement britannique impliquant 10 000 personnes, pour construire un ordinateur; cet effort est connu sous le nom de projet Ultra. Le Robinson sert à décoder les messages secrets des Allemands au cours de la Deuxième guerre mondiale, messages encodés par une machine appelée Enigma inventée par les Nazis. énigma

1941 : Arnold Fast, un mathématicien aveugle engagé par Konrad Zuse, va programmer le premier ordinateur binaire programmable du monde, le Z3.

1943 : En 1937, Howard Aiken de la Harvard University a proposé à IBM de construire un calculateur géant à partir des dispositifs mécaniques et électromécaniques utilisés pour les machines à cartes perforées. La machine a été complétée en 1943. Elle pesait 5 tonnes, comportait des accumulateurs capables de 72 décimales et elle pouvait multiplier deux nombres de 23 unités en 6 secondes. Elle était contrôlée par des instructions sur un ruban de papier perforé. Malgré les connaissances acquises depuis Babbage, elle ne possédait pas la capacité de faire des branchements conditionnels. Mais elle ajoutait à la machine de Babbage une horloge pour contrôler les séquences d'opérations, et des registres, sortes de mémoires temporaires pour enregistrer des informations.

1.2. La 1ère génération d'ordinateurs : le passage à l'électronique

«Je pense qu'il y a un marché mondial pour environ 5 ordinateurs.»

Thomas Watson, président d'IBM, 1943.

Le passage à l'électronique, dans la mise au point d'ordinateurs, s'est fait grâce à l'invention du tube à vide. Celui-ci permettait de produire un courant direct d'électrons dans un tube sous vide capable de générer deux états : ON/OFF. Le tube à vide était donc prédestiné au calcul binaire. A l'aide de "contacteurs"(interrupteurs) fermés pour "vrai" et ouverts pour "faux" il était possible d'effectuer des opérations logiques en associant le nombre "1" pour "vrai" et "0" pour "faux". Ce codage de l'information est nommé base binaire. C'est avec ce codage que fonctionnent les ordinateurs modernes.

Le bit : Bit signifie "binary digit", c'est-à-dire 0 ou 1 en numérotation binaire. C'est la plus petite unité d'information manipulable par une machine numérique. Il est possible de représenter physiquement cette information binaire :

- par un signal électrique ou magnétique, qui, lorsqu'elle atteint une certaine valeur, correspond à la valeur 1.
- par des aspérités géométriques dans une surface.
- grâce à des bistables, c'est-à-dire des composants électroniques qui ont deux états d'équilibre (un correspond à l'état 1, l'autre à 0)

Avec un bit il est ainsi possible d'obtenir deux états: soit 1, soit 0. 2 bits rendent possible l'obtention de quatre états différents (2x2) : 00, 01, 10, 11. Avec 3 bits il est possible d'obtenir huit états différents (2x2x2) : 000, 001, 010, 011, 100, 101, 110, 111. Pour un groupe de n bits, il est possible de représenter 2^n valeurs.

L'octet : L'octet (en anglais byte) est une unité d'information composée de 8 bits. Il permet de stocker un caractère, un chiffre... Ce regroupement de nombres par série de 8 permet une lisibilité plus grande, au même titre que l'on apprécie, en base décimale, de regrouper les nombres par trois pour pouvoir distinguer les milliers. Par exemple le nombre 1 256 245 est plus lisible que 1256245.

Le mot : Une unité d'information composée de 16 bits est généralement appelée mot (en anglais word). Une unité d'information de 32 bits de longueur est appelée double mot (en anglais double word, d'où l'appellation dword). Pour un octet, le plus petit nombre est 0 (représenté par huit zéros 00000000), le plus grand est 255 (représenté par huit chiffre "un" 11111111), ce qui représente 256 possibilités de valeurs différentes.

KiloOctets, MégaOctets :

- Un kilo-octet (Kio) ne vaut pas 1000 octets mais 1024 octets (2^{10} octets).
- Un méga-octet (Mio) vaut 1024 Kio = 1 048 576 octets (2^{20} octets).
- Un giga-octet (Gio) vaut 1024 Mio = 1 073 741 824 octets (2^{30} octets).
- Un téra-octet (Tio) vaut 1024 Gio = 1 099 511 627 776 octets (2^{40} octets).

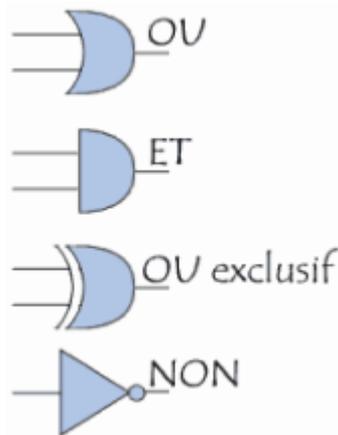


Fig. 1.8: Différentes portes logiques

Codage d'informations : La mémoire de l'ordinateur conserve toutes les données sous forme numérique. Il n'existe pas de méthode pour stocker directement des informations quelconques. Il faut définir un codage. Par exemple, pour les caractères, on a défini le code ASCII. Chaque caractère possède donc son équivalent en code numérique : c'est le code ASCII (American Standard Code for Information Interchange traduit "Code Américain Standard pour l'Échange d'Informations"). Le code ASCII de base représentait les caractères sur 7 bits (c'est-à-dire 128 caractères possibles, de 0 à 127). Le code ASCII a été mis au point pour la langue anglaise, il ne contient donc pas de caractères accentués, ni de caractères spécifiques à une langue. Pour coder ce type de caractère il faut recourir à un autre code. Le code ASCII a donc été étendu à 8 bits (un octet) pour pouvoir coder plus de caractères (on parle d'ailleurs de code ASCII étendu...). Ce code attribue les valeurs 0 à 255 (donc codées sur 8 bits, soit 1 octet) aux lettres majuscules et minuscules, aux chiffres, aux marques de ponctuation et aux autres symboles (caractères accentués dans le cas du code iso-latin1).

1904 : John Flemming invente le premier tube à vide. Il s'agit d'un diode pouvant prendre deux états ON/OFF.

1907 : A partir de la diode de John Flemming, L. De Forest met au point la « triode », un amplificateur d'intensité électrique. La première « lampe » réellement utilisable a été mise au point en France par H. Abraham et utilisée par l'ensemble des Alliés pendant la première guerre mondiale, d'où son nom : Lampe TM (pour Télégraphie Militaire).

1919 : Invention du basculeur d'Eccles et Jordan à partir de deux triodes. Plus connu maintenant sous le nom de flip-flop ou circuit bi-stable.

1939 : Le mathématicien et physicien John Vincent Atanasoff, applique l'idée d'utiliser des tubes à vide pour faire du calcul numérique, idée soutenue par un ami de Zuse dans une thèse de doctorat. Il construit, avec un de ses étudiants du Iowa State College, Clifford Berry, une machine capable de résoudre des équations complexes en physique, le ABC (Atanasoff Berry Computer).

Exercice : Avec des composants capables de changer leur état binaire, il est possible de fabriquer des portes logiques pour effectuer des opérations logiques. Grâce au code des nombres en binaire, les opérations logiques peuvent aussi être utilisées pour faire du calcul. Considérons par exemple les fonctions logiques $S = A+B$ et $R = A \times B$. Donner les tables de vérité des fonctions S et R en fonction de A et de B. Expliquer dans quelle mesure ces fonctions permettent de calculer la somme S de deux chiffres binaires en gérant la retenue R. De tels additionneurs élémentaires peuvent être mis en série pour réaliser les additions de nombres plus grands.

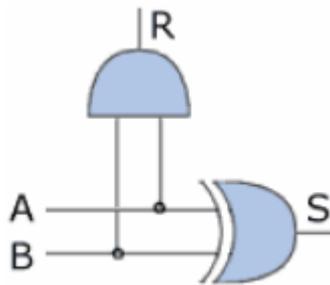


Fig. 1.9: Un additionneur avec deux portes logiques

1943 : Le Colossus est mis au point en Angleterre, durant la Deuxième guerre mondiale, dans le cadre du projet Ultra. Les Allemands ayant perfectionné leur système d'encodage de messages, le Robinson ne suffit plus à la tâche. Le Colossus doit prendre la relève. On doit sa mise au point à Max M. A. Newman, professeur de mathématiques à la Cambridge University, à T. H. Flowers, ingénieur au Post Office Research Station et à Alan Turing, un Hongrois installé en Angleterre depuis 1936. Ce dernier inventa une machine universelle sur papier appelée la machine de Turing qui jette les bases de ce que sera l'ordinateur moderne. C'est également Turing qui adopte l'algorithme pour le calcul de certaines fonctions. Le Colossus, qui comporte 2 000 tubes à vide, peut lire des rubans perforés à la vitesse de 5 000 caractères à la seconde. Un seul ruban à la fois est donné à la machine. Les comparaisons sont effectuées électroniquement, ce qui est une innovation majeure pour l'époque, et les résultats sont conservés dans une mémoire de triodes thyatron remplies de gaz. Pour éliminer les erreurs, une pulsation d'horloge générée par une cellule photosensible permet de marquer la cadence des opérations. La programmation se fait à l'aide de cartes enfichables. Le Colossus II, complété en 1944, aura 2 400 lampes et sera cinq fois plus rapide que le Colossus I.

Plusieurs autres versions du Colossus verront le jour jusqu'à la fin de la guerre.

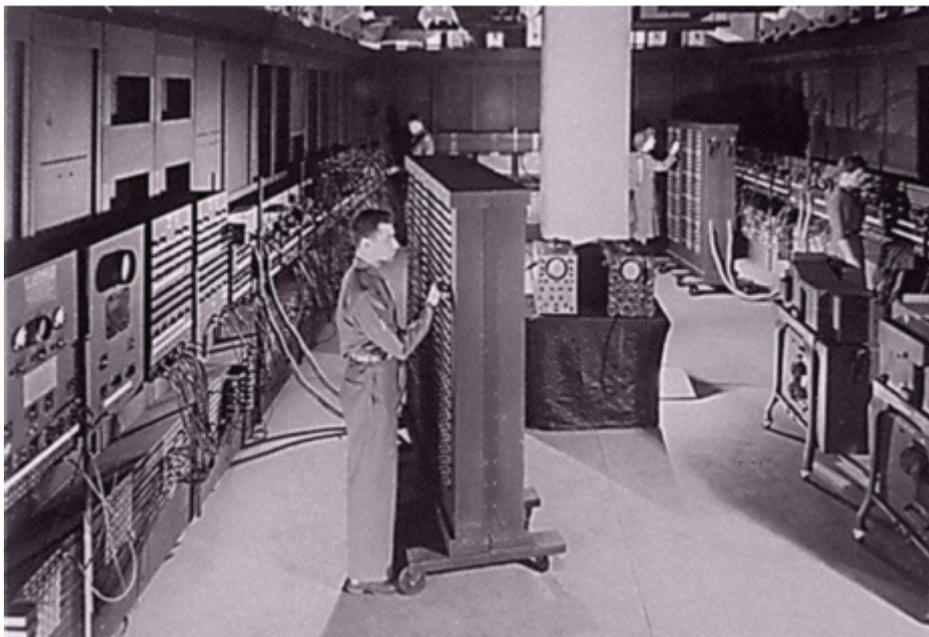


Fig. 1.10: ENIAC

1946 : A la Pennsylvania University, John Mauchly, J. Presper Eckert et leurs collègues ont entrepris la construction de l'ENIAC (Electronic Numerical Integrator And Calculator) en 1943, avec le

financement du Ballistic Research Laboratory. Cette machine était destinée aux calculs balistiques, mais c'était en réalité une machine à usage général. Elle était programmable grâce à des interconnexions enfichables. Elle était capable de 5 000 opérations arithmétiques à la seconde et était 1 000 fois plus rapide que le Harvard Mark I. Il lui fallait une puissance électrique de près de 200 KW pour fonctionner, car l'ENIAC comportait quelques 19 000 lampes. Il pesait 30 tonnes et occupait un espace de 3 000 pieds cubes soit 160 m² au sol. Sa capacité de mémoire était seulement de 20 mots de 10 unités chacun et elle ne pouvait traiter que des programmes d'instructions ne dépassant pas 300 mots. C'est durant les travaux de mise au point de l'ENIAC que John Tuckey créa le terme BIT ou binary digit, unité binaire d'information, soit 0 ou 1. Un jour, en 1947, l'ENIAC tomba en panne sans que ses constructeurs ne sussent pourquoi. Après exploration, on constata qu'un insecte s'était logé dans un relais; le technicien qui a fait la découverte s'est écrié : there is a bug in the machine. Le nom bug est resté pour désigner une erreur de matériel ou de programmation.

1946 : Le BINAC (Binary Automatic Computer) est l'oeuvre de Mauchly et de Eckert. Il s'agit en réalité de deux ordinateurs réunis qui calculent simultanément, leurs résultats étant ensuite comparés. On disait que le BINAC travaillait en temps réel car il fournissait des résultats dans l'immédiat. Le BINAC était un ordinateur extrêmement fiable.

1947 : Après le ENIAC, il y a eu le EDVAC auquel John von Neumann a travaillé. Il s'agit d'une machine binaire sérielle qui nécessite beaucoup moins de lampes que le ENIAC. Elle possède aussi une mémoire interne beaucoup plus grande (100 fois plus grande). Avec le EDVAC, l'invention de l'ordinateur numérique de première génération est à toute fin complétée. Ses plans, largement diffusés, ont donné lieu à plusieurs autres projets d'ordinateurs vers la fin des années 40 jusqu'en 1953. Mentionnons le SEAC, le SWAC, les Harvard Mark III et IV, le IAS Machine, les UNIVAC I et 1103, le WHIRLWIND et les IBM 701, 702. De 1953 à 1959, période reconnue comme étant celle de la première génération avancée d'ordinateurs, les IBM 650, 704, 705, 709, les UNIVAC II, 1103A, SS80, les Burroughs 205, 220, les NCR 120, 200, le Datamatic 1000, le RCA Bizmac ont vu le jour.

1947 : Parmi les projets d'ordinateurs de la fin des années 1940, il faut mentionner le SSEC (Selective Sequence Electronic Calculator), première machine de IBM à programme enregistré. Depuis 1944, John von Neumann travaille avec l'équipe qui a mis au point le ENIAC. Il imagine cet ordinateur à programme enregistré. La mémoire de l'ordinateur ne contient pas seulement des données mais également des instructions que l'ordinateur peut modifier pour choisir sa séquence de calcul, d'où son nom. Le SSEC était composé de 13 500 tubes à vide, 21 000 relais et il pouvait additionner 3 500 nombres de 14 décimales par seconde. Il multipliait 100 fois plus rapidement que le Mark 1.

1.3. La 2ème génération d'ordinateurs : l'ère du transistor

«Les ordinateurs du futur ne pèseront pas moins d'une tonne et demi.»

