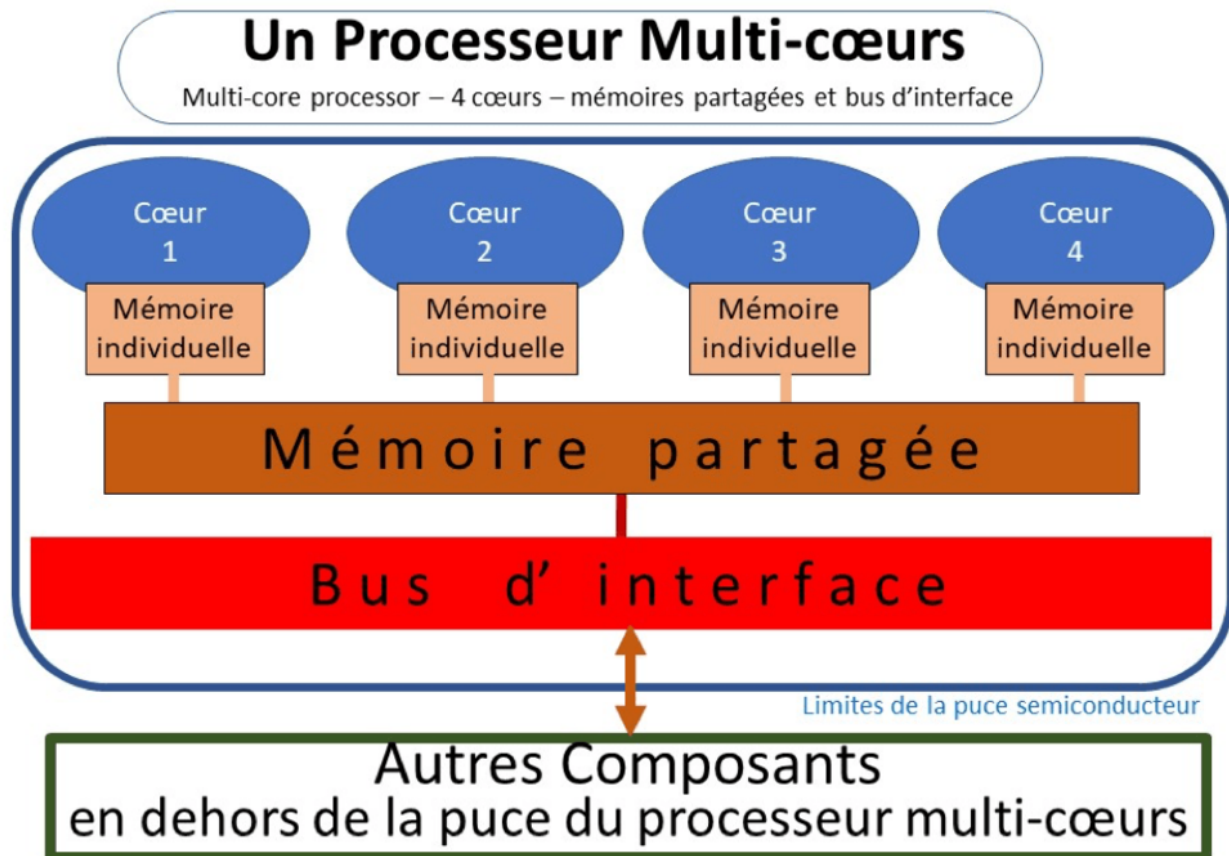


Machines et personnalités : L'architecture des ordinateurs parallèles (4/10 année 2019)

Publié le 30 avril 2019



Schémas traduit de :
Anil Sethi, Himanshu Kushwah, (University of Delhi, India), MULTICORE PROCESSOR TECHNOLOGY- ADVANTAGES AND CHALLENGES, IJRET, Volume: 04-09, 2015, (p. 87-89)

Résultant d'un dialogue entre Jean Ricodeau et Philippe Denoyelle, ACONIT

Photo d'en-tête : Schéma de principe d'un processeur multi-cœurs

Au cours de nos recherches destinées à mettre en lumière comment était constituées les anciennes machines informatiques, nous nous sommes interrogés sur le sens réel du terme « multi-cœurs ». Ce terme nous a interpellé. En effet, qu'est-ce qu'un cœur en informatique ? Et est-ce que ce terme est compréhensible pour le plus grand nombre ? Après nous être informés dans le cadre de nos activités à l'ACONIT, nous avons trouvé un début de réponse satisfaisante. Et cette réponse nous a ouvert de nouvelles perspectives sur la vie secrète des ordinateurs.

En effet, une telle question revient à nous interroger sur le sens réel de la programmation parallèle, qui est couramment utilisée de nos jours lorsqu'il est demandé à une machine informatique de mener plusieurs activités en simultanée.

