# INTRODUCTION

De nombreuses decenies apres la naissance des ordinateurs qui  
realisaient deja la plus part des taches pour lesquels ils etaient  
construit, au meme titre que leur minuaturisation vers laquelle les  
concepteurs ou alors les grandes firmes se penchaient, l’on desirait realiser dans une machine plusieurs taches simultanees et “independantes” ce qui la rendra plus rapide et performante. C’est de là qu’est nee le parallelisme avec l’apparition dans les annees 80, de la fameuse **“connexion machine”** qui était un gros ordinateur composé deplusieurs processeurs, coûtant jusqu'à un million de dollars. Mais, ce n’est que dans les annees 2000 avec l’avement des processeurs multicoeurs que l’on integre reellement le parallelisme dans nos differents equipements avec l’execution des taches de manière simultannee.  
A ce niveau nait une panoplie d’interogations a savoir **dans quels  
cadres le parallelisme pourrait nous etre utile**? Pour etre un peu  
plus pragmatique, **quels sont les differents projets et applications realises uilisant le calcul parallele dans le quotidien**?l’interet etant saisie, **comment fonctionne le calcul parallel**? Plus loin, **en quoi est-il avantagieux par rapport au paradigme sequentiel traditionnel** et **quels en sont donc les avantages et les limites**? Pour ainsi finir, **si l’on voudrait le realiser, comment procederons nous**?  
Tout ce chapitre sera axé sur ces questions.Nous presenterons tout d’abord le principe et les applications du calcul prallel, puis les differentes classfications et modeles du calcul parallel, apres quoi viendra la presentation de l’interface MPI necessaire dans l’architecture parrallele; ensuite, nous realiserons quelques programmes parallels avec OpenMP et en fin presenterons les avantages et limites du parallelisme.

# PRINCIPE ET APPLICATION DU CALCUL PARALLEL

1. Definitions

Avant de donner le principe du calcul parallel, il est important que nous comprenions et fixions

certains termes qui d’ailleurs tout au long de ce chapitre nous serons tres utiles.

**La Mémoire :** c’est un dispositif électronique servant à stocker des informations. Elle est dite partagée lorsque plusieurs processeurs peuvent y accédez simultanément et distribuée lorsqu’elle est répartie en plusieurs nœuds, chaque portion n'étant accessible qu'à certains processeurs.

**Un processeur** quant à lui est un composant présent dans de nombreux dispositifs électroniques, ayant pour rôle d’exécuter les instructions machines et les programmes informatiques. Il est dit **multi-cœurs** s’il possède plusieurs cœurs physiques fonctionnant simultanément ; un cœur physique étant tout simplement un ensemble de circuit capable d’exécuter des programmes de façon autonome(realisant les instruction de Lecture, de decodage, d’ecriture et de rangement).

**Une architecture en informatique** désigne l’organisation des différents éléments d’un système  
informatique (logiciel, matériel, humain, information...), ainsi que les relations entre eux.

**Un processus** est l’instance d’un programme en cours d’exécution. Et un **Thread** n’est rien d’autre qu’un processus léger.

**Une tache** quant a elle est une liste d’instructions réalisant un certain travail et utilisant des  
données en entrée dans un ensemble E (ensemble des variables d’entrées) et qui modifient un ensemble de variable S (variables de sortie). Selon que la tâche soit petite ou grosse (en terme de nombre d’instructions), on parlera de granularité fine ou grossière.

**Le Parallélisme en informatique** consiste à mettre en œuvre des architectures d’électroniques  
numériques permettant de traiter les informations de manière simultanée ainsi que les algorithmes spécialisés pour celles-ci

**Un calcul** est une opération ou un ensemble d’opérations réalisées sur des grandeurs (les nombres par exemple). Il est dit **parallèle** lorsque l’on lui applique les principes du parallélisme, et ceux dans l’optique d’obtenir plus rapidement le résultat.

**Ordinateur parallèle** : c’est un ordinateur composé de plusieurs processeurs qui coopèrent à la  
solution d’un même problème.

**Un système distribué (ou reparti)** : est un système composé de plusieurs unités de calculs impliquées dans la résolution d’un ou plusieurs problèmes.

**Le partitionnement** : C’est le découpage du problème complet en tâches, avec l’espoir que certaines de ces tâches pourront être résolues en parallèle. Souvent, dans les cas SPMD (Small Program Multiple Data), le Partitionnement revient à découper le domaine de calcul (c’est l’ensemble des variables d’entrée qui est fractionné). On parle de **« Domain décomposition »**.

**Variable partagée** : c’est une variable dont le nom permet d’accéder au même bloc de stockage au sein d’une région parallèle entre tâches.

**Variable privée** : c’est une variable dont le nom permet d’accéder à différents blocs de stockage

suivant les tâches, au sein d’une région parallèle.

