

DE LA RECHERCHE À L'INDUSTRIE



ARCHITECTURES HYBRIDES UN INCONTOURNABLE ?

Asprom 2014 | Guillaume Colin de Verdière

30 SEPTEMBRE 2014

GCdV

www.cea.fr

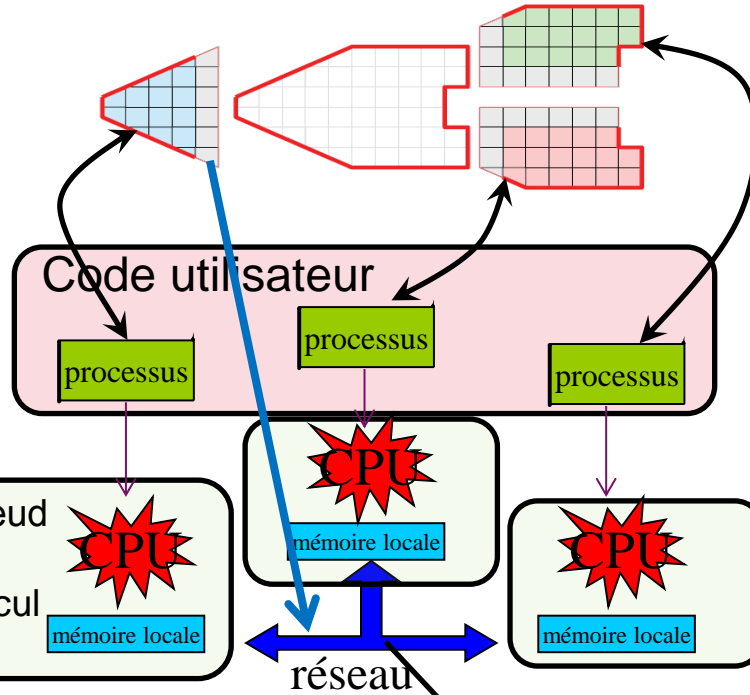
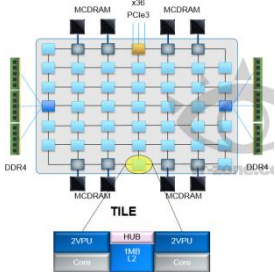
9h30-10h : Architectures hybrides, un incontournable ?

***Par Guillaume COLIN DE VERDIERE,
CEA/DAM***

Résumé: L'étude du récent TOP500 montre que l'architecture des calculateurs HPC évolue en donnant de plus en plus d'importance aux architectures hybrides. Nous présenterons donc ici ce que sont ces architectures en tentant de faire quelques projections pour l'avenir. Cet exposé pourra être aussi vu comme une introduction aux présentations suivantes.

Les architectures sont de plus en plus complexes

- Architecture interne du nœud toujours plus complexe



• Rechercher la réduction du nombre de tâches MPI

Nombre croissant de nœuds (≥ 5000)

• Architecture hybride

- Vectorisation
- Multithreading
- Grand nombre de cœurs
- Effets NUMA multiples
- Faible mémoire par cœur

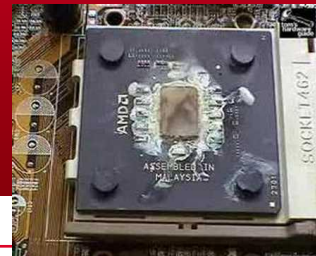
• Opérations collectives prohibitives

• Plus d'asynchronisme

• Débits croissants

• Peu de progrès sur la latence

Contrainte : la puissance dissipée: $P = P_d + P_s$



- $P_d = C_e \times F \times V^2$ **dynamique**
- $P_s = V \times I_f$ **statique**

- I_f : courant de fuite
- C_e : capacité équivalente
- V : tension
- F : fréquence [fonction de V aussi]