Popular Mechanics, 1949.



Fig. 1.11: Transistor

1947 : John Bardeen, Walter Brattain et William Shockley inventent le transistor aux Bell Laboratories du New Jersey, en 1947. Le transistor ressemble à un petit sandwich qui contient de la matière capable de conduire l'électricité à un voltage donné. Le transistor fait le même travail que la lampe à vide des premiers ordinateurs mais il est beaucoup plus petit, moins cher à fabriquer et beaucoup plus fiable. Cependant, en raison de problèmes de production, ce n'est qu'à la fin des années 1950 que le transistor s'est répandu.

1956 : Le premier ordinateur à base de transistors, le Leprechaun, a été construit dans les laboratoires Bell.

1959 : Le 1401 de IBM est un ordinateur à transistors très populaire. IBM en a livré plus de 10 000 à la petite et moyenne entreprise.

1960 : Le STRETCH, d'IBM, comporte 150 000 transistors et a une vitesse de traitement de 200 instructions à la seconde. C'est le plus rapide de son époque, affirme IBM C'est à l'occasion de la mise au point du STRETCH que le mot octet est employé pour la première fois. Un octet égale 8 unités binaires. Un octet désigne communément un caractère alphanumérique car chaque lettre ou chaque nombre est représenté par une série de 8 zéro ou 1. Un Kio (kilo-octet) = 1 024 octets. Un Mio (méga-octet) = 1 024 Kio. Un Gio (giga-octet) = 1 024 Mio.

1961 : Le premier système de travail à temps partagé sur ordinateur, le CTSS (Compatible Time Sharing System) a été mis au point au M.I.T. (Massachusetts Institute of Technology) à Boston, sous la direction du professeur F. Corbato. Il réunissait des ordinateurs IBM 709 et 7090. C'est un concept majeur pour l'époque, qui aura des retombées importantes : il s'agit de permettre à plusieurs usagers en même temps mais chacun à leur tour, sans qu'il s'en aperçoive, d'avoir accès à la puissance de calcul d'un gros ordinateur à partir de terminaux à distance.

1959-1964 : La multiplication des transistors sur le marché a fait qu'ils ont complètement remplacé les lampes à vide dans les ordinateurs, à partir de 1959. Les ordinateurs de deuxième génération possédaient des mémoires principales magnétiques à tores (le tore est un anneau magnétique). Quelques-uns d'entre eux possédaient comme mémoire auxiliaire des tambours magnétiques et des rubans magnétiques. Nommons le Philco 2000, les CDC 1604, 3600, les IBM 7000, 1400, les RCA 302, 501, le Honeywell 800, les UNIVAC III, 1107, les Ferranti Atlas.

1.4. La 3ème génération d'ordinateurs : l'ère des circuits intégrés

«A quoi ça peut-il bien servir?»

Ingénieur chez IBM à qui l'on présentait une puce électronique, 1968.

L'invention du transistor va vite appeler le développement d'une technologie qui permettra de rapetisser les autres composants de l'ordinateur et surtout de les intégrer, car les liaisons électriques multiples qui doivent se faire entre chaque transistor sont complexes, coûteuses à réaliser, pas assez rapides et peu fiables, en tout cas pas assez pour les militaires qui intègrent les composants électroniques dans le guidage des avions. Le circuit imprimé va résoudre ce problème puis le circuit intégré. Dans les premiers ordinateurs, les éléments des circuits, tels les lampes, étaient reliés entre eux par des réseaux extrêmement complexes de fils. Le premier développement important avec l'apparition du transistor a été de monter les transistors sur une même plaque de circuits et de graver les fils qui les reliaient dans la plaque; on a appelé cela des circuits imprimés.

Par la suite, grâce d'abord aux travaux d'un ingénieur de la Texas Instruments, Jack Kilby, et ensuite, d'un ingénieur de Fairchild Semiconductors, Robert Noyce, en 1959, on a été capable de relier entre eux tous les éléments du circuit, transistors, diodes, condensateurs, fils, etc. dans des circuits dits complètement intégrés, manufacturés en une seule opération. Si, à ce momentlà, cette percée technologique n'a pas connu tout le succès qu'elle méritait, on prétendait que le coût en était trop élevé, le lancement du Spoutnik par les Russes, l'inquiétude qui s'en est suivie chez les Américains d'être surpassés et le désir du président d'alors, John Kennedy, d'envoyer un homme sur la lune, acheva de convaincre, car il fallait pour réaliser cette conquête de la lune, des ordinateurs assez petits pour tenir dans une capsule spatiale. Par exemple, l'ordinateur de la cabine Apollo ne pesait que 54 livres, soit le centième du poids de certains gros ordinateurs de l'époque. (Palfreman et Swade, 1991).

évidemment, ces efforts de la NASA et de l'armée américaine également, qui avait besoin de petits ordinateurs pour guider ses fusées Minuteman, ont permis de financer la recherche et développement sur les circuits intégrés et d'en réduire sensiblement les coûts de fabrication, si bien que les circuits imprimés se sont retrouvés non seulement dans les nouveaux ordinateurs mais également dans toutes sortes d'appareils électroniques y compris les appareils domestiques, radio, télévision, horloges, etc. C'est donc grâce aux circuits intégrés qu'on a pu construire une nouvelle génération d'ordinateurs appelée mini-ordinateurs dont les coûts étaient très inférieurs aux gros mais encore trop chers pour le commun des mortels.

Les circuits intégrés sont miniaturisés et déposés sur une pastille de quelques millimètres carrés, appelée chip ou puce, faite de matériau semi-conducteur, généralement du silicium. Il y a des degrés quant au nombre de composants électroniques intégrés qu'on peut imprimer sur une puce : on parle de SSI (Small Scale Integration) pour quelques dizaines de composants, de MSI (Medium Scale Integration) jusqu'à 500 composants, de LSI (Large Scale Integration) de 500 à quelques milliers, de VLSI (VeryLarge Scale Integration) jusqu'à 10 000 composants, soit un microprocesseur complet capable de remplir toutes les fonctions de l'unité centrale de traitement d'un ordinateur, et de SLSI (Super Large Scale Integration) où sont regroupés sur une même puce, des microprocesseurs interconnectés représentant plus de 50 000 transistors. En 1983, on était capable de mettre un million de composants sur un microprocesseur. Cette technologie fera également diminuer sensiblement les coûts d'un ordinateur car le coût de revient d'un circuit imprimé en 1970, représente le 1/250e de celui de 1962 si bien que dans les années 1980, cela représentait le 1/260 000e du coût.

Les années 1964 à 1975 ont vu l'arrivée de plusieurs mini-ordinateurs de différentes grandes compagnies. Ces ordinateurs dits de troisième génération sont caractérisés par l'utilisation massive de circuits intégrés. Parmi les ordinateurs de début de troisième génération, mentionnons le Burroughs B5500, les CDC 6000 et 3300, les IBM 360, l'UNIVAC 1108, les Honeywell 200, le RCA Spectra 70, le NCR Century et les G.E. 400 et 600. De 1969 à 1975, les ordinateurs IBM de la série 370, le CDC Cyber 70, les Honeywell 2000 et 6000, les UNIVAC 1110 et 9400, les Burroughs B6700 et 1700 ont vu le jour. Plusieurs d'entre eux sont encore en opération.

C'est aussi l'époque des mini-ordinateurs dont l'unité centrale de traitement manipule des mots de 16 ou de 32 bits. Mentionnons les plus connus : IBM 360, DEC VAX 11/780, Data General MV/8000, la série des PDP de DEC, etc.

1964 : Jusqu'alors, les ordinateurs de marques différentes et même d'une même marque n'étaient pas compatibles. On devait réécrire les programmes quand on changeait de type de machine. De plus, la plupart des périphériques étaient propres à une machine en particulier. IBM décida donc de mettre au point une famille d'ordinateurs à travers laquelle les programmes et les périphériques seraient interchangeables, la famille des IBM/360. Le plus petit des ordinateurs de la famille pouvait faire 33 000 additions à la seconde et le plus gros, 2 500 000. La famille offrait en tout 19 combinaisons de vitesse de calcul et de capacité de mémoire sans compter les quelques 44 périphériques compatibles. Les IBM/360 étaient à base de circuits intégrés produits grâce à une méthode mise au point par IBM qui leur permettait de placer de minuscules circuits sur des modules de céramique de 1 cm.

1.5. La 4ème génération d'ordinateurs : l'ère des microprocesseurs

«Il n'y a aucune raison que des gens veuillent un ordinateur à la maison.»

Ken OLSON, PDG et fondateur de DEC, 1977

La mise au point des microprocesseurs va entraîner la miniaturisation des composants d'ordinateurs et partant, l'apparition de deux nouveaux types d'ordinateurs : le super ordinateur et le micro-ordinateur ou ordinateur personnel. La miniaturisation va aussi permettre l'invention des calculatrices de poches, des montres à affichage numérique, des contrôles numériques d'appareils domestiques comme le four à micro-ondes, la machine à laver, etc...

1968 : Robert Noyce et Gordon Moore quittent la Fairchild pour fonder Intel avec chacun un investissement de 250 000\$. Intel Corp. va devenir le plus gros fabricant de puces et de microprocesseurs au monde.

1969 : La société AMD est créée par W.J. Sanders III de Fairchild et sept autres personnes pour compétitionner Intel sur le marché des microprocesseurs. Cette compagnie aura des haut et des bas.

1970 : C'est en 1970, que Bob Abbott, sous la direction de Les Vadasz, met au point chez Intel, la première puce à mémoire dynamique, la DRAM (Dynamic Read Access Memory), INTEL 1103, qui deviendra à compter de 1972, la plus vendue à travers le monde. Elle n'avait qu'un seul kilo-octet. Sa vitesse était de 300 nanosecondes. Il faudrait aujourd'hui 65 000 de ces puces 1103 pour faire 8 Mo de mémoire. Les puces ne sont pas fabriquées à l'unité. Elles sont à peine grandes comme des timbres-poste. Après l'étape du design, on fabrique des masques par processus photographique répétés de telle sorte à obtenir jusqu'à 200 fois l'image réduite du circuit de la puce sur une même gaufre. Par conséquent, plus la puce est réduite, plus on peut en mettre sur une gaufre et plus le coût de revient est bas.

1971 : Intel crée la première puce à mémoire morte qu'on peut programmer électroniquement, la EPROM (Electronic Programmed Read Only Memory) Intel 1702.

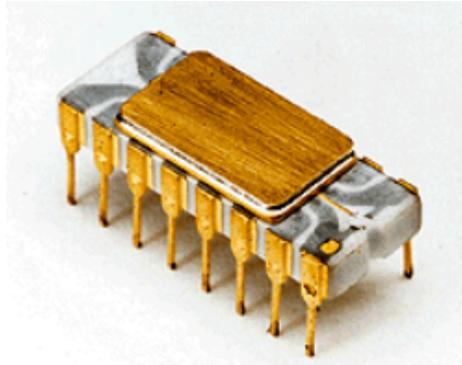


Fig. 1.12: Microprocesseur Intel 4004

1971 : La compagnie Intel, fondée en 1968 et elle progresse très rapidement, en particulier grâce au travail de Ted Hoff, jeune ingénieur à Intel et de son équipe formée de Stan Mazor, Federico Faggin et Masatoshi Shima. Cette équipe est la première à concevoir le design d'un ordinateur sur une puce, autrement dit un microprocesseur. Le premier microprocesseur a été fabriqué pour répondre à une commande spéciale d'une compagnie japonaise, la Busicom, qui voulait intégrer une série de 6 puces très complexes à son nouveau modèle de calculatrice. Hoff leur fait savoir que leur design est beaucoup trop complexe à réaliser; il se met alors au travail devant son miniordinateur PDP-8 et il se demande pourquoi une simple calculatrice devait être si complexe. Il fait alors le design du microprocesseur qui n'a pas réussi à impressionner les Japonais. Il en rachètera les droits à Busicom au coût de 60 000\$ US. Mais Bob Noyce, le fondateur d'Intel avec Gordon Moore, continue à supporter le projet de Hoff. Un microprocesseur à 4 bits voit le jour, le 4004 : 2 300 transistors de 10 microns sur une puce de 12 mm carré. Le 15 novembre 1971, le 4004 est lancé sous le slogan A micro-programmable computer on a chip! dans une publicité parue dans le Electronic News de New York. Le Intel 4004 tournait à la vitesse de 108 KHz, soit 0,06 MIPS (million d'instructions par seconde). Son bus de mémoire était de 4 Ko. sa mémoire adressable était de 640 octets. Il fallait 16 griffes pour le fixer.

1973 : A. Truong, fondateur de R2E, une petite société Française, crée le premier microordinateur du monde : le Micral-N.

1973 : Le prototype Alto, de XEROX, pensé pour devenir le bureau du futur, est un condensé des idées proposées par les chercheurs réunis par XEROX au Palo-Alto Research Center (PARC). Il est le premier à introduire l'idée de fenêtres et d'icônes que l'on peut gérer grâce à une souris. Principalement, en raison de son coût, cet ordinateur ne connaîtra qu'un succès d'estime. 1975 : Ed Roberts fonde son entreprise, la MITS, qui met en marché l'Altair 8800, premier micro-ordinateur en kit.



Fig. 1.13: Altair 8800

1977 : Apple, fondée par Steve Jobs et Steve Wozniak avec l'aide d'un ex-ingénieur millionnaire de Intel, Mike Markula, qui aide les jeunes capitalistes à se partir en affaires, lance en grande pompe en avril 1977 (les livraisons commenceront en juin de la même année), le Apple II, un ordinateur à 4 Ko de mémoire vive (RAM) extensible à 16 Ko, 16 Ko de mémoire morte (ROM) et à processeur MOS 6502 de 1 MHz, le même que le Pet de Commodore et un clavier intégré de 52 touches.

Au départ, il comportait un lecteur-enregistreur de cassettes puis à compter de mars 1978 d'un lecteur de disquettes 5,25 po. On le branchait à un téléviseur noir et blanc ou couleur ou à un écran monochrome séparé qui affiche en 40 colonnes et 24 lignes, bien que le Apple II soit déjà capable d'afficher des graphiques haute résolution en couleur, ce qui constitue une innovation. Il coûtait 1 295\$ US initialement mais deux ans plus tard, son prix baissera à 970\$ pour un modèle à 2 Ko de mémoire alors qu'un modèle à 48 Ko coûtait 1 795\$ US.

