

« de l'achat à l'exploitation des serveurs de calcul »

Autrans, Octobre 2009



Architect of an Open World™

Benchmarking

Ludovic Saugé

Ingénieur Benchmarking – Bull Extreme Computing BU

LIBERATE IT

Plan de la présentation

- Introduction – Foire aux questions
 - Benchmarkeur
 - Benchmark
- Benchmark
 - Composants élémentaires d'un nœud
 - CPU, Mémoire, Chipset ...
 - Benchmarks « standards » et synthétiques
 - HPL, HPCC, NAS ...
 - Réseau d'interconnexion rapide
 - Stockage, IOs
- Conseils et cas concrets



- Travaux Pratiques

La foire aux questions !

LIBERATE IT

Qui suis-je ?

- Ingénieur dans l'équipe benchmarking de la BU Extreme Computing de Bull (Grenoble)
 - Benchmarker, benchmarqueur, bencher, bencheur et pis ...
- Les benchmarqueurs sont des experts applicatifs HPC avec une excellente connaissance sur soft et du hard
- Ce sont des généralistes qui s'appuient sur des experts dans tous les domaines du HPC (soft, hard...)
 - Grenoble est le « centre de compétences HPC » de Bull et le centre de R&D HPC.
- La mission première du benchmarqueur
 - La réponse à l'AO ; deliverable : « *le rapport de benchmark* »
 - Autre mission : support, livraison, étude de perf., training, service ...