• Pour limiter P

■ Limitation des tensions

- ex: Pentium 4 = 1.7V; Nehalem = 1.247V

■ Limitation des fréquences voire les diminuer

- Nehalem = 2.8GHz, GPU = 1.1GHz

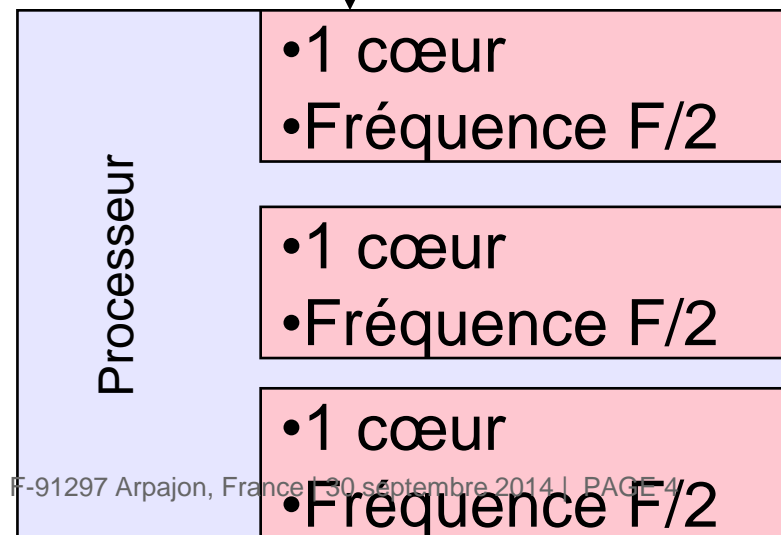
• Pour augmenter la capacité calcul à puissance constante

■ Accroître le nombre de cœurs de calcul

• Conséquence

■ La programmation doit être parallèle

- 1 processeur / cœur
- Fréquence F



Plus de performance = multiplier les cœurs

- Des possibilités explorées en fonction des contraintes de marché et de l'historique des constructeurs

- Emergence de calculateurs hybrides
- Top500 : 8 hybrides dans le top 20
 - #1 et #2 sont hybrides

- Une tentative d'architecture hybride : Road Runner

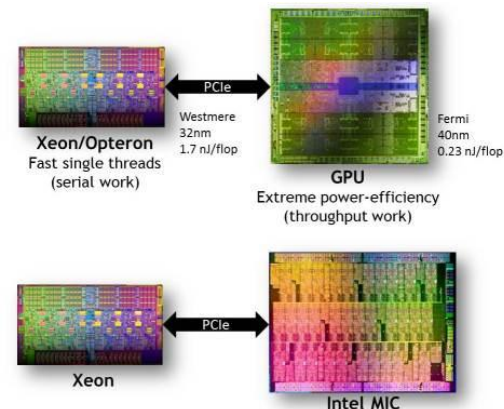
- Utilisation d'un processeur de console de jeu

- Approche « GPU » **NVIDIA, AMD**

- Vient d'un monde spécialisé : jeu et graphique
 - Elargir son périmètre
- Un très grand nombre de cœur simplifiés (un cœur = un pixel)

- Approche « Manycore » **INTEL**

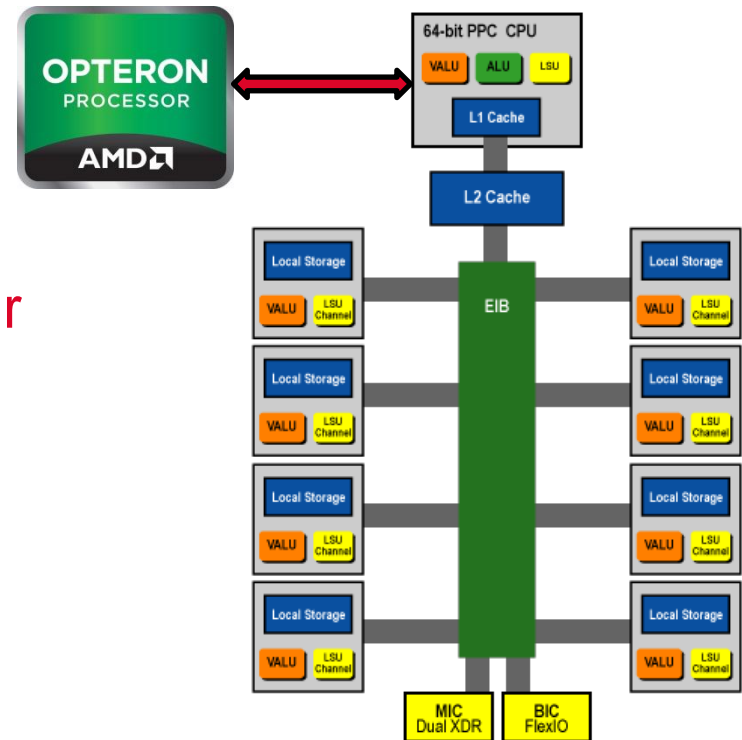
- Vient d'un monde « généraliste »
 - Pérenniser un écosystème logiciel préexistant
- Un nombre maîtrisé de cœurs de puissance moyenne



