

Formation en Calcul Scientifique - LEM2I

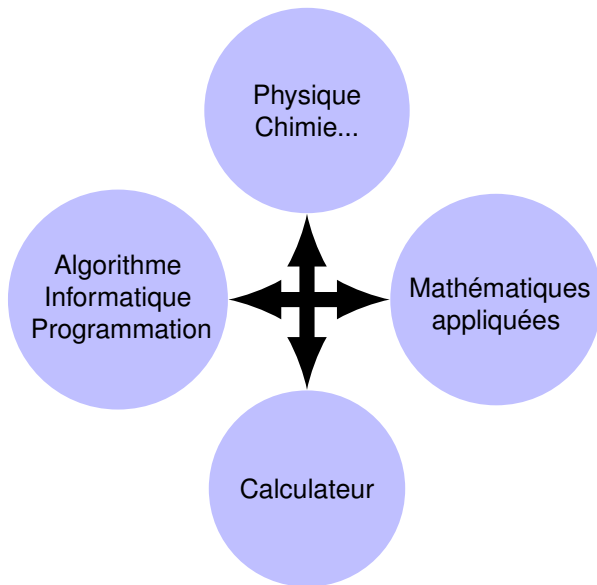
---

# Introduction à l'architecture des calculateurs

Violaine Louvet <sup>1</sup>

<sup>1</sup>Institut Camille Jordan - CNRS

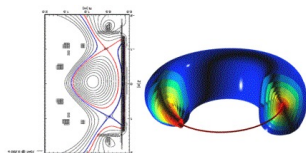
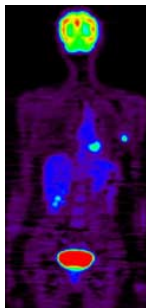
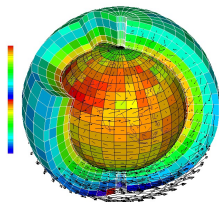
12-13/09/2011



# Simulation numérique ...

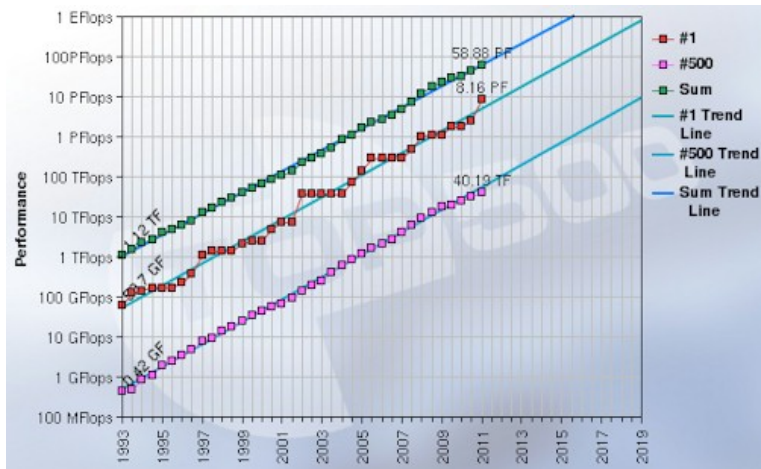
**Complexification** des modèles, **augmentation** de la taille et de la précision des domaines de calcul, **nouvelles disciplines** ... les besoins en simulation numérique et en calculs ne font qu'augmenter :

- Climat
- Médecine
- Sismologie
- Fusion nucléaire ...

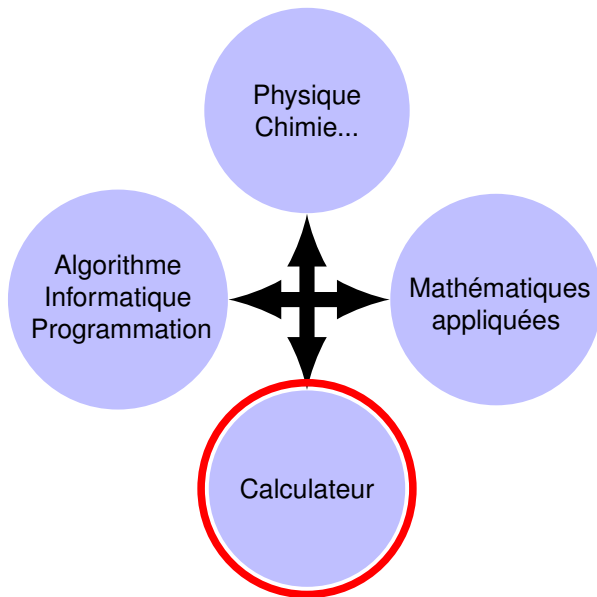


# et puissance de calcul ...

Parallèlement, la puissance de calcul poursuit sa **croissance exponentielle**



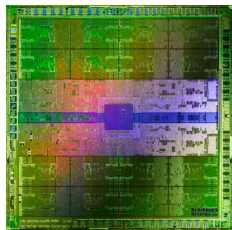
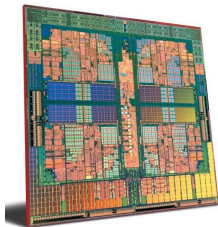
# Ce qui nous intéresse ici ...