## **Principe du calcul parallel**

Tout d’abord pourquoi le parallelisme? Nous avons dans l’introduction donne une ebauche des raisons qui poussent a entreprendre des calculs paralleles.En effet, comme dit, ils sont fortement connexes aux architectures:

* Premièrement, le parallélisme est présent au plus profond de la  
  microarchitecture du processeur. Autrefois, les processeurs exécutaient des  
  programmes en répétant ce qu'on appelle le cycle d'instructions, une  
  séquence de quatre étapes : (i**) lecture et décodage d'une instruction** ;  
  (ii) **trouver les données nécessaires pour traiter l'instruction** ; (iii) **le  
  traitement de l'instruction** ; et (iv**) écrire le résultat dans la  
  memoire**. Étant donné que l'étape (ii) a introduit de longs retards dus à  
  l'arrivée des données, une grande partie de la recherche s'est  
  concentrée sur des conceptions qui ont réduit ces retards et de cette  
  façon augmenter la vitesse d'exécution effective des programmes. Au fil  
  des années, cependant, l'objectif principal est devenu la conception  
  d'un processeur capable d'exécuter des instructions simultanément. Le fonctionnement d'un tel processeur a permis détection et exploitation du parallélisme inhérent à l'exécution des instructions. Ces processeurs ont permis des vitesses d'exécution des  
  programmes encore plus élevées, quel que soit le fréquence du processeur et de la mémoire.
* Deuxièmement, tout ordinateur commercial, tablette et smartphone contient un processeur avec plusieurs cœurs, chacun étant capable d'exécuter sa propre instruction flux. Si les flux sont conçus de manière à ce que les cœurs collaborent à l'exécution d'unel'application, cette derniere s'exécute en parallèle et peut être considérablement accélérée.
* Image de representation
* Troisièmement, de nombreux serveurs contiennent plusieurs processeurs  
  multicœurs. . Un tel serveur est capable d'exécuter un service en parallèle,  
  ainsi que plusieurs services en parallèle.

Image de representation

* Enfin, même les ordinateurs grand public contiennent des processeurs  
  graphiques capables d’exécuter des centaines voire des milliers de threads  
  en parallèle. Processeurs capables de faire face à un si grand parallélisme  
  sont nécessaires pour prendre en charge l'animation graphique.
* Image de representation

Il y a de ce fait plusieurs raisons qui ont pousse a ce que l’on parallelise nos  
machines :

* Tout d’abord, il n'est pas possible d'augmenter indéfiniment les fréquences  
  du processeur et de la mémoire, du moins pas avec la technologie actuelle  
  à base de silicium. Par conséquent, pour augmenter la puissance de calcul  
  des ordinateurs, de nouvelles architectures et organisations des notions  
  sont nécessaires.

Nous pouvons en occurrence enumererle projet[**Word community grid**](https://www.worldcommunitygrid.org/)de **IBM**, qui aux travers des appareils(smart phones, tablettes, ordinateurs) de volontaires executent en arriere plan jeux des experiences et tests dans le cadre de la covid-19.





* Enfin, le parallélisme est devenu une partie intégrante de tout ordinateur et  
  il est probable qu'il restera inchangé du fait de la simple inertie : le parallélisme peut se faire et il se vend bien.

Le calcul parallel est donc tout simplement la realisation des differentes taches d’une operation de maniere simultanee.Ici, l’on decoupe le probleme initial en plusieurs sous problemes plus ou moins simple à resoudre confié chacun à differentes unités de traitement (celle-ci peuvent être soit des processeurs, soit des cœurs) pouvant se partager les donnees ou pas et qui sont d’une manière ou d’une autre liées entre elles afin que le résultat global attendu de l’exécution soit la resolution du probleme initial. Pour donc realiser cette execution sur differentes unites de traitement, plusieurs types de parallelisme relatifs aux architectures des composants sont mis sur pied.On en distingue plusieurs type de parallelismes:

* **Le parallelisme des taches** dans lequel un ensemble de taches sont realisees par differents UC
* **Le parallélisme de données** dans lequel un ensemble de tâches est répété sur plusieurs données.

# I. Applications du parallelisme

Tout d’abord, il est necessaire de prendre connaissance ce qui a déjà etet realise en utlisant le parallelelisme.