Machine hybride
Machine qui utilise
plusieurs types
d'éléments de calcul,
classiquement un
CPU + un
accélérateur relié
entre eux par un PCI-
Express

- Juin 2008 : Numéro 1 mondial
- 1,026 petaflop/s
 - Serait #36 au top500 de juin 2014 !
- 6562 AMD dual-core Opteron
- 12240 Cell Broadband Engine
- Architecture hybride aussi

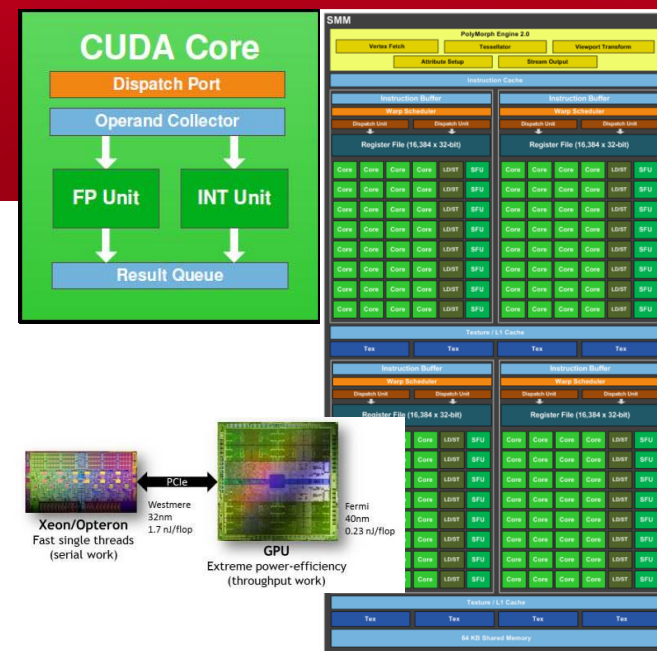
- GROS effort de programmation pour obtenir de la performance



The CELL Architecture

La voie issue des « GPU »

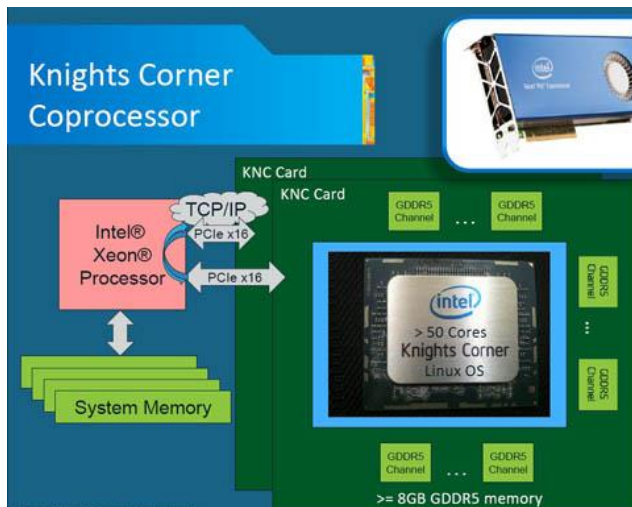
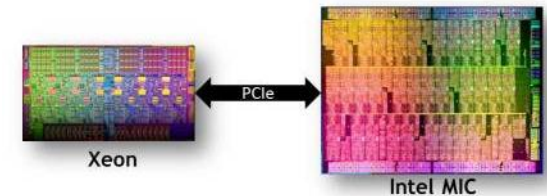
- Voie proposée par NVIDIA et AMD
- Le traitement des pixels favorise un parallélisme massif
- Repose sur une programmation mettant en œuvre un grand nombre de threads
- Le matériel introduit des contraintes sur la programmation
 - Ex: le traitement des branchements, les synchronisations,...
- Peut être vu comme de grosses machines vectorielles (data parallel)
- **Les performances potentielles sont importantes**
- Au prix d'une programmation adaptée pour les algorithmes qui s'y prêtent.
- Nécessite un CPU + mémoire (généralement deux fois celle du GPU) et l'utilisation d'un lien externe
 - Surcoût matériel non négligeable
 - Met l'accent sur la localisation des données
 - Accroît le coût des mouvements de données
- **Des réalisations probantes existent au CEA**