A la fin de 1979, on en avait vendu 35 000. C'était un favori des bricoleurs car il possédait de nombreuses fentes d'expansion, il était facile à programmer grâce à son langage Basic intégré en mémoire morte et il suffisait de lever le couvercle pour y avoir accès. Il deviendra rapidement l'ordinateur favori du monde scolaire car c'est un ordinateur ouvert avec des fentes d'expansion et il comprend le langage Basic en mémoire morte. Plusieurs petites compagnies se mettent à fabriquer des périphériques de toutes sortes et des logiciels très diversifiés pour le Apple II. En septembre 1979, les ventes du Apple II étaient de plus de 35 000 unités par année. Avec les années, il connaîtra plusieurs générations successives : le Apple II Plus, le Apple IIe, le Apple IIc et le Apple II GS.

1978 : C'est le 8 juin 1978, qu'Intel lance le premier microprocesseur à 16 bits, le 8086 à 4,77 MHz (0,33 MIPS) , et 8 (0,66 MIPS) ou 10 MHz (0,75 MIPS), qui deviendra, à l'époque, un standard de l'industrie. Il comprenait 29 000 transistors de 3 microns soit six fois plus que le 8080. Il permettait la division et la multiplication accélérant ainsi les calculs de précision. La performance du 8086 est dix fois celle de son ancêtre à 8 bits, le 8080. Son bus est de 16 bits et sa mémoire adressable est de 1 Mo.

1980 : Motorola introduit un microprocesseur 32 bits à bus interne de 16 bits et bus de mémoire à 24 bits qui équipera les premiers LISA de Apple et les premiers Macintosh mis en marché en 1984. Il équipera également les Amiga de Commodore et les ST de Atari et quelques stations de travail UNIX.



Fig. 1.14: Apple II

1981 : Basé sur le microprocesseur Z80A, le ZX81 est le premier micro-ordinateur à moins de 1000F. Il utilise l'écran du poste de télévision.



Fig. 1.15: IBM PC

1981 : IBM, sentant Apple devant elle dans un nouveau marché, annonce en août 1981, qu'il se lance sur le marché du micro-ordinateur avec le concept d'ordinateur personnel qui prend la forme du IBM PC construit autour du microprocesseur Intel 8088 à 8-16 bits et 4,77 MHz. Il avait 16 Ko de mémoire vive, extensible à 64 Ko , un seul lecteur de disquettes simple face de 5,25 po d'une capacité de stockage de 160 Ko. Il était doté d'un écran monochrome vert de 12 po. Il était offert à 2 880\$ US, une aubaine. Un compétiteur géant pour Apple venait de voir le jour. Le PC a été créé par une équipe de jeunes informaticiens dirigée par Philip D. Estridge; cett équipe comprenait Bill Sydnes, ingénieur chef, Dan Wilkie, responsable de la fabrication et H.L. Sparks, responsable des ventes. Il fonctionne sous un système d'exploitation mis au point par Microsoft sous contrat avec IBM, le MS DOS et le PC DOS. Ce système est un héritage de cerveaux d'informaticiens qui ne se soucient pas de convivialité. Le DOS se présente comme un langage d'ordinateur à la syntaxe capricieuse et rebutante. C'est une machine qui nécessite la formation de l'usager. Malgré cette absence de convivialité, le PC prend une bonne part du marché; on en vend 35 000 en 1981 seulement soit cinq fois plus que ce qui avait été prévu au départ, aidé en cela par une campagne de publicité bien orchestré autour du personnage de Charlie Chaplin des Temps modernes, le sérieux de la grosse machine bleue et par le fait que plusieurs compagnies ont pu commercialiser des compatibles (ou des clones) sous licence.

1982 : Compaq lance le premier clone de l'IBM PC.

1984 : Apple lance le Macintosh. Basé sur le projet LISA , c'est l'ordinateur convivial par excellence : Son utilisation est très simple grâce à la souris et à la qualité de ses graphismes. Il devient au fil des années et des version, l'autre grand standard (avec le PC d'IBM) du monde de la micro-informatique.

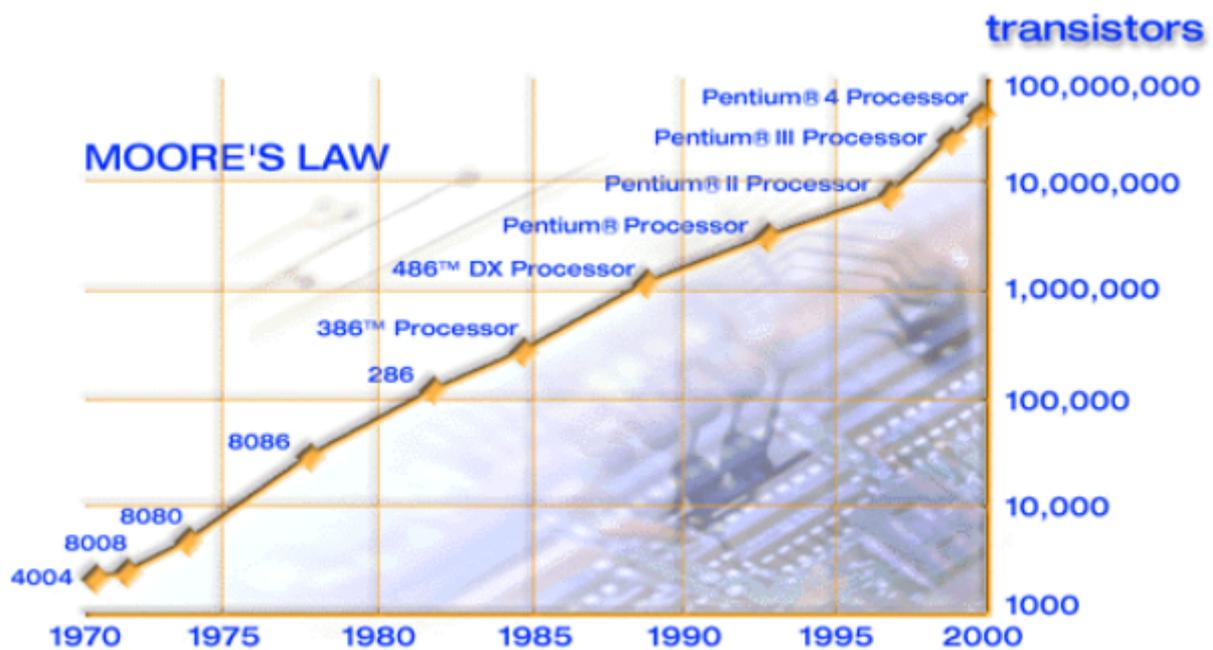


Fig. 1.16: Evolution du nombre de transistors dans les microprocesseurs Intel. Jusqu'ici, il a doublé tous les 18 mois (loi de Moore)

1984+ : Les techniques évoluent et les batailles commerciales font rage, mais exploitent toujours le même principe de base, comme les premiers ordinateurs : l'architecture Von Neumann. Dans le chapitre suivant, nous détaillons les différents composants d'un micro-ordinateur moderne.

2. Les micro-ordinateurs modernes

2.1. Introduction à la notion d'ordinateur

La compréhension du vocabulaire informatique représente généralement la principale difficulté à laquelle se heurtent les utilisateurs débutants d'ordinateurs personnels. Ce chapitre n'a pas pour but de donner un sens à toutes les abréviations informatiques (dans la mesure où de nombreux constructeurs ont leurs propres technologies) mais il a vocation à donner de la visibilité sur les principaux composants d'un ordinateur, d'en expliquer le fonctionnement et d'en donner les principales caractéristiques à connaître.

2.1.1. Présentation de l'ordinateur

Un ordinateur est un ensemble de circuits électronique permettant de manipuler des données sous forme binaire, c'est-à-dire sous forme de bits. Le mot « ordinateur » provient de la firme IBM. Celle-ci demanda en 1954 à un professeur de lettres à Paris de trouver un mot pour désigner ce que l'on appelait vulgairement un « calculateur » (traduction littérale de computer en anglais).

2.1.2. Types d'ordinateurs

Toute machine capable de manipuler des informations binaires peut être qualifiée d'ordinateur.

Toutefois, la plupart des personnes pensent à un ordinateur personnel (PC, abréviation de personal computer), le type d'ordinateur le plus présent sur le marché, toutefois il existe beaucoup d'autres types d'ordinateurs (la liste suivante est non exhaustive) :

- Amiga;
- Atari;
- Apple Macintosh;
- stations Alpha;
- stations SUN;
- stations Silicon Graphics.

Nous nous intéresserons dans la suite qu'aux ordinateurs de type PC, appelés aussi ordinateurs compatible IBM, car IBM est la firme qui a créé les premiers ordinateurs de ce type et a longtemps (jusqu'en 1987) été le leader dans ce domaine, à un tel point qu'elle contrôlait les standards, copiée par les autres fabricants.

2.1.3. Constitution de l'ordinateur

Un ordinateur est un ensemble de composants électroniques modulaires, c'est-à-dire des composants pouvant être remplacés par d'autres composants ayant éventuellement des caractéristiques différentes. Ces composants sont architecturés autour d'une carte principale comportant de nombreux circuits intégrés (soudés sur la carte) et un grand nombre de connecteurs; cette carte est appelée carte mère.

La carte mère est logée dans un boîtier, comportant des emplacements pour les périphériques de stockage sur la face avant, ainsi que des boutons permettant de contrôler la mise sous tension de l'ordinateur et un certain nombre de voyants permettant de vérifier l'état de marche de l'appareil et l'activité des disques durs. Sur la face arrière le boîtier propose des ouvertures en vis-à-vis des cartes d'extension et des interfaces d'entrée-sortie connectées sur la carte mère. Enfin le boîtier héberge une alimentation, chargée de fournir un courant électrique stable à l'ensemble des éléments constitutifs de l'ordinateur.

On appelle unité centrale l'ensemble composé du boîtier et de l'ensemble des éléments qu'il embarque. L'unité centrale doit être connectée à un ensemble de périphériques externes. Un ordinateur est généralement composé au minimum d'une unité centrale, un écran (moniteur), d'un clavier et d'une souris, mais il est possible de connecter une grande diversité de périphériques externes sur les interfaces d'entrée-sortie (ports séries, port parallèle, port USB, port firewire, ...) :

- imprimante;
- scanner;
- périphérique de stockage externe;
- appareil photo ou caméra numérique;
- assistant personnel (PDA);
- ...

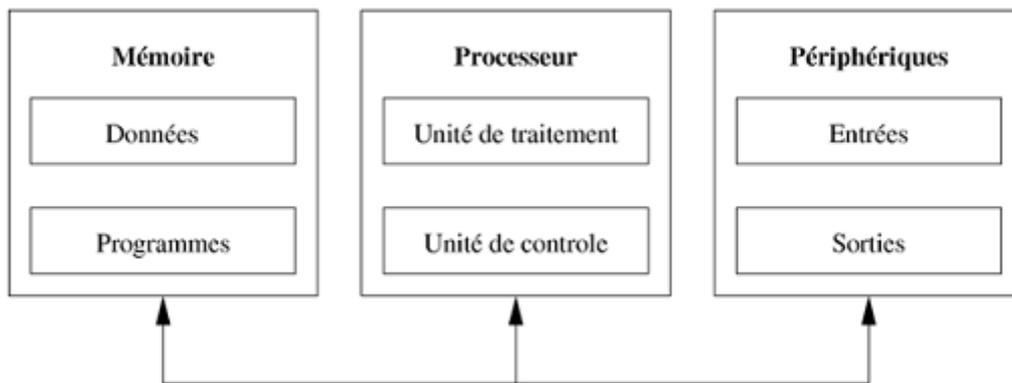


Fig. 2.1: L'architecture de Von Neumann

Depuis les années 70, les micro-ordinateurs sont devenus beaucoup plus puissants, mais leur principe de base est resté le même. Un micro-ordinateur utilise une architecture Von Neumann.

Ils utilisent tous :

- la mémoire qui stocke les données et les programmes;
- l'unité logique qui traite et modifie les données;
- l'unité de contrôle qui organise le fonctionnement interne de la machine;
- les différents organes d'entrée et de sortie (roues dentées, cartes perforées, claviers, écrans, imprimantes etc).

Un bus de données permet l'échange d'informations entre tous ces composants.

2.2. La carte mère

2.2.1. Présentation de la carte-mère

L'élément constitutif principal de l'ordinateur est la carte-mère, c'est sur cette carte que sont connectés ou soudés l'ensemble des éléments essentiels de l'ordinateur.

La carte-mère contient des éléments embarqués (intégrés à la carte) :

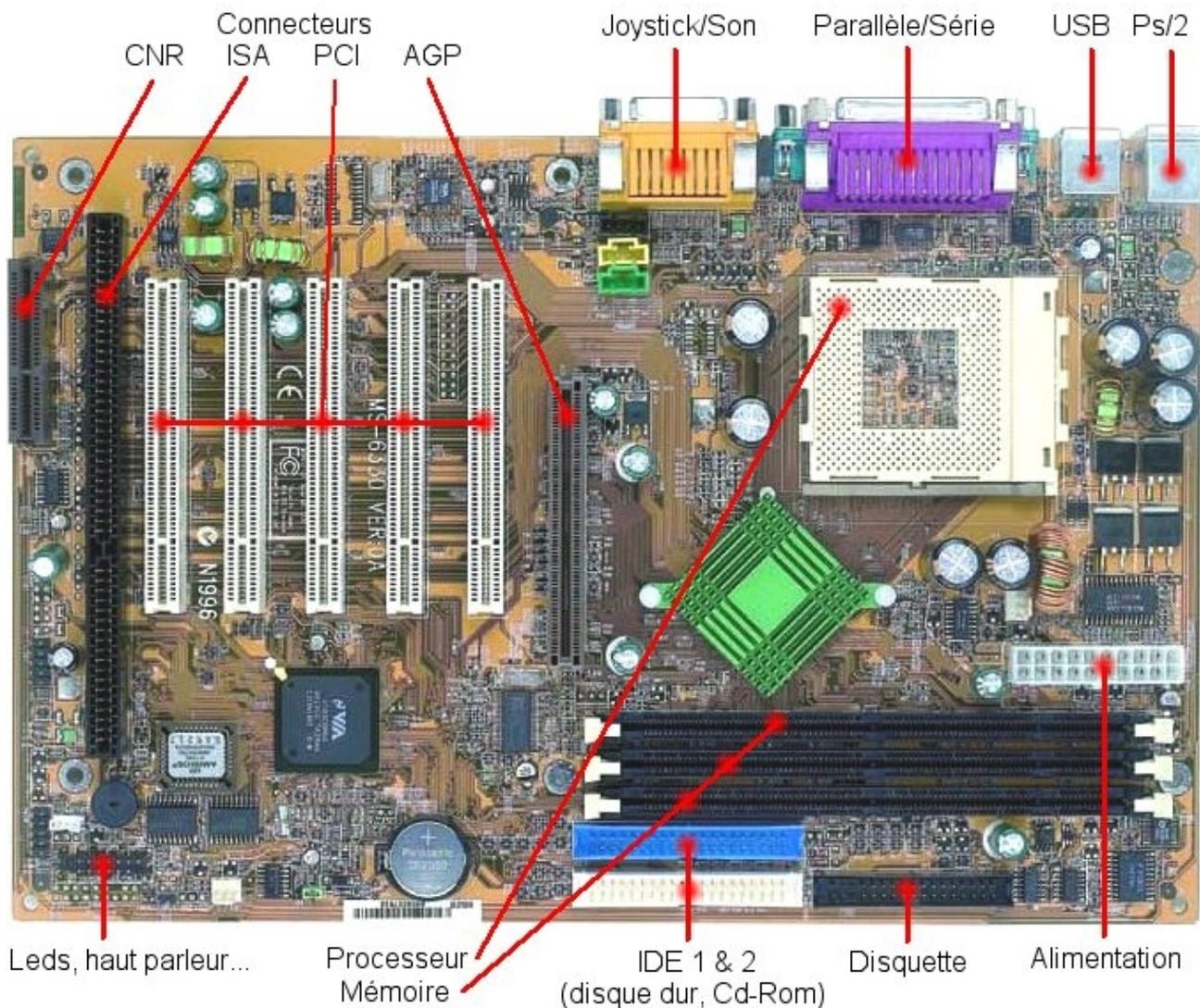


Fig. 2.2: Carte mère

- le chipset, circuit qui contrôle la majorité des ressources (interface de bus du processeur, mémoire cache et mémoire vive, slots d'extension, ...);
- l'horloge et la pile du CMOS;
- le BIOS.