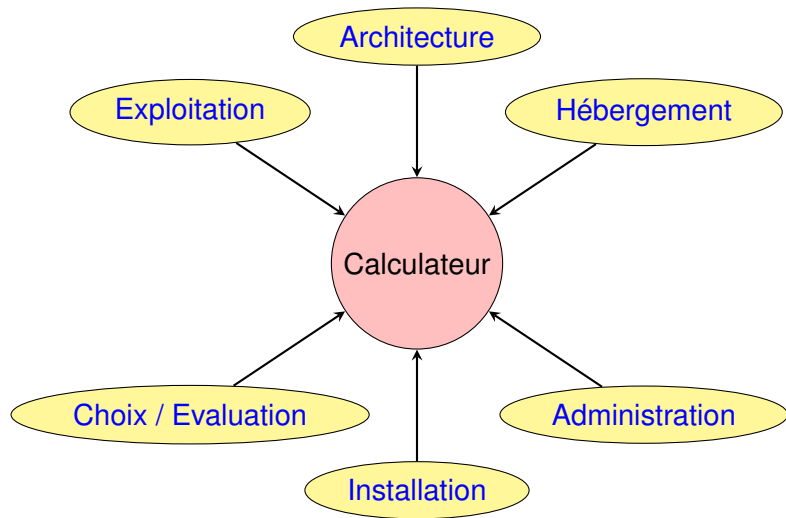
# Importance des aspects matériels

La connaissance des infrastructures et architectures de calcul permet :

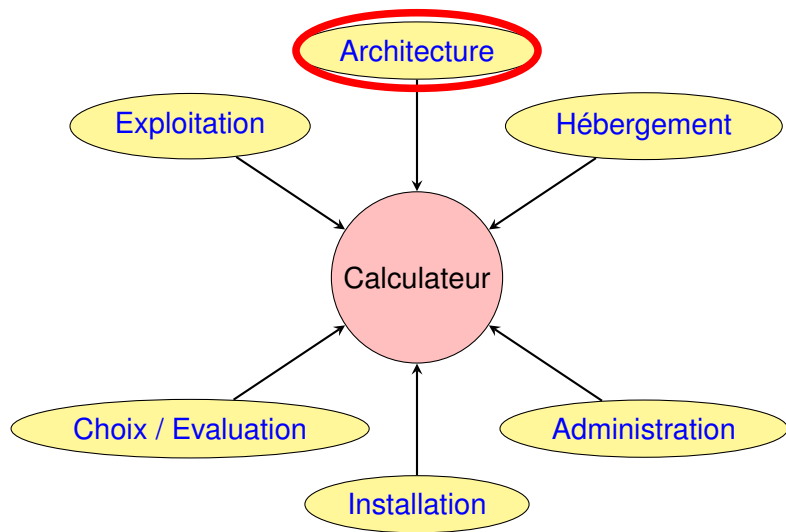
- de **choisir et d'adapter** son infrastructure en fonction de ses besoins
- de la **dimensionner** correctement en retenant le meilleur compromis : architecture équilibrée
- de **comprendre** le comportement d'un programme
- d'**adapter** les méthodes numériques, les algorithmes, la programmation



# Ce n'est pas si simple ...



# Ce n'est pas si simple ...





## 1 Architectures de calcul actuelles

## 2 Architecture d'un serveur

- Processeur
- Mémoire

## 3 Evolutions technologiques et problématiques associées

# Architectures actuelles

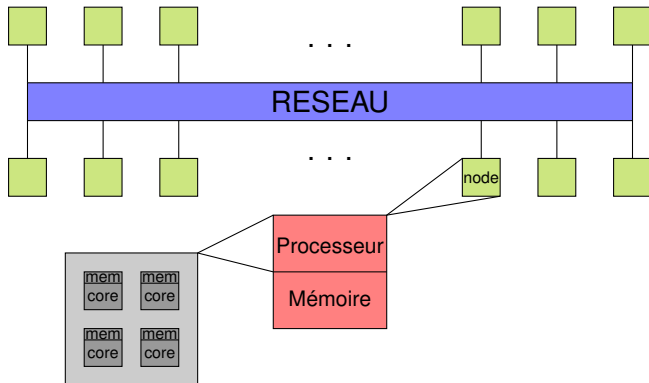
Majoritairement des **clusters** : regroupement de serveurs connectés par un réseau

- Nombre et type des **unités de calcul** (processeurs, cœurs, GPU ...)
- Taille de la **mémoire**
- Type du **réseau** (infiniband, ethernet ...)