La machine ROMEO de l'URCA
 #184 au Top500 06/14(384TFlop/s)
 #6 au Green500
 NVIDIA K20X

La voie « Manycore »

- Voie proposée par Intel
- Conserver la compatibilité avec l'architecture x86 omniprésente
- Affirmer une simplicité de programmation
 - Adhésion d'Intel au standard OpenMP seulement
- Considéré aujourd'hui par Intel comme une architecture jalon pour l'Exascale.
 - Evolution de tout l'écosystème logiciel



GCdV

Tianhe-2, NUDT, Chine
#1 Top500 06/14
33,86 petaflops
32000 Ivy-Bridge
48000 Xeon Phi (KNC)



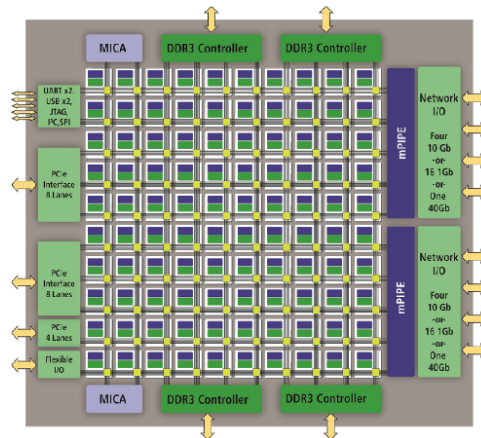
Vers une convergence des architectures ?

L'enjeu = réduire le coût énergétique du déplacement des données et du calcul

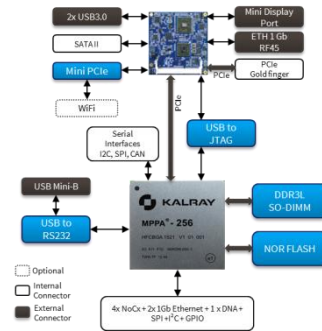
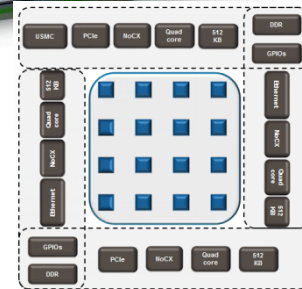
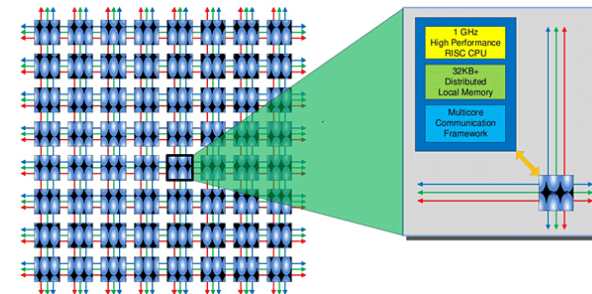
La multiplication des cœurs de calcul se généralise