Il existe plusieurs façons de caractériser une carte-mère :

- son facteur d'encombrement;
- son chipset;
- son type de support de processeur.

2.2.2. Facteur d'encombrement d'une carte-mère

On entend généralement par facteur d'encombrement, la géométrie et les dimensions de la carte-mère. Afin de fournir des cartes-mères pouvant s'adapter dans différents boîtiers de marques différentes, des standards ont été mis au point :

- AT baby;
- AT full format;
- ATX;
- LPX;
- NLX.

2.2.3. Le chipset

Le chipset (traduisez jeu de composants) est un circuit électronique chargé de coordonner les échanges de données entre les divers composants de l'ordinateur (processeur, mémoire, ...). Dans la mesure où le chipset est intégré à la carte-mère, il est important de choisir une carte-mère embarquant un chipset récent afin de garantir à votre PC un maximum de chance de pouvoir évoluer.

Certains chipsets intègrent parfois une puce graphique ou une puce audio (généralement sur les PC bas de gamme), ce qui signifie qu'il n'est pas nécessaire d'installer une carte graphique ou une carte son. Toutefois, étant donné la piètre qualité de ces composants intégrés, il est généralement conseillé de les désactiver (lorsque cela est possible) dans le setup du BIOS et d'installer des cartes d'extension dans les emplacements prévus à cet effet!

2.2.4. L'horloge et la pile du CMOS

L'horloge temps réel (parfois notée RTC, ou real time clock) est un circuit chargé de la synchronisation des signaux du système. Elle est constituée d'un cristal qui, en vibrant, donne des impulsions (appelés tops d'horloge) afin de cadencer le système. On appelle fréquence de l'horloge (exprimée en MHz) le nombre de vibrations du cristal par seconde, c'est-à-dire le nombre de tops d'horloge émis par seconde. Plus la fréquence est élevée, plus il y a de tops d'horloge et donc plus le système pourra traiter d'informations.

Lorsque vous mettez votre ordinateur hors tension, l'alimentation cesse de fournir du courant à la carte-mère. Or, lorsque vous le rebranchez, votre système d'exploitation est toujours à l'heure bien que l'unité centrale n'était plus alimentée pendant un certain temps. En réalité même lorsque votre PC est débranché ou qu'une panne d'électricité intervient, un circuit électronique appelé CMOS (Complementary Metal-Oxyde Semiconductor, parfois appelé BIOS CMOS) conserve certaines informations sur le système, y compris l'heure et la date système. Le CMOS est continuellement alimenté par une pile (au format pile bouton) située également sur la carte-mère.

Ainsi, les informations sur le matériel installé dans l'ordinateur (comme par exemple le nombre de pistes, de secteurs de chaque disque dur) sont conservées dans le CMOS. Dans la mesure où le CMOS est une mémoire lente, certains systèmes recopient parfois le contenu du CMOS dans la RAM (mémoire rapide), le terme de memory shadow est utilisé pour décrire ce processus de copie en mémoire vive.

Le «complementary metal-oxyde semiconductor» est une technologie de fabrication de transistors, précédée de bien d'autres, comme la TTL (Transistor-transistor-logique), ou la TTLS (TTL Schottky) (plus rapide),... Avant la CMOS, il y avait également la NMOS (canal négatif) et la PMOS (canal positif). La CMOS, qui a permis de mettre des canaux complémentaires sur une même puce, a ainsi été une grande avancée. Par rapport à la TTL ou TTLS, la CMOS est beaucoup moins rapide, mais a le grand avantage de consommer infiniment moins d'énergie, d'où son emploi dans les horloges d'ordinateurs, qui sont alimentées par des piles. Ainsi le terme de CMOS est

parfois utilisé abusivement pour désigner l'horloge des ordinateurs.

Ainsi, si vous constatez que votre PC a tendance à oublier l'heure, où que l'horloge prend du retard, pensez à en changer la pile!

2.2.5. Le BIOS

Le BIOS (Basic Input/Output System) est le programme basique servant d'interface entre le système d'exploitation et la carte-mère. Le BIOS est stocké dans une ROM (mémoire morte, c'est-à-dire une mémoire en lecture seule), ainsi il utilise les données contenues dans le CMOS pour connaître la configuration matérielle du système.

Il est possible de « configurer » le BIOS grâce à une interface (nommée BIOS setup, traduisez configuration du BIOS) accessible au démarrage de l'ordinateur par simple pression d'une touche (généralement la touche Suppr. En réalité le setup du BIOS sert uniquement d'interface pour la configuration et les données sont stockées dans le CMOS. Pour plus d'informations n'hésitez pas à vous reporter au manuel de votre carte-mère).

2.2.6. Le processeur

Le processeur (aussi appelé microprocesseur) est le cœur de l'ordinateur, car il exécute les instructions des programmes grâce à un jeu d'instructions. Le processeur est caractérisé par sa fréquence, c'est-à-dire la cadence à laquelle il exécute les instructions. Ainsi, de manière grossière, un processeur cadencé à 600 MHz effectuera 600 millions d'opérations par seconde.

La carte-mère possède un emplacement (parfois plusieurs dans le cas de cartes-mères multiprocesseurs) pour accueillir le processeur. On distingue deux catégories de supports :

- slot : il s'agit d'un connecteur rectangulaire dans lequel on enfiche le processeur verticalement;
- socket : il s'agit d'un connecteur carré possédant un grand nombre de petits connecteurs sur lequel le processeur vient directement s'enficher.

Dans la mesure où le processeur rayonne thermiquement, il est nécessaire d'en dissiper la chaleur pour éviter que ses circuits ne fondent. C'est la raison pour laquelle il est généralement surmonté d'un dissipateur thermique, un matériau ayant une bonne conduction thermique, chargé d'augmenter la surface d'échange thermique du microprocesseur. Le dissipateur thermique comporte une base en contact avec le processeur et des ailettes afin d'augmenter la surface d'échange thermique. Un ventilateur accompagne généralement le dissipateur pour améliorer la circulation de l'air autour du dissipateur et améliorer l'échange de chaleur. C'est le ventilateur du boîtier qui est chargé d'extraire l'air chaud du boîtier et permettre à l'air frais provenant de l'extérieur d'y entrer.

2.2.7. La mémoire-cache

La mémoire-cache permet au processeur de se « rappeler » les opérations déjà effectuées auparavant. En effet, elle stocke les opérations effectuées par le processeur, pour qu'il ne perde pas de temps à recalculer des choses qu'il a déjà faites précédemment. La taille de la mémoire-cache est généralement de l'ordre de 512 Ko. Sur les ordinateurs récents ce type de mémoire est directement intégré dans le processeur.

2.2.8. La mémoire vive

La mémoire vive (RAM pour Random Access Memory) permet de stocker des informations pendant tout le temps de fonctionnement de l'ordinateur, son contenu est par contre détruit dès lors que l'ordinateur est éteint ou redémarré, contrairement à une mémoire de masse comme le disque-dur qui garde les informations même lorsqu'il est hors-tension.

Pourquoi alors se servir de mémoire alors que les disques durs sont moins chers ? Car elle est extrêmement rapide comparé aux périphériques de stockage de type disque dur (de l'ordre de quelques dizaines de nanosecondes: environ 70 pour la DRAM, 60 pour la RAM EDO, et 10 pour la SDRAM).

2.2.9. Les connecteurs d'extension

Les connecteurs d'extension (en anglais slots) sont des réceptacles dans lesquels il est possible d'enficher des cartes d'extension, c'est-à-dire des cartes offrant de nouvelles fonctionnalités ou de meilleures performances à l'ordinateur. Il existe plusieurs sortes de connecteurs :

- connecteur ISA (Industry Standard Architecture) : permettant de connecter des cartes ISA, les plus lentes fonctionnant en 16-bit;
- connecteur VLB (Vesa Local Bus): Bus servant autrefois à connecter des cartes graphiques;
- connecteur PCI (Peripheral Component InterConnect) : permettant de connecter des cartes PCI, beaucoup plus rapides que les cartes ISA et fonctionnant en 32-bit;
- connecteur AGP (Accelerated Graphic Port): un connecteur rapide pour carte graphique;
- connecteur AMR (Audio Modem Riser): ce type de connecteur permet de brancher des mini-cartes sur les PC en étant équipés.

2.2.10. Le bus système

On appelle bus, le canal permettant de transférer des données entre deux éléments. Le bus système est le canal (pistes de la carte-mère) reliant le microprocesseur à la mémoire vive du système. Un bus est caractérisé par sa largeur, c'est-à-dire le nombre de bits pouvant être simultanément transmis, et par sa fréquence, c'est-à-dire la cadence à laquelle les paquets de bits peuvent être transmis. Des caractéristiques du bus système dépendent donc les caractéristiques générales du système. La fréquence du microprocesseur est égale à la fréquence du bus système multiplié par un facteur. Ainsi un PC tournant à 400 MHz sera plus rapide s'il est basé sur un bus système cadencé à 133 MHz (3 x 133 MHz) que si la carte-mère a un bus dont la fréquence est 100 MHz (la fréquence du processeur étant alors égale à 4 x 100 MHz).

2.3. Le processeur

2.3.1. Qu'est-ce qu'un processeur?

Le processeur (CPU: Central Processing Unit) est un circuit électronique cadencée au rythme d'une horloge interne, c'est-à-dire un élément qui envoie des impulsions (que l'on appelle top). A chaque top d'horloge les éléments de l'ordinateur accomplissent une action. La vitesse de cette horloge (le nombre de battements par secondes) s'exprime en Mégahertz, ainsi un ordinateur à 200 MHz a donc une horloge envoyant 200.000.000 de battements par seconde (un cristal de quartz soumis à un courant électrique permet d'envoyer des impulsions à une fréquence précise). A chaque top

d'horloge (pour les instructions simples) le processeur :

- lit l'instruction à exécuter en mémoire
- effectue l'instruction
- passe à l'instruction suivante

Le processeur est en fait constitué :

- d'une unité de commande qui lit les instructions et les décode
- d'une unité de traitement (UAL unité arithmétique et logique) qui exécute les instructions

Lorsque tous les éléments d'un processeur sont regroupés sur une même puce, on parle alors de microprocesseur.

2.3.2. A quoi ressemble une instruction?

Les instructions (opération que le processeur doit accomplir) sont stockées dans la mémoire principale. Une instruction est composée de deux champs :

- le code opération: c'est l'action que le processeur doit accomplir
- le code opérande: c'est les paramètres de l'action. Le code opérande dépend de l'opération, cela peut être une donnée ou bien une adresse d'un emplacement mémoire.

Une instruction peut être codée sur un nombre d'octets variant de 1 à 4 suivant le type de données.

2.3.3. Les registres

Lorsque le processeur traite des données (lorsqu'il exécute des instructions) le processeur stocke temporairement les données dans de petites mémoires de 8, 16 ou 32Ko (qui ont l'avantage d'être très rapides) que l'on appelle registres. Suivant le type de processeur le nombre de registres peut varier entre une dizaine et plusieurs centaines.

Les registres les plus importants sont :

- le registre accumulateur : il permet de stocker les résultats des opérations arithmétiques et logiques
- le registre tampon : il permet de stocker temporairement une des opérandes
- le registre d'état : il permet de stocker les indicateurs
- le registre instruction : il contient l'instruction en cours de traitement
- le compteur ordinal : il contient l'adresse de la prochaine instruction à traiter
- le registre tampon : il permet de stocker temporairement une donnée provenant de la mémoire

2.3.4. Les signaux de commande

Les signaux de commande sont des signaux électriques qui permettent au processeur de communiquer avec le reste du système (le signal Read/Write lecture/écriture permet notamment de signaler à la mémoire qu'il désire lire ou écrire une information.

2.3.5. Qu'est-ce qu'un microprocesseur ?

Le premier microprocesseur (Intel 4004) a été inventé en 1971. Depuis, la puissance des microprocesseurs augmente exponentiellement. Le processeur (CPU) est le cerveau de l'ordinateur, c'est lui qui coordonne le reste des éléments, il se charge des calculs, bref il exécute les instructions qui ont été programmées. Toutes ces opérations sont des informations numériques. Les microprocesseurs utilisent des petits transistors (équivalents transistors) pour faire des opérations de base; il y en a plusieurs millions sur un seul processeur.

Les éléments principaux d'un microprocesseur sont :

- une horloge qui rythme le processeur. A chaque TOP d'horloge le processeur effectue une instruction, ainsi plus l'horloge a une fréquence élevée, plus le processeur effectue d'instructions par seconde (MIPS : Millions d'instruction par seconde). Par exemple un ordinateur ayant une fréquence de 100 MHz effectue 100 000 000 d'instructions par seconde
- une unité de gestion des bus qui gère les flux d'informations entrant et sortant
- une unité d'instruction qui lit les données arrivant, les décode puis les envoie à l'unité d'exécution
- une unité d'exécution qui accomplit les tâches que lui a donné l'unité d'instruction

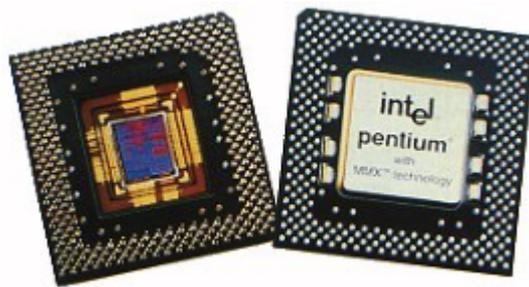


Fig. 2.3: Microprocesseur Intel Pentium

Le processeur travaille en fait grâce à un nombre très limité de fonctions (ET logique, Ou logique, addition ...), celles-ci sont directement câblées sur les circuits électroniques. Il est impossible de mettre toutes les instructions sur un processeur car celui-ci est limité par la taille de la gravure, ainsi pour mettre plus d'instructions il faudrait un processeur ayant une très grande surface, or le processeur est constitué de silicium et le silicium coûte cher, d'autre part il chauffe beaucoup. Le processeur traite donc les informations compliquées à l'aide d'instructions simples.