Mes activités

Portage
Benchmark

A

Démo

Visite

Accès

Staging

Acceptance

Fast Start

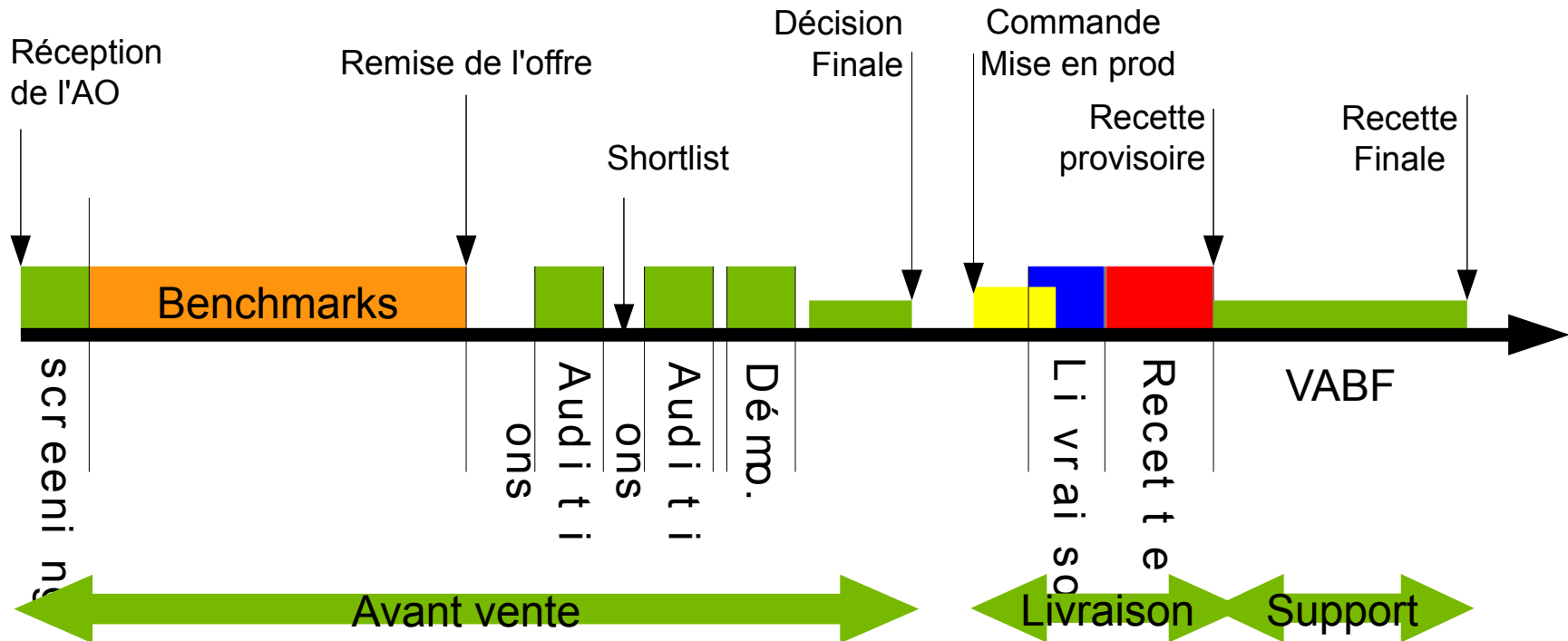
Training

A

Optimisation

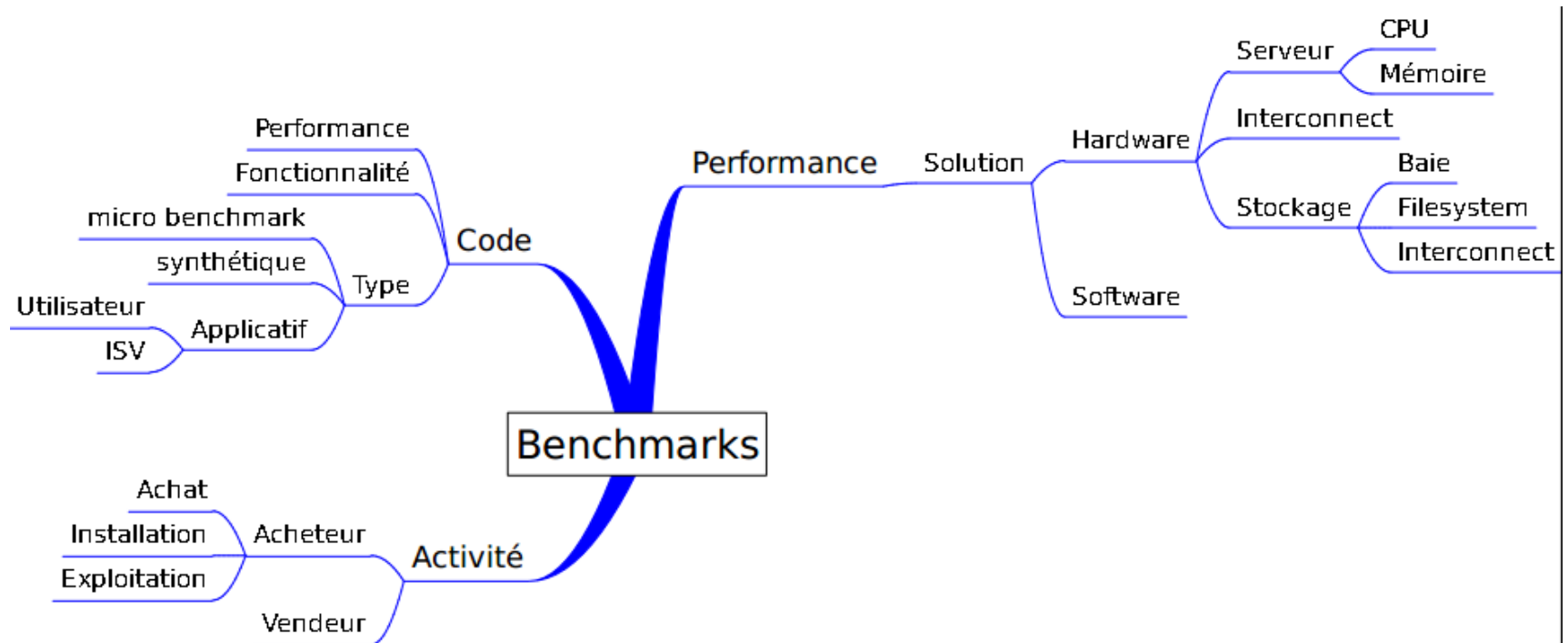
Support L3

Processus



- Y'a-t-il un processus standard ? Plus ou moins ...
- Les échelles de temps : ça dépend et rien n'est vraiment rationnel

Benchmarks



Qui « benches » et pourquoi ?