Mais prenons un exemple historique :

Au temps du calcul à la mine de plomb et à la plume d'oie

Vers le XVIII^e siècle, le besoin de tables de navigation va entraîner la première demande de calcul « intensif ». Chaque ligne de la table (logarithme, trigonométrie...) nécessite une suite fastidieuse d'opérations pour atteindre la précision recherchée, et il y a des milliers de lignes à calculer.

Le temps de calcul est énorme... Mettons 40 personnes à calculer en même temps, 80 s'il le faut... Chacun prend en charge une ligne et on divise par autant le temps de calcul...

Une terminologie simplifiée:

*Nous désignerons comme « **processus** » la séquence de calcul qu'exécute chacune des personnes. C'est une suite d'opérations de base (addition, multiplication...) qui réalise une certaine recette mathématique (autrement dit, pour les puristes, un algorithme ou une heuristique).*

*On désignera par « **processeur** » ce qui exécute le processus, ici un homme...*

Le système est efficace, réel, mais sa limite est évidente : il faut que les calculs soient indépendants, que personne n'ait à attendre le résultat d'un autre calcul... Les informaticiens aiment à rappeler que l'on n'a jamais pu réduire le temps de gestation d'une mère en mettant plusieurs femmes en parallèle...

Les ordinateurs

Les premiers ordinateurs ont fait suite aux machines à calculer géantes destinées aux calculs des tables de tir pendant la guerre de 1940-45, mais ceci est une autre histoire. Ces ordinateurs avaient l'énorme avantage de pouvoir effectuer les calculs les plus divers, sans fatigue et sans erreurs (enfin presque). Mais finalement, ils n'étaient pas si rapides que ça ! Les premières machines avaient des temps d'instruction de l'ordre de la milliseconde et surtout, elles s'arrêtaient quasiment pour les opérations d'entrée/sortie... Si l'on devait envoyer 100 caractères sur la machine à écrire de l'opérateur, à 10 c/s (caractères par seconde), l'ordinateur ne faisait rien d'autre pendant 10 secondes ! Et ne parlons pas de l'attente sur les entrées d'information !

Le premier perfectionnement fut d'ajouter une « unité d'entrée/sortie » à l'unité « centrale ». L'unité centrale indiquait à l'unité d'E/S l'adresse mémoire de la chaîne de caractères à sortir ainsi que le nombre de caractères à traiter, et pouvait reprendre le calcul sans attendre. On voit que l'on a ici deux processeurs séparés, mais avec des fonctions bien spécialisées.

Mais malgré les progrès de la technologie, les très grosses machines classiques (appelées « main frames ») ont inévitablement commencé à plafonner en vitesse. Pour y remédier, leurs concepteurs ont eu l'idée de mettre plusieurs organes de calcul spécialisés (additionneurs, multiplieurs) côte à côte pour exécuter simultanément quelques opérations de base. La machine emblématique de cette approche nouvelle fut le Control Data 6600

de Seymour Cray, développé vers 1970. Les Américains parlent à leur sujet de « multi threading ». C'est à peu près à cette époque qu'est apparu le parallélisme élémentaire, avec des ordinateurs à structure double, comme l'Iris 80 de CII.

Encore un peu de terminologie :

*Voici un nouveau terme, anglais, « **thread** ». Littéralement c'est un fil. Disons simplement que c'est une suite d'opérations indissociables. On parlera généralement en français de « **tâche** ».*

L'arrivée des microprocesseurs

Les progrès des semi-conducteurs vont conduire rapidement aux circuits intégrés comportant de plus en plus de transistors, et enfin au micro-processeur : la partie centrale d'un ordinateur sur une seule puce. C'est bien ici un « processeur » dans un seul boîtier.

Et l'histoire se répète : il faut libérer ce processeur des entrées/sorties. On va développer des circuits intégrés spécialisés en E/S : les « UART », les « DMA ».

Et arrivent les « coprocesseurs », associant deux processeurs en parallèle. D'abord le micro-processeur principal associé à un processeur spécialisé pour les calculs arithmétiques et la « virgule flottante* ». Ensuite sont arrivés les processeurs spécialisés « graphiques » poussés en avant par les jeux sur ordinateur !

Une nouvelle limite : thermique !

Très peu de temps après sont arrivés sur le marché américain les microprocesseurs, induisant une nouvelle course à la vitesse, car la réduction de la taille de ces éléments permettait d'accélérer considérablement les horloges (réduction des temps de transmission et des capacités parasites).

Vers 1995, pourtant, un nouveau problème s'est posé : on constate en effet que ces microprocesseurs de plus en plus rapides chauffent trop. D'où l'idée de ne plus augmenter la vitesse, mais de mettre deux processeurs complets dans un même boîtier. On va alors parler de « cœur » pour ce nouveau type de configuration, puisque le terme 'processeur' est indument appliqué au boîtier lui-même.

Halte ! Terminologie !