2.3.6. Le parallélisme

Le parallélisme consiste à exécuter simultanément sur des processeurs différents des instructions relatives à un même programme. Cela se traduit par le découpage d'un programme en plusieurs processus qui seront traités par des processeurs différents dans le but de gagner en temps d'exécution. Cela nécessite toutefois une communication entre les différents processus. C'est le même principe de fonctionnement que dans une entreprise : le travail est divisé en petits processus traités par des services différents et qui ne servent à rien si la communication entre les services ne fonctionne pas (ce qui est généralement le cas dans les entreprises...).

2.3.7. Le pipelining

Le pipelining est un principe simple à comprendre. Un programme comporte généralement des

portions de code (plus ou moins grandes) qui sont traitées de nombreuses fois par le processeur. Le pipelining consiste donc à éviter d'avoir à réitérer de nombreuses fois des instructions que l'on a déjà traitées en fournissant directement le résultat !

2.3.8. L'architecture CISC

L'architecture CISC (Complex Instruction Set Computer, ce qui signifie « ordinateur avec jeu d'instructions complexes ») est utilisée par tous les processeurs de type x86, c'est-à-dire les processeurs fabriqués par Intel, AMD, Cyrix,... Les processeurs basés sur l'architecture CISC peuvent traiter des instructions complexes, qui sont directement câblées sur leurs circuits électroniques, c'est-à-dire que certaines instructions difficiles à créer à partir des instructions de base sont directement imprimées sur le silicium de la puce afin de gagner en rapidité d'exécution sur ces commandes. L'inconvénient de ce type d'architecture provient justement du fait que des fonctions supplémentaires sont imprimées sur le silicium, d'où un coût élevé.

D'autre part, les instructions sont de longueurs variables et peuvent parfois prendre plus d'un cycle d'horloge ce qui les rend lentes à l'exécution étant donné qu'un processeur basé sur l'architecture CISC ne peut traiter qu'une instruction à la fois !

2.3.9. L'architecture RISC

Contrairement à l'architecture CISC, un processeur utilisant la technologie RISC (Reduced Instruction Set Computer, dont la traduction est « ordinateur à jeu d'instructions réduit ») n'a pas de fonctions supplémentaires câblées. Cela impose donc des programmes ayant des instructions simples interprétables par le processeur. Cela se traduit par une programmation plus difficile et un compilateur plus puissant. Cependant vous vous dites qu'il peut exister des instructions qui ne peuvent pas être décrites à partir des instructions simples... En fait ces instructions sont tellement peu nombreuses qu'il est possible de les câbler directement sur le circuit imprimé sans alourdir de manière dramatique leur fabrication.

L'avantage d'une telle architecture est bien évidemment le coût réduit au niveau de la fabrication des processeurs l'utilisant. De plus, les instructions, étant simples, sont exécutées en un cycle d'horloge, ce qui rend l'exécution des programmes plus rapides qu'avec des processeurs basés sur une architecture CISC.

De plus, de tels processeurs sont capables de traiter plusieurs instructions simultanément en les traitant en parallèle.

2.3.10. CISC ou RISC?

A comparer les spécificités des deux types d'architecture on pourrait conclure que les processeurs basés sur une architecture de type RISC sont les plus utilisés... Cela n'est malheureusement pas le cas... En effet les ordinateurs construits autour d'une architecture RISC nécessitent une quantité de mémoire plus importante que les ordinateurs de type CISC.

2.4. La mémoire

2.4.1. Rôle de la mémoire vive (RAM)

La mémoire vive, généralement appelée RAM (Random Access Memory, traduisez mémoire à accès aléatoire), est la mémoire principale du système, c'est-à-dire qu'il s'agit d'un espace permettant de stocker de manière temporaire des données lors de l'exécution d'un programme.

En effet le stockage de données dans la mémoire vive est temporaire, contrairement au stockage de données sur une mémoire de masse telle que le disque dur (mémoire avec laquelle les novices la confondent généralement), car elle permet uniquement de stocker des données tant qu'elle est alimentée électriquement. Ainsi, à chaque fois que l'ordinateur est éteint, toutes les données présentes en mémoire sont irrémédiablement effacées.



Fig. 2.4: Barette de mémoire

La mémoire morte, appelée ROM pour Read Only Memory (traduisez mémoire en lecture seule) est un type de mémoire permettant de conserver les informations qui y sont contenues même lorsque la mémoire n'est plus alimentée électriquement. A la base ce type de mémoire ne peut être accédée qu'en lecture. Toutefois il est désormais possible d'enregistrer des informations dans certaines mémoires de type ROM.

2.4.2. Fonctionnement de la mémoire vive

La mémoire vive est constituée de centaines de milliers de petits condensateurs emmagasinant des charges. Lorsqu'il est chargé, l'état logique du condensateur est égal à 1, dans le cas contraire il est à 0, ce qui signifie que chaque condensateur représente un bit de la mémoire.

Étant donné que les condensateurs se déchargent, il faut constamment les recharger (le terme exact est rafraîchir) à un intervalle de temps régulier appelé cycle de rafraîchissement (d'une durée d'environ 15 nanosecondes (ns) pour une mémoire DRAM).

Chaque condensateur est couplé à un transistor (de type MOS) permettant de « récupérer » ou de modifier l'état du condensateur. Ces transistors sont rangés sous forme de tableau (matrice), c'est-à-dire que l'on accède à une « case mémoire » (aussi appelée point mémoire) par une ligne et une colonne.

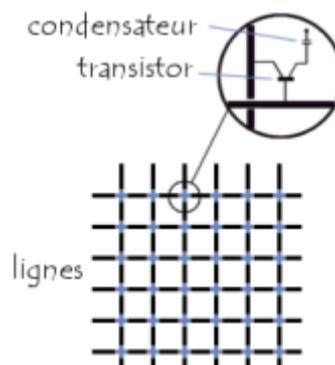


Fig. 2.5: Table de mémoire

Chaque point mémoire est donc caractérisé par une adresse, correspondant à un numéro de ligne et

un numéro de colonne. Or cet accès n'est pas instantané et s'effectue pendant un délai appelé temps de latence. Par conséquent l'accès à une donnée en mémoire dure un temps égal au temps de cycle auquel il faut ajouter le temps de latence.

Ainsi, pour une mémoire de type DRAM, le temps d'accès est de 60 nanosecondes (35ns de délai de cycle et 25ns de temps de latence). Sur un ordinateur, le temps de cycle correspond à l'inverse de la fréquence de l'horloge, par exemple pour un ordinateur cadencé à 200 MHz, le temps de cycle est de 5ns ($1/(200.10^6)$).

Par conséquent un ordinateur ayant une fréquence élevée et utilisant des mémoires dont le temps d'accès est beaucoup plus long que le temps de cycle du processeur doit effectuer des cycles d'attente (en anglais wait state) pour accéder à la mémoire. Dans le cas d'un ordinateur cadencé à 200 MHz utilisant des mémoires de types DRAM (dont le temps d'accès est de 60ns), il y a 11 cycles d'attente pour un cycle de transfert. Les performances de l'ordinateur sont d'autant diminuées qu'il y a de cycles d'attentes, il est donc conseillé d'utiliser des mémoires plus rapides.

2.4.3. La correction d'erreurs

Certaines mémoires possèdent des mécanismes permettant de pallier les erreurs afin de garantir l'intégrité des données qu'elles contiennent. Ce type de mémoire est généralement utilisée sur des systèmes travaillant sur des données critiques, c'est la raison pour laquelle on trouve ce type de mémoire dans les serveurs.

Bit de parité

Les barrettes avec bit de parité permettent de s'assurer que les données contenues dans la mémoire sont bien celles que l'on désire. Pour ce faire, un des bits de chaque octet stocké en mémoire sert à conserver la somme des bits de données. Le bit de parité vaut 0 lorsque la somme des bits de données est impaire et 1 dans le cas contraire.

De cette façon les barrettes avec bit de parité permettent de vérifier l'intégrité des données mais ne permettent pas de corriger les erreurs. De plus pour 8 Mo de mémoire, seulement 7 serviront à stocker des données, dans la mesure où le dernier méga-octet conservera les bits de parité.

Barrettes ECC

Les barrettes de mémoire ECC (Error Correction Coding) sont des mémoires possédant plusieurs bits dédiés à la correction d'erreur (on les appelle ainsi bits de contrôle). Ces barrettes, utilisées principalement dans les serveurs, permettent de détecter les erreurs et de les corriger.

2.4.4. La mémoire morte (ROM)

Il existe un type de mémoire permettant de stocker des données en l'absence de courant électrique, il s'agit de la ROM (Read Only Memory, dont la traduction est mémoire en lecture seule) appelée parfois mémoire non volatile, car elle ne s'efface pas lors de la mise hors tension du système.

Ce type de mémoire permet notamment de conserver les données nécessaires au démarrage de l'ordinateur. En effet, ces informations ne peuvent être stockées sur le disque dur étant donné que les paramètres du disque (essentiels à son initialisation) font partie de ces données vitales à l'amorçage. Différentes mémoires de type ROM contiennent des données essentielles au démarrage, c'est-à-dire :

- le BIOS est un programme permettant de piloter les interfaces d'entrée-sortie principales du système, d'où le nom de BIOS ROM donné parfois à la puce de mémoire morte de la carte-

mère qui l'héberge;

- le chargeur d'amorce : un programme permettant de charger le système d'exploitation en mémoire (vive) et de le lancer. Celui-ci cherche généralement le système d'exploitation sur le lecteur de disquette, puis sur le disque dur, ce qui permet de pouvoir lancer le système d'exploitation à partir d'une disquette système en cas de dysfonctionnement du système installé sur le disque dur
- le Setup CMOS, c'est l'écran disponible à l'allumage de l'ordinateur permettant de modifier les paramètres du système (souvent appelé BIOS à tort...)
- le Power-On Self Test (POST), programme exécuté automatiquement à l'amorçage du système permettant de faire un test du système (c'est pour cela par exemple que vous voyez le système « compter » la RAM au démarrage)

Étant donné que les ROM sont beaucoup plus lentes que les mémoires de types RAM (une ROM a un temps d'accès de l'ordre de 150 ns tandis qu'une mémoire de type SDRAM a un temps d'accès d'environ 10 ns), les instructions contenues dans la ROM sont parfois copiées en RAM au démarrage, on parle alors de shadowing (en français cela pourrait se traduire par ombrage, mais on parle généralement de mémoire fantôme).

2.4.5. Les types de ROM

Les ROM ont petit à petit évoluées de mémoires mortes figées à des mémoires programmables, puis reprogrammables.

ROM

Les premières ROM étaient fabriquées à l'aide d'un procédé inscrivant directement les données binaires dans une plaque de silicium grâce à un masque. Ce procédé est maintenant obsolète.

PROM

Les PROM (Programmable Read Only Memory) ont été mises au point à la fin des années 70 par la firme Texas Instruments. Ces mémoires sont des puces constituées de milliers de fusibles pouvant être « grillés » grâce à un appareil appelé programmeur de ROM, envoyant un fort courant (12V) dans certains fusibles. Ainsi, les fusibles grillés correspondent à des 0, les autres à des 1.

EPROM

Les EPROM (Erasable Programmable Read Only Memory) sont des PROM pouvant être effacées. Ces puces possèdent une vitre permettant de laisser passer des rayons ultra-violet. Lorsque la puce est en présence de rayons ultra-violet d'une certaine longueur d'onde, les fusibles sont reconstitués, c'est-à-dire que tous les bits de la mémoire sont à nouveau à 1. C'est pour cette raison que l'on qualifie ce type de PROM d'effaçable.

EEPROM

Les EEPROM (Electrically Erasable read Only Memory) sont aussi des PROM effaçables, mais contrairement aux EPROM, celles-ci peuvent être effacées par un simple courant électrique, c'est-à-dire qu'elles peuvent être effacées même lorsqu'elles sont en position dans l'ordinateur. Ces mémoires sont aussi appelées mémoires flash (ou ROM flash), et l'on qualifie de flashage l'action consistant à reprogrammer une EEPROM.

2.5. Les périphériques externes

Les ports d'entrée-sortie sont des éléments matériels de l'ordinateur, permettant au système de communiquer avec des éléments extérieurs, c'est-à-dire d'échanger des données, d'où l'appellation d'interface d'entrée-sortie (notée parfois interface d'E/S). Les périphériques externes sont comme leur nom l'indique connectés à l'extérieur du PC, c'est-à-dire sur les ports de communication (COM1, COM2, COM3 ..) ou le(s) port(s) imprimante (LPT1, LPT2 ...). Il s'agit principalement:

- de la souris
- du clavier
- des scanner
- des imprimantes
- des modems externes

2.5.1. Les ports série

Les ports série représentent les premières interfaces ayant permis aux ordinateurs d'échanger des informations avec le « monde extérieur ». Le terme série désigne un envoi de données via un fil unique: les bits sont envoyés les uns à la suite des autres (reportez-vous à la section transmission de données pour un cours théorique sur les modes de transmission).

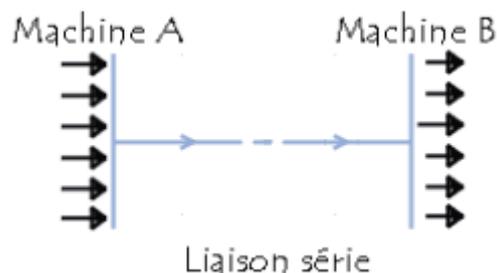


Fig. 2.6: Transmission sur un port série

A l'origine les ports série permettait uniquement d'envoyer des données, mais pas d'en recevoir, c'est pourquoi des ports bidirectionnels ont été mis au point (ceux qui équipent les ordinateurs actuels le sont); les ports séries bidirectionnels ont donc besoin de deux fils pour effectuer la communication.

La communication série se fait de façon asynchrone, cela signifie qu'aucun signal de synchronisation (appelé horloge) n'est nécessaire: les données peuvent être envoyées à intervalle de temps arbitraire. En contrepartie, le périphérique doit être capable de distinguer les caractères (un caractère a une longueur de 8 bits) parmi la suite de bits qui lui est envoyée... C'est la raison pour laquelle dans ce type de transmission, chaque caractère est précédé d'un bit de début (appelé bit START) et d'un bit de fin (bit STOP). Ces bits de contrôle, nécessaires pour une transmission série, gaspillent 20% de la bande passante (pour 8 bits envoyés, 2 servent à assurer la réception).