- Voir l'exposé de B. de Dinechin

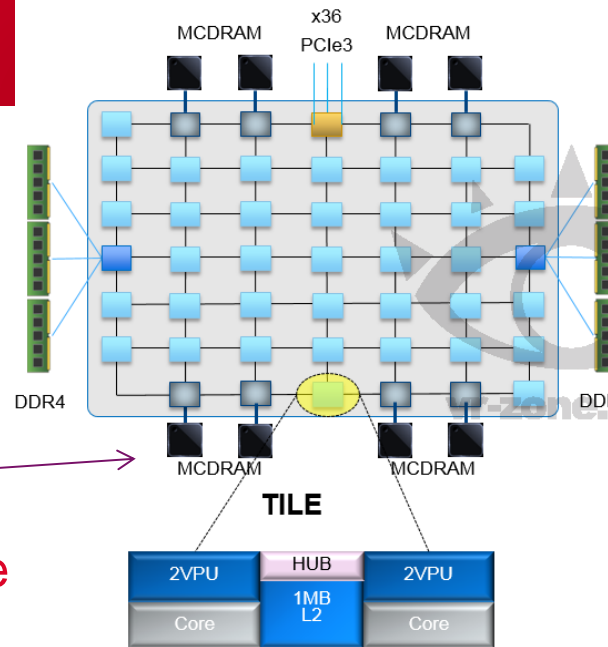
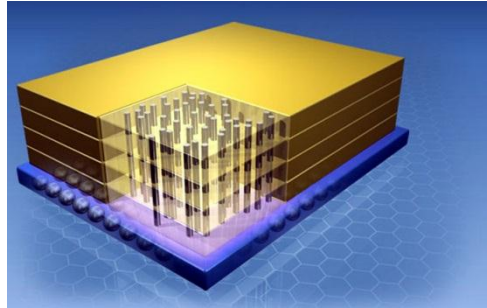
- Tiler
■ 100 cœurs



- Adapteva Epiphani
■ 64 cœurs



Mémoire stackée



Up to 72 Intel Architecture cores based on Silvermont (Intel® Atom processor)

- Four threads/core
- Two 512b vector units/core
- Up to 3x single thread performance improvement over KNC generation

Full Intel® Xeon processor ISA compatibility through AVX-512 (except TSX)

6 channels of DDR4 2400 MHz -up to 384GB

36 lanes PCI Express* Gen 3

8/16GB of high-bandwidth on-package MCDRAM memory >500GB/sec

200W TDP

- Une architecture innovante

- 72 cœurs + 4 threads / cœur

- multithreading incontournable

- Introduction de mémoire stackée (B/W ↑, latence ≡)

- Impact certain sur les codes (contrôle du placement des tableaux en mémoire)