Nous nageons ici dans la confusion :

- en langage technico-commercial, on confond maintenant 'processeur' et 'boîtier microprocesseur' - et on parle de 'cœur' pour la micro-structure interne, qui est en réalité un 'processeur' au sens propre du terme.

Un microordinateur portable Mac Book contient un 'processeur' à deux 'cœurs'. Pour comparaison, le Mac Pro, professionnel, contient pour sa part deux 'processeurs' à quatre 'cœurs'. Ce genre de machine est souvent équipé d'outils permettant d'observer la répartition de la charge de travail auprès des différents cœurs. Comment se répartit cette charge ? Le système d'exploitation, Mac OS 11.6, dans l'environnement de ces deux exemples, a de nombreuses tâches à effectuer et les répartit entre les cœurs : sur le MacBook précédemment cité, à l'instant où j'écris, elle est divisée en 1051 'threads'/tâches.



Evolution de la configuration des ordinateurs parallèles (photos ACONIT)
 Photo de gauche : Face avant du MégaNode - Photo de droite : Face arrière
 du MPP12000

Du parallélisme au parallélisme massif

Pendant que s'opéraient les évolutions évoquées plus haut, les études ont continué sur le parallélisme. Ainsi, on est monté progressivement en capacité, ce dont témoignent plusieurs items de la collection de l'ACONIT : le Volvox comportait 12 processeurs, le MégaNode, développé en partie à Grenoble, en comportait 64×8 , soit 512 processeurs, tout comme le MPP12000, dans une configuration légèrement différente (16×32). Puis on est monté progressivement vers le massivement parallèle, qui s'est aujourd'hui généralisé pour les plus gros centres de calcul.

À côté de chaque système parallèle est adjoint un ordinateur de pilotage qui fournit les tâches complètes à chaque processeur (un DECstation 5000 était utilisé pour le MPP12000). Ce système s'applique particulièrement bien aux calculs tels que ceux produits pour les cellules d'atmosphère pour la météorologie, ou pour les cellules simulant une réaction atomique : dans ces cas, chaque processeur calcule ce qui se passe dans sa cellule et échange le résultat avec les cellules adjacentes.



Photo : une des cartes supportant les micro-processeurs travaillant en parallèle (photos ACONIT)

Et les ordinateurs neuronaux ?

Dans le même temps (soit au début des années 1980) est apparue l'architecture neuronale. En quoi consiste-t-elle ? Dans le principe, il s'agit de machines qui tentent d'imiter le fonctionnement des neurones du cerveau. Chacun de ces neurones simulés reçoit les résultats des autres neurones à travers des 'synapses' (sortes d'interfaces du système nerveux) auxquelles sont attribués certaines pondérations. Cette approche nécessite donc une phase d'apprentissage pour chaque ordinateur, phase durant laquelle on fait varier les pondérations afin d'obtenir le résultat recherché pour un certain jeu d'entrées... Les chercheurs ont commencé par simuler les neurones par des circuits analogiques. Puis ils sont passés à des simulations réalisées par des armées de microprocesseurs : le fameux Mind 1024 (voir l'article Echosciences du 21/12/2018 qui lui est consacré) en contient 64 qui contrôlent effectivement 1024 neurones.

Mais dans ce cas, ce n'est plus du parallélisme. L'approche qui s'en dégage est complètement différente et s'adapte particulièrement bien à la reconnaissance d'images, au calcul des jeux (échec, Go...) et à bien d'autres activités de traitement d'information en nombre, mais avec un inconvénient majeur : il est devenu impossible de savoir ce qui se passe réellement entre les 3 ou 4 couches de neurones qu'on empile maintenant les unes au-dessus des autres...

CONCLUSION :

Toutes ces évolutions ont conduit à ce qu'est devenue l'informatique aujourd'hui : une puissance de calcul phénoménale. Son histoire ne s'arrête pourtant pas là, et la prochaine étape se dessine déjà avec les projets de recherche autour des ordinateurs quantiques. Mais ceci est une autre histoire, comme disait Rudyard Kipling.

Note :

*La **virgule flottante** est une méthode d'écriture des nombres réels utilisée dans les ordinateurs, qui consiste à décomposer un nombre réel en deux parties, par exemple un octet va contenir P « l'exposant » et les 3 octets suivants vont contenir M « la mantisse » et le nombre sera interprété par l'ordinateur comme : $M \times 2^{\text{puissance } P}$.

C'est l'équivalent en base 2 de la « notation scientifique » en base 10 : $1,23456 \times 10^{\text{puissance } 12}$, couramment noté 1,23456 E12

Quelques pistes **pour en savoir plus** :

voir les fiches machines des ordinateurs multi-tâches :

- CDC-6600 (pas d'exemplaire à ACONIT)
- AC 20239 (Volvox)
- AC 18434 (Meganode)
- AC 12169 (MPP12000)
- AC 20253-01 (MIND 1024)
-