Fig. 2.7: Connecteur DB9

Les ports série sont généralement intégrés à la carte-mère, c'est pourquoi des connecteurs présents à l'arrière du boîtier, et reliés à la carte-mère par un nappe de fils, permettent de connecter un élément extérieur. Les connecteurs séries possèdent généralement 9 ou 25 broches et se présentent sous la forme suivante (respectivement connecteurs DB9 et DB25).

Un ordinateur personnel possède généralement entre deux et quatre ports séries (certains de ces ports possèdent des connecteurs DB9, d'autres des connecteurs DB25).

2.5.2. Les ports parallèle

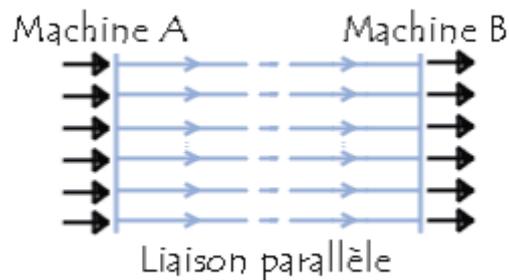


Fig. 2.8: Transmission sur un port parallèle

La transmission de données en parallèle consiste à envoyer des données simultanément sur plusieurs canaux (fils). Les ports parallèle présents sur les ordinateurs personnels permettent d'envoyer simultanément 8 bits (un octet) par l'intermédiaire de 8 fils.

Les premiers ports parallèles bidirectionnels permettaient d'atteindre des débits de l'ordre de 2.4Mb/s. Toutefois des ports parallèles améliorés ont été mis au point afin d'obtenir des débits plus élevés :

- le port EPP (Enhanced Parallel Port, port parallèle amélioré) a permis d'atteindre des débits de l'ordre de 8 à 16 Mbps;
- le port ECP (Enhanced Capabilities Port, port à capacités améliorées), mis au point par Hewlett Packard et Microsoft. Il reprend les caractéristiques du port EPP en lui ajoutant un support Plug and Play, c'est-à-dire la possibilité pour l'ordinateur de reconnaître les périphériques branchés.



Fig. 2.9: Connecteur DB25

Les ports parallèle sont, comme les ports série, intégrés à la carte-mère. Les connecteurs DB25 permettent de connecter un élément extérieur (une imprimante par exemple).

2.5.3. Les ports USB

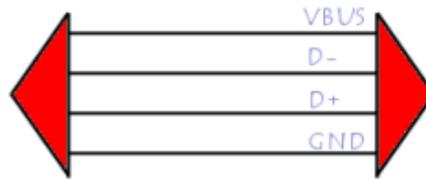


Fig. 2.10: Le câble USB

Les ports USB (Universal Serial Bus, ports séries universels) sont, comme leur nom l'indique, basés sur une architecture de type série. Il s'agit toutefois d'une interface entrée-sortie beaucoup plus rapide que les ports série standards. L'architecture qui a été retenue pour ce type de port est en série pour deux raisons principales :

- l'architecture série permet d'utiliser une cadence d'horloge beaucoup plus élevée qu'une interface parallèle, car celle-ci ne supporte pas des fréquences trop élevées (dans une architecture à haut débit, les bits circulant sur chaque fil arrivent avec des décalages, provoquant des erreurs);
- les câbles séries coûtent beaucoup moins chers que des câbles parallèles.

Ainsi, dès 1995, le standard USB a été élaboré. Il propose deux modes de communication (12 Mbps en mode haute vitesse et 1.5 Mbps à basse vitesse) pour la connexion d'une grande variété de périphériques. L'architecture USB a pour caractéristique de fournir l'alimentation électrique aux périphériques qu'elle relie. Elle utilise pour cela un câble composé de quatre fils (la masse GND, l'alimentation VBUS et deux fils de données appelés D- et D+).

La norme USB permet le chaînage des périphériques, en utilisant une topologie en bus ou en étoile. Les périphériques peuvent alors être soit connectés les uns à la suite des autres, soit ramifiés. La ramification se fait à l'aide de boîtiers appelés hubs (ou concentrateurs), comportant une seule entrée et plusieurs sorties. Certains sont actifs (fournissant de l'énergie électrique), d'autres passifs.

La communication entre l'hôte (l'ordinateur) et les périphériques se fait selon un protocole (langage de communication) basé sur le principe de l'anneau à jeton (token ring). Cela signifie que la bande passante est partagée temporellement entre tous les périphériques connectés. L'hôte émet un signal de début de séquence chaque milliseconde (ms), intervalle de temps pendant lequel il va donner simultanément la « parole » à chacun d'entre-eux. Lorsque l'hôte désire communiquer avec un périphérique, il émet un jeton (un paquet de données, contenant l'adresse du périphérique, codé sur 7 bits) désignant un périphérique. Si ce dernier reconnaît son adresse dans le jeton, il envoie un paquet de données en réponse. Sinon, il fait suivre le paquet aux autres périphériques connectés à lui. Puisque l'adresse est codée sur 7 bits, 128 périphériques (2^7) peuvent être connectés simultanément à un port de ce type. Il convient en réalité de ramener ce chiffre à 127 car l'adresse 0 est une adresse réservée (cf plus loin). A raison de 5m de câble maximum entre deux périphériques, il est possible de créer une chaîne longue de 636m !

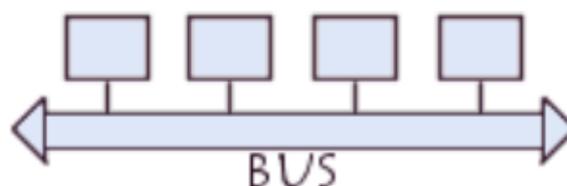


Fig. 2.11: Topologie en bus des ports USB

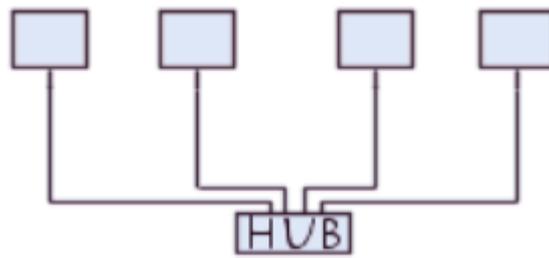


Fig. 2.12: Topologie en étoile des ports USB

Les ports USB supportent le Hot plug and play. Ainsi, il est possible de brancher les périphériques sans éteindre l'ordinateur (branchement à chaud). Lors de la connexion du périphérique à l'hôte, ce dernier détecte l'ajout du nouvel élément grâce au changement de la tension entre les fils D+ et D-. A ce moment, l'ordinateur envoie un signal d'initialisation au périphérique pendant 10 ms, puis lui fournit du courant grâce aux fils GND et VBUS (jusqu'à 100mA). Le périphérique est alors alimenté en courant électrique et récupère temporairement l'adresse par défaut (l'adresse 0). L'étape suivante consiste à lui fournir son adresse définitive (c'est la procédure d'énumération). Pour cela, l'ordinateur interroge les périphériques déjà branchés pour connaître la leur et en attribue une au nouveau, qui en retour s'identifie. L'hôte, disposant de toutes les caractéristiques nécessaires est alors en mesure de charger le pilote approprié...

2.5.4. Le port FireWire

Afin de fournir un système d'interconnexion permettant de faire circuler des données à haute vitesse en temps réel, le bus FireWire (appelé IEEE 1394, nom de la norme à laquelle il fait référence) a été mis au point à la fin de l'année 1995.

Le bus IEEE 1394 suit à peu près la même structure que le bus USB. si ce n'est qu'il utilise un câble composé de six fils (deux paires pour les données et pour l'horloge, et deux fils pour l'alimentation électrique) lui permettant d'obtenir un débit de 400 Mbps (il devrait atteindre prochainement 1Gbps). Ainsi, les deux fils dédiés à une horloge montrent la différence majeure qui existe entre le bus USB et le bus IEEE 1394 : ce dernier peut fonctionner selon deux modes de transfert :

- le mode de transfert asynchrone
- le mode isochrone

Le mode de transfert asynchrone est basé sur une transmission de paquets à intervalles de temps variables. Cela signifie que l'hôte envoie un paquet de données et attend de recevoir un accusé de réception du périphérique. Si l'hôte reçoit un accusé de réception, il envoie le paquet de données suivant, sinon le paquet est à nouveau réexpédié au bout d'un temps d'attente. Le mode de transfert isochrone permet l'envoi de paquets de données de taille fixe à intervalle de temps régulier (cadencé grâce aux deux fils d'horloge). De cette façon aucun accusé de réception n'est nécessaire, on a donc un débit fixe et donc une bande passante garantie. De plus, étant donné qu'aucun accusé n'est nécessaire, l'adressage des périphériques est simplifié et la bande passante économisée permet de gagner en vitesse de transfert.

Autre innovation du standard IEEE 1394 : la possibilité d'utiliser des ponts, systèmes permettant de relier plusieurs bus entre-eux. En effet, l'adressage des périphériques se fait grâce à un identificateur de $n/2$ ud (c'est-à-dire de périphérique) codé sur 16 bits. Cet identificateur est scindé en deux champs : un champ de 10 bits permettant de désigner le pont et un champ de 6 bits

spécifiant le n 1/2 ud. Il est donc possible de relier 1023 ponts, sur chacun desquels il peut y avoir 63 n 1/2 uds, il est ainsi possible d'adresser 65535 périphériques! Le standard IEEE 1394 permet aussi le Hot plug'n play, mais alors que les ports USB sont réservés à l'utilisation de périphériques peu gourmands en ressources (souris ou clavier par exemple), la bande passante de l'IEEE 1394 la destine à des utilisations multimédias sans précédents (acquisition vidéo, ...).

2.6. Le système d'affichage

Le système d'affichage d'un PC est son interface de sortie principale. Il est composé d'un moniteur et d'une carte graphique.

2.6.1. Le moniteur

Le moniteur à tube cathodique

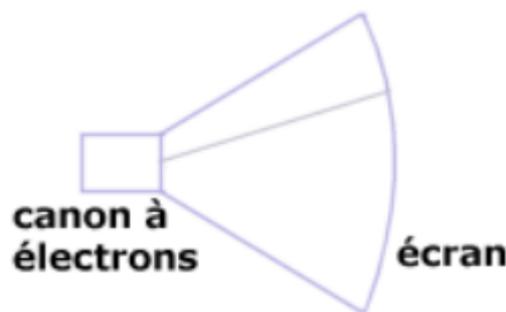


Fig. 2.13: Tube et canon à électrons

Les moniteurs (écrans d'ordinateur) sont la plupart du temps des tubes cathodiques, c'est à dire un tube en verre dans lequel un canon à électrons émet des électrons dirigés par un champ électrique vers un écran sur lequel de petits éléments phosphorescents (luminophores) constituent des points (pixels). Ces luminophores émettent de la lumière lorsque les électrons viennent les heurter.

Le champ magnétique dévie les électrons de gauche à droite afin de créer un balayage, puis vers le bas une fois arrivé en bout de ligne.

Ce balayage n'est pas perçu par l'œil humain grâce à la persistance rétinienne, essayez par exemple d'agiter votre main devant votre écran pour visualiser ce phénomène : vous voyez votre main en plusieurs exemplaires ...

Le moniteur couleur

Un moniteur noir et blanc permet d'afficher des dégradés de couleur (niveaux de gris) en variant l'intensité du rayon.

Pour les moniteurs couleur, trois faisceaux d'électrons sont utilisés simultanément en visant chacun un point d'une couleur spécifique : un rouge, un vert et un bleu (RGB: Red/Green/Blue ou en français RVB: Rouge/vert/bleu). Cependant ces luminophores sont situés de façon tellement proche que l'œil n'a pas un pouvoir séparateur assez fort: il voit une couleur composée de ces trois couleurs. Essayez de mettre une minuscule goutte d'eau sur le verre de votre moniteur: celle-ci faisant un effet de loupe va vous faire apparaître les luminophores.

Il existe deux grandes catégories de tubes :

- les tubes FST-Invar et Cromaclear dont les luminophores sont ronds (grâce à une grille

appelée masque);

- les tubes Trinitron dont le le masque est constitué de fentes verticales, laissant passer plus de lumière.

Les moniteurs à cristaux liquides

Cette technologie est basée sur un écran composé de deux plaques transparentes entre lesquelles il y a une fine couche de liquide dans laquelle il y a des molécules (cristaux) qui ont la propriété de s'orienter lorsqu'elles sont soumises à du courant électrique. L'avantage majeur de ce type d'écran est son encombrement réduit, d'où son utilisation sur les ordinateurs portables.

Les caractéristiques

Les moniteurs sont souvent caractérisés par les données suivantes :

- La définition: c'est le nombre de points qu'il peut afficher, ce nombre de points est actuellement compris entre 640x480 (640 points en longueur, 480 points en largeur) et 1600x1200
- la taille: Il ne faut pas confondre la définition de l'écran et la taille de l'écran. En effet un écran d'une taille donnée peut afficher différentes définitions, cependant, généralement, les écrans de grande taille (celle-ci se calcule en mesurant la diagonale de l'écran et est exprimée en pouces, c'est-à-dire 2.54 cm) possèdent une meilleure définition que les écrans de petite taille
- la résolution: Elle détermine le nombre de pixels par unité de surface (pixels par pouce linéaire (en anglais DPI: Dots Per Inch, traduisez points par pouce). Une résolution de 300 dpi signifie 300 colonnes et 300 rangées de pixels sur un pouce carré ce qui donnerait donc 90000 pixels sur un pouce carré. La résolution de référence de 72 dpi nous donne un pixel de $1''/72$ (un pouce divisé par 72) soit 0,353mm, correspondant à un point pica (unité typographique anglo saxonne)
- le pas de masque: C'est la distance qui sépare deux points, plus celle-ci est petite plus l'image est précise
- la fréquence de balayage: C'est le nombre d'images qui sont affichées par seconde, on l'appelle aussi rafraîchissement, elle est exprimée en Hertz. Plus cette valeur est élevée meilleur est le confort visuel (on ne voit pas l'image scintiller), il faut donc qu'elle soit supérieure à 67 Hz (limite inférieure à partir de laquelle l'œil remarque véritablement l'image « clignoter »).

2.6.2. La carte graphique

Les cartes accélératrices 2D

Les cartes 2D n'ont pas changé de principe depuis leur création. Chaque puce possède de nombreux circuits qui permettent d'exécuter de nombreuses fonctions :

- déplacement des blocs (curseur de la souris par exemple)
- tracé de lignes
- tracé de polygones



Fig. 2.15: Carte graphique

Ainsi, les performances des cartes 2D n'évoluent plus depuis quelques temps.