– Le vendeur

- Les benchmarks sont a priori un processus itératif et sont une contrainte à prendre à compte par l'architecte au même terme que :
 - Les contraintes économiques
 - Les contraintes environnementales (consommation, place...)
 - Autres ...
- Ces dernières contraintes sont elles-mêmes des contraintes pour le benchmarks
- Dans le but de publier (white-papers...)

Qui « benches » et pourquoi ?

– L'acheteur

- **Achat** : Dans la phase de préparation de l'AO
 - Connaître les caractéristiques de la solution existante
 - Préparer la partie benchmark de l'AO
 - Logistique
 - Analyse des résultats soumis par les vendeurs
- **Installation** :
 - Phase de recette (« acceptance tests »). Vérification des engagements fonctionnels et de performances.
- **Exploitation** :
 - Suivre tout au long de la durée de vie de ça machine les performances et déceler d'éventuels problèmes ou possibles régressions.

Mais est-il vraiment nécessaire de benchmarker ?

Est-il vraiment nécessaire de benchner ?

- Aad van der Steen (NCF, Hollande)
 - « One of the top capability machines in the world ». (IBM, BG/P)
 - « HP BladeSystem Servers lead with 35% of systems in the TOP500 ». (HP)
 - « Appro is uniquely positioned to support High-Performance Computing markets ». (Appro)
 - « The Cray XT5TM system combines unprecedented sustained application performance... ». (Cray)
 - « Bull is capable of offering Extreme Computing solutions that feature the very best available technologies at every level in the chain ». (Bull)
 - They are ALL the best you can get. So why worry?

Est-il vraiment nécessaire de benchner ?

- Publié dans Linux Magazine (Août 2009) par Jeffrey B. Layton : « *Lies, Damn Lies and File System Benchmarks* »

- « Bench... the poi... ering to
- « ***There***... .) they or for
- « How... single graphs or tables with little or no explanation of them. In some cases (...) benchmarks are published with little or no information about what was done solely to promote or detract from a particular product. In other words, benchmarks have become marketing material instead of usable information. »



Extrapolation

- Les benchmarks sont-ils une angoisse pour vous ?
- Qu'est-il intéressant de mesurer pour être sur de faire le bon choix ?
- Y'a-t-il un seul indice de performance ?
 - Le temps de restitution d'une application.
- Existe-t-il le benchmark universel ?
 - Celui qui donne une valeur unique, facile à comparer ...
 - Celui qui tourne vite ...

HPL (Top500) ? Spec2006 ? HPCC ?