Sur quel type de machine calculez vous ?

# Hierarchie des architectures



- 1 Architectures de calcul actuelles
- 2 Architecture d'un serveur
  - Processeur
  - Mémoire
- 3 Evolutions technologiques et problématiques associées

- **bit (b)** : binary digit, plus petite unité d'information d'un composant en informatique. Vaut 0 ou 1
  - 1 bit permet donc d'avoir 2 états (0 ou 1), 2 bits permettent d'avoir 4 ( $2^2$ ) états (00,01,10,11),... ,  $n$  bits permettent ainsi d'avoir  $2^n$  états
- **octet (o) ou byte (B)** = composé de 8 bits

## Normalisation 1998

Depuis 1998 les **préfixes binaires** associés aux octets ou bytes ont été normalisés (kibi = kilo binaire, mébi = méga binaire ...) :

- $1 \text{ kibio} = 2^{10} \text{ octets} = 1024 \text{ octets} = 8192 \text{ bits}$
- $1 \text{ ko} = 1000 \text{ octets}$

Mais attention, de façon courante : 1 kilo-octet =  $1 \text{ Ko} = 1024 \text{ octets}$

## Usage

- **Processeurs et mémoires** opèrent sur des **octets**
- **Débits des réseaux et des bus** exprimés en **bits par seconde**

- 1 Architectures de calcul actuelles
- 2 Architecture d'un serveur
  - Processeur
  - Mémoire
- 3 Evolutions technologiques et problématiques associées

# Caractéristiques d'un processeur

- **Fréquence d'horloge (MHz)** : vitesse de fonctionnement du processeur = nombre de millions de **cycles** que le processeur est capable d'effectuer par seconde
  - **Cycle** = plus petite unité de temps au niveau du processeur. Chaque opération/instruction nécessite au minimum un cycle, et plus souvent plusieurs
  - $1\text{GHz} = 10^9\text{Hz} = 10^9\text{cycle/s}$
- **Largeur (32 ou 64 bits)** : notamment du bus de données et des registres internes. Bon indicateur de la quantité d'information que celui-ci peut gérer en un temps donné
- **Jeu d'instructions** : ensemble des opérations qu'un processeur peut exécuter, plus ou moins complexes

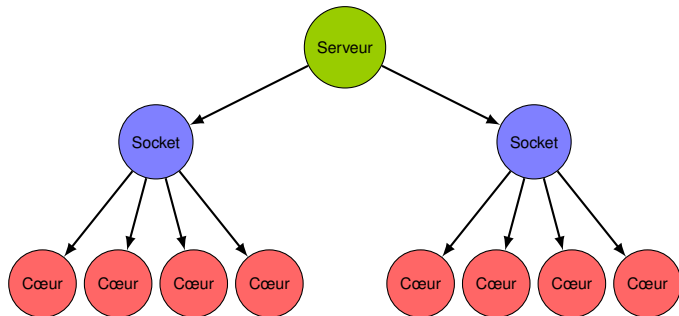
## Adressage mémoire

Les **processeurs 32 bits** ne peuvent pas adresser plus de 4 giboctets ( $2^{32}$ ) de mémoire centrale, tandis que les **processeurs 64 bits** peuvent adresser 16 exboctets ( $2^{64}$ ) de mémoire.

# Cœurs, sockets

On évite de parler de CPUs : il faut distinguer le **support** (**socket**) de l'**unité de calcul** elle-même (**cœur**).

Le socket ou slot est le connecteur qui interfère entre la carte mère d'un ordinateur et le processeur lui-même.



Serveur bi-sockets quadri-cœurs

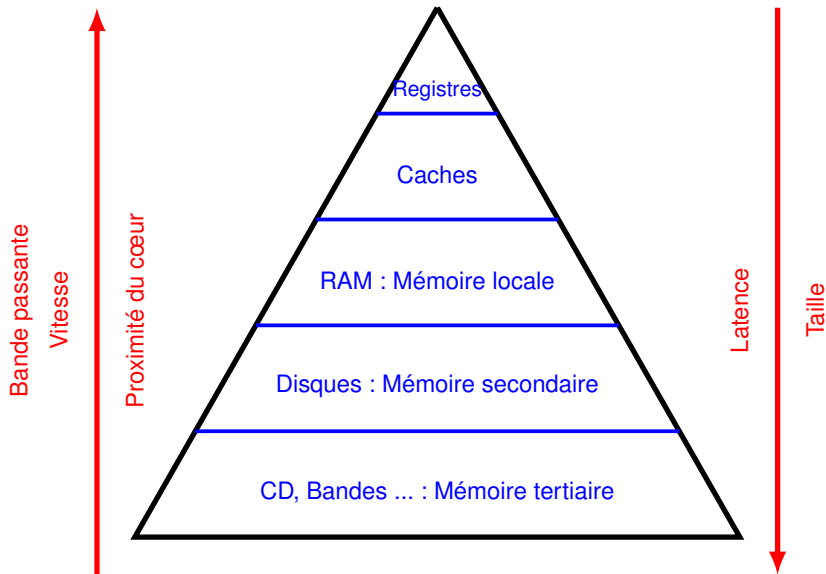


- 1 Architectures de calcul actuelles
- 2 Architecture d'un serveur
  - Processeur
  - Mémoire
- 3 Evolutions technologiques et problématiques associées