Leurs performances sont tributaires du type de mémoire utilisée sur la carte (les mémoires SGRAM ou WRAM, mémoires vidéo spécifiques à 10 ns, donnent des résultats bien meilleurs que la mémoire EDO (60 ns))

La fréquence du RAM-DAC (RAM Digital Analogic Converter), ainsi que la quantité de mémoire vidéo ne permettent en rien d'avoir de meilleures performances, elles permettent juste d'avoir un meilleur taux de rafraîchissement (nombre d'images par seconde) et de pouvoir accéder à des résolutions plus grandes.

Les cartes accélératrices 3D

Le domaine de la 3D est beaucoup plus récent, donc plus porteur. On arrive à des puissances de calculs sur PC supérieures à celles de certaines stations de travail.

Le calcul d'une scène 3D est un processus qui se décompose grossièrement en quatre étapes :

- le script : mise en place des éléments
- la géométrie : création d'objets simples
- le setup : découpage en triangles 2D
- le rendering : C'est le rendu, c'est-à-dire le plaquage des textures.

Ainsi, plus la carte accélératrice 3D calcule elle-même ces étapes, plus l'affichage est rapide. Les premières puces n'effectuaient que le rendering, laissant le processeur s'occuper du reste. Depuis, les cartes possèdent un « setup engine » qui prend en charge les deux dernières étapes. A titre d'exemple, un Pentium II à 266 MHz qui calcule les trois premières étapes peut calculer 350 000 polygones par secondes, lorsqu'il n'en calcule que deux, il atteint 750 000 polygones par seconde.

Cela montre à quel point ces cartes déchargent le processeur.

Le type de bus est lui aussi déterminant. Alors que le bus AGP n'apporte aucune amélioration dans le domaine de la 2D, les cartes utilisant ce bus plutôt que le bus PCI sont beaucoup plus performantes. Cela s'explique par le fait que le bus AGP est directement relié à la mémoire vive, ce qui lui offre une bande passante beaucoup plus grande que le bus PCI.

Ces produits de haute technologie ont maintenant besoin de la même qualité de fabrication que les

processeurs, ainsi que des gravures allant de 0.35 à 0.25 microns.

2.7. Les disques durs

Le disque dur est l'organe du PC servant à conserver les données de manière permanente, contrairement à la RAM, qui s'efface à chaque redémarrage de l'ordinateur. Il a été inventé au début des années 50 par IBM.

2.7.1. Le fonctionnement interne



Fig. 2.16: Disque dur

Un disque dur est constitué non pas d'un seul disque, mais de plusieurs disques rigides (en anglais hard disk signifie disque dur) en métal, en verre ou en céramiques empilés les uns après les autres à une très faible distance les uns des autres. Ils tournent très rapidement autour d'un axe (à plusieurs milliers de tours par minute actuellement) dans le sens inverse des aiguilles d'une montre.

Un ordinateur fonctionne de manière binaire, il faut donc stocker les données sous forme de 0 et de 1, c'est pourquoi les disques sont recouverts d'une très fine couche magnétique de quelques microns d'épaisseur, elle-même recouverte d'un film protecteur.

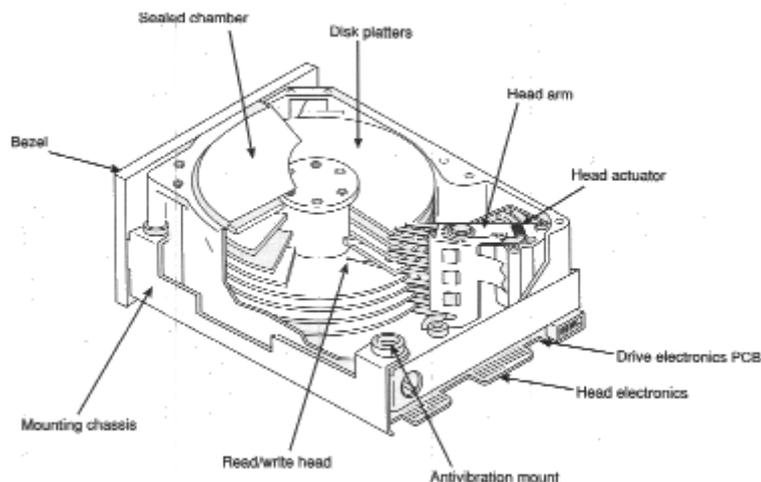


Fig. 2.17: Structure d'un disque dur (têtes de lecture)

La lecture et l'écriture se font grâce à des têtes (head) situées de part et d'autre de chacun des plateaux (un des disques composant le disque dur). Ces têtes sont des électroaimants qui se baissent et se soulèvent (elles ne sont qu'à quelques microns de la surface, séparées par une couche d'air provoquée par la rotation des disques qui crée un vent d'environ 250km/h) pour pouvoir lire

l'information ou l'écrire. De plus ces têtes peuvent balayer latéralement la surface du disque pour pouvoir accéder à toute la surface...

Cependant, les têtes sont liées entre-elles et seulement une seule tête peut lire ou écrire à un moment donné. On parle donc de cylindre pour désigner l'ensemble des données stockées verticalement sur la totalité des disques.

L'ensemble de cette mécanique de précision est contenue dans un boîtier totalement hermétique, car la moindre particule peut détériorer l'état de surface du disque. Vous pouvez donc voir sur un disque des opercules permettant l'étanchéité, et la mention « Warranty void if removed » qui signifie littéralement « la garantie expire si retiré » car seul les constructeurs de disques durs peuvent les ouvrir (dans des salles blanches : exemptes de particules).

2.7.2. La lecture et l'écriture

Les têtes de lecture/écriture sont dites « inductives », c'est-à-dire qu'elles sont capables de générer un champ magnétique. C'est notamment le cas lors de l'écriture, les têtes en créant des champs positifs ou négatifs viennent polariser la surface du disque en une très petite zone, ce qui se traduira lors du passage en lecture par des changements de polarité induisant un courant dans la tête qui sera ensuite transformé par un convertisseur analogique numérique (CAN) en 0 et en 1 compréhensibles par l'ordinateur.

Les têtes commencent à inscrire des données à la périphérie du disque (piste 0), puis avancent vers le centre. Les données sont organisées en cercles concentriques appelés « pistes », créées par le formatage de bas niveau.

Les pistes sont séparées en quartiers (entre deux rayons) que l'on appelle secteurs, c'est la zone dans laquelle on peut stocker les données (512 octets en général).

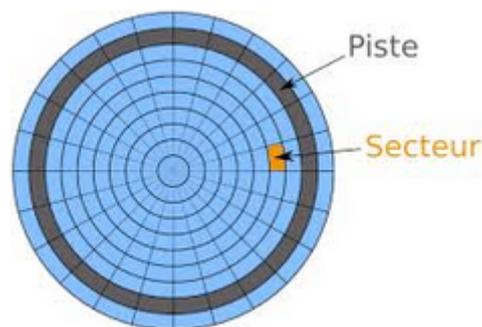


Fig. 2.18: Pistes et secteur d'un disque dur

On appelle cylindre l'ensemble des données situées sur une même piste de plateaux différents (c'est-à-dire à la verticale les unes des autres) car cela forme dans l'espace un « cylindre » de données.

On appelle cluster la zone minimale que peut occuper un fichier sur le disque. En effet le système d'exploitation exploite des blocs qui sont en fait plusieurs secteurs (entre 1 et 16 secteurs). Un fichier minuscule devra donc occuper plusieurs secteurs (un cluster).

2.7.3. Le mode bloc des disques durs

Le mode bloc et le transfert 32 bits permettent d'exploiter pleinement les performances de votre disque dur. Le mode bloc consiste à effectuer des transferts de données par bloc, c'est-à-dire par

paquets de 512 octets généralement, ce qui évite au processeur d'avoir à traiter une multitude de minuscules paquets d'un bit. Le processeur a alors du « temps » pour effectuer d'autres opérations. Ce mode de transfert des données n'a malheureusement une véritable utilité que sous DOS car Windows 95 et Windows NT utilisent leur propres gestionnaires de disque dur, ce qui rend ce gestionnaire obsolète.

Une option du BIOS (IDE HDD block mode ou Multi Sector Transfer, ...) permet parfois de déterminer le nombre de blocs pouvant être gérés simultanément. Ce nombre se situe entre 2 et 32. Si vous ne le connaissez pas, plusieurs solutions s'offrent à vous :

- consulter la documentation de votre disque dur
- rechercher les caractéristiques de votre disque sur Internet
- le déterminer expérimentalement en effectuant des tests :
 - exécuter scandisk sur votre ordinateur pour éliminer les erreurs
 - augmenter progressivement le nombre de blocs puis faire une copie et lancer scandisk
 - si des erreurs apparaissent remettre la valeur précédente...sinon continuer.

Le mode bloc peut toutefois générer des erreurs sous Windows 3.1 (à cause d'une redondance de gestionnaire de disque dur) ou bien lors d'une gravure de CD (le tampon se vide). La solution consiste alors à désactiver l'un des deux gestionnaires :

- la gestion logicielle du mode 32-bit sous Windows
- le mode bloc dans le BIOS

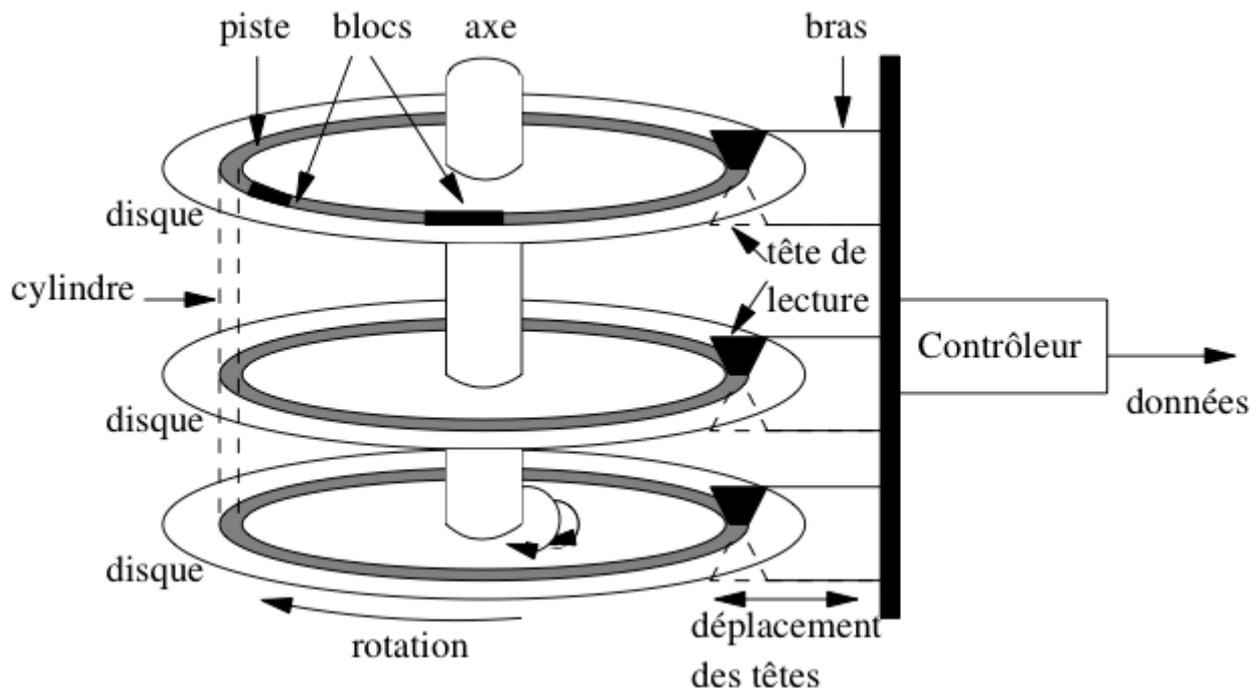


Fig. 2.20: Cylindres d'un disque dur

2.7.4. Le mode 32 bits des disques durs

Le mode 32 bits (par opposition au mode 16 bits) est caractérisé par un transfert des données sur 32

bits (Rappel : un ordinateur fonctionne avec des données binaires, c'est-à-dire avec des zéros ou des 1, schématiquement une porte qui s'ouvre ou bien qui se ferme. Le transfert sur 32 bits correspond à 32 portes qui s'ouvrent et se ferment simultanément. En mode 16 bits on a deux mots (ensemble de bits) de 16 bits qui sont transmis successivement, puis assemblés).

Le gain de performance relatif au passage du mode 16 bits au mode 32 bits (pour les disques durs) est généralement insignifiant. Quoi qu'il en soit il n'est la plupart du temps plus possible de choisir le mode, car la carte-mère détermine seule le type de mode à adopter en fonction du type de disque dur branché sur l'interface E-IDE.

La détermination automatique du mode 32 bits peut toutefois ralentir les lecteurs de CD-ROM IDE dont la vitesse est supérieure à 24x lorsqu'ils sont seuls sur une nappe IDE. En effet, dans le cas où le lecteur de CD-ROM est seul sur le port, le BIOS peut ne pas détecter sa compatibilité avec le mode 32 bits (puisqu'il cherche un disque dur) auquel cas il passe en mode 16 bits. Le taux de transfert est alors en dessous du taux de transfert annoncé par le constructeur d'où une grande déception de son possesseur... Heureusement, il existe une solution : brancher sur la même nappe que le lecteur de CD-ROM un disque dur supportant le mode 32 bits, ce qui aura pour effet d'activer le mode .

2.7.5. L'interface SCSI

L'interface SCSI est une interface qui permet la prise en charge d'un nombre important d'unités (disques durs, CD-ROM, Graveur, scanner, ...), c'est-à-dire plus d'une dizaine simultanément. Elle est beaucoup utilisée pour sa stabilité notamment au niveau du taux de transfert. En effet, c'est un adaptateur SCSI (carte adaptatrice sur un emplacement PCI ou ISA ou bien directement intégré sur la carte-mère pour les configurations haut de gamme) qui se charge de la gestion et du transfert des données avec un microprocesseur dédié. Le microprocesseur central est alors relégué de ses activités concernant le flux de données, il ne communique qu'avec la carte SCSI. Ainsi chaque contrôleur SCSI a ses propres caractéristiques (fréquence, ...), le BIOS n'a donc aucune influence sur les performances de l'interface SCSI étant donné qu'elle possède elle-même son propre BIOS. Il est toutefois possible d'optimiser cette interface en faisant évoluer le BIOS de la carte SCSI.

2.7.6. Les caractéristiques du disque

Le taux de transfert est la quantité de données qui peuvent être lues ou écrites sur le disque en un temps donné. Il s'exprime aujourd'hui en Méga-Octets par seconde.