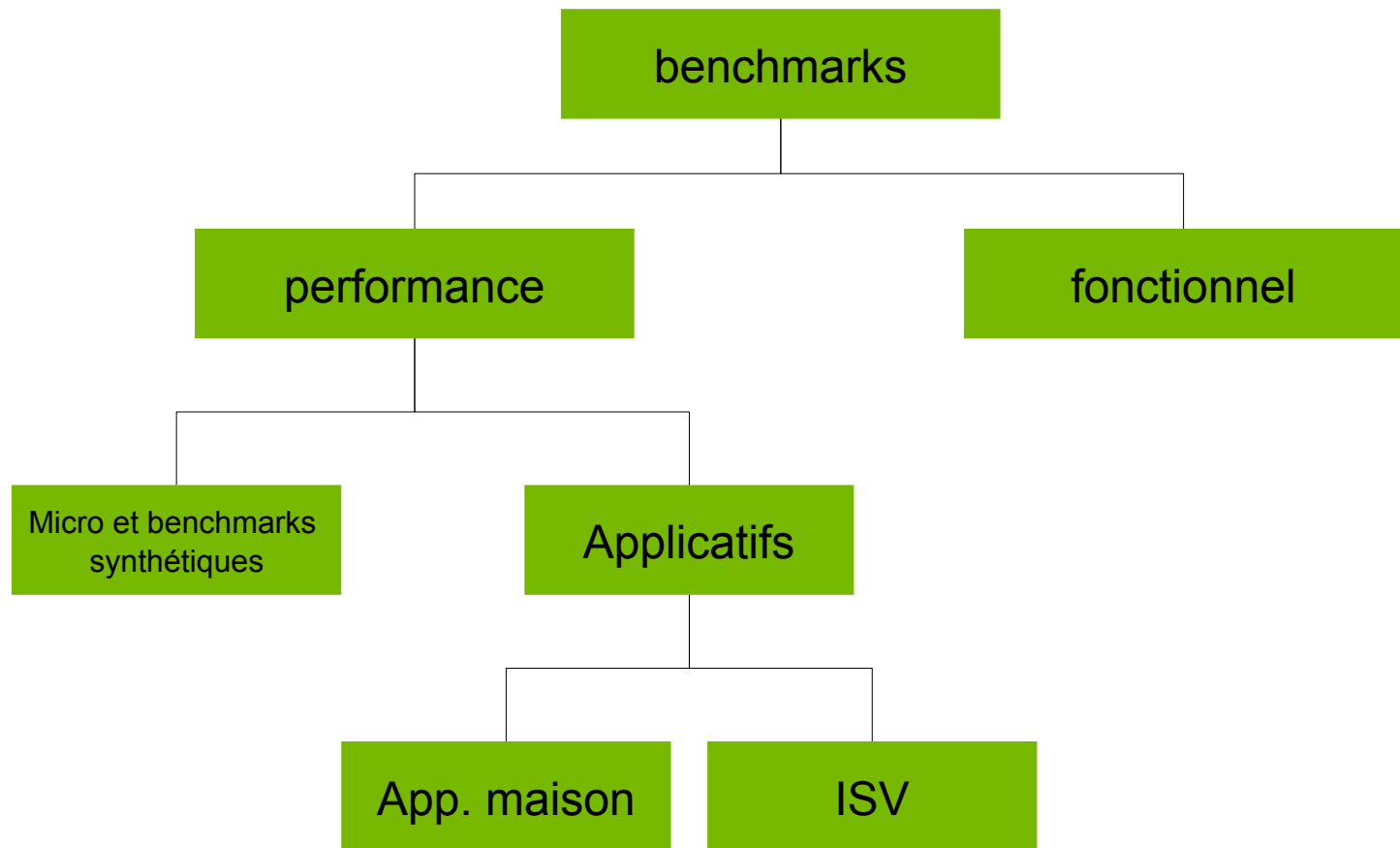
Benchmarks

LIBERATE IT

Qu'est-ce qu'un benchmark en HPC ?

- Par le terme « benchmarks » on entend :
 - La procédure faisant partie intégrante du process d'AO
 - Un ensemble de codes permettant de tester la fonctionnalité et les performances d'une solution HPC.dans son ensemble.

Benchmarks



Benchmarks

- Les micro-benchmarks et les benchmarks synthétiques fournissent une idée du comportement d'une solution élément à élément. C'est une référence, un point de comparaison entre les technologies
- Les applications sont là pour donner une idée du comportement **GLOBAL** d'une solution
 - La fin des fins d'une machine'est d'effectuer des calculs utilisateurs
 - Doivent être réalistes d'une situation de production.

Benchmarks

- Une observable est une propriété du système. Elle est mesurable et la mesure obtenue lors d'une expérience donne lieu à une quantité dénombrable (avec une unité).
- Avec cette définition, on est peut-être pas plus avancé mais ceci montre qu'il est nécessaire :
 - De bien comprendre et définir ce que l'on veut mesurer (pourquoi veut-on le mesurer ?)
 - Qu'il est nécessaire de mettre au point des **outils adaptés**, une expérience pour obtenir la mesure → le **benchmark** ...
 - Ex. « capacité d'un noeud de calcul d'effectuer des opérations mémoires aléatoires »
 - On veut donc déterminer la latence mémoire ou des « Gups »
 - On peut utiliser un bench style 'latmem' ou HPCC s-RandomAccess ...
 - Ces benches donnent des mesures différentes car ils ne mesurent pas exactement la même chose ...
 - Lequel est-il le plus pertinent pour mon application ?