- Introduction de 2 unités vectorielles par cœur

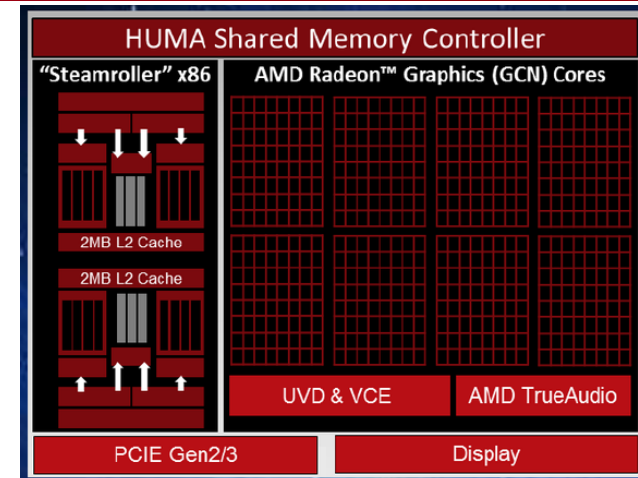
- Délivrent toute la puissance du KNL

- Programmation vectorielle obligatoire

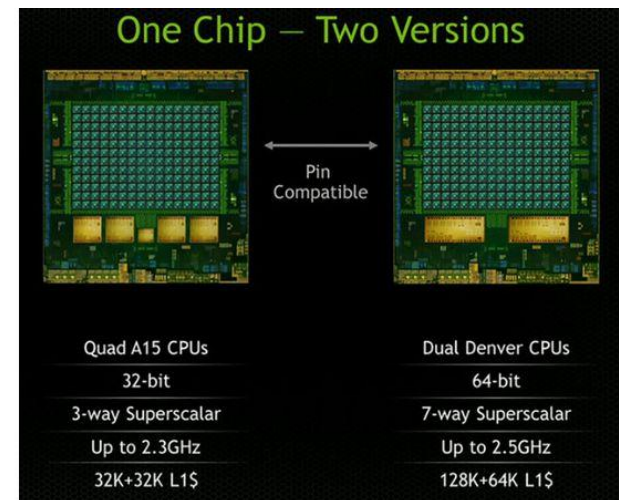
- Apparition de comportements NUMA

L'architecture Fusion selon AMD L'architecture Tegra selon NVIDIA

- Objectif : rapprocher significativement le GPU des CPUs
- Suppression des communications via PCI-Ex
- Faciliter l'usage des ressources matérielles par le programmeur.
- Contraintes
 - Adaptation du système d'exploitation nécessaire
 - Avoir un eco-système logiciel adapté
 - Quels standards ?
 - AMD pousse OpenCL, OpenMP 4.0
 - NVIDIA pousse CUDA, OpenACC
- Solutions encore inadaptées au HPC
 - Initialement marché des laptops, tablettes, smartphones
 - Évolution rapide à prévoir