Le temps de latence (aussi appelé délai rotationnel) représente le temps entre lequel le disque a trouvé la piste et où il trouve les données.

Le temps d'accès est le temps que met la tête pour aller d'une piste à la piste suivante (elle doit être la plus petite possible).

Le temps d'accès moyen est le temps que met le disque entre le moment où il a reçu l'ordre de fournir des données et le moment où il les fournit réellement.

La densité radiale est le nombre de pistes par pouce (tpi : Track per Inch).

La densité linéaire est le nombre de bits par pouce sur une piste donnée (bpi : Bit per Inch).

La densité surfacique est le rapport de la densité linéaire sur la densité radiale (s'exprime en bit par pouces carré).

2.8. Les disques compacts

Le CD-ROM (Compact Disc Read Only Memory) est un disque optique de 12 cm de diamètre et de 1mm d'épaisseur, permettant de stocker des informations numériques, c'est-à-dire correspondant à 650 Mo de données informatiques (correspondant à 300000 pages dactylographiées) ou bien jusqu'à 78 min de données audio. Le Compact Disc a été inventé par Sony et Philips.

2.8.1. La composition d'un CD-ROM

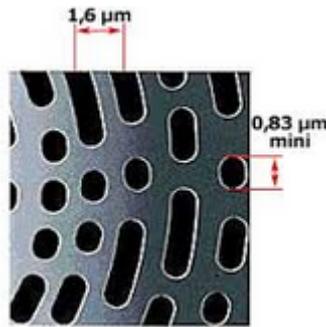


Fig. 2.21: surface d'un CD-ROM

Le CD est constitué de matière plastique, recouvert d'une fine pellicule métallique sur une des faces. Les pistes sont gravées en spirales, ce sont en fait des alvéoles d'une profondeur de 0,83 microns et espacées de 1,6 microns. ces alvéoles forment un code binaire, une alvéole correspond à un 0, un espace à un 1.

Exemple : prenons la séquence suivante : 110010101. Celle-ci correspond sur le CD-ROM à deux espaces, deux trous, un espace, un trou, un espace, un trou, un espace. séquence binaire d'un CD-ROM

On a ainsi une séquence binaire que le lecteur parcourt grâce à un laser; celui-ci est réfléchi lorsqu'il rencontre un espace, il ne l'est pas lorsqu'il rencontre une alvéole.

2.8.2. Le lecteur de CD-ROM

C'est une cellule photoélectrique qui permet de capter le rayon réfléchi, grâce à un miroir semi réfléchissant. Un chariot permet de déplacer le miroir de façon à pouvoir accéder au CD-ROM en entier.

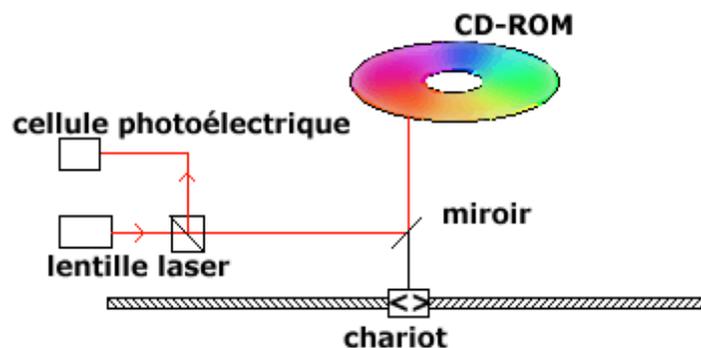


Fig. 2.22: Fonctionnement d'un lecteur de CD-ROM

Il est ainsi possible de stocker sur ce support des musiques, des images, des vidéos, du texte et tout ce qui peut être enregistré de façon numérique.

Le lecteur CD-ROM est caractérisé :

- par sa vitesse : celle-ci est calculée par rapport à la vitesse d'un lecteur de CD-Audio (150 Ko/s). Un lecteur allant à 3000Ko/s sera caractérisé de 20X (20 fois plus vite qu'un lecteur 1X);
- par son temps d'accès. C'est le temps moyen qu'il met pour aller d'une partie du CD à une autre;
- par son type : ATAPI (IDE) ou SCSI.

2.8.3. Le DVD-ROM

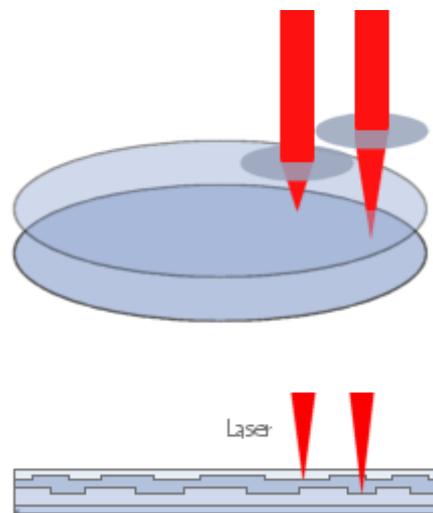


Fig. 2.23: Fonctionnement d'un lecteur de DVD-ROM

Le DVD-ROM (Digital Versatile Disc Read Only Memory) est une variante du CD-ROM dont la capacité est largement plus grande que celle du CD-ROM. En effet, les alvéoles du DVD sont beaucoup plus petite (0,4 microns et un espacement de 0.74 microns), impliquant un laser avec une longueur d'onde beaucoup plus faible.

Les DVD existent en version « double couche », ces disques sont constitués d'une couche transparente à base d'or et d'une couche réflexive à base d'argent. Pour aller lire ces deux couches le lecteur dispose de deux intensités pour le laser. Avec une intensité faible, le rayon se réfléchit sur la surface dorée. Lorsqu'on augmente cette intensité le rayon traverse la première couche et se réfléchit sur la surface argentée.

L'intérêt du DVD touche en priorité le stockage vidéo qui demande une place de stockage importante. Un DVD de 4,7 Go permet de stocker plus de deux heures de vidéo compressées en MPEG-2 (Motion Picture Experts Group), un format qui permet de compresser les images tout en gardant une très grande qualité d'image.

2.8.4. Les zones

Les DVD Vidéo sont conçus pour n'être consultables que dans certaines régions du monde : c'est le découpage en zone (qui « empêche » le piratage). Il est ainsi théoriquement impossible de lire un DVD d'une zone en étant dans une autre. Heureusement, les lecteurs de DVD pour PC peuvent les

lire grâce à des utilitaires.

Type de support	Capacité	Temps musical équivalent	Nb. CD équivalent
CD	650Mo	1h18 min	1
DVD simple face simple couche	4.7Go	9h30	7
DVD simple face double couche	8.5Go	17h30	13
DVD double face simple couche	9.4Go	19h	14
DVD double face double couche	17Go	35h	26

Tab. 2.1: Différents types de DVD

3. Le système d'exploitation

3.1. Description

Pour qu'un ordinateur soit capable de faire fonctionner un programme, il faut que la machine puisse effectuer un certain nombre d'opérations préparatoires afin d'assurer les échanges entre l'unité centrale, la mémoire, et certains périphériques.

Or, les périphériques varient d'un ordinateur à un autre, particulièrement sur un PC. Il faut donc des pilotes de périphériques (instructions servant à piloter un périphérique, en anglais drivers) différents selon les types d'ordinateur et les types de périphériques. C'est le système d'exploitation (en anglais operating system, souvent abrégé en OS) qui assure ces tâches de liaison entre le matériel, l'utilisateur et les applications (traitement de texte, jeu, ...). Il permet de "dissocier" les programmes et le matériel, ce qui simplifie grandement le développement de logiciels.

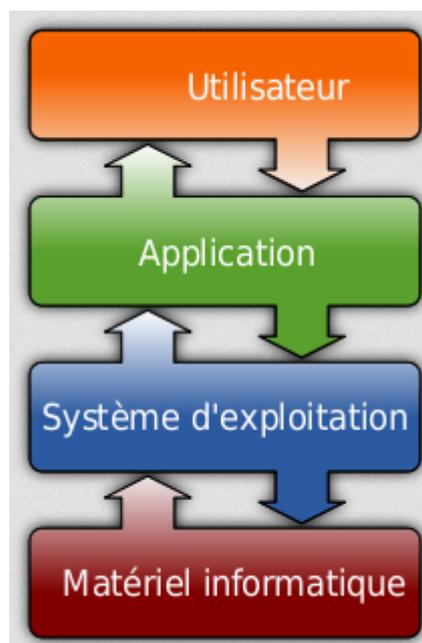


Fig. 3.1: Système d'exploitation

D'autre part, le système d'exploitation fournit un certain nombre d'outils pour gérer la machine. Il assure l'initialisation du système après une mise sous tension. Grâce à des routines (drivers ou gestionnaires de périphériques) il peut gérer les périphériques, en assurant des opérations aussi simple que l'affichage des caractères à l'écran ou bien la lecture du clavier, mais aussi le pilotage d'une imprimante ou d'un scanner...

En effet, les périphériques d'entrée-sortie (par exemple les cartes d'extension) varient d'un modèle d'ordinateur à un autre, il faut donc un système qui puisse unifier l'écriture des instructions gérant le matériel. Ainsi lorsqu'un programme désire afficher des informations à l'écran, il n'a pas besoin d'envoyer des informations spécifiques à la carte graphique (il faudrait que chaque programme prenne en compte la programmation de chaque carte...), il envoie les informations au système d'exploitation, qui se charge de les transmettre au périphérique concerné...

La communication avec le système d'exploitation s'établit par l'intermédiaire d'un langage de commandes et un interpréteur de commandes, cela permet à l'utilisateur de piloter les périphériques en ignorant tout des caractéristiques du matériel qu'il utilise, de la gestion des adresses physiques...

3.1.1. Systèmes multi-tâches

Les systèmes d'exploitation multi-tâches permettent de partager le temps du processeur pour plusieurs programmes, ainsi ceux-ci sembleront s'exécuter simultanément.

Pour réaliser ce processus, les applications sont découpées en séquence d'instructions que l'on appelle tâches ou processus. Ces tâches seront tour à tour actives, en attente, suspendues ou détruites, suivant la priorité qui leur est associée ou bien séquentiellement.

Un système est dit préemptif lorsqu'il possède un ordonnanceur (aussi appelé planificateur), qui répartit, selon des critères de priorité le temps machine entre les différentes tâches qui en font la demande.

Le système est dit à temps partagé lorsqu'un quota de temps est alloué à chaque processus par l'ordonnanceur. C'est notamment le cas des systèmes multi-utilisateurs qui permettent à plusieurs utilisateurs d'utiliser simultanément sur une même machine des applications similaires (le système est alors dit "système transactionnel") ou différentes. Le système alloue alors à chaque utilisateur une tranche de temps.

3.2. Systèmes multi-processeurs

Ces systèmes sont nécessairement multi-tâches puisqu'on leur demande d'une part de pouvoir exécuter simultanément plusieurs applications, mais surtout d'organiser leur exécution sur les différents processeurs (qui peuvent être identiques ou non). Ces systèmes peuvent être soit architecturés autour d'un processeur central qui coordonne les autres processeurs, soit avec des processeurs indépendants qui possèdent chacun leur système d'exploitation, ce qui leur vaut de communiquer entre eux par l'intermédiaire de protocoles.

3.3. Les types de systèmes d'exploitation

On distingue plusieurs types de systèmes d'exploitation, selon qu'ils sont capables de gérer simultanément des informations d'une longueur de 16 bits, 32 bits, 64 bits ou plus.

Systeme	Codage	Mono-util.	Multi-util.	Mono-tâche	Multi-tâche
MS-DOS	16 bits	X		X	
Windows 95/98/Me	32 bits	X			coopératif
Windows NT/2000/XP	32 bits		X		coopératif
Unix	32 bits		X		préemptif
MAC/OS	32 bits		X		coopératif
MAC/OS X	32 bits		X		préemptif
VMS	32 bits		X		préemptif

Tab. 3.1: Différents systèmes d'exploitation

3.4. Organisation des données persistantes

3.4.1. Les fichiers

Un fichier est une suite d'informations binaires, c'est-à-dire une suite de 0 et de 1. Ce fichier peut être stocké pour garder une trace de ces informations. Un fichier texte est un fichier composé de caractères stockés sous la forme d'octets.

Ce fichier est enregistré sous la forme « nom_du_fichier.ext ». La partie .ext représente l'extension c'est un moyen de reconnaître le type de programme avec lequel ce fichier peut être ouvert (attention cela ne garantit pas le type de fichier : lorsque l'on change l'extension on ne change pas le type de fichier). La longueur du nom et de l'extension peut varier suivant le système d'exploitation :

- 8 caractères pour le nom et 3 pour l'extension sous DOS et Windows 3.1
- 256 caractères pour le nom et l'extension sous Unix, Windows 95, 98 et NT

De plus, un fichier contient un en-tête qui permet de stocker des informations supplémentaires, comme le type de fichier et surtout la taille du fichier. Il contient aussi un caractère de fin de fichier signalant que les informations situées au-delà de ce caractère ne font plus partie du même fichier.

3.4.2. Les répertoires

Un répertoire est un objet informatique qui contient des fichiers. Imaginez une grande commode qui contient des tiroirs dans lesquels pourraient se trouver des fichiers et d'autres tiroirs... Un répertoire peut donc contenir :

- des fichiers
- d'autres répertoires



Fig. 3.2: Répertoires sous MacOS X, Unix et Windows

Si l'on reprend notre exemple de la commode, la plus grande entité contenant d'autres entités est la commode : elle ne peut pas se trouver dans un tiroir! Dans le cas de l'informatique on appelle cette entité la racine : c'est l'entité de plus bas niveau, car elle peut contenir des fichiers ou des répertoire mais ne peut pas se trouver dans un répertoire elle-même. On la note \ dans la plupart des systèmes d'exploitation (pour y aller sous le shell DOS on tapera `cd \`. Il en existe une seule par volume de données (ou du moins par partition ...).

Il est à noter que le volume correspondant à la disquette est noté A:, et que le volume correspondant au disque dur est noté C:. La racine à partir de laquelle tous les répertoires et fichiers de la disquette sont stockés est donc notée A:\. La racine du disque dur est C:\. Pour chaque nouvelle unité de stockage, on utilisera une nouvelle lettre. Souvent, la racine des CD-ROM est par exemple notée E:\.

Un répertoire qui en contient un autre est dit "répertoire parent". Lorsque d'un répertoire on veut aller au répertoire parent, celui-ci est désigné par .. sur la plupart des systèmes (on tapera donc `cd ..` sous DOS ou sous UNIX pour accéder à un répertoire parent).

Considérons La représentation ci-dessus d'un système de répertoires sous Windows. Ici, « Programmes » est parent de « Adobe ». Vis-à-vis de « Adobe », « Programmes » pourra être notée « .. » La racine (C:\), vis-à-vis de « Adobe », pourra être notée « ..\.. » car deux relations de parenté les séparent.