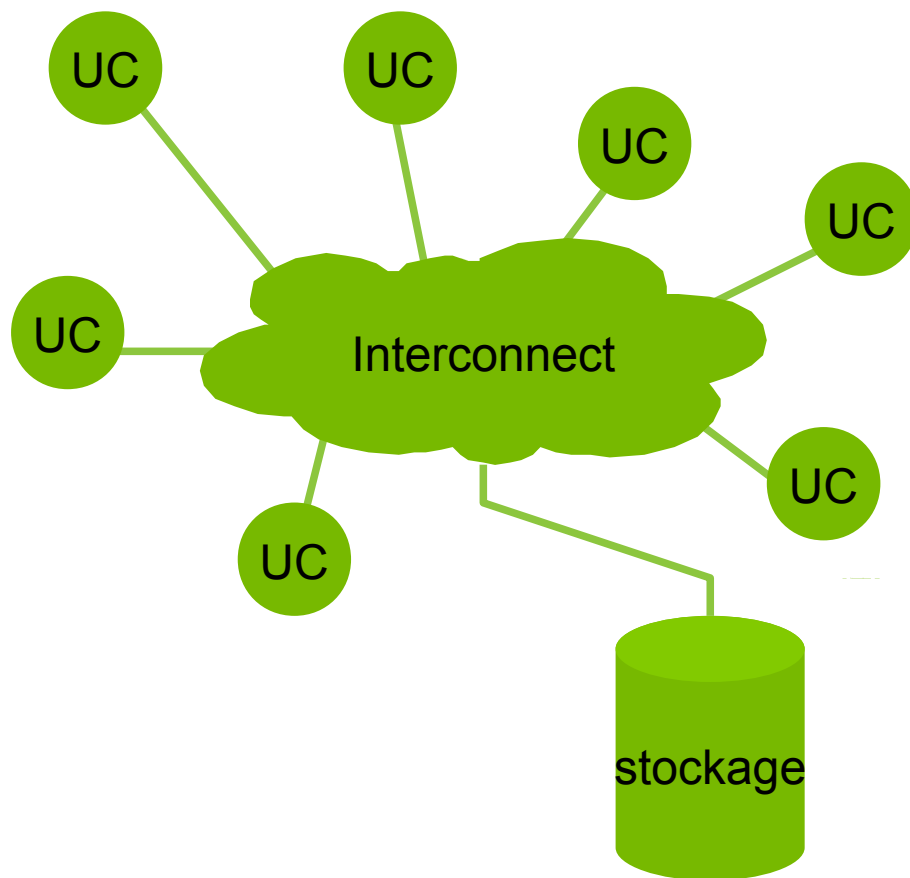
Le fantôme du benchmarkeur

- A partir d'un certain nombre d'observables et des mesures obtenues à l'aide de μ -benchmarks, c'est prédire sur ce système le comportement d'une application et en particulier le temps de restitution comme une combinaison (linéaire) des mesures w_i .

$$T = a_1 w_1 + a_2 w_2 + a_3 w_3 + \dots + a_n w_n$$

- Nécessite de connaître son code en profondeur et ses comportements et dépendances en fonction de ces observables \rightarrow détermination des a_i ...
- Est-ce que ça marche ? des fois ...
 - Nécessaire dans le cas des extrapolations
 - Petit nombre d'observables
ex. HPL