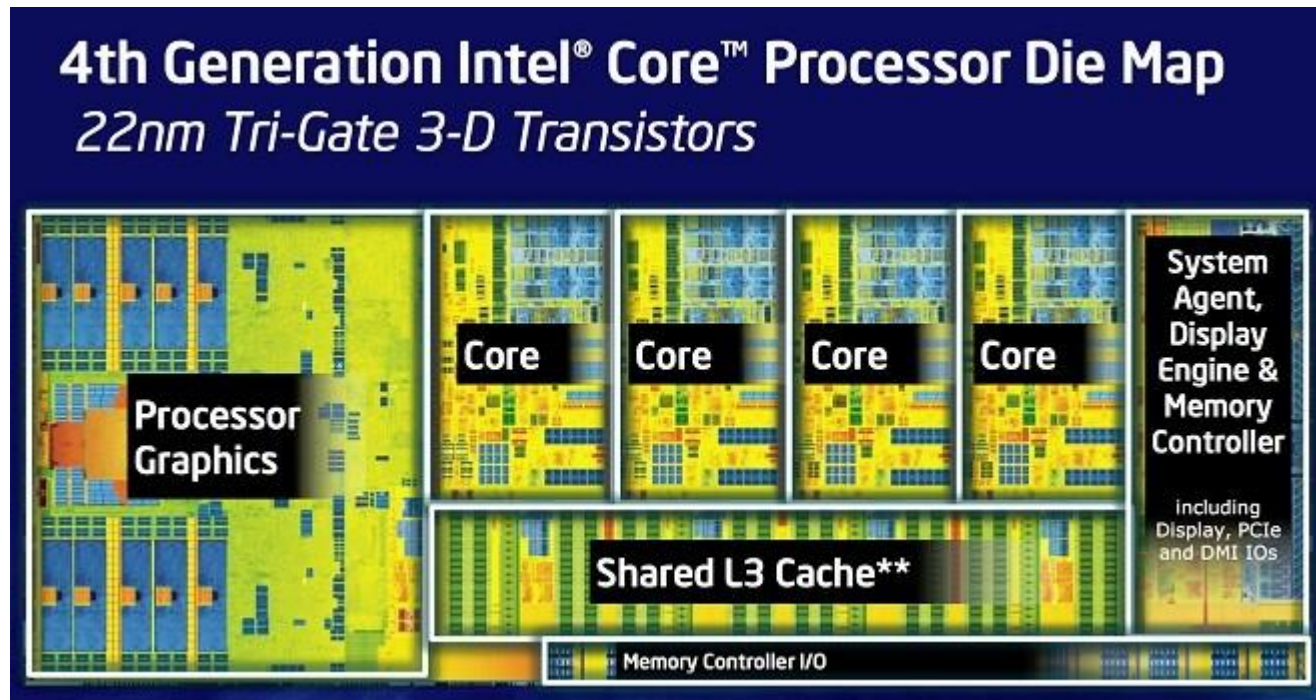


AMD Kaveri NVIDIA Tegra k1



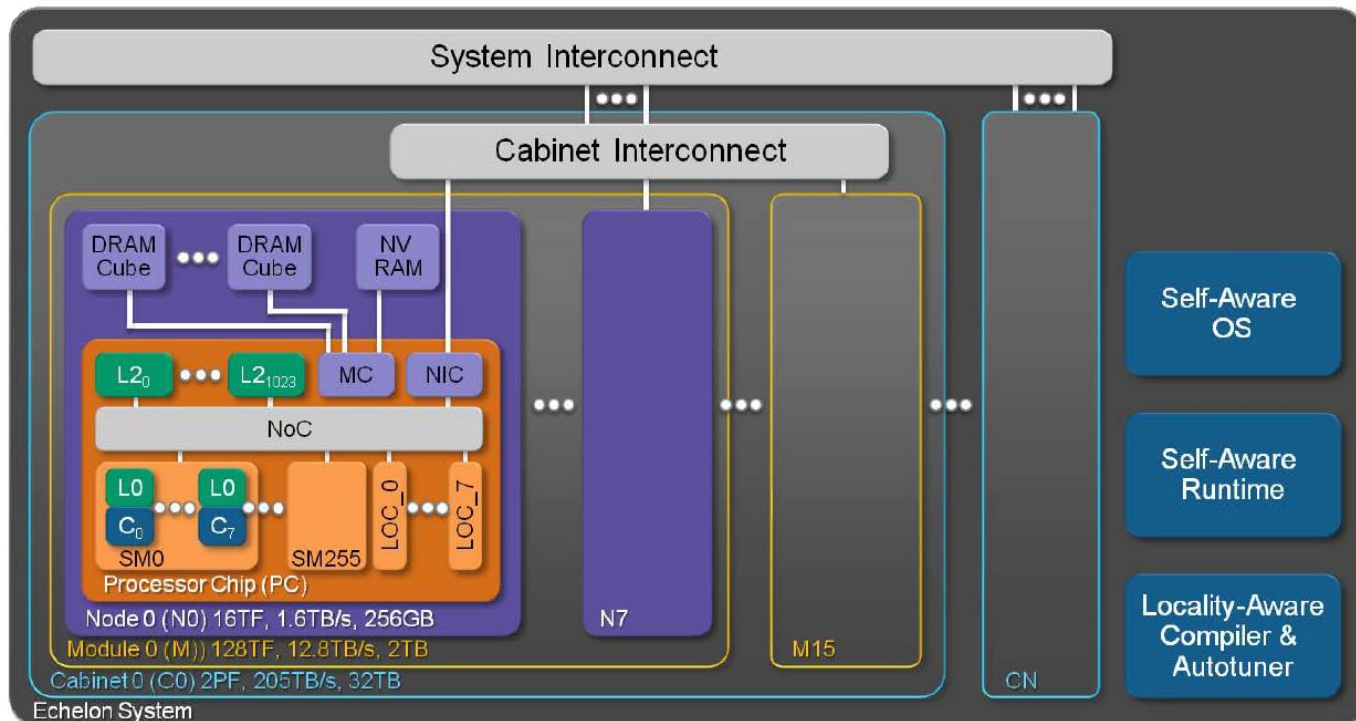
Remarque : chez Intel, les processeurs classiques possèdent aussi (déjà) ce type d'architecture

- Le GPU intégré est efficace en OpenCL (1.2)
- Haswell desktop
- Intel pousse OpenMP 4.0
- Et l'écosystème logiciel est déjà bien fourni



Concept de machine exaflopique (NVIDIA)

- **Assemblage dense des composants spécialisés**
 - LC = Latency core : CPU classique pour les opérations séquentielles
 - SM = Symmetric Multicore : Moteur SIMT de type GPU pour le massivement parallèle
 - NoC = Network on Chip : assure la cohérence des données entre les unités fonctionnelles
- **La question du logiciel (standards) reste posée**



Conclusions

- **Matérielles**

- Réduire la consommation de la partie calcul
 - Toujours plus de cœurs à fréquence maîtrisée
 - Usage systématique d'unités spécialisées SIMD/SIMT
- Réduire la consommation du mouvement des données
 - Rapprochement inévitable des CPUs et des unités SIMD puissantes
 - Intégration des interfaces réseau rapide
 - Mémoire stackée
 - Plus de threads pour masquer les latences

- **Logicielles**

- Évolution nécessaire des standards pour prendre en compte les nouveaux matériels
- Mise à niveau des écosystèmes associés
 - Compilateurs, débogueurs, profilers

- **Les challenges**

- Programmer ces « processeurs » c.f. B. de Dinechin et F. Bodin
- Optimiser les programmes pour utiliser à plein le matériel c.f. W. Jalby