## 1. Le parallelisme dans la geologie

La géologie est la [science](https://fr.wikipedia.org/wiki/Science) dont le principal objet d'étude est la [Terre](https://fr.wikipedia.org/wiki/Terre), et plus particulièrement la [lithosphère](https://fr.wikipedia.org/wiki/Lithosphère). [Discipline](https://fr.wikipedia.org/wiki/Discipline_(spécialité)) majeure des [sciences de la Terre](https://fr.wikipedia.org/wiki/Sciences_de_la_Terre), elle se base en premier lieu sur l'observation, puis établit des hypothèses permettant d'expliquer l'agencement des roches et des structures les affectant afin d'en reconstituer l'histoire et les processus en jeu. Ainsi, la detection sismique par exemple prenant en compte de nombreux parametres sur un temps relativement long, necessite de grandes puissances de calculs pour l’analyse et l’interpretation des metadonnees (la startigraphie, la poreusite du sol, les differents nimearux constitutis de la region etudie, les pressions,etc)suivant leurs diferrentes projection a l’aide des macines vectorielle qui utilisent des algorithmes parallels(Demontre par un groupe de chercheur de l’universite d’Algarve au Portugal utilisant des )

Nous pouvons donc ici enumerer la platefome [Dug Software](http://dug.com/) qui a l’aide de son super ordinateur (d ‘ailleur l’un des plus grand su monde avec des equipements depoye sur plus de la motie d’un stade de football) analyse a partir d’algorithmes parallels et donne des prediction sismiques sur des reliefs.

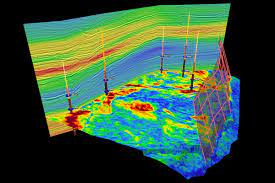
Figure 1: Interpretation sismique de Dug Software

Figure 2: Equipement de Dug Software

## **2. DepBlue et le paralleleisme**

Dans le contexte des jeux avec ordinateur oul celui-ci est appele a interagir rapidement, il necessite grande puissances de calcul et une rapidite extraordinaire.Ces caracteristique lors du jeux peuvent etre realisee grace a l’execution en parallel des taches dans l’optique d’avoir des resultats rapidement.Pour le visuliser prenons le cas des **jeux d’echecs**.

Dans ce jeux, a tout instant de reactions, l’on doit evaluer un ensemble de combinaisons possibles de deplacements pour ganger en un temps tres court.Alors, a partir de chaque pion, l’on pourrait evaluer simultanement et les combinaisons possibles pour la victoire.

Un super-ordinateur déjà realise pour cela est [**Deep Blue**](https://www.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/) d’IBM qui, pouvant calculer plus de 200 millions de combinaisons a un moment, a puis battre en 1997 le champion en tire des jeux d’echecs Garry Kasparov.

Figure 3: Garry Kasparov Vs Deep Blue

3. Le parallelisme dans les marches boursiers et le Business Intelligence(BI)

Dans les differents marches boursiers, les bourses, les investisseurs desirent une certaine assurance de benefice pour pouvoir acheter des actions.Alors, il devra anlysser des montagnes de donnees (les résultats et performances boursières des entreprises, informations financières…) pour debusquer le bon coup avant le reste du marchei) ; c’est la qu’intervient les algorithmes paralleles dont les instamces analyserons chacun un partir des donnees pour une grande rapidite.

De la est nee annees dans les annee 2010 le logiciel [**Lexicon**](https://mendoza.nd.edu/news/new-dow-jones-lexicon-analyzes-patterns-in-the-news-to-create-more-powerful-trading-strategies/)  lance par le  **Dow jones ,** un service d’information financière pour les investisseurs professionnels.Ce service anlyse et prend des decisons sur les differentes actions a acheter en quelques microsecondes, d’ou le terme « **flash trading »**

Par ailleurs, pour des entreprises desirant par exemples lancer des games de produits, doivent etudier les consomations des potentiels clients suivant differents aspects et meme les pouvoir de la concurence.Ceci revient a analyser des metadonnees en peu de temps, de la intervient des calculs paralleles.

Figure 5: Bourses de Wall Street

Figure 4: Analyses dans les marches bousiers

4. Le pareallelisme dans la meteorologie

La difficute ici est l;analyse massive des donnees sur le temps.En effet, pour realiser des predictions meteorologiques, l’on doit prende en compte des parametres telsque les vents, les courants marins, les tempreatures, la position des nuages, leur evolution, etc et ceux suivant la direction, les regions et bien d’autres.Comme precedement, il necissate beaucoup de puissance de calculs qui fourni par des super oridinateurs comme [**Fukagu**](https://www.fujitsu.com/global/about/innovation/fugaku/specifications/)Utilisant des algorithmes paralles.

5. Simulation des dans le domaine de l’aero spatiale

Pour faire simuller des evenements dans l’espace les institutions comme SI, Space X, et la NASA font recours au algorithmes paralleles.Chque instance du programme pouvant gerer un aspect du probleme a resoudre. C’est ainsi que certainnes simulations ont permis recement de decouvrir qu’un asteroide devait heurter la terre ce qui a permi de devier sa trajectoire.

On retirnt que ces applications tournent toutes autour de vitesse d’analyses de donnees massives.