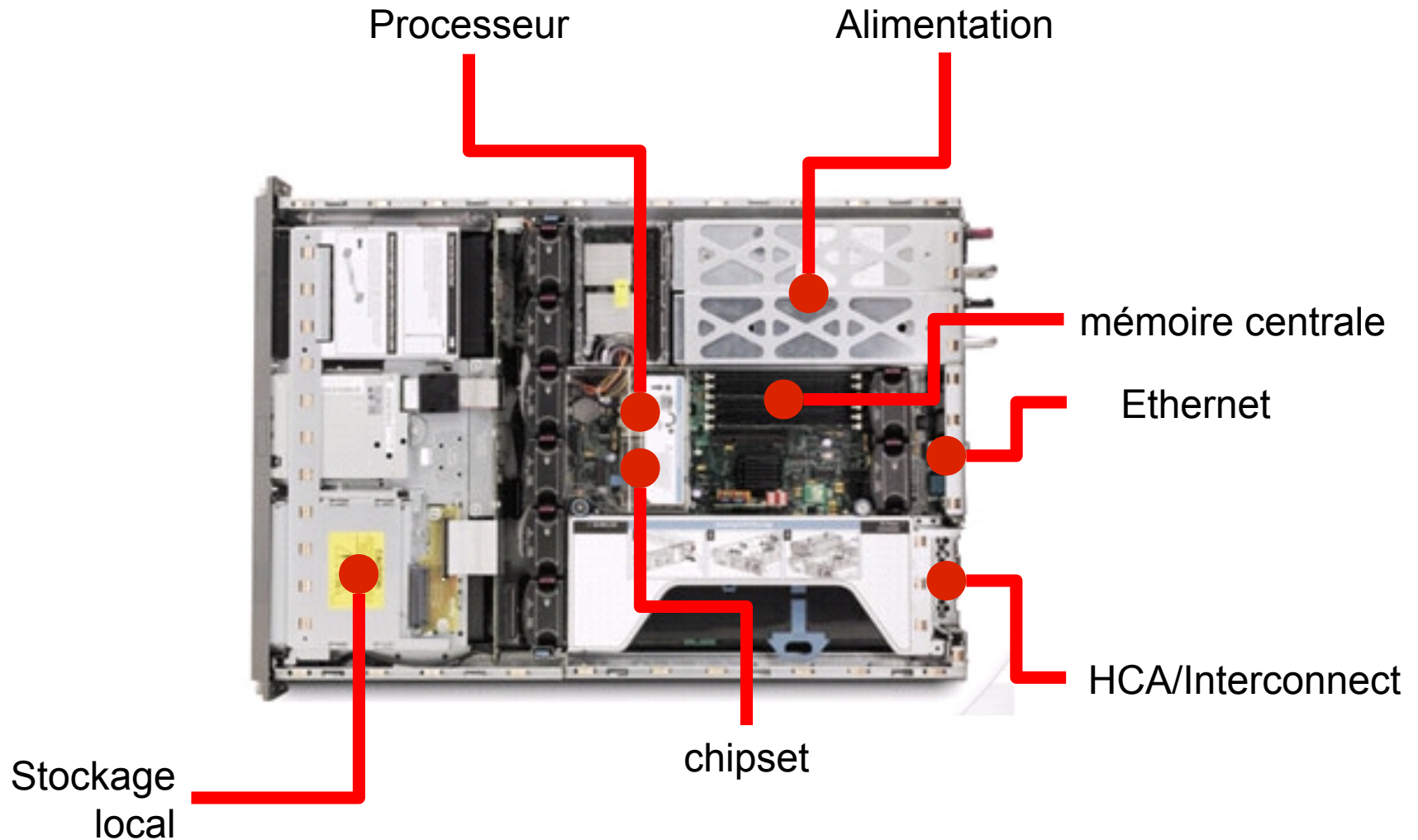
Une solution HPC



Element d'un cluster

- Réseau rapide d'interconnexion (interconnect)
 - Topologie
 - Routage
- Ensemble de nœuds de calculs
 - CPU
 - Unité de calculs (coeur) ; largeur, fréquence, instructions/cycle ...
 - Hiérarchie de cache ; nombre, quantité, organisation ...
 - Mémoire
 - Organisation/Topologie : UMA, NUMA ...
 - Technologie, fréquence (latence, bande passante,...)
 - Chipset (north/south bridge)
 - Une communication entre tout ça (bus, liens)
- Stockage
 - Les baies (type de disques, bande passante,...)
 - Systèmes de fichiers

Elements matériels



Pourquoi sommes nous là ?

- Comment le tel ou tel choix affecte les performances de machines
 - Ces effets sont-ils mesurables ? Comment les mesurer ?
 - Ex. interconnect
 - Latence, bande passante
 - Performance MPI
 - Code de l'utilisateur ...
- Ne pas oublier l'objectif final : la machine est destinée à faire tourner des codes utilisateurs/ISV en production et non des benchmarks ...