# Quelques définitions

- **Bande passante** = débit d'informations ; d'un périphérique : mémoire, disque dur ... ou d'un médium de communication : réseau, bus ....
  - Mesurée généralement en **octets (byte)** par seconde (o/s, ou B/s) ou en **bits** par seconde (bit/s ou bps)
  - **Exemple** : USB 2.0 → 480 Mbits par seconde. Le transfert d'une vidéo de 900 Mo mettra donc 15 secondes ( $900Mo = 8 \times 900Mbits = 7200Mbits$  ;  $7200/480 = 15s$ ).
- **Latence** = temps minimum d'établissement de la connexion : indépendant de la quantité de données à transporter
  - **Exemple** : latence réseau = temps nécessaire à un paquet de données pour passer de la source à la destination à travers un réseau.
  - Latence de l'ADSL : 30 ms, latence de l'infiniband QDR : 100 ns

# Hiérarchie Mémoire



- 1 Architectures de calcul actuelles
- 2 Architecture d'un serveur
  - Processeur
  - Mémoire
- 3 Evolutions technologiques et problématiques associées

# Kilo, Méga, Giga, Téra, Péta, Exa ...

Top 500 Bench Linpack	Puissance soutenue en Gflops	Puissance crête en Gflops	Nombre de processeurs ou de cœurs
Juin 1993	59.7	131	1024
Juin 2011	8 262 000	8 773 630	548 352

1997 : année du **Teraflops** ( $10^{12}$  Flops)



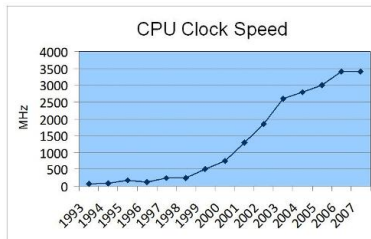
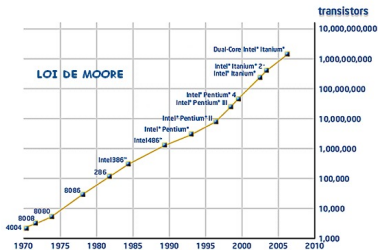
2008 : année du **Petaflops** ( $10^{15}$  Flops)



2017 : année de l'**Exaflops** ( $10^{18}$  Flops) ???

# Comment faire des calculateurs plus rapides ?

- 1 Améliorer la vitesse du processeur : augmenter la **fréquence d'horloge**.
- 2 Augmenter le **nombre de transistors**, donc augmenter le nombre de cœurs.
- 3 Utiliser des **processeurs spécialisés** : GPU notamment



## Consommation électrique

- Augmente de façon **exponentielle** en fonction de la **fréquence d'horloge** :

$$P \propto f^3$$

- $P$  : puissance (Watt)
- $f$  : fréquence
- Mais pas que !! **tous les éléments** sont consommateurs : mémoires, carte mère, alimentation inefficace ...
- *Pour fonctionner à une vitesse d'un exaflops avec ses composants actuels, le super-calculateur Tianhe nécessiterait une puissance de 1,6 milliard de watts, soit un peu plus que la puissance de la centrale nucléaire de Belleville dans le Cher !*

# Les principales limitations

## Dissipation thermique

- Directement liée à la **puissance consommée** : plus on augmente la fréquence, plus la dissipation thermique est importante
- Problème du **refroidissement** au niveau des chips, et au niveau de l'infrastructure

## Finesse de gravure

- Actuellement **32 nm** et bientôt 22 nm. Feuille de route Intel : 4 nm en 2022.
- Plus de transistors donc **plus de cœurs**, et moins de dissipation thermique (plus c'est petit moins ça dissipe)
- Mais des **défis technologiques** pour les fondeurs et des **technologies de plus en plus coûteuses** !
- **Jusqu'à quand ??**



# Conclusions

Le choix de son architecture de calcul n'est pas anodin :

- Importance d'une **architecture équilibrée** : nombre de cœurs, accès mémoire, réseau ...
- Importance d'une bonne compréhension des **contraintes liées aux infrastructures**
- Importance d'une bonne connaissance des **différents composants** pour choisir, installer, administrer et exploiter